Preface

Special Issue dedicated to the 70th Birthday of Professor Imre J. Rudas – Part I

Technology development unfolds in medatrands these decades, transforming individuals and the society alike. The Internet of Things, Cloud Robotics, Industry 4.0, agile Cyber-Physical Systems show the way along which these transformations occur. The fundamental engineering concepts behind are rooted in academic research, often presented at the major scientific conferences of the community. The IEEE International Conference on Intelligent Engineering Systems (INES) is one of those, a prestigeous series established by Bánki Donát Polytechnic and Budapest Tech as the predecessors of Óbuda University. The IEEE INES series was started in 1997, and in 2019, reaches its 23rd edition, dedicated to a special occasion, the celebration of Professor Imre J. Rudas' birthday. The conference was held April 25-27, 2019 in Gödöllő, in the magnificent Royal Palace, favorite summer residence of Queen Elizabeth in the 1840s. Over 140 attendees were present from 22 countries, delivering 63 talks and 2 keynote lectures in a variety of related research topics.

The current Acta Polytechnica Hungarica issue is a collection of the newest research results based on the selected presentation at the 23rd IEEE INES. The articles span across a wide range of intelligent engineering, focusing on control and applications. The intelligence relates to the characteristics of how uncertainty, i.e., unmodeled dynamics can be handled, or what kind of solutions can be used in order to operate in an uncertain environment with proper sensing techniques. The presented solutions discuss the autonomous or automatic ways to execute given sequences or tasks without human user interaction (without detailed instructions). Arguably, the next few years will be dominated by the topics of artificial intelligence. The control engineering perspectives of soft computing gave new sights of control engineering (e.g., Tensor-Product-based control), and expanded the range of applicability in different fields: robotics, medical applications, data mining, cybersecurity, computer networks, accounting, agriculture, chemical processes, military applications, etc. Theories leading to applications, feeding into product developments for the benefit of all.

As it was clearly articulated duing the confernece, the whole series and probably this journal could not have lived withour Imre J. Rudas. Every such major project has its "power engine", distinguished senior who channels the connections, mediates the subfields, and keeps this linked scientific community at a high international standard. Imre J. Rudas, just turned 70 during the IEEE INES 2019.

His wide spectrum of professional interest is reflected in the variety of the conference contributions, and this Acta Polytechnica Hungarica special issue is dedicated to him. The current volume is a selection of his most renowned collaborators and friends throughout the world, as a special tribute to his whole academic carrier. 70 years of outstanding achievement truly deserves a reflection.

Imre J. Rudas graduated from Bánki Donát Polytechnic (Budapest) in 1971, received Master Degree in mathematics from Eötvös Lóránd University (Budapest) and received Ph.D. in robotics from the Hungarian Academy of Sciences in 1987. Achieved the Doctor of Science degree from the Hungarian Academy of Sciences in 2004. He received his first Doctor Honoris Causa degree from the Technical University of Košice (Slovakia) (2001), followed by the "Politehnica" University of Timişoara (Romania) in 2005, Óbuda University (2014) and the Slovak University of Technology (Bratislava) in 2016. He is Honorary Professor (2013) and Ambassador (2016) of the Technical University of Wrocław. His research activity is related to robotics, computational cybernetics with special emphasis on robot control, soft computing, computer-aided process planning, and fuzzy control and fuzzy sets. He has published more than 850 scientific publications, 125 journal papers, and authored 4 scientific books. The number of his independent citations is over 5000, with h-index of 32. He served as Rector of Budapest Tech from August 1, 2003 and became the founding Rector of Óbuda University in 2010 until 2014. He founded the University Research and Innovation Center of Óbuda University in 2012 and the Acta Polytechnica Hungarica journal in 2004, where he still serves as editor-in-chief.

He has a rich history with IEEE as well, started in 1991. He became IEEE Fellow in 2001, member of the IEEE Board of Directors Section/Chapter Support Committee (1998). He was mainly active in two IEEE Societies: the IEEE Industrial Electronics Society he was Administrative Committee Member and Senior Member (1996) and Vice-President of the Society (2000-2001). However, even bigger devotion presented to the IEEE System, Man and Cybernetics Society, where he was several times member of the Board of Governors (2007-2009, 2012, 2014), Vice-President (2013-2016), and currently he is President-Elect of the Society (2019). He is the founding chair of the Computational Cybernetics Technical Committee and co-chair of the Cyber-Medical Systems Technical Committee of the IEEE SMC Society. In the meantime, Imre J. Rudas served and made his footprint in the IEEE Hungary Section as well, where he was first Treasurer (1994-1998), then Vice Chair (2003-2009), and finally Chair (2009-2013). He was founding chair at IEEE Hungary Section level of the IEEE Chapter of Computational Intelligence Society (2003-2008), IEEE Chapter of System, Man and Cybernetics Society (2003-2008) and IEEE Joint Chapter of Industrial Electronics and Robotics & Automation Societies (1997-2002).

Among the several scientific awards received, he is laureate of the John von Neumann Award (2006), the Dennis Gábor Award (2006), the civil division of the Hungarian Order of Merit (2009), Pro Óbuda Award (2014), International Fuzzy

System Association Fellow (2016) and first awardee of the Rudolf Kalman Professor title given jointly by IEEE Hungary Section and Óbuda University (2017).

In the past decades, thousands of students, hundreds of colleagues and dozens of friends gained inspiration and professional support from Imre J. Rudas. His academic work was groundbreaking in Hungary, and the institutions and organization he established became a prominent part of the nation's engineering heritage. His professional network and outreach cover continents, and people around the globe are building on his results.

The editors are grateful to Imre J. Rudas for all his achievements care and support, and thankful to the authors for their excellent work composing this volume.

Budapest, Hungary
August 2019

Levente Kovács
Tamás Haidegger
Anikó Szakál

The $\mathcal{H}_\infty$ Control Performance Group

József Bokor and Zoltán Szabó

Institute for Computer Science and Control, Hungarian Academy of Sciences, Kende u. 13-17, 1111 Budapest, Hungary; e-mail: szabo.zoltan@sztaki.mta.hu

Abstract: Conventional robust control design algorithms generate only one solution that fulfils the suboptimal $\mathcal{H}_\infty$ norm criterion and thus, leaves no room for further controller tuning. Often, the designed controller is not suitable, because it is either unstable or some structural properties needs to be also satisfied. Then, the designer has to modify the original control problem and to perform the entire synthesis again. This paper proposes a method for improving the $\mathcal{H}_\infty$ control synthesis, by introducing extra flexibility into the design process. Based on the formulation of all controllers belonging to a given performance level and Lyapunov function candidate, the paper reveals the group structure, corresponding to performance problem. Based on this group structure, efficient systematic algorithms can be developed for $\mathcal{H}_\infty$ controller tuning.

Keywords: performance blending; robust control; geometry

1 Introduction

The most typical robust performance problem, can be cast as a suboptimal normalized $\mathcal{H}_\infty$ design, where for a fix (given) generalized plant description P we seek all controllers K that internally stabilize the loop and achievesthe performance guarantee $\| \mathfrak{F}_l(P, K) \| < 1$. Through a practical design problem often it would be desirable to perform a search on a set of controllers that guarantee a given performance level in order to select a suitable one for a specific implementation goal. A typical example is to find a stable controller, or a stable controller that achieves a closed loop performance that was included in the $\mathcal{H}_\infty$ design specification. This problem leads to an iterative design process. In order to implement such an iterative algorithm, a controller blending method is needed which keeps invariant the stability of the loop and the prescribed $\mathcal{H}_\infty$ performance level.

It is a fact, that by applying the Youla parametrization, the closed-loop will be an affine expression $\mathfrak{F}_l(\bar{P}, Q)$, defined by the stable parameter Q and the stable matrix $\bar{P} = \begin{pmatrix} n_{zw} & n_{zu} \\ \tilde{n}_{yw} & 0 \end{pmatrix}$. Recall that the Youla parametrization, provided as $\mathcal{K}_{stab} = \{K = \mathfrak{M}_{\Sigma_P}(Q) \mid Q \in \mathbb{Q}, (V + NQ)^{-1} \text{exists}\}$, where $\mathbb{Q} = \{Q \mid Q \text{ stable}\}$

and $\mathfrak{M}_{\Sigma_P}(Q) = (U + MQ)(V + NQ)^{-1}$, is induced by a double coprime factorization of the plant, i.e., we have stable matrices such that

$$\begin{pmatrix} \tilde{V} & -\tilde{U} \\ -\tilde{N} & \tilde{M} \end{pmatrix} \begin{pmatrix} M & U \\ N & V \end{pmatrix} = \tilde{\Sigma}_P \Sigma_P = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \tag{1}$$

with $P = NM^{-1} = \tilde{M}^{-1}\tilde{N}$ and a stabilizing controller $K_0 = UV^{-1} = \tilde{V}^{-1}\tilde{U}$. For a recent work that covers most of the known control system methodologies using a unified approach based on the Youla parameterization, see [7].

With a further simplification, i.e., an inner(co-inner)-outer factorization we can consider a parametrization where n_{zu} and n_{yw} are isometries. Then we have the invariance relation $\| \mathfrak{F}_l(\bar{P}, Q_1) - \mathfrak{F}_l(\bar{P}, Q_2) \| = \| Q_1 - Q_2 \|$ of the Euclidean distance. However, this is not the invariance we are interested in.

The starting point of this paper is the fact that solutions of the suboptimal $\mathcal{H}_\infty$ design are parametrized by the elements of the unit ball. One of the most well-known approach to arrive to this conclusion assumes either left or right invertibility of P and uses the scattering framework by augmenting the plant, if necessary, to obtain a well defined Potapov-Ginsburg transform $\hat{P}$, see [1, 9] for details. Then, a J-inner outer factorization $\hat{P} = \widehat{\Theta}_a \hat{R}$, with a block tridiagonal structure of the outer factor that corresponds to the structure of the augmentation, solves the problem. The controllers are given by $\mathfrak{M}_{\hat{R}^{-1}}(H_a)$ with $\quad H_a = \begin{pmatrix} 0 & 0 \\ 0 & H \end{pmatrix}$, $\| H \| < 1$, while the closed loop is given by $\mathfrak{M}_{\widehat{\Theta}_a}(H_a)$. Recall that Θ_a is an inner function, thus

$$\| \mathfrak{F}_l(P, K) \| = \| \mathfrak{M}_{\widehat{\Theta}_a}(H_a) \| = \| \mathfrak{F}_l(\Theta_a, H_a) \| < 1 \tag{2}$$

For the details on J-inner and J-lossless functions see [2] and [9].

These facts motivate our interest in the unit corresponding ball: if we would like to blend controllers and guarantee a prescribed performance level, we should blend elements of the unit ball. One possible approach is to consider the action of the J-unitary operators on this ball – they obviously form a group considering the composition of operators– and to express the desired operation as a group homomorphism. This is the same idea (the indirect approach) that we follow with the addition of the Youla parameters to blend stable controllers:

$$K = \mathfrak{M}_{\Sigma_P}((\mathfrak{M}_{\tilde{\Sigma}_P}(K_1) + \mathfrak{M}_{\tilde{\Sigma}_P}(K_2))) \tag{3}$$

We can formulate this process in more technical terms as follows: considering the parameter space $\mathbb{Q}$, the group of automorphisms associated to this space is formed by simple translations $Q \mapsto \tau_Q$, with $\tau_Q = \begin{pmatrix} I & Q \\ 0 & I \end{pmatrix}$, $\quad \tau_{Q_1}\tau_{Q_2} = \tau_{Q_1+Q_2}$.

In this particular case, the group homomorphism between the composition of translations and the addition of parameters is trivially combined with the Möbius transform that defines the Youla parametrization. The only obstruction might

appear for non-strictly proper plants, where some of the non-strictly proper parameters, are out-ruled. While this approach does not provide an exhaustive characterization of the topic, one can define a blending, that preserves stability and it is defined directly in terms of the plant and controller, without the necessity to use any factorization, see [17, 18].

The group actions that correspond to the addition of stable plants seen for the Youla parametrization are the hyperbolic motions of the unit ball, determined by the J-unitary operators. Therefore, to fulfil our program for the $\mathcal{H}_\infty$ problem, a suitable parametrization is needed that relates the J-unitary operators to the elements of the unit ball. Moreover, due to the increase in the plant order, we might encounter serious difficulties. While most of the results presented in this paper remain valid in a more general, operator valued, setting, here we restrict our attention to the state space solutions and blending of full order $\mathcal{H}_\infty$ controllers.

We cannot define directly, an operation on the unit ball, in a trivial way, that bears a nice algebraic structure. The map

$$\varphi_a(z) = a + \sqrt{1 - |a^*|^2}\, z(1 + a^*z)^{-1}\sqrt{1 - |a|^2}$$

is called a translation in the unit disc $\mathbb{D}$. It can be shown that φ_a is an analytic automorphism of $\overline{\mathbb{D}}$. Moreover, $\varphi_a^{-1} = \varphi_{-a}$. In the general case, the analytic automorphisms $\varphi_a(z)$ are called Möbius-Potapov-Harris transformations, see [11, 3, 5, 6]. The elementary Blaschke transformation defines the hyperbolic translations, the Möbius addition $a \oplus z = \varphi_a(z)$, like the translation group on the Euclidean plane. However, elementary translations of the hyperbolic plane do not form a group. Moreover, Möbius addition in the disc is neither commutative nor associative.

One can introduce the concept of "gyrator" $\mathrm{gyr}\colon \mathbb{D} \times \mathbb{D} \to \mathrm{Aut}(\mathbb{D}, \oplus)$, that measures the extent to which Möbius addition, deviates from associativity and commutativity:

$$\mathrm{gyr}[a, b]z = \ominus\,(a \oplus b) \oplus \{a \oplus (b \oplus z)\},$$

$$a \oplus b = \mathrm{gyr}[a, b](b \oplus a), \qquad (\mathrm{gyro-commutative\ law}),$$

i.e., gyrations represent rotations of the disc $\mathbb{D}$ about its center. Thus, in terms of elementary translations and rotations the group structure of the hyperbolic transformations can be characterized by using the concept of the gyrogroups, that was introduced and applied mainly in the context of Einstein's special relativity, see, e.g., [20, 19] and the references cited therein. The group operation, the Blaschke group, can be expressed as $(a, \alpha) \odot (b, \beta) = (a \oplus \alpha b, \mathrm{gyr}[a, \alpha b]\alpha\beta)$. Elements of the set $\mathbb{D} \times \mathrm{Aut}(\mathbb{D}, \oplus)$ are called motions of the gyrogroup in the sense that each element $(a, \phi) \in \mathbb{D} \times \mathrm{Aut}(\mathbb{D}, \oplus)$ gives rise to the motion $(l_a, \phi)z \mapsto z \oplus \phi a$. Moreover, every biholomorphic mapping h is of the form $h = \varphi_{h(0)}(uxv) = u\varphi_{h^{-1}(0)}(x)v$, where u and v are unitary operators. The metric defined as

$$\rho(a, b) = \ln\frac{1+\|\varphi_a(b)\|}{1-\|\varphi_a(b)\|} = \operatorname{arctanh}(\|\ \varphi_a(b)\ \|)$$

is invariant with respect to biholomorphic automorphisms and provides an extension of the Poincaré disk model of the hyperbolic geometry to the operator ball [8].

It turns out that when we consider the solution of different quadratic performance problems by using a state space description and LMI techniques, the solution sets are parametrized by elements of a matrix unit ball, see [14, 15, 16]. This paper presents in details an explicit parametrization of these suboptimal $\mathcal{H}_\infty$ controllers and the corresponding induced operation on the parameter space. In contrast to the operator valued case, in this context one can implement the necessary operations easily.

Concerning the structure of the presentation: for the sake of completeness in Section 2 we summarize the basic results related to the LMI-based suboptimal $\mathcal{H}_\infty$ controller synthesis problem, while Section 3 presents the result that provides all the solutions of the problem that correspond to a fixed Lyapunov matrix. Additional standard facts and notations are summarized in the Appendix. As a counterpart of the indirect approach for the controller blending based on the Youla parameters for stability, Section 4 presents the main result of the paper for performance problems by providing a parametrization of the J-unitary matrices and the group operation of this parameter space that corresponds to the hyperbolic motions defined by these J-unitary matrices.

2   LMI-based $\mathcal{H}_\infty$ Synthesis for LTI Systems

In this section we recall the main steps of LMI-based robust control synthesis. The synthesis starts from the state-space model of the augmented plant comprising the nominal plant model and all necessary weighting functions:

$$\begin{pmatrix} \dot{x} \\ z \\ y \end{pmatrix} = \begin{pmatrix} A & B_p & B \\ C_p & D_p & E_p \\ C & F_p & 0 \end{pmatrix} \begin{pmatrix} x \\ w \\ u \end{pmatrix} \tag{4}$$

Here u is the control input, y is the measured output, z is the performance output and w collects the external (performance) inputs, such as noises, disturbances, reference signals, etc. The controller is a finite dimensional, linear time invariant system described as

$$\begin{pmatrix} \dot{x}_c \\ u \end{pmatrix} = \begin{pmatrix} A_c & B_c \\ C_c & D_c \end{pmatrix} \begin{pmatrix} x_c \\ y \end{pmatrix} \tag{5}$$

With this controller, the closed loop system admits the following description:

$$\begin{pmatrix} \dot{\xi} \\ z \end{pmatrix} = \begin{pmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{pmatrix} \begin{pmatrix} \xi \\ w \end{pmatrix} \qquad \text{where}$$

$$\begin{pmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{pmatrix} = \begin{pmatrix} A + BD_cC & BC_c & B_p + BD_cF_p \\ B_cC & A_c & B_cF_p \\ C_p + E_pD_cC & E_pC_c & D_p + E_pD_cF_p \end{pmatrix}$$

$$= \begin{pmatrix} A & 0 & B_p \\ 0 & 0 & 0 \\ C_p & 0 & D_p \end{pmatrix} + \begin{pmatrix} 0 & B \\ I & 0 \\ 0 & E_p \end{pmatrix} \begin{pmatrix} A_c & B_c \\ C_c & D_c \end{pmatrix} \begin{pmatrix} 0 & I & 0 \\ C & 0 & F_p \end{pmatrix} \qquad (6)$$

The aim of the control design is to minimize the induced $\mathcal{L}_2$ norm between w and z of $T_{zw} = \mathcal{D} + \mathcal{C}(sI - \mathcal{A})^{-1}\mathcal{B}$ of, i.e., to find a stable controller (5) so that the closed loop (6) satisfies the performance relation

$$\int_0^\infty \begin{pmatrix} w(t) \\ z(t) \end{pmatrix}^T \begin{pmatrix} -\gamma^2 I & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} w(t) \\ z(t) \end{pmatrix} dt \le -\varepsilon \int_0^\infty w(t)^T w(t) dt, \varepsilon > 0 \qquad (7)$$

where the performance bound $\gamma > 0$ is minimized to be as small as possible. If $\mathcal{X}$ defines a quadratic storage function $V(x) = x^T \mathcal{X} x$ the dissipativity relation

$\frac{dV(x)}{dt} + \begin{pmatrix} w \\ z \end{pmatrix}^T \begin{pmatrix} -\gamma^2 I & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} w \\ z \end{pmatrix} < 0$ leads to the matrix inequality $\qquad \mathcal{X} > 0,$

$$\begin{pmatrix} I & 0 \\ 0 & I \\ \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{pmatrix}^T \begin{pmatrix} 0 & 0 & \mathcal{X} & 0 \\ 0 & -\gamma^2 I & 0 & 0 \\ \mathcal{X} & 0 & 0 & 0 \\ 0 & 0 & 0 & I \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & I \\ \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{pmatrix} < 0 \qquad (8)$$

which is nonlinear (quadratic) in the unknown variables. To render it linear, $\mathcal{X}$ is partitioned as $\mathcal{X} = \begin{pmatrix} X & U \\ U^T & * \end{pmatrix}$ and $\mathcal{X}^{-1} = \begin{pmatrix} Y & V \\ V^T & * \end{pmatrix}$, where $\dim X = \dim A$ and $\dim * = \dim A_c$.

If we consider $\ker \begin{pmatrix} 0 & I & 0 \\ B^T & 0 & E_p^T \end{pmatrix} = \begin{pmatrix} \Phi^1 \\ 0 \\ \Phi^2 \end{pmatrix}$ and $\ker \begin{pmatrix} I & 0 & 0 \\ 0 & C & F_p \end{pmatrix} = \begin{pmatrix} 0 \\ \Psi^1 \\ \Psi^2 \end{pmatrix}$, then, by an application of the elimination lemma, (8) is equivalent to the following set of LMIs:

$$\begin{pmatrix} Y & I \\ I & X \end{pmatrix} > 0 \qquad (9)$$

$$(*)^T \begin{pmatrix} 0 & X & 0 & 0 \\ X & 0 & 0 & 0 \\ 0 & 0 & -\gamma^2 I & 0 \\ 0 & 0 & 0 & I \end{pmatrix} \begin{pmatrix} I & 0 \\ A & B_p \\ 0 & I \\ C_p & D_p \end{pmatrix} \Psi < 0 \qquad (10)$$

$$(*)^T \begin{pmatrix} 0 & Y_\gamma & 0 & 0 \\ Y_\gamma & 0 & 0 & 0 \\ 0 & 0 & -I & 0 \\ 0 & 0 & 0 & \gamma^2 I \end{pmatrix} \begin{pmatrix} -K^T & -C_p^T \\ I & 0 \\ -B_p^T & -D_p^T \\ 0 & I \end{pmatrix} \Phi > 0 \tag{11}$$

where $\Phi = \begin{pmatrix} \Phi^1 \\ \Phi^2 \end{pmatrix} = \ker\begin{pmatrix} B^T & E_p^T \end{pmatrix}$ and $\Psi = \begin{pmatrix} \Psi^1 \\ \Psi^2 \end{pmatrix} = \ker\begin{pmatrix} C & F_p \end{pmatrix}$ and $Y_\gamma = \gamma^2 Y$.

Once we have determined X, Y and the minimal performance level γ_*, the corresponding Lyapunov matrix $\mathcal{X}_*$ can be computed as follows: compute full rank U, V such that $UV^T = I - XY$ by using an SVD decomposition and set $\mathcal{X}_* = \begin{pmatrix} Y & V \\ I & 0 \end{pmatrix}^{-1} \begin{pmatrix} I & 0 \\ X & U \end{pmatrix}$ to obtain the desired closed-loop Lyapunov matrix.

The last step of the synthesis procedure is the construction of a stable controller for the previously determined Lyapunov matrix and performance bound. By substituting $\mathcal{X}_*$ and γ_* in (8) one can easily recognize that (8) – due to the special structure (6) of the closed loop system – has exactly the same structure as the LMI in the Elimination Lemma. As a consequence, one possible controller candidate can be determined by using the basiclmi procedure.

3   Parameterization of the Controllers

In what follows we present an approach for characterizing all solutions of the design equations based on the following results:

Lemma 3 () *Let $P \in \mathbb{R}^{(m+n)\times(m+n)}$ be a given symmetric (Hermitian) matrix with inertia $in(P) = (m, 0, n)$. Let the matrix M be defined such that $P = M^* J M$, where, $J = diag(-I_m, I_n)$. Then all solutions $Z \in \mathbb{R}^{n\times m}$ of inequality*

$$\begin{pmatrix} I \\ Z \end{pmatrix}^* P \begin{pmatrix} I \\ Z \end{pmatrix} < 0 \tag{12}$$

can be expressed as $Z = T_{M^{-1}}(H)$, where H is an arbitrary contraction: $H^T H < I$.

Theorem 1 *Consider the quadratic matrix inequality*

$$\begin{pmatrix} I \\ AKB + C \end{pmatrix}^T P \begin{pmatrix} I \\ AKB + C \end{pmatrix} < 0 \tag{13}$$

in the unstructured unknown K. Assume C is of dimension $n \times m$, P has inertia $(m, 0, n)$ and assume that A has full column- and C has full row rank, respectively. If the solvability conditions are satisfied then all solutions of (13) can be characterized as follows:

$$K = V_a \Sigma_a^{-1} Z \Sigma_b^{-1} U_b^T, \quad Z = T_N(H) \tag{14}$$

where V_a, Σ_a, Σ_b, U_b and N are constant matrices determined by A, B, C, P and H is an arbitrary contraction.

Remark 1 *The rank conditions on A and B have been introduced to ease the discussion. By slightly modifying the proof and the final formula* (14) *they can be relaxed.*

Proof. Suppose (13) has a solution, i.e., the solvability conditions hold. Compute first the SVD-decomposition of A and B:

$$A = U_a \begin{pmatrix} \Sigma_a \\ 0 \end{pmatrix} V_a^T, \qquad B = U_b(\Sigma_b \quad 0)V_b^T$$

Σ_a, Σ_b are diagonal matrices collecting the nonzero singular values of A and B. Then we have

$$AXB = U_a \begin{pmatrix} \Sigma_a \\ 0 \end{pmatrix} V_a^T K U_b(\Sigma_b \quad 0)V_b^T$$

$$= U_a \begin{pmatrix} \Sigma_a & 0 \\ 0 & 0 \end{pmatrix} \widetilde{K} \begin{pmatrix} \Sigma_b & 0 \\ 0 & 0 \end{pmatrix} V_b^T = U_a \begin{pmatrix} \Sigma_a \widetilde{K} \Sigma_b & 0 \\ 0 & 0 \end{pmatrix} V_b^T$$

Introducing $Z = \Sigma_a \widetilde{K} \Sigma_b$ (13) reads as

$$(*)^T P \begin{pmatrix} I & 0 \\ C & I \end{pmatrix} \begin{pmatrix} I \\ U_a \begin{pmatrix} Z & 0 \\ 0 & 0 \end{pmatrix} V_b^T \end{pmatrix} < 0$$

Multiplying it from left and right by V_b^T and V_b we get

$$(*)^T P \begin{pmatrix} I & 0 \\ C & I \end{pmatrix} \begin{pmatrix} V_b \\ U_a \begin{pmatrix} Z & 0 \\ 0 & 0 \end{pmatrix} \end{pmatrix} < 0$$

which is the same as

$$(*)^T P \begin{pmatrix} I & 0 \\ C & I \end{pmatrix} \begin{pmatrix} V_b & 0 \\ 0 & U_a \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & I \\ Z & 0 \\ 0 & 0 \end{pmatrix} < 0$$

The next step is reordering the rows of the rightmost matrix. For this, a permutation matrix Π is introduced:

$$\Pi = \begin{pmatrix} I & 0 & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & 0 & I \end{pmatrix} \qquad \Pi \begin{pmatrix} I & 0 \\ 0 & I \\ Z & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} I & 0 \\ Z & 0 \\ 0 & I \\ 0 & 0 \end{pmatrix}$$

Then (13) amounts to

$$(*)^T P \begin{pmatrix} I & 0 \\ C & I \end{pmatrix} \begin{pmatrix} V_b & 0 \\ 0 & U_a \end{pmatrix} \Pi^T \begin{pmatrix} I & 0 \\ Z & 0 \\ 0 & I \\ 0 & 0 \end{pmatrix} < 0$$

Denoting the inner matrix product by $\tilde{P}$ and partitioning it according to the blocks of the outer terms we arrive at the following inequality:

$$
\begin{pmatrix} I & 0 \\ Z & 0 \\ 0 & I \\ 0 & 0 \end{pmatrix}^T \begin{pmatrix} \tilde{P}_{11} & \tilde{P}_{12} \\ * & \tilde{P}_{22} \end{pmatrix} \begin{pmatrix} I & 0 \\ Z & 0 \\ 0 & I \\ 0 & 0 \end{pmatrix} < 0
$$

or, equivalently

$$
\begin{pmatrix} \begin{pmatrix} I \\ Z \end{pmatrix}^T \tilde{P}_{11} \begin{pmatrix} I \\ Z \end{pmatrix} & \begin{pmatrix} I \\ Z \end{pmatrix} \tilde{P}_{12} \begin{pmatrix} I \\ 0 \end{pmatrix} \\ * & \begin{pmatrix} I \\ 0 \end{pmatrix}^T \tilde{P}_{22} \begin{pmatrix} I \\ 0 \end{pmatrix} \end{pmatrix} < 0
$$

If the analysis equation has a solution (which is assumed), then the bottom-right block is negative definite, i.e., $\bar{P}_{22} = \begin{pmatrix} I \\ 0 \end{pmatrix}^T \tilde{P}_{22} \begin{pmatrix} I \\ 0 \end{pmatrix} < 0$.   Schur   complement theorem can be applied now to transform the LMI to the form of (12):

$$
\begin{pmatrix} I \\ Z \end{pmatrix}^T \left[\tilde{P}_{11} - \tilde{P}_{12} \begin{pmatrix} I \\ 0 \end{pmatrix} P_{22}^{-1} \begin{pmatrix} I \\ 0 \end{pmatrix}^T \tilde{P}_{12}^T \right] \begin{pmatrix} I \\ Z \end{pmatrix} < 0 \tag{15}
$$

Using this form Lemma 3 can be applied to generate all solutions of (15): denoting by $\bar{P} = M^* J M$ the inner matrix if one picks a particular solution given by the matrix $Z = T_{M^{-1}}(H)$, then the original unknown controller variable K can be computed as $K = V_a \Sigma_a^{-1} Z \Sigma_b^{-1} U_b^T$.

If we apply Theorem 1 to the synthesis inequality (8) evaluated at the previously constructed Lyapunov matrix $\mathcal{X}$ and performance level $\gamma = \gamma_*$ values then we can see that the controllers that guarantee the given performance level can be parameterized as follows:

$$
K = \begin{pmatrix} A_c & B_c \\ C_c & D_c \end{pmatrix} = V_a \Sigma_a^{-1} Z \Sigma_b^{-1} U_b^T \tag{16}
$$

with $Z = T_N(H)$ and H a contractive matrix. Throughout this paper it is assumed that the domain of the Möbius transform T_N is the entire contractive ball.

Remark 2 *An analogous result can be obtained along the classical two Riccati based approach, where the set of the controllers is described by a linear fractional transform defined on the set of the contractive transfer functions, for the details see, e.g., [21]. Then, by restricting the set of parameters on the set of contractive matrices, we obtain an analogous starting point as for the LMI case.*

4   The Matrix Blaschke Group

As we have already shown, for performance problems the parametrization of the solutions provides an immediate blending possibility by following the indirect approach. In contrast to the stabilization problem [17], the identification of the elements of this approach is not trivial. In what follows we present the group structure and a parametrization of the automorphism group of the unit ball.

Setting $J = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}$ we consider the associated group of J-unitary matrices Φ, i.e., those matrices for which $\Phi^* J \Phi = J$. There is a correspondence between the contractive ball and the J-unitary matrices: for every contraction H the matrix

$$\Phi_H = \begin{pmatrix} N_H & 0 \\ 0 & N_{H^*} \end{pmatrix} \begin{pmatrix} I & -H^* \\ -H & I \end{pmatrix}$$

is J-unitary. It is convenient to introduce the following notations: $D_H = (I - H^*H)$ and $N_H = D_H^{-1}$. Observe that we have the following properties:

$$N_H = N_H^*, \quad N_{(-H)} = N_H, \quad HN_H = N_{H^*}H,$$

$$N_{UH} = N_H, \quad N_{HU}U^* = U^*N_H,$$

for any unitary U. It is immediate that $\Phi_H = \Phi_H^*$ and that $\Phi_H^{-1} = \Phi_{-H}$.

Concerning the geometric content, recall that J-unitary matrices define the movements, i.e., hyperbolic translations, on the matrix unit ball that preserve the hyperbolic distance. Their Möbius transform defines the multidimensional generalization of the elementary Blaschke products:

$$B_H(Z) = \mathfrak{M}_\Phi(\underline{Z}) = N_{H^*}(Z - H)(I - H^*Z)^{-1}D_H =$$

$$-H + D_{H^*}Z(I - H^*Z)^{-1}D_H = \mathfrak{F}_l(\Psi, Z)$$

with $\Psi = \begin{pmatrix} -H & D_{H^*} \\ D_H & H^* \end{pmatrix}$. The elementary Blaschke products $B_H(Z)$ are biholomorphic automorphisms of the unit ball $\mathcal{B}$ and $\| B_H(Z) \| \leq B_{\|H\|}(\| Z \|)$. Moreover, every biholomorphic mapping h is of the form $h = B_{h(0)}(UZV) = UB_{h^{-1}(0)}(Z)V$, where U and V are unitary operators. The metric defined as

$$\rho(A, B) = \ln \frac{1 + \| B_A(B) \|}{1 - \| B_A(B) \|} = arctanh(\| B_A(B) \|)$$

is invariant with respect to biholomorphic automorphisms and provides an extension of the Poincaré disk model of the hyperbolic geometry to the operator ball. For details see, e.g., [4, 8, 10].

Note that

$$B_H(0) = -H, \quad B_H(H) = 0, \quad B_{-H}(0) = H \tag{17}$$

$$B_H \circ B_{-H} = B_{-H} \circ B_H = I \tag{18}$$

In contrast to the Euclidean geometry, where elementary translations form a group, in the hyperbolic world we do not have this property. This fundamental difference makes things more complicated: we cannot define a group structure merely on the contractive ball. However, based on the observation that every J-unitary matrix can be expressed as an elementary translation and a block diagonal unitary action, there is a remedy.

Theorem 2 *Every J-unitary matrix can be expressed as $\Phi = W_{U,V}\Phi_H$, where H is a suitable contraction and U and V are unitary matrices, with $W_{U,V} = diag\{U,V\}$.*

For the result in the general, operator valued context, see, e.g., [2]. Its proof relies on the existence and uniqueness properties of the polar decomposition. The following commutation formula is the basic observation for our purposes.

$$\Phi_H W_{U,V} = W_{U,V}\Phi_{V^*HU} \tag{19}$$

Its importance relies in the derivation of the formula that relates the action of the J-unitary group in terms of the three parameters (U,V,H). Observe that

$$\Phi_1\Phi_2 = W_{U_1,V_1}\Phi_{H_1}W_{U_2,V_2}\Phi_{H_2} = W_{U_1,V_1}W_{U_2,V_2}\Phi_{V_2^*H_1U_2}\Phi_{H_2} = W_{U,V}\Phi_H,$$

i.e., $\Phi_{(U_1,V_1,H_1)}\Phi_{(U_2,V_2,H_2)} = \Phi_{(U,V,H)}$

The operation $(U,V,K) = (U_1,V_1,H_1) \circ (U_2,V_2,H_2)$ defined by this homomorphism is obviously a group, called the Blaschke group. If we would like to provide an explicit expression of this homomorphism, we need to provide a formula for the product $\Phi_{H_1}\Phi_{H_2}$ of the elementary Blaschke factors, i.e., for $(U,V,H) = (I,I,H_1) \circ (I,I,H_2)$.

As a first step, observe that by definition we have

$(U,V,H) = (U,V,0) \circ (I,I,H)$

$(U_1U_2,V_1V_2,0) = (U_1,V_1,0) \circ (U_2,V_2,0)$

and we have already shown that

$$(U_1,V_1,H_1) \circ (U_2,V_2,H_2) = (U_1U_2,V_1V_2,0) \circ (I,I,V_2^*H_1U_2) \circ (I,I,H_2) \tag{20}$$

Before arriving to the final formula, we need some relations that are interesting in their own right. First observe that by using the J-unitary property of Φ_H and the definition of B_H we have $\begin{pmatrix} I \\ B_H(Z) \end{pmatrix}^* J \begin{pmatrix} I \\ B_H(Z) \end{pmatrix} = (\star)J\begin{pmatrix} I \\ Z \end{pmatrix}(I - H^*Z)D_H$, i.e., the defect can be expressed as $D_{B_H(Z)}^2 = I - B_H^*(Z)B_H(Z) = Q_H^*(Z)Q_H(Z)$, with the factor $Q_H(Z) = D_Z(I - H^*Z)^{-1}D_H$.

Thus, for the unitary matrix $E_H(Z)$ we get the expression $D_{B_H} = E_H^*(Z)Q_H(Z)$, i.e.,

$E_H(Z) = Q_H(Z)N_{B_H(Z)}$, and, by a direct verification results that

$$B_H(Z) = -B_Z(H)E_H(Z) \tag{21}$$

Now we can formulate one of the main results of this section:

Theorem 3 *The product of elementary J-unitary matrices is the J-unitary matrix given by*

$$\Phi_{H_1}\Phi_{H_2} = W_{U,V}\Phi_H,$$

where the contractive term and the unitary factor can be computed as

$$H = B_{-H_2}(H_1), \quad U = E_{-H_2}(H_1), \quad V = E_{-H_2^*}(H_1^*)$$

Proof: Indeed, from the identity $B_{H_1}(B_{H_2}(Z)) = VB_H(Z)U^*$ and applying (17) we get $0 = VB_H(B_{-H_2}(H_1))U^*$, *i.e.*, $H = B_{-H_2}(H_1)$. It also follows that $B_{H_1}(-H_2) = -VHU^*$. Thus

$$D^2_{B_{H_1}(-H_2)} = UD^2_H U^* = UD^2_{B_{-H_2}(H_1)}U^*$$

$$D^2_{B^*_{H_1}(-H_2)} = VD^2_{H^*}V^* = VD^2_{B^*_{-H_2}(H_1)}V^*$$

Putting together all these results we can obtain the expressions of the unitary factors as $U = E_{-H_2}(H_1), \quad V = E_{-H_2^*}(H_1^*)$, as it was claimed. Finally, we have that

$$(E_{-H_2}(H_1), E_{-H_2^*}(H_1^*), B_{-H_2}(H_1)) = (I, I, H_1) \circ (I, I, H_2)$$

Combining Theorem 3 with (20) we have obtained the explicit formula for the desired blending operation that defines the group homomorphism

$$\Phi_{(U_1,V_1,H_1)}\Phi_{(U_2,V_2,H_2)} = \Phi_{(U_1,V_1,H_1)\circ(U_2,V_2,H_2)} = \Phi_{(U,V,H)}$$

as follows:

Theorem 4 *Corresponding to our notations, the operation given by*

$$(U,V,H) = (U_1,V_1,H_1) \circ (U_2,V_2,H_2) =$$
$$(U_1U_2E_{-H_2}(V_2^*H_1U_2), V_1V_2E_{-H_2^*}(U_2^*H_1^*V_2), B_{-H_2}(H_1)) \tag{22}$$

defines a group structure.

Remark 3 *In the performance problem considered in this paper we are interested only in the contraction part, see (16). One might think that the map $(H_1, H_2) \to B_{-H_2}(H_1)$ is sufficient to define the blending, and that the unitary part does not play any role. Thus, it seems that in the matrix case, for practical purposes one needs only the elementary Blaschke maps according to $T_{\Phi_H}(0) = -H$.*

Remember, however, that $\Phi_{H_1}\Phi_{H_2} = W_{U,V}\Phi_H$, in general. Thus, the elementary Blaschke maps are not enough to define an automorphism group structure and we should use the formula $T_{\Phi_{H_1}}T_{\Phi_{H_2}}(0) = T_{\Phi_H}(0) = -VHU^*$, where the parameters are given by Theorem 4. At this point recall, that the controller is given by (16),

where $Z = T_N(H) = (C + DH)(A + BH)^{-1}$. Thus, in an iterative process, the additional unitary factors may be used to maintain some structural constraints through the iteration. As an example, taking a generalized SVD of the pair (A, B), one can simplify the computation of the inverse during the iteration.

Conclusions

This paper proposes a method for improving the $\mathcal{H}_\infty$ control synthesis, which provides a starting point for developing algorithms that uses some sort of iteration. The paper is based on the observation that solutions of the quadratic performance problems, e.g., a suboptimal $\mathcal{H}_\infty$ design, are parametrized by the elements of the unit ball. Based on the formulation for all controllers belonging to a given performance level and Lyapunov function candidate, the paper reveals the group structure corresponding to the control performance problem.

The paper presents, in detail, an explicit parametrization of the hyperbolic motions of the matrix unit ball and the corresponding induced operation, on this parameter space. The obtained formula, leads to an indirect blending algorithm, for controllers, that guarantees a given performance level. In contrast to the operator valued case, in this context, one can implement the necessary operations easily. Based on this group structure, efficient systematic algorithms can be developed for $\mathcal{H}_\infty$ controller tuning.

Acknowledgement

References

[1]   J. A. Ball, J. W. Helton, and M. Verma. A factorization principle for stabilization of linear control systems. *Journal of Nonlinear and Robust Control*, 1:229-294, 1991

[2]   H. Dym. *J–contractive matrix functions, reproducing kernel Hilbert spaces and interpolation*. AMS, Providence, R. I., 1989

[3]   A. V. Efimov and V. P. Potapov. J-expanding matrix functions and their role in the analytical theory of electrical circuits. *Russian Math. Surveys*, 28(1):69-140, 1973

[4]   L. A. Harris. *Bounded symmetric homogeneous domains in infinite dimensional spaces*, pp. 13-40, Number 364 in Lecture Notes in Mathematics. Springer- Verlag, New York, 1973

[5]   L. A. Harris. Bounded symmetric homogeneous domains in infinite dimensional spaces. In T. L. Hayden and T. J. Suffridge, editors, *Proceedings on Infinite Dimensional Holomorphy*, Volume 364 of *Lecture Notes in Mathematics*, pp. 13-40, Springer Berlin Heidelberg, 1974

[6]   L. A. Harris. A generalization of C^*-algebras. *Proceedings of the London Mathematical Society*, 42(2):331-361, 1981

[7]   L. Kevickzy and Cs. Bányász. *Two-Degree-of-Freedom control systems: the Youla parameterization approach*. Academic Press, 2015

[8]   V. A. Khatskevich. Generalized Poincaré metric on the operator ball. *Mathematical notes of the Academy of Sciences of the USSR*, 17(4):321-323, 1984

[9]   H. Kimura. *Chain-scattering approach to $\mathcal{H}_\infty$ control*. Birkhauser, 1997

[10]   M. I. Ostrovskii, V. S. Shulman, and L Turowska. Unitarizable representations and fixed points of groups of biholomorphic transformations of operator balls. *Journal of Functional Analysis*, 257(8):2476-2496, 2009

[11]   V. P. Potapov. The multiplicative structure of J-contractive matrix functions. *Amer. Math. Soc. Transl.*, 15:131-243, 1960

[12]   C. Scherer and S. Weiland. *Linear Matrix Inequalities in Control*. Lecture Notes, 2005

[13]   C. W. Scherer. *Recent Advances on LMI Methods in Control*, chapter Robust Mixed Control and LPV Control with Full Block Scalings, pp. 187-208, SIAM, 2000

[14]   Z. Szabó, Zs. Biró, and J. Bokor. All controllers for an LPV robust control problem. In *7th IFAC Symposium on Robust Control Design, Aalborg, Denmark*, 2012

[15]   Z. Szabó, Zs. Biró, and J. Bokor. Möbius transform and efficient LPV synthesis. In *Proc. of the 51th IEEE Conference on Decision and Control, Maui, Hawaii*, 2012

[16]   Z. Szabó, Zs. Biró,. Gáspár, and J. Bokor. Elimination lemma and contractive extensions. In *Proceedings of the 52nd IEEE Conference on Decision and Control, 2013, Firenze, Italy*, 2013

[17]   Z. Szabó and J. Bokor. Controller blending – a geometric approach. In *European Control Conference 2015, Linz, Austria*, 2015

[18]   Z. Szabó and J. Bokor. Jordan algebra and control theory. In *European Control Conference 2016, Aalborg, Denmark*, 2016

[19]   A. A. Ungar. *Hyperbolic Triangle Centers*, volume 166 of *Fundamental Theories of Physics*. Springer Netherlands, 2010

[20]   A. A. Ungar. *Analytic Hyperbolic Geometry and Albert Einstein's Special Theory of Relativity*. World Scientific, 2008

[21]   K. Zhou, J. C. Doyle, and K. Glover. *Robust and Optimal Control*. Prentice Hall, Englewood Cliffs, New Jersey, 1996

Appendix

Notations and basic results

The notations used in the paper are fairly standard. The kernel of a matrix M is denoted by $M_\perp$ and is interpreted as $MM_\perp = 0$. The inertia of a matrix M is denoted by $\text{in}(m, k, n)$ where m, k, n are the number of positive, zero and negative eigenvalues of M. The Möbius transformation of matrix K with respect to the matrix N is denoted by $T_N(K)$ and is defined by $\qquad T_N(K) = (C + DK)(A + BK)^{-1}$,

where $N = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$.

Lemma 2 (Projection lemma) For arbitrary A, B and a symmetric P, the LMI

$$K^T X B + B^T X^T A + P < 0 \tag{30}$$

in the unstructured X has a solution if and only if

$$A_\perp^T P A_\perp < 0 \quad and \quad B_\perp^T P B_\perp < 0, \tag{31}$$

where $A_\perp = \ker(A)$ and $B_\perp = \ker(B)$.

If (31) is satisfied then one particular solution X of (30) can be determined by the numerical algorithm implemented in basiclmi MATLAB routine.

Lemma 3 (Elimination lemma) Consider the quadratic matrix inequality

$$\begin{pmatrix} I \\ AXB + C \end{pmatrix}^T P \begin{pmatrix} I \\ AXB + C \end{pmatrix} < 0 \tag{32}$$

in the unstructured unknown X. Assume C is of dimension $n \times m$ and P has inertia $(m, 0, n)$. Then (32) has a solution if and only if

$$B_\perp^T \begin{pmatrix} I \\ C \end{pmatrix}^T P \begin{pmatrix} I \\ C \end{pmatrix} B_\perp < 0, \quad and \quad A_\perp^T \begin{pmatrix} -C^T \\ I \end{pmatrix}^T P^{-1} \begin{pmatrix} -C^T \\ I \end{pmatrix} A_\perp > 0, \tag{33}$$

where $A_\perp = \ker(A)$ and $B_\perp = \ker(B)$.

Note, that solution of the H_∞ problem uses the Projection lemma, which is a special case of the Elimination lemma when $P = \begin{pmatrix} Q & S \\ S^* & 0 \end{pmatrix}$.

Möbius transformation and basic properties

Definition 1 *Let $M \in \mathbb{F}^{(m+n) \times (m+n)}$ ($F = \mathbb{R}$ or $\mathbb{C}$) be partitioned as $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$. The Möbius transformation T_M is defined by the equation*

$T_M(X) = (C + DX)(A + BX)^{-1}$ *for $X \in \mathbb{F}^{n \times m}$ where $(A + BX)^{-1}$ exists. Denote by* $\text{dom}(T_M) = \{X \in \mathbb{F}^{n \times m} : \exists (A + BX)^{-1}\}$ *the domain of T_M.*

The dual Möbius transformation is defined by $T_M^d(Z) = (ZB + D)^{-1}(ZA + C)$,

and $\text{dom}(T_M^d) = \{Z \in \mathbb{F}^{n \times m} : \exists (ZB + D)^{-1}\}$.

Theorem 5 *Let $M \in \mathbb{F}^{(m+n) \times (m+n)}$. Then*

$$X \in dom(T_M^d) \quad \Leftrightarrow \quad X^* \in dom(T_{L^*M^*L}).$$

Moreover $T_M^d(X) = T_{L^*M^*L}^*(X^*)$, where $L = \begin{pmatrix} 0 & I_m \\ I_n & 0 \end{pmatrix}$. If $M \in \mathbb{F}^{(m+n) \times (m+n)}$ is a nonsingular matrix, then $T_M(X) = -T_{M^{-1}}^d(-X)$.

Corollary 1 $-T_M^*(X) = T_{L^*M^{-*}L}(-X^*)$.

Let us consider the composition of two Möbius transformations.

Definition 2 *Let M and N be partitioned as $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \quad N = \begin{pmatrix} E & F \\ G & H \end{pmatrix}$.*

Composition of the transformations T_M and T_N is $(T_N \circ T_M)(X) = T_N(T_M(X))$.

Lemma 4 $(T_N \circ T_M)(X) = T_N\big(T_M(X)\big) = T_{NM}(X)$, *with $X \in dom(T_M)$ and*

$T_M(X) \in \text{dom}(T_N)$. *If M is nonsingular, $Z \in \text{dom}(T_M)$ and $T_M(X) = K$ then $K \in \text{dom}(T_{M^{-1}})$ and $T_{M^{-1}}(K) = X$, i.e., $\text{dom}(T_M) = \text{Range}(T_{M^{-1}})$.*

Comparison of Various Data Mining Classification Techniques in the Diagnosis of Diabetic Retinopathy

Spandana Vadloori[1], Yo-Ping Huang[1,2,*] and Wei-Chi Wu[3,4]

[1]Department of Electrical Engineering, National Taipei University of Technology, Taipei, Taiwan 10608; yphuang@ntut.edu.tw

[2]Department of Computer Science and Information Engineering, National Taipei University, New Taipei City, Taiwan 23741

[3]Department of Ophthalmology, Chang Gung Memorial Hospital, Linkou, Taoyuan, Taiwan 33333

[4]College of Medicine, Chang Gung University, Taoyuan, Taiwan 33333

*Corresponding author

Abstract: Diabetic retinopathy (DR) has been the most frequently occurring complication in the patients suffering from a long-term diabetic condition, that ultimately leads to blindness. Early detection of the disease through biomarkers and effective treatment has been proposed to prevent/delay its occurrence. Several biomarkers have been explored, to help understand the incidence and progression of DR. These included the presence of microaneurysms, exudates, hemorrhages, etc. in the retina of the patients, which contributes to the disease. Investigation of the retinal images from time to time has been proposed as a strategy to prevent blindness. Evaluating the retinal images manually is time-consuming and demands great expertise in the diagnosis of DR. To circumvent such issues computer-aided diagnosis are very promising in the detection of DR. In the present study, we used a DR dataset and applied different classification algorithms in machine learning to predict the occurrence of the DR. The classifiers employed herein, included K-nearest neighbor, random forest classifier, support vector machine, regression tree classifier, logistic regression and the Naïve Bayes theorem. Our results showed that the random forest classification model provided the significant detail of attributes in terms of their importance in the diagnosis of the DR. More importantly, our supervised classification models provided the prediction accuracy of the disease and Naïve Bayes classifier demonstrated highest accuracy of 80.15% in the prediction of DR compared to the others. Additionally, receiver operating characteristics (ROC) analysis, with the classifiers and the area under curve (AUC) represented the fitting results of each classifier. The presented approach can prove to be a potential tool for the ophthalmologist in the early diagnosis tool for DR.

1   Introduction

Diabetic Retinopathy (DR) is the most common, yet a serious condition that occurs
in patients who are suffering from chronic, severe diabetes and may lead to
damaged retinas, causing blindness if not diagnosed and treated in the early stages
[1-3]. DR is commonly observed in 80% of the individuals who are suffering due
to diabetes for more than 20 years [1]. Statistics reveal that the occurrence of DR is
increasing at an alarming rate, and as stated by WHO (World Health Organization),
DR reports for 4.8% of the 37 million cases of blindness that take place globally
[4]. By 2030 more than 366 million people are estimated to suffer from DR [2].

DR is a progressive disease that develops gradually. The development of DR has
been grouped into four stages, (1) mild-NPDR (non-proliferative diabetic
retinopathy), (2) moderate-NPDR, (3) severe-NPDR and (4) proliferative diabetic
retinopathy (PDR) [5, 6]. Mild-NPDR is the early stage of DR characterized by
the swelling of the retinal blood vessels. These balloon-like swollen blood vessels
appear as dark red dots called microaneurysms (MAs) whose sizes range from 20
to 200 microns [7, 8]. Leakage of MAs is a noticeable and essential sign for the
detection of DR. In the second stage of DR, known as moderate-NPDR, the retinal
blood vessels swell and get blocked that leads to a state called diabetic macular
edema (DME). This fluid further increases in the macular zone of the retina
leading to the third stage called severe-NPDR where the DR symptoms become
worse. Initially mild-, moderate- and severe-NPDR were classified as three
different groups. However, later they were combined into a single group called
NPDR. The final stage of the DR is called PDR, where more blood vessels are
generated inside the retina and a fluid called vitreous gel is filled in the eyes. Such
blood vessels are very fragile and prone to fluid leakage and bleeding
consequently forming a scar tissue that leads to the retinal detachment. In addition
to MAs, the other biomarkers for DR are exudates (EXU). These exudates are
round or oval-shaped fatty protein-based particles found in the nerve fiber layers
of the retina that are formed in the NPDR state due to fluid leakage from the
damaged retinal blood vessels and had been linked to blindness [9]. Depending on
their appearance EXUs were divided into two types, hard EXUs and soft EXUs,
which appear as the hard waxy patches and softer EXUs respectively. These soft
EXUs are also called cotton wool EXUs (Figure 1). Hard EXUs are located near
MAs or at the edges of the retinal edema. Regular screening of the retinal images
in diabetic patients has been proposed for early diagnosis of the DR diseases,
which can prevent people from becoming blind [10, 11].

Data mining is a process of analyzing the raw information from a large information database and turning it into useful information where meaningful patterns and trends of the data emerge [12]. Data mining has been potentially useful in various areas of healthcare [13, 14]. Today, the healthcare industry produces copious data about patients, clinical symptoms, disease detection techniques, etc. Extracting healthcare data could be very useful to perform medicinal evaluations to diagnose or cure disease. Particularly data mining has gained momentum in identifying the risk of cardio-vascular diseases [15, 16], lung diseases [17], cancer [18], diabetes [19], etc.
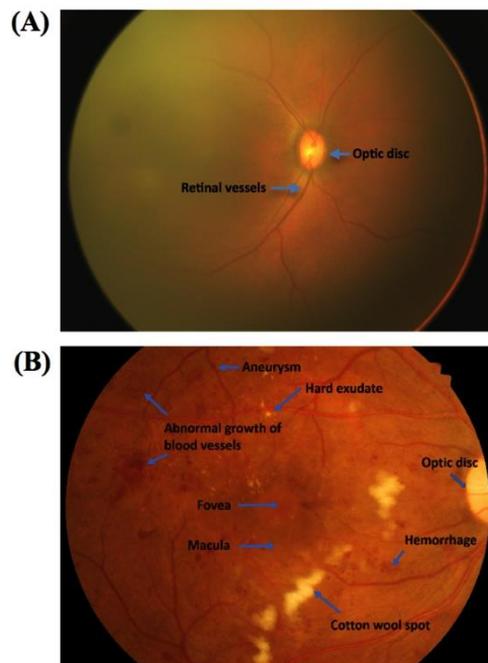
Figure 1
Comparison of the retinal images from (A) normal and (B) with the lesion of NPDR

In the present study, we focused on evaluation of the DR dataset by using various data mining classification techniques. This strategy enabled us to determine the prediction accuracy of the DR. We further evaluated the performance of the classifier models by calculating the sensitivity and specificity of these models. Our results showed that the Naïve Bayes classifier was the most efficient classification model compared to the others and also addressed the significant contribution of attributes in the diagnosis of DR.

This paper has been organized as follows. Dataset used to test different classifiers were introduced in Section 2. Test results and discussions were presented in Section 3. The conclusions and future work were given in Section 4.

2   Detailed Attributes of Dataset

In the present study, we tested the Diabetic Retinopathy Debrecen (DRD) dataset accessible from UCI Machine Learning Repository [20]. The dataset comprised of the attribute features extracted from Messidor images. In light of this dataset, we anticipated whether the image has indications of having DR or no DR. Based on Hidden Markov Random fields (HMRF) [21] the quality of the images in the dataset was assessed. The retinal images were characterized during the initial pre-screening step and all the attributes features were extracted [22]. This information was accessible from the dataset and we performed analysis on this dataset. But the problem is that there is no report, on which those attributes, that are critical for the annotation of patients with DR are available. The detailed features and attributes in the dataset and image description are given in Table 1 and Table 2, respectively.

Table 1

Details of the featured indices of the dataset

Dataset characteristics	
Number of instances	1151
Attribute characteristics	Integer, Float
Number of attributes	20
Associated tasks	Classification

Table 2

Dataset attributes and their descriptions

Attribute	Description
1	Binary result of quality assessment 0: bad 1: sufficient quality
2	Binary result of pre-screening 0: Lack of retinal abnormality 1: presence of retinal abnormality
3-8	Results of Micro aneurysm detection revealing the number of Micro aneurysms at varying confidence levels alpha =0.5 to 1
9-16	The exudates contained the same information as microaneurysms (3-8), where they are represented by a set of points instead of number of pixels constructing the lesions. These features were normalized by dividing the number of lesions with the diameter of the ROI to compensate different image sizes
17	Euclidean distance between the center of the macula and the center of the optic disc
18	Optic disk diameter
19	The binary result of AM/FM classification
20	Presence or absence of DR 1: Presence of DR (accumulation of 1,2,3 stages in Messidor) 0: Healthy

MAs were identified by preprocessing method and candidate extractor ensembles [23], while EXUs were detected by optimal combinations of ensemble-based system through voting system [24]. They did not mention among the candidate factors which MAs and EXUs were critical in determining DR. In this study, the dataset containing 1151 samples was cleaned by removing bad quality samples from the dataset. Furthermore, we also preprocessed for outliers using Python program by setting a threshold to 3. The final dataset contained 983 samples, which were later divided into 60% for training and 40% for prediction dataset. Figure 2 illustrates the framework of the methodology used.
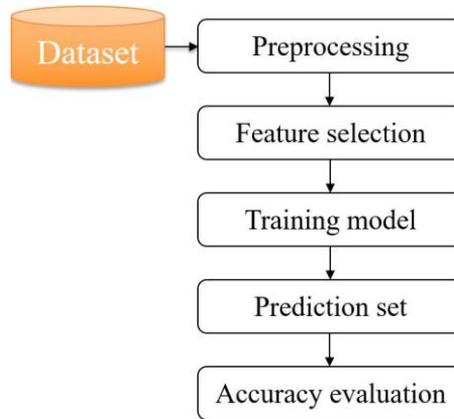
Figure 2
Framework for classification methodology

The performance of all these models was evaluated by calculating the sensitivity, specificity and accuracy based on the formulas given below:

$$Sensitivity = \frac{TP}{TP + FN} \times 100\% \tag{1}$$

$$Specificity = \frac{TN}{FP + TN} \times 100\% \tag{2}$$

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \times 100\% \tag{3}$$

where TP: true positive; TN: true negative, FN: false negative, FP: false positive.

Additionally, we performed Area Under the Curve (AUC) and Receiver Operating Characteristics (ROC) curve to display the performance of the models. In AUC, the value always lies between 0 and 1 [25]. This is useful to visualize the classification problems to distinguish between the classes.

3   Results and Discussions

In this study, we analyzed a total of 983 samples from the dataset. This dataset included 51.3% of patients diagnosed with DR (indicated by 1), while the remaining 48.7% were not diagnosed with DR (indicated by 0). Attributes with higher variance were proposed to have valuable information [26]; therefore, we performed attribute scores (Table 3) and removed the low-scoring attributes which were almost zero, namely EC.DIST. and OPT.DIA.

The dataset was characterized by using various classifiers with a motivation to predict the occurrence of DR, derive the rules in diagnosis of DR and understand the importance of attributes in detection of the DR. Six classification models were evaluated for the disease prediction accuracy, which included K-NN, random forest, regression tree, SVM, logistic regression, and Naïve Bayes. All the experiments were carried out in XLSTAT.

Initially, the classification model was trained and the prediction results from the test dataset were analyzed by the classifier using the DR attributes, MAs and EXUs, etc. The prediction accuracy of each classifier model was evaluated based on the result that determines the presence or absence of the disease. Later, we evaluated all the classification models tested in the study, and compared the accuracy of each of the models in detection of DR. The attributes and their statistical representations are given in Table 3.

Table 3
Summary statistics (training/quantitative) of the dataset

Attribute	Minimum	Maximum	Mean	Std. deviation	Score
MAœ=0.5	1.00	151.00	38.43	25.62	0.47
MAœ=0.6	1.00	132.00	36.91	24.11	2.91
MAœ=0.7	1.00	120.00	35.14	22.81	14.02
MAœ=0.8	1.00	105.00	32.30	21.11	38.75
MAœ=0.9	1.00	97.00	28.75	19.51	60.79
MAœ=1.0	1.00	89.00	21.15	15.10	66.31
EXU1	0.35	403.90	64.10	58.49	4.60
EXU2	0.00	167.10	23.09	21.60	29.69
EXU3	0.00	106.10	8.71	11.57	18.88
EXU4	0.00	59.77	1.84	3.92	1.03
EXU5	0.00	51.42	0.56	2.48	37.38
EXU6	0.00	20.10	0.21	1.06	23.40
EXU7	0.00	5.94	0.09	0.40	16.46
EXU8	0.00	3.09	0.04	0.18	7.36
EC. DIST.	0.37	0.59	0.52	0.03	0.00
OPT.DIA.	0.06	0.22	0.11	0.02	0.00
AM/FM	0.00	1.00	0.34	0.47	0.13

In K-NN classification, the class of the query is obtained based on the nearest neighbors to that of the example query. K-NN searches the pattern to the closest data and then assigns to data that is unknown [27]. In this classification, we used Euclidean distance as a criterion along with 10-fold cross validation to predict the class label from the prediction dataset. With K-NN classification system, 71.25% of accuracy was achieved from the prediction dataset as shown in Table 4. The results based on prediction class showed 205 samples without DR, while 188 samples showed the occurrence of DR. The results were given in Table 5, where the first 10 predicted results in each class were shown as representation for all the test cases from the dataset. Here, the terminology PredObs indicates with the sample number from the prediction dataset. For example, PredObs1 means that the 1st out of 393 patients was predicted to be no DR. Similarly, PredObs7 indicates that the 7th patient was predicted to have DR. In K-NN, we obtained AUC value of 0.707 (Figure 3).

Table 4

K-NN based classification

Confusion matrix (prediction dataset)				
from \ to	0	1	Total	%
0	146	54	200	73.00
1	59	134	193	69.43
Total	205	188	393	71.25

Table 5

K-NN prediction results by class

Class	0	1
Objects	205	188
	PredObs1	PredObs7
	PredObs2	PredObs10
	PredObs3	PredObs11
	PredObs4	PredObs14
	PredObs5	PredObs16
	PredObs6	PredObs17
	PredObs8	PredObs19
	PredObs9	PredObs28
	PredObs12	PredObs30
	PredObs13	PredObs32
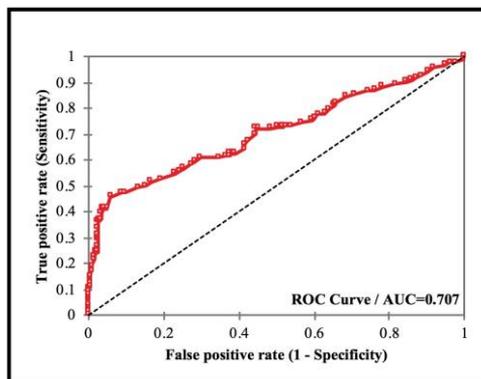
Figure 3
ROC for K-NN model

The next classification system we tested was random forest. In this type of classification, many small decision-trees are merged into a forest that displays the classification table of well-classified observations. Usually this classifier is fast and robust to noise and has better explanation and visualization of its output [28]. This classifier does eliminate the possibility of overfitting the data. Importantly, random forest classifier in our study has identified the most important attributes from the training dataset. Here, we used bagging approach to obtain more accurate results. We observed MAs and EXUs were two critical attributes in detecting DR. Among all the MAs, MA0.5 was seen as a critical attribute. In the case of EXUs, EXU7 was critical followed by EXU1 (Figure 4). In this classification model an accuracy of 72.71% (Table 6) was achieved in the training set and 71.76% with the prediction set. Here, we obtained AUC value of 0.715 (Figure 5).

Table 6
Accuracy as determined by using random forest classifier

Confusion matrix –Training set				
From \ to	0	1	Total	Accuracy (%)
0	210	88	298	70.47
1	73	219	292	75.00
Total	283	307	590	72.71
Confusion matrix –Prediction set				
From \ to	0	1	Total	Accuracy (%)
0	147	49	196	75.00
1	62	135	197	68.53
Total	209	184	393	71.76
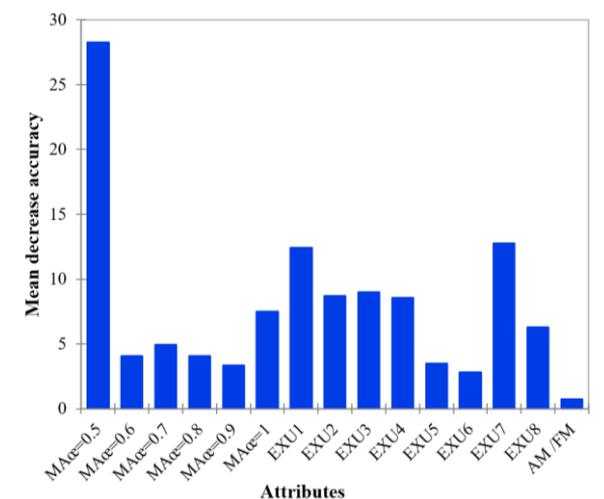
Figure 4
Significance of attributes classified by random forest classifier

We also tested the classification by the regression tree method. This classifier uses the trained dataset and generates itself correctly in order to generate a decision tree. Such decision trees are quite easy to understand and analyze. In this model the decisions are predicted by following the decisions from the root node and to the leaf node. The response of occurrence of the DR is present in the leaf node. Depending on the working process of learning, any new input data would be classified in generation of decision tree [29].
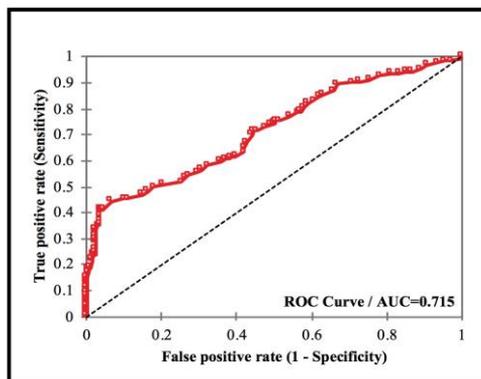
Figure 5
ROC for random forest model

This classification system also generated the rules with the critical attribute relating to the possibility of occurrence of DR (Table 7). The rules predict the number of cases with and without DR, based on the individual attributes and also

the combination of different attributes. These rules show 80% of the cases were without DR when the attribute EXU7 $\leq$ 0.01029 and 10% of the case were with DR when the EXU7 $\geq$ 0.136501. 27.1% of the cases are with no DR when the EXU7 $\leq$ 0.01029 and MAα=0.5 $\leq$ 18. 7.8% of the cases are no DR when the EXU7 $\leq$ 0.01029 and the value of MAα=0.5 $\leq$ 18. 17.5% of the cases are with DR when the value of the EXU7 $\leq$ 0.01029 and the value of the MAα=0.5 is between 18 and 38. 14.7% of the cases are with DR when the value of EXU7 $\leq$ 0.01029 and MAα=0.5 is between 46 and 61, etc. (Table 8). Moreover, this system also evaluates the most possible prediction using the combination of the attributes. This process involves several dimensions including splitting criterion, stopping rules, branch condition, etc. In this study, an accuracy of 72.71% on training dataset and 72.77% on the prediction dataset was achieved (Table 8) in classifying the occurrence/non-occurrence of DR. By this model AUC of 0.749 was obtained (Figure 6).

Table 7

Rules of the attributes generated by regression tree classifier in diagnosis of DR

DR (Pred)	Rules
0	If EXU7 $\leq$ 0.01029 then CLASS LABEL = 0 in 80% of cases
1	If EXU7 (0.01029, 0.136501] then CLASS LABEL = 1 in 10% of cases
1	If EXU7 > 0.136501 then CLASS LABEL = 1 in 10% of cases
0	If EXU7 $\leq$ 0.01029 and MAœ=0.5 $\leq$ 18 then CLASS LABEL = 0 in 27.1% of cases
1	If EXU7 $\leq$ 0.01029 and MAœ=0.5 (18, 38] then CLASS LABEL = 1 in 17.5% of cases
0	If EXU7 $\leq$ 0.01029 and MAœ=0.5 (38, 46] then CLASS LABEL = 0 in 7.8% of cases
1	If EXU7 $\leq$ 0.01029 and MAœ=0.5 (46, 61] then CLASS LABEL = 1 in 14.7% of cases
1	If EXU7 (0.01029, 0.136501] and EXU3 $\leq$ 9.74204 then CLASS LABEL = 1 in 6.3% of cases
0	If EXU7 (0.01029, 0.136501] and EXU3 (9.74204, 19.4151] then CLASS LABEL = 0 in 2.0% of cases
1	If EXU7 (0.01029, 0.136501] and EXU3 > 19.4151 then CLASS LABEL = 1 in 1.7% of cases
0	If EXU7 > 0.136501 and EXU5 $\leq$ 0.275714 then CLASS LABEL = 0 in 0.3% of cases
1	If EXU7 > 0.136501 and EXU5 > 0.275714 then CLASS LABEL = 1 in 9.7% of cases
1	If EXU7 $\leq$ 0.01029 and MAœ=0.5 (18, 38] and MAœ=0.7 MAœ=0.8 $\leq$ 16 then CLASS LABEL = 1 in 1.2% of cases
0	If EXU7 $\leq$ 0.01029 and MAœ=0.5 (18, 38] and MAœ=0.7 MAœ=0.8 > 16 then CLASS LABEL = 0 in 16.3% of cases

1	If EXU7 ≤ 0.01029 and MAœ=0.5 (38, 46] and MAœ=0.9 ≤ 35 then CLASS LABEL = 1 in 2.7% of cases
0	If EXU7 ≤ 0.01029 and MAœ=0.5 (38, 46] and MAœ=0.9 > 35 then CLASS LABEL = 0 in 5.1% of cases
1	If EXU7 ≤ 0.01029 and MAœ=0.5 (46, 61] and EXU4 ≤ 0.747636 then CLASS LABEL = 1 in 8.8% of cases
0	If EXU7 ≤ 0.01029 and MAœ=0.5 (46, 61] and EXU4 > 0.747636 then CLASS LABEL = 0 in 5.9% of cases
1	If EXU7 ≤ 0.01029 and MAœ=0.5 > 61 and EXU1 ≤ 34.5626 then CLASS LABEL = 1 in 7.6% of cases
0	If EXU7 ≤ 0.01029 and MAœ=0.5 > 61 and EXU1 (34.5626, 66.7516] then CLASS LABEL = 0 in 4.6% of cases
1	If EXU7 ≤ 0.01029 and MAœ=0.5 > 61 and EXU1 > 66.7516 then CLASS LABEL = 1 in 0.7% of cases
1	If EXU7 (0.01029, 0.136501] and EXU3 ≤ 9.74204 and MAœ=0.9 ≤ 41 then CLASS LABLE = 1 in 3.4% of cases
0	If EXU7 (0.01029, 0.136501] and EXU3 ≤ 9.74204 and MAœ=0.9 (41, 45] then CLASS LABEL = 0 in 0.7% of cases
1	If EXU7 (0.01029, 0.136501] and EXU3 ≤ 9.74204 and MAœ=0.9 > 45 then CLASS LABLE = 1 in 2.2% of cases
0	If EXU7 (0.01029, 0.136501] and EXU3 (9.74204, 19.4151] and OPTDIA ≤ 0.089971 then CLASS LABEL = 0 in 0.5% of cases
1	If EXU7 (0.01029, 0.136501] and EXU3 (9.74204, 19.4151] and OPTDIA (0.089971, 0.100454] then CLASS LABEL = 1 in 0.3% of cases
0	If EXU7 (0.01029, 0.136501] and EXU3 (9.74204, 19.4151] and OPTDIA > 0.100454 then CLASS LABEL = 0 in 1.2% of cases
1	If EXU7 ≤ 0.01029 and MAœ=0.5 > 61 then CLASS LABEL = 1 in 12.9% of cases

Table 8

Classification of dataset based on regression tree

Confusion matrix –Training set				
From \ to	0	1	Total	Accuracy (%)
0	241	38	279	86.38
1	123	188	311	60.45
Total	364	226	590	72.71
Confusion matrix –Prediction set				
From \ to	0	1	Total	Accuracy (%)
0	159	66	225	70.67
1	41	127	168	75.60
Total	200	193	393	72.77
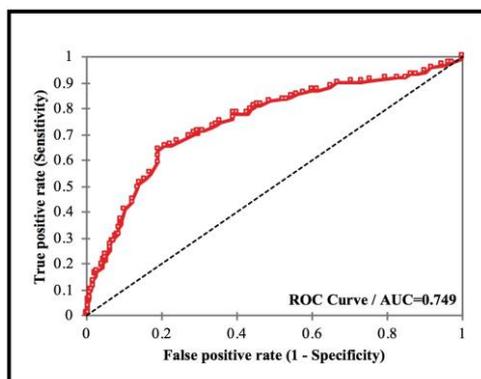
Figure 6
ROC for regression tree model

Next, we performed SVM classification, which is a supervised learning method and belongs to family of linear classification. It is one of the powerful methods in classification where a decision boundary is created through which the class labels are predicted from the feature vectors. SVM is good at recognizing patterns in complex datasets. However, there is no particular "best" kernel to recognize the patterns. The only way to select the best kernel is by trial and error [30]. We used linear Kernel and preprocessed by rescaling. We performed data validation using 150 samples for better fitting of the model since the model resulted in over-fitting. In this model, we obtained the accuracy of 74.32% on training set, 70.23% on prediction set (Table 9) and 64.67% on validation set. Performance metrics from this classifier indicated the sensitivity (recall) and specificity (Table 10). AUC was 0.809 with SVM model (Figure 7).

Table 9
Classification of dataset based on SVM

Confusion matrix –Training set				
From \ to	0	1	Total	Accuracy (%)
0	201	27	228	88.16
1	86	126	212	59.43
Total	287	153	440	74.32
Confusion matrix –Validation set				
From \ to	0	1	Total	Accuracy (%)
0	55	7	62	88.71
1	46	42	88	47.73
Total	101	49	150	64.67
Confusion matrix –Prediction set				
From \ to	0	1	Total	Accuracy (%)
0	170	19	189	89.95

| 1 | 98 | 106 | 204 | 51.96 |
| Total | 268 | 125 | 393 | 70.23 |

Table 10
Performance metrics (Class label of DR 0 / 1)

Statistic	Training set (%)	Validation set (%)
Accuracy	0.743	0.647
Precision	0.700	0.545
Recall	0.882	0.887
F-score	0.781	0.675
Specificity	0.286	0.280
FPR	0.714	0.720
Prevalence	0.457	0.367
NER	0.518	0.413
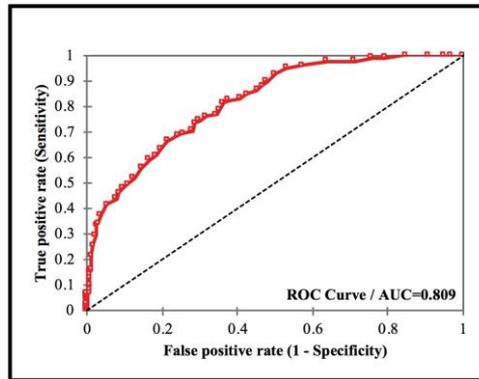
Figure 7
ROC for support vector machine

Logistic regression, a machine learning algorithm and a kind of linear regression classification model, for predictive analysis based on probability that depends on the linear measurement of the samples. It is the most commonly used statistical classification when the dependent variable is dichotomous, i.e., either positive or negative. This regression model inspects the bond between the independent and dependent variables of binary outcome and is extensively used in applications like medical and biomedical research, to predict the outcome. The logistic regression equation was used to estimate the possibility of specified consequence [31]. In the present study, we found that the logistic regression showed an accuracy of 74.41% for training set and 73.28% for the prediction set (Table 11). Here, we obtained AUC value of 0.764 (Figure 8), signifying 76.4% of the chance this model can distinguish correctly between those with or without DR. Our results show a higher AUC representing good performance of the model.

Table 11
Accuracy from logistic regression classification

Confusion matrix –Training set				
From \ to	0	1	Total	Accuracy (%)
0	234	49	283	82.69
1	102	205	307	66.78
Total	336	254	590	74.41
Confusion matrix –Prediction set				
From \ to	0	1	Total	Accuracy (%)
0	164	32	196	83.67
1	73	124	197	62.94
Total	237	156	393	73.28
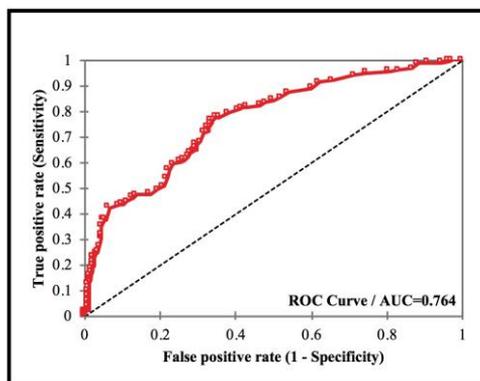
Figure 8
ROC for logistic regression model

Finally, we examined Naïve Bayes classification model for analyzing the dataset. Naïve Bayes is a supervised machine learning algorithm that classifies the observations based on the instructions set by the algorithm itself. Naïve Bayes is one of the efficient and effective classifier that works on the principle of Bayes theorem. Naïve Bayes has proven to be robust and simple probabilistic and was known for its best performance in the classification of medical data [32]. Compared to other classifiers this classification was found to be more effective computationally, where a small training dataset is sufficient enough for more accurate prediction of disease [33]. It was also reported to diagnose the disease just like a physician, considering the available attributes for the prediction analysis [34]. In this classification, the system was initially trained with the set of inputs from the training dataset that were further refined and classified for the prediction dataset. Here, we used 10-fold cross validation to predict the classes from the prediction dataset. Therefore, in the present study, 590 cases were considered as training dataset (Table 12) and the remaining cases were taken as test dataset.

Based on this classification system, the class label results predicted 212 cases without DR, while 181 cases were with DR. The first 10 predicted results in each class were shown as representation for all the test cases from the dataset (Table 13). In this study, Naïve Bayes method showed 83.56% of accuracy on training set and 80.15% of accuracy on prediction set in determining the occurrence of DR. Here, we obtained AUC value of 0.816 (Figure 9). When comparing the results shown in Table 13 to Table 5, we found some inconsistencies from the classifications. For example, PredObs1 and PredObs3 were classified to Class 0 by K-NN but were group to Class 1 by Naïve Bayes classifier. By contrast, PredObs2 was classified to Class 0 by both classifiers.

Table 12

Naïve Bayes classification on training set

Confusion matrix –Training set				
From \ to	0	1	Total	Accuracy (%)
0	257	22	279	92.11
1	75	236	311	75.88
Total	332	258	590	83.56
Confusion matrix – Prediction set				
From \ to	0	1	Total	Accuracy (%)
0	167	33	200	83.50
1	45	148	193	76.68
Total	212	181	393	80.15

Table 13

Prediction accuracy from Naïve Bayes classification on test dataset

Class	0	1
Objects	212	181
	PredObs2	PredObs1
	PredObs4	PredObs3
	PredObs5	PredObs9
	PredObs6	PredObs10
	PredObs7	PredObs11
	PredObs8	PredObs13
	PredObs12	PredObs14
	PredObs15	PredObs18
	PredObs16	PredObs25
	PredObs17	PredObs28
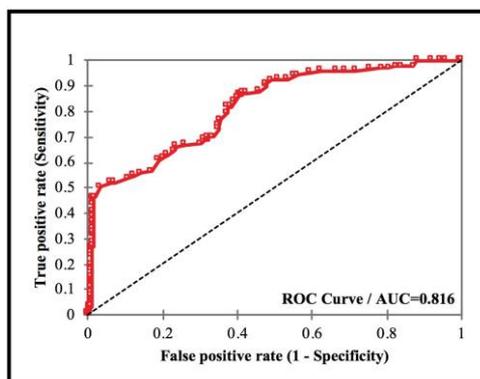
Figure 9
ROC for Naïve Bayes model

Among all the six classifiers that were studied thoroughly in the current study, our results suggested that the Naïve Bayes model of classification displayed the best accuracy followed by logistic regression model (Table 14).

Table 14
Accuracy achieved by different classification models

No.	Classification model	Accuracy (%)
1	K-NN	71.25
2	Random forest	71.76
3	Regression tree	72.77
4	SVM	70.23
5	Logistic regression	73.28
6	Naïve Bayes	80.15

Conclusions

Diabetic retinopathy is the main cause of blindness for patients suffering from diabetes mellitus. In spite of the fact that early identification of retinal images for the disease symptoms have been proposed and could prevent or delay its occurrence, the approach falls short, due to the limited availability of human expertise and lack of infrastructure to detect DR manually. Nevertheless, data mining serves as an essential tool to carry out classification and diagnose the disease. In the present study, we assessed the Messidor DR dataset, employed various classification models and evaluated their prediction accuracy in diagnosis of DR. We determined the significant role of attributes individually and in combination with other attributes crucial in the development of DR. Naïve Bayes performed best, among the six classifiers, in terms of accuracy and performance in evaluation.

In the future, we aim to develop deep learning algorithms to automatically pre-screen images for the diagnosis of DR.

References

[1]   M. M. Nentwich and M. W. Ulbig: Diabetic retinopathy-ocular complications of diabetes mellitus, World Journal of Diabetes, Vol. 6, No. 3, Apr 2015, pp. 489-499

[2]   Y. Zheng, M. He and N. Congdon: The worldwide epidemic of diabetic retinopathy, Indian Journal Ophthalmology, Vol. 60, Sep-Oct 2012, pp. 428-431

[3]   S. K. Lynch and M. D. Abramoff: Diabetic retinopathy is a neurodegenerative disorder, Vision Research, Vol. 139, Oct 2017, pp. 101-107

[4]   J. L. Leasher, R. R. A. Bourne, S. R. Flaxman, J. B. Jonas, J. Keeffe, N. Naidoo, K. Pesudovs, H. Prince, R. A. White, T. Y. Wong, S. Resnikoff and H. R. Taylor: Global estimates on the number of people blind or visually impaired by diabetic retinopathy: a meta-analysis from 1990-2010, Diabetes Care, Vol. 39, Nov 2016, pp. 2096-2096

[5]   K. C. Jordan, M. Menolotto, N. M. Bolster, I. A. T. Livingstone and M. E. Giardini: A review of feature-based retinal image analysis, Expert Review of Ophthalmology, Vol. 12, No. 3, Mar 2018, pp. 207-220

[6]   J. Amin, M. Sharif and M. Yasmin: A review on recent developments for detection of diabetic retinopathy, Scientifica, Vol. 2016, May 2016, pp. 1-20

[7]   M. Dubow, A. Pinhas, N. Shah, R. F. Cooper, A. Gan, R. C. Gentile, V. Hendrix, Y. N. Sulai, J. Carroll, T. Y. Chui, J. B. Walsh, R, Weitz, A. Dubra and R. B. Rosen: Classification of human retinal microaneurysms using adaptive optics scanning light ophthalmoscope fluorescein angiography, Investigative Ophthalmology and Visual Science, Vol. 55, Mar 2014, pp. 1299-1309

[8]   L. Dai, R. Fang, H. Li, X. Hou, B. Sheng, Q. Wu and W. Jia: Clinical report guided retinal microaneurysm detection with multi-sieving deep learning, IEEE Trans. on Medical Imaging, Vol. 37, No. 5, Jan 2018, pp. 1149-1161

[9]   E. M. Kohner, I. M. Stratton, S. J. Aldington, R. C. Turner and D. R. Matthews: Microaneurysms in the development of diabetic retinopathy (UKPDS 42), Diabetologia, Vol. 42, issue 9, Sep 1999, pp. 1107-1112

[10]   M. V. Jimenez-Baez, H. Marquez-Gonzalez, R. Barcenas-Contreras, C. Morales-Montoya and L. F. Espinosa-Garcia: Early diagnosis of diabetic retinopathy in primary care, Colombia Medica, Vol. 46, Jan-Mar 2015, pp. 14-18

[11]   B. Corcostegui, S. Duran, M. O. Gonzalez-Albarran, C. Hernandez, J. M. Ruiz-Moreno, J. Salvador, P. Udaondo and R. Simo: Update on diagnosis and treatment of diabetic retinopathy: a consensus guideline of the working group of ocular health (Spanish Society of Diabetes and Spanish Vitreous and Retina Society), Journal of Ophthalmology, Vol. 2017, Jun 2014, p. 10

[12]   A. K. Pujari: Data Mining Techniques, Universities Press, India, Feb 2001, p. 44

[13]   M. Ilayaraja and T. Meyyappan: Mining medical data to identify frequent diseases using Apriori algorithm, in proc. of Int. Conf. on IEEE Int. Symp. on Pattern Recognition, Informatics and Mobile Engineering, Salem, Apr 2013, pp. 194-199

[14]   J. C. Prather, D. F. Lobach, L. K. Goodwin, J. W. Hales, M. L. Hage, and W. E. Hammond: Medical data mining: knowledge discovery in a clinical data warehouse, in proc. of Int. Conf. AMIA Annu Fall Symp. 1997, pp. 101-105

[15]   T.-H. Cheng, C.-P. Wei and V. S. Tseng: Feature selection for medical data mining: Comparisons of expert judgment and automatic approaches, in Proc. of Int. Conf. on 19th IEEE Int. Symp. on Computer-Based Medical Systems, Salt Lake City, UT, 2006, pp. 165-170

[16]   W. Xue, Y. Sun and Y. Lu: Research and application of data mining in traditional Chinese medical clinic diagnosis, in Proc. of Int. Conf. on IEEE Int. Symp. On Signal Processing, Beijing, 2006, pp. 1-4

[17]   S. M. Khan, R. Islam and M. U. Chowdhury: Medical image classification using an efficient data mining technique, on Proc. of Int. Conf. on Machine Learning and Applications, Louisville, Kentucky, USA, 2004, pp. 397-402

[18]   P.-H. Tang and M. Tseng: Medical data mining using BGA and RGA for weighting of features in fuzzy k-NN classification, in Proc. of Int. Conf. on Machine Learning and Cybernetics, Hebei, 2009, pp. 3070-3075

[19]   S. Balakrishnan, R. Narayanaswamy, N. Savarimuthu and R. Samikannu: SVM ranking with backward search for feature selection in type II diabetes databases, in Proc. of Int. Conf. on IEEE Int. on Systems, Man and Cybernetics, Singapore, 2008, pp. 2628-2633

[20]   UCI, Diabetic Retinopathy Debrecen dataset, [Online]. Available:https://archive.ics.uci.edu/ml/datasets/Diabetic+Retinopathy+De brecen+Data+Set, Accessed on. Sep 2018

[21]   G. Kovács and A. Hajdu: Extraction of vascular system in retina images using averaged one-dependence estimators and orientation estimation in hidden Markov random fields, in Proc. of Int. Conf. on IEEE Int. Symp. on Biomedical Imaging: From Nano to Macro, Chicago, IL, USA, Mar-Apr 2011, pp. 693-696

[22]   B. Antal, A. Hajdu, Z. Maros-Szabo, Z. Torok, A. Csutak and T. Peto: A two-phase decision support framework for the automatic screening of digital fundus images, Journal of Computational Science, Vol. 3, Sep 2012, pp. 262-268

[23]   B. Antal and A. Hajdu: An ensemble-based system for microaneurysm detection and diabetic retinopathy grading, IEEE Trans. on Biomedical Engineering, Vol. 59, No. 6, Jun 2012, pp. 1720-1726

[24]   B. Nagy, B. Harangi, B. Antal and A. Hajdu: Ensemble-based exudate detection in color fundus images, in Proc. of Int. Conf. on 7th Int. Symp. on Image and Signal Processing and Analysis, Dubrovnik, Croatia, Sep 2011, pp. 700-703

[25]   T. Fawcett: An introduction to ROC analysis, Pattern Recognition Letters, Vol. 27, Jun 2006, pp. 861-874

[26]   A. Yousefpour, H. N. A. Hamed, U. H. H. Zaki, K. A. M. Khaidzir: Feature subset selection using mutual standard deviation in sentiment mining, in Proc. of Int. conf. on IEEE Int. on Big Data and Analytics, Kuching, Malaysia, Nov 2017, pp. 13-18

[27]   P. Rani: A review of various KNN techniques, International Journal for Research in Applied Science & Engineering Technology, Vol. 5, Aug 2017, pp. 1174-1179

[28]   E. Goel and E. Abhilasha: Random forest: A review, International Journal of Advanced Research in Computer Science and Software Engineering, India, Vol. 7, Jan 2017, pp. 251-251

[29]   K. A. Kaur and L. Bhutani: A Review on Classification Using Decision Tree, International Journal of Computing and Technology, Vol. 2, Feb 2015, pp. 42-46

[30]   S. Huang, N. Cai, P. P. Pacheco, S. Narandes, Y. Wang and W. Xu: Applications of support vector machine (SVM) learning in cancer genomics, Cancer Genomics & Proteomics, Canada, 2018, pp. 41-51

[31]   S. R. N. Kalhori, M. Nasehi, X.-J. Zeng: A logistic regression model to predict high risk patients to fail in tuberculosis treatment course completion, IAENG International Journal of Applied Mathematics, Vol. 40, No. 2, May 2010

[32]   I. Rish: An empirical study of the naïve bayes classifier, in Proc. of Int. Conf. on International Joint Conferences on Artificial Intelligence, Watson Research Center, Hawthorne, Jan 2001, pp. 41-46

[33]   K. M. Al-Aidaroos, A. A. Bakar and Z. Othman: Naïve Bayes variants in classification learning, in Proc. of Int. Conf. on IEEE Int. on Information Retrieval & Knowledge Management, Shah Alam, Selangor, May 2010, pp. 276-281

[34]   K. M. A1-Aidaroos, A. A. Bakar and Z. Othman: Medical data classification with Naïve Bayes approach, Information Technology Journal, Vol. 11, No. 9, 2012, pp. 1166-1174

Fuzzy Behavior Description Language:
A Declarative Language for
Interpolative Behavior Modeling

Imre Piller, Szilveszter Kovács

University of Miskolc, Department of Information Technology,
Miskolc-Egyetemváros, 3515, Miskolc Hungary,
e-mail: piller@iit.uni-miskolc.hu, szkovacs@iit.uni-miskolc.hu

Abstract: The behavior-based system (BBS) is a hierarchical structure built upon behavior components, behavior coordination and behavior fusion. The goal of this paper, is to recall the concept of the interpolative fuzzy behavior-based system and to introduce a declarative language especially designed for supporting its implementation and configuration into embedded applications. The suggested Fuzzy Behavior Description Language (FBDL) aids the definition of fuzzy rule-based systems and their connections to form behavior components and behavior coordination as fuzzy state-machines. The suggested language also assists the fuzzy rule definition with variable consequent, to help the creation of behavior fusion functions. For simplifying the definition of hierarchical rule-bases, the structure of rule-base dominancy is also introduced in the FBDL. According to the suggested embedded application concept, the FBDL code, as a parameter configuration, can directly "run" on a built in fuzzy state machine controller, called "FRI Behavior Engine". This case the behavior of the agent controlled by the FRI Behavior Engine, can be directly modified by changing the FBDL code, without reprogramming other parts of the agent controller software.

Keywords: Behavior Based Control; Fuzzy Rule Interpolation; Fuzzy State Machine; Declarative Language; FBDL: Fuzzy Behavior Description Language

1 Introduction

The behavior-based system (BBS) is a hierarchical structure built upon independent and parallel behavior components, and a behavior coordination, which can determine the usefulness (weights) of the behavior components in handling a given situation. The task of the behavior fusion is the combination of the behavior component actions to form the control action of the BBS. A behavior component could be a reactive (stateless) function, e.g. in case of a fuzzy BBS, a fuzzy rule-based system, or a compound behavior (a whole BBS itself). In common sense the function of the behavior fusion could be a convex combination

of the behavior component actions according to the corresponding component weights, but practically to be able to handle alternative, or contradictive component actions, it could be a (fuzzy) rule-based system again. The role of the behavior coordination is the situation awareness, the determination, which behavior component action with what level is required in handling a given situation. Considering the BBS to be a model based reflexive agent, the behavior coordination should have states, i.e. in case of an interpolative fuzzy BBS, it must be an interpolative fuzzy state machine.

There are numerous adaptations of the BBS concept [1] [2]. The suggested Fuzzy Behavior Description Language (FBDL) follows the fuzzy rule-based systems, fuzzy rule interpolation (FRI) [12] and fuzzy state machines based adaptation [3] [4] [5] [6] [7] [8]. See for example Fig. 1 (see Section 3 for notation in details).
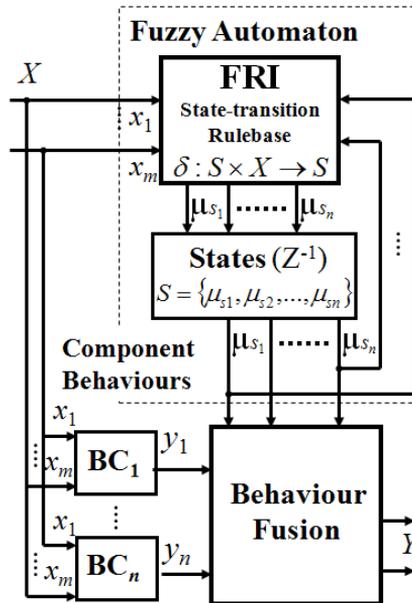
Figure 1

The adapted Fuzzy Rule Interpolation (FRI) based BBS structure with Fuzzy Automaton acting as Behavior Coordination, Behavior Components (BC) and Behavior Fusion

The goal of this paper is to introduce a declarative language especially designed for supporting embedded behavior-based applications, by offering a common framework for defining a BBS (the behavior components, the behavior coordination and the behavior fusion) as an interpolative fuzzy rule-based system, as Fuzzy Rule Interpolation based BBS (FRI BBS).

The interpolative fuzzy rule-based knowledge representation is an important issue in the suggested methodology. The rule-based structure makes the knowledge representation human readable and self-explanatory. The fuzziness, and its

"Linguistic Term" fuzzy set concept even strengthen the human readability in case of variables defined on continuous universes. The applied Fuzzy Rule Interpolation (FRI) reasoning methodology, simplifies the fuzzy rule-base definition by relaxing some constraints of the fuzzy rule-base, the FRI can handle sparse fuzzy rule-bases too [32].

The unified rule-based human readable system construction FRI BBS framework makes the suggested FBDL to be a suitable platform to support the implementation of human (expert) knowledge to a working system. The FBDL code can "run" on a system directly or, having some additional observed data, can serve as an object for machine learning parameter optimization methods. Some examples for FRI BBS applications are appearing in ethological model based human-robot interaction applications [9] [10] [11] and Ethorobotics [33], e.g. for expressing human readable emotions [34] for robots. In case of ethological models, the expert's knowledge is based on real animal observations and represented as a descriptive verbal model built upon a series of facts and action-reaction rules. Moreover, the verbal models can be incomplete, or contain some expert domain specific implicit knowledge, which is missing from the verbal description, e.g. as "well-known" facts. These requirements fit well the fuzzy rule-based knowledge representation and the FRI reasoning of the suggested FBDL description. See the concept of FRI with some application examples more detailed in [12].

For creating the FBDL the goal was forming a language which supports function definition by linguistic rules similar like fuzzy systems. This case the fuzzy rules act like "fuzzy points" of a fuzzy function and the fuzzy reasoning method acts as a fuzzy function definition. In classical fuzzy reasoning, the fuzzy rule-base has to be complete (i.e. they need fully defined rule-base (e.g. the Zadeh-Mamdani-Larsen Compositional Rule of Inference (CRI) (Zadeh [13]) (Mamdani [14]) (Larsen [15]) or the Takagi-Sugeno fuzzy inference (Sugeno [16], Takagi-Sugeno [17])). For releasing this condition, to be able to handle sparse fuzzy rule-base (where not all the possible rules are defined), the concept of the FRI was adapted. In case of FRI, the fuzzy reasoning method is a fuzzy interpolation, where the fuzzy rules are the fuzzy node points of the fuzzy interpolating function. (An axiomatic approach of the fuzzy interpolation can be found in [18]).

There are numerous FRI methods that exist in the literature. For the FRI Behavior Engine implementation any of them is adaptable, which can handle multidimensional antecedent spaces. For the current implementation [31], because of its simplicity, the FRI "FIVE" [19] [20] [21] [22] were adapted.

2   Related Works

The current work is the continuation of [23]. The Fuzzy Behavior Description Language and the FRI Behavior Engine have been improved but the aims remained unchanged. A framework was developed, where an expert could define the behavior of an agent. The applied FBDL is technically a declarative programming language. Therefore, it is necessary to make some comparison with the available agent based modeling methods too. This section briefly summarizes some works related to other declarative behavior definition languages and modeling methods.

For dynamic declarative agent configuration in [24], the authors show a plan generation mechanism in a declarative manner. They emphasize the importance of adaptation of the agent. The approach is the same as the role of fuzzy interpolation in the suggested FRI BBS concept, i.e. the agent should be able to succeed in unseen situations.

Other authors recognize that the usage of a simplified programming toolchain for non-technical users only viable solution for a short term [25]. For avoiding the fast increment of software complexity they introduced a new programming paradigm (*Targets-Drives-Means*). Their construction is similar to the proposed FBDL behavior description method. The main difference is the usage of fuzzy reasoning and a dedicated description language which is interpreted by the FRI behavior engine directly.

Authors in [26] also suggests that the simplified imperative way is also viable for simple behaviors. They choose a component centric approach where the reasoning system tries to find a solution which match with the requirements of the designer and the runtime system.

For agent based modeling and simulation, the SESAM [27] provides a useful visual programming environment, as UML-like activity diagrams. On the other hand, for larger models, and parameter optimization, the textual description could be more concise and practical.

The comparison of the previously mentioned behavior description methods can be seen in the Table 1.

Table 1

Comparison of AgentSpeak [24], Target-Drives-Means (TDM) [25], Agent Based Modeling (ABM) [26] and SeSAmUML [27] description languages

	AgentSpeak	**TDM**	**ABM**	**SeSAmUML**	**FBDL**
representation	textual	textual	textual	visual	textual
base language	STRIPS	-	JSON	UML	-
paradigm	declarative	TDM	declarative	procedural	declarative
inference	layered planning	score calculation	backtracking	transition rules	fuzzy interpolation
hierarchy	STRIPS operators	priority control	subcomponents	rules	dominancy

The first aspect of the comparison is the representation method of the behavior. Most of the considered languages use textual representation. It is important to note that, the textual and visual representation are always interchangeable. This comparison focuses on the primary usage of the behavior description method. It means that the choice of visual and textual representation is depends on the preferred usage regardless the implementation difficulties.

The behavior description is a special kind of knowledge representation. Therefore, in many cases the behavior description language is a refined version of a general purpose language. Three of the mentioned languages are based on STRIPS, JSON and UML languages. The TDM has an own semi-graphical definition language, while the proposed FBDL language uses a simple, natural language-like syntax.

The paradigms behind the behavior description languages show many differences. The AgentSpeak is a declarative language where the user can define operators and their pre- and post-conditions. The TDM itself is a new programming paradigm, which organizes the behaviors to small, trigger activated behavior components. The ABM proposes a fully declarative model in the sense that the JSON-like description defines the required, higher level actions instead of low-level commands. The SeSAmUML follows the standard UML method for defining a final state machine. The FBDL defines the set of production rules which have organized to rule-bases. It is also a declarative approach which is similar to the AgentSpeak because the antecedent parts of the rules are similar to the preconditions of the standard STRIPS language.

A common difficulty of the behavior description systems is how they can resolve the conflicts and contradictions of the behavior descriptions. They have to use some kind of inference for calculating the most suitable action for the given situation. The AgentSpeak solves the problem by a layered planning approach. It tries to dynamically evaluate the lower and higher level plans. The TDM calculates score for any action of its behavior components. After, it can choose the appropriate action according to these values. The ABM tries to find a proper solution which fulfils all constraints by backtracking. For the unspecified properties, it uses interpolation. The SeSAmUML uses different abstraction for the actions and for the resolution of conflicting cases. It has specific rules for selecting and terminating activities. The FBDL solves the conflicting situations by fuzzy rule interpolation. It obtains the required action by interpolating the consequent values.

For simplification purposes, it is reasonable hierarchically organize or prioritize the actions of the behavior descriptions. The last aspect of the comparison considers the preferred way of this kind of hierarchy. In the AgentSpeak language, the domain expert can use STRIPS operators to express higher priority plans. The TDM contains priority management in its model. It assigns priorities for the different behaviors and prioritize the actions after the calculation of behavior scores. The ABM organizes the behavior description to a tree of components.

Each component has constraints, therefore, the component preference can be coded into this strict hierarchical structure. The SeSAmUML solves the conflict resolution problem on the level of rules. After an action has activated it does not terminate until the terminating rule has not fired. The FBDL organizes the rule-bases into a hierarchy which automatically denotes their priorities.

3 Fuzzy State Machine Model

Having FRI models in Behavior Component definitions and FRI state transition-rules in Behavior Coordination, most part of the BBS model forms a Fuzzy State Machine [8]. There are various understandings of the fuzzy state machine can be found in the literature (for summary see [28] [29]). Most of them are extending the classical finite state automaton by applying fuzziness for the state transitions, while the state remains discrete (crisp, one of the predefined ones). On the other hand, the fuzzy model suggested for the FRI Behavior Engine adapts the concept of "fuzzy state", where the state is a vector of membership values. According to the fuzzy state concept, the system could be in all its states in the same time, but with different membership levels. This view fits well the FRI BBS concept, as the actual fuzzy state can be easily interpreted as the usefulness (weights) of the behavior components in handling a given situation. For example, the fuzzy state of the behavior coordination can directly control the behavior fusion.

The fuzzy state machine adapted for the FRI Behavior Engine is an extended version of the Fuzzy Finite-state Automaton. It extends the finite set of input symbols to finite dimensional input values and the finite set of states to finite dimensional state values [8]. This case the fuzzy state machine can be defined by a tuple:

$$\tilde{F} = (S, X, \delta, P, Y, \omega) \tag{1}$$

where S is a finite n length of fuzzy states, $S = \{\mu_{s_1}, \mu_{s_2}, \dots, \mu_{s_n}\}$, μ_{s_i} is the membership value of the ith dimension of the n dimensional fuzzy state. X is a finite m dimensional input vector, $X = \{x_1, x_2, \dots, x_m\}$. $P \in S$ is the fuzzy initial state of $\tilde{F}$. Y is a finite l dimensional output vector, $Y = \{y_1, y_2, \dots, y_l\}$. $\delta: S \times X \to S$ is the fuzzy state-transition function which is used to map the current fuzzy state into the next fuzzy state upon an input value. $\omega: S \times X \to Y$ is the output function which is used to map the fuzzy state and input to the output value. See e.g. on Fig. 2.
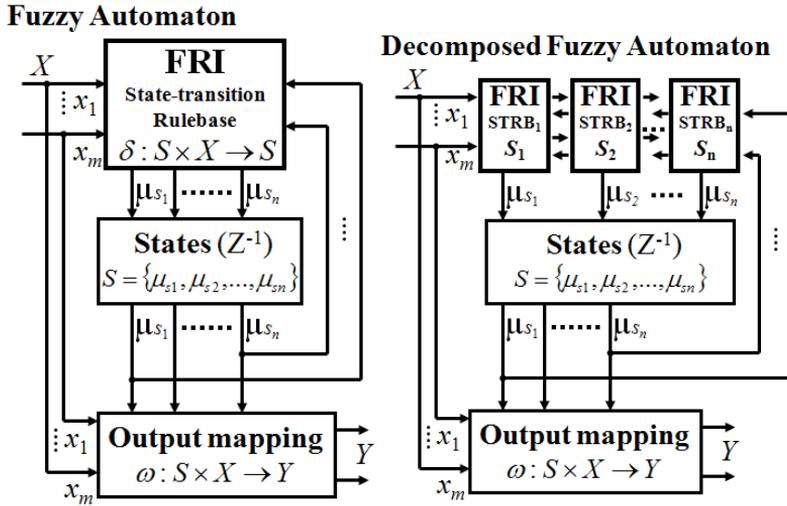
Figure 2
FRI based Fuzzy Automaton and its decomposition

In case of fuzzy rule-based representation of the state-transition function $\delta: S \times X \rightarrow S$, the rules have $n + m$ dimensional antecedent space, and n dimensional consequent space.

It is important to note, that the state S is a vector of membership values, the actual state is a point in the n dimensional unit hypercube. The state-transition rule-base moves this point in each discrete time step.

The FRI state-transition rule-base defines the state-transition function, which is a $R^{n+m} \rightarrow R^n$ mapping, having $n + m$ rule antecedent and n rule consequent dimensions. For simplifying the rule-base definition, this rule-base is decomposed to n pieces of single consequent rule-bases $R_i^{n+m} \rightarrow R_i$, $i \in [1, \dots, n]$. See e.g. in Fig. 2.

For evaluating the FRI state-transition rule-base the FRI FIVE [22] was applied. The main idea of the FIVE FRI is based on the fact that most of the control applications serves crisp observations and requires crisp conclusions from the controller. Adopting the concept of the Vague Environment (VE) [30], FIVE can handle the antecedent and consequent fuzzy partitions of the fuzzy rule-base by scaling functions [30] turning the fuzzy interpolation to crisp interpolation. The idea of a VE is based on the indistinguishability of elements. In VE the fuzzy membership function $\mu_A(x)$ is indicating level of similarity of x to a specific element a which is a representative or prototypical element of the fuzzy set $\mu_A(x)$, or, equivalently, as the degree to which x is indistinguishable from a (see e.g. on Fig. 3) [30]. Two values in a VE are ε-indistinguishable if their distance is less or equal than ε. The distances in a VE are weighted distances (Eq. 2). The weighting factor or function is called scaling function (factor) [30]:

$$\delta_s(a,b) = \left| \int_a^b s(x)dx \right| \leq \varepsilon \qquad (2)$$

where $\delta_s(a,b)$ is the scaled distance of the values a, b and $s(x)$ is the scaling function on X.
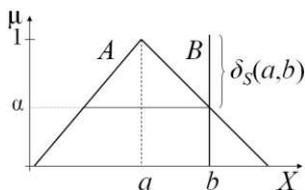
Figure 3

The α-cuts of $\mu_A(x)$ contain the elements that are $(1-\alpha)$-indistinguishable from a

If the VE of a fuzzy partition (the scaling function or at least the approximate scaling function [19] [20] [21]) exists, the member sets of the fuzzy partition can be characterized by points in that VE (see e.g. scaling function s in Fig. 4).
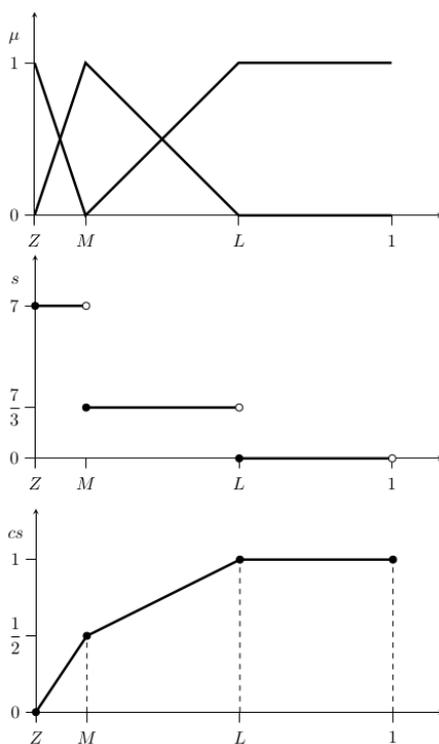
Figure 4

A "Ruspini" (0.5-covering) fuzzy partition built upon three linguistic terms,
namely the Z (Zero), M (Middle), L (Large) fuzzy sets,
its scaling function s, and its normalized cumulative scaling function cs

Having the VE concept and the scaling function based similarity calculation, any crisp interpolation, extrapolation, or regression method can be adapted very simply for FRI [7] [8] [9]. Because of its simple multidimensional applicability, in FIVE the Shepard operator based interpolation (first introduced in [35]) is adapted.

In case of singleton rule consequents (c_k) the fuzzy rule R_k has the following form:

If $x_1 = A_{k,1}$ And $x_2 = A_{k,2}$ And ... And $x_m = A_{k,m}$ Then $y = c_k$ (3)

Adapting the VE concept and the scaling function based similarity calculation to the Shepard operator based interpolation, the conclusion of the FRI can be obtained as:

$$y(x) = \begin{cases} c_k & if\, x = a_k\ for\ some\ k, \\ \left(\sum_{k=1}^r \frac{c_k}{\delta_{s,k}^\lambda}\right)\left(\sum_{k=1}^r \frac{1}{\delta_{s,k}^\lambda}\right) & otherwise, \end{cases}$$ (4)

where $\delta_{s,k}$ are normalized scaled distances:

$$\delta_{s,k} = \delta_s(a_k, x) = \sqrt{\sum_{i=1}^m \left(cs_i(x_i) - cs_i(a_{k,i})\right)^2/m}$$ (5)

$cs_i(x_i)$ is the normalized cumulative scaling function (see e.g. on Fig. 4) of s_i:

$$cs_i(x_i) = \int_0^{x_i} s_i(x_i)dx_i / \int_0^1 s_i(x_i)dx_i$$ (6)

and s_i is the ith scaling function of the m dimensional antecedent universe, x is the m dimensional crisp observation and a_k are the cores of the m dimensional fuzzy rule antecedents A_k.

According to Eq. 4 the proposed interpolation method calculates the conclusions in the following steps.

1) Determine the normalized Euclidean distances of the observations from the rules on all the antecedent universes according to Eq. 5.

2) The rule weights are calculated as the reciprocal value of the distances corrected by the Shepard power p.

3) The consequent of the rule-base is calculated as the convex combination of the rule consequences weighted by the rule weights according to Eq. 4.

Considering the implemented fuzzy state machine model to be a discrete time system, the connections between the decomposed state-transition rule-bases (see Fig. 2) can serve as the time delays (Z^{-1}) temporarily storing the state variables. This case the fuzzy state machine can be defined as a recurrent network of multiple input single output state-transition rule-bases (see e.g. on Fig. 5), where the state variables are the values of the connections between them. On Fig. 5 the OBS prefix denotes observations and the RB denotes rule-bases. The number of the input is equivalent with the number of incoming edges.
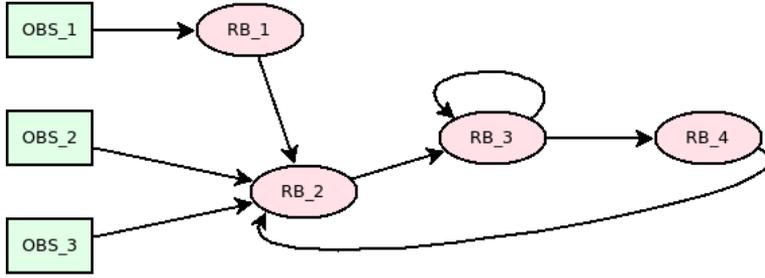
Figure 5

Example of a behavior description with recurrent connections

(OBS denotes observations, RB denotes rule-bases)

4 Behavior Fusion

Comparing the suggested FRI BBS structure (see Fig. 1) with the FRI state machine (see Fig. 2), the network of the Behavior Components and the Behavior Fusion is corresponding to the Output mapping of the Fuzzy Automaton.

A behavior component could be a reactive (stateless) function, e.g. an FRI rule-based system, or a compound FRI BBS. The problematic part is the Behavior Fusion. The function of the behavior fusion could be a convex combination of the behavior component actions according to the corresponding component weights, but to be able to handle alternative, or contradictive behavior component actions, it should be a FRI rule-based system again.

For simplifying the fusion, the continuous conclusions of the behavior components, for the FRI BBS, the application of a special fuzzy rule format is suggested. The basic rule structure is similar to the rule format applied in the Takagi-Sugeno fuzzy inference (Sugeno [16], Takagi-Sugeno [17]). In Takagi-Sugeno fuzzy inference, the consequent of a fuzzy rule is a function of the input variables. In the suggested FRI BBS, a rule of the Behavior Fusion rule-base has a consequence, which is a conclusion of another (behavior component) rule-base. This case the conclusions of the behavior component rule-bases are appearing as consequences of the behavior fusion rules.

From the viewpoint of the FBDL these are fuzzy rules with variable rule consequents, where the variable could be an observation, or a conclusion of a rule-base. The variable consequent fuzzy rule R_k has the following suggested form:

If $x_1 = A_{k,1}$ And $x_2 = A_{k,2}$ And ... And $x_m = A_{k,m}$ Then $y = y_{V_k}$ (7)

where $y_{V_k} = x_i$ is an observation, or $y_{V_k} = y_{RB_i}$ the conclusion of the ith rule-base RB_i.

5   The Syntax of the Behavior Description Language

In this section the proposed behavior description language will be introduced. It is a declarative language which has a SQL-like syntax. (The assumption behind is that the verbosity of the language and the avoidance of special characters makes easier to learn its usage for non-technical users.)

The formal specification of the language is provided by extended Backus-Naur form. By using its common notation, the following language definition has obtained.

```
behavior ::= universe+ rulebase+ [init]
universe ::= 'universe' string ['description' string]
symbol+ 'end'
symbol ::= string number number
rulebase ::= 'rulebase' string ['description' string] rules
'end'
rules ::= rule+
rule ::= 'rule' ['description' string] ['use'] string
['when' predicates] 'end'
predicates ::= predicate ('and' predicate)*
predicate ::= string 'is' string
init ::= 'init' ['description' string] (string (string |
number))+ 'end'
```

In this formalism the `string` as a terminal symbol is a string literal with quotas, for example `string literal example`. The `number` is also a literal, which describes a floating point value, for instance 12.34.

The behavior description is a `behavior` non-terminal symbol. It contains at least one universe definition and a rule-base definition. After them, the expert can define an initialization block, where the initial state of the state machine can be set.

The definition of the universe uses the `universe` keyword. It followed by the name of the universe. Its name must be unique in a behavior description. The definition (similarly to the `rulebase`, `rule` and `init` definition) can be contains an optional description. It is introduced by the `description` keyword.

The `symbol` is a named point in the enclosing universe. The expert defines all of the language variables in this way. The symbol name is followed by the associated points of the antecedent and consequent side.

As an example, let define a universe called `distance`:

```
universe "distance"
description "The distance from the target."
    "close" 0 0
```

```
    "middle" 10 0.8
    "far" 50 1
end
```

As it can be seen in the description part of the definition block, the purpose of this universe is to represent the distance from a target position. There are three symbols: close, middle and far in the defined universe. The measurement unit of the distance is not necessary for the universe. According to the definition, the 0 distance is close, the 10 unit distance, is considered as middle and 50 units is seen as far. The second parameters after the symbol names are required for scaling. It means that the distance on interval [0, 10] changes quickly, while there is only a slight difference on the interval [10, 50].

For defining rules, it is necessary to define a further universe:

```
universe "tiredness"
description "The measure of tiredness."
    "low" 0 0
    "small" 0.3 0.5
    "middle" 0.7 0.5
    "high" 1 1
end
```

The universes of distance and tiredness are sufficient for the antecedent space. For the output side let define the universe approach:

```
universe "approach"
description "The speed how the agent approaching the
target."
    "low" 0 0
    "high" 1 1
end
```

At this point there are two universes: one for the observations and one for the consequence. The rule-bases defines the connection between the inputs and the output of the defined behavior. The rule-base encloses its rules. The general form of a rule-base is the following:

```
rule <consequent-name> when <predicates> end
```

The name of the consequent must match with the name of the universe. Let assume that we would like to define the behavior which fulfils the following rules:

1) The agent approaches the target, when the target is close and the agent is not tired.

2) The agent does not approach the target, when the target is far.

3) The agent does not approach the target, when the agent is tired.

These rules can be formalized via the proposed description language by the following way:

```
rulebase "approach"
    rule "high" when
        "distance" is "close" and "tiredness" is "low"
    end
    rule "low" when
        "distance" is "far"
    end
    rule "low" when
        "tiredness" is "high"
    end
end
```

5.1   Variable Valued Consequent

The language makes possible to use a recently calculated variable value as a consequent instead of a symbol of the universe. It is notated by the `use` keyword and the rule-base name which represents the variable, for example:

```
rulebase "first" ... end

rulebase "second" ... end

rulebase "output"
    rule use "first" when "need_first" is "true" end
    rule use "second" when "need_first" is "false" end
end
```

In the suggested FRI BBS, the variable valued consequents are the tools of the behavior fusion. The behavior component rule-bases are calculating the conclusions of the behavior components as variables. Then a rule-base with the corresponding variable valued consequents fuses them as behavior fusion.

The variable valued consequent has improved the flexibility of the language significantly. In fact, it makes available the calculation of weighted summation of rule-base outputs, where the weights are also calculated according to the user defined rules.

5.2   Initialization

The behavior model is iterative; therefore it requires an initial state. The initialization is compulsory if the variable (a value of a rule-base conclusion) is used as an input. The observations are also needed to be initialized, but it must be done by the environment. For variable initialization the language uses the `init`

keyword. It contains a pair of variables (rule-base names) and values. It is possible to use both symbols and values for setting the initial values.

```
init
    "value" "low"
    "result" 20
end
```

The integrity of the initial values are also checked. It means, that the following statements must be fulfilled:

- All universe names must have corresponding universe definitions.
- The symbol values of a given universe must be in the discourse bounds of the universe.
- All input values must be defined.

The initialization is an essential part of the behavior model. The initial values, as initial state is an integral part of the fuzzy state machine model.

5.3   Evaluation of the Consequences

If we consider our fuzzy state machine to be a connectionist structure, a rule-base is a node which represents a $R^n \rightarrow R$ function.

At first the behavior engine must calculate the distance of the observations from the symbols on all the antecedent universes. Let see the following example:

```
universe "distance"
    "zero" 0 0
    "close" 1 0.1
    "far" 5 1
    "max" 10 1
end
```

In this example the distance is given directly in meter. The consequent value is on the [0,1] interval. (The normalization on consequent side is optional but a good practice in most cases.)

Above 5 meter the rule-base does not distinguish the values. All values are "far" up to 10 meter.

Let see the following rule-base:

```
rulebase "speed"
    rule "high" when "distance" is "far" and "curiosity" is
"high" end
    rule "low" when "distance" is "close" end
```

```
    rule "low" when "curiosity" is "low" end
end
```

where the appropriate universes are:

```
universe "curiosity"
    "low" 0 0
    "high" 1 1
end

universe "speed"
    "low" 0 0
    "high" 1 100
end
```

It is necessary to initialize the observations and calculate the distance of the observations from the symbols on all the antecedent universes. For instance, the values are initialized as the followings:

```
init
    "distance" 3
    "curiosity" 0.4
end
```

Consider the first rule and calculate the distance of 3 from "far" on the "distance" universe. The distance is determined from the difference of the cumulative scaling function values of the two points.

The value of the cumulative scaling function at 3 is 0.55, at "far" is 1, therefore their cumulative scaled distance is 0.45 (see Fig. 6). Similarly we can calculate the distance of 0.4 and "high" on "curiosity" interval. The result is 0.6.
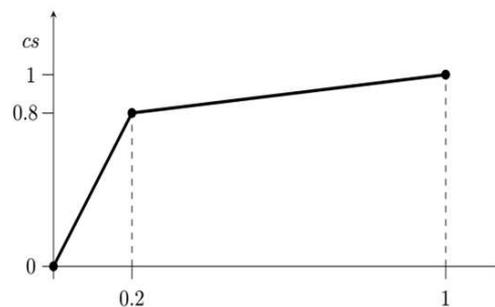
Figure 6
The non-linear scaling of the "distance" universe

The distance of the rule is given by the Euclidean norm of the distances by dimensions divided by the square of the number of antecedents. (It is necessary

because this way the distance is on [0,1] is independent from the scaling of the universe.) The rule distance for the first rule is calculated as:

$$d_{rule1} = \frac{\sqrt{0.45^2 + 0.6^2}}{\sqrt{2}} \approx 0.5303 \tag{8}$$

Similarly, the engine calculates the distance of the second and the third rule:

$$d_{rule2} = \frac{\sqrt{0.45^2}}{\sqrt{2}} \approx 0.3182, d_{rule3} = \frac{\sqrt{0.4^2}}{\sqrt{2}} \approx 0.2828 \tag{9}$$

If there is a 0 distance from some rules, then the consequent is calculated as the mean value of the corresponding consequent symbol values, or variable values.

The weights are the reciprocals of the distances on the p Shepard-power, $w_i = \frac{1}{d_i^p}$

If $p = 1$ then the weights of the rules are:

$$w = [1.8856, 3.1427, 3.5355] \tag{10}$$

The consequent values of the rules are:

$$c = [1,0,0] \tag{11}$$

The consequence can be obtained as the sum of rule consequents weighted by the rule weights:

$$C = \frac{\sum_i w_i \cdot c_i}{\sum_i w_i} \approx \frac{1.8856}{8.5638} \approx 0.2202 \tag{12}$$

The consequent value of the "speed" rule-base is approximately 22.0183.

6 Dominancy

In many cases the domain expert's heuristically descriptive verbal model can be incomplete, or can contain some expert domain specific implicit knowledge. This domain specific implicit knowledge is inherently missing from the verbal description, e.g. as "well-known" facts, supposing that these are already known (explicitly given earlier).

On the other hand, having a rule-based knowledge representation, the lack of information means some missing cardinal rules from the rule-base. Let see a simple example. We want to describe the connection between the distance and the interest. We give a single expert's rule, that the interest is high when the distance is close. This case the natural assumption is that the interest is low when the distance is far. However, this assumption is not explicitly stated as a rule. It can be deduced, as a straightforward fact, only in the case, if we take into consideration, the domain knowledge. Otherwise, this rule is meaningless. But the rule-base is formally incomplete, because we do not say anything what should happen, when

the distance is far. Hence the previously considered model with one rule results high interest value independently from the distance. The FRI (fuzzy interpolation based) behavior model designed for handling undefined situation only in case if it can be deduced from the existing knowledge. A single rule means that only one output value is given, therefore it should be chosen independently from the input value. The other example is a kind of contradiction because of the unequal priorities of the rules. The first rule is "Stay home", the second is "Go to the garden if the sunshine is moderate". This case the "Stay home" is a kind of "default" rule requesting staying home if nothing special is happening. The second rule has more specialty (priority), requesting "Go to the garden" in a special case of "the sunshine is moderate". These two rules have contradictive conclusions in the case if "the sunshine is moderate". One solution could be the extension of the contradictive rules to make them to be the same specialty e.g. "Stay home if the sunshine is not moderate". On the other hand, this solution could be problematic, if the linguistic term "not moderate" is not defined (or it is impossible to define).

As a possible heuristic to solve the above problems, we suggest the concept of rule dominancy. The suggested rule dominancy heuristic is implemented as an extension of the FBDL, which enables the expression of rule relations in the form of rule dominancy.

6.1   Dominancy of the Rules

The suggested language extension provides elements for expressing hierarchical relation between rule-bases and between rules. We can express hierarchical rule relation without language extension, only by modifying the suppressed rule antecedents in the FBDL, but usually this is a tedious task. The suggested language extension automatizes this burden. Instead of try to cover the antecedent space by hand written rules, we can save this task by defining dominancy on the rule level. For supporting this, the grammar is slightly modified in the following way:

```
rules ::= rule+ ['dominates' rules 'end']
```

The dominancy means that the user can select dominant and dominated rules. The dominant rule consequent suppresses the conclusion of the dominated rules according to the fulfilment of the dominant rule antecedent. The consequent of the dominated rule is viable in a high level only when the fulfilment of the dominant rule antecedent is low. The dominancy can be applied in a hierarchical manner.
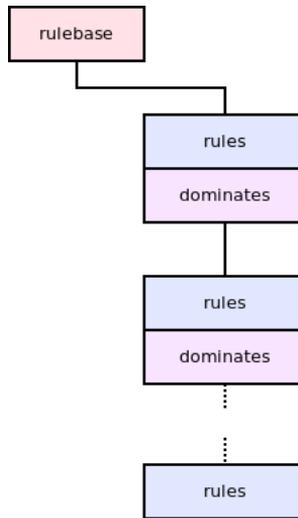
Figure 7
The semantical hierarchy of dominated rules

Let see the following example for multilevel dominancies:

```
rulebase "sample"
description "rule-base dominancy with examples"
   rule
      "low" when "stay_at_home" is "high"
   end
   dominates
      rule
         "high" when "playmates_exist" is "true"
      end
      dominates
         rule # Default rule
            "low"
         end
      end
   end
end
```

The dominancy also gives sense for defining default rules (see e.g. last rule of the rule-base "sample" above). It has no antecedent part; its conclusion is always selected. Having a default rule dominated by the others, all the undefined situations are automatically covered.

7 Sample Behavior

The following behavior shows a complete example of the previously mentioned features.

The model describes a one dimensional world, where an agent can walk between the target and the rest position. The target can make noise. The agent can move forward (to the target) and backward (to the rest position). The *interest* is the internal state of the agent. When the target makes noise, the agent interest is increasing. While the agent stays near the target, the interest value is decreasing. The measure of approaching mood to the target is also depending on the distance from the target. If the level of tiredness is high, the agent goes to the rest position. The FBDL language definition of this sample behavior is the following.

```
universe "has_noise"
description "Has the target noise?"
    "false" 0 0
    "true" 1 1
end

universe "distance"
description "The distance from the target."
    "close" 0 0
    "far" 1 1
end

universe "tiredness"
description "The measure of tiredness."
    "low" 0 0
    "high" 1 1
end

universe "speed"
description "The speed of the agent."
    "forward" 1 1
    "stop" 0 0
    "backward" -1 -1
end

universe "approach"
description "The speed how the agent approaching the
target."
    "low" 0 0
    "high" 1 1
end
```

```
universe "go_to_rest"
description "The speed how the agent go to the rest
position."
    "low" 0 0
    "high" -1 -1
end

universe "interest"
description "The measure how the agent interested about the
target."
    "low" 0 0
    "high" 1 1
end

rulebase "interest"
description "The agent is interested when the target has
noise."
    rule "high" when "has_noise" is "true" end
    rule "high" when "interest" is "high" end
    rule "low" when "distance" is "close" end
    dominates
        rule "low" end
    end
end

rulebase "approach"
description "Approach when there is a closer interesting
target."
    rule "high" when "distance" is "far" end
    dominates
        rule "low" end
    end
end

rulebase "go_to_rest"
description "Go to rest position when the target is tired."
    rule "high" when "tiredness" is "high" end
    dominates
        rule "low" end
    end
end

rulebase "speed"
```

```
description "The speed fusioned from approach and go_to_rest
values."
    rule use "approach" when "interest" is "high" end
    rule use "go_to_rest" when "tiredness" is "high" end
    dominates
        rule "stop" end
    end
end
```
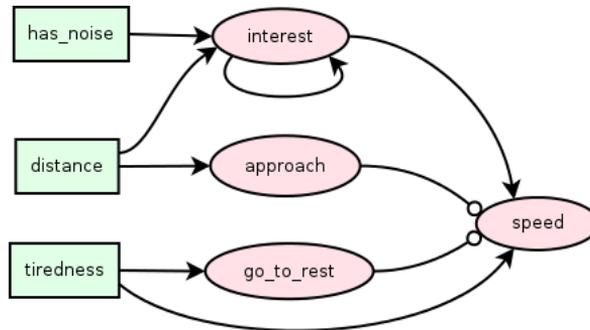

Figure 8
The behavior description of the agent

The behavior description can be checked by testing the following scenario in the simulation framework [31]:

1) Move the agent to the rest position by setting the *distance* value to 1.

2) Make noise by setting the *has_noise* observation to 1 and after back to 0. As a result, the interest value remains high, however the noise already stopped.

3) The agent speed is 1 which means that the agent wants to go to the target. Simulate the motion by setting the *distance* value to 0. At that point the interest level has decreased and the speed is near 0.

4) Set the tiredness to 0.5 - The agent now wants to move to the rest position. Start to increase the *distance* value, supposing the agent is moving away from the target. Near 0.55 the agent stops, because the interest level and the tiredness are close to each other.

5) Set the tiredness value to 0.8 - The *speed* value becomes negative, therefore the agent wants to go to the rest position. By increasing the *distance* further, the direction of speed remains, which means that the agent goes back to the rest position.

8   JavaScript Library and Framework

According to the suggested concept, the FBDL code is "running", as a parameter configuration, directly on a FRI state machine called "FRI Behavior Engine". In the current implementation, the FRI Behavior Engine was written in JavaScript. The instantiated FRI Behavior Engine loads and interprets the FBDL code and according to the fetched parameters provides interface for accessing its functions. For the implementation of the FRI Behavior Engine the JavaScript version was chosen. It makes testing easier and can run in a browser without installation. Moreover, it can be easily embedded to an online simulation environment too.

The FBDL web page, the FBDL library, the online FBDL simulator and their sources are available at [31].

The behavior engine is a software component which is able to evaluate the inputs based on the defined rules and calculates the internal state and the output, which will be available via its interface.

The engine instance can be applied according to the following sample:

```
engine.set("observation", 10);
engine.step();
var distance = engine.get("result");
```

The *set* method in the first line sets the *observation* value to 10. After, the *step* method calculates the new state. The variable values in the new state are accessible via the *get* method.

Conclusion and Future Work

The suggested Fuzzy Behavior Description Language (FBDL) is an easy-to-learn declarative language for non-programmer users, to define Fuzzy Rule Interpolation based Behavior-based System (FRI BBS) applications. Its inference model is based on the Fuzzy Rule Interpolation. Among the basic parameter and topology configuration the suggested FBDL also supports the concept of variable valued consequent and rule dominancy definitions.

The FBDL describes a fuzzy state machine having a massively parallel structure. The rule-bases and even the rules of the rule-bases can be evaluated parallel independently from each other. This parallel structure can be implemented on a massively parallel architecture speeding up the evaluation time and supporting the implementation of more complex models.

The FBDL is also serves embedded applications, where the FRI Behavior Engine is implemented on an embedded system. Instead of storage consuming look-up tables, or time consuming calculation of complex models, the FBDL and the FRI Behavior Engine can be an effective way for implementing complex behavior models. Moreover, the modification of the agent behavior can be simply achieved by modifying the FBDL description.

The proposed behavior description method, basically targets the non-programmer domain experts, where the definition of complex behaviors is a requirement from behavioral related studies or agent development reasons. The proposed Behavior Description Language and its Behavior Engine can be control physical and virtual robots, where there is a feedback from the environment and the actions are available. The FBDL description is in fact a special kind of configuration. It can be applied where the user has predefined inputs and outputs for the system. For instance, it is appropriate for the high level configuration, of IoT devices. The presented FBDL language can be regarded as a structured, semi-natural way of automatized robot control.

Future research will focus on the complexity of the defined behaviors. There is an assumption that the quantitative metrics of the behavior description, shows a strong correlation with the complexity of the modeled behavior. Its validation requires more experimentation, within the various types of behaviors.

Acknowledgement

References

[1]   Arkin, R.C. et al.: *Behavior-based robotics*. MIT press, 1998

[2]   P. Pirjanian, P.: *Behavior Coordination Mechanisms - State-of-the-art*. Techreport IRIS-99-375, Institute for Robotics and Intelligent Systems, School of Engineering, University of Southern California, October, 1999, 49 p.

[3]   Kovács, Sz. et al.: *Behaviour based techniques in user adaptive kansei technology*. In: Proceedings of the VSMM2000, 6th international conference on virtual systems and multimedia, october. 2000, pp. 3-6

[4]   Kovács, S.: *Interpolative fuzzy reasoning and fuzzy automata in adaptive system applications*. In: Proceedings of the IIZUKA. 2000, pp. 1-4

[5]   Kovács, Sz.: *Fuzzy behaviour-based control techniques in adaptive system applications*. In: Proceedings of the ieee international conference on computational cybernetics, ICCC, 2003, pp. 29-31

[6]   Kovács, Sz., Kóczy, L. T.: *Application of interpolation-based fuzzy logic reasoning in behaviour-based control structures*. In: 2004 ieee international conference on fuzzy systems (ieee cat. no. 04CH37542) IEEE, 2004, pp. 1543-1548

[7]   Kovács, Sz.: *Interpolative fuzzy reasoning in behaviour-based control*. In: Computational intelligence, theory and applications. Springer, 2005, pp. 159-170

[8]   Kovács, Sz.: *Fuzzy rule interpolation in embedded behaviour-based control*. In: 2011 IEEE international conference on fuzzy systems (Fuzz-IEEE 2011) IEEE, 2011, pp. 436-441

[9]   Kovács, Sz. et al.: Interpolation based fuzzy automaton for human-robot interaction. *IFAC Proceedings Volumes*, **42** (16) 2009, pp. 317-322

[10]   Kovács, Sz. et al.: *Ethologically inspired robot behavior implementation*. In: 2011 4th international conference on human system interactions, hsi 2011. IEEE, 2011, pp. 64-69

[11]   Vincze, D. et al.: A novel application of the 3d virca environment: Modeling a standard ethological test of dog-human interactions. *Acta Polytechnica Hungarica*, **9** (1), 2012, pp. 107-120

[12]   The Fuzzy Rule Interpolation (FRI) web page, the FRI-Toolbox, the Sparse Fuzzy Model Identification Toolbox and their sources can be found at: http://fri.uni-miskolc.hu/

[13]   Zadeh, L. A.: Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Transactions on systems, Man, and Cybernetics*, (1), 1973, pp. 28-44

[14]   Mamdani, E., Assilian, S.: An experiment in linguistic synthesis with a fuzzy logic controller. *International journal of human-computer studies*, **51** (2), 1999, pp. 135-147

[15]   Larsen, P. M.: Industrial applications of fuzzy logic control. *International Journal of Man-Machine Studies*, **12** (1), 1980, pp. 3-10

[16]   Sugeno, M.: An introductory survey of fuzzy control. *Information sciences*, **36** (1-2), 1985, pp. 59-83

[17]   Takagi, T., Sugeno, M.: *Fuzzy identification of systems and its applications to modeling and control*. In: Readings in fuzzy sets for intelligent systems. Elsevier, 1993, pp. 387-403

[18]   Perfilieva, I.: Fuzzy function as an approximate solution to a system of fuzzy relation equations. *Fuzzy sets and systems*, **147** (3), 2004, pp. 363-383

[19]   Kovács, Sz.: *New aspects of interpolative reasoning*. In: Proceedings of the 6th international conference on information processing and management of uncertainty in knowledge-based systems, Granada, Spain, 1996, pp. 477-482

[20]   Kovacs, Sz., Koczy, L. T.: *Approximate fuzzy reasoning based on interpolation in the vague environment of the fuzzy rulebase*. In: Proceedings of IEEE international conference on intelligent engineering systems. IEEE, 1997, pp. 63-68

[21]   Kovács, Sz., Kóczy, L. T.: The use of the concept of vague environment in approximate fuzzy reasoning. *Fuzzy Set Theory and Applications, Tatra Mountains Mathematical Publications, Mathematical Institute Slovak Academy of Sciences, Bratislava, Slovak Republic*, **12**, 1997, pp. 169-181

[22]   Kovács, Sz.: *Extending the fuzzy rule interpolation' five' by fuzzy observation*. In: Computational intelligence, theory and applications. Springer, 2006, pp. 485-497

[23]   Piller, I. et al.: *Declarative language for behaviour description*. In: Emergent trends in robotics and intelligent systems. Springer, 2015, pp. 103-112

[24]   Meneguzzi, F., Luck, M.: Declarative planning in procedural agent architectures. *Expert Systems with Applications*, **40** (16), 2013, pp. 6508-6520

[25]   Berenz, V., Suzuki, K.: Targets-drives-means: A declarative approach to dynamic behavior specification with higher usability. *Robotics and Autonomous Systems*, **62** (4), 2014, pp. 545-555

[26]   Borenstein, D. B.: A composite constraints approach to declarative agent-based modeling. *arXiv preprint arXiv:1503.08880*, 2015

[27]   Klügl, F. et al.: *From simulated to real environments: How to use sesam for software development*. In: German conference on multiagent system technologies. Springer, 2003, pp. 13-24

[28]   Doostfatemeh, M., Kremer, S. C.: New directions in fuzzy automata. *International Journal of Approximate Reasoning*, **38** (2), 2005, pp. 175-214

[29]   Stamenković, A. et al.: Different models of automata with fuzzy states. *Facta Universitatis, Series: Mathematics and Informatics*, **30** (3), 2015, pp. 235-253

[30]   Klawonn, F.: Fuzzy sets and vague environments. *Fuzzy Sets and Systems*, **66** (2), 1994, pp. 207-221

[31]   The FBDL web page, the FBDL library, the online FBDL simulator and their sources can be found at: http://fbdl.iit.uni-miskolc.hu/

[32]   Kóczy, L. T., Hirota, K.: *Interpolative reasoning with insufficient evidence in sparse fuzzy rule bases*, Information Sciences, **71**, 1992, pp. 169-201

[33]   Miklosi, A., Korondi, P., Matellan, V., Gacsi, M.: *Ethorobotics: A New Approach to Human-Robot Relationship*, Frontiers in Psychology, **8** Paper: 958, 2017, p. 8

[34]   Lakatos, G., Gácsi, M., Konok, V., Brúder, I., Bereczky, B., Korondi, P., Miklósi, Á.: Emotion Attribution to a Non-Humanoid Robot in Different Social Situations, PLOS ONE, **9**, 12, Paper: e114207, 2014, p. 32

[35]   Shepard, D.: *A two dimensional interpolation function for irregularly spaced data*, Proc. 23rd ACM Internat. Conf., 1968, pp. 517-524

Reducing Energy Consumption with IoT Prototyping

Mika Saari*, Pekka Sillberg*, Jere Grönman*, Markku Kuusisto*, Petri Rantanen*, Hannu Jaakkola*, Jaak Henno**

*Tampere University, Faculty of Information Technology and Communication Sciences, Pohjoisranta 11A, 28101 Pori, Finland, mika.saari@tuni.fi, pekka.sillberg@tuni.fi, jere.gronman@tuni.fi, markku.kuusisto@tuni.fi, petri.rantanen@tuni.fi, hannu.jaakkola@tuni.fi

**Tallinn Technical University, School of Information technologies, Ehitajate tee 5, 19086 Tallinn, Estonia, jaak.henno@ttu.ee

Abstract: Nowadays, energy consumption and especially energy saving, are topics of great importance. Recent news regarding global warming has increased the need to save energy. In Finland, one of the major sources of energy consumption is housing. Furthermore, the heating of residential buildings accounts for up to 68% of housing energy consumption. Therefore, it is not surprising that apartment energy consumption and ways to save energy in housing are a popular research topic in Finland. In this paper, two different research areas are introduced: First, a literature survey is presented on the research subjects of energy saving in the area of real estate and housing. The goal is to gain overall knowledge of the current state of energy saving research. The overall conclusion is that knowledge of energy consumption improves efforts toward energy saving. Second, rapid prototyping with off-the-shelf devices and open source software are described. These devices are cheap to install, and a wide range of sensors are available. Consequently, it is important to deal with these topics together. The former studies provide knowledge about the usage of open hardware, open software, and open architectures with the development of prototype systems for gathering data. The literature survey gives us new information on the specialties of energy consumption measuring, offering a new area for modeling and developing prototype systems. These experiences will be taken forward and utilized in energy saving and environmentally sustainable solutions, such as Green Computing.

Keywords: IoT; Prototyping; Energy saving

1   Introduction

In the modern world, energy saving has become an important issue, in almost every aspect of life. Global warming is forcing people to search for low-energy solutions. It is important to be aware of the living comfort when thinking about the

low-energy solutions. For example, most people want the living temperature to be comfortable - not too low and not too high. Furthermore, the awareness of one's energy consumption has been proven to reduce overall energy usage. Thus, in the context of this paper, the research problem can be formulated as:

How to reduce energy consumption by collecting and serving suitable data?

For this problem, we are looking for a solution for two questions.

1) How to categorize the energy consumption related studies?

2) How to utilize free and open solutions in the energy consumption context preserving adequate living conditions?

In our use cases, we are especially looking for solutions that utilize open-source components and open hardware, architectures and interface specifications. This study belongs to the Internet of Things (IoT) research area and to studies focusing on Wireless Sensor Networks (WSN). In addition, one of the focus areas of this paper is rapid prototyping in the IoT world by using off-the-shelf devices. An example of rapid prototyping method was described by [1] for the automotive industry.

This paper introduces the application architectures and system models for IoT prototyping. Furthermore, sensors and sensor networks that collect data into the cloud are discussed, and more specifically, wireless sensor network (WSN) systems that can be utilized in testing data collection in rapid prototyping are of interest. In our use cases, the prototypes are built using off-the-shelf devices and tools. Additionally, Green ICT (Information and Communication Technology) should be part of the developing process when either the goal is to save energy or make systems which help to save energy.
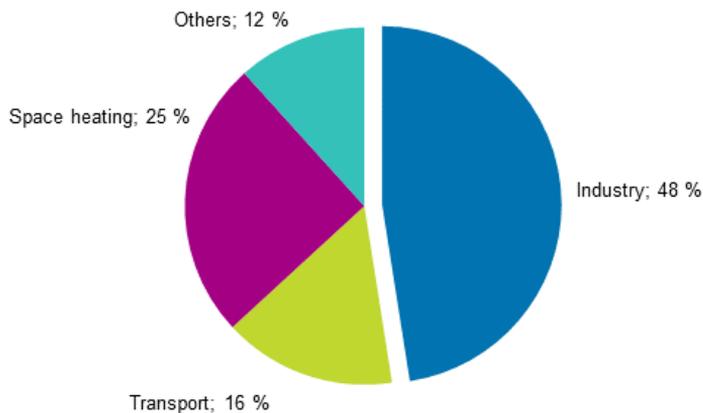
Figure 1
Finland's energy consumption by sectors in 2018 [2]

According to official statistics, collected and published by Statistics Finland, energy in Finland is produced mostly in three ways: wood, oil, and nuclear fission. These three sources combined add up to 66% of the energy produced in Finland. Various other sources of energy production include but are not limited to: coal, gas, water, peat, and wind. In their report, Statistics Finland [2] profile the Finnish energy consumption as shown in Fig. 1: Industry uses the most energy (48%) while heating comes in second place with 25% of energy consumed. Traffic is also a major consumer with 16% of the total energy used in Finland. Other sources then add up to the remaining 12%.

This research is focused on Finland (and further applicable in other northern countries), in which energy is often used for heating, instead of cooling (as is common in many other countries). The *"Cold weather raised energy consumption in housing in 2016"* report by Statistics Finland [3] shows that heating residential buildings consumed 46 TWh of energy in Finland during 2016. Furthermore, the heating of residential buildings was reported to account for up to 68% of the total energy consumption of housing with the second largest consumer of energy being heating water, accounting for 15%. Other notable energy consumers in Finnish households were electrical appliances, saunas, and lighting. The most common source of energy for heating was electricity, at 34%. The next most common source of energy was district heating (29%) and the third most common heating energy source was wood, at 22%, followed by heat pumps, at 9%. The usage of heat pumps in Finland has grown significantly since the start of the millennium because of their efficiency, saving energy and money compared to direct heating sources. All together, these four sources of energy made up about 95% of the energy used for heating in Finland. The remainder was mostly heating oil at approx. 5%, with other technologies accounting for less than 1%.

Our former research focus has been IoT and prototyping. This preliminary research will show how existing studies could be applied to a new research area. The structure of paper follows the research process: Section 2 includes a brief introduction to studies related in energy consumption. Section 3 continues with further analysis and categorization of energy consumption papers. Section 4 will present our studies and those findings, which could be combined with energy consumption monitoring. Further, the combined ideas of reducing energy consumption and prototype developing are introduced. Finally, Section 5 concludes the study.

2   Related Studies in Energy Consumption

This section deals with studies related to energy consumption. One important point of view is the awareness aspect of energy consumption. In [4], it was found that dormitory residents reduced electricity consumption when exposed to real-

time visual feedback and incentives. This study examined electricity and water usage. In the study, two dormitories were equipped with automated monitoring systems that provided high-resolution, real-time feedback. The study showed that the residents' awareness, knowledge, and behavior regarding energy saving improved after they were provided with relevant information and exposed to campaigns.

The study [5] examined the effects of energy saving, by analyzing the changes in the awareness and behavior of apartment residents after the promotion of energy-saving activities and their proper usage, and the provision of relevant information. In this study, the questionnaire included topics such as energy awareness and the knowledge and practice of energy conservation. In addition, this study performed an additional survey, which was conducted for women who were given energy-saving information and asked to participate in energy-saving activities after submitting the initial questionnaire. The results showed that energy-saving behavior improved after being provided with relevant information.

In the third study [6], the focus was on the meaning of comfort and comfort practices, barriers to and motivators for saving energy, and knowledge about the heating system. Data were collected from social housing tenants and university staff using surveys, interviews, and monthly energy meter readings. This study showed that warmth was mentioned most often as the meaning of comfort. In addition, comfort practices were to a large extent defined as temperature-related actions that were low in energy consumption. This study also found that willingness to change behavior was the greatest when the motivation was to save money.

The study [7] focused on energy-saving awareness, by using In-Home Display (IHD) devices. These devices provide real-time data about the use of electricity in specific appliances. Also, the costs of these devices were shown, and the users had the opportunity to reduce their electricity consumption. The result of this study was that the direct feedback provided by IHDs encouraged consumers to make more efficient use of energy. In addition, active IHD users were able to reduce their electricity consumption by about 7%, on average.

All these studies show that knowledge of energy consumption improves efforts toward energy saving.

3   Literature Survey

The introduction posed the research question: how to categorize the energy consumption related studies? To answer this research question, a literature review was performed, in order to map the existing knowledge in this domain.

3.1   Research Approach

The literature review used the Systematic Literature Review (SLR) method for collecting relevant primary studies and followed the guidelines given by Kitchenham and Charters [8]. For the SLR, an electronic literature search was executed. The databases used were IEEE Xplore Digital Library (IEEE) and Google Scholar. The survey was started by using the main search term: "Energy consumption". During the pilot study and related research [4-7], several other research terms arose such as "Temperature comfort", "Learning temperature comfort", "Apartment temperature comfort", "Smart home communication", "Real-time energy consumption monitoring," and "Energy apartment sensor". With a combination of these keywords, a good coverage of potential studies was obtained. The target amount of related studies was a total of fifty publications, as this amount would provide enough information for categorization and determination of research trends. Of these fifty publications a small number of papers were selected, which were considered to include the most relevant papers for the energy consumption or energy savings.

3.2   Categories for Existing Studies

To get an overview of the existing studies, the papers included in the study were analyzed for common topics. Most of the papers were relatively distinctive in terms of research objective, methodology, and application. Ultimately, based on the analysis of the research papers, we selected four categories taking into consideration the variations in research themes. The reason for choosing a relatively small amount of categories was to enable the examination of the details of research papers falling under the same category systematically. Selecting too many categories would have made it difficult to compare the trends or research methodologies. It is worth noting that some of the papers could be classified into more than one category. The research categories identified from the source material are:

- Comfort

- Retrofitting

- Network APIs

- IoT

The categories are listed according to the importance of the background research. The category 'Comfort' contains studies that discuss the basic elements for living comfort, which are often considered to be more important than energy saving. In general, comfort is an important aspect of energy saving. Too much saving means that the comfort of the living environment, such as thermal comfort and humidity, decreases. The most important factor is thermal comfort, which is taken into

account in several studies [9-14] in this category. Most of the research addressed previous studies, but [9] in particular reviewed thermal comfort research work and discussed the implications for the energy efficiency of buildings.

In our use cases, focus is on existing building stock and therefore the 'Retrofitting' category contains the research on applications or solutions installed in existing buildings. A different approach is used for monitoring energy consumption monitoring in new buildings and old buildings. In new buildings, monitoring applications and systems are included in the design phase of the building. For example, the heating system could be selected by weighing up the energy aspects. In old buildings, the main structure (e.g. the heating system) already exists, and the monitoring must fit this structure. This category consists of studies [15-20] where the presented application or solution was installed in existing buildings.

The study [15] focused on the problems of buying or renting a house. The potential purchaser or renter of the property does not know its living comfort factors such as temperature and lighting. This study introduced IoT sensors for the evaluation of the comfort levels of real estate properties. Another study [16] focused on studying and determining the cost-optimal renovation measures to decrease both the supplied and primary energy consumption of the building. This study encouraged apartment building owners to conduct thorough renovations toward nearly zero-energy apartment buildings.

The third category focuses on Application Programming Interfaces (APIs) and other methods that allow remote control or management of devices over networks. In addition, devices including a network API can provide (web) services usable by application developers or by client devices. A RESTful API is an architectural style for communications used in web service development, which was mentioned in [21] although the usage was not described in detail. The second study [22] present four RESTful services: one developed in Arduino and three mobile applications. A third study [23] integrated smart power outlets into the web and facilitated the development of extensions and novel features. They were implemented in a web user interface and a mobile phone interface for demonstration purposes. In addition, this was confirmed with a 12-month pilot deployment.

The study [24] described the construction of a smart outlet network as a system for automated energy-aware services utilizing humidity, temperature and light sensors, and motion sensor data. The sensors were installed on smart outlets and the appliances were under policy-based automatic control. This study also presented the deployed system in real-life environments.

The last category, 'IoT', includes the studies which do not fit in any of the other categories, but are nevertheless related to our focus area. This category is the widest and most of the papers could be included in it. Therefore, this research only introduces studies which: (*i*) collect the data in some way; (*ii*) save the data; and (*iii*) the saved data are then used or processed.

The survey [25] explored state-of-the-art control systems in buildings. The ref. [26] focused on intelligent control systems for energy and comfort management in smart energy buildings. The study [27] presented the wireless, smart comfort sensing system that they developed. This system consists of sensor nodes, which send data to a sink node that sends data to a PC. Another, lower-cost implementation was presented and discussed in [28], describing the hardware IoT infrastructure providing real-time monitoring in multiple school buildings. The sensor nodes and gateway node were based on Arduino boards or similar. A further study [29] also used low-cost devices in their HVAC and sensor system. IoT is also discussed in several studies [10], [13] and [21], which have been mentioned above.

Table 1
Breakdown of the papers reviewed

Category	Selected Studies	Author's Studies	Number of Studies
Comfort	[9-14]	[30]	7
Retrofitting	[15-20]		6
Network APIs	[21-24]	[30-36], [38]	12
IoT	[10], [13], [21], [25-29]	[30-31], [33], [37-40]	15

The results of the literature survey and the selected categories (Comfort, Retrofitting, Network APIs and IoT) can be seen in Table 1. The table also shows how the authors' own contribution related to the categories.

4   Prototype Systems and Models

This section gives a brief summary of our earlier studies related to rapid prototype development. The proof-of-concept demonstrations and prototype applications have been developed to illustrate how to utilize cost-effective, open, and modular solutions. The studies have been chosen based on their potential for including methods or technologies that could be transferred or exploited in the energy consumption monitoring or energy saving context.

4.1   Rapid Prototyping

In the context of rapid prototype development (and in the context of IoT devices in general), a working solution for gathering data needs:

- Hardware – a device or devices running the software

- Software to work with the data – collect, save, and transmit

- Technologies–choosing the right technologies for a use case makes things easier for both the developer and the user.

In our use cases, the prototype development has had more of software than hardware orientation. Data are gathered with embedded software, which controls the action of sensor devices. The data transfer to the cloud can be made in various ways and requires applicable software to control the sending and receiving of data. The WSN and sensor networks have several possible technologies for data transfer, for example: Ethernet [30], WiFi, ZigBEE [31] and LoRa [32]. In addition, power saving algorithms for WSN [33] and network topology related issues of Portable Fog Gateways [34] can be considered important topics.

The prototype systems gathered data which was saved to cloud-based services. In a basic example, the cloud service could be implemented with a Linux-based server and database [35], which has been modeled in [36].

Software development was carried out in several areas: data gathering software, data processing software, visualization of results, etc. The software development consisted of small-to-medium sized applications written in C/C++, Java, JavaScript or Python. The operating systems were generally chosen from the Open Source selection. For example, the Raspberry Pi is usually equipped with Linux–based operating systems (e.g. the Debian-based Raspbian). Also other software, such as databases, communication and web server software, was typically Open Source software.

Hardware development can be an integral part of prototype system development, but in our use cases the prototypes used off-the-shelf devices. In the past few years the price of microcontrollers, small computers and sensors has become much lower. At the same time, more and more features have been added to the off-the-shelf devices. These factors have made utilizing off-the-shelf devices both cheaper and easier, and it has also reduced the need to construct (or design) sensor or device packages from the ground-up using basic electronic components. Often used off-the-shelf devices include:

- Smartphones and tablets

- Single-board computers: Raspberry Pi, Beagle Bone, Intel Galileo, etc.

- Single-board microcontroller: Arduino Uno

- Sensors: Heat, humidity, pressure, movement, position, etc.

Using these off-the-shelf devices for the manufacturing and up scaling the number of prototype devices is more rapid than implementing a prototype based on printed circuit board design. In addition, the Raspberry Pi has been shown to be good choice for research projects and is a widely used device [37].

Furthermore, nowadays mobile phones have the ability to act as sensor devices. Even the basic Android smartphone has several of the following sensors: light, proximity, camera, microphone, touch, position (GPS, WiFi, Cellular), accelerometer, gyroscope, pressure, temperature, humidity. The data collection and processing can be handled in a smartphone. In addition, a basic smartphone

usually has more than adequate communication features: Bluetooth, WiFi, GSM, GPRS, 3G, 4G, etc. are often available.

4.2 Data Gathering with Sensor Network–Modeling, Piloting, and Testing

The sensor networks can be modeled as is illustrated in Fig. 2 [31]. The sensor nodes gather data and send it without processing to the master node. The master node may validate the received data, it may also process it, and send the data to the cloud. The data are usable from the cloud for various purposes.
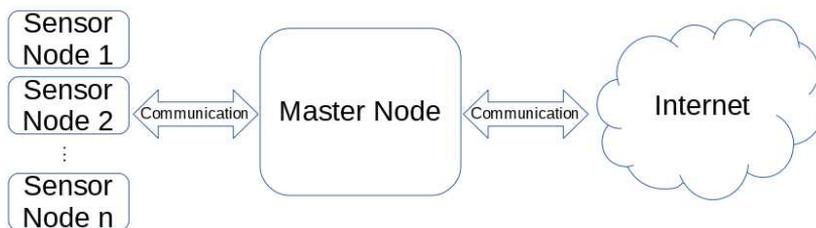
Figure 2
Basic model of sensor network [32]

This model was tested during the study [31], and a proof-of-concept solution was implemented and presented. Based on a survey of prototyping solutions that utilize Raspberry Pi the commonly used solutions were observed to adhere to this basic model even when no specific model was described in the studies [37]. However, the model shown in Fig. 2 has to be modified if smartphones are used as sensor nodes. Fig. 3 shows a combined presentation of the sensor node and master node model.
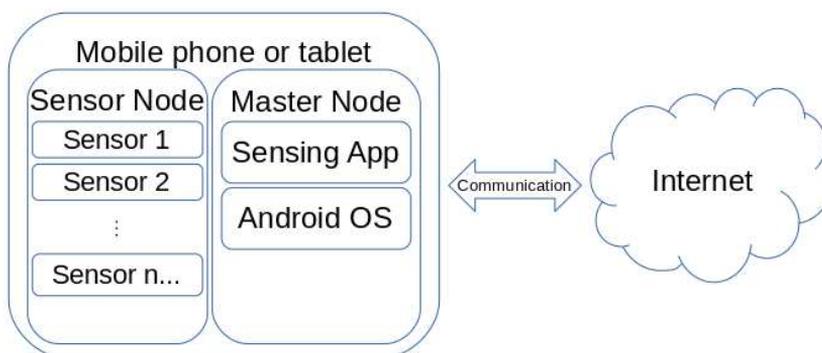
Figure 3
The combined sensor node—master node model for data gathering [36]

The model presented in Fig. 3 was developed especially for data collection with smartphones. The smartphone includes the necessary sensors, data storage, and communication channels for the data gathering prototype system. In addition, the Android operating system (OS) was used, which has enough capabilities to gather and store data. Also, the commonly used communication protocols are directly supported by the APIs provided by the OS. [36]

The studies [30-31] [36-37] show several important results:

- Study [30] introduced an example of how a cost-efficient single-board computer (SBC) can be used to gather sensory data, and how this data can be provided to the client over the public Internet. In addition, the use of standard protocols makes development easier, but not all development boards support all standards (in this case the I2C protocol).

- Study [31], mentioned that master nodes often have access to a constant power source, but one should carefully choose which components to use in remote sensors to minimize power consumption. In addition, most of the energy is consumed in the wireless transmission of data and consequently it is important to only send what is required (optimization of the nodes). The energy consumption issue was handled more specifically by [33].

- The survey about prototyping with Raspberry Pi was introduced in [37]. This paper shows that there is a lack of formalized approaches, methods, and tools in the research studies. Often only a single use case and a single system are described in the paper with a minimal use of testing practices and methods. The commonly used testing methods are software testing, software performance testing, and validation of data tests.

The conclusion from the results of the papers [30-31] [35-36] is that rapid prototyping with off-the-shelf devices is possible, but requires guidelines that include an architecture model of components—both software and hardware.

4.3   Prototype System: Road Condition Analysis and Visualization

Nowadays, almost everyone has a mobile phone and even the most basic smartphones often come embedded with a variety of sensors. In [38], smartphones were utilized to collect road condition data. The smartphone application developed during the research collects data from the phone's built-in sensors. The application can be installed in a common Android smartphone. This collected data could be further refined into more specific data, such as reports of bumps in the road, uneven road surfaces, roadworks, and so on. The data are sent to the cloud where they are processed. Fig. 4 shows the visualization of the captured data and the routes where the data were collected.
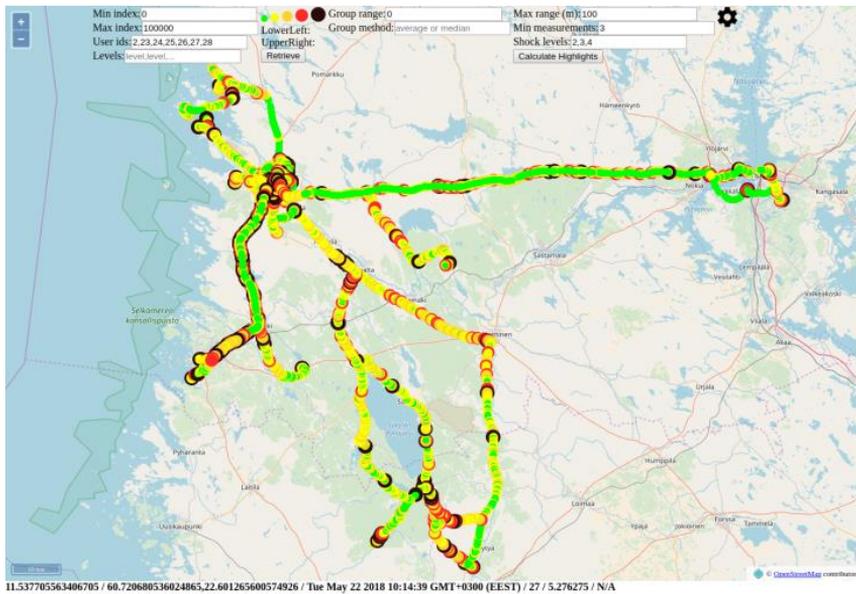
Figure 4
Visualization of the routes driven [38]

This research shows that it is possible to use a commonly used consumer product for data collecting. However, it turned out that, even though modern smartphones/devices are fairly similar by functionality, hardware differences can cause unexpected problems for implementation. Further, the embedded sensors are often not "calibrated" across devices and manufacturers. This can cause variances in the results and therefore comparison of data can be difficult if accuracy is of high concern. In addition, non-system-related effects and interference (environmental factors) may affect the final results e.g., when measuring shocks or vibrations different vehicles provide slightly different results. In addition, it is often necessary to perform pre-processing and filtering on-device, versus a fully service-implemented analysis.

A further result of this research is connected to the visualization of collected data. This is often no minor issue when measuring the quality of the user experience. Also, a fluent execution of visualization of a large dataset can be challenging, especially on a web browser.

4.4   Prototype: Approach to Image Data Collection

Customer complaints can be resolved by means of image and data collection. The research [39] introduces two prototypes installed in vehicles and a cloud service for autonomous collection of data. The first prototype—an Android application—

was implemented for a smartphone to take pictures of a bus as it approaches the bus stop. The second prototype was implemented for the Raspberry Pi single-board computer by using off-the-shelf devices such as a camera, GPS sensor, and 3G/4G wireless modem. The prototype was installed in a garbage truck to take pictures of recycling areas, as shown in Fig. 5.

Figure 5

Three pictures taken of a recycling area. Left: in daylight; center: at night; right: blocking obstacle [40]

The prototypes use a camera and GPS. The collected data—picture, location, time, etc.—were sent to the cloud server. The paper [39] discusses the differences and challenges faced in designing and implementing the two prototypes for different platforms.

The main conclusions were that mobile platforms (i.e., smartphones, tablets) can work as a quick starting point for rapid prototyping. These have embedded sensors, proper documentation, and the availability of examples, all of which support rapid prototyping. On the other hand, small computers like Raspberry Pi and microcontrollers offer a better option for use cases requiring remote management. Of course this has disadvantages, such as requiring more "hands-on" labor, and being more difficult to find examples or production quality code. In addition, both mobile platforms and small computers highlight the importance of environmental factors—such as the availability of electricity, telecommunications, and installation of the prototype [39].

4.5   Prototype: Counting Passengers from Image Data

The research [40] was the result of a real-life need for counting passengers. In the summer of 2018 a large public event was organized in the city of Pori, Finland. The event had free-to-ride buses and the organizer wished to collect statistics about the bus passengers: Where they got in and where they got out. The use case utilized cost-effective and off-the-shelf components such as the Raspberry Pi 3 computer, position sensors, and cameras. In this use case, the software used was Open Source Computer Vision Library version 3.

Figure 6
An example of the detection area of the bus, as seen by the device [40]

During the research, a prototype system was developed, consisting of hardware and software components. The prototype takes pictures, as shown in Fig. 6. The pictures are processed by the system, which was based on image analysis and shape detection. The data are processed in the Raspberry Pi and the results of the processed data are sent to the cloud server. [40]

4.6 Toward Reducing Energy Consumption with IoT Prototyping

An important part of achieving energy usage reductions is a reliable way of collecting data about current environmental conditions. The research presented in this section (Section 4) illustrated simple models that could be used when implementing a sensor network for collecting data. Furthermore, Section 4.2 illustrated certain pitfalls related to currently used approaches and highlighted the lack of existing model for rapid prototyping in the IoT domain. Sections 4.3 and 4.4 showed advantages of using smartphones as tools for data collection. Modern smartphones contain a huge variety of built-in sensors and the available devices range for low-cost affordable models to more expensive high-end devices. Today, almost everyone already has a smartphone, and thus, the cost of using smartphones for environmental monitoring can be negligible. Additionally, even the low-end devices are capable of running simple applications, that can be used to show statistics about current living conditions, and at least in theory, to provide the user with interfaces for controlling the environment. Unfortunately, there are challenges related to installing devices to real-life scenarios, such as, creating solid, durable packaging for the sensors and the availability of electricity and

telecommunications. Specifically when dealing with rapid prototype development and actual locations, there can be unexpected challenges, even when not considering the interoperability issues with existing structures and systems. More advanced scenarios can be realized with customizable devices. Section 4.5 described how Raspberry Pi could be used to monitor passenger ridership, an approach that could be easily expanded to energy consumption domain. Detecting whether rooms or buildings are occupied can have huge effect on the cooling and heating requirements. Furthermore, all of the presented prototypes use free and open software and low-cost modular components proofing that rapid prototyping with off-the-shelf devices is possible.

Conclusions

One of the initial research questions for this study was "How to categorize the energy consumption related studies?" Based on the literature review carried out, the existing studies can be roughly divided into four distinct categories: studies related to measuring and ensuring occupant comfort in buildings; research on how to extend existing systems with modern sensor and optimization solutions (retrofitting); studies on the usage and description of network-based APIs; and studies on IoT-based devices in general. All of these categories—comfort, retrofitting, network APIs, and IoT—include a wide array of existing research and provide numerous examples of applications and systems for monitoring and optimizing energy consumption. Several conclusions can be drawn from the results of the literature review, and from our previous experience in prototype development in the various research projects presented in this paper.

Our second research question was "How to utilize free and open solutions in the energy consumption context preserving adequate living conditions?" In the scope of this paper, the solutions for this question answered more on the basic technical problems. The paper gave insights on available software and hardware options, but the aspect of preserving living conditions was given less focus, and would require more extensive research.

In existing studies, IoT often consists more of "proof-of-concept" style research. The studies present a use case, various testing methods, and results, but often no formal model for testing or benchmarking is described. Without further studies it is difficult to say why there is an apparent lack of a standardized or *de facto* model for rapid IoT prototype development, but research on developing such a model or applying an existing model for the IoT context could be one potential direction for future studies.

Mobile devices (i.e., smartphones, tablets) can work as a good starting point for prototype development—they are ubiquitous, and they come embedded with various built-in sensors. Documentation and application examples are, in general, easy to find, and the utilization of mobile devices can be combined with off-the-shelf devices to create more complex systems. Off-the-shelf products—such as the Raspberry Pi single-board computer and wide multitude of available sensors—

have become much cheaper in recent years and offer adequate performance with a relatively good set of features and expansion capabilities. The market has also seen an increase in cheaper commercial sensor products targeted at consumers (end users). This price and market development has caused an increase in research utilizing cost-effective off-the-shelf devices as opposed to building and designing devices (e.g., sensor nodes) from the "ground up". Additionally, the increase in commercial products has enabled people with lesser technical knowledge to buy and set up sensor devices in their homes. Unfortunately, the interoperability of existing systems (air conditioning systems installed in older buildings, commercial products lacking proper interfaces or APIs, etc...) is often less than seamless and connecting the systems to available off-the-shelf devices can be challenging. With more barebone devices (Raspberry Pi, Arduino, etc.), packaging, designing a case, and installing the sensor node in a real-life environment or for outdoor use can pose further difficulties.

Finally, the paper attempted to answer the question: "How to reduce energy consumption by collecting and serving suitable data?" Based on the existing studies, the availability of energy consumption information can have a huge effect on people's habits, and properly presented usage statistics can lead to energy savings. In existing systems, the information is often limited to simple statistics (numerical details, graphs). Unfortunately, meaningful visualization can be challenging: How to select what is "meaningful"? How detailed should the statistics be? And how should the information be presented? In some cases, the user cannot affect the energy consumption and occupant comfort as desired. The user may not have access to the building's air conditioning or the building may not have devices capable of altering the indoor air quality (i.e., CO_2 levels, humidity, temperature, etc.)—should these statistics still be shown to the user? Furthermore, a building seldom has only a single occupant, and taking the possibly conflicting preferences of the users fully into account may in practice even be impossible. One potential research topic could be how to tackle the aforementioned issues, perhaps by utilizing A.I. or modern smart devices.

Acknowledgements

References

[1]   Dubar, I. G., Bogdan, R., & Popa, M. (2017) External rapid prototyping validation system for the automotive development cycle. Acta Polytechnica Hungarica, 14(6), 41-57, https://doi.org/10.12700/APH.14.6.2017.6.3

[2]   Official Statistics of Finland (OSF): Energy supply and consumption [e-publication] ISSN 1799-7976. 4th quarter 2018, Helsinki: Statistics Finland [referred: 24.6.2019] Access method: http://www.stat.fi/til/ehk/2018/04/ehk_2018_04_2019-03-28_tie_001_en.html

[3]   Official Statistics of Finland (OSF): Energy consumption in households [e-publication] ISSN 2323-329X. 2016. Helsinki: Statistics Finland [referred: 24.6.2019]   Access   method: http://www.stat.fi/til/asen/2016/ asen_2016_2017-11-17_tie_001_en.html

[4]   J. E. Petersen, V. Shunturov, K. Janda, G. Platt, and K. Weinberger, "Dormitory residents reduce electricity consumption when exposed to real-time visual feedback and incentives," International Journal of Sustainability in Higher Education, Vol. 8, No. 1, pp. 16-33, 2007

[5]   N. N. Kang, S. H. Cho, and J. T. Kim, "The energy-saving effects of apartment residents' awareness and behavior," Energy and Buildings, Vol. 46, pp. 112-122, 2012

[6]   G. M. Huebner, J. Cooper, and K. Jones, "Domestic energy consumption - What role do comfort, habit, and knowledge about the heating system play?" Energy and Buildings, Vol. 66, pp. 626-636, 2013

[7]   A. Faruqui, S. Sergici, and A. Sharif, "The impact of informational feedback on energy consumption - A survey of the experimental evidence," Energy, Vol. 35, No. 4, pp. 1598-1608, 2010

[8]   B. Kitchenham and S. Charters, "Guidelines for performing Systematic Literature Reviews in Software Engineering Version 2.3," EBSE Tech. Rep. EBSE-2007-01, 2007

[9]   L. Yang, H. Yan, and J. C. Lam, "Thermal comfort and building energy consumption implications – A review," Applied Energy, Vol. 115, pp. 164-173, 2014

[10]   D. Marković, D. Vujicic, Z. Jovanovic, U. Pesovic, S. Randik, and D. Jagodic, "Concept of IoT system for monitoring conditions of thermal comfort," in International Scientific Conference "UNITECH 2016," 2016, November

[11]   M. Taleghani, M. Tenpierik, S. Kurvers, and A. van den Dobbelsteen, "A review into thermal comfort in buildings," Renew. Sustain. Energy Rev., Vol. 26, pp. 201-215, Oct. 2013

[12]   F. Salamone, L. Belussi, C. Curro, L. Danza, M. Ghellere, G. Guazzi, B. Lenzi, V. Megale, and I. Meroni, "Integrated Method for Personal Thermal Comfort Assessment and Optimization through Users," Sensors, 2018

[13]   L. Ciabattoni, F. Ferracuti, G. Ippoliti, S. Longhi, and G. Turri, "IoT based indoor personal comfort levels monitoring," in 2016 IEEE International Conference on Consumer Electronics (ICCE), 2016, December 2015, pp. 125-126

[14]   A. Ghahramani, F. Jazizadeh, and B. Becerik-Gerber, "A knowledge based approach for selecting energy-aware and comfort-driven HVAC

temperature set points," Energy and Buildings, Vol. 85, pp. 536-548, Dec. 2014

[15]   Y. Obuchi, T. Yamasaki, K. Aizawa, S. Toriumi, and M. Hayashi, "Measurement and evaluation of comfort levels of apartments using IoT sensors," IEEE International Conference on Consumer Electronics (ICCE), 2018, pp. 1-6

[16]   T. Niemelä, R. Kosonen, and J. Jokisalo, "Cost-effectiveness of energy performance renovation measures in Finnish brick apartment buildings," Energy Build., Vol. 137, pp. 60-75, Feb. 2017

[17]   B. E. Medina and L. T. Manera, "Retrofit of air conditioning systems through a Wireless Sensor and Actuator Network: An IoT-based application for smart buildings," IEEE 14th International Conference on Networking, Sensing and Control (ICNSC), 2017, pp. 49-53

[18]   S. R. West, J. K. Ward, and J. Wall, "Trial results from a model predictive control and optimisation system for commercial building HVAC," Energy and Buildings, Vol. 72, pp. 271-279, Apr. 2014

[19]   S. Gupta, M. S. Reynolds, and S. N. Patel, "ElectriSense: single-point sensing using EMI for electrical event detection and classification in the home," in UbiComp, 2010

[20]   C. A. Björkskog, G. Jacucci, L. Gamberini, T. Nieminen, T. Mikkola, C. Torstensson, and M. Bertoncini, "EnergyLife: Pervasive Energy Awareness for Households," in Proceedings of the 12th ACM International Conference Adjunct Papers on Ubiquitous Computing - Adjunct, 2010, pp. 361-362

[21]   M. Weiss, A. Helfenstein, F. Mattern, and T. Staake, "Leveraging smart meter data to recognize home appliances," IEEE International Conference on Pervasive Computing and Communications, 2012, pp. 190-197

[22]   L. Özgür, V. K. Akram, M. Challenger, and O. Dağdeviren, "An IoT based smart thermostat," 5th International Conference on Electrical and Electronic Engineering (ICEEE), 2018, pp. 252-256

[23]   M. Weiss and D. Guinard, "Increasing Energy Awareness Through Web-enabled Power Outlets," in Proceedings of the 9th International Conference on Mobile and Ubiquitous Multimedia, 2010, p. 20:1-20:10

[24]   N. Morimoto, Y. Fujita, M. Yoshida, H. Yoshimizu, M. Takiyamada, T. Akehi, and M. Tanaka, "Smart Outlet Network for Energy-Aware Services Utilizing Various Sensor Information," 27th International Conference on Advanced Information Networking and Applications Workshops, 2013, pp. 1630-1635

[25]   B. L. Risteska Stojkoska and K. V. Trivodaliev, "A review of Internet of Things for smart home: Challenges and solutions," J. Clean. Prod., Vol. 140, pp. 1454-1464, Jan. 2017

[26] P. H. Shaikh, N. B. M. Nor, P. Nallagownden, I. Elamvazuthi, and T. Ibrahim, "A review on optimized control systems for building energy and comfort management of smart sustainable buildings," Renewable and Sustainable Energy Reviews, Vol. 34, pp. 409-429, Jun. 2014

[27] A. Kumar and G. P. Hancke, "An Energy-Efficient Smart Comfort Sensing System Based on the IEEE 1451 Standard for Green Buildings," IEEE Sensors Journal, Vol. 14, No. 12, pp. 4245-4252, Dec. 2014

[28] L. Pocero, D. Amaxilatis, G. Mylonas, and I. Chatzigiannakis, "Open source IoT meter devices for smart and energy-efficient school buildings," HardwareX, Vol. 1, pp. 54-67, Apr. 2017

[29] S. Godo, J. Haase, and H. Nishi, "Air conditioning control using selfpowered sensor considering comfort level and occupant location," in IECON 2015 - 41st Annual Conference of the IEEE Industrial Electronics Society, 2015, pp. 002497-002502

[30] M. Saari, P. Sillberg, P. Rantanen, J. Soini, and H. Fukai, "Data collector service - practical approach with embedded linux," 38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2015, pp. 1037-1041, IEEE

[31] M. Saari, A. M. Baharudin, P. Sillberg, P. Rantanen, and J. Soini, "Embedded Linux controlled sensor network," 39th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2016, pp. 1185-1189, IEEE

[32] M. Saari, A. M. bin Baharudin, P. Sillberg, S. Hyrynsalmi, and W. Yan, "LoRa — A survey of recent research trends," 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2018, May, pp. 0872-0877, IEEE

[33] A. M. bin Baharudin, M. Saari, P. Sillberg, P. Rantanen, J. Soini, and T. Kuroda, "Low-energy algorithm for self-controlled Wireless Sensor Nodes," International Conference on Wireless Networks and Mobile Communications (WINCOM), 2016, pp. 42-46

[34] A. M. bin Baharudin, M. Saari, P. Sillberg, P. Rantanen, J. Soini, J. Jaakkola, W. Yan, "Portable Fog Gateways for Resilient Sensors Data Aggregation in Internet-less Environment," Eng. J., Vol. 22, No. 3, pp. 221-232, Jun. 2018

[35] I. Ahmad, P. Rantanen, P. Sillberg, J. Laaksonen, S. Liu, "VisualLabel: An Integrated Multimedia Content Management and Access Framework," in Proceedings of the 27th International Conference on Information Modelling and Knowledge Bases, EJC 2017, 2017, pp. 332-353

[36] P. Sillberg, M. Saari, J. Grönman, P. Rantanen and M. Kuusisto, "Interpretation, Modeling and Visualization of Crowdsourced Road Condition Data", IS-TRIA2019, Submitted

[37] M. Saari, A. M. bin Baharudin, and S. Hyrynsalmi, "Survey of prototyping solutions utilizing Raspberry Pi," 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2017, pp. 991-994, IEEE

[38] P. Sillberg, J. Gronman, P. Rantanen, M. Saari, and M. Kuusisto, "Challenges in the Interpretation of Crowdsourced Road Condition Data," International Conference on Intelligent Systems (IS), 2018, pp. 215-221

[39] J. Grönman, P. Rantanen, M. Saari, P. Sillberg, and H. Jaakkola, "Lessons Learned from Developing Prototypes for Customer Complaint Validation," in CEUR Workshop Proceedings, 2018

[40] J. Grönman, P. Sillberg, P. Rantanen, and M. Saari, "People Counting in a Public Event—Use Case: Free-to-Ride Bus," 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2019, IEEE

Results on Tensor Product-based Model Transformation of Magnetic Levitation Systems

Elena-Lorena Hedrea, Radu-Emil Precup, Claudia-Adina Bojan-Dragos

Department of Automation and Applied Informatics, Politehnica University of Timisoara, Bd. V. Parvan 2, RO-300223 Timisoara, Romania
E-mail: elena.constantin@student.upt.ro, radu.precup@upt.ro, claudia.dragos@upt.ro

Abstract: In this paper the TP-based model transformation method is used in order to obtain a Tensor Product-based model of magnetic levitation systems which approximates the behavior of the plant, but exhibiting a numerical approximation error. In order to test the derived TP model, the behavior of the TP model is compared to the laboratory equipment behavior taking into consideration five testing scenarios. Experimental results show that approximation errors are generally low, but depend on model parameters.

Keywords: LTI systems; qLPV models; Tensor Product; magnetic levitation systems; transformation spaces

1 Introduction

The Tensor Product-based Model transformation (TPM) technique is a numerical, non-heuristic method that is capable of transforming a dynamic system model, given over a bounded domain, into parameter-varying weighted combination of parameter independent (constant) system models under the form of Linear Time-Invariant (LTI) systems. More precisely the TPM starts with Linear Parameter–Varying (LPV) dynamic models and derivates Linear Time-Invariant (LTI) systems as shown, for example, in the seminal papers and book (Baranyi, 2004) [1], (Petres et al., 2007) [2] and (Baranyi et al., 2013) [3].

TPM has the advantage of allowing linear matrix inequality (LMI) and parallel distributed compensation (PDC) frameworks to be applied immediately to the resulting affine models. This leads to tractable and improved control system performance.

The derivations of TP-based model transformation design approaches for different application plants such as models of diabetes mellitus and nonlinear insulin-

glucose dynamics, a nonlinear flexible joint robot system, a multi-tank system, etc., are given in the specialized literature in (Korondi, 2006) [4], (Galambos et al., 2015) [5] and (Hedrea et al., 2018) [6]. The combination with Proportional–Integral–Derivative controller tuning is treated in (Kuti and Galambos, 2018) [7]. The book (Baranyi, 2016) [8] and the papers (Szöllösi and Baranyi, 2016) [9] and (Szöllösi and Baranyi, 2016) [10] are important as they prove that the manipulation of the TPM is neccesary in control deisgn like PDC. The latest results where the number of variables or inputs may differ are presented in (Baranyi, 2018) [11], and also in (Baranyi, 2019) [12] if the TPmodel starts from $\mathbf{f}(\mathbf{x}, \mathbf{u}, \mathbf{p})$, where the matrix structure is unknown.

The Magnetic Levitation System (MLS) is a laboratory equipment (LabEq) used for experiments. It is an important benchmark to test linear and nonlinear modeling and control approaches applied to various areas including transportation systems. Some recent modeling solutions proposed for MLS include neural networks reported in (Rubio et al., 2017) [13], evolving fuzzy models reported in (Precup et al., 2017) [14] and Euler–Lagrange method reported in (Sun et al., 2017) [15]. The evolving fuzzy models prove to be popular recently and the results related to MLS can be considered as belonging to the hot fields of transportation systems and automotive technology as exemplified by Precup et al. (2017) in [16].

This paper is an extended version of the paper (Hedrea et al., 2019) [17], where the derivation of a TP–based model (TPmodel) using TPM was recently proposed. The TPmodel is then tested and its validation is improved using two testing scenarios. The topic at hand should be of interest to many engineers hoping to apply the TPmodel as a numerical modeling approach. The main contributions of this paper, which required restructuring in all sections including authors team, are pointed out as follows: the authors use the same main steps as the ones presented in [17] in order to obtain the TPmodel of the stabilized reduced order linearized model of a magnetic levitation system (referred to as stMaglev) and discussed in (Inteco, 2008) [18] and (Bojan-Dragos et al., 2018) [19]. However, the derived model is tested using four new testing scenarios. More precisely four control inputs (signals), namely a staircase control input, a sine control input, a chirp control input and a Pulse–Width modulation (PWM) control input, were applied to both stMaglev LabEq and TPmodel of stMaglev and their corresponding outputs were compared.

A part of the results given in both [17] and the current paper represent a sample of the continuation of the fruitful cooperation with the team of the Óbuda University (Budapest, Hungary). The excellent scientific contributions and management activity of Prof. Imre J. Rudas are kindly acknowledged. Some representative well-accepted joint papers in this regard are given in (Pozna et al., 2010 [20], (Haidegger et al., 2012) [21], (Precup et al., 2012) [22] and (Takács et al., 2015) [23].

The paper treats these topics: Section 2 gives the steps of TPM and the derivation of the TPmodel for stMaglev. Section 3 illustrates the four testing scenarios used for testing the derived TPmodel for stMaglev and Section 4 highlights the conclusions.

2 The TPmodel Derivation for stMaglev

2.1 TP-based Model Design Approach

When creating TPmodels for system representations, it is always useful for the reader to understand the base system equations of the physical system, preferably in a continuous–time state–space representation. That is the reason why such details are given in the section.

The TP-based model transformation is a numerical non-heuristic method which was first introduced by Baranyi (2004) in [1]. This method uses the high order singular value decomposition (HOSVD) technique in order to generate convex polytopic forms starting with the LPV models. In order to derivate a TPmodel uring the TPM technique the six steps with the diagram illustrated by Hedrea et al. (2019) in [17] and detailed in the following paragraphs are used.

In the first step the Transformation Space (TSp) is defined. Let $\mathbf{p} = [p_1 \ p_2 \dots p_n]^T \in \boldsymbol{\Omega}$ be a parameter vector and n the number of parameters. Therefore, $\boldsymbol{\Omega} = [a_1, b_1] \times [a_2, b_2] \times \dots \times [a_n, b_n] \subset \Re^n$ is the TSp with the bounds of the intervals $[a_i, b_i]$, $i = 1 \dots n$ chosen according to the plant specifications. A TSp $\boldsymbol{\Omega} = [a_1, b_1] \times [a_2, b_2]$ for two parameters is illustrated in (Hedrea et al., 2019) [17].

In the second step the Dicretization Grid (DG) is defined. Let $M_i, M_i \in \mathbf{N}, M_i \geq 2$ be the number of the discretization points from each interval $[a_i, b_i]$, $i = 1 \dots n$, including the ends of the intervals, which are computed using the technique described in (Baranyi et al., 2013) [3]. Therefore, the DG is given as:

$$\mathbf{M} = \{\mathbf{g}_{m_1, m_2, \dots m_n} \in \boldsymbol{\Omega}\}, m_i = 1 \dots M_i, i = 1 \dots n,$$
$$|\mathbf{M}| = M_1 \cdot M_2 \cdot \dots \cdot M_n, \tag{1}$$

where $\mathbf{g}_{m_1, m_2, \dots m_n} \in \boldsymbol{\Omega}$ is a discretization point. An example of DG with $|\mathbf{M}| = M_1 \cdot M_2 = 8 \cdot 6$, where $n = 2$, for $M_1 = 8$ and $M_2 = 6$ is given in (Hedrea et al., 2019) [17].

In the third step the discretized Tensor (dTens) is determined. Using the LPV model of the plant as shown in (Baranyi et al., 2013) [3] and (Hedrea et al., 2018) [6], the System matrix (Sm) can be defined as:

$$\mathbf{S}(\mathbf{p}) = \begin{bmatrix} \mathbf{A}(\mathbf{p}) & \mathbf{B}(\mathbf{p}) \\ \mathbf{C}(\mathbf{p}) & \mathbf{D}(\mathbf{p}) \end{bmatrix} \in \mathfrak{R}^{(l+q)\times(m+q)},$$ (2)

$$\mathbf{S}(\mathbf{p}) = [s_{ij}(\mathbf{p})]_{i=1\dots(l+q),\, j=1\dots(m+q)}.$$

Considering the parameter vector equal to the discretization point $\mathbf{p} = \mathbf{g}_{m_1,m_2,\dots m_n} = [g_{1,m_1} \quad g_{2,m_2} \quad \cdots \quad g_{n,m_n}]^T \in \mathbf{M}$ the Discretized System matrix (DSm) is given as:

$$\mathbf{S}^D_{m_1,m_2,\dots m_n} = \mathbf{S}(\mathbf{g}_{m_1,m_2,\dots m_n}) \in \mathfrak{R}^{(l+q)\times(m+q)}$$
$$\mathbf{S}^D_{m_1,m_2,\dots m_n} = [s_{ij}(\mathbf{g}_{m_1,m_2,\dots m_n})]_{i=1\dots(l+q),\, j=1\dots(m+q)}$$ (3)

and the dTens $\mathbf{S}^D$ is defined as:

$$\mathbf{S}^D = [\mathbf{S}^D_{m_1,m_2,\dots m_n}]_{m_1=1\dots M_1, m_2=1\dots M_2, \dots m_n=1\dots M_n} \in \mathfrak{R}^{M_1\times M_2\times\dots\times M_n\times(l+q)\times(m+q)}.$$ (4)

A particular example of a dTens computed for two parameters $p_1 \in [a_1, b_1]$ and $p_2 \in [a_2, b_2]$ with the TSp $\Omega = [a_1, b_1] \times [a_2, b_2]$ and the DG $|\mathbf{M}| = M_1 \cdot M_2 = 8 \cdot 6$ has the following expression:

$$\mathbf{S}^D = \begin{bmatrix} \mathbf{S}^D_{1,1} & \mathbf{S}^D_{1,2} & \dots & \mathbf{S}^D_{1,6} \\ \mathbf{S}^D_{2,1} & \mathbf{S}^D_{2,2} & \dots & \mathbf{S}^D_{2,6} \\ \dots & \dots & \dots & \dots \\ \mathbf{S}^D_{8,1} & \mathbf{S}^D_{8,2} & \dots & \mathbf{S}^D_{8,6} \end{bmatrix} \in \mathfrak{R}^{8\times6\times(l+q)\times(m+q)}.$$ (5)

In the fourth step the HOSVD is applied in order to obtain the singular values of the dTens $\mathbf{S}^D \in \mathbf{R}^{M_1\times M_2\times\dots\times M_n\times(l+q)\times(m+q)}$, which can be expressed as $\mathbf{S}^D = \mathbf{S} \overset{N}{\underset{n=1}{\otimes}} \mathbf{U}_n$ (Baranyi et al., 2013) [3] where $\mathbf{U}_n$, $\mathbf{S}$ and $\otimes$ are expressed in (Baranyi et al., 2013) [3] and (Hedrea et al., 2019) [17].

The n–mode matrix $\mathbf{S}^D_{(n)} \in \mathfrak{R}^{M_n\times(M_{n+1}M_{n+2}\dots(m+q)M_1M_2\dots(l+q))}$ can be given as $\mathbf{S}^D_{(n)} = [\mathbf{s}^D_r]$, where $\mathbf{s}^D_r \in \mathfrak{R}^{M_n}$ denote the column vectors of the M_n dimension of tensor $\mathbf{S}^D$ and $r = 1\dots R$, with $R = M_{n+1}M_{n+2}\dots(m+q)M_1M_2\dots(l+q)$.

In order to compute the HOSVD of the tensor $\mathbf{S}^D$ n singular value decompositions (SVD) made for all the n–mode matrices $\mathbf{S}^D_{(n)}$ are made using the theorem given in (Hedrea et al., 2019) [17] and (Lathauwer et al., 2000) [25], whose proof is given by Lathauwer et al. (2000) in [25].

Using this theorem given in (Lathauwer et al., 2000) [25], the SVD (with the three steps a), b) and c) detailed in [16]) of the n–mode matrix $\mathbf{S}_{(n)}^{D}$ can be given as $\mathbf{S}_{(n)}^{D} = \mathbf{U}_n \mathbf{\Sigma}_n \mathbf{V}_n^T$ (Hedrea et al., 2019) [17].

Finally the matrices $\mathbf{U}_n$ and $\mathbf{V}_n$ are computed following the steps taken from (Hedrea et al., 2019) [17].

In the fifth step the numerical values of the weighting functions are determined. The column vectors $\mathbf{u}_{n,I_n}$ in the matrix $\mathbf{U}_n$ are called weighting vectors and they contain the values of the w.f. $\mathbf{w}_n(\mathbf{p}_{m_1,m_2,..m_n})$ for $\mathbf{p}_{m_1,m_2,..m_n} = (g_{1,m_1},\cdots g_{n,m_n})$ (Baranyi et al., 2013) [3]:

$$\mathbf{w}_n(\mathbf{p}_{m_1,m_2,..m_n}) = \mathbf{u}_{n,I_n}. \tag{6}$$

In the final step the core tensor $\mathbf{S}_f$ is computed using the dTens $\mathbf{S}^D$ and the matrix $\mathbf{U}_N$ from the above steps (Baranyi et al., 2013) [3]:

$$\mathbf{S}_f = \mathbf{S}^D \overset{N}{\underset{n=1}{\otimes}} \mathbf{U}_N^T \tag{7}$$

The core tensor $\mathbf{S}_f$ is defined as $\mathbf{S}_f = \sum_{m_1=1}^{M_1}\sum_{m_2=1}^{M_2}\cdots\sum_{m_n=1}^{M_n}\prod_{n=1}^{N}\mathbf{w}_n(\mathbf{p}_{m_1,m_2,..m_n})\mathbf{S}_{m_1,m_2,...m_n}^{LTI}$ with the equivalent notation $\mathbf{S}(\mathbf{p}(t)) = \mathbf{S}_f \otimes \mathbf{w}_n(\mathbf{p}_{m_1,m_2,..m_n})$ presented in (Baranyi, 2004) [1].

2.2   Derivation of TPmodel for stMaglev

The modelled plant considered in this paper is a laboratory system based on the magnetic levitation principle, which includes a metallic frame with one upper electromagnet, Electromagnet1, and one lower electromagnet, Electromagnet2, between which a ferromagnetic sphere levitates as shown in Figure 1. The position of the ferromagnetic sphere is measured using position sensors. In order to ensure the communication between the hardware and the software components one computer interface is used.

The base system equations for MLS are (Inteco, 2008) [18]:

$$\dot{x}_1(t) = v(t),$$

$$\dot{v}(t) = -\frac{i_{EM1}^2(t) \cdot F_{emP1} \cdot \exp[-x_1(t)/F_{emP2}]}{m \cdot F_{emP2}} + g$$

$$+ \frac{i_{EM2}^2(t) \cdot F_{emP1} \cdot \exp[-(x_d - x_1(t))]/F_{emP2}}{m \cdot F_{emP2}}, \qquad (8)$$

$$i_{EM1}(t) = \frac{k_i \cdot u_{EM1}(t) + c_i - i_{EM1}(t)}{(f_{iP1}/f_{iP2}) \cdot \exp[-x_1(t)/f_{iP2}]},$$

$$i_{EM2}(t) = \frac{k_i \cdot u_{EM2}(t) + c_i - i_{EM2}(t)}{(f_{iP1}/f_{iP2}) \cdot \exp[-(x_d - x_1(t))/f_{iP2}]},$$

$$y(t) = k_m \cdot x_1(t),$$

where the ferromagnetic sphere position (m) is $x_1 \in [0, 0.0016]$, the speed of the ferromagnetic sphere (m/s) is $v \in \Re$, the current of Electromagnet1 (A) is $i_{EM1} \in [0.03884, 2.38]$, the current of Electromagnet2 (A) is $i_{EM2} \in [0.03884, 2.38]$, the control signals applied to Electromagnet1 and Electromagnet2, respectively (V) are $u_{EM1} \in [0.005, 1]$ and $u_{EM2} \in [0.005, 1]$ and the measured output of the process (m) is denoted by y. The process parameters are: m=0.0571 [kg] – the mass of ferromagnetic sphere, F_{emP1}=1.7521·10^{-2} [H], F_{emP2}=5.8231·10^{-3} [m], k_i=0.0243 [A], c_i=2.5165 [A], f_{iP1}=1.4142·10^{-4} [ms], f_{iP2}=4.5626·10^{-3} [m] (Bojan-Dragos et al., 2018) [19].
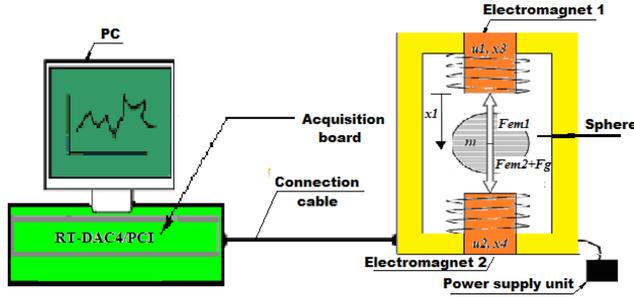
Figure 1
Experimental setup for MLS

In order to determine the qLPV model of the process, which is later used in the derivation of the TPmodel a stabilizing control solution was designed (Bojan-Dragos et al., 2018) [19] resulting the stabilized linearized model for MLS (stMaglev):

Therefore, the qLPV model representation of stMaglev is expressed as

$$\dot{\mathbf{x}} = \mathbf{A}_x(\mathbf{p})\mathbf{x} + \mathbf{b}_{1x}(\mathbf{p})u_{1x},$$
$$y = \mathbf{c}^T(\mathbf{p})\mathbf{x}, \tag{9}$$
$$\mathbf{x} = [x_1 \quad v \quad i_{EM1}]^T, \ \mathbf{p} = [x_1 \quad i_{EM1}]^T,$$

where the matrices $\mathbf{A}_x(\mathbf{p})$, $\mathbf{b}_{1x}(\mathbf{p})$ and $\mathbf{c}^T(\mathbf{p})$ are (Bojan-Dragos et al., 2018) [19]

$$\mathbf{A}_x(\mathbf{p}) = \begin{bmatrix} 0 & 1 & 0 \\ a_{21}(\mathbf{p}) & 0 & a_{23}(\mathbf{p}) \\ a_{31}(\mathbf{p}) & a_{32}(\mathbf{p}) & a_{33}(\mathbf{p}) \end{bmatrix}, \ \mathbf{b}_{1x}(\mathbf{p}) = \begin{bmatrix} 0 \\ 0 \\ b_{31}(\mathbf{p}) \end{bmatrix}, \mathbf{c}^T(\mathbf{p}) = [1 \ 0 \ 0], \tag{10}$$

$$\mathbf{A}_x(\mathbf{p}) \in \Re^{3x3}, \ \mathbf{b}_{1x}(\mathbf{p}) \in \Re^{3x1}, \ \mathbf{c}^T(\mathbf{p}) \in \Re^{1x3}, \ u_{1x} \in \Re,$$

with the elements:

$$a_{21}(\mathbf{p}) = \frac{\mathbf{p}(2)^2}{m}\frac{F_{emP1}}{F_{emP2}^2}e^{-\frac{\mathbf{p}(1)}{F_{emP2}}}, \ a_{23}(\mathbf{p}) = -\frac{2\mathbf{p}(2)}{m}\frac{F_{emP1}}{F_{emP2}}e^{-\frac{\mathbf{p}(1)}{F_{emP2}}},$$

$$a_{31}(\mathbf{p}) = -(k_iu_{1x} + c_i - \mathbf{p}(2))\frac{\mathbf{p}(1)}{f_{iP1}} \cdot e^{\frac{\mathbf{p}(1)}{f_{iP2}}} + 66.33 \cdot k_i \cdot \frac{f_{iP2}}{f_{iP1}} \cdot e^{\frac{\mathbf{p}(1)}{f_{iP2}}}, \tag{11}$$

$$a_{32}(\mathbf{p}) = 1.62 \cdot k_i \cdot \frac{f_{iP2}}{f_{iP1}} \cdot e^{\frac{\mathbf{p}(1)}{f_{iP2}}}, a_{33}(\mathbf{p}) = -\frac{f_{iP2}}{f_{iP1}} \cdot e^{\frac{\mathbf{p}(1)}{f_{iP2}}} - 0.15 \cdot k_i \cdot \frac{f_{iP2}}{f_{iP1}} \cdot e^{\frac{\mathbf{p}(1)}{f_{iP2}}},$$

$$b_{31}(\mathbf{p}) = k_i \cdot \frac{f_{iP2}}{f_{iP1}} \cdot e^{\frac{\mathbf{p}(1)}{f_{iP2}}}.$$

where $\mathbf{p}$ is vector of the parameters which contains the state variable $\mathbf{p}(1)$ – the position of the ferromagnetic sphere and the state variable $\mathbf{p}(2)$ – the top electromagnet current, v is the speed of the ferromagnetic sphere, u_{1x} is the plant input, y is the measured output of the process.

Introducing in (9) the Sm $\mathbf{S}(\mathbf{p}) = [\mathbf{A}_x(\mathbf{p}) \quad \mathbf{b}_{1x}(\mathbf{p})] \in \Re^{3\times4}$, the model is transformed in the qLPV state–space form

$$\dot{\mathbf{x}} = \mathbf{S}(\mathbf{p})[\mathbf{x}^T \quad u_{1x}]^T, \tag{12}$$
$$y = \mathbf{c}^T(\mathbf{p})\mathbf{x}.$$

with the following LTI models (Hedrea et al., 2017) [26]:

$$\dot{\mathbf{x}} = \mathbf{S}(\mathbf{p}) \overset{N}{\underset{n=1}{\otimes}} \mathbf{w}_n(\mathbf{p}_n)[\mathbf{x}^T \quad u_{1x}]^T$$
$$= \sum_{m_1=1}^{M_1}\sum_{m_2=1}^{M_2} w_{1,m_1}(p_1)w_{2,m_2}(p_2)\mathbf{S}_{m_1,m_2}[\mathbf{x}^T \quad u_{1x}]^T, \tag{13}$$
$$y = \mathbf{c}^T(\mathbf{p})\mathbf{x},$$

The LTI Sms contain the matrices $\mathbf{A}_{x_{m_1,m_2}}$ and $\mathbf{b}_{1x_{m_1,m_2}}$ from the state–space model

$$\dot{\mathbf{x}} = \sum_{m_1=1}^{3}\sum_{m_2=1}^{3} w_{1,m_1}(p_1)w_{2,m_2}(p_2)(\mathbf{A}_{x_{m_1,m_2}}\mathbf{x} + \mathbf{b}_{1x_{m_1,m_2}}u_{1x}), \tag{14}$$

$$y = \mathbf{c}^T(\mathbf{p})\mathbf{x},$$

3   Experimental Results

Using the TP Tool, with its operation mode described by Nagy et al. (2007) in [27], the matrices $\mathbf{S}_{m1,m2}$ obtained for stMaglev are given in (14) and the w.f.s are presented in Figure 2 for the two parameters, namely the sphere position and the top electromagnetic current.
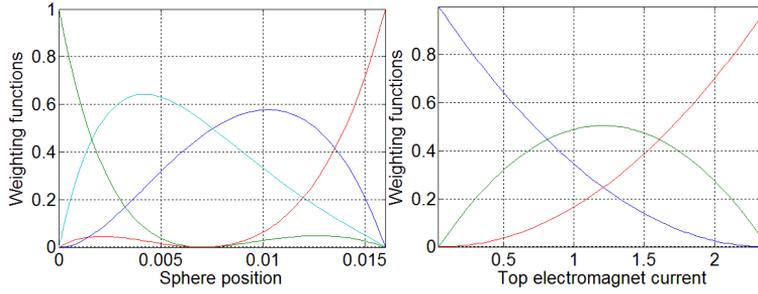
Figure 2

W.f.s obtained by TPM of stMaglev (sphere position and top electromagnetic current)

$$S_{1,1} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 709.3 & -0.1 & -60.7 & 0 \\ -3413.2 & -248.6 & 84 & -153.5 \end{bmatrix}, \quad S_{2,1} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 157 & -0.1 & -8 & 0 \\ 60074 & 4385 & -1482 & 2707 \end{bmatrix},$$

$$S_{3,1} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1411.7 & -0.1 & -127.5 & 0 \\ 1795.1 & 131.5 & -44.4 & 81.2 \end{bmatrix}, \quad S_{4,1} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -183 & -0.1 & 24 & 0 \\ 36927 & 2699 & -912 & 1666 \end{bmatrix},$$

$$S_{1,2} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 76.5 & -0.1 & -2 & 0 \\ -3282.4 & -248.6 & 84 & -153.5 \end{bmatrix}, \quad S_{2,2} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 72 & -0.1 & -8 & 0 \\ 58602 & 4385 & -1482 & 2707 \end{bmatrix},$$

$$S_{3,2} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 83.8 & -0.1 & -4.1 & 0 \\ 1795.1 & 131.5 & -44.4 & 81.2 \end{bmatrix}, \quad S_{4,2} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -68 & -0.1 & 1 & 0 \\ 36305 & 2699 & -912 & 1666 \end{bmatrix}, \quad (15)$$

$$S_{1,3} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 24496 & -0.1 & -120 & 0 \\ -3544 & -249 & 84 & -153.5 \end{bmatrix}, \quad S_{2,3} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 3356 & -0.1 & -16 & 0 \\ 61546 & 4385 & -1482 & 2707 \end{bmatrix},$$

$$S_{3,3} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 51329 & -0.1 & -251 & 0 \\ 1795 & 132 & -44 & 81.2 \end{bmatrix}, \quad S_{4,3} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -9650 & -0.1 & 48 & 0 \\ 37549 & 2699 & -912 & 1666 \end{bmatrix}.$$

In order to test the derived TPmodel, five testing scenarios are presented in this extented paper and are detailed in the followings. The same testing signal was applied both to the stMaglev LabEq and to the TPmodel derived for stMaglev on the time frame of 20 s and their corresponding outputs, y_{MLS_j} and y_{TP_j}, $j \in \{PRBS,$ STAIRS, SINE, CHIRP, PWM\}, were compared (Figure 3). The initial state vector matching the experiments was $\mathbf{x}_0 = [0.0083 \quad 0 \quad 0]^T$.
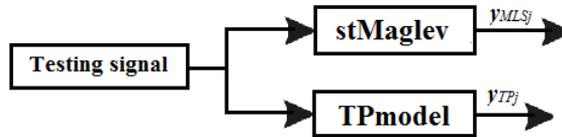
Figure 3
Testing block diagram for stMaglev LabEq and TPmodel

The first testing scenario is the same as the first one used by Hedrea et al. (2019) in [17] and consists in applying a Pseudo Random Binary Signal (PRBS) with a 0.008 m amplitude as control input with the corresponding plot of the sphere position versus time illustrated in Figure 4.
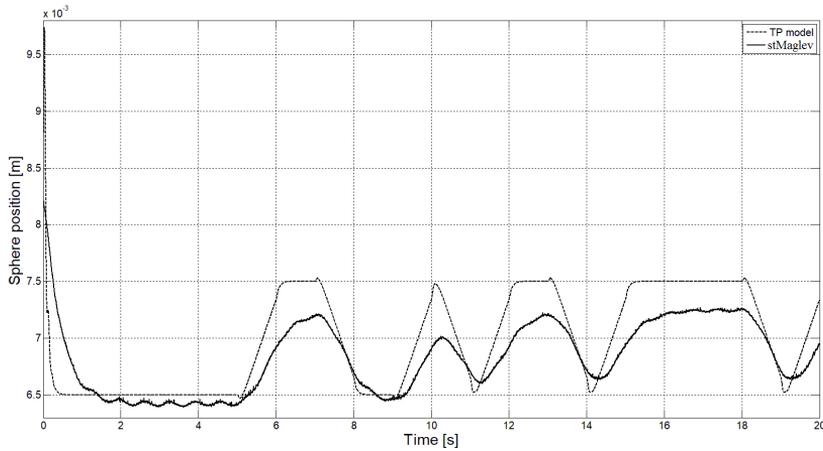
Figure 4

The time response of TPmodel and stMaglev with PRBS control input

The next four testing scenarios consist in applying four new control inputs (signals), namely a staircase control input, a sine control input, a chirp control input and a PWM control input, to both stMaglev LabEq and TPmodel of stMaglev.

In the first new testing scenario, the plot of the sphere position versus time obtained after applying a staircase control input with a $R1=0.006$ m, $R2=0.008$ m and $R3=0.007$ m amplitude as control input is illustrated in Figure 5.

In the second new testing scenario, the plot of the sphere position versus time obtained after applying a sine control input with a 0.001 m amplitude as control input is illustrated in Figure 6.
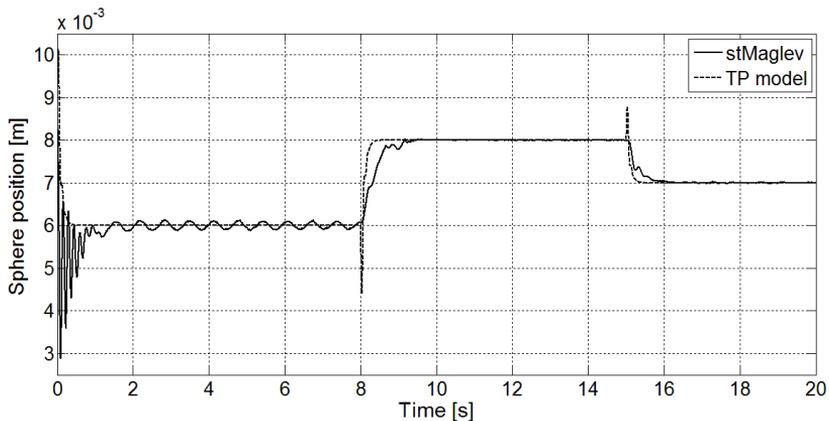
Figure 5

The time response of TPmodel and stMaglev with PRBS control input
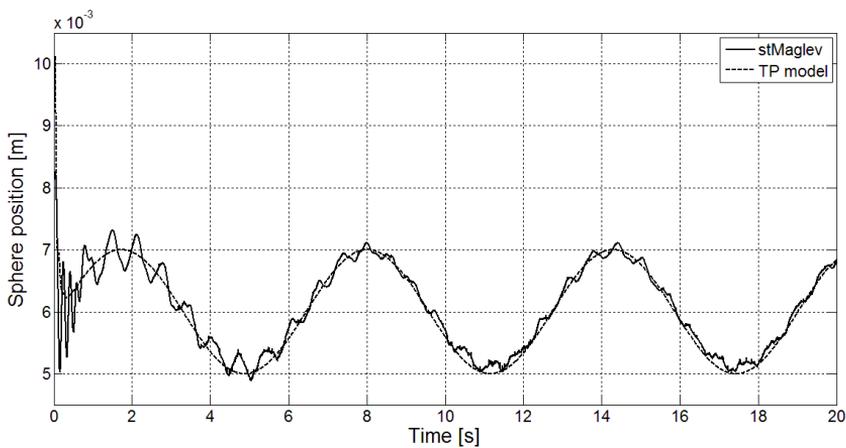
Figure 6

The time response of TPmodeland stMaglev with sine control input

In the third testing scenario the plot of the sphere position versus time obtained after applying a Chirp control input with a 0.1 initial frequency as control input is illustrated in Figure 7.

In the fourth testing scenario the plot of the sphere position versus time obtained after applying a Pulse–width modulation (PWM) control signal with a 0.0012 m amplitude, a 50% pulse width as control input is illustrated in Figure 8.
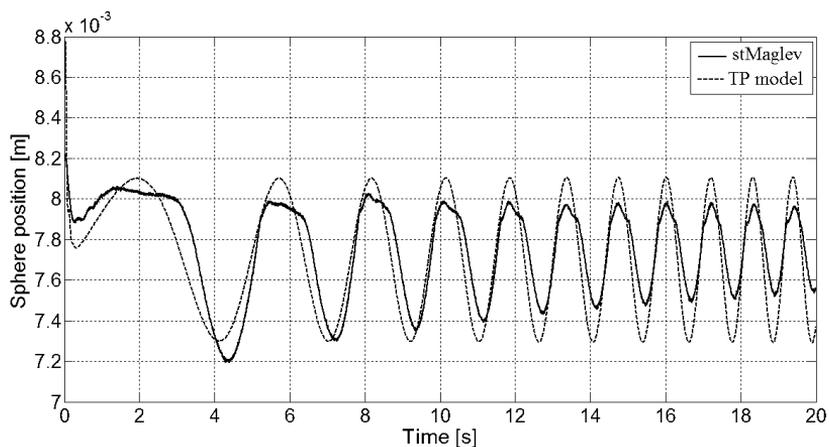
Figure 7

The time response of TPmodel and stMaglev with chirp control input
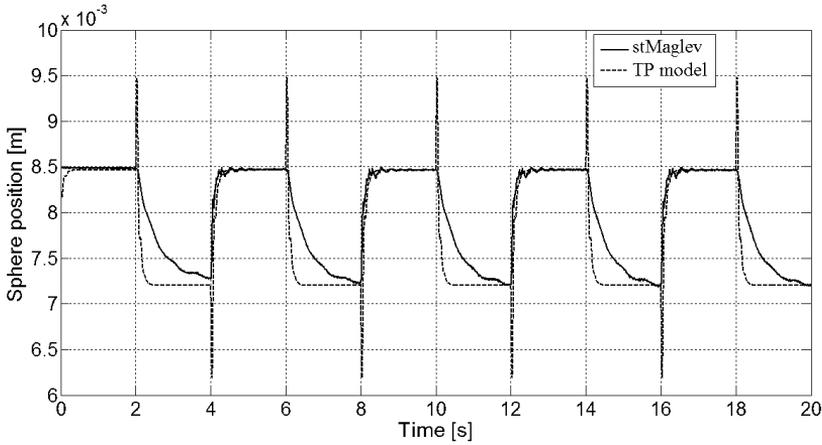
Figure 8
The time response of TPmodel and stMaglev with PWM control input

In order to better highlight the performances of the TPmodel derived for stMaglev in all testing scenarios the following performance indices were computed: the modeling errors, the mean square error and the percent relative modeling error.

The modeling errors were computed as the difference between the output responses of the real–world stMaglev (experimenting on the LabEq) and the TPmodel of stMaglev:

$$e_j = y_{MLS_j} - y_{TP_j},$$ (16)

The mean square error (MSE) was also calculated as:

$$\text{MSE}_j = \frac{1}{N} \sum_{t_d=1}^{N} (e_j(t_d))^2,$$ (17)

where e_j results from (16), N=80000 is the number of records. The following numerical values of MSE were obtained: $\text{MSE}_{PRBS} = 7.2096 \cdot 10^{-8}$, $\text{MSE}_{STAIRS} = 3.5279 \cdot 10^{-8}$ in case of PRBS control input, $\text{MSE}_{SINE} = 1.1904 \cdot 10^{-7}$ in case of sine control input, $\text{MSE}_{CHIRP} = 1.8753 \cdot 10^{-8}$ in case of chirp control input and $\text{MSE}_{PWM} = 8.9096 \cdot 10^{-8}$ in case of Pulse–width modulation (PWM) control input. The MSE numerical values are small because the ranges of stMaglev and TPmodel outputs are less than 10 mm.

The percent relative modeling errors have the following expressions:

$$e_{rj}[\%] = |e_j| / |y_{MLS_j}| \cdot 100.$$ (18)

The plot of the percent relative modeling errors in case of the PRBS control signal is illustrated in Figure 9. The plots of the percent relative modeling errors in all new testing scenarios are illustrated in Figures 10-13.

The large values of the percent relative modeling errors in the initial phase of system responses are caused by neglecting the fourth state variable of stMaglev, the bottom electromagnet current, in the design of stMaglev and next the derivation of TP-m.
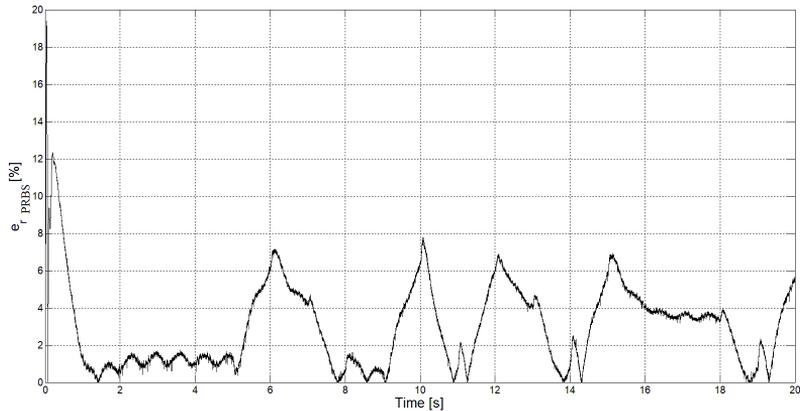
Figure 9

Percent relative modeling error with PRBS control input
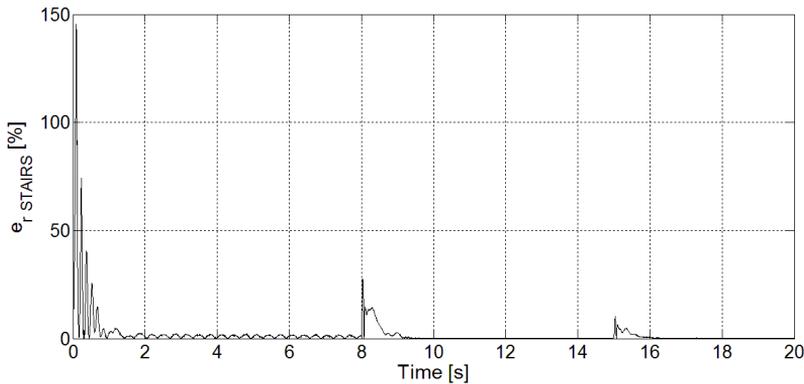
Figure 10

Percent relative modeling error with staircase control input
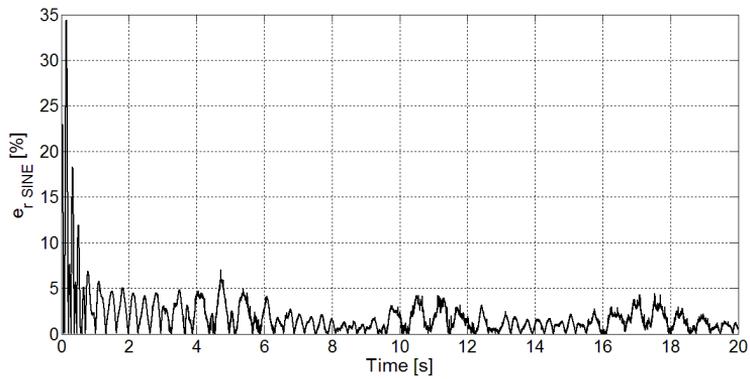
Figure 11

Percent relative modeling error with sine control input
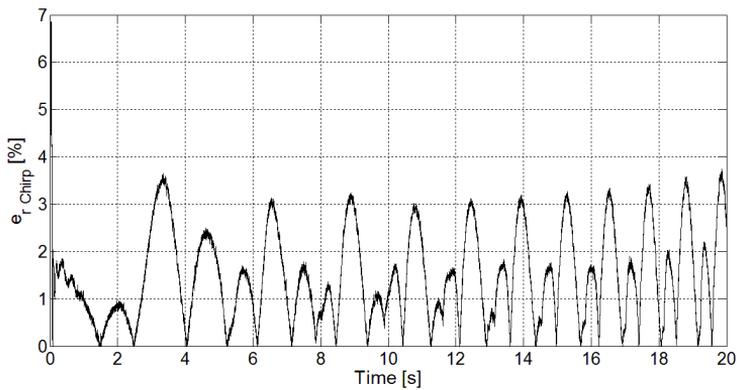
Figure 12

Percent relative modeling error with chirp control input
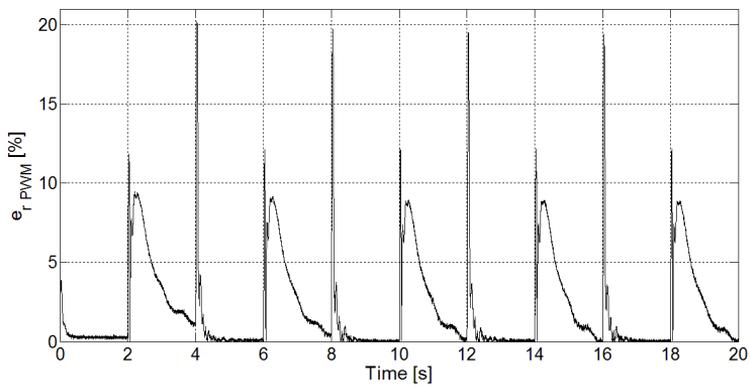
Figure 13

Percent relative modeling error with PWM control input

Conclusions

This paper proposed an extension of the ideas suggested in Hedrea et al. (2019) [17] by means of four new testing scenarios and adding useful information on the nonlinear plant that is subjected to the attractive nonlinear modeling and control technique built around TP. The new testing scenarios are important as a nonlinear plant is controlled and the TPmodel proposed by Hedrea et al. (2019) in [17] and this paper and will next be used in model-based control requires adequate validation. Various operating regimes were considered in this respect, and three performance indices, namely the modeling error, the mean square error and the percent relative modeling error, were also computed.

The experimental results show that the derived TPmodel approximates the behavior of the plant, but exhibiting a numerical approximation error which depends on the model parameters. The numerical values of the performance indices show that the TPmodel ensures good performance in terms of mean square error and percent relative modeling error in all testing scenarios, which are relevant to the real process operation. Experimental results also show that approximation errors are generally low, but depend on the control input.

Future research will be focused on finding what options are available to further reduce the approximation errors of the derived TPmodels or, in other words, to analyze what parameters do the approximation errors depend on. Future research will also include a part of the next directions already identified and proposed in (Hedrea et al., 2019) [17]: the derivation of other TPmodels for different plants, and the adaptation of results from other models and application areas. Such promising and also challenging plants and applications include robotics [28-31], fuzzy models and control [32-38], neural networks [39], medicine [40-42], servo systems and engines [43, 44], supervisory control [45], and various modern optimization algorithms [46-51] applied to controller tuning and system model identification as well.

Acknowledgement

References

[1]   P. Baranyi: TP Model Transformation as a Way to LMI-based Controller Design, IEEE Transactions on Industrial Electronics, Vol. 51, No. 2, 2004, pp. 387-400

[2]   Z. Petres, P. Baranyi, P. Korondi, H. Hashimoto: Trajectory Tracking by TP Model Transformation: Case Study of a benchMark Problem, IEEE Transactions on Industrial Electronics, Vol. 54, No. 3, 2007, pp. 1654-1663

[3]   P. Baranyi, Y. Yam, P. Varlaki: TP Model Transformation in Polytopic Model-based Control. Boca Raton, FL: Taylor & Francis, 2013

[4]   P. Korondi: Tensor Product Model Transformation-based Sliding Surface Design, Acta Polytechnica Hungarica, Vol. 3, No. 4, 2006, pp. 23-36

[5]   P. Galambos, J. Kuti, P. Baranyi, G. Szögi, I. J. Rudas: Tensor Product-based Convex Polytopic Modeling of Nonlinear Insulin-Glucose Dynamics, in Proceedings of 2015 IEEE International Conference on Systems, Man and Cybernetics, Hong Kong, 2015, pp. 2597-2602

[6]   E.-L. Hedrea, R.-E. Precup, C.-A. Bojan-Dragos, C. Hedrea: Tensor Product-based Model Transformation Technique Applied to Modeling Vertical Three Tank Systems, in Proceedings of the 12th IEEE International Symposium on Applied Computational Intelligence, Timisoara, Romania, 2018, pp. 63-68

[7]   J. Kuti, P. Galambos: Tensor Product Model-based PID Controller Optimisation for Propofol Administration, IFAC-PapersOnLine, Vol. 51, No. 4, 2018, pp. 645-650

[8]   P. Baranyi: TP-Model Transformation-Based-Control Design Frameworks. Cham: Springer International Publishing Switzerland, 2016

[9]   A. Szöllösi, P. Baranyi: Influence of the tensor product model representation of qLPV models on the feasibility of linear matrix inequality, Asian Journal of Control, Vol. 18, No. 4, 2016, pp. 1328-1342

[10]   A. Szöllösi, P. Baranyi: Influence of the tensor product model representation of qLPV models on the feasibility of linear matrix inequality based stability analysis, Asian Journal of Control, Vol. 20, No. 1, 2018, pp. 531-547

[11]   P. Baranyi: Extension of the multi-TP model transformation to functions with different numbers of variables, Complexity, Vol. 2018, 2018, pp. 1-9

[12]   P. Baranyi: Extracting LPV and qLPV structures from state-space functions: a TP model transformation based framework, IEEE Transactions on Fuzzy Systems, 2019, DOI: 10.1109/TFUZZ.2019.2908770

[13]   J. J. Rubio, L. Zhang, E. Lughofer, P. Cruz, A. Alsaedi, T. Hayat: Modeling and control with neural networks for a magnetic levitation system, Neurocomputing, Vol. 227, 2017, pp. 113-121

[14]   R.-E. Precup, C.-A. Bojan-Dragos, E.-L. Hedrea, M.-D. Rarinca, E. M. Petriu: Evolving fuzzy models for the position control of magnetic levitation systems, in Proceedings of 2017 IEEE Conference on Evolving and Adaptive Intelligent Systems, Ljubljana, Slovenia, 2017, pp. 1-6

[15]   Y. Sun, W. Li, J. Xu, H. Quiang, C. Chen: Nonlinear dynamic modeling and fuzzy sliding-mode controlling of electromagnetic levitation system of

low-speed maglev train, Journal of Vibroengineering, Vol. 19, No. 1, 2017, pp. 329-342

[16]   R.-E. Precup, S. Preitl, C.-A. Bojan-Dragos, M.-B. Radac, A.-I. Szedlak-Stinean, E.-L. Hedrea, R.-C. Roman: Automotive applications of evolving Takagi-Sugeno-Kang fuzzy models, Facta Universitatis, Series: Mechanical Engineering, Vol. 15, No. 2, 2017, pp. 231-244

[17]   E.-L. Hedrea, R.-E. Precup, C.-A. Bojan-Dragos, O. Tanasoiu: Tensor product-based model transformation technique applied to modeling magnetic levitation systems, in Proceedings of the 23rd IEEE International Conference on Intelligent Engineering Systems, Gödöllő, Hungary, 2019, pp. 179-184

[18]   Inteco Ltd., Magnetic Levitation System 2EM (MLS2EM), User's Manual (Laboratory Set) Krakow, Poland: Inteco Ltd., 2008

[19]   C.-A. Bojan-Dragos, M.-B. Radac, R.-E. Precup, E.-L. Hedrea, O.-M. Tanasoiu: Gain-Scheduling Control Solutions for Magnetic Levitation Systems, Acta Polytechnica Hungarica, Vol. 15, No. 5, 2018, pp. 89-108

[20]   C. Pozna, R.-E. Precup, J. K. Tar, I. Škrjanc, S. Preitl: New results in modelling derived from Bayesian filtering, Knowledge-based Systems, Vol. 23, No. 2, 2010, pp. 182-194

[21]   T. Haidegger, L. Kovács, S. Preitl, R.-E. Precup, B. Benyó, Z. Benyó: Controller design solutions for long distance telesurgical applications, International Journal of Artificial Intelligence, Vol. 6, No. S11, 2011, pp. 48-71

[22]   R.-E. Precup, M. L. Tomescu, S. Preitl, E. M. Petriu, J. Fodor, C. Pozna: Stability analysis and design of a class of MIMO fuzzy control systems, Journal of Intelligent and Fuzzy Systems, Vol. 25, No. 1, 2013, pp. 145-155

[23]   Á. Takács, L. Kovács, I. Rudas, R.-E. Precup, T. Haidegger: Models for Force Control in Telesurgical Robot Systems, Acta Polytechnica Hungarica, Vol. 12, No. 8, 2015, pp. 95-114

[24]   Z. Petres: Polytopic Decomposition of Linear Parameter-Varying Models by Tensor-Product Model Transformation, PhD Thesis, Budapest University of Technology and Economics, Budapest, Hungary, 2006

[25]   L. Lathauwer, B. De Moor, J. Vandewalle: A multilinear singular value decomposition, SIAM Journal of Matrix Analysis and Applications, Vol. 21, No. 4, 2000, pp. 1253-1278

[26]   E.-L. Hedrea, C.-A. Bojan-Dragos, R.-E. Precup, R.-C. Roman, E. M. Petriu, C. Hedrea: Tensor product-based model transformation for position control of magnetic levitation systems, in Proceedings of 26th International Symposium on Industrial Electronics, Edinburgh, UK, 2017, pp. 1-6

[27] S. Nagy, Z. Petres, P. Baranyi: TP Tool – A Matlab Toolbox for TP Model Transformation, in Proceedings of 8th International Symposium on Computational Intelligence and Informatics, Budapest, Hungary, 2007, pp. 483-495

[28] J. Vaščák, M. Rutrich: Path planning in dynamic environment using fuzzy cognitive maps, in Proceedings of 6th International Symposium on Applied Machine Intelligence and Informatics, Herľany, Slovakia, 2008, pp. 5-9

[29] S. Blažič: On periodic control laws for mobile robots, IEEE Transactions on Industrial Electronics, Vol. 61, No. 7, 2014, pp. 3660-3670

[30] B. Kovács, G. Szayer, F. Tajti, M. Burdelis, P. Korondi: A novel potential field method for path planning of mobile robots by adapting animal motion attributes, Robotics and Autonomous Systems, Vol. 82, 2016, pp. 24-34

[31] C. Pozna, R.-E. Precup: An approach to the design of nonlinear state-space control systems, Studies in Informatics and Control, Vol. 27, No. 1, 2018, pp. 5-14

[32] R.-E. Precup, S. Preitl: Popov-type stability analysis method for fuzzy control systems, in Proceedings of Fifth European Congress on Intelligent Technologies and Soft Computing (EUFIT'97), Aachen, Germany, 1997, Vol. 2, pp. 1306-1310

[33] R.-E. Precup, S. Preitl: Fuzzy Controllers. Timisoara: Editura Orizonturi Universitare, 1999

[34] P. Angelov, N. Kasabov: Evolving computational intelligence systems, in Proceedings of 1st International Workshop on Genetic Fuzzy Systems, Granada, Spain, 2005, pp. 76-82

[35] Zs. Cs. Johanyák: A simple fuzzy logic based power control for a series hybrid electric vehicle, in Proceedings of 9th IEEE European Modelling Symposium, Madrid, Spain, 2015, pp. 207-212

[36] I. Dzitac, F. G. Filip, M. J. Manolescu: Fuzzy logic is not fuzzy: World-renowned computer scientist Lotfi A. Zadeh, International Journal on Computers Communications and Control, Vol. 12, No. 6, 2017, pp. 748-789

[37] F. Olivas, F. Valdez, O. Castillo, C. I. González, G. E. Martinez, P. Melin: Ant colony optimization with dynamic parameter adaptation based on interval type-2 fuzzy logic systems, Applied Soft Computing, Vol. 53, 2017, pp. 74-87

[38] V. Nikolić, M. Milovančević, D. Petković, D. Jocić, M. Savić: Parameters forecasting of laser welding by the artificial intelligence techniques, Facta Universitatis, Series: Mechanical Engineering, Vol. 18, No. 2, 2018, pp. 193-201

[39] I. Dumitrache, N. Constantin, M. Drăgoicea: Retele neurale: identificarea si conducerea proceselor. Bucharest: Matrix Rom, 1999

[40]   H. Costin, C. Rotariu: Medical image analysis and representation using a fuzzy and rule-based hybrid approach, International Journal on Computers Communications and Control, Vol. 1, No. S, 2006, pp. 156-162

[41]   L. Kovács: Linear Parameter Varying (LPV) based robust control of type-I diabetes driven for real patient data, Knowledge-Based Systems, Vol. 122, 2017, pp. 199-213

[42]   R.-E. Precup, T.-A. Teban, A. Albu, A.-I. Szedlak-Stinean, C.-A. Bojan-Dragos: Experiments in incremental online identification of fuzzy models of finger dynamics, Romanian Journal on Information Science and Technology, Vol. 21, No. 4, 2018, pp. 358-376

[43]   R.-E. Precup, S. Preitl: Development of fuzzy controllers with non-homogeneous dynamics for integral-type plants, Electrical Engineering, Vol. 85, No. 3, 2003, pp. 155-168

[44]   R. Andoga, L. Főző, J. Judičák, R. Bréda, S. Szabo, R. Rozenberg, M. Džunda: Intelligent situational control of small turbojet engines, International Journal of Aerospace Engineering, Vol. 2018, article ID 8328792, 2018, pp. 1-16

[45]   R. E. Haber, J. R. Alique, S. Ros Torrecillas, C. R. Peres Ferreira: Fuzzy supervisory control of end milling process, Information Sciences, Vol. 89, No. 1, 1996, pp. 95-106

[46]   B. Niu, Y., Fan, H. Wang, L. Li, X. Wang: Novel bacterial foraging optimization with time-varying chemotaxis step, International Journal of Artiffical Intelligence, Vol. 7, No. A11, 2011, pp. 257-273

[47]   R.-E. Precup, M.-L. Tomescu, C.-A. Dragos: Stabilization of Rössler chaotic dynamical system using fuzzy logic control algorithm, International Journal of General Systems, Vol. 43, No. 5, 2014, pp. 413-433

[48]   T. V. Tran, Y. N. Wang: Artificial chemical reaction optimization algorithm and neural network based adaptive control for robot manipulator, Control Engineering and Applied Informatics, Vol. 19, No. 2, 2017, pp. 61-70

[49]   M. Shams, E. Rashedi, S. M. Dashti, A. Hakimi: Ideal gas optimization algorithm, International Journal of Artiffical Intelligence, Vol. 15, No. 2, 2017, pp. 116-130

[50]   S. Vrkalovic, E.-C. Lunca, I.-D. Borlea: Model-free sliding mode and fuzzy controllers for reverse osmosis desalination plants, International Journal of Artiffical Intelligence, Vol. 16, No. 2, 2018, pp. 208-222

[51]   E. Osaba, X. S. Yang, I. Fister Jr, J. Del Ser, P. Lopez-Garcia, A. J. Vazquez-Pardavila: A discrete and improved bat algorithm for solving a medical goods distribution problem with pharmacological waste collection, Swarm and Evolutionary Computation, Vol. 44, 2019, pp. 273-286

Visceral versus Verbal:
Can We See Depression?

Xuanying Zhu, Tom Gedeon, Sabrina Caldwell, Richard Jones

Research School of Computer Science, The Australian National University, Canberra, Australia 2601

E-mail: xuanying.zhu@anu.edu.au, tom@cs.anu.edu.au, sabrina.caldwell@anu.edu.au, richard.jones@anu.edu.au

Abstract: Depression widely affects global populations and is one of the leading causes of disability and suicide. Despite its prevalence, traditional diagnosis for depression is exceedingly associated with misidentification and over-estimation, due to its subjective nature. With advances in affective computing, computational approaches make it possible to discern depression through second party physiological indicators; people observing the behaviour of depressed individuals have measurable changes in their physiological signals. We explored Blood volume pulse (BVP), Galvanic Skin Response (GSR), Skin Temperature (ST) and Pupillary Dilation (PD) from observers as valid sources to indicate depression in others. The behaviour of individuals suffering from four levels of depression was shown in 16 videos to 12 experimental observers whose physiological signals were recorded. We found that depression provokes visceral physiological reactions in observers that we can measure, resulting in neural network classification of 94% accuracy. In contrast, we also found that depression does not provoke strong conscious recognition ('verbal') in observers, which is only slightly over a chance level, at 27%.

Keywords: depression detection; physiological signals; observers; galvanic skin response; skin temperature; blood volume pulse; pupillary dilation, affective computing

1   Introduction

Major depressive disorder, or 'depression' for short, is a common but serious mental disorder that widely affects populations around the world [1]. Its cause is believed to be a combination of genetics [2] and environmental factors [3], such as, major life changes, trauma, or long-lasting exposure to difficulties. It usually presents with persistent depressed mood, loss of interest and enjoyment, feelings of sadness, guilt or low self-esteem, poor concentration, and at its worst, suicidal actions [4]. According to the World Health Organization (WHO) [1], depression is one of the leading causes of disability, affecting more than 300 million people.

Since depression comes with some observable behavioural symptoms regarding the normal expression of emotions and general functioning [5], traditional diagnostic approaches for depression rely on subjective measures of behaviours. These methods are typically involved with self-reported questionnaires such as the Beck Depression Index (BDI) [6], or clinician-assisted interview style assessments such as the Hamilton Rating Scale for Depression (HAMD) [7], which score patients' depression level by the severity of their symptoms. However, meta-analyses of depression diagnosis have indicated the wide-spread existence of both over- and under-recognition [8] [9]. The central problem is that these diagnostic tools are subjective and biased, as they are heavily associated with patients' sensitivity to symptoms and willingness to honestly reveal the symptoms [10]. Given that the accuracy of depression diagnosis correlates with reassessments and longer consultation time [8], these approaches can be time-consuming. To better serve the needs of the patient, medical profession and community, it is desirable that we find simpler and less subjective methods of depression diagnosis.

To tackle the unreliable issues of subjective assessments of depression and other emotions, research has explored the possibility of measuring emotions objectively via human physiological signals along with self-assessment reports [11], based on the demonstration that physiological signals are highly correlated with subject assessments [12]. These measures are typically automated and involve affective sensors to study changes in galvanic skin response (GSR) [13], blood volume and heart rate activities [13]–[16], pupillary sizes and eye movements [17].

That emotion can be distinguished based on physiological signals relies on the human peripheral nervous system, which consists of, the somatic nervous system (SoNS), which controls voluntary body movement, and the autonomic nervous system (ANS). The ANS is responsible for involuntary activities and, without conscious awareness, it automatically regulates bodily functions such as heart rate, respiratory rate, and pupillary responses. The ANS consists of the sympathetic nervous system (SNS) and the parasympathetic nervous system (PNS). The PNS is responsible for activities during resting and digesting states. The SNS, dominates when a person is threatened or under stress. To get the body ready to cope with danger, it expands pupils, allowing more light to enter the eyes for better vision, and increases the respiratory rate and heart rate, providing better oxygenation and easier blood flow throughout the body. It is often known as the 'fight or flight' system. Thus, we expect that SNS activation will change multiple physiological signals in the individuals undergoing such emotion.

Given that physiological responses maintained by the ANS can indicate individuals' inner states without conscious awareness, physiological signals have been used as biomarkers for depression. For example, depressed patients have different eye gaze behaviours [18], lower Galvanic Skin Response (GSR) [19], reduced Heart Rate Variability (HRV) [20], and depressed brain activities as captured by Near Infrared Spectroscopy (NIRS) [21] and Electroencephalogram (EEG) [22]. These signals provide more objective and quantitative criteria, and

when combined with machine learning technologies, can play an essential role in providing an objective assessment for depression.

Despite the physiological correlates of depressed patients, since individuals with depression tend to withdraw from social activities [1] [8], and so leaving them with less chance to access these facilities, our goal is to investigate physiological signals of *observers* to identify others' depression level. Our previous work demonstrated the feasibility of using observers' physiological signals as indicators of other individuals' depression [23] using neural networks. Subtle cues could be noticed by observers, which are reflected in observers' physiological signals. We found that neural networks trained with physiological features can recognise other individuals' depression levels with 92% accuracy. We extend this methodological and analytical approach, to ascertain whether neural networks trained on observers' BVP and associated heart rate (HR) and heart rate variability (HRV) signals can identify other individuals' depression. The identification of universal physiological indicators from observers watching depressed individuals could assist with earlier diagnosis, which, combined with known effective treatments, would decrease the burden for individuals and society. The use of physiological signals could also be applicable in other domains such as engineering [24].

This paper examines whether observers' BVP and associated heart rate (HR) and heart rate variability (HRV) signals respond to depressed individuals and whether a computational model trained with single BVP signal, as well as, trained with a hybrid of four physiological signals (GSR, BVP, ST, and PD), could better recognise other individuals' depression level. It details an experiment conducted to collect multiple physiological response signals from experiment participants who watched videos of people with various levels of depression and includes selecting optimally useful features from the response signals. The paper concludes with a summary of the findings and suggests directions for future work.

2   Experimental Design

Our aim is to detect other individuals' depression using observers' Blood Volume Pulse (BVP), Galvanic Skin Response (GSR), Skin Temperature (ST) and Pupillary Dilation (PD) signals, both singly and in combination. Following a similar experimental design to our previous work [23], we selected sixteen videos from the 2014 Audio-Visual Emotion Challenge (AVEC 2014) dataset as stimuli [25] in which individuals with four depression severities read aloud a paragraph in German. We recruited 12 participants as observers to watch the video stimuli, while we recorded their BVP, GSR, ST, and PD. We also collected observers' conscious subjective depression prediction of the individuals in the videos via a survey. A schematic diagram of the equipment setup is provided in Figure 1.
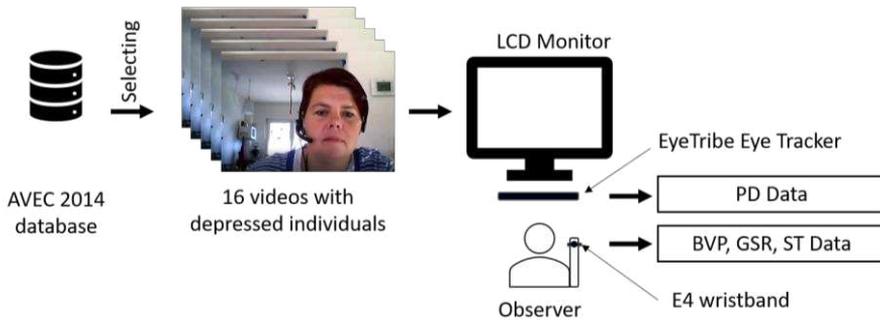
Figure 1

A schematic diagram of the equipment setup

2.1   Stimuli

We used videos from the Northwind category of the AVEC 2014 dataset [25]. The Northwind category consists of 150 webcam video recordings of 4 categories of participants individually reading aloud a paragraph in German. Each recording was labelled with a single depression level derived from involved participants self-reported depression level indicated by the Beck Depression Inventory – II (BDI-II) [6]. This index gives depression scores ranging from 0 to 63 and groups the scores into four depression categories:

- 0 - 13   Indicates no or minimal depression
- 14 - 19   Mild depression
- 20 - 28   Moderate depression
- 29 - 63   Severe depression

We chose 16 videos (see Table 1) with similar durations, from 36 s to 50 s (Ave = 41.2, Standard Deviation = 3.8), evenly across the four depression categories.

Table 1

Stimuli videos selected from the testing set of Northwind tasks in AVEC 2014

Video name	Duration (sec)	Depression level	Category
210_2, 249_1, 341_1, 240_3	43, 42, 39, 41	1, 4, 7, 11	no
220_3, 242_1, 315_3, 214_3	39, 42, 40 43	15, 16, 17, 18	mild
245_3, 218_3, 325_2. 250_1	40, 39, 39, 41	21	moderate
226_2, 359_1. 315_2, 237_1	41, 45, 58, 47	30	severe

2.2   Participants

Fourteen students who do not understand German and do not have prior training in depression recognition took part in the experiment. They were recruited to watch German-language depression videos as §2.4Procedure below describes. Ethics Approval was obtained from the Australian National University Human Research Ethics Committee. Two subjects were excluded based on the predefined exclusion criteria for having a history of cardiovascular disease or technical failures of the sensors. The final sample consisted of 12 participants, six males, and six females, from 18 to 27 years in age (mean = 21.1, standard deviation = 2.8) with normal or corrected-to-normal vision and hearing. This sample size of participants is normal for publications as a preliminary study in medicine [26].

2.3   Measures and Sensors

2.3.1   Blood Volume Pulse (BVP)

Blood Volume Pulse (BVP) indicates the volume of blood running through the vessels over time [27]. It can be measured by a photoplethysmographic (PPG) sensor using infra-red light through the skin surface and measures the reflected light. With every beat, the heart pushes a volume of blood causing a wave which travels from the heart, and returns to the heart. As the surge of blood dissipates, the signal falls. The direct pulse wave then bounces back from the lower body, causing a secondary wave, which appears as a second rise in the signal. The signal then drops until the next heartbeat. A typical BVP signal is illustrated in Figure 2, which consists of the systolic peak (Figure 2-I), dicrotic notch (2-II), diastolic peak (2-III) and diastolic point (2-IV).
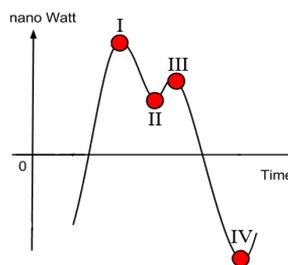
Figure 2
A typical BVP waveform

BVP can provide information on changes in SNS activation, which are influenced by emotional context. For example, BVP is negatively correlated with stress and positively correlated with sadness [28]. We can derive other cardiovascular measures from BVP, such as heart rate (HR) and heart rate variability (HRV). HR is a useful predictor of emotional valence and can distinguish between positive and negative emotions [29]. HRV refers to the temporal, beat-to-beat variations, in the consecutive heartbeats, and can indicate mental effort and emotions [28].

We placed an Empatica E4 wristband on the wrist of the non-dominant hand of observers [30], which recorded BVP with a sampling rate of 64 Hz [31].

2.3.2   Galvanic Skin Response (GSR)

Galvanic Skin Response (GSR) measures electrical conductivity of the skin, which varies due to the amount of sweat [32]. Stress or danger stimulates glands to produce salty sweat, which increases skin conductivity [33]. GSR is composed of two separate electro dermal activities: the tonic component is slow-moving and shows the general activity of the perspiratory glands caused by body or external temperature, while the phasic component is a faster distinctive waveform in the signal, and is considered to be linearly correlated with the intensity of arousal in mental state [11]. In this study, we recorded participants' wrist GSR using an Empatica E4 wristband with a sampling rate of 4 Hz [31].

2.3.3   Skin Temperature (ST)

Skin Temperature (ST) fluctuates due to vasodilatation of peripheral blood vessels induced by increased activity of the SNS. It is negatively correlated with unpleasant emotions such as stress [34] and fear [35] because blood is redirected to vital organs as a protection measure. In this study, we recorded participants' wrist ST using an Empatica E4 wristband with a sampling rate of 4 Hz [31].

2.3.4   Pupillary Dilation (PD)

Pupillary Dilation (PD) provides indications of changes in mental states and of mental activities [36]. Pupil size was found to respond to emotionally stimuli. The pupil is significantly bigger after positively or negatively arousing stimuli than after neutral stimuli [37]. We used The EyeTribe, an affordable, non-intrusive and precise eye tracker [38], to record pupil size at 60 Hz. Python code was written to analyse data collected by the EyeTribe SDK software [39].

2.4   Procedure

The experiment was conducted with each participant in the same quiet experiment room. Participants were given a written set of instructions and guidance from the experiment instructor before they provided written informed consent. An Empatica E4 sensor [31] was attached to the wrist of each participant's non-dominant hand [30], and eye gaze calibration for the eye tracker was performed.

Participants then filled in a questionnaire to collect demographic and health characteristics that may affect cardiovascular and pupillary responses. Each participant then watched 16 videos and was asked at the end of each video to respond to a question of "How would you like to rank the patient's depression level?" on a four-item scale of "None, Mild, Moderate, Severe" that matches with the BDI-II [6] scale. A five-second gap was provided between videos. The videos were presented in an order balanced way to avoid the effects of presentation order. At the end of the experiment, participants filled in the BDI-II [6] survey assessing their depression level. In total, the experiment took approximately forty minutes.

3 Methodology

Following our previous work methodology [23], we first pre-processed the physiological responses of observers to remove noise and individual bias. We then computed features for the four recorded physiological signals before we trained neural networks with the most significant features selected by a genetic algorithm. An overall structure of our depression recognition system is illustrated in Figure 3.
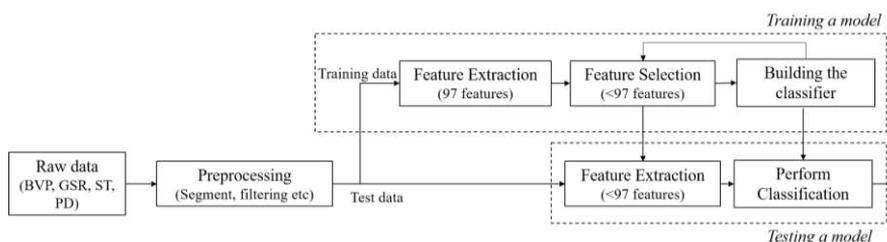
Figure 3
An overall structure of our depression recognition system

3.1 Pre-processing

For all four physiological signals, we first extracted the raw signal for all observers when they were watching the full set of 16 videos, removing the noise caused by the movement of observers, which mostly happened at the beginning and the end of the recording when they were filling in the demographic questionnaire and post-experiment survey. We then applied a cubic spline interpolation to the missing pupil size data caused by occasional eye blinks. This procedure was employed on the left and the right pupil data separately.

To reduce the between-participant differences, we then separately normalised BVP, GSR, ST, left and right PD to the range between 0 and 1 with a min-max normalisation scaler as shown in (1):

$$S_{normalised} = \frac{S - \min(S)}{\max(S) - \min(S)} \tag{1}$$

Where $S_{normalised}$ is the min-max scaled data of signal S, and $\max(S)$ and $\min(S)$ are the maximum and minimum value of signal S.

After normalisation, to remove noise artefacts, we applied a lowpass Butterworth filter to BVP, GSR and ST with an order of 6 and a cut-off frequency of 0.5 Hz, 0.2 Hz [13] and 0.3 Hz [40] to form the Low Pass (LP) BVP, LP GSR, and LP ST data, respectively. We also filtered the left and right PD data with a 10-point Hann moving window. The average pupillary size of the normalised left and right pupil data was then calculated.

Following this, we further segmented both the normalised and filtered signals by each video watching session, so that each segmented physiological data set corresponds to one observer's physiological state invoked by his or her experience of watching one video.

3.2   Feature Extraction

After pre-processing the raw signals, we generated time- and frequency-domain features that characterise the changes in the physiological signals over the time observer participants spent on watching each video.

3.2.1   Blood Volume Pulse (BVP) Features

According to the literature of using BVP for emotion recognition [41], we first calculated the following six time-domain features from the LP BVP.

1) Minimum	4) Standard deviation
2) Maximum	5) Variance
3) Mean	6) Root mean square

Let R_i be the ith systolic peak, RR_i be the interval between peak R_{i+1} and R_i, and $RR_{diff}\,i$ be the differences between intervals RR_{i+1} and RR_i (as Figure 4 shows). Heartbeats defined as the systolic peaks of the LP BVP as illustrated as R_n in Figure 4 were then identified adapting a peak detection technique devised by Van Gent et al [42]. We calculated a moving average using a window of 0.8 seconds before and after each data point. Regions of Interest (ROI) are then marked between two diastolic points where the amplitude of the signal is larger than the moving average. Systolic peaks were detected at the maximum of each ROI.
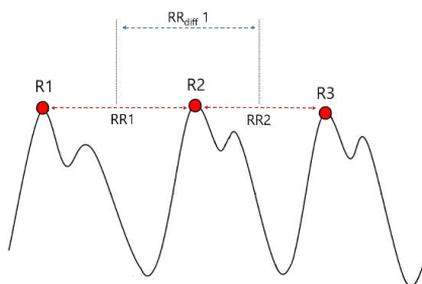
Figure 4

BVP signal with systolic peaks, peak intervals and differences between intervals annotated

To extract time-domain heart rate features, or heart rate variability, we computed the following 8 time-domain features that were previously shown to be correlated with external stimuli and mental states [43]–[45].

Inter beats interval (IBI) $IBI = \overline{RR}$ (2)

Ave. beats per min (BPM) $BPM = (60 \times Sampling\ Rate)/IBI$ (3)

Standard deviation of intervals between heart beats (SDNN)

$$SDNN = \sqrt{\frac{\sum_{n=1}^{N}(RR_i - \overline{RR})}{N-1}}$$ (4)

Standard deviation of differences of adjacent R-R intervals (SDSD)

$$SDSD = \sqrt{\frac{\sum_{n=1}^{N}(RR_{diff_i} - \overline{RR_{diff}})^2}{N-1}}$$ (4)

Root mean square of differences of adjacent R-R intervals (rMSSD)

$$rMMSSD = \sqrt{\frac{\sum_{n=1}^{N}(RR_{diff_i})^2}{N-1}}$$ (5)

Percentage of the differences greater than 20 ms (pNN20)

Percentage of the differences greater than 50 ms (pNN50)

Proportion of differences greater than 50 ms / 20 ms (pNN50/pNN20)

As indicated in [46], the spectral features of heart rate signal are more robust for short time durations and less sensitive to missing heartbeats, so we also included 5 frequency-domain features after we performed a Fast Fourier Transform (FFT) over the peak intervals to convert the signal into the frequency domain.

High frequency power (HFP): ranging from 0.15 to 0.5 Hz

Low frequency power (LFP): in a range between 0.04 and 0.15 Hz

Very low frequency power (VLFP) in a range between 0.003 and 0.04 Hz

LF/HF ratio (LFHF): the ratio of LFP over HFP

Respiratory Rate (RSP): max power in frequency range of 0.1-0.25 Hz

3.2.2   Galvanic Skin Response (GSR) Features

Sixteen time-domain features were calculated from both the normalised GSR and LP GSR separately, based on the following statistical methods:

a) Minimum

b) Maximum

c) Mean

d) Standard deviation

e) Variance

f) Root mean square

g) Mean of absolute values of first difference

h) Mean of absolute values of second difference

GSR consists of a tonic component (also called DC level) and a phasic component (also called the skin conductance response, SCR) [11]. DC level shows the long-term slow variation in the signal, indicating general activity of perspiratory glands caused by body or external temperature, while SCR reflects relatively faster responses to external stimuli. To extract the DC level component, we used a very low pass Butterworth filter with a cut-off frequency of 0.08 Hz to obtain the Very Low Pass signal (VLP). We further acquired a detrended SCR signal without DC component by removing continuous piecewise linear trend in both LP and VLP signal. Afterward, we calculated the following frequency-domain features:

- Number of SCR occurrences for VLP, LP and normalised GSR
- Mean of amplitudes of SCRs for VLP, LP and normalised GSR
- Ratio of SCR occurrences in VLP to occurrences in LP

3.2.3   Skin Temperature (ST) Features

For ST, we used a similar feature extraction approach as for the GSR signal. We calculated 16 time-domain features which include the minimum, maximum, mean, standard deviation, variance, root mean square, means of the absolute values of the first and second difference of the normalised and LP ST signal. Subsequently, we applied a very low pass Butterworth filter with a cut-off frequency of 0.08 Hz to the normalised ST signal to form the VLP ST signal. We finally calculated the numbers and amplitudes of peak occurrences for VLP and LP ST signals as well as the ratio of peak occurrences in VLP to those in LP as features.

3.2.4   Pupillary Dilation (PD) Features

Similar to the GSR and ST signals, for PD, we first calculated the following 8 features from the normalised left PD, normalised right PD, and average PD separately: minimum, maximum, mean, standard deviation, variance, root mean square, means of the absolute values of the first difference and means of the absolute values of the second difference. We then applied a very low pass Butterworth filter with a cut-off frequency of 0.08 Hz to the normalised left, right and average PD signal to form the left VLP PD, right VLP PD, and average VLP PD. Numbers and amplitudes of peak occurrences for left, right and average VLP and LP PD signals, as well as the ratio of peak occurrences in VLP to those in LP for the left, right and average signals, were subsequently extracted as features. Thus, we collected a total of 104 features from the four physiological signals: 19 (BVP) + 23 (GSR) + 23 (ST) + 39 (PD).

3.3   Feature Selection

Our previous work [23] indicated that neural networks (NNs) trained with subsets of features selected by Genetic Algorithm (GA) [47] perform better at depression prediction than without feature selection. A full set of features may include

redundant / irrelevant features that outweigh more useful features. To make a direct comparison to our prior work, we used a GA feature selection method.

The initial population for the GA was set to use all features. A candidate chromosome was defined as a binary string where the index for a bit represented a feature, and the bit value indicated whether the feature was used for classification. The presence (1) or absence (0) of every possible feature was determined based on a fitness function, which is the depression recognition performance of an NN. An example of such representation is demonstrated in Figure 5. All settings for the GA used in the hybrid classification system can be found in Table 2.

Vector of best features selected by GA	**1**	0	**1**	**1**	…
Vector of derived features	**0.2**	0.3	**0.5**	**0.1**	…
Vector for best features selected by GA	**0.2**	0	**0.5**	**0.1**	…

Figure 5

GA representation of features

Table 2

GA settings

GA Parameter	Value
Population size	100
Crossover rate	0.8
Mutation rate	1/(length of the chromosome)
Crossover type	Uniform crossover
Mutation type	Uniform mutation
Selection type	Stochastic uniform selection

3.4 Neural Network Classifiers

In this study, we were interested in determining the depression recognition capability achieved by a combination of BVP, GSR, ST, and PD measurements as monitored signals. Assessment of overall usefulness of signals is also important as fewer sensors are required if only a single signal is needed to achieve similar recognition capability. Thus, we trained five NN classifiers with the following five conditions: 1) BVP+GSR+ST+PD: using a subset of features selected by GA from all features extracted from all four signals; 2-5) each of BVP/GSR/ST/PD singly: using a subset of features selected by GA from features extracted from the

BVP/GSR/ST/PD signal. We note that when each physiological signal was used, the classifier was retrained and retested using the same validation scheme.

All five NNs performed a 4-class classification indicating 4 depression severities using 4 output neurons. The first NN was set to have a sigmoid hidden layer of 100 neurons after we tested the first NN with different hidden neuron size from 10 to 200 and found 100 to be optimal. With a similar approach, the other four NNs use 50 neurons. All NNs were trained with a commonly used optimizer, the Adam optimizer [48] using backpropagation with the Cross-Entropy loss function.

As noted in [23], that in the context of continuous physiological data, training a classifier on random splits of data is not appropriate, we used the leave-one-participant-out validation method. For each run, we took physiological features from one observer as the testing set, and those from the remaining participants as the training set. We repeated this process for all observers, each time leaving out physiological features from a different observer as the testing set. We averaged the performance as the final results reported.

3.5   Evaluation Measures

To validate the effectiveness of our models, we used *precision*, *recall,* and *F1-score* as evaluation measures. For a specific depression level L, *precision* is defined as the proportion of individuals that are correctly predicted with depression level L and actually have that level of depression; *recall* is the percentage of depressed individuals that are correctly predicted with depression level L among all individuals labelled with depression level L; and *F1-score* takes the harmonic mean of precision and recall defined as $2 \times \frac{\text{Precision}_L \times Recall_L}{Precision_L + Recall_L}$.

As multiclass depression labels were also predicted by our models, we calculated the average precision, recall, and F1 score for all depression levels as a whole, to give a view on the general prediction performance. Also, we computed the overall accuracy, which is the number of individuals correctly predicted with their corresponding depression levels by the model over total number of individuals.

4   Results and Discussion

4.1   Observers Subjective Prediction

As Table 3 shows, observers are not good at *consciously* identifying the depression severity of other individuals in videos. The overall accuracy was 27%, which is slightly over the *prima-facie* chance level of 25% since there were four options for observers over balanced numbers of video stimuli. This is consistent

with earlier findings about the accuracy of people's conscious judgments on the veracity of smiles [49], anger [50] and deceiving behaviours [51], which are all only marginally higher than chance level. Thus a prediction barely over chance in general is not surprising. However, the low recognition accuracy could be exacerbated by our observer participants being recruited from a naïve population who had not previously received any training regarding depression diagnosis. Future research should explore the accuracy of conscious judgments from psychologists who are trained to diagnose depression patients.

Table 3

Results of depression prediction from observers' subjective verbal responses

Depression level	Subjective Prediction		
	Precision	*Recall*	*F1 score*
None	0.31	0.33	0.32
Mild	0.18	0.21	0.19
Moderate	0.23	0.25	0.24
Severe	0.42	0.29	0.35
Average	0.29	0.27	0.28
Overall Accuracy	0.27		

The average ratios of consciously identifying healthy individuals and severely depressed individuals correctly were 33% and 29% respectively, higher than those of identifying depressed individuals in the middle ranges, at 21% and 25%. This could imply that people are better at identifying healthy individuals and depressed patients with severe symptoms, but worse at differentiating depression levels.

4.2   Classification based on All Physiological Signals

All features derived from observers' BVP, GSR, ST, and PD signals were provided to NNs with GA for feature selection. Performance of the classifications was calculated based on the average results of 10 runs and shown in Table 4.

Table 4

Results of depression prediction from NN trained with BVP, GSR, ST, and PD features selected by GA compared with our previous study trained with GSR, ST and PD features selected by GA

Depression level	Our previous study [23]			This study		
	Precision	*Recall*	*F1 score*	*Precision*	*Recall*	*F1 score*
None	0.92	0.95	0.94	0.90	0.98	0.94
Mild	0.93	0.89	0.91	0.94	0.92	0.93
Moderate	0.88	0.90	0.89	0.95	0.92	0.93
Severe	0.95	0.95	0.95	0.98	0.96	0.97
Average	0.92	0.92	0.92	0.94	0.94	0.94
Overall Accuracy	0.92			0.94		

As can be seen above, the overall prediction accuracy across all four depression levels was much higher than observers' conscious judgments, at 94%. Statistical analysis was conducted on the results using the Student's t-test since models trained with different physiological features share normality and equality of variances across comparison groups. In accordance with the Student's t-test, the model trained with BVP, GSR, ST, and PD features produced significantly better depression recognition rates than the conscious evaluation of observers ($p < 0.005$). This could indicate that although humans are not good at consciously detecting the depression severity of others, they can emotionally sense depression in others. Their physiological changes provoked by depression from others can be effectively detected by computational classifiers such as neural networks. The superior detecting ability of human unconscious physiological responses over conscious judgments is also found in other research in which the realness of two basic emotions are examined [22] [23]. Taken together, it could suggest that unconscious responses from instinctive human ability, which has been adaptively evolved by natural selection, can make effective use of cues to identify depressed individuals without being influenced by conscious biases.

Compared to our previous work, where only three physiological signals, GSR, ST, and PD were examined, our improved model received a statistically significantly better overall accuracy ($p < 0.01$). Our improved model also outperformed in recognising depression at mild, moderate, and severe levels with higher F1 scores ($p < 0.01$). This may imply that BVP and associated HR and HRV signals improve identifying depression severity of depressed individuals when combined with other physiological signals.

However, this model obtained a slightly less precision rate when the video individuals do not have depression, meaning that fewer video individuals with no depression were correctly identified with no depression and thus more video individuals with no depression were overestimated to have some levels of depression. It could indicate that while BVP improves the recognition rate of depressed individuals, it may also over-recognise depression in healthy individuals. This should be investigated and overcome for clinical applications.

Further, similar to clinical depression diagnosis where the middle levels of depression are harder to correctly identify [52], our model performed slightly worse in predicting depression levels of individuals with mild and moderate depression severity than those with none and severe depression level, reflected by the lower F1-scores of mild and moderate depression level. This result is also found in observers' subjective predictions and our previous work [23] [53], indicating the difficulty of accurately recognising middle levels of depression. Future research can consider exploring the feasibility of analysing more complex physiological signals, such as brain activity tracking with electroencephalogram (EEG) [54]–[56] or functional near-infrared spectroscopy (fNIRS) [52], to detect more subtle cues in observed stimuli.

4.3   Classification based on Individual Physiological Signal

To evaluate the classification capability of models with fewer physiological signals, features derived from individual physiological signals were provided to the NN model with GA feature selection. Performance of the classifications were calculated based on the average results of 10 runs and presented in Table 5.

Table 5

Results of depression prediction from NN trained with single physiological signal with GA

Depression level	BVP			GSR			ST			PD		
	Precision	*Recall*	*F1 score*	*Precision*	Recall	F1 score	*Precision*	Recall	F1 score	*Precision*	*Recall*	*F1 score*
None	0.90	0.81	0.85	**0.92**	**0.94**	**0.93**	0.87	0.83	0.85	0.91	0.92	0.91
Mild	0.84	**0.92**	0.88	0.89	0.81	0.85	0.84	0.83	0.84	**0.92**	0.91	**0.91**
Moderate	0.90	**0.95**	0.92	0.87	0.87	0.87	0.80	0.81	0.81	**0.93**	0.92	**0.92**
Severe	0.90	0.86	0.88	0.87	**0.93**	0.90	0.84	0.86	0.85	**0.95**	0.92	**0.94**
Average	0.89	0.89	0.88	0.89	0.89	0.89	0.84	0.83	0.84	**0.93**	**0.92**	**0.93**
Overall Accuracy	0.89			0.89			0.84			**0.93**		

When features from only one physiological signal were available, depression patterns were best recognised from PD features, with an average F1 score and an overall accuracy of 93%. These results were ≥4% higher than models trained with other individual signals. BVP features and GSR features contributed similarly to depression recognition in general, while ST features were less accurate. We also performed statistical analyses on the average F1 score and the overall accuracy of each pair of models trained with different single physiological signals separately. The Student's t-test has shown a significant difference between models trained with PD and models trained with BVP, GSR, or ST ($p < 0.01$). This is consistent with the literature [57] where pupil size was prominent among other signals in detecting stress, revealing some physiological signals convey more informative features to classifiers. It also indicated that models trained with ST features were significantly less likely to predict depression level ($p < 0.01$). No significance was found between models trained with BVP and GSR features ($p > 0.1$).

The contribution of each signal to the prediction of each depression level can also be seen in Table **5** based on the precision, recall, and F1 score of each depression level. PD was the best in recognising all depression levels except the "None" category, achieved by having the highest precision and F1 score ($p<0.005$). On the other hand, when identifying individuals with no depression levels, GSR obtained the best result across precision, recall, and F1 score ($p<0.05$). Taken together, it possibly shows that while PD is a valid indicator of other individuals' depression state, GSR is more useful to recognise healthy individuals.

Although models trained with PD-only features obtained an acceptable recognition performance, when compared with the models trained with a hybrid of four physiological signals, BVP, GSR, ST and PD, it performed slightly worse,

reflected by lower precision rates, recall rates and F1 scores across all four depression levels. Statistical significance was found for recognition performances of mild, moderate, and severe depression between PD and a hybrid of 4 signals ($p < 0.05$) but no significance was found when observed individuals were healthy. This phenomenon also happened in [58] [59] where a combination of multiple signals outperforms models trained with individual signals separately. It is probably because we as human beings use a combination of different modalities in our body to express emotions and thus our physiological signals have been evolving and favoured by natural selection to function as a whole [60].

6   Limitations, Future Work and Conclusions

Observers in this study are naïve individuals who do not have depression diagnosis experience and do not understand German, which is the language spoken by the individual in the videos. Future research should recruit clinicians who are skilled in diagnosing depression and German speakers to evaluate the effect of domain knowledge and language understanding on depression prediction. Stronger conclusions could be drawn in subsequent studies, with more observers involved. Different neural network settings and structures could also be explored. Other directions for future work include studying how generalizable our results are in more realistic environments such as in daily social interactions. Finally, use of more complex physiological signals, such as, brain activity tracking, with EEG and fNIRS, could also be investigated, to perform better recognition.

Conclusions

In this article, we explored the use of physiological signals from observers, to detect depression severity, of other individuals. We investigated the utility of BVP, GSR, ST, and PD from observers (singly and in combination) to predict the depression level of individuals they observed in videos. The results show that the combination of these four signals achieved an NN classification accuracy of 94%, outperforming models trained in our previous work, which did not include BVP and associated HR and HRV signals and models trained using individual signals separately. We also identified PD as the most promising physiological source for depression recognition, as well as, the potential of GSR for recognising healthy individuals among depressed patients. Future research and implementation of the findings in this area are likely to be beneficial in assisting with more objective and earlier depression diagnosis, which combined with the use of known effective treatments, could decrease the burden of depression for individuals and society.

References

[1]   M. Marcus, M. T. Yasamy, M. van Ommeren, D. Chisholm, and S. Saxena, "Depression: A Global Public Health Concern," *WHO Dep. Ment. Heal. Subst. Abus.*, 2012

[2]   R. Chandra *et al.*, "Reduced Slc6a15 in nucleus accumbens D2-neurons underlies stress susceptibility," *J. Neurosci.*, Vol. 37, No. 27, pp. 6527-6538, 2017

[3]   A. T. Beck and B. A. Alford, *Depression: Causes and treatment*. University of Pennsylvania Press, 2009

[4]   K. Hawton, C. C. i Comabella, C. Haw, and K. Saunders, "Risk factors for suicide in individuals with depression: a systematic review," *J. Affect. Disord.*, Vol. 147, No. 1-3, pp. 17-28, 2013

[5]   J. M. Girard and J. F. Cohn, "Automated audiovisual depression analysis," *Curr. Opin. Psychol.*, Vol. 4, pp. 75-79, 2015

[6]   A. T. Beck, R. A. Steer, and G. K. Brown, "Beck depression inventory-II," *San Antonio*, Vol. 78, No. 2, pp. 490-498, 1996

[7]   M. Hamilton, "A rating scale for depression," *J. Neurol. Neurosurg. Psychiatry*, Vol. 23, No. 1, p. 56, 1960

[8]   A. J. Mitchell, A. Vaze, and S. Rao, "Clinical diagnosis of depression in primary care: a meta-analysis," *Lancet*, Vol. 374, No. 9690, pp. 609-619, 2009

[9]   J. A. Bilello, "Seeking an objective diagnosis of depression," *Biomark. Med.*, Vol. 10, No. 8, pp. 861-875, 2016

[10]   J. Joshi *et al.*, "Multimodal assistive technologies for depression diagnosis and monitoring," *J. Multimodal User Interfaces*, Vol. 7, No. 3, pp. 217-228, 2013

[11]   J. Kim and E. André, "Emotion recognition based on physiological changes in music listening," *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 30, No. 12, pp. 2067-2083, 2008

[12]   J. Tomaka, J. Blascovich, R. M. Kelsey, and C. L. Leitten, "Subjective, physiological, and behavioral effects of threat and challenge appraisal.," *J. Pers. Soc. Psychol.*, Vol. 65, No. 2, p. 248, 1993

[13]   K. Hercegfi, "Heart rate variability monitoring during human-computer interaction," *Acta Polytech. Hungarica*, Vol. 8, No. 5, pp. 205-224, 2011

[14]   A. Haag, S. Goronzy, P. Schaich, and J. Williams, "Emotion recognition using bio-sensors: First steps towards an automatic system," in *Tutorial and research workshop on affective dialogue systems*, 2004, pp. 36-48

[15]   S. Khoór *et al.*, "Heart Rate Analysis and Telemedicine: New Concepts & Maths," *Acta Polytech. Hungarica*, Vol. 5, No. 1, 2008

[16]   G. Förd\Hos, I. Bosznai, L. Kovács, B. Benyó, and Z. Benyó, "Sensor-net for monitoring vital parameters of vehicle drivers," *ACTA Polytech. hungarica*, Vol. 4, No. 4, pp. 25-36, 2007

[17] P. C. Schmid, M. S. Mast, D. Bombari, F. W. Mast, and J. S. Lobmaier, "How mood states affect information processing during facial emotion recognition: an eye tracking study," *Swiss J. Psychol.*, 2011

[18] S. Scherer *et al.*, "Automatic behavior descriptors for psychological disorder analysis," in *Automatic Face and Gesture Recognition (FG), 2013 10th IEEE International Conference and Workshops on*, 2013, pp. 1-8

[19] Y.-T. Chen, I.-C. Hung, M.-W. Huang, C.-J. Hou, and K.-S. Cheng, "Physiological signal analysis for patients with depression," in *Biomedical Engineering and Informatics (BMEI), 2011 4th International Conference on*, 2011, Vol. 2, pp. 805-808

[20] F. A. Jain *et al.*, "Heart rate variability and treatment outcome in major depression: a pilot study," *Int. J. Psychophysiol.*, Vol. 93, No. 2, pp. 204-210, 2014

[21] T. Fekete, F. D. C. C. Beacher, J. Cha, D. Rubin, and L. R. Mujica-Parodi, "Small-world network properties in prefrontal cortex correlate with predictors of psychopathology risk in young children: A NIRS study," *Neuroimage*, Vol. 85, pp. 345-353, 2014

[22] N. V Thakor and S. Tong, "Advances in quantitative electroencephalogram analysis methods," *Annu. Rev. Biomed. Eng.*, Vol. 6, pp. 453-495, 2004

[23] X. Zhu, T. Gedeon, S. Caldwell, and R. Jones, "Detecting emotional reactions to videos of depression," in *IEEE International Conference on Intelligent Engineering Systems*, 2019

[24] J. Katona and A. Kovari, "A Brain--Computer Interface Project Applied in Computer Engineering," *IEEE Trans. Educ.*, Vol. 59, No. 4, pp. 319-326, 2016

[25] M. Valstar *et al.*, "AVEC 2014: 3D Dimensional Affect and Depression Recognition Challenge," *Proc. 4th ACM Int. Work. Audio/Visual Emot. Chall. (AVEC '14)* pp. 3-10, 2014

[26] R. Simon, "Optimal two-stage designs for phase II clinical trials," *Contemp. Clin. Trials*, Vol. 10, No. 1, pp. 1-10, 1989

[27] K. Gouizi, F. Bereksi Reguig, and C. Maaoui, "Emotion recognition from physiological signals," *J. Med. Eng. Technol.*, Vol. 35, No. 7, pp. 300-307, 2011

[28] N. Sharma and T. Gedeon, "Modeling a stress signal," *Appl. Soft Comput.*, Vol. 14, pp. 53-61, 2014

[29] P. Ekman, "An argument for basic emotions," *Cogn. Emot.*, Vol. 6, No. March 2014, pp. 37-41, 2008

[30] M. E. Dawson, A. M. Schell, and D. L. Filion, "The electrodermal system," *Handb. Psychophysiol.*, Vol. 2, pp. 200-223, 2007

[31]   Empatica,   "E4   wristband."   [Online]   Available:
        https://www.empatica.com/research/e4/ [Accessed: 30-May-2018]

[32]   R. M. Stern, W. J. Ray, and K. S. Quigley, *Psychophysiological recording*.
        Oxford University Press, USA, 2001

[33]   W. Boucsein, *Electrodermal activity*. Springer Science & Business Media,
        2012

[34]   F. Al-Shargie, T. B. Tang, and M. Kiguchi, "Mental stress grading based on
        fNIRS signals," in *Proceedings of the Annual International Conference of
        the IEEE Engineering in Medicine and Biology Society, EMBS*, 2016, Vol.
        2016-Octob, pp. 5140-5143

[35]   J. E. LeDoux and S. G. Hofmann, "The subjective experience of emotion: a
        fearful view," *Curr. Opin. Behav. Sci.*, Vol. 19, pp. 67-72, Feb. 2018

[36]   M. Porta, S. Ricotti, and C. J. Perez, "Emotional e-learning through eye
        tracking," in *Global Engineering Education Conference (EDUCON) 2012
        IEEE*, 2012, pp. 1-6

[37]   B. Laeng, S. Sirois, and G. Gredebäck, "Pupillometry: a window to the
        preconscious?," *Perspect. Psychol. Sci.*, Vol. 7, No. 1, pp. 18-27, 2012

[38]   E. Dalmaijer, "Is the low-cost EyeTribe eye tracker any good for
        research?," 2014

[39]   TheEyeTribe,   "The   EyeTribe."   [Online]   Available:
        http://theeyetribe.com/theeyetribe.com/about/index.html

[40]   V. Xia, N. Jaques, S. Taylor, S. Fedor, and R. Picard, "Active learning for
        electrodermal activity classification," in *Signal Processing in Medicine and
        Biology Symposium (SPMB) 2015 IEEE*, 2015, pp. 1-6

[41]   A. Kushki, J. Fairley, S. Merja, G. King, and T. Chau, "Comparison of
        blood volume pulse and skin conductance responses to mental and affective
        stimuli at different anatomical sites," *Physiol. Meas.*, Vol. 32, No. 10, p.
        1529, 2011

[42]   P. Van Gent, H. Farah, N. van Nes, and B. van Arem, "Analysing Noisy
        Driver Physiology Real-Time Using Off-the-Self Sensors: Heart Rate
        Analysis Software from the Taking the Fast Lane Project.," *J. Open Res.
        Softw.*, 2018

[43]   B. M. Appelhans and L. J. Luecken, "Heart rate variability as an index of
        regulated emotional responding," *Rev. Gen. Psychol.*, Vol. 10, No. 3, pp.
        229-240, 2006

[44]   M. Züger and T. Fritz, "Interruptibility of software developers and its
        prediction using psycho-physiological sensors," in *Proceedings of the 33rd
        Annual ACM Conference on Human Factors in Computing Systems*, 2015,
        pp. 2981-2990

[45]   C. G. Beevers, A. J. Ellis, and R. M. Reid, "Heart rate variability predicts cognitive reactivity to a sad mood provocation," *Cognit. Ther. Res.*, Vol. 35, No. 5, pp. 395-403, 2011

[46]   J. Zhai, A. B. Barreto, C. Chin, and C. Li, "Realization of stress detection using psychophysiological signals for improvement of human-computer interactions," in *SoutheastCon, 2005. Proceedings. IEEE*, 2005, pp. 415-420

[47]   J. Yang and V. Honavar, "Feature subset selection using a genetic algorithm," in *Feature extraction, construction and selection*, Springer, 1998, pp. 117-136

[48]   D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv Prepr. arXiv1412.6980*, 2014

[49]   M. Z. Hossain and T. Gedeon, "Classifying Posed and Real Smiles from Observers' Peripheral Physiology," in *11th International Conference on Pervasive Computing Technologies for Healthcare*, 2017

[50]   L. Chen, T. Gedeon, M. Z. Hossain, and S. Caldwell, "Are you really angry?: detecting emotion veracity as a proposed tool for interaction," in *Proceedings of the 29th Australian Conference on Computer-Human Interaction*, 2017, pp. 412-416

[51]   X. Zhu, Z. Qin, T. Gedeon, R. Jones, M. Z. Hossain, and S. Caldwell, "Detecting the Doubt Effect and Subjective Beliefs Using Neural Networks and Observers' Pupillary Responses," in *International Conference on Neural Information Processing*, 2018, pp. 610-621

[52]   X. Li, B. Hu, S. Sun, and H. Cai, "EEG-based mild depressive detection using feature selection methods and classifiers," *Comput. Methods Programs Biomed.*, Vol. 136, pp. 151-161, 2016

[53]   J. F. Plested, T. D. Gedeon, X. Y. Zhu, A. Dhall, and R. R. Geocke, "Detection of universal cross-cultural depression indicators from the physiological signals of observers," in *2017 Seventh International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)* 2017, pp. 185-192

[54]   K. A. Buza and J. Koller, "Classification of electroencephalograph data: A hubness-aware approach," *Acta Polytech. Hungarica*, Vol. 13, No. 2, pp. 27-46, 2016

[55]   J. Katona and A. Kovari, "Examining the Learning Efficiency by a Brain-Computer Interface System," *Acta Polytech. Hungarica*, Vol. 15, No. 3, pp. 251-280, 2018

[56]   M. Tariq, P. M. Trivailo, Y. Shoji, and M. Simic, "Comparison of Event-related Changes in Oscillatory Activity During Different Cognitive Imaginary Movements Within Same Lower-Limb," *Acta Polytech.*

Hungarica, Vol. 16, No. 2, pp. 77-92, 2019

[57] J. Zhai and A. Barreto, "Stress detection in computer users through non-invasive monitoring of physiological signals," *Blood*, Vol. 5, No. 0, 2008

[58] O. Barral, I. Kosunen, and G. Jacucci, "No need to laugh out loud: Predicting humor appraisal of comic strips based on physiological signals in a realistic environment," *ACM Trans. Comput. Interact.*, Vol. 24, No. 6, p. 40, 2018

[59] O. AlZoubi, S. K. D'Mello, and R. A. Calvo, "Detecting naturalistic expressions of nonbasic affect using physiological signals," *IEEE Trans. Affect. Comput.*, Vol. 3, No. 3, pp. 298-310, 2012

[60] L. ten Brinke, D. Stimson, and D. R. Carney, "Some evidence for unconscious lie detection," *Psychol. Sci.*, Vol. 25, No. 5, pp. 1098-1105, 2014

On the Implementation of Fixed Point Iteration-based Adaptive Receding Horizon Control for Multiple Degree of Freedom, Higher Order Dynamical Systems

Hamza Khan[†], József K. Tar[⋆]

[†]Doctoral School of Applied Informatics and Applied Mathematics,
[⋆]University Research, Innovation and Service Center,
Antal Bejczy Center for Intelligent Robotics (ABC iRob),
Óbuda University, H-1034 Budapest, Bécsi út 96/B, Hungary
E-mail: hamza.khan@phd.uni-obuda.hu, tar.jozsef@nik.uni-obuda.hu

Abstract: Based on the observation that it is very difficult to combine the mathematical frameworks of the prevailing Lyapunov function-based adaptive controllers and the traditional, optimal control-based "Model Predictive Controllers (MPC)", a novel adaptive solution was introduced to improve the operation of the "Receding Horizon Controllers (RHC)". Because at the local optima the gradient of the auxiliary function is zero, Lagrange's original "Reduced Gradient Method (RGM)" was replaced by a "Fixed Point Iteration (FPI)"-based algorithm that directly drove this gradient toward zero. According to "Banach's Fixed Point Theorem" the convergence of the method was guaranteed by a contractive function that generated the iterative sequence. The greatest modeling burden in this approach was the need for the calculation of the Jacobian of the problem, i.e. the gradient of the gradient of the auxiliary function. In the first simulations only a single "Degree of Freedom (DoF)" 2nd order nonlinear system, a van der Pol oscillator was investigated. The attempts that were made to evade the calculation of the Jacobian were finished with the conclusion that at least a rough numerical approximation of this Jacobian generally must be utilized. Though the MPC approach allows the use a great variety of cost functions and dynamical models, mathematically well established results are available only for quadratic cost functions and "Linear Time-invariant (LTI)" models. For other cost functions and models careful numerical analysis is needed. In this paper the use of non-quadratic cost functions is numerically investigated in the FPI-based adaptive RHC control of 2 DoF 2nd order nonlinear system that consists of two, nonlinearly coupled van der Pol oscillators, is considered. To guarantee lucid calculations simple functions are introduced that map the active parts of the horizon under consideration to the elements of the gradient of the auxiliary function that are calculated analytically. For the calculation of the Jacobian only a rough numerical estimation is applied. The simulation results reveal certain limitations of the suggested method.

Keywords: Optimal Control; Model Predictive Control; Receding Horizon Control; Adaptive Control; Fixed Point Iteration-based Adaptive Control; Banach Space; van der Pol Oscillator

1   Introduction

The *"Optimal Controllers (OC)"* mathematically can be formulated in strict analogy with finding the local extremum of the action functional in Classical Mechanics. The implementation of the variational calculus leads to the Hamilton – Jacobi – Bellman equations, that numerically can be carried out by *"Dynamic Programming (DP)"* introduced by Bellman in the fifties of the past century [1, 2]. Though the computational power of the computers has been tremendously increased by our days, for control applications this approach means a great computational burden, so alternative methods are in use, too. The most popular one is the heuristic RHC introduced in the late seventies of the past century [3] in which the time domain is approximated by a discrete grid, the optimization process is limited to consecutive finite fragments of this grid that are referred to as "horizons", and it is carried out by Lagrange's RGM method he invented for the use in Classical Mechanics about 1811 [4]. This approach is often referred to as *"Nonlinear Programming (NP)"*. It allows the use of various cost functions and dynamic models e.g. in chemistry [5, 6], and for its realization for "limited size problems" the professional "Solver" package of Microsoft's Excel is generally available. While its use in economic calculations is very popular (e.g. [7, 8, 9]), in optimal control problems it is less frequent (e.g. [10] also related to economic and educational issues). Based on these antecedents in [11] the dynamic control of a 1st order single variable hypothetical system, and in [12] its possible use in the model-based treatment of the illness type 1 diabetes mellitus was investigated.

It was observed in the past century [13, 14] that in the case of LTI dynamic models and quadratic cost functions considerable mathematical simplifications are generally available. In the so introduced *"Linear Quadratic Regulator (LQR)"* instead of the original RGM either the differential or the algebraic form of the Riccati equation has to be solved. (This special form can be achieved by seeking the solution of the problem in a special product form so "decoupling" two variables in the first step.) In 1724 Riccati realized that the solution of a special scalar quadratic differential equation can be obtained by solving a linear one [15]. In the case of matrix equations similar possibilities are available by the use of the method of "Schur complements" introduced by Haynsworth in 1968 [16]. The solution of these linear matrix equations can be tackled by the method of *"Linear Matrix Inequalities (LMI)"* for the realization of which efficient software packages are available [17]. For more complicated cases "state-dependent Riccati equations" are in use (e.g. [18]). The LQR method is widely used in nonlinear dynamic systems in which the available nonlinear models are often linearized for the sake of the use of this special formalism. Furthermore, the quadratic structure of the cost functions also means significant limitations.

Since it is a special variant of the *"Model Predictive Control (MPC)"*, the applicability of the RHC also depends on the reliability of the available system models.
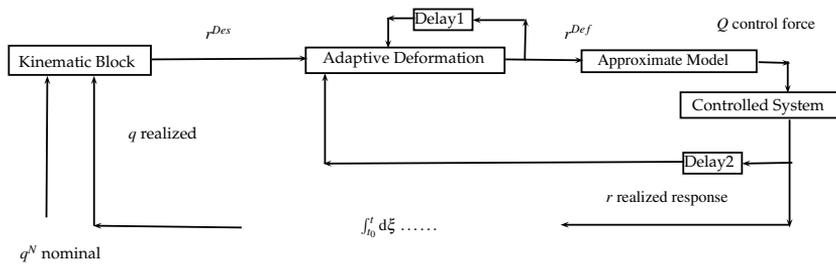
Figure 1
Schematic structure of the *"Fixed Point Transformation-based Adaptive Controller"* after [19] (the adaptive deformation can be realized by the use of various fixed point transformations, and the system's response *r* can be an arbitrary order time-derivative of the generalized coordinates of the controlled system)

The general difficulties of this approach consist in the lack of reliable system models, and limited possibilities for state observation. Illustrative examples related to various subject areas as robotics (e.g. [20, 21]), turbojet engine modeling (i.e. [22, 23, 24, 25]), life sciences as modeling the glucose-insulin system (e.g. [26, 27, 28, 29]), anaesthesia control (e.g. [30, 31, 32, 33, 34, 35]) can be mentioned. To compensate the effect of modeling imprecisions a possibility is the application of structurally precisely known analytical models with unknown parameters that can be tuned via observations by *"Adaptive Controllers (AC)"*. All the prevailing approaches are based on Lyapunov's dissertation on the stability of motion he defended in 1892 [36, 37] and appeared in the adaptive control in the nineties of the past century (e.g. [38, 39]). It does not seem to be easy to combine the mathematical framework of the Lyapunov function-based adaptive controllers with that of the cost function-based optimal controllers.

An alternative approach to adaptive control design was suggested in 2009 in [40] outlined in Fig. 1 where the controlled system's response $r(t) \equiv q^{(n)}(t)$, $n \in \mathbb{N}$, i.e. it is an order n time-derivative of its generalized coordinate q, that is the state variable consists of the components $\{q, \dot{q}, \ddot{q}, \ldots, q^{(n-1)}\}$. On the basis of purely kinetic/kinematic considerations a *"Desired Response"* $r^{Des}(t) \equiv q^{(n)^{Des}}(t)$ can be calculated that is able to drive the tracking error, i.e. the difference between the *"Nominal Trajectory"* $q^N(t)$ and the realized trajectory $q(t)$ to 0 as $t \to \infty$ if it is precisely realized. Such a strategy can be formulated e.g. by the use of the *"Integrated Tracking Error"* with a positive constant Λ as in (1), in which t_0 denotes the starting time of the control:

$$e_{Int}(t) \stackrel{def}{=} \int_{t_0}^{t} \left[q^N(\xi) - q(\xi) \right] \mathrm{d}\xi \; , \tag{1a}$$

$$\left(\frac{\mathrm{d}}{\mathrm{d}t} + \Lambda \right)^{n+1} e_{Int}(t) \equiv 0 \Rightarrow \tag{1b}$$

$$q^{(n)^{Des}}(t) = q^{N^{(n)}}(t) + \sum_{\ell=1}^{n+1} \left(\frac{(n+1)!}{\ell!(n+1-\ell)!} \right) \Lambda^{\ell} \left(\frac{\mathrm{d}}{\mathrm{d}t} \right)^{n+1-\ell} e_{Int}(t) \; . \tag{1c}$$

In the special case of $n = 2$ (1) corresponds to a PID-type error-feedback with the

proportional coefficient $K_P = 3\Lambda^2$, integral coefficient $K_I = \Lambda^3$, and derivative coefficient $K_D = 3\Lambda$. By feeding back the integrated error, if it is precisely realized, (1) makes the tracking error asymptotically converge to 0. However, if it is only *approximately realized*, the integrated tracking error can diverge. In this case the generalization of the PD-type feedback is more successful as defined in (2) as

$$e(t) \overset{def}{=} q^N(t) - q(t) \ , \tag{2a}$$

$$\left(\frac{d}{dt} + \Lambda\right)^n e(t) \equiv 0 \Rightarrow \tag{2b}$$

$$q^{(n)Des}(t) = q^{N^{(n)}}(t) + \sum_{\ell=1}^{n} \left(\frac{n!}{\ell!(n-\ell)!}\right) \Lambda^{\ell} \left(\frac{d}{dt}\right)^{n-\ell} e(t) \ . \tag{2c}$$

According to [41, 42] this feedback strategy was successfully applied for the control of ships in the first half of the past century. In the *"Kinematic Block"* of the FPI-based adaptive controller described in Fig. 1 a combined application of (1) and (2) can be successful: at the beginning, when the tracking error is still big, the controller can use (2), and later it can turn to (1) when the use of the integrated term causes more precise tracking instead of divergence.

Returning back to the scheme in Fig. 1, to compensate the effects of the modeling errors, before using the available *"Approximate Model"* for the calculation of the force or the other appropriate control signal that is necessary for producing $q^{(n)Des}(t)$, its "deformed version" $q^{(n)Def}(t)$ is calculated in the block *"Adaptive Deformation"*. In the case of digital controllers this calculation happens in discrete control cycles, and the two *"Delay"* blocks correspond to a single control cycle: in the given cycle the system "learns" from the deformed signal applied, and the system's response obtained in the previous cycle. It is important to see that in the case of an order n physical system $q^{(n)}$ can have very fast variation because it is directly set by the control force $Q(t)$, but the "desired" value $q^{(n)Des}(t)$ varies only slowly for moderate $q^{N^{(n)}}(t)$ because it is constructed of the integrated tracking error, and the order $\{0, 1, \ldots, n-1\}$ time-derivatives of the tracking error. In the block *"Adaptive Deformation"* a function of the form $q_{i+1}^{(n)Def} = G\left(q_i^{(n)Def}, q_i^{(n)}, q_{i+1}^{(n)Des}\right)$ is applied ($q_i^{(n)Def}$ is fed back at the input on the top) so that for constant $q^{(n)Des}$ it could generate a sequence of deformed signals as $\{q_1^{(n)Def}, q_2^{(n)Def} = \Phi\left(q_1^{(n)Def}\right), \ldots, q_{i+1}^{(n)Def} = \Phi\left(q_i^{(n)Def}\right), \ldots\}$ since in a given state the controlled system's response depends on $q^{(n)Def}$. With other words, during a digital control step only one step of the iteration can be realized. The possible convergence of the iteration was investigated on the basis of Banach's *"Fixed Point Theorem"* [43] according to which in a linear, normed, complete metric space a contractive map generates a sequence that converges to the unique fixed point of this map. For this purpose the contractivity of the function $\Phi\left(q^{(n)Def}\right)$ must be guaranteed. In [40, 44, 45] various solutions were suggested for the function G.

It must be emphasized that this approach may provide satisfactory results if the generated sequence has fast convergence and the slowly varying "desired" value $q^{(n)^{Des}}$ "pulls with itself" the limit point of the convergence. The applicability of this approach in each particular case needs preliminary investigations via simulation before trying any real application. For instance, in the adaptive control of a brushless electric DC motor the method was found to be successful via simulations and experimental investigations [46]. In the case of an inverted pendulum-type electric cart via simulations it was found that this approach may be successful for controlling the tilting angle, but in the control of its linear displacement it could cause not tolerable, tremendous spinning up [47].

Returning to the question of adaptive MPC or adaptive RHC in [11] a simple proposition was done. While in the original version of the RHC the problem of the modeling errors is tackled by the application of the available approximate system model for the calculation of the control signals, and normally only a small part of the calculated horizon is actually utilized, furthermore, the control is frequently redesigned by the introduction of novel horizons in the beginning of which the actually measured system variables are used as initial conditions, in [11] the following idea was suggested: instead using the control forces generated by the RHC controller, the controller can adaptively track the "optimized trajectory". In this approach the mathematical frameworks of the optimal and the adaptive controllers can be integrated without any technical difficulty.

A further step towards the FPI-based adaptive RHC was based on the observation that Lagrange's RGM may be replaced by a fixed point iteration in the optimization phase of the calculations [48, 49]. The idea was generated by considering analogies with the problem of solving the inverse kinematic task for redundant robot arms.

The prevalent literature in robotics applies an augmented Jacobian for obstacle avoidance (e.g. [50, 51]) and uses the Moore-Penrose pseudoinverse [52, 53] for the disambiguation of the otherwise ambiguous solution of this task. To utilize the advantages of the ambiguity of the solution to this special choice some elements of the null space of the Jacobian are added later (e.g. [54, 55, 56]) that makes the problem of the continuity of the so obtained solution arise. Furthermore, this pseudoinverse is calculated by the utilization of the traditional inverse of a quadratic matrix that is not invertible in the singular configurations, and is ill conditioned in their vicinity. To treat this problem the singular matrix is completed by an addendum that makes it invertible, that corresponds to a "deformation" of the original task: *"Damped Least Squares (DLS)"* [57, 58]. Desperate attempts were invented in the early nineties of the past century to tackle the problem of singularities as considering 2nd order differential inverse kinematics [59] or the complex extension of the real generalizes coordinates in [60]. In [61] D. Drexler suggested a witty idea of "action point transformation" for the regularization of the Jacobian, and in [62] an implicit 2nd order integration was suggested.

Realizing the fact that the Moore-Penrose pseudoinverse can simply be generalized by the use of different weights for the quadratic cost function contributions that play similar role as the "joint activation factors" when certain elements of the null space

of the Jacobian are mixed into the solution after computing the pseudoinverse, a matrix inversion-free, FPI-based solution of the inverse kinematic task was proposed by Csanádi et al. in [63]. The forward kinematic task is expressed by the function $f : \mathbb{R}^n \mapsto \mathbb{R}^m$ in which $n, m \in \mathbb{N}$, and $n > m$. The "initial position" $x_{ini} = f(q_{ini})$ is known, and a function of the scalar variable $t \in \mathbb{R}$, $q(t)$ has to be found for which $x(t) = f(q(t))$, $x(t_{ini}) = x_{ini}$, and $q(t_{ini}) = q_{ini}$. The *differential form* of this task is

$$\dot{x}(t) = \frac{\partial x(q(t))}{\partial q} \dot{q}(t) \equiv J(q(t))\dot{q}(t) \ . \tag{3}$$

Instead of computing the generalized inverse of $J(q)$, Csanádi introduced a finite time grid with the resolution Δt, considered a given configuration at time i as $x(t_i), q(t_i)$, computed the next position as $x(t_i + \Delta t)$, and commenced an "internal iteration" between the neighboring grid points as

$$q(t_i + \Delta t, s + 1) = G\left(q(t_i + \Delta t, s), f\left(q(t_i + \Delta t, s)\right), x(t_i + \Delta t)\right) \ ,$$
$$q(t_i + \Delta t, 1) = q(t_i) \tag{4}$$

that is inserted the task of finding the differential solution into the FPI-based adaptive control scheme depicted in Fig. 1. It soon became clear that a common Jacobian generally will not produce convergent sequence, so the task was modified as

$$J^T(q(t))x(t) = J^T(q(t))f(q(t)) \ , \text{ or on the grid} \tag{5a}$$

$$J^T(q(t_i))x(t_{i+1}) \approx J^T(q(t_i)) \left[f(q(t_i)) + \frac{\partial f}{\partial q}\bigg|_{q(t_i)} (q(t) - q(t_i)) \right] \ , \tag{5b}$$

in which a 1st order Taylor series approximation of function $f(q)$ was taken in $q(t_i)$. This task corresponds to the iteration

$$J^T(q(t_i)) \left[x(t_{i+1}) - f(q(t_i)) \right] \approx J^T(q(t_i))J(q(t_i))(q - q(t_i)) \ . \tag{6}$$

In a nonsingular configuration the positive definiteness of the matrix $J^T(q(t_i))J(q(t_i))$ guaranteed the convergence of the internal iteration, while in and in the vicinity of the singular configurations it automatically caused the "stagnation" of the appropriate axles until the vicinity of the singularity was left. That is, this approach automatically treated the problem of the singular configurations without the use of any complementary trick or task deformation, and it yielded nice solutions without risking the application of big joint coordinate time-derivatives.

If the same task is tackled on the basis of Lagrange's RGM method, the optimization under constraints can be formulated in (7)

$$\Phi \overset{def}{=} \sum_i P_i \dot{q}_i^2 = \text{ extremum under the constraints} \tag{7a}$$

$$\dot{x}_k - \sum_\ell J_{k\ell}(q)\dot{q}_\ell = 0 \ \ \forall k \ , \tag{7b}$$

in which the weights $P_i > 0$ express the "penalty" for rotating the axle i in the solution. As is well known, the *auxiliary function* of the above problem is

$$\Psi(\dot{q}, \lambda) = \sum_i P_i \dot{q}_i^2 + \sum_k \lambda_k \left(\dot{x}_k - \sum_\ell J_{k\ell} \dot{q}_\ell \right) \quad , \tag{8}$$

in which the symbols λ_k denote the Lagrange multipliers, and in the local optima $\forall i, j$ it holds that $\frac{\partial \Psi}{\partial \dot{q}_i} = 0$ and $\frac{\partial \Psi}{\partial \lambda_j} = 0$.

In [48, 49] it was recognized that driving the array $\nabla \Psi$ toward 0 is a task very similar to that formulated in (4) and (5). It was also recognized that for the validity of this structure *it is not necessary to have quadratic contributions in the cost function*. The question that $\nabla \Psi = 0$ in both the local minima and maxima of the cost function was evaded by the following argumentations:

1. In the calculation of the Moore-Penrose generalized inverse, as well as in the case of the modified version in (7), immediately matrix equations are provided after computing $\nabla \Psi$. These equations can be solved without minding if the solution belongs to a local minimum or maximum of Φ. However, it is known that the global minimum of Φ is zero that belongs to $\dot{q} = 0$ that generally does not satisfy the constraint equations in (7b). Also, it is clear that Φ does not have local maximum, it can diverge to ∞ for infinitely large $\dot{q}$ components. Therefore the iteration that is commenced from a low $\|\dot{q}\|$ will increase this norm until the reduced gradient of the problem, i.e. $\nabla \Psi$ achieves 0.

2. In Classical Thermodynamics (e.g. [64]) in the state of the thermal equilibrium the entropy of an isolated system must have its maximum under the constraint of fixed internal energy, and all the other constraints determined by the internal thermodynamic walls that partly isolate the subsystems of the big isolated system under consideration (the *"Entropy Maximum Principle"*). The entropy is monotonic increasing, unbounded function of the internal energy. Again, in finding the thermal equilibrium the zero value of the appropriate $\nabla \Psi$ is found, in which the Lagrange multipliers have definite physical interpretation.

3. Again, in Classical Thermodynamics, the *"Energy Minimum Principle"* states that in the state of thermal equilibrium the internal energy of the big isolated system takes its minimum value under the main constraint of the fixed entropy and all the other constraints determined by the internal thermodynamic walls that partly isolate the subsystems of the big isolated system under consideration. Again, the fixed entropy does not allow to reach the absolute minimum of the energy that in principle is allowed by Classical Physics. The internal energy does not have upper bound. In the solution the Lagrange multipliers have definite physical interpretation, and if a certain subsystem behaves as a reservoir, i.e. it keeps the values of certain Lagrange multipliers constant, in the formalism automatically appear the appropriate Legendre transform of the internal energy that offers the introduction of the thermodynamic potentials as the enthalpy, the Helmholtz or the Gibbs potentials. Again, the question of local minima or maxima generally is not pondered.

In the field of Optimal Control the tracking errors at the internal points of the horizon, and at the last point of the horizon, the control force components, and the constraint equations that express the dynamic properties of the controlled system together with the Lagrange multipliers provide the components of the gradient of the appropriate auxiliary function. The norm of the tracking error does not have theoretical upper bound, but it has the value 0 as the theoretical lower limit. In principle the control force components can have infinite values, so the situation is similar to that of the above considered general examples.

In [48, 49] the components of $\nabla\Psi$ and that of the Jacobian, i.e. $\nabla\nabla\Psi$ were analytically computed for single DoF systems that was found to be a great programming burden. In [65] an attempt was made to evade the calculation of the Jacobian but it lead to the conclusion that at least some numerical estimation of this matrix practically is needed to achieve convergence.

In the present paper the method is further investigated by numerical simulations for a strongly nonlinear, 2nd order, 2 DoF dynamical system and non-quadratic cost functions. Only the components of $\nabla\Psi$ are analytically calculated, the Jacobian is obtained by rough numerical approximations, due to the complexity of the programming task. In the sequel the dynamic model of the controlled system, the generation of the nominal trajectory, the cost functions, and the FPI-based adaptive RHC are investigated via numerical simulations.

2   The Dynamic Model of the Two Coupled van der Pol Oscillators

The model of the van der Pol oscillator [66] originally was developed to describe the nonlinear oscillations of an externally excited triode in 1927. For the sake of simplicity in this paper it is interpreted as a "mechanical system" with the equations of motion given in (9)

$$\ddot{q}_1 = (-k_1(q_1 - L_{10}) - b_1((q_1 - L_{10})^2 - a_1^2)\dot{q}_1 + k_{12}\,g(q_1, q_2) + F_1)/m_1 \ , \qquad (9a)$$

$$\ddot{q}_2 = (-k_2(q_2 - L_{20}) - b_2((q_2 - L_{20})^2 - a_2^2)\dot{q}_2 - k_{12}\,g(q_1, q_2) + F_2)/m_2 \ , \qquad (9b)$$

$$g(q_1, q_2) = |q_2 - q_1 - L_{12}|^{\sigma}\,\text{sign}(q_2 - q_1 - L_{12}) \ , \qquad (9c)$$

with the parameter values given in Table 1. The generalized coordinates are $q_1\,[m]$ and $q_2\,[m]$, and $F_1\,[N]$ and $F_2\,[N]$ denote the control forces. The different signs of the coupling forces in (9a) and (9b) expresses the fact that they are each other's reaction forces. The *"excitation/damping separator"* parameters separate from each other the excited and the damped regions. If no interaction occurs between the oscillators (i.e. $k_{12} = 0$) the $q_1 = L_{10}$, $q_2 = L_{20}$, $\dot{q}_1 = 0$, and $\dot{q}_2 = 0$ state corresponds to an unstable equilibrium point of these oscillators. Even an infinitesimally small perturbation that kicks the system out of this equilibrium point causes nonlinear oscillations that are bounded because for too big amplitudes the excitation terms turn into damping factors. The motion of the free oscillators is depicted in Figs. 2

and 3. In the control simulation the trajectory depicted in Fig. 2 was in use as the "Nominal Trajectory".

Table 1
The applied exact and approximate model parameters

Parameter	Exact	Approximate
$m_1\,[kg]$ mass	1.0	$0.5 \cdot m_1$
$m_2\,[kg]$ mass	2.0	$1.2 \cdot m_2$
$k_1\,[N \cdot m^{-1}]$ spring constant	10.0	$1.2 \cdot k_1$
$k_2\,[N \cdot m^{-1}]$ spring constant	15.0	$0.8 \cdot k_2$
$L_{10}\,[m]$ zero force position	-5.0	$1.3 \cdot L_{10}$
$L_{20}\,[m]$ zero force position	3.0	$0.9 \cdot L_{20}$
$b_1\,[N \cdot s \cdot m^{-1}]$ viscous damping/excitation coeff.	1.2	$1.2 \cdot b_1$
$b_2\,[N \cdot s \cdot m^{-1}]$ viscous damping/excitation coeff.	1.3	$1.2 \cdot b_2$
$a_1\,[m]$ excitation/damping separator	$\sqrt{0.4}$	$0.5 \cdot a_1$
$a_2\,[m]$ excitation/damping separator	$\sqrt{0.5}$	$1.2 \cdot a_2$
σ [nondim.] nonlinearity exponent	1.2	$1.2 \cdot \sigma$
$k_{12}\,[N \cdot m^{-\sigma}]$ nonlinear stiffness	16.0	$1.1 \cdot k_{12}$
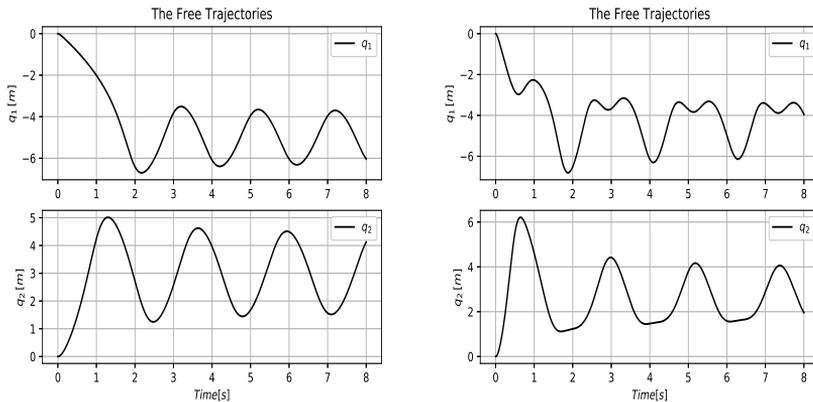
Figure 2
The trajectories of the free motion ($F_1 \equiv F_2 \equiv 0$)) of the decoupled oscillators (i.e. the $k_{12} = 0$ case) (LHS) and that of the coupled one obtained by Euler integration of (9) using the "Exact" parameters and a fixed time-step $\delta t = 10^{-3}\,[s]$

3 The Adaptive RHC Controller

In the numerical simulations in a horizon 6 grid points were applied of which the 1st and the 2nd one determined the *initial conditions* of the 2nd order dynamical system. The constraints within the horizon were formulated by the approximation as $(q_H(i+2) - 2q_H(i+1) + q_H(i))/\delta t^2 - \ddot{q}_H(i) = 0$, i.e. the optimization task
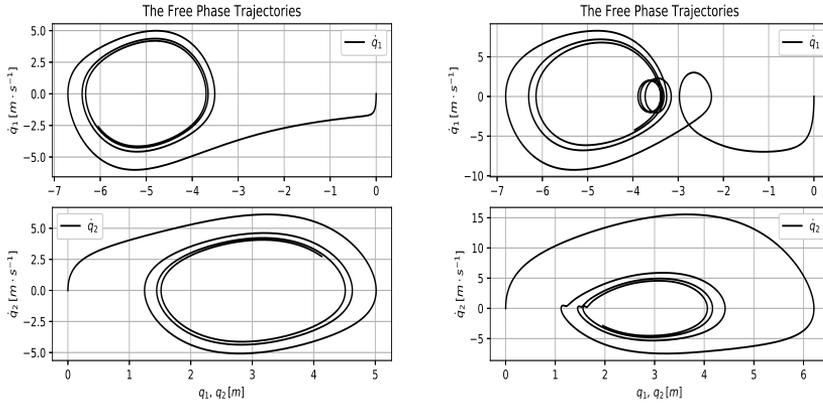
Figure 3
The phase trajectories of the free motion ($F_1 \equiv F_2 \equiv 0$)) of the decoupled (LHS) and the coupled (RHS) oscillators (i.e. when k_{12} takes the "Exact" value in Table 1) obtained by Euler integration of (9) using the "Exact" parameters and a fixed time-step $\delta t = 10^{-3}\,[s]$

had 24 variables within a horizon as from x_1 to x_4 as $q_{H1}(3),\ldots,q_{H1}(6)$, from x_5 to x_8 as $q_{H2}(3),\ldots,q_{H2}(6)$, from x_9 to x_{12} as $F_{H1}(1),\ldots,F_{H1}(4)$, from x_{13} to x_{16} as $F_{H2}(1),\ldots,F_{H2}(4)$, from x_{17} to x_{20} as $\lambda_{H1}(1),\ldots,\lambda_{H1}(4)$, and finally from x_{21} to x_{24} as $\lambda_{H2}(1),\ldots,\lambda_{H2}(4)$. For feeding the input of the model in (9) the numerical differentiator functions as $der1(q_H(i+1),q_H(i)) = (q_H(i+1)-q_H(i))/\delta t$ were introduced in a Julia [67] code. The cost functions were defined for the horizon points $i = 3,4,5$ as $C(q^N,q^O) = \log\left(1+M\left|\frac{q^N-q^O}{A_C}\right|^{\alpha_C}\right)$, for the terminal point $i = 6$ as $T(q^N,q^O) = \log\left(1+M\left|\frac{q^N-q^O}{A_T}\right|^{\alpha_T}\right)$ (the superscript N denotes the nominal trajectory, while O belongs to the optimized trajectory), and for the control force components for $i = 1,2,3,4$ the cost function is $P(F^O) = B_u\log\left(1+M\left|\frac{F^O}{A_F}\right|^{\alpha_F}\right)$ in which the parameters have lucid geometric interpretations as A_C is responsible for the *"strictness" of tracking*: if $\alpha_C > 1$ the penalty is small for $|q^N - q^O| < A_C$, and strongly increases for $|q^N - q^O| > A_C$. The parameter α_C determines the *"sharpness"* of the increase of the penalty function. Similar interpretation can be given to the parameters that limit the applicable control force components. The role of parameter $M > 0$ is "softening" the penalty function for the too big error or force components within the function $log(\cdot)$.

In the simulations different adaptive deformation functions were applied in the appropriate block of Fig. 1. Both versions were introduced by Dineva in [44, 45] as

$F : \mathbb{R} \mapsto \mathbb{R}$ defined as $F(x) = \text{atanh}(\tanh(x+D_{ad})/2)$,

$F(x_\star) = x_\star$, $h_n \overset{def}{=} r_n - r_{n+1}^{Des}$,   (10)

$r_{n+1}^{Def} = (F(x_\star + A_c\|h_n\|) - x_\star)\dfrac{h_n}{\|h_n\| + \varepsilon} + r_n^{Def}$

that was used for the optimization of the horizon, and a similar one in which in the role of the scalar function $F(x) = \frac{x}{2} + D_{ad}$ was applied for the adaptive tracking. In both functions $D_{ad} = 0.3$ was set. For the optimization, a very cautious value $A_c = -10^{-3}$, and for the adaptive trajectory tracking $A_c = -1.0$ was chosen. The role of the small parameter $\varepsilon = 10^{-10}$ was the evasion of division by zero in the calculation of the unit vector in (10). In the sequel simulation results are presented. The Jacobian was approximated by separately increasing the components of x with $\delta x = 10^{-10}$ and the partial derivatives were estimated accordingly. The last two points of a given horizon were used as the first two initial points of the next horizon. In the optimization process the Lagrange multipliers and the control forces started from the 0 values, and the optimized trajectories were initiated with the nominal ones.

4   Simulation Results

In the 1st step it is investigated that how the cost functions can cooperate with the finite element approximation of the constraints, i.e. the dynamic model of the controlled system. The tracking parameters were $A_C = A_T = 10^{-2}$, $\alpha_C = \alpha_T = 6.0$, the penalization of the control forces was switched off by $B_u = 0$, the "Approximate" model parameters were identical with the "Exact" ones an in both of them $k_{12} = 0$ was set. Only one step was made in the internal iteration, and the adaptive dynamic tracking was switched off. In the cost functions $M = 10^{-1}$ taming parameter was used. In this case it was expected that the tracking of the nominal trajectory could be possible. The tracking results in Fig. 4 reveal that an accumulated delay was produced in the tracking of the nominal trajectory that partly might be caused by the finite element approximation of the 2nd time-derivatives in the constraints, and partly by the taming parameter in the cost function. Following the removal of "cost function taming" (i.e. directly using the form $C(q^N, q^O) = \left| \frac{q^N - q^O}{A_C} \right|^{\alpha_C}$) practically the same result were obtained. *This fact reveals a significant weak point of the RHC-based approach in the case of higher order dynamical systems, because it roots in the finite element approximation of the higher order time-derivatives.*

Observing that the "optimized" trajectories remain in the vicinity of the "realized ones" since the optimization is initiated from the actual initial values, in the adaptive tracking control instead of the "optimized" ones the trajectories estimated by the dynamic model of the optimization (denoted in the figures as $q^{RealEst}$) can be adaptively tracked. These signals are far smoother than the "optimized" trajectory.

The effect of adaptivity in the dynamic trajectory tracking is revealed by Fig. 5. In this case the "Approximate Model Values" were restored, as well as the exact and approximate values of k_{12}. (It has to be remarked that in the beginning a PD-type tracking defined in (2) was applied that later was switched to a PID-type tracking defined in (1) with $\Lambda = 6.0 \left[s^{-1} \right]$, because in the initial part the PID-type strategy always caused divergence.) The significance of the coupling force and the modeling errors as well as that of the adaptive dynamic tracking are well illustrated by these figures. Figure 6 shows that the iteration reduced the norm of $\|\nabla \Psi\|$, and quite con-

siderable difference can be revealed between the "optimized" and the "adaptively set" control forces.
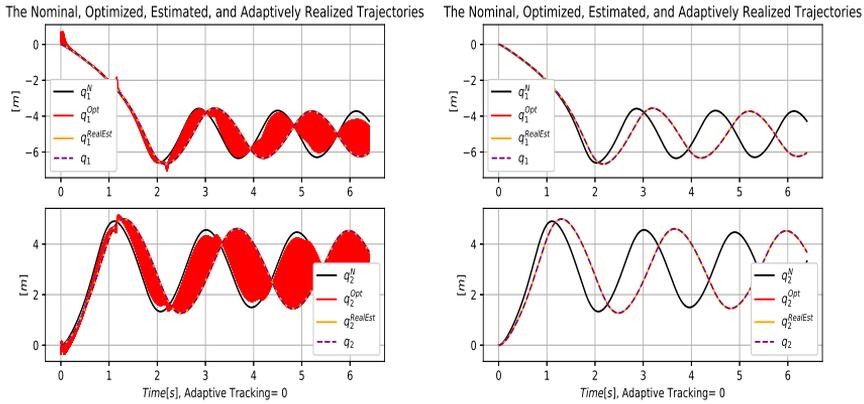
Figure 4
The test trajectory tracking for the decoupled oscillators with (LHS) and without (RHS) showing the optimized trajectory
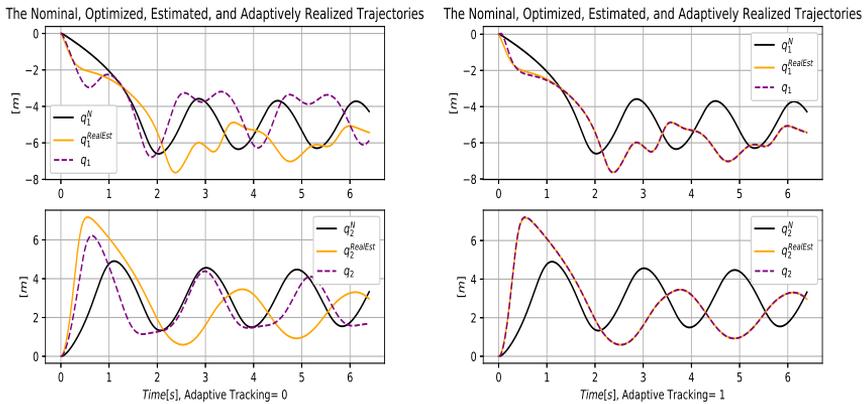
Figure 5
Tracking of the trajectories estimated by the dynamic model of the optimization without (LHS) and with (RHS) adaptivity

In Fig. 7 simulation results are given for $B_u = 10^4$, $A_F = 1.0$, $\alpha_F = 6.0$, and 100 steps in the internal iteration (in which after 10 steps the Jacobian was re-estimated). It can well be seen that the adaptive RHC controller's parameters behave according to the qualitative expectations. Figure 8 testifies that the adaptive deformation of the 2nd time-derivatives worked correctly, and that the eigenvalues of the estimated $J^T J$ matrix vary over a huge range that results in slow convergence for the smaller eigenvalues.

Finally the adaptive tracking with the modified parameters $A_C = 10^{-3}$, $\alpha_C = 8.0$,
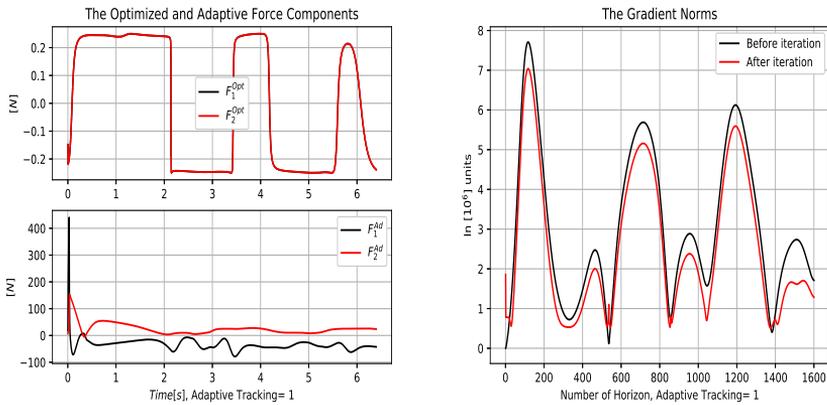
Figure 6
The "optimized" and the adaptively set control forces (LHS), and $\|\nabla\Psi\|$ before and after the optimizing iteration
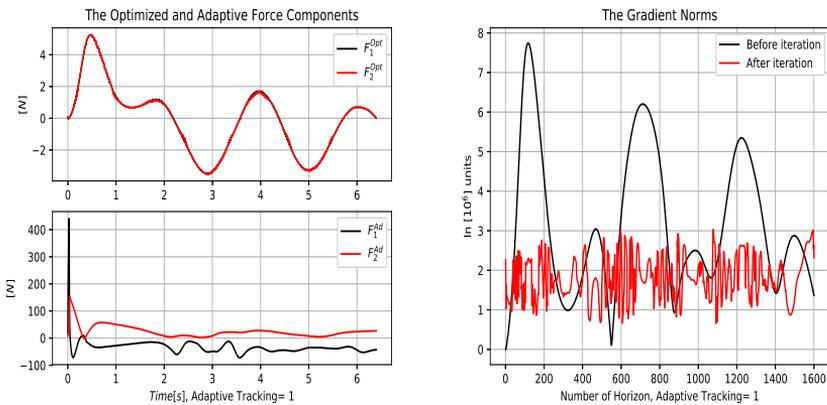
Figure 7
The "optimized" and the adaptively set control forces (LHS), and $\|\nabla\Psi\|$ before and after 100 steps of the optimizing iteration

$A_T = 10^{-3}$, $\alpha_T = 8.0$, $A_F = 10^2$, $\alpha_F = 6.0$, $B_u = 10^4$ and 21 internal steps in the iteration were investigated. Figure 9 reveals that $\|\nabla\Psi\|$ can be reduced essentially during 21 steps, and good adaptive tracking can be achieved. Figure 10 provides information on the Lagrange multipliers, the "optimized" and the adaptive force components.

Conclusions

In this paper the operation of a novel, fixed point iteration-based adaptive RHC was investigated for a 2 DoF 2nd order dynamical system consisting of two, nonlinearly coupled van der Pol oscillators. The method uses essentially the same iteration (with
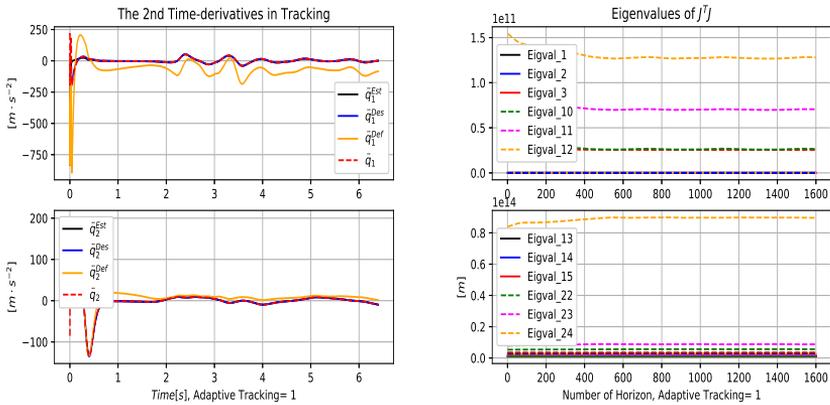
Figure 8
The 2nd time-derivatives of the adaptive tracking (LHS), and the 24 eigenvalues of the estimated $J^T J$ matrix
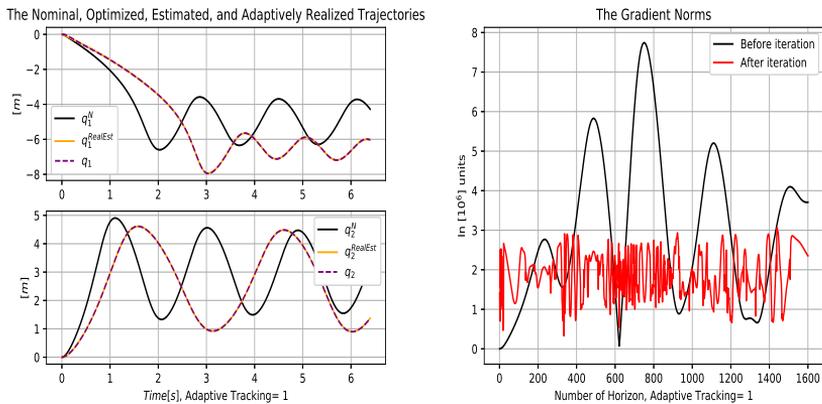
Figure 9
Adaptive trajectory tracking (LHS), and $\|\nabla\Psi\|$ before and after 21 steps of the optimizing iteration

different parameters) for the calculation of the "optimized" trajectory and for the adaptive tracking of the trajectory that was calculated by the approximate dynamic model also used in the optimization.

In the simulations only the gradient of the auxiliary function of the problem was analytically calculated by the use of simple internal functions that map the active (i.e. variable) quantities of the horizon to the elements of this gradient and vice versa. To evade programming complications the Jacobian of the problem was only "roughly", numerically estimated.

It was found via simulations that the main limitation of the present approach roots in the finite element approximation of the 2nd time-derivatives of the optimized tra-
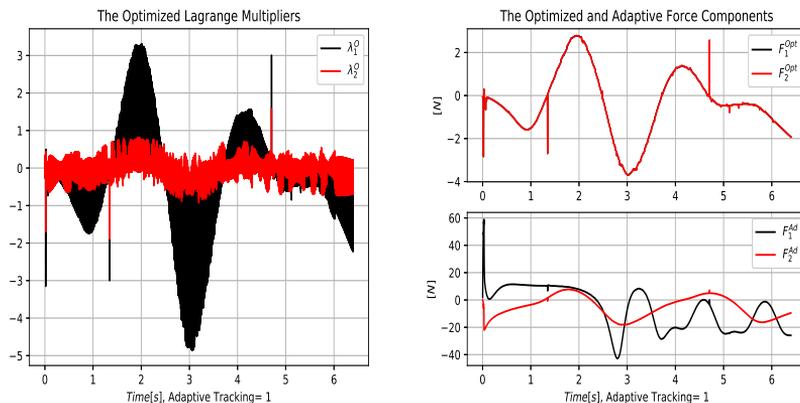
Figure 10
The Lagrange multipliers of the adaptive trajectory tracking (LHS), and the control forces (RHS) using 21 steps in the optimizing iteration

jectory used in the constraint terms of the optimization. With the exception of this problem the simulations supported the qualitative expectations based on the mathematical structure of the suggested method.

The present simulations were restricted to a relatively "short" horizon consisting of 6 grid points of which only 4 points were "active". It can be assumed that by increasing the horizon length the optimization gets more possibility for improving the "delay" caused by the applied numerical tackling of the constraints. The method deserves more numerical investigations.

Acknowledgement

References

[1]   R.E. Bellman. Dynamic programming and a new formalism in the calculus of variations. *Proc. Natl. Acad. Sci.*, 40(4):231–235, 1954.

[2]   R.E. Bellman. *Dynamic Programming*. Princeton Univ. Press, Princeton, N. J., 1957.

[3]   J. Richalet, A. Rault, J.L. Testud, and J. Papon. Model predictive heuristic control: Applications to industrial processes. *Automatica*, 14(5):413–428, 1978.

[4]   J.L. Lagrange, J.P.M. Binet, and J.G. Garnier. *Mécanique analytique (Eds. J.P.M. Binet and J.G. Garnier)*. Ve Courcier, Paris, 1811.

[5]   S. Rohani, M. Haeri, and H.C. Wood. Modeling and control of a continuous crystallization process Part 2. Model predictive control. *Computers & Chem. Eng.*, 23:279, 1999.

[6]   N. Moldoványi. *Model Predictive Control of Crystallisers*. PhD Thesis, Department of Process Engineering, University of Pannonia, Veszprém, Hungary, 2012.

[7]   J. Houston. Economic optimisation using Excel's Solver: A macroeconomic application and some comments regarding its limitations. *Computers in Higher Education Economics Review*, 11(1):2–5, 1997.

[8]   J.I. Silva and A. Xabadia. Teaching the two-period consumer choice model with EXCEL-SOLVER. *Australasian Journal of Economics Education*, 10(2):24–38, 2013.

[9]   A. Ionescu. Microsoft Office Excel 2010 operational research in Excel 2010. *International Journal of Computer and Information Technology*, 02(05):1026–1064, 2013.

[10] E. Nævdal. Solving continuous-time optimal-control problem with a spreadsheet. *Journal of Economic Education*, 34(2):99–122, 2003.

[11] H. Khan, Á. Szeghegyi, and J.K. Tar. Fixed point transformation-based adaptive optimal control using NP. *In Proc. of the 2017 IEEE 30th Jubilee Neumann Colloquium, November 24-25, 2017, Budapest, Hungary*, pages 35–40, 2017.

[12] H. Khan, J.K. Tar, I. Rudas, L. Kovács, and Gy. Eigner. Receding Horizon Control of Type 1 Diabetes Mellitus by using Nonlinear Programming. *Complexity*, `https://doi.org/10.1155/2018/4670159`, 2018.

[13] R.E. Kalman. Contribution to the theory of optimal control. *Boletin Sociedad Matematica Mexicana*, 5(1):102–119, 1960.

[14] Brian D.O. Anderson and John B. Moore. *Optimal Control: Linear Quadratic Methods*. Prentice – Hall International, Inc., A Division of Simon & Schuster, Englewood Cliffs, NJ 07632, 1989.

[15] J. Riccati. Animadversiones in aequationes differentiales secundi gradus (observations regarding differential equations of the second order). *Actorum Eruditorum, quae Lipsiae publicantur, Supplementa,*, 8:66–73, 1724.

[16] E.V. Haynsworth. On the Schur complement. *Basel Mathematical Notes*, BMN 20:17, 1968.

[17] S. Boyd, L.E. Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in Systems and Control Theory*. SIAM books, Philadelphia, 1994.

[18] Tayfun Çimen. State-dependent Riccati equation in nonlinear optimal control synthesis. *In the Proc. of the Special International Conference on Complex Systems: Synergy of Control, Communications and Computing - COSY 2011, Hotel Metropol Resort, Ohrid, Republic of Macedonia, September, 16 – 20, 2011*, pages 321–332, 2011.

[19] H. Redjimi and J.K. Tar. On the effects of time-delay on precision degradation in fixed point transformation-based adaptive control. *In the Proc. of the 2017*

IEEE 30th Jubilee Neumann Colloquium, November 24-25, 2017, Budapest, Hungary, pages 125–130, 2017.

[20] B. Armstrong, O. Khatib, and J. Burdick. The explicit dynamic model and internal parameters of the PUMA 560 arm. *Proc. IEEE Conf. On Robotics and Automation 1986*, pages 510–518, 1986.

[21] P.I. Corke and B. Armstrong-Helouvry. A search for consensus among model parameters reported for the PUMA 560 robot. *Proc. IEEE Conf. Robotics and Automation, 1994*, pages 1608–1613, 1994.

[22] L. Madarász, R. Andoga, L. Főző, and T. Lázár. *Situational control, modeling and diagnostics of large scale systems, In I. Rudas, J. Fodor, J. Kacprzyk (eds.) Towards Intelligent Engineering and Information Technology pp. 153-164*. Springer Verlag, Heidelberg, 2009.

[23] R. Andoga, L. Madarász, L. Főző, T. Lazar, and V. Gašpar. Innovative approaches in modeling, control and diagnostics of small turbojet engines. *Acta Polytechnica Hungarica*, 10(5):81–99, 2013.

[24] L. Főző, R. Andoga, and R. Kovács. Thermo-dynamic cycle computation of a micro turbojet engine. *In Proc. of the 17th IEEE International Symposium on Computational Intelligence and Informatics (CINTI 2016), 17-19 November 2016, Budapest, Hungary*, pages 75–79, 2016.

[25] K. Beneda and L. Főző. Identification of small scale turbojet engine with variable exhaust nozzle. *In Proc. of the IEEE 15th International Symposium on Applied Machine Intelligence and Informatics (SAMI 2017), January 26-28, 2017, Herl'any, Slovakia*, pages 73–78, 2017.

[26] R.N. Bergman, Y.Z. Ider, C.R. Bowden, and C. Cobelli. Quantitative estimation of insulin sensitivity. *Am. J. Physiol. Endocrinol. Metab.*, 236:E667–E677, 1979.

[27] C. Dalla Man, R.A. Rizza, and C. Cobelli. Meal simulation model of glucose-insulin system. *IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING*, 54(10):1740–1749, 2007.

[28] Gy. Eigner, P. Horváth, J. Tar, I. Rudas, and L. Kovács. Application of Robust Fixed Point control in case of T1DM. *In Proc. of the IEEE International Conference on Systems, Man, and Cybernetics 2015 (IEEE SMC 2015), Hong Kong, China*, pages 2459–2464, 2015.

[29] Gy. Eigner, J.K. Tar, I.J. Rudas, and L. Kovács. LPV-based quality interpretations on modeling and control of diabetes. *Acta Polytechnica Hungarica*, 13(1):171–190, 2016.

[30] T.W. Schnider, C.F. Minto, P.L. Gambus, C. Andersen, D.B. Goodale, S.L. Shafer, and E. Youngs. The influence of method of administration and covariates on the pharmacokinetics of propofol in adult volunteers [clinical investigations]. *Anesthesiology*, 88(5):1170–1182, 1998.

[31] A. Morley, J. Derrick, P. Mainland, B.B. Lee, and T.G. Short. Closed loop control of anaesthesia: an assessment of the bispectral index as the target of control. *Anaesthesia*, 55:953–959, 2000.

[32] F. Padula, C. Ionescu, N. Latronico, M. Paltenghi, A. Visioli, and G. Vivacqua. A gain-scheduled PID controller for Propofol dosing in anesthesia. *In: Preprints of the 9th IFAC Symposium on Biological and Medical Systems, The International Federation of Automatic Control Berlin, Germany, Aug. 31 - Sept. 2, 2015*, pages 545–550, 2015.

[33] M.M.R.F. Struys, H. Vereecke, A. Moerman, E.W. Jensen, D. Verhaegen, N. De Neve, F.J.E. Dumortier, and E.P. Mortier. Ability of the Bispectral Index, autoregressive modelling with exogenous input-derived auditory evoked potentials, and predicted Propofol concentrations to measure patient responsiveness during anesthesia with Propofol and Remifentanil. *Anesthesiology*, 99(4):802–812, 2003.

[34] T. Zikov, S. Bibian, G.A. Dumont, M. Huzmezan, and C.R. Ries. Quantifying cortical activity during general anesthesia using wavelet analysis. *IEEE Transactions on Biomedical Engineering*, 53(4):617–632, 2006.

[35] N. Ferdous and A. Kiber. Estimating depth of anesthesia from EEG signals using wavelet transform. *International Journal of Intelligent Information Systems*, 3(4):40–44, 2014.

[36] A.M. Lyapunov. *A General Task about the Stability of Motion. (in Russian)*. Ph.D. Thesis, University of Kazan, Tatarstan (Russia), 1892.

[37] A.M. Lyapunov. *Stability of Motion*. Academic Press, New-York and London, 1966.

[38] Jean-Jacques E. Slotine and W. Li. *Applied Nonlinear Control*. Prentice Hall International, Inc., Englewood Cliffs, New Jersey, 1991.

[39] R. Isermann, K.H. Lachmann, and D. Matko. *Adaptive Control Systems*. Prentice-Hall, New York DC, USA, 1992.

[40] J.K. Tar, J.F. Bitó, L. Nádai, and J.A. Tenreiro Machado. Robust Fixed Point Transformations in adaptive control using local basin of attraction. *Acta Polytechnica Hungarica*, 6(1):21–37, 2009.

[41] S. Bennett. Nicholas Minorsky and the automatic steering of ships. *IEEE Control Systems Magazine*, 4(4):10–15, 1984.

[42] S. Bennett. *A history of control engineering, 1930-1955. (In: IEE Control Engineering Series 47)*. Peter Peregrinus Ltd. on behalf of the Institution of Electrical Engineers, 1993.

[43] S. Banach. Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales (About the Operations in the Abstract Sets and Their Application to Integral Equations). *Fund. Math.*, 3:133–181, 1922.

[44] A. Dineva, J.K. Tar, A.R. Várkonyi-Kóczy, and V. Piuri. Generalization of a Sigmoid Generated Fixed Point Transformation from SISO to MIMO systems. *In Proc. of the IEEE* 19th *International Conference on Intelligent Engineering Systems, September 3-5, 2015, Bratislava, Slovakia (INES 2015)*, pages 135–140, 2015.

[45] A. Dineva, J.K. Tar, A. Várkonyi-Kóczy, and V. Piuri. Adaptive control of underactuated mechanical systems using improved "Sigmoid Generated Fixed Point Transformation" and scheduling strategy. *In Proc. of the* 14th *IEEE International Symposium on Applied Machine Intelligence and Informatics, January 21-23, 2016, Herl'any, Slovakia*, pages 193–197, 2016.

[46] T. Faitli. *Investigation of Control Methods for a Speed-controlled Electric Motor (BSc Thesis, Supervisor: J.K. Tar)*. Óbuda University, Donát Bánki Faculty of Mechanical and Safety Engineering, Institute of Mechatronics and Autotechnics, Budapest, Hungary, 2018.

[47] H. Khan, T. Faitli, T. Szili, and J.K. Tar. Preliminary investigation on the possible adaptive control of an inverted pendulum-type electric cart. *In Proc. of the 18th IEEE International Symposium on Computational Intelligence and Informatics (CINTI 2018), Nov. 21-22, 2018, Budapest, Hungary*, pages 109–113, 2018.

[48] H. Khan, J.K. Tar, I.J. Rudas, and Gy. Eigner. Adaptive model predictive control based on fixed point iteration. *WSEAS Transactions on Systems and Control*, 12:347–354, 2017.

[49] H. Khan, J.K. Tar, I.J. Rudas, and Gy. Eigner. Iterative solution in Adaptive Model Predictive Control by using Fixed-Point Transformation method. *International Journal of Mathematical Models and Methods in Applied Sciences*, 12:7–15, 2018.

[50] O. Egeland. Task-space tracking with redundant manipulators. *IEEE Journal on Robotics & Automation*, 3(5):471–475, 1987.

[51] P. Chiacchio, S. Chiaverini, L. Sciavicco, and B. Siciliano. Closed-loop inverse kinematics schemes for constrained redundant manipulators with task space augmentation and task priority strategy. *International Journal of Robotics Research*, 10(4):410–425, 1991.

[52] E.H. Moore. On the reciprocal of the general algebraic matrix. *Bulletin of the American Mathematical Society*, 26(9):394–395, 1920.

[53] R. Penrose. A generalized inverse for matrices. *Proceedings of the Cambridge Philosophical Society*, 51:406–413, 1955.

[54] N. Mansard, O. Khatib, and A. Kheddar. A unified approach to integrate unilateral constraints in the stack of tasks. *IEEE Transactions on Robotics*, 25(3):670–685, 2009.

[55] T. Petric and L. Zlajpah. Smooth continuous transition between tasks on a kinematic control level: Obstacle avoidance as a control problem. *Robotics & Autonomous Systems*, 61(9):948–959, 2013.

[56] E. Simetti and G. Casalino. A novel practical technique to integrate inequality control objectives and task transitions in priority based control. *Journal of Intelligent & Robotic Systems*, 84(1-4):1–26, 2016.

[57] K. Levenberg. A method for the solution of certain non-linear problems in least squares. *Quarterly of Applied Mathematics*, 2:164–168, 1944.

[58] S. Chiaverini, O. Egeland, and R.K. Kanestrom. Achieving user-defined accuracy with damped least squares inverse kinematics. *1991 IEEE International Conference on Robotics and Automation, June 19-22, 1991, Pisa, Italy*, page 672, 1991.

[59] E.D. Pohl. A new method of robotic rate control near singularities. *1991 IEEE International Conference on Robotics and Automation, June 19-22, 1991, Pisa, Italy*, page 405, 1991.

[60] E.D. Pohl and H. Lipkin. Complex robotic inverse kinematic solutions. *J. Mech. Des.*, 115(3):509–514, 1993.

[61] D.A. Drexler. *New Methods for Solving The Inverse Kinematics Problem of Serial Robot Manipulators (PhD Thesis, Supervisor: I. Harmati)*. Budapest University of Technology and Economics, Department of Control Engineering and Information Technology, Budapest, Hungary, 2014.

[62] D.A. Drexler. Closed-loop inverse kinematics algorithm with implicit numerical integration. *Acta Polytechnica Hungarica*, 14(1):147–161, 2017.

[63] B. Csanádi, J.K. Tar, and J.F. Bitó. Matrix inversion-free quasi-differential approach in solving the inverse kinematic task. *In Proc. of the* 17th *IEEE International Symposium on Computational Intelligence and Informatics (CINTI 2016), 17-19 November 2016, Budapest, Hungary*, pages 61–66, 2016.

[64] H.B. Callen. *Thermodynamics and an Introduction to Thermostatistics (Second Edition)*. John Wiley & Sons, New York, Chichester, Brisbane, Toronto, Singapore, 1960.

[65] Hamza Khan and József K. Tar. On replacing Lagrange's "Reduced Gradient Algorithm" by simplified fixed point iteration in adaptive model predictive control. *In Proc. of the* 23rd *IEEE International Symposium on Intelligent Engineering Systems (INES 2019), Gödöllő, Hungary, April 25-27, 2019*, pages 201–206, 2019.

[66] B. van der Pol. Forced oscillations in a circuit with non-linear resistance (reception with reactive triode). *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 7(3):65–80, 1927.

[67] J. Bezanson, A. Edelman, S. Karpinski, and V.B. Shah. Julia. `https://julialang.org`, Last time visited: 5 May 2019.

Application of a System-Level Synthesis Tool in Industrial Process Control Design

Péter Arató, Dezső Nagy, György Rácz

Department of Control Engineering and Information Technology, Budapest University of Technology and Economics, Magyar tudósok krt. 2, 1117 Budapest, Hungary; arato@iit.bme.hu; nagy.dezso@iit.bme.hu; gyuriracz@iit.bme.hu

Abstract: Complex digital systems usually demand some kind of a multiprocessing architecture. The requirements to be fulfilled (energy and communication efficiency, speed, pipelining, parallelism, the number of component processors, cost, etc.) and their definable priority order may cause conflicts. Therefore, the best choice of the component processors (beside general-purpose CPUs also DSPs, GPUs, FPGAs and other custom hardware) is very important. Such resulting architectures are called heterogeneous multiprocessing architectures (HMA). The system-level synthesis (SLS) methodology can be applied beneficially in designing the HMAs. In this way, the design procedure can get rid of the most intuitive trial and error steps including also the partly reusing of existing structures. Therefore, the SLS methods help to optimize HMAs by reducing the intuitive steps. A high degree of similarity can be observed between HMAs and modern distributed industrial process control systems (DCS). This paper illustrates the procedure of adapting and applying an SLS tool in redesigning of an existing DCS as a benchmark for analyzing, evaluating and comparing the results. Through this adaptation, all such SLS functions become executable on the traditional and standardized documentation form of a DCS.

Keywords: system-level synthesis; high-level synthesis; heterogeneous multiprocessing system; industrial process control

1   Introduction

Multiprocessing can be considered the most characteristic common property of complex digital systems. Due to the more and more complex tasks to be solved for fulfilling often conflicting requirements (cost, speed, energy and communication efficiency, pipelining, parallelism, the number of component processors, etc.), the so called heterogeneous multiprocessing architectures (HMA) have become unavoidable. The component processors of such systems may be not only general purpose CPUs or cores, but also, DSPs, GPUs, FPGAs and other custom hardware. A subtask must be defined for each component processor depending on the requirements and their desired priority order [1]. Prioritizing in fulfilling the

requirements becomes critical at the highest abstraction level in the design process. [2] Thus, it is important to predict the consequences of such decisions already on the highest abstraction level before executing the rest of the design process.

The cost and performance of the whole system is strongly influenced by the definition of the subtasks, i.e. the decomposition of the task. Systematic algorithms are very helpful to the designer in comparing and evaluating the effects of different decompositions into subtasks in order to approach the optimal decisions already in the system-level synthesis phase. Existing solutions are often extended and reused intuitively in HMA design in order to shorten design time even though this usually does not guarantee advantageous results. Such evaluations of intuitive solutions generally cannot deliver unambiguous directions for the necessary changes in the architecture and trial-and-error experiments could not be avoided. This practice usually results in unnecessarily expensive and redundant system architectures.

In contrast, the system-level synthesis methods may be able to support the designer in finding, optimizing and evaluating the proper HMAs. Meanwhile, the intuitive steps in the synthesis procedure can be eliminated in a great extent. By variously allocating subtasks to different component processors, the system-level synthesis methods of HMAs may result in several different but acceptable solutions. Thus, efficiency checking, evaluating and comparing these different solutions are also supported by SLS methods already on the system-level abstraction.

Industrial Distributed Control Systems can be considered as a special case of HMA, where the component processors and the communication buses might be limited to certain types. Therefore, the existing SLS methods can also be adapted to help in the design of such systems as well. The aim of this paper is to propose such an adaptation.

2   Related System-Level Synthesis Tools

Most commercial SLS tools [3] can only be considered as high-level synthesis (HLS) tools, because they are only capable to convert a high level language (usually C) description into a hardware description and/or machine codes for several predefined architectures, usually for FPGAs or FPGA and CPU based SoC (System on Chip) platforms.

The recognized commercial tools are, for example, the Mentor Graphics Catapult HLS [4] and the Xilinx Vivado Suite [5]. There are other free and open source tools, in contrast to the commercial ones. These are mostly created for academic and research purposes with applicable documentations. Some of them have

capabilities the same as or even better than the commercial ones. LegUp [6] for example is one of the most well-known free tools. Besides being an HLS tool, it is capable to synthesize heterogeneous architectures as well. However, it can only do so by using predefined templates and by considering the communication time only between components based on those templates. Most commercial HLS tools support only a restricted subset of the given high level language. However, LegUp supports all ANSI C syntax elements including pointers, structures and global variables, the only exceptions being recursion and dynamic memory allocation. The final result of LegUp is a synthesizable Verilog code for several Altera FPGAs and one specific Altera SoC module.

The SLS methods may apply several different HLS tools and algorithms [3, 7] supporting also the design of pipeline systems. Such tools usually start from a task description formalized by a dataflow-like graph or by a high level programming language [3] [4]. These algorithms can also be utilized after suitable modifications for HMA design and in case of hardware-software co-design [8]. The latter problem can also be considered as a special case of decomposition [8].

Table 1 summarizes some properties of several SLS tools.

Table 1
Overview of several SLS tools

Tool	Input format	Applies preliminary decomposition	Exchangeable algorithms	Considers communication time between components	Priority order of requirements is variable?
PIPE [7]	Dataflow graph	NO	**YES** (multiple schedulers selectable)	**YES**	NO
XILINX Vivado suite [5]	C, C++, SystemC	NO	NO	NO, communication is only a calculated parameter	**YES, either the latency or the restart time can be prioritized.**
SYLVA [9]	Dataflow graph	**YES**	NO	NO, communication is only a calculated parameter	NO
LEGUP [6]	C, C++	NO	NO	NO	NO
Mentor Graphics Catapult HLS [4]	C, C++, SystemC	NO	NO	NO, communication is only a calculated parameter	**YES, area, restart time, power consumption can be prioritized.**
Microsemi Synphony Model Compiler [10]	MATLAB	NO	NO	NO	NO
DECHLS [2]	C, Dataflow graph	**YES**	**YES**	**YES**	**YES, restart time, bus communication time and cost can be prioritized.**

2.1   Motivation of Choosing the SLS Tool DECHLS

Based on Table 1 the DECHLS can be considered as the most suitable tool for adaptation because it accepts dataflow graphs as its input, already includes a preliminary decomposition phase, its algorithms are accessible as well as variable and the priority order of the various requirements in its cost function is also variable. In the field of industrial control, the communication time between components can be similar of even longer than the execution time of some operations. Therefore, it is also required to consider the communication time which is not considered in the case of e.g. [6]. Modification is usually not allowed for commercial products such as [4, 5, 10]. Availability and potential modifiability of the DECHLS are the most important aspects to choose it. Also a motivation for choosing DECHLS is that it has been developed at the department of the authors.

The chosen SLS tool provides the results in form of XML files or dataflow graphs that suit the highest abstraction level. At the same time, the chosen application field (industrial process control) demands several special requirements also in the lower abstraction levels. Therefore, special adaptation and modification procedures are required in order to utilize the given results. Besides, the application field has its special widespread traditional, even standardized description and design methodology. In industrial process control, the component processors may be placed in large distances from each other. Therefore, the communication time between them can be significant and must not be neglected at adapting the SLS tool.

3   Adapting the DECHLS Tool

3.1   Industrial Process Control Systems as Special Cases of HMA

Modern industrial Distributed Process Control Systems (DCS) usually consist of many different intelligent modules. The modules of a DCS are usually connected by hierarchical and standardized bus systems. A DCS system performs a well-defined set of subfunctions distributed between several special programmable modules (often called programmable logic controllers, PLCs). Multiprocessing is a characteristic property of these systems, because the PLCs perform their functions concurrently with each other's and repeatedly at prescribed cycle times.

A noticeable similarity exists between the DCSs and general HMA systems as shown in Fig. 1.
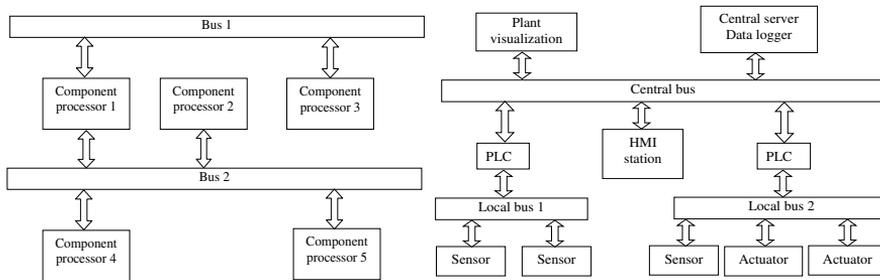
Figure 1
The general architecture of a HMA system (left) and that of a typical DCS (right)

The most problematic phase of a DCS's design process is the task decomposition. An additional special difficulty may arise, if the number of component processors is also prescribed. However, the DECHLS tool can be utilized in an adapted form to automate and optimize the design process of the DCS. The application strategy of the DECHLS-DCS adaptation is shown in Fig. 2.
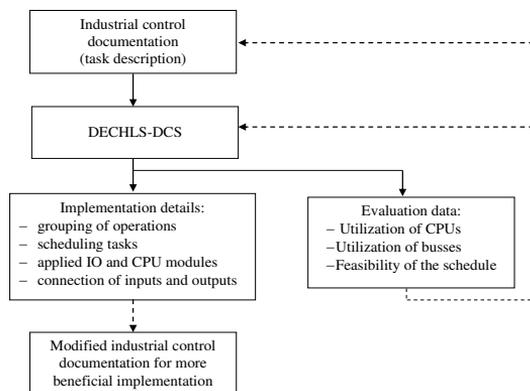
Figure 2
Strategy of adapting DECHLS for DCS design

3.2   Special Requirements of DCS

In an industrial process control project, many designer specialists usually must work together (e.g. architects, mechanical technologists, power engineers and control engineers) [11]. The complete documentation of the project must be understandable and unambiguous to all participants. Therefore, a traditional documentation form is already in use and it is standardized as IEC 61131 [11].

A new SLS approach to the DCS design must be adapted to this documentation form. This task description standard basically contains all inputs and outputs of the designed system and a formal description of the control algorithms to be

implemented. These control algorithms can also be designed by various graph-based methods. For example, [12] is based on a CP-graph model. Usually the control algorithm itself is already available as a dataflow-like graph. According to the [11] standard, the control algorithms are implemented by visual programming languages such as the Sequential Function Chart (SFC) or the Functional Block Diagram (FBD) [13]. Neither of these programming languages can be directly applied in the existing SLS tools. In order to adapt the DECHLS, a new dataflow graph model has been developed that is able to contain all information required by the standard description and it is also usable in DECHLS.

This new Functional Dataflow Graph (FDFG) can be summarized as follows:

$$FDFG = \{V, E, F\} \tag{1}$$

where V is the set of nodes, E is the set of directed edges and F is the set of subfunctions.

Every node in V represents an operation in the task description and a tuple of natural numbers is assigned to each node as:

$$\forall v_i \in V, v_i \rightarrow \{t_i, m_i\} \tag{2}$$

where t_i means the execution time of v_i and m_i is the required redundancy of v_i.

Every directed edge of the graph must have a source and a destination node and also must have a natural number assigned to it:

$$\forall e_i \in E, e_i = \{v_a, v_b\}, v_a, v_b \in V, e_i \rightarrow c_i, c_i \in N \tag{3}$$

where c_i is the number of data bits used in the communication between the two nodes belonging to the edge. The value of c_i can also be 0, in this case there is no data communication, only timing dependency between v_s and v_d.

Set F consists of f_i sets, each of them representing a subfunction. These f_is are disjoint sets of nodes:

$$F = \{f_1, f_2, ..., f_n\} \tag{4}$$

where

$$f_i = \{..., v_j, ...\}, \forall v_j \in V, f_i \in F \tag{5}$$

and

$$\forall f_i \in F, f_i \rightarrow (T_i, P_i) \tag{6}$$

where T_i is a natural number that means the maximum possible execution time limit of f_i. If this subfunction does not have a defined maximum execution time, then T_i should be 0. The P_i is a Boolean value; it is true if all parts of this subfunction must be allocated to the same processor.

3.3   Transforming the Task Description

As it is mentioned earlier, the FBD and SFC are not directly suitable as FDFG for DECHLS because of two problems. The first problem is with the order of the variable reads and writes. The second problem is that these graphs may have invalid execution orders or contain loops. These problems, however, can be solved by two simple algorithms as follows.

3.3.1   Handling Variable Reads and Writes

The PLC programs may contain variable reads and writes as elementary operations. These operations are used to access either actual data in temporary memory, or the physical inputs and outputs of the controller hardware. These read and write operations must be included into the resulting FDFG as part of the transformation process. The general rule of PLC software execution is that the program runs in a cycle. This must begin by reading all the input variables, then following by executing all the operations of the program. Finally, the modified variables should be written back. According to the IEC 61131-3 standard [11], variables should not be modified during a program unit's execution. The reason of this is to prevent execution hazards. It also means that variable reads and writes must be handled separately. In this way, operations writing and then reading the same variable cannot be connected in the dataflow graph representation. This means that read operations never have inputs and they must be source nodes in the FDFG. Likewise, writes never have outputs and they must be destination nodes in the FDFG.

If more than one operation within a subfunction writes to the same variable, then only the latest write will be valid. Since the order of the writes is known already at the transformation stage, the invalid writes must be removed from the resulting FDFG. Therefore, the occurrence of such invalid writes can be considered as a possible error in the task description, and the transformation algorithm should warn the designer about this fact. Most commercial PLC development environments also perform this check and issue a warning. However, in case of alternative data paths and conditional execution of operations, the order of writes is determined by input data at runtime and cannot be determined during the transformation step. In these cases, the transformation algorithm should preserve all write operations.

3.3.2   Handling Prescribed Execution Sequence

An FDFG can have multiple valid and unambiguous execution orders. The rule to create an unambiguous execution order is that operations can only be executed if all their input data are present. In other words, after all the nodes having a directed path to this node are already executed. This rule allows for many different execution sequences considered valid ones. The output data yielded by any valid

execution sequence are the same. Therefore, both the designer, both the scheduling algorithm can freely choose each such valid sequence. It is easy to prove that an FDFG containing a cycle must not have any valid execution sequence.

An FBD program always prescribes the execution order of function blocks (i.e. the elementary operations). After transforming the FBD into an FDFG, at least one of the valid execution orders in the FDFG must be the same as the prescribed execution order of the FBD.

If this prescribed execution sequence is also valid in the FDFG, then the graph needs not be modified further. However, if the prescribed execution sequence is not valid, then the FDFG must be modified into a form that makes the prescribed execution sequence also valid.

The following simple algorithm can be used to perform the aforementioned modification on the FDFG, as illustrated in Fig. 3:

1) Find an edge $e_i = (v_s \rightarrow v_d)$, where v_d precedes v_s according to the execution order.

2) Create a new temporary variable.

3) Create a write operation for the new variable and create an edge leading from v_s to this operation.

4) Create a read operation for the new variable and create an edge from v_d to this operation.

5) Delete edge e_i

6) Continue with step 1 until all the problematic edges have been tested.
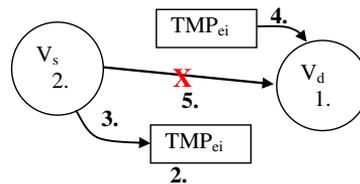
Figure 3
Illustration of the steps needed to enforce the prescribed execution order

This algorithm can also be applied to solve the problem of cycles in the FBD. It is easy to prove that cycles in the FDFG will always contain at least one edge that will be found by the aforementioned algorithm. By performing the algorithm, this edge will be eliminated and the cycle will cease to exist. The effect of execution order on the resulting FDFG is illustrated in Fig. 4, where the FBD on the left and the FBD on the right differs only in the execution sequence order of the TO_IN and the TO_RE function blocks. The FDFG resulted from this valid execution order is seen on the left.

It can be observed that the execution sequence on the right means that the FDFG has two disjoint parts connected only by variables. Therefore, the TO_RE operation uses the previous output of the TO_IN operation that was saved from the previous cycle. It must be noted that this transformation also introduces one cycle of intentional signal delay between the affected blocks.

Fig. 5 illustrates the extended DECHLS flowchart adapted to DCS design.
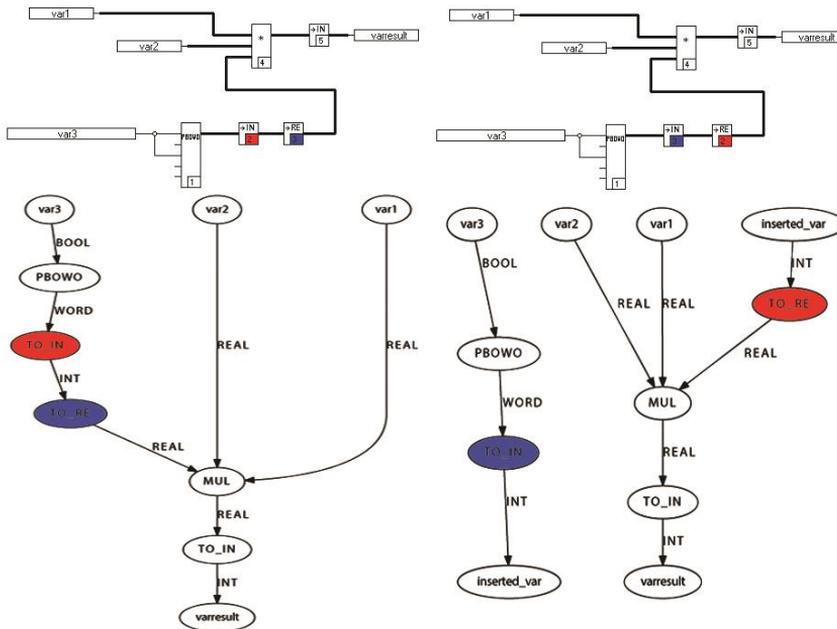
Figure 4
Effect of execution order of the FBD on the FDFG

3.4   The Adaptation Algorithm

The proposed adaptation algorithm consists of the following phases:

- Transformation phase, that produces the FDFG from the standard task description formalism of the DCS (given in SFC or FBD languages). [14]

- Preliminary Functional Decomposition phase (PFD) prevents separating the user-defined logically coherent functions, in contrast to usual decomposition [15] algorithms.

- Multirate Function Scheduler phase (MFS) can enforce different cycle times (latency times) for some subfunctions in the FDFG [16].

- Extended Allocation phase (EA) that is also capable to handle additional replication constraints arising after the decomposition. The aim of this step is to ensure safety-critical redundancy and availability, if any.
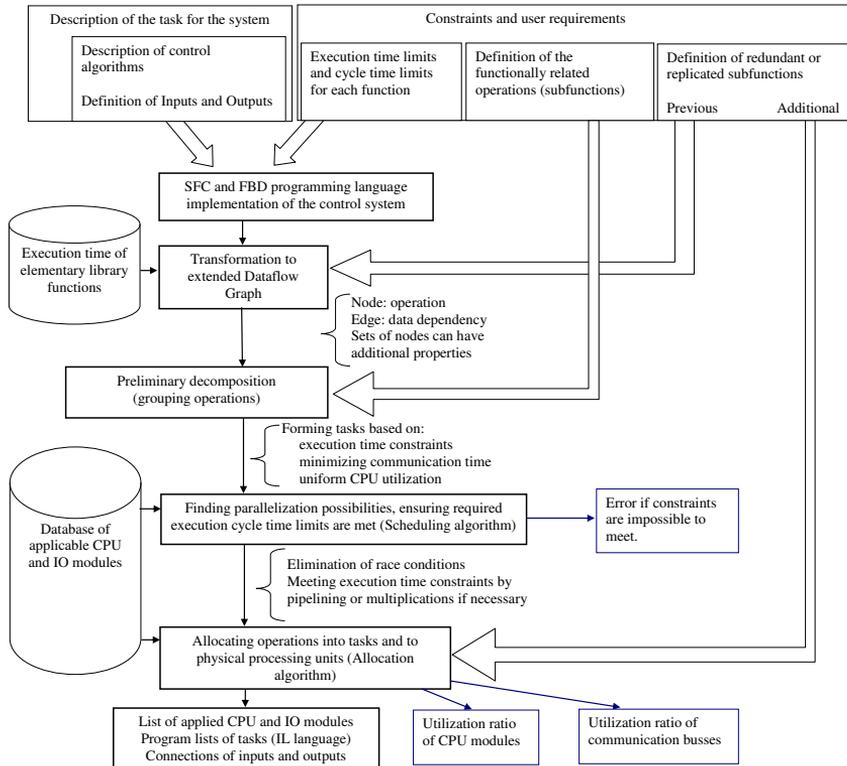
Figure 5

Flowchart of the extended DECHLS tool adapted to DCS design

3.4.1   The Decomposition Phase

The main goal of the original preliminary decomposition in DECHLS is to reduce the number of elementary operations for the scheduler and allocation phases. In this way, the performance of those may be improved [2]. An additional benefit of the primary decomposition is that some special requirements can be taken into account with a higher priority. In case of DCS, the main goal of the decomposition is to distribute the workload uniformly along with also minimizing communication between segments. Segments will be treated as atomic in the scheduler. This means that operations assigned to the same segment will not be overlapping in time and cannot be allocated between multiple component

processors. Therefore, the decomposition can easily ensure that elementary operations of those subfunctions that must be allocated together due to the serviceability constraint (P_i = true) should never overlap in time. Each of such elementary operation will form a separate segment, independently of the other segments.

There can be more decomposition algorithms used in DECHLS. The spectral clustering based algorithm has proved to be the most suitable for industrial control system tasks because it is capable to a simpler parameterization [14, 2].

3.4.2   The Scheduling Phase

The scheduler algorithm used in DECHLS is a modified force directed one [7]. The scheduler can be simplified because pipeline execution in DCS systems is usually not allowed.

In this case, the scheduler only needs to determine the starting time of the segments created already in the decomposition step. This starting time of most segments can range up to the maximum cycle time (the latency time, L), which is an input parameter to the scheduler. A minimum latency (L_{min}) can also be determined based on the shortest execution path in the FDFG. The scheduler cannot find a valid solution when $L < L_{min}$, therefore it must stop in this case. Such minimum latencies also exist for each subfunction (subgraph) of the FDFG.

Some functions (segments) must be completed faster than L, because their prescribed cycle time is $T_i < L$. The force directed algorithm should be modified to take into account the T_i of each subfunction by implementing a multi-rate scheduler as in [16].

The multi-rate scheduler has different queues for different operation groups as seen in Fig. 6. Operations belonging to subfunctions without a defined execution time limit will be scheduled in the general queue. Subfunctions, having longer or equal execution time limits than the given latency, are also scheduled in this general queue because their execution time limits will be trivially met. Subfunctions having a smaller time limit than L must be handled in different queues. The queue of function f_i with time limit T_i will have a multi-rate latency of L_i, where

$$L_i = \frac{L}{k_i} \qquad\qquad k_i = \left\lceil \frac{L}{T_i} \right\rceil$$

(7)

All the separate queues will have the same force function, and operations potentially overlapping in any queues will increase the force. Because of this, the scheduler will attempt to prohibit overlapping operations as long as possible in order to reduce the required number of processors. There is one more important task of the scheduler. If any subfunctions' minimum latency ($L_{min,i}$) defined by its

FDFG is smaller than its specified T_i, then the scheduler must report an error and stop. In this case, the execution time constraints are too strict and cannot be met by this algorithm.
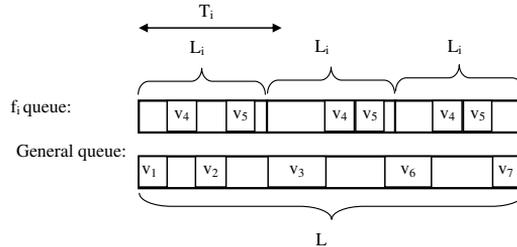
Figure 6
Multi-rate scheduler

3.4.3   The Allocation Phase

The allocation phase is the simplest among the modifications. It should be able to allocate m_i copies of every operation. These multiple copies must be allocated into different processors, irrespective of their timings. This can be done even before the actual allocation phase as a previous step, or by slightly modifying an existing allocation algorithm. In case of the graph-coloring based algorithm in DECHLS, this modification can be done in the following way.

The existing allocation algorithm builds a conflict graph, based on the scheduled dataflow graph. The nodes of this conflict graph represent the nodes of the scheduled dataflow graph, but the edges in the conflict graph represent the time overlapping between nodes. If two operations are overlapping in time, they cannot be executed by the same component processor. In this sense, the allocation basically means the coloring of the conflict graph, by the least amount of colors. Thus, any two adjacent nodes are colored differently.

In case of DCS, the redundancy criteria may require the replication of certain nodes. Those nodes are already replicated before scheduling and they will appear as two separate nodes in the scheduled dataflow graph. Since multiple copies of the same logical node are scheduled independently, it may happen that they are not overlapping. However, they are not allowed to be allocated into the same processing unit because their multiplication would not result a real redundancy.

Such a scenario is seen in Fig. 7, where nodes 5, 6 and 7 are replicated (5', 6' and 7'). The scheduler placed some of the replicated nodes into different time slots for minimizing the resource usage. In this way, only 3 processors will be required.
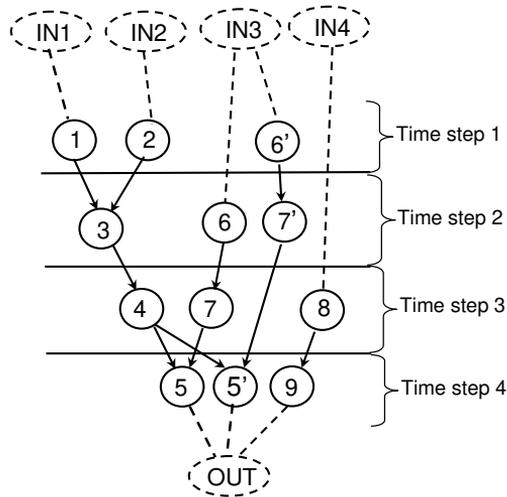
Figure 7
An example for a scheduled dataflow graph with replicated nodes

In order to solve the problem, additional edges are needed in the conflict graph between redundant copies of the same node. In Fig. 8 these edges are marked as double lines for easy identification. Otherwise in the remaining steps of the graph coloring algorithm, these additional edges are treated the same as regular edges.

This way, the redundancy can be safely handled, by the allocation algorithm.
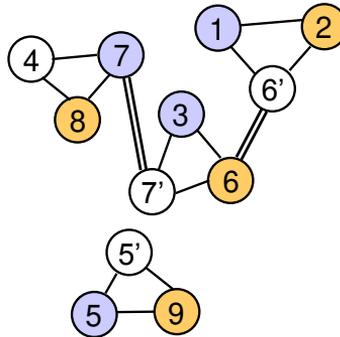
Figure 8
A conflict graph with the additional edges and a possible coloring

4   Benchmark Results

The chosen benchmark application is an existing DCS that was programmed in the FBD language by using the ABB Freelance development environment [17]. The benchmark consists of 29 program blocks. Each of the program blocks are logically coherent functions. Some blocks are scheduled to operate with 1000 ms, while others with 500 ms cycle times. The program blocks are distributed between two separate ABB AC800F3 PLCs. There are two special functions that redundantly implement the same safety critical function (burner control). These two tasks must never be allocated to the same PLC for safety reasons. This is the main reason why two PLCs were needed in this implementation.

The first step is the transformation of the existing FBD task description into FDFG form. Only an essential part of the original description (one subfunction) and the resulting FDFG is illustrated in Fig. 9 and Fig. 10.

The FDFG was then given to the preliminary decomposition phase and the algorithm divided it into 97 subfunctions instead of the original 29. For example, the #27 subfunction was divided into 6 parts as illustrated in Fig. 11. Increasing the number of subfunctions provides more freedom in the later phases, namely the scheduling and the allocation.

The purpose of Figs. 9, 10 and 11 is only to illustrate the structures of the original and the transformed graphs, the text fields in the blocks are not important in this sense.

After the decomposition, the scheduler and allocation phases of the experimental DECHLS-DCS tool have determined the required number of PLCs. The average utilization of PLCs as well as the average communication bus utilization at several different cycle times were also computed and these are shown in Fig. 12.

The cycle times indicated in the Fig. 12 are the longest ones in each case, some functions have shorter cycle times. The reason of this is the fixed 500 ms cycle time of the critical functions in the original implementation. For the rest of the functions, there are no cycle time prescriptions. If the actual cycle time is shorter than the prescribed one, the whole system runs at the actual cycle time. If the actual cycle time is longer, then the critical functions must still run at least as fast as the prescribed cycle time. The diagram has markers at specific cycle time values to show where practical solutions are obtained. Between 400 ms and 450 ms there are many markers, because the tool was restarted many times in this interval in order to find the shortest possible cycle time allowing the task to be solved by only 2 processors.

Figure 9
FBD form of subfunction #27 in the benchmark

Figure 10
FDFG form of subfunction #27 in the benchmark
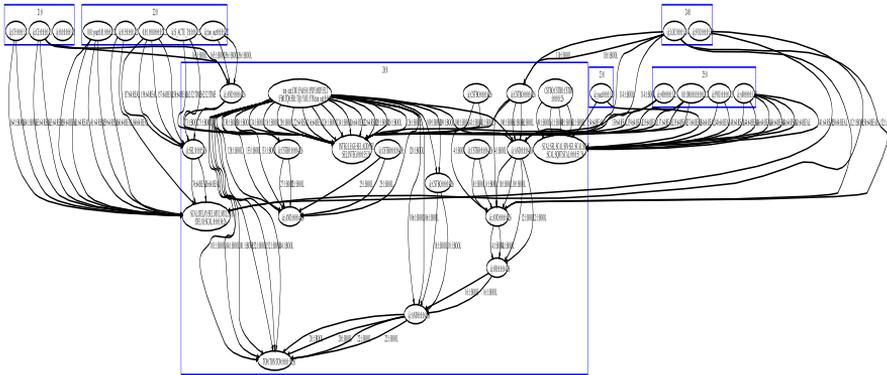
Figure 11
The 6 new subfunctions resulted from subfunction #27
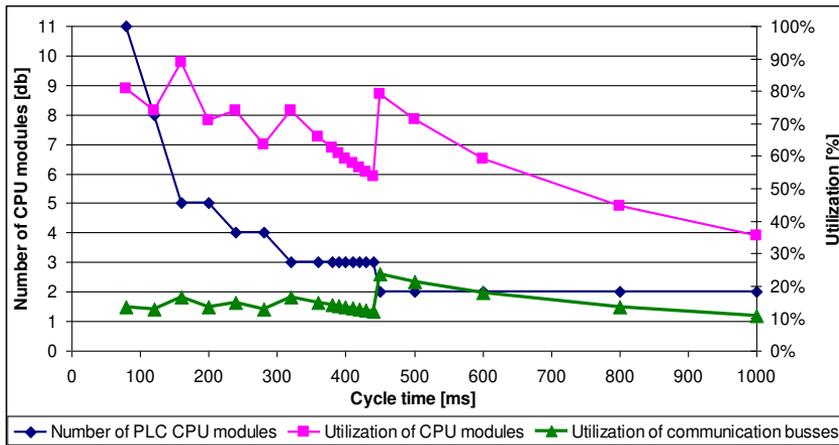
Figure 12
The resulted number of PLCs and their utilization levels at different cycle times

Fig. 12 shows in the left axis that only 2 PLCs are required at 450 ms cycle time (blue line). In this case, the average utilization (purple line) is only about 80%, as seen on the right axis. All functions are still able to run at this speed. Less than 2 PLCs are never enough to solve the task, because of the redundancy constraint. It can be observed that prescribing 1000 ms cycle time was not necessary in the existing implementation.

The DECHLS-DCS also shows that the average utilization of CPUs and communication busses in the existing system were only around 37% and 11%. The official ABB Freelance development tool cannot deliver such estimations without completely building the system, followed by downloading, running and profiling the whole software.

Summarizing above simulation results:

- A 450 ms cycle time could be achieved safely by utilizing the same number of CPUs

- 79% CPU utilization was achieved instead of 37%

- 22% communication bus utilization was achieved instead of 11%

Conclusions

This paper has illustrated how an SLS tool (DECHLS) can be adapted, modified and utilized in designing a specific form of HMA (an industrial process control system). For this proper adaptation, DECHLS have been modified in order to be capable to handle the application-specific standardized task input graph descriptions (SFC or FBD). A converting algorithm has been presented for this extension. The additional necessary extensions for adaptation have been also presented in the paper: a special decomposition algorithm, a multirate function scheduler algorithm, an extended allocation algorithm handling safety-critical redundancy. By these modifications and extensions, DECHLS became a capable tool for providing various resulting designs to compare and evaluate them already on the SLS level without any lower level implementations.

The benchmark presented in the paper illustrates on comparing with existing solutions that the preliminary decomposition in DECHLS increased the number of subfunctions. Hereby, more freedom remains for the scheduling and allocation phases. It can also be observed that a proper systematic scheduling could lead to higher processor utilization even at high communication times between processors.

Consequently, such an adaptation of the DECHLS tool, helps to compare and evaluate various resulting HMAs exclusively on the SLS level, without implementing the whole system.

References

[1]   Gy. Rácz, P.Arató, „A System-Level Synthesis Approach to Industrial Process Control Design", 23rd IEEE International Conference on Intelligent Engineering Systems, Gödöllő, Hungary, 2019

[2]   Rácz, Gy.; Arató, P. "A decomposition-based system level synthesis method for heterogeneous multiprocessor architectures," 2017 30th IEEE International System-on-Chip Conference (SOCC), Munich, Germany, 2017, pp. 381-386

[3]   W. Meeus, K. Van Beeck, T. Goedemé, J. Meel, D. Stroobandt, "An overview of today's high-level synthesis tools", Design Automation for Embedded Systems, 16(4), 31-51, 2012

[4]   Mentor   Graphics   Catapult   C   software   manual: https://www.mentor.com/hls-lp/catapult-high-level-synthesis/

[5]   Xilinx   Vivado   design   Suite   Manual:
        https://www.xilinx.com/products/design-tools/vivado.html

[6]   A. Canis, J. Choi, B. Fort, B. Syrowik, R. L. Lian, Y. T. Chen, H. Hsiao, J.
        Goeders, S. Brown, J. H. Anderson, "LegUp High-Level Synthesis,"
        chapter in FPGAs for Software Engineers, Springer, 2016

[7]   P. Arató, T. Visegrády, I. Jankovits, "High Level Synthesis of Pipelined
        Datapaths", John Wiley & Sons, New York, ISBN: 0 471495582 4, 2001

[8]   N. Govil, S. R. Chowdhury, "GMA: a high speed metaheuristic algorithmic
        approach to hardware software partitioning for Low-cost SoCs", 2015
        International Symposium on Rapid System Prototyping (RSP), Amsterdam,
        2015

[9]   L. Shuo, "System-Level Architectural Hardware Synthesis for Digital
        Signal Processing Sub-Systems." PhD thesis, Stockholm, 2015

[10]   Microsemi   Synphony   Model   Compiler   User   Guide:
        https://www.microsemi.com/product-directory/dev-tools/4899-
        synphony#documents

[11]   IEC, IEC 61131 en:2003, Programmable Controllers

[12]   K. M. Hangos, F. Friedler, J. B. Varga, L. T. Fan, "A graph-theoretic
        approach to integrated process and control system synthesis", IFAC
        Proceedings Volumes, 27(7), 1994, 61-66

[13]   IEC, IEC 61131-3 en:2003, Programmable Controllers - Part 3:
        Programming languages, 2018

[14]   Gy. Rácz, P. Arató, "Adapting the system level synthesis methodology to
        industrial control design", Proceedings of the Workshop on the Advances
        of Information Technology: WAIT 2018, Budapest, 2018, 131-136

[15]   P. Arató, D. Drexler, Gy. Rácz, "Analyzing the Effect of Decomposition
        Algorithms on the Heterogeneous Multiprocessing Architectures in System
        Level Synthesis" Scientific Buletin of Politechnica University of Timisoara
        Transactions on Automatic Control and Computer Science, 60(74) 39-46,
        2015

[16]   X. Y. Zhu, M. Geilen, T. Basten and S. Stuijk, "Static Rate-Optimal
        Scheduling of Multirate DSP Algorithms via Retiming and Unfolding,"
        2012 IEEE 18th Real Time and Embedded Technology and Applications
        Symposium, Beijing, 2012, 109-118

[17]   ABB   Freelance   Engineering,   https://new.abb.com/control-
        systems/essential-automation/freelance/engineering

Multi-Thread Implementation of Tool Tip Tracking for Laparoscopic Surgical Box-Trainer Intelligent Performance Assessment System

Janos L. Grantner, Aous H. Kurdi, Mohammed Al-Gailani and Ikhlas Abdel-Qader

Department of Electrical and Computer Engineering, Western Michigan University, 1903 West Michigan Ave, Kalamazoo, MI 49008, USA

janos.grantner@wmich.edu; aoushammad.kurdi@wmich.edu; mohammedyasser.algailani@wmich.edu; ikhlas.abdelqader@wmich.edu

Robert G. Sawyer, M.D. and Saad Shebrain M.D.

Department of Surgery, Homer Stryker M.D. School of Medicine, Western Michigan University, 300 Portage Street, Kalamazoo, MI 49007, USA

robert.sawyer@med.wmich.edu; saad.shebrain@med.wmich.edu

Abstract: Laparoscopic surgery has revolutionized medicine, yet it requires specific skills not traditionally taught to surgeons. Early in training, surgical trainees frequently use low-fidelity "box-trainers" to increase their skills. Evaluation of performance, however, is crude, frequently focusing on speed alone or subjective observations. There is a research collaboration between the Department of Electrical and Computer Engineering and the Department of Surgery, of the Homer Stryker M.D. Medical School, at WMU. The objective of this research is to develop an intelligent box-trainer system which incorporates sensory devices and high definition digital video cameras along with a fuzzy logic-based performance assessment system. Two of the key research problems of this work are the implementation of the tool tip tracking task and the associated performance assessment method. This system will more consistently evaluate trainees and their performance in real time without needing the presence of an independent observer. Using fuzzy logic to capture expert knowledge and fusing it with sensory data for performance assessment purposes is a new approach in the area of laparoscopic surgery training.

Keywords: Intelligent Laparoscopic Surgical Box-Trainer; Laparoscopic Surgical Tool Tip Tracking; Fuzzy Logic-Based Performance Assessment System

1   Introduction

Laparoscopic surgery, one form of minimally invasive surgery, is performed through a series of small incisions, through which, plastic ports are placed, allowing the introduction of a camera and specially designed instruments. The avoidance of a single large incision, reduces pain and recovery time. The surgeon's interface occurs through a video monitor that displays images from the camera inside the abdomen. Early training for laparoscopic surgery is more complex than for more traditional (open) surgery, leading to a more specified training curriculum known as Fundamentals of Laparoscopic Surgery (FLS). The FLS program (SAGES/ACS, FLS Program, Los Angeles, CA, USA) [1] [2] consists of five basic laparoscopic tasks that are non-procedure specific and are performed and tested on a surgical box-trainer, a low cost *in vitro* system that uses laparoscopic instruments in a laboratory, bench-top setting [3]. These tasks emphasize ambidexterity, depth perception, eye-hand coordination and controlled instrument movement. The development of excellent eye-hand coordination and cognitive capabilities are crucial for the safe execution of laparoscopic procedures.

When measuring surgical skills during simulation, the users' performance is, generally speaking, assessed by monitoring the precision and the speed of the execution of the tasks. One major disadvantage when using the universally accepted and standard FLS Box-Trainer device [4] [5] is that, outside of crude time to task completion, neither qualitative nor quantitative performance data are measured and recorded electronically. As a consequence, the test results are not easily available for the users in a format which can be reviewed by computer or teaching faculty, who, instead, must depend on subjective assessments performed in real time.

Several attempts have been made to objectively measure and record performance during laparoscopic training exercises (reviewed by Reily et al. [6]). Oropesa et al. [7] proposed an EVA (Endoscopic Video Analysis) tracking system for extracting the motion of laparoscopic instruments based on non-obtrusive video tracking. The feasibility of using EVA in laparoscopic settings has been tested in a box trainer setup. EVA makes use of an algorithm that employs information of the laparoscopic instrument's shaft edges in the image, the instrument's insertion point and the camera's optical center to track the three-dimensional position of the instrument tip. A validation study of EVA comprised a comparison of the measurements achieved with EVA and the TrEndo tracking system. EVA was successfully validated in a BT setup showing the potential of endoscopic video analysis to assess laparoscopic psychomotor skills.

Horeman et al. [8] investigated the added value of force parameters with respect to commonly used motion and time parameters such as path length, motion volume, and task time. Two new dynamic bimanual positioning tasks were developed that not only require adequate motion control but also appropriate force control during simulated tissue manipulation. Their study concluded that it is possible to

distinguish the skill levels of a novice versus an expert with an accuracy up to 100%. The results indicated that the manipulation forces applied by novices often exceeded the established threshold value set for producing tissue damage. This means that a student's focus on task time and instrument motion alone during a skills training exercise may have a negative influence on tissue handling skills. The relatively high forces used by the intermediates in combination with the apparent lack of correlation between force and motion parameters argues for the inclusion of specific training in and assessment of force application in tissue handling in laparoscopic skills training programs.

Although progress has been made in terms of objectively quantifying laparoscopic surgical proficiency, a fair amount of assessment still depends on subjective evaluations by experienced surgeons. To this end, the introduction of fuzzy logic to skills evaluation may help lessen variability in testing. Grantner et al. [9] proposed a fuzzy logic based intelligent decision support system for children, in a similar context. Further, Grantner et al. [10] proposed an intelligent box-trainer system that is enhanced by adding several sensors and high definition digital video cameras along with a fuzzy logic-based performance assessment system. Thus, the combination of technologically advanced measurement of instrument control with fuzzy logic evaluation of performance must be considered a promising new area in the field of surgical skill evaluation.

One challenge to producing objective evaluations based on imaging is the fact that modern imaging systems can generate a massive amount of data in a short time span defined by the frame rate. In order to provide real-time intelligent assessment capabilities, the dynamic performance of the fuzzy logic-based system must be able to maintain a pace that matches that of image data generation. We, therefore, propose a hardware accelerator framework for fuzzy logic edge detection.

This paper is organized as follows: Section II describes the concepts of tool tip tracking for the FLS BT tests. Section III presents the fuzzy logic knowledge base for the tool tip tracking system. Section IV reports on the software design and its implementation. Section V outlines simulation and live test results. Section VI lays out a proposed framework for developing a hardware accelerator for edge detection. Conclusions and plans for further research are given in Section VII.

2   Tool Tip Tracking System for FLS Box-Trainer Tests

The following tasks should be carried out during the FLS tests: Peg Transfer, Precision Cutting, Endoloop, Extracorporeal Suture and Intracorporeal Suture [1]. For assessment purposes, in all five of them, it is very important that the surgical tool tips exhibit smooth movements in a limited three-dimensional space. The simplified block diagram of the tool tip tracking section and the test execution

monitoring section of the Intelligent Box-Trainer System [10] is depicted in Fig. 1. The capacitive force sensors mounted on the graspers along with their interface electronics, the START/STOP push button and the microcontroller system which sends the force measurement data and the execution time to the Main PC are omitted.
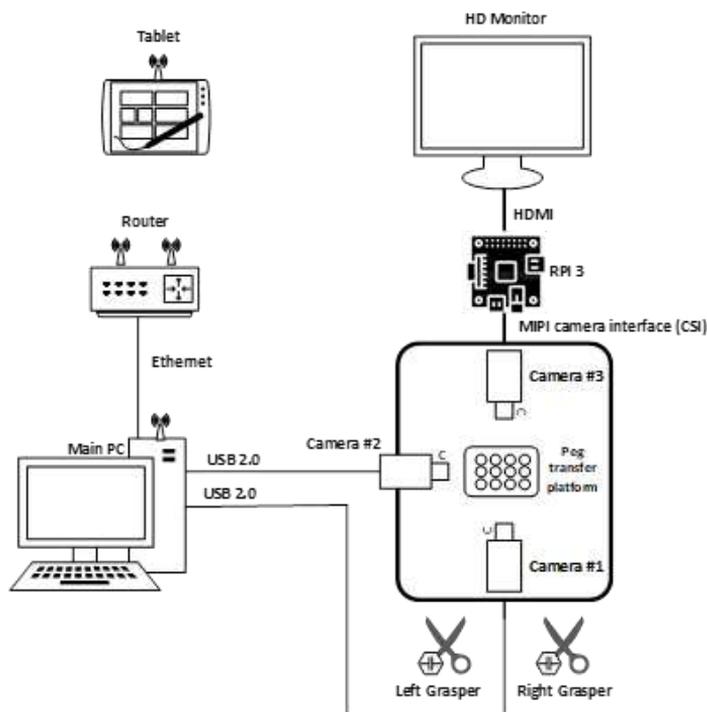
Figure 1
Tool tip tracking system

The tool tip tracking system encompasses two image sensors, a computer system and specific laparoscopic surgical tools suited for a particular FLS test. Currently, each image sensor is a 5-megapixel USB 2.0 camera with a variable-focus lens. The cameras are positioned perpendicularly inside the BT to facilitate a three-dimensional tracking of the tool tips movements. The center of each camera lens is calibrated to point to the center of the actual test platform inside the BT. The cameras are configured to produce a video stream at the rate of approximately 30 frames per seconds with resolution of 1028x720 pixels per frame. The BT is retrofitted with extra LED strips for better lighting conditions. The back side of the BT is closed with a white piece of plastic to avoid any outlier tracking of objects outside the BT. During the tests the video streams produced by Cameras 1 and 2 are recorded.

Camera 3 along with the Raspberry Pi 3 (RPi3) microcomputer provide visual feedback for the test taker on the HD monitor. This video stream is not recorded. The tablet is used by the supervising medical personnel to register the user for the test and to select the type of test to be taken as well as the skill level. During the test the tablet displays the results of the performance assessment system in real time. Those results are created by programs running on the Main PC. The wireless communications between the tablet and the Main PC are implemented through the Router. At the conclusion of the test, the supervisor may accept, or reject, the assessment outcome by the intelligent system and, in addition, may add comments by using the tablet. Those comments will be included in the text file in which the assessment results are stored.

3   Knowledge Base for Tool Tip Tracking System

The concept to track and assess the movements of the tool tips during laparoscopic surgery skill tests using a BT is new. At present time there is no other, published, comparable intelligent BT system available than that which is presented in this paper. As a consequence, no public data sets are available to work with. However, there is expert surgeon knowledge about the desirable tool tip movements. Using camera images, there are various algorithms and software tools available to track the tool tip movements. In this research the color tracking feature has been chosen because it is less computationally intensive than other approaches and the objective of this project is to provide real-time skill assessment capabilities. The tracking is carried out in pixel space but the results can be converted into distance measurements in a coordinate system.

Fuzzy set theory provides a mathematical framework, in which, expert knowledge can be fused with measured data in an uncertain environment. Even in the case of non-existent experimental data sets by interviewing domain experts a fuzzy knowledge base can be created.

The discussion which follows is based upon the Peg Transfer test. The right-hand grasper tool tip is painted red while the left-hand grasper tool tip is painted yellow for object movement tracking. To make the color tracking more robust the red color may be replaced by a more composite one, like cyan or magenta, in the future. A crucial aspect of executing a laparoscopic surgery procedure is that tool tip movements should be confined into a three-dimensional (3D) space which should be of limited size. If the tips were moving without imposed 3D constraints, they might cause unintended damage to internal organs near to the target of the procedure. The idea for creating a fuzzy model for tool trip tacking is illustrated in Fig. 2. The test platform is surrounded by three virtual 3D boxes. The physical sizes of these boxes depend upon the chosen skill level of the user and are decided by expert surgeons. The actual sizes are subject of further research and will be fine-tuned for each skill level category to reflect the opinions of expert surgeons.

The coordinate system is set up as follows: X stands for horizontal movements, Y stands for vertical movements and Z stands for depth movements (away from the test platform), respectively. The origin of the virtual coordinate system is fixed to the geometric center of the test platform.

The virtual 3D spaces are divided into symmetrical left and right sub-spaces. The input and output fuzzy sets for tool tip tracking are depicted in Figs. 3a-3c. Both the recorded left-hand and the right-hand tool tip movements are fuzzified in the same way. The membership functions for the X and Y movements have the same properties for the linguistic labels Excellent, Good and Bad, respectively. The labels for the Z movements are defined in a somewhat different fashion.
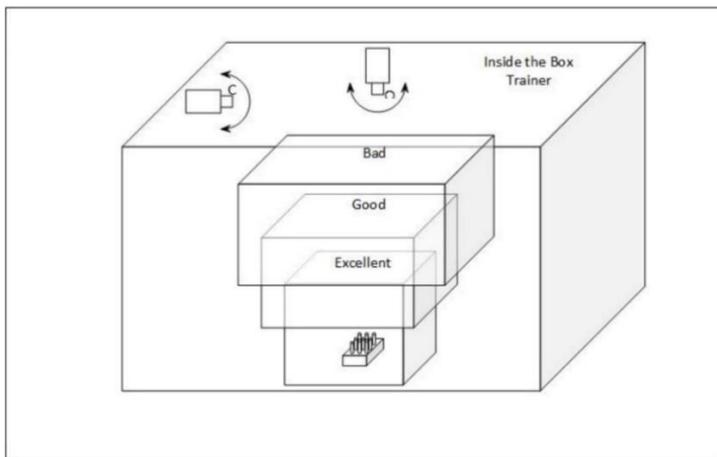
Figure 2
Virtual 3D spaces for tool trip tracking
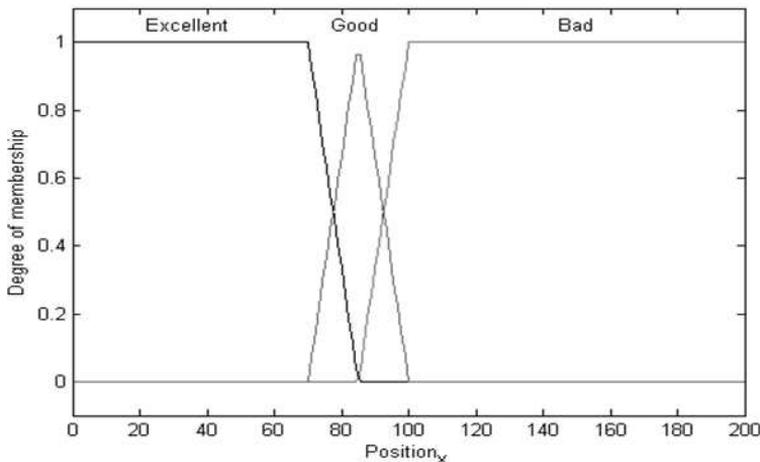
Figure 3a
Membership functions for X movements

The membership functions for performance assessment using letter grades are shown in Fig. 4. A stands for the top grade and E stands for the failing grade. The fuzzy IF-THEN rules for the tool-tip tracking performance assessment system are given in the form of Fuzzy Associate Memory (FAM) matrices which are depicted in Table 1.
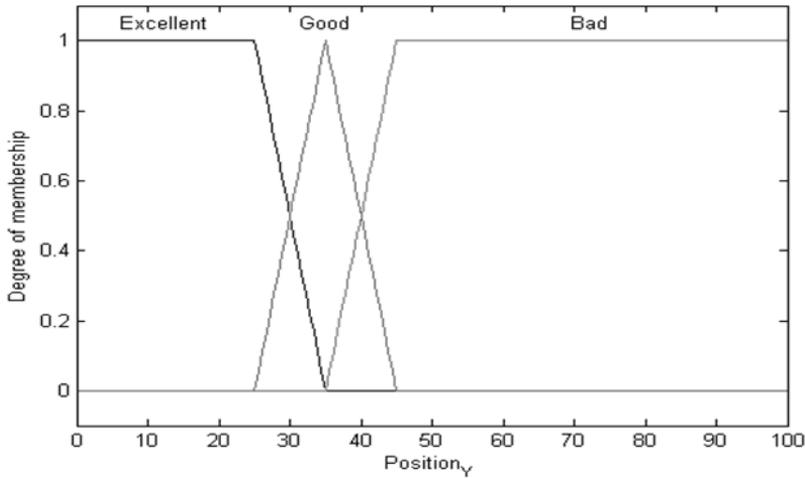
Figure 3b
Membership functions for Y movements
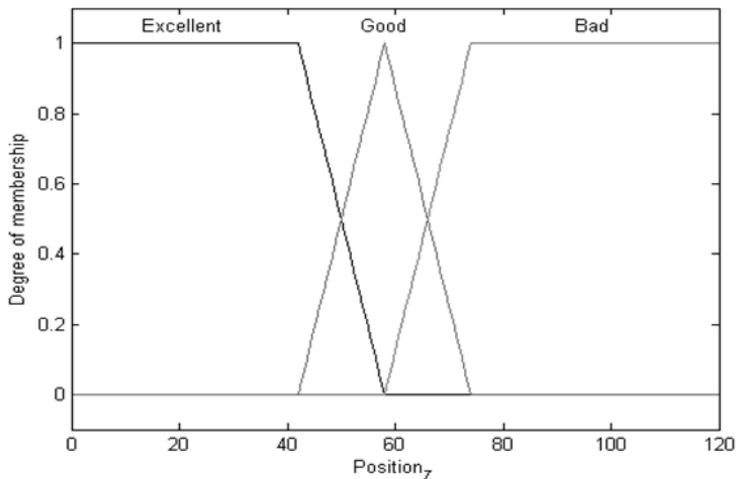
Figure 3c
Membership functions for Z movements
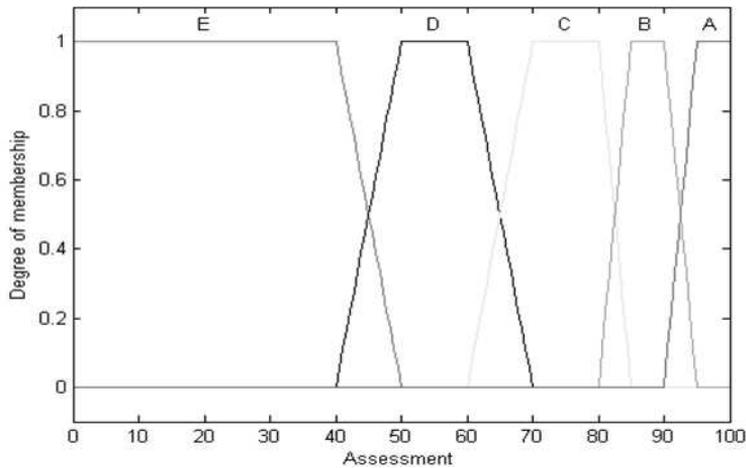
Figure 4

Membership functions for performance assessment

The membership functions and the fuzzy knowledge base have been devised by collecting data from peg transfer tests executed by first year residents and also by interviewing expert surgery faculty about their perceptions on the residents' performance. Hence, the membership functions and the FAM matrices included in this paper reflect the current status of the research project. As more test data, along with expert surgeons' performance assessment, becomes available, the membership functions and the FAM matrices will be fine-tuned to improve the quality of the assessment system. In addition, fuzzy membership functions and associated FAM matrices will be developed for users of other skill levels as well, like novices, second and third year residents and expert surgeons.

The research work will continue to extend the intelligent assessment system to support all five standard tests which are established in the FLS program.

Table 1

FAM Matrices for tool-trip tracking performance assessment

Input			Output
X	**Y**	**Z**	**Assessment**
Bad	Bad	Bad	E
Bad	Bad	Good	E
Bad	Bad	Excellent	E
Bad	Good	Bad	E
Bad	Good	Good	D
Bad	Good	Excellent	D
Bad	Excellent	Bad	E
Bad	Excellent	Good	D
Bad	Excellent	Excellent	C

Input			Output
X	**Y**	**Z**	**Assessment**
Good	Bad	Bad	E
Good	Bad	Good	E
Good	Bad	Excellent	C
Good	Good	Bad	D
Good	Good	Good	C
Good	Good	Excellent	C
Good	Excellent	Bad	D
Good	Excellent	Good	B
Good	Excellent	Excellent	B
Input			**Output**
X	**Y**	**Z**	**Assessment**
Excellent	Bad	Bad	E
Excellent	Bad	Good	D
Excellent	Bad	Excellent	C
Excellent	Good	Bad	D
Excellent	Good	Good	C
Excellent	Good	Excellent	B
Excellent	Excellent	Bad	C
Excellent	Excellent	Good	B
Excellent	Excellent	Excellent	A

4   Software Design and Implementation

The software running on the main PC workstation consists of three major modules as follows: data acquisition and recording, data processing and decision making. The data acquisition and recording software module connects to the two USB HD video cameras and records the two video feeds by using the open source C# library Aforge.net [11].

The video cameras run at the rate of approximately 30fps with the resolution of 1028x720 pixels per frame. The format mpeg4 is used to record the video streams with .avi file extension.

Tracking the movements of the laparoscopic tool tips in the defined virtual 3D spaces is the main task to be carried out by the data processing module. The Euclidean color filtering algorithm [12] is employed for tool tip tracking. The actual color tracking is implemented by properly setting up RGB filters in the relevant Aforge.net function calls.

The tool tip tracking program is implemented as a multi-thread environment. The main, or parent thread is the one where all other threads are created and maintained while the system is running. Each time a camera has sent a new frame an event-based thread is launched. In this thread, the new frame is placed in a waiting queue for processing. There are two waiting queues in the system, each of them is assigned to one particular camera.

Two similar threads are tasked to process the frames for identifying the tool tip positions, one for each camera. The thread is designed to implement the functions as follows: check the status of the waiting queue for a new frame to process, execute the algorithm to find the dedicated tool tip position (either right or left, respectively) in the 3D spaces and save the result in a queue. After the completion of these functions, the thread is scheduled to go into sleep mode for 10 ms. After that it executes the same algorithm again, all the time.

A third thread is designed to check the status of the queues holding the tool tip position results and feed the tool tip position data as crisp inputs to the fuzzy logic-based assessment system. The fuzzy logic-based assessment system carries out the fuzzification task for those inputs, performs Mamdani-type [13] inference operations using the fuzzy knowledge base, executes the Center of Gravity defuzzification algorithm and, finally, generates crisp assessment output values in the form of a percentage scale.

There are additional supporting threads as well. Those threads implement utility functions as follows: recoding the video feeds to the PC hard-drives, controlling the communications between the main PC and the supervisor tablet and providing console information for the supervisor medical personal on the system status. The block diagram of these threads are shown in Fig. 5.

The supervisor tablet software is also based upon the multi-thread paradigm. Its main thread controls a set of threads such as the Graphical User Interface (GUI) update thread, wireless communication with the main PC thread and the performance assessment results archiving on the main PC thread.
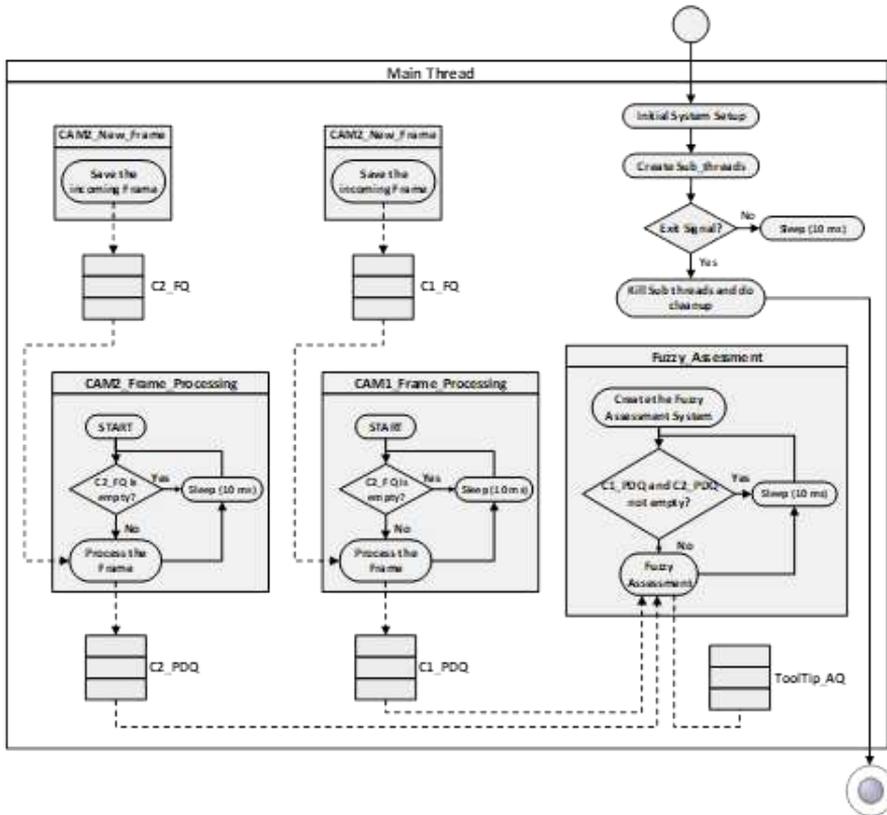
Figure 5
Software threads

5 Simulations and Live Test Results

The fuzzy logic-based tool tip movement assessment system was simulated using MATLAB. The multiple-input-single-output (MISO) system consists of three fuzzy inputs, a type-1 Mamdani [13] fuzzy inference system, and a single output. The knowledge base in the fuzzy inference system was represented by 27 fuzzy IF-THEN rules. The MATLAB system simulation architecture using the Fuzzy Logic Toolbox is shown in Fig. 6.
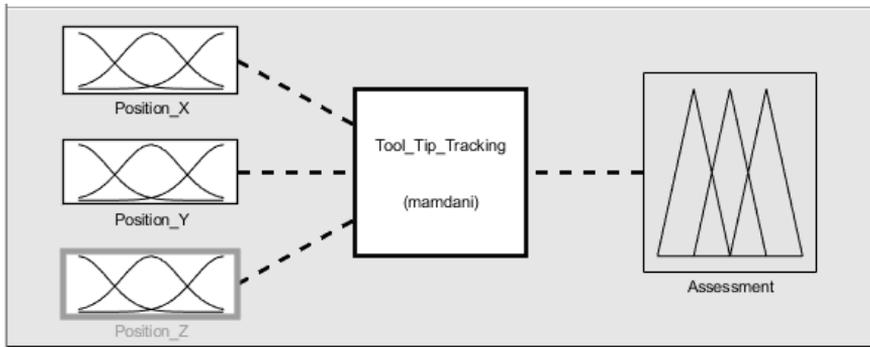
Figure 6
MATLAB simulations architecture

The intelligent BT system in a laboratory environment is depicted in Fig. 7. It is shown in preparation for a live experiment. During a test, a snapshot of the real time assessment results are displayed on the supervisor tablet in Fig. 8. After a test session has been completed the assessment results are stored in a text file as given in Fig. 9. RHTA and LHTA stand for Right Hand Tool Assessment and Left Hand Tool Assessment, respectively. In addition, the recorded video streams by both cameras are also made available for review. A monitor screenshot of the Camera 1 replay is shown in Fig. 10.

Regarding the tool tip tracking assessment capability provided by the intelligent BT system, Department of Surgery faculty are very much satisfied by the performance of the new intelligent BT device. The test outcome assessments are basically in agreement with their own. Since no similar device is available for use at this point, the results delivered by the intelligent assessment system cannot be compared with anything else. The fact that the system allows the recording (both camera video streams and also the fuzzy assessment results in a text file) of the performance of the residents during their tests opens a new chapter in teaching laparoscopic surgery skills. If requested, the residents could be provided with the recorded session results on a flash drive. That will allow them to study the tests' outcomes, by reviewing the records on their own computers. As an expected result, the quality of the laparoscopic surgery training program will significantly improve. That should benefit surgery patients and the society for the long run.

In addition, the approach to develop a fuzzy logic-based performance assessment support system will allow another major benefit. As more test data and experts' opinions will have been collected over time, the assessment results will become more consistent and more objective. It will represent a great improvement over the current, potentially subjective assessment methods.

Figure 7
Intelligent BT system
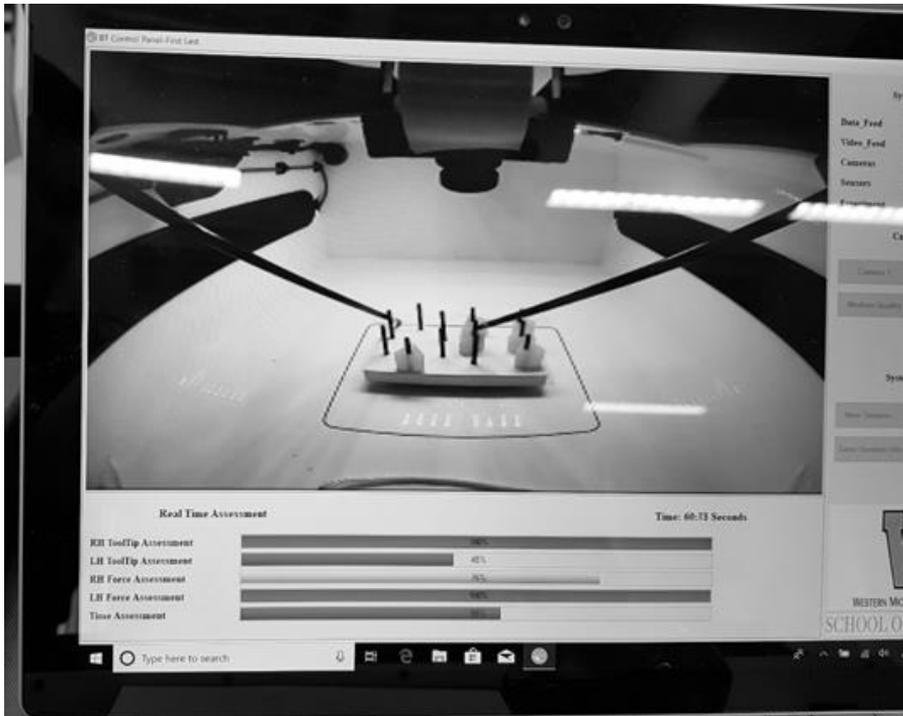
Figure 8
Real time display of assessment results during a test

```
------------------------------------------------------------------------
|                     Box Trainer Session Report                       |
------------------------------------------------------------------------
Date: Wednesday, August 7, 2019 1:50 PM
Trainee Name: Me Me
Experiment: PEG_TRANSFER
Level of Experience: VeryHigh
------------------------------------------------------------------------
Experiment Time: 82.86 Seconds
Right Hand Tool Tip Assessment: 60%
Left Hand Tool Tip Assessment: 28%
Right Hand Force Assessment: 78%
Lef Hand Force Tip Assessment: 77%
Time Assessment: 33%
Final Assessment: 45%
Comments: Accepted
Assessment Approved: Yes
------------------------------------------------------------------------
Time         RHFA          LHFA          RHTA          LHTA        TIMEA
------------------------------------------------------------------------
0.03         100%          100%           0%            0%          100%
0.33         100%          100%          100%          100%         100%
0.63         100%          100%          100%          100%         100%
0.93         100%          100%          100%           0%          100%
1.23         100%          100%          100%          100%         100%
1.53         100%          100%          100%          100%         100%
1.83         100%          100%          100%          100%         100%
2.13         100%          100%          100%          100%         100%
2.43         100%          100%           97%          100%         100%
2.73         100%          100%           91%           0%          100%
3.03         100%          100%           88%           0%          100%
3.33         100%          100%           94%           0%          100%
3.63         100%          100%           93%           0%          100%
3.93          86%           84%           79%           0%          100%
4.23          94%           82%           61%           0%          100%
4.53          75%           75%           24%          100%         100%
4.83          70%           74%           24%           0%          100%
5.13          64%           73%           23%          100%         100%
5.43          68%           68%           23%          100%         100%
5.73          75%           68%           23%           0%          100%
6.03          74%           74%           23%           0%          100%
6.33          78%           74%            0%          100%         100%
6.63          78%           75%            0%          100%         100%
6.93          79%           74%           23%          100%         100%
7.23          81%           75%           23%           0%          100%
7.53          83%           75%           23%          100%         100%
7.83          86%           77%           23%          100%         100%
8.13          86%           75%           23%           0%          100%
8.43          93%           76%           23%           0%          100%
8.73          93%           76%           23%           0%          100%
9.03          91%           76%            0%           0%          100%
9.33          94%           77%           23%           0%          100%
9.63          93%           77%           23%           0%          100%
9.93          93%           77%           23%           0%          100%
10.23         88%           77%           26%          100%         100%
```

Figure 9
BT session report

Figure 10
Replay of recorded video stream by Camera 1

6   Hardware Accelerator Design for Fuzzy Logic Based Edge Detection

Another possible approach for tool tip tracking, is based upon the edge detection method. The proposed fuzzy logic edge detection system is based upon a Mamdani-type MISO fuzzy system [13]. The system has four fuzzy inputs and one crisp output. The fuzzy knowledge base is made up of seven IF-THEN rules. Inputs one and two are the gradients regarding the X-axis and the Y-axis out of a kernel of 3×3 pixels. The respective fuzzy sets are denoted as GX, and GY.

$$GX = \sum_{x=1}^{3} \sum_{y=1}^{3} \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \cdot I(x,y) \tag{1}$$

$$GY = \sum_{x=1}^{3} \sum_{y=1}^{3} \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \cdot I(x,y) \tag{2}$$

$$LP = \sum_{x=1}^{3} \sum_{y=1}^{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \cdot \frac{1}{9} \cdot I(x,y) \tag{3}$$

$$HP = \sum_{x=1}^{3} \sum_{y=1}^{3} \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix} \cdot \frac{1}{9} \cdot I(x,y) \tag{4}$$
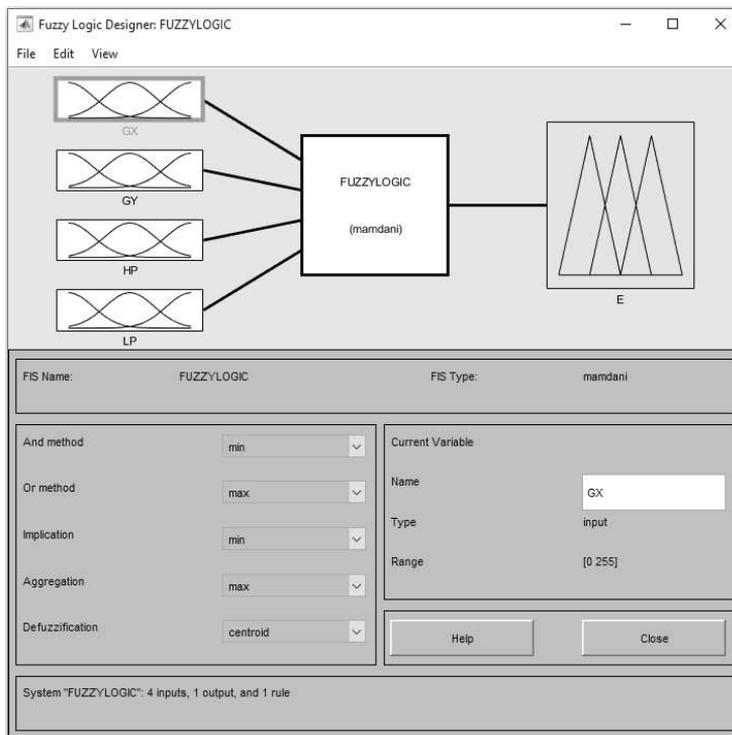
Figure 11
MATLAB simulations architecture for edge detection

Input three to the system is the output of a low-pass filter. The associated fuzzy set is denoted as LP. Input four to the system is the output of a high-pass filter and the associated fuzzy set is denoted as HP. The preprocessing system calculates GX, GY, LP and HP, respectively, using Equations (1) to (4). The performance of this system was simulated using the MATLAB Fuzzy Logic toolbox, shown in Fig. 11.

The proposed hardware accelerator architecture for this system consists of four modules as follows: preprocessing subsystem where raw grayscale values of the input pixels pass through four different linear spatial filters, fuzzification subsystem in which four fuzzy input variables are created by a fuzzification algorithm, fuzzy inference subsystem and defuzzification subsystem by which a single crisp output is generated. The knowledge base is implemented using seven IF-THEN rules. Various, different defuzzification methods were experimented with to assess the performance of the system.

The proposed pipelined hardware accelerator was designed and implemented on a number of Xilinx 7000-series Field Programmable Gate Array (FPGA) devices using the Xilinx Vivado Design Suite. For speedup, the architecture is made up of seven pipeline stages. The preprocessor subsystem is implemented by the first three stages, the inference subsystem utilizes two stages and the fuzzification and defuzzification subsystems, respectively, take just one each. Each stage requires just one clock cycle of execution time. The dynamic performance of the system was simulated using various Xilinx 7 Series FPGA devices.

Simulation results using 11 ns clock cycle time (i.e., the clock rate is slightly lower than the nominal maximum external clock frequency of 100 MHz) showed that the system needs 77 ns (7 cycles) to fill up the pipeline and after that it takes just 11 ns of execution time to process a pixel. For a pure software performance comparison, a MATLAB program-based implementation was running on a PC with an Intel Core i7 processor and 8 GB of memory. The software tests required 1.3178 millisecond execution time per pixel.

Conclusions and Future Research

The concepts of the tool tip tracking section of an intelligent, fuzzy logic-based performance assessment support system for the FLS BT device are presented. We propose tracking of the tool tip movements over the various test platforms in three dimensional virtual spaces of constrained sizes. By taking this approach the performance of surgery residents executing the certified FLS tests can be assessed in a quantifiable, objective and consistent way. The system provides files of recorded video clips from both cameras and also measured and processed sensory data. By using those files, the residents may review their performance which will help them to improve their laparoscopic surgery skills.

The quality of the assessment results can be improved by fine-tuning the membership functions and the fuzzy logic rule base, with the help of more test data and expert surgeon opinions. The plan is that in a research collaboration with

the Department of Surgical Research and Techniques, of Semmelweis University (SU), and the Antal Bejczy Center for Intelligent Robotics, of Obuda University, a fully equipped Intelligent BT System will be rented to SU. SU has a large number of residents in their surgery program and that will help in collecting large amount of data in a relatively short amount of time.

In the continuation of this project, we will be working on adding the Pattern Cutting Test and the three Suturing Tests, to the intelligent performance assessment system. The latter will be a very challenging project. Also in future research, hardware accelerators based on System-on-Chip (SoC) devices [14] will be investigated to potentially replace the main PC which is currently being used to perform the computationally extensive functions of the system. Using an SoC based system would facilitate the implementation of more complex and custom algorithms, at much higher processing rates.

Acknowledgement

References

[1]   Fundamentals of Laparoscopic Surgery, developed by Daniel J. Scott, MD, FACS – UT Southwestern Medical Center, Dallas, TX and E. Matt Ritter, MD, FACS – Uniformed Services University, Bethesda, MD, pp. 1-5, http://emorysurgery.org/docs/technicalcurriculum.pdf, approved by the SAGES FLS Committee 4/25/06

[2]   M. C. Vassiliou, G. A. Ghitulescu, L. S. Feldman, D. Stanbridge, K. Leffondre, H. H. Sigman, G. M. Fried, The MISTELS program to measure technical skill in laparoscopic surgery, Surg Endosc (2006) 20: pp. 744-747, DOI: 10.1007/s00464-005-3008-y © Springer Science+Business Media, Inc. 2006

[3]   S. L. de Montbrun and H. MacRae, Simulation in surgical education, Clinics in Colon and Rectal Surgery, 25(3), 2012, Sept., pp. 156-165

[4]   User Guide, Fundamentals of Laparoscopic Surgery Trainer System & Accessories by Limbs and Things Inc., pp. 1-12, Issue 5, Oct 2017

[5]   S. Muller et all, Simulation Training in Laparoscopic and Robotic Surgery, DOI 10.1007/978-1-4471-2930-1_1, © Springer-Verlag London, 2012

[6]   Carol E. Reily, Henry C. Lin, David D. Yuh and Gregory D. Hager, Review of Methods for Objective Surgical Skill Evaluation, Surg Endosc (20111) 25: pp. 356-366

[7]   Ignacio Oropesa, Patricia Sánchez-González, Magdalena K. Chmarra, Pablo Lamata, Alvaro Fernández, Juan A. Sánchez-Margallo, Frank Willem Jansen, Jenny Dankelman, Francisco M. Sánchez-Margallo, Enrique J.

Gómez, EVA: Laparoscopic Instrument Tracking Based on Endoscopic Video Analysis for Psychomotor Skills Assessment, *Surgical Endoscopy, 27*(3), pp. 1029-1039. DOI:10.1007/s00464-012-2513-z

[8]   Horeman, T., Dankelman, J., Jansen, F. W., & Dobbelsteen, J. J. (2014) Assessment of Laparoscopic Skills Based on Force and Motion Parameters, *IEEE Transactions on Biomedical Engineering, 61*(3), pp. 805-813, DOI:10.1109/tbme.2013.2290052

[9]   Janos L. Grantner, George A. Fodor, Ramakrishna Gottipati, Norali Pernalete and Sandra Edwards, Intelligent Decision Support System, IFSA 2007 World Congress, Cancun, Mexico, 18-21 June 2007, in Advances in Soft Computing, Vol. 41, Analysis and Design of Intelligent Systems, Patricia Melin, Oscar Castillo, Eduardo Gomez Ramirez, Janusz Kacprzyk and Witold Pedrycz (Eds.), ISBN: 978-3-540-72431-5, pp. 549-557

[10]   Janos L. Grantner, Aous H. Kurdi, Mohammed Al-Gailani and Ikhlas Abdel-Qader Robert G. Sawyer, M.D. and Saad Shebrain M.D., Intelligent Performance Assessment System for Laparoscopic Surgical Box-Trainer, Electronic Proceedings of the 2018 WCCI/FUZZ-IEEE Conference, July 8-13, 2018, Rio de Janeiro

[11]   Kirillov, A. (2013) Aforge.net framework, 2010-03-02, 2010-12-20

[12]   Tomasi, Carlo, and Roberto Manduchi. "Bilateral filtering for gray and color images." In *Computer Vision, 1998. Sixth International Conference on*, pp. 839-846, IEEE, 1998

[13]   E. H. Mamdani, Application of Fuzzy Logic to Approximate Reasoning Using Linguistic Systems, IEEE Transactions on Systems, Man, and Cybernetics, 1976, 26(12) pp. 1182-1191

[14]   Aous H. Kurdi, Janos L. Grantner, and Ikhlas M. Abdel-Qader, "Fuzzy Logic Based Hardware Accelerator with Partially Reconfigurable Defuzzification Stage for Image Edge Detection," International Journal of Reconfigurable Computing, Vol. 2017, Article ID 1325493, 13 pages, 2017, https://doi.org/10.1155/2017/1325493

Creating Randomness with Games

Jaak Henno*, Hannu Jaakkola**, Jukka Mäkelä***

*Tallinn Technical University, School of Information Technologies: Department of Software Sciences, Ehitajate tee 5, 19086 Tallinn, Estonia, jaak.henno@ttu.ee

**Tampere University, Pori Campus, P.O. Box 300, FI-28101 Pori, Finland, hannu.jaakkola@tuni.fi

*** University of Lapland, Rovaniemi Finland, jukka.makela@ulapland.fi

Abstract: In our increasingly connected and open World, randomness has become an endangered species. We may soon not have anything private, all out communication, interaction with others becomes publicly available. The only method to secure (temporarily) communication is mixing it with randomness – encoding it with random keys. But massive reuse of the same sources of randomness and rapid development of technology often reveals that used sources were not perfectly random. The Internet security is top-down, based on higher-level certificates, but we can never be quite certain with 'given from above' products in their quality – in order to beat each other producers are 'cutting corners' and even the high-level security certificates are available on Internet dark markets. This clearly shows in tremendous increase of all kind of security accidents, so there is an urgent need for new, independent sources of randomness. Mathematical treatment of randomness is based on infinite concepts, thus useless in practice with devices with finite memory (humans, computers, Internet Of Things). Here is introduced a definition for randomness based on devices with finite memory – k-randomness; it is shown, how this allows to create new randomness in computer games; numerous tests show, that this source is quite on par with established sources of randomness. Besides algorithmically-generated randomness is in computer games present also human-generated randomness - when competing players try to beat each other they invent new moves and tactics, i.e. introduce new randomness. This randomness appears in the sequence of players moves and when combined with the sequences of moves of other players can be used for generating secret keys for symmetric encryption in multi-player game communication system. The method does not use public-key step for creation of shared secret (the key), thus the encryption system does not need any upper-level security authorities.

Keywords: entropy; randomness; encryption; digital games; finite-state machines; human behavior; cyclic order; k-random sequences; player's actions combination

1   Introduction

Traditional fields of human activity – agriculture, manufacturing and construction are currently producing only 35% of all values consumed by humanity [1], the rest is produced in mental/information sphere, where the input for production of new values is data. Most important source of new data are we self and all producers are trying to capture as much as possible data concerning us.

Thus, it is becoming increasingly important to safeguard our privacy, our 'self', our data and our communication and for this we need randomness. Randomness has become a commercial product, several countries are introducing new random number generators [2] and with rapid increase of communication and data we have growing need for new randomness for encryption. Encryption ciphers are based on modifying messages using random data. But encryption is only temporary measure - when some encryption method/cypher is broken it becomes worthless and the randomness used in it also becomes worthless. Data breaches are increasing by more than 20 percent in a year [3], they have become the most worrying feature of Internet [4] and every breach is decreasing the value of randomness used. Growth of 'big data' inevitably increases amount of randomness needed to establish ownership for these 'big data' items. Thus, we constantly need new sources of randomness. New computing environments - Internet of Things (IoT), virtual/cloud servers etc. all increase the need for randomness.

To satisfy this continually growing need for randomness there are emerging dedicated services to serve random data [5]. For delivering this data was proposed a special new protocol 'Entropy as a Service' [6]. But for delivery this data also should be encrypted, thus it is a new source needing 'fresh' entropy. So it is not clear, whether this kind of 'top-down' service will reduce the need for entropy or contrary, increase it.

Trust in the current top-down security practices, based on higher-level security authorities issuing and controlling security certificates is decreasing – the high-level certificates are on sale on 'Dark Web' for $260 … $1,600 [7]. Security should be 'bottom-up' (Neighborhood Security) and entropy/randomness created just where it is needed (like in blockchain). Everything is simpler and safer if the entropy/randomness is generated where it is used.

Randomness is quite an infeasible concept. Mathematical treatment of randomness is based on infinite concepts, thus not applicable in real-world practice, where all information is handled by devices with limited memory - humans, computers, devices in IoT. Here is introduced a definition of *k*-randomness applicable to devices with finite memory and shown, how this can be used to produce with games of chance random integer sequences; tests show, that the created randomness is quite on level with established sources of randomness. As a practical use of generated with game randomness is introduced herein a method to create secure encryption keys for symmetric encryption.

2   Types of Randomness

It is impossible to generate random values using a computer's basic operations –
binary operations conjunction $\&$ (AND), disjunction $\vee$ (OR) and the unary
negation $\neg$ (NOT) – all combinations of these connectives return single
determined value (if not, then the computer is broken). John von Neumann
commented on this: "Anyone who attempts to generate random numbers by
deterministic means is, of course, living in a state of sin". But he did not elaborate:
"why?" "what are the 'non-sinful means' of creating randomness?" and what
actually is randomness

Computers are deterministic, orderly; randomness is the opposite of order, the
absence of any pattern. The current understanding is that 'true' randomness can be
extracted only from physical processes, which have rich deep inner structure –
entropy, e.g. thermal fluctuations in processor, pixels found by mouse sensor
when user makes some rapid random strokes, atmospheric disturbances [8] etc.
These sources are 'Pure Randomness Generators' (PRG), but they are often not
rich enough e.g. for network servers which do not have external devices.

Operating systems send 'new-born' messages, thus should have means for
securing them with 'new-born' randomness/entropy and for this all operating
systems maintain an entropy pool. The first versions of Linux kernel created
entropy from the third derivative of differences in timings of user actions; this
information is stored in two files */dev/random* and */dev/urandom*. This method
turned out to be too slow and currently uselow-order bits (lest significant, i.e.
changing most rapidly) from timing of user actions on keyboard, mouse
movements, IDE requests; extracting entropy from audio [9] and video [10] data is
also studied.

Programming language's compilers need methods to create random values [11],
[12]. All compilers work under an Operation System (OS) and get their
randomness from OS, e.g. in Windows environment randomness to all
programming languages comes from the same source as to the Microsoft C/C++
compiler (and the Intel compiler [13] or newer [14]) - they use the random values
generated in Common Language Runtime [15], using the entropy produced by
processor. But there have been found several problems for Intel processors [16],
[17] thus specialists distrust randomness produced by Intel processors [18], e.g. in
Linux kernel it is only one of many inputs into the random pool. Research has
shown that even processors built-in functions (PRG-s) for generating random
values can be compromised [19] and processors and microchips may have built-in
hardware rojans [20] which can leak information leading to successful key
recovery attacks. After the NSA (U.S. National Security Agency) leaks by Edward
Snowden, many engineers have lost faith in hardware randomness [21].

The hardware entropy pool decrease every time random numbers are generated
from it Requesting many random numbers may starve them; this is a practical
issue on servers without input devices. Other PRG sources also decrease, e.g. the

online source of randomness *Random.org* [22] limits its daily available amount of free random bytes (currently 10^6 bits). Thus PRG sources do not suffice, for random number generation are needed also computer algorithms.

Computers are finite devices and after a while 'fall into loop', start to repeat computed values. Thus 'calculated randomness' is pseudo-randomness produced by pseudo-random number generators (PRNG). All PRNG-s are loops, which after their period repeat produced values.

The first value in the loop is created using a random seed, i.e. comes from other, usually PRG source. The next value is calculated from the previous one by some recurrent function; common method is to use linear (for speed) recurrent functions with reduction by modulus. For these Congruential Generators (CG) is the period (length of the loop) the most important measure of security of such a generator. For the C language it should be at least $2^{32} = 32767$ [23] - a rather small number for current CPU-s and its use (installing the Microsoft or GNU suite of compilers) requires decent computer skills. A 'high-end' PRNG–s have much bigger period, e.g. period of the 'mersenne twister' is the Mersenne prime $2^{19937} - 1$, but use of these requires good computer skills and good hardware.

Many PRNG-s which at their introduction were considered 'good enough' have later become obsolete. For example, John von Neumann used for generation of random numbers the 'middle-square' method [24] – for the recurrence step earlier produced number was squared and then the middle digits were sliced out. This mix of number's semantics (squaring) and syntax (use only middle digits) was used already in 13th century [25] and seems good, since un-computability results (e.g. the Rice theorem [26]) indicate, that most semantic properties are undecidable from syntax. However, research revealed that with *n*-bit seed the length of generated cycle is $\leq 8^n$ and with many seeds even much shorter, e.g. $3792 \to 79^2 = 6241 \to 24^2 = 0576 \to 57^2 = 3249 \to 24^2$ - a cycle.

Many PRNG-s have similar fate. The RC4 (Rivest Cipher 4) was used in several commercial encryption protocols and standards (e.g. in the TLS - Transport Layer Security – the base of all traffic in WWW), but is currently prohibited; widely known was periodicity in the random function of Microsoft PHP translator. Already in 1999 were presented general methods for prediction of CG-s [27], [28].

For assessment of quality of new PRNG-s have been constructed several suites of statistical tests – the NIST (the U.S. National Institute of Standards and Technology) suite [29], the Dieharder (Marsaglia) suite [30], ENT [31] etc. These tests check presented samples for some common regularities in everyday data, e.g. the Dieharder 3.20 implements 26 tests.

We tested with the ENT suite several established sources:

1. The first 7 KB part of the 2.1 GB file */dev/urandom* from Ubuntu 16.04.3 (a three months old installation, used mainly for making music)

2. 10000 decimal digits downloaded from the *Random.org* (randomness from atmosphere);

3. 10000 decimal digits created using the function *window.crypto.getRandomValues( );*

4. 10000 decimal digits created by Wolfram Mathworld with function *RandomInteger[ ]* using the default method *Rule30CA*

In the following table are shown three characteristics from the test with the ENT suite of statistical tests: entropy (bits per bit), possible compression (randomness can't be compressed) and serial correlation coefficient.

Table 1.
Some characteristics of established sources of randomness

	Entropy	Compression	Correlation
/dev/urandom	0.988577	1%	0.035161
Random.org	0.919040	8%	0.060193
Windows	0.974450	2%	-0.010378
Wolfram	0.974448	2%	-0.010948

The results are rather similar except a bit weaker performance of atmosphere processes – the random sequence downloaded from the site *Random.org*.

However, statistical tests cannot guarantee randomness and the results of these tests do not tell the whole truth. Although the randomness from Linux performed best, visual inspection (the 'Statistics' tool from the free hex editor HxD) reveals, that distribution of frequencies in */dev/urandom* contains a surprising peak.
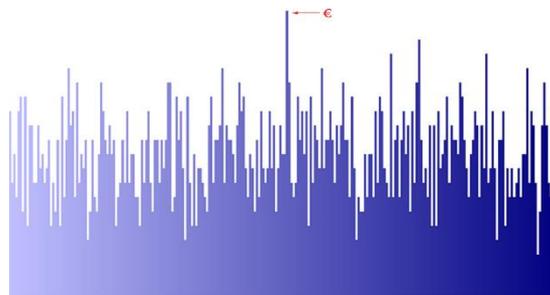
Figure 1.
Distribution of frequencies of bytes in the first 7 KB from the */dev/urandom* file from Ubuntu 16.04.3;
the sharp peak in the middle is the code for the € (Euro) symbol

Thus, statistical tests are also uncertain method for evaluation of randomness sources. There could be surprising dependencies in data - the above peak in € symbol code come from computer, which is rather new (in use only for several months) and was never used for any kind of financial data handling.

3   Randomness from Games

Randomness is an essential ingredient in most games and for utilizing this source have been proposed different methods [32] [33]. In [34] authors presented a method for producing on-line in real playtime binary random strings from simple repeated games; here the principles of the proposed method are applied to produce from gameplay *m*-ary ($m > 2$) random sequences.

In the (economics - based) texts on games the game decision mechanism is usually not detailed – it is determined by unpredictable markets. In video games decisions are deterministic, thus we follow computer science tradition (see e.g. [35]) and use for the decision mechanism finite automaton. Games considered here are 'games of luck', where both players have equal chance to win and the best strategy (the Nash equilibrium) for both players is total randomness, i.e. in the game payoff matrix (economics-based format) the sums of all rows and columns are equal, such games are e.g. the rock-paper-scissors and odd-even.

If one player is human or some established source of randomness and the other – the computer algorithm, then the (statistical) result of numerous repeated plays is also an assessment for the quality of computer-created randomness. The length of all considered here random sequences/plays is 10000, following the suggestion: "A reasonable estimate (for humanly interesting cases) reckons that some 10,000 digits would suffice" [36]

Thus in the following game is a structure $G = <P_1, P_2, M, R, A>$, where

P_1, P_2 are (two) players;

$M = \{0, 1, ..., m-1\}$, $m > 2$ - the set of legal moves (actions) of players (the same set for both players); in every round both players apply simultaneously one action which initiates some change in automaton A

$R = [r_1, r_2]$ - player's utilities (points); at the start $r_1 = r_2 = 0$

A - a finite automaton, deciding the output (move) and payoffs to players. Here are considered simultaneous (synchronous) games, where players produce their actions (moves) simultaneously at the same time, thus the input for the automaton A are pairs (m_{1i}, m_{2i}), where m_{1i} is the i move the first player, m_{2i} - i move second player; denote $(m_{1i}, m_{2i})^{-1} = (m_{2i}, m_{1i})$ - actions of players switched.

Automaton's (possible) outputs are "1" (player P_1 won, $r_1 += 1$), "-1" (player P_2 won, $r_2 += 1$), "0" – draw. Thus, the automaton has four distinguished states:
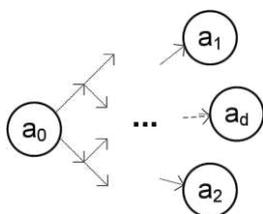
Figure 2.

Game automaton with distinguished states: a_0 – the start state, a_1 – first player won, a_2 – second player won, a_d – draw

Here a_0 is the game start state, a_1 - here automaton outputs, that the first player won, a_2 - the second player won, a_d - draw; in any other state (not shown) automaton does not produce output; $F = \{a_1, a_2, a_d\}$ is the set of final states

Automaton does not have cycles, thus the graph of the automaton is a tree (with possible loops with limited length at some nodes) and all rounds are finite. The length $D(A)$ of the longest round (the depth of the tree) is the depth of the game. Game is repeated and after some fixed number (here 10000) of moves automaton announces if the result of the game is draw or who won.

Automaton is deterministic, i.e. in any state $a \in A$, $a \notin F$ and for any move (m_i, m_j) there is a single transition $a(m_i, m_j) \rightarrow a' \in A \setminus \{a_0\}$

Transitions are in natural way extended to words from
$$M^{2D(A)} = \{(m_i, m_j)\}^+ = \{(m_{11} m_{21})...(m_{1t} m_{2t}), \ t < D(A)\}$$

$$a((m_{11}, m_{21})(m_{12}, m_{22})...(m_{1t}, m_{2t})) = $$
$$(a(m_{11}, m_{21}))((m_{12}, m_{22})...(m_{1t}, m_{2t}))$$

Action of words on states of automaton A creates partition of the set $M^{2D(A)}$ of words into three sub-languages:

$$L_1 = \{w \mid a_0 w = a_1\}$$

$$L_{2=} = \{w \mid a_0 w = a_2\}$$

$$L_d = \{w \mid a_0 w = a_d\}$$

Call all words $w \in M^{2D(A)}$ plays.

4   Symmetry-based Non-Learnable Games

In games of chance, nobody wants to have worst chances by design of the game and all actions of players should be significant, i.e. could change the result, thus these games obey the following symmetry principle:

in any move (m_i, m_j) from any play both players have equal chances, i.e. if all other moves in the play remain the same they can change their action so that expectation of outcomes a_1, a_2 is the same, i.e. 0.5.

Since there are m^2 possibly moves it follows that if m is odd, the set L_0 can't be empty – otherwise $|L_1| = |L_2|$ is impossible.

Therefore the games with $D(A) = 1$ (one round, i.e. every single state of the automaton) should satisfy the following conditions.

1. All games are zero-sum, i.e. the involution $\alpha : (m_i, m_j) \rightarrow (m_i, m_j)^{-1} = (m_j, m_i)$ produces an automorphism of the automaton A , i.e. $\alpha(L_1) \subseteq L_2$, $\alpha(L_2) \subseteq L_1$, $\alpha(L_d) \subseteq L_d$.

2. Any substitution

$$\beta : \{m_0, ..., m_{k-1}\} \overset{1-1}{\rightarrow} \{m_0, ..., m_{k-1}\}$$

of actions produces automorphism of automaton A , which does not break the partition $\{L_1, L_2, L_d\}$.

3. The sublanguage L_d contains all words (m_i, m_i), $m_i \in M$ and is minimal – it should not contain words which could be moved into L_1 or L_2 without breaking conditions 1,2.

Proposition. Conditions 1-3 define for given m unique (up to isomorphism) game payoff function.

Proof. Consider the set $M_1 = \{(m_1, m_i), m_i \in M \setminus \{m_1\}\}$, i.e. moves, where the first player selects action m_1. From the condition 3. it follows, that the set L_d can contain at most one of them, otherwise we could pairwise move them one to L_1, another to L_2 without breaking conditions 1.-2.

From the condition 2. it follows, that there should be equal number of elements from the set M_1 inside sets L_1, L_2. If these sets were of different size then substitutions which keep m_1 fixed, but move other actions will break the condition 2.

According to condition 2. actions inside L_1 could be re-arranged so that $(m_1, m_2), (m_1, m_3), ..., (m_1, m_{k1}) \in L_1$, $\quad k1 = \lfloor m/2 \rfloor$. Using the substitution $\chi : m_i \rightarrow m_{i+k1}$ and the property 2. we get that all sets $\{(m_i, m_{i+1}), (m_i, m_{i+2}), ..., (m_i, m_{i+k1})\}$ should belong to L_1. If m is odd, then all moves are now evenly divided between sets L_1 and L_2. If m is even, then from the above discussion it follows that the moves $(m_i, m_{i+k/2})$ should belong both to L_2 and L_1, i.e. they should be moved to set L_d.

Thus, a game with properties 1.-3.- has an unique (up to involution α) payoff function, based on cyclic order [37] on moves: if moves of players are $P_1(m) = m_i$, $P_2(m) = m_j$, then output from the automaton A is:

$$\text{sgn}((m_i - m_j) \bmod m - m/2)$$

Table 2.

Decision table of cyclic 5-ary order; e.g. $(0, 2), (1, 2) \in L_1$, but $(0, 3), (2, 1) \in L_2$

	0	1	2	3	4
0	0	1	1	-1	-1
1	-1	0	1	1	-1
2	-1	-1	0	1	1
3	1	-1	-1	0	1
4	1	1	-1	-1	0

In case $m = 3$ this is isomorphic to the well-known game rock-paper-scissors, which appeared in China at the beginning of Current Era. Apparently, Chinese know how to use symmetry groups for inventing amusing games.

There are variants of this game with greater cycle length e.g. movements of fighters can be *punch, kick, grab, push* (the next one stronger than the previous), in some games even more than ten with non-linear order [38]. The pay-offs could be any (increasing) sequence of numbers, e.g. in the above 5-ary game the payoffs (ordinals) could be integers [-2,-1,0,1,2] calculated by $(m_1 - m_2) \bmod m - m$.

Table 3.

Payoff table for the first player in a cyclic 5-ary game (payoffs for the second player – multiply by -1 – game is zero-sum)

	0	1	2	3	4
0	0	-1	-2	2	1
1	1	0	-1	-2	2
2	2	1	0	-1	-2
3	-2	2	1	0	-1
4	-1	-2	2	1	0

5   Randomness for Finite Devices – k-randomness

Games where all actions produce similar reward (payoff) are non-learnable even for Tensorflow [39]. Nevertheless there are professional players of rock-paper-scissors [40], in USA is a league of professional players of the "Rock, Paper, Scissors" game [41], regular tournaments [42] and programming competition [43]. These professionals win their opponents not since they have learned the game (impossible!), but they have learned to learn their human opponents, i.e. create better randomness than their opponents.

Randomness is an evasive concept to define. The widely accepted definition is the Kolmogorov- Martin-Löf definitions: [44] [45]

sequence is random if it can't be compressed - expressed by any algorithm or device which can be described using less symbols than what are in the sequence.

This definition and other consequent definitions, e.g. [46] are using infinite concepts ('any algorithm') and apply to infinite sequences, thus useless in practice for evaluating quality of a source of randomness, where all actors/devices are finite (have finite memory) and produce finite sequences. All deterministic devices inevitably go to cycle after enough time (they do not have any new states and have to repeat already used states). Thus if deterministic devices A_1, A_2 with finite memory (humans or computers) interact, the deciding factor (who can predict/learn whom or contrary, cannot learn/predict and concludes, that the other produces random output) depends on available memory of these interacting devices [47], [48]. If A_1 has (sufficiently) more memory than A_2 so that it can remember the whole cycle produced by A_2, then when A_2 goes into cycle A_1 can always predict the next response of A_2 (A_1 has learned, 'pwnd' A_2), but since A_2 cannot store in its smaller memory the whole cycle produced by A_1, it has to conclude, that A_1 is creating random output. This insight is the base for the following definition:

a finite sequence of integers is k-random if its length > k and it can't be created as the sequence of outputs by any deterministic finite automaton with less than k states.

This definition can be expressed also in terms of the Zif-Lempel compression [49] thus this is a (particular case) of the Kolmogorov's definition:

a finite sequence of length > k is k-random if it can't be compressed using dictionary with item length < k.

When $k \rightarrow \infty$ this definition yields the presented above definition. All PRNG-s are interactive (input is the seed) deterministic finite automata – with the same

seed they produce the same output and a PRNG with cycle length k produces (maximally) a k-random sequence.

6   The Cycle Disruption Algorithm

The k-randomness of a sequence (with length $> k$) actually means, that the sequence does not have a loop at its end. Any finite deterministic automaton with k states and m input symbols produces a periodic sequence [50], i.e. 'goes into loop', if the length of its input is longer than $k \times m$ - there are no new possibilities for the pair (*state, input*), thus a deterministic automaton produces the same subsequence which already occurred when it first arrived at (*state, input*). The evolutionary game theory of bounded rationality [51] of human players also predicts cyclic patterns in playing behavior [52]. Thus for successful play one has to find when the loop begins, i.e. automaton repeats its moves and then disrupt the loop. This is the idea of the loop disruption algorithm for creating random sequences:

scan the sequence of stored moves (input-output pairs) and when you find a situation similar to the current one (see that the sequence of last moves already appeared earlier) make the move that in the previous situation would be winning.

Suppose the sequence of moves in a game up to now is

$$a_0(m_{11}, m_{21})(m_{12}, m_{22}),...,(m_{1n}, m_{2n}),...(m_{1k}, m_{2k})(m_{1k+1}, m_{2k+1})...,$$
$$(m_{1n}, m_{2n}),...(m_{1k}, m_{2k})$$

and $(m_{1n}, m_{2n}),...,(m_{1k}, m_{2k})$ is the longest repeated subsequence of moves (looking from the current state backwards). Then algorithm should select the move which wins in the state (m_{1k+1}, m_{2k+1}).

For instance, in the following situation from a real play of 3-ary game (moves follow in pairs, first human then computer, e.g. on the second move human played '1', computer – '2') computer discovered a repeated sequence (underlined), thus its next move will be '2':

1, 1, 1, 2, 2, 2, 0, 1, 0, 1, 1, 1, 1, 0, 1, 1, 0, 2, 2, 2, 2, 1, 0, 0, 1, 1, 0, 2, 0, 0, 2, 2, 1, 1, 2, 2, 2, 1, 1, 1, 2, 2, 2, 0, 2, 2, 0, 1, 1, 1, 1, 2, 0, 0, 2, 2, 2, 2, 1, 1, 0, 0, 1, 2, <u>1, 0, 0, 2, 2, 0, 2, 0,</u> 2, 0, 0, 0, 0, 2, 0, 0, 2, 0, <u>1, 0, 0, 2, 2, 0, 2, 0</u>

The above sequence shows 45 moves (there are 90 symbols); length of the repeated subsequence is 4 moves; thus, the all sequence is 41-random.

The algorithm has been implemented in several browser games [53].

7   Tests

We tested the algorithm using it as the computer opponent in several games of luck (odd-even, rock-paper-scissors etc.) in many plays. Against human players (students from the Tallinn University of Technology and others) computer was in most cases already winning if the length of the game was >30. Humans are not sufficiently random to beat computer, especially if the memory requirements (length of the game) grows; it seems that here works the famous human short-term memory size principle – human memory is also finite [54].

As opponent players for testing were used several well-established sources of randomness: JavaScript's functions *Math.random()* and *window.crypto.getRandomValues()*, random numbers produced by Wolfram's *Mathematica* and a table of 10000 random integers downloaded from *https://www.random.org/*.

Tests indicated that the algorithm plays quite well against all these common sources of 'computed' randomness, i.e. its own randomness is on the same level. Below is a table of results from three tests, each a 10 series of plays, each play 10000 rounds with $m = 3$. Player P_1 is in the first test random numbers produced by the JavaScript function *Math.random()*, in the second – random numbers produced by the function *RandomInteger*[] of Wolfram's *Mathematica* (using the default rule *Rule30CA* in Mathematica for creating pseudorandom sequences) and in the third – random numbers produced by function *window.crypto.getRandomValues()*; player P_2 is our algorithm; L was the length of the longest cycle in the sequence of player's moves (i.e. the repeated sequence in above example). The last row indicates how many times each player won and length of the longest repeated sequence.

Table 4.

Results of tests

P_1	P_2	L	P_1	P_2	L	P_1	P_2	L
3350	3365	16	3403	3289	16	3356	3287	18
3396	3237	16	3369	3242	20	3277	3285	16
3328	3332	16	3392	3286	16	3281	3351	18
3428	3209	18	3392	3317	18	3342	3305	18
3310	3377	16	3512	3163	16	3299	3405	16
3369	3365	16	3424	3278	18	3366	3259	16
3360	3345	16	3440	3316	18	3367	3263	16
3315	3402	16	3355	3265	18	3283	3446	20
3322	3412	18	3409	3301	19	3383	3354	16
3294	3364	16	3330	3453	16	3324	3314	16
4	6	18	9	1	20	6	4	20

These results show, that used in tests sequences were (at least) 9980-random according to the above definition – they did not contain repeated sequences longer than 20 moves.

In the following table are discretized results (showing not actual results, but showing how many times player was better than the opponent) from 10×10000 series of tests against random numbers table from *Random.org* (the first column in all three sub-partitions), JavaScript function *Math.Random()* (the second column in all three sub-partitions) and the function *window.crypto.getRandomValues()* (the third column); the last row is the summary of results.

Table 5.
Discretized results of tests

Better P_1			Better P_2			Draw		
4	7	2	4	3	8	2	0	0
6	4	7	4	3	3	0	3	0
5	7	1	3	2	9	2	1	0
6	3	1	3	7	9	1	0	0
2	5	2	7	4	8	1	1	0
3	4	4	5	5	5	2	1	1
4	6	4	4	4	6	2	0	0
4	4	3	5	5	7	1	1	0
7	4	3	3	2	7	0	4	0
4	4	4	4	4	3	2	2	3
85	100	75	88	74	115	27	26	10

As seen from this table, our algorithm was nearly on the same level against *Math.Random()*, slightly outperformed the randomness from *Random.org* and slightly lost to *window.crypto.getRandomValues()*.

As output (new randomness) could be used two sequences – the sequence of 'full' moves (pairs of moves from player and computer) or the sequence of only computer-generated moves (twice shorter). We tested both as the source of random sequence against our computer's algorithm. In the following table are results from 10 series of plays, each 10000 rounds with $m = 3$; player P_1 is in the first series (the first three columns of the table) generated in a previous game (10000 moves against JavaScript *Random()*) sequence of full moves (pairs), in the second (the last three columns) – sequence of computer moves; player P_2 is our algorithm.

According to Table 6 the created in the game randomness already mostly outperformed our algorithm, its results are better than that of commonly established sources. When the generation process was iterated, i.e. generated randomness was used as input for the next play, it become more difficult to predict and our algorithm started to loose.

Table 6.
Tests against randomness, created in game

P_1	P_2	L	P_1	P_2	L
3298	3344	16	3403	3275	16
3377	3351	16	3328	3337	16
3439	3303	16	3391	3342	16
3419	3284	16	3375	3297	18
3328	3376	16	3490	3272	18
3471	3212	16	3408	3273	18
3360	3294	20	3379	3343	20
3367	3314	16	3342	3370	16
3513	3250	16	3376	3288	16
3416	3362	16	3362	3316	18
7	3	20	8	2	20

In the following table are results of play against randomness, created on third iteration, i.e. after three rounds of 10x50000 moves; player P_1 is in the first column the table of full moves (pairs), in the second – sequence of computer-generated moves.

Table 7.
Tests with iterated randomness

P_1	P_2	L	P_1	P_2	L
16811	16740	22	16617	16834	24
16840	16550	18	16636	16759	18
16785	16599	20	16729	16592	20
16779	16701	18	16703	16641	20
16777	16412	18	16601	16720	18
16928	16672	18	16801	16682	20
16904	16610	20	16757	16589	20
16902	16599	18	16787	16547	22
17017	16445	22	16581	16702	20
16680	16655	20	16458	16854	18
10	0	22	5	5	20

8   Use in Practice - Creating Encryption Keys with the Move Sequences Combination

Participants of online multiuser communities (multiplayer games, social networks) often want to establish also a direct communication with fellow players (chat).

This communication/chat system should not burden the game server, thus has to be implemented as a separate sub-process.
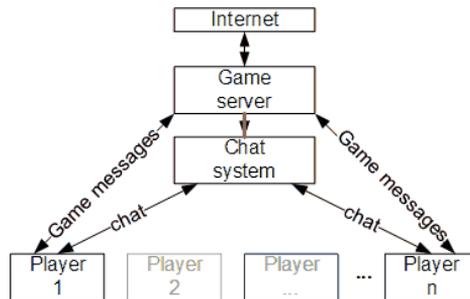
Figure 3.
Information flows in a multiplayer game with communication (chat system) for players

To ensure security of game and players communication (this may involve exchange of substantial game values) the communication system should be 'sand-boxed', should be encrypted and should not reveal any information to outside/Internet. This makes undesirable the commonly used first phase of encryption key creation – use of public-key encrypted communication, which requires security certificates from outside.
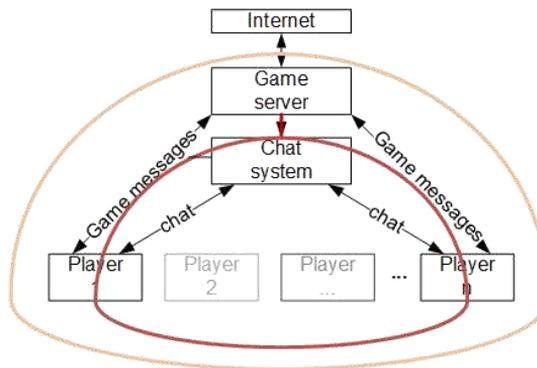
Figure 4.
Security surfaces of information flows in a multiplayer game with communication (chat system)

Statement "Players of games create randomness" is similar to many other non-provable statements. Faith in its correctness comes from long history of game-playing – nobody would play games where everything is pre-determined, just randomness makes games enjoyable. We play them more and more, thus in gameplay is created randomness and this randomness could be utilized.

A multiplayer game is a communication system where players constantly generate new randomness with their moves, thus for key generation could be used

randomness from player's moves; for greater security could be added also a computer-directed player, who for its play uses the algorithm presented above.

The server records sequence of player's moves, e.g. for a game with two players *Alice* and *Bob* this sequence of their moves could be $m_{11}m_{21}m_{12}m_{22},...,m_{1t}m_{2t},...,m_{1l}m_{2l}$; here m_{1t}, m_{2t} are respectively moves of *Alice* and *Bob* in gameplay move/moment *t*.

To generate a key server sends to players the sequence of all moves from which the player's own moves are removed, e.g. server sends to Alice the sequence $*m_{21}*m_{22},...,*m_{2t},...,*m_{2l}$ - this information with holes does not give to an eavesdropper any information (it is assumed, that the game server communication with players is secure;, here is the only time when the game communication is used for the chat system). When players replace holes in the received sequences with their own moves they all get the same random sequence which could be used as the secure random key for symmetric encryption.
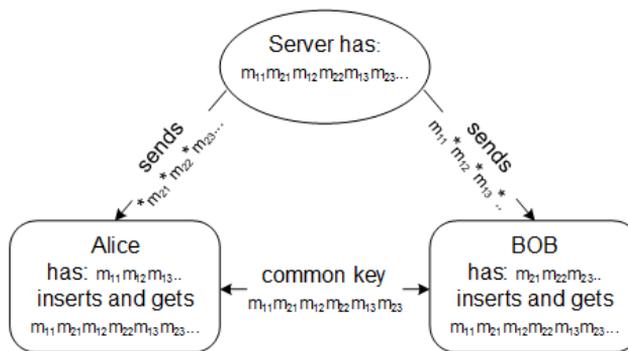
Figure 5.

Key generation combining a sequence with holes from server with sequence of player moves

This 'move sequences combination method' for symmetric key generation could be applied in any multiplayer game where players send their moves to game server (it is not essential, that moves alternate as in the above example). It has many desirable properties: key may be created for any subset of players (for any pair of them or for the whole player's community), after the first (secure) communication players could easily switch to a new key (without announcing the server) just with message „From now on use moves from time moment t_0 to t_1 " etc. To increase security of the key server could use some filters, e.g. remove all moves certain properties (produced certain result); to speed up the game could be used multi-moves, i.e. participants send in every move a fixed-length sequence of moves etc.; several test applications are in implementation.

Conclusions

This work analyzed the concept of randomness (*k*-randomness), applicable for use in devices with finite memory humans orcomputers. A method was presented (the loop disruption algorithm) for creating random sequences in gameplay; the quality of the created randomness was tested in a series of plays against established sources of randomness. Tests show, that the randomness is quite on a par with established sources of random numbers. As a practical use of game-created randomness is shown how this could be used for generating secure encryption keys for symmetric encryption without using the open-key procedure, typically used for creating common random sequence; the introduced 'moves combination method' is currently under implementation. Using a generated in game randomness for symmetric encryption makes such a communication systems very secure – they do not depend on any 'upper-level' security principals or certificates for key creation.

References

[1]   FAO 2019. The future of food and agriculture: Trends and challenges. Food and Agriculture Organization of the United Nations, ISSN 2522-7211

[2]   Sophia Chen. Why are countries creating public random number generators? Science, Jun 28, 2018

[3]   BSA. Encryption: Why It Matters. Retrieved May 9, 2018 from http://encryption.bsa.org/

[4]   Fortinet 2018. Data Breaches Are A Growing Epidemic. How Do You Ensure You're Not Next? Retrieved May 08,2018 from https://www.fortinet.com/blog/threat-research/data-breaches-are-a-growing-epidemic--how-do-you-ensure-you-re-n.html

[5]   Nist 2016. Entropy as a Service. https://csrc.nist.gov/projects/entropy-as-a-service

[6]   EaaSP 2018. EaaSP - Entropy as a Service Protocol. https://github.com/usnistgov/EaaS

[7]   David Maimon et all 2019. SSL/TLS Certificates and Their Prevalence on the Dark Web – Venafi https://www.venafi.com/sites/default/files/2019-02/Dark-Web-WP.pdf

[8]   https://www.random.org

[9]   vanheusden.com 2018. audio entropy daemon. https://vanheusden.com/aed

[10]   vanheusden.com 2018. video_entropyd. https://vanheusden.com/ved/

[11]   Generating Random Data in Python. https://realpython.com/python-random/

[12] How to generate random numbers, characters, and sequences in Scala. https://alvinalexander.com/scala/how-to-generate-random-numbers-characters-sequences-in-scala

[13] CryptGenRandom. https://docs.microsoft.com/en-us/windows/desktop/api/wincrypt/nf-wincrypt-cryptgenrandom

[14] RtlGenRandom function. https://docs.microsoft.com/en-us/windows/desktop/api/ntsecapi/nf-ntsecapi-rtlgenrandom

[15] Common Language Runtime (CLR) overview. https://docs.microsoft.com/en-us/dotnet/standard/clr

[16] A Provable-Security Analysis of Intel's Secure Key RNG. https://eprint.iacr.org/2014/504.pdf

[17] Gagallium: How I found a bug in Intel Skylake processors. http://gallium.inria.fr/blog/intel-skylake-bug/

[18] The Register. Torvalds shoots down call to yank 'backdoored' Intel RdRand in Linux crypto. Sept 10, 2013, https://www.theregister.co.uk/2013/09/10/torvalds_on_rrrand_nsa_gchq/

[19] G. T. Becker, F. Regazzoni, C. Paar, W. P. Burleson. Stealthy Dopant-Level Hardware Trojans. Journal of Cryptographic Engineering, April 2014, Volume 4:1, pp 19-31

[20] M. Ender, S. Ghandali, A. Moradi, C. Paar. The First Thorough Side-Channel Hardware Trojan. https://eprint.iacr.org/2017/865.pdf

[21] arstechnica. "We cannot trust" Intel and Via's chip-based crypto, FreeBSD developers say. https://arstechnica.com/information-technology/2013/12/we-cannot-trust-intel-and-vias-chip-based-crypto-freebsd-developers-say/

[22] https://www.random.org

[23] ISO/IEC 9899:2011. https://www.iso.org/standard/57853.html

[24] John von Neumann, "Various techniques used in connection with random digits," in A. S. Householder, G. E. Forsythe, and H. H. Germond, eds., Monte Carlo Method, National Bureau of Standards Applied Mathematics Series, vol. 12 (Washington, D.C.: U.S. Government Printing Office, 1951): pp. 36-38

[25] I. Ekeland. The Broken Dice, and Other Mathematical Tales of Chance. University of Chicago Press 1993, pp. 1-190, ISBN: 9780226199924

[26] H. G. Rice. "Classes of Recursively Enumerable Sets and Their Decision Problems". Trans. Amer. Math. Soc. 74 1953, pp. 358-366

[27] H. Krawczyk. How to predict congruential generators. Journal of Algorithms, Vol. 13:4, December 1992, pp. 527-545

[28] J. Stern. Secret linear congruential generators are not cryptographically secure. 28th Annual Symposium on Foundations of Computer Science (sfcs 1987), DOI: 10.1109/SFCS.1987.51

[29] NIST SP 800-22. A Statistical Test Suite for Random. https://nvlpubs.nist.gov/nistpubs/legacy/sp/nistspecialpublication800-22r1a.pdf

[30] R. G.Brown. Dieharder: A Random Number Test Suite. https://webhome.phy.duke.edu/~rgb/General/dieharder.php

[31] ENT. A Pseudorandom Number Sequence Test Program. http://www.fourmilab.ch/random/

[32] R. Halprin, M. Naor. Games for Extracting Randomnes. SOUPS '09 Proceedings of the 5th Symposium on Usable Privacy and Security 2009

[33] M. Alimomeni, R. Safavi-Naini. Human Assisted Randomness Generation Using Video Games. https://eprint.iacr.org/2014/045.pdf

[34] Henno, J., Jaakkola, H., Mäkelä, J. Using games to understand and create randomness. SQAMIA2018 - Proceedings of the 7th Workshop on Software Quality Analysis, Monitoring, Improvement, and Applications, Vol. 2217, CEUR-WS, http://ceur-ws.org/Vol-2217/

[35] E. Ben-Porath. Repeated Games with Finite Automata. Journal of Economic Theory 59:1, 1993, pp. 17-32

[36] Sergio B. Volchan. What Is a Random Sequence? https://www.maa.org/sites/default/files/pdf/upload_library/22/.../Volchan46-63.pdf

[37] A Quilliot. Cyclic Orders. European Journal of Combinatorics 10:5, 1989, pp. 477-488

[38] Gang Beasts Controls Guide. https://www.gameskinny.com/ly5jv/gang-beasts-controls-guide

[39] Shai Ben-David et all 2019. Learnability can be undecidable. Nature Machine Intelligencevolume 1, pp 44-48, https://www.nature.com/articles/s42256-018-0002-3

[40] The World Rock Paper Scissors Association. https://www.wrpsa.com/

[41] USA Rock Paper Scissors League. https://myspace.com/usarps

[42] Rock Paper Scissors tournament rules. https://do317.com/p/rpsrules

[43] Rock   Paper   Scissors   Programming   Competition.
      http://www.rpscontest.com/

[44] Kolmogorov, A. N. (1965) Three Approaches to the Quantitative Definition
      of Information. Problems Inform. Transmission. 1(1), pp. 1-7

[45] Martin-Löf, P. (1966) The definition of random sequences. Information and
      Control. 9 (6): 602-619

[46] Schnorr, C. P. (1971) A unified approach to the definition of a random
      sequence. Mathematical Systems Theory. 5 (3), pp. 246-258

[47] Jaak Henno (2015) Information and Information Security. Information
      Modelling and Knowledge Bases XXVII, pp. 103-120

[48] B. A. Trakhtenbrot, Ya. M. Barzdin 1973. Finite Automata. Behavior and
      Synthesis. North-Holland, p. 211

[49] Ziv, J.; Lempel, A. (1978) Compression of individual sequences via
      variable-rate coding. IEEE Transactions on Information Theory. 24 (5): 530

[50] A. Cobham, Uniform tag sequences, Math. Systems Theory, 6 (1972), pp
      164-192

[51] H. Matsushima. Bounded Rationality in Economics: A Game Theorist's
      View. The Japanese Economic Review (1997), 48:3, pp 293-306

[52] Z. Wang, B. Xu, H. Zhou. Social cycling and conditional responses in the
      Rock-Paper-Scissors   game.   eprint   arXiv:1404.5199,
      https://ui.adsabs.harvard.edu/#abs/arXiv:1404.5199

[53] http://staff.ttu.ee/~jaak/games

[54] GA. Miller. The magical number seven, plus or minus two: Some limits on
      our capacity for processing information. Psychological Review. (1956) 63,
      pp. 81-97

An Extended TAM Model, for Evaluating eLearning Acceptance, Digital Learning and Smart Tool Usage

Andrea Tick

Óbuda University, Keleti Faculty of Business and Management, Institute of Enterprise Management, Tavaszmező u. 17, 1084 Budapest, Hungary, Tick.Andrea@kgk.uni-obuda.hu

Abstract: The learning habits among students in higher education, has radically changed in the last 20 years, partially due to the features of the information and digital society, wide scale broadband internet access, proliferation of smart devices and consequently, available online mobile applications. As a result, the use of eLearning systems, in higher education, is a must in the 21st Century. As hardware and software developments periodically foster each other's progression, the technological developments, including more and more sophisticated eLearning platforms and available mobile applications, triggers a multiplicative, radical change in educational practices and methodologies. This paper presents an extended version of the Technology Acceptance Model (xTAM), applying Structural Equation Modeling (SEM) with the AMOS program. It focuses on the motivation and usage intention of eLearning systems, among early Z generation students, in higher education and highlights the digital learning aspects and smart tool usages in the Hungarian environment. The evaluation of the above external factors illustrates the behavior of students, when using eLearning systems.

Keywords: IT security awareness; xTAM; SEM; eLearning systems; digital learning; Z generation; smart tools

1 Introduction

Modern digital technologies shift educational practices and methodologies toward collaborative, online and offline computer-supported learning. A radical teaching-learning methodological change is still underway in the higher educational practice in Hungary. Universities and higher educational institutions are revising their teaching practices and integrate more and more offline and online eLearning possibilities, furthermore, they strive to involve students' digital skills gained by internet and smart device usages. The extensive use of smart phones and mobile devices, the transition in the methodology, methods and ways of learning from a traditional form through blended learning to a digital form, as well as, the shift of

learning from offline to online platforms is unquestionable not only in case of individual learning but within organizational frameworks as well. Students as well as adults in higher education turn to a great extent to massive online educational platforms and utilize self-directed learning using their own smart and mobile devices.

The turn of the century brought about the digital information society, in which, the acceleration of economic and digital development originating from the 20th Century further changes and speeds up the life of individuals. As a result, the requirements in the job market, the lifelong learning phenomenon, the tuition fees and expenses of university studies all require students to have a job parallel to their studies as well as continuously improve their skills and gain novel knowledge to become successful in the job market. Consequently, the time spent on studying shortens, students require such forms of learning that gives the freedom of time and space and, at the same time, creates such a framework that enables more concentrated way of learning. Students find the way to use up their idle time frames, for instance time spent on travelling, queuing or waiting. The integration of smart devices as tools for eLearning boosts the process even further since smart devices are always at hand and, by now, broadband internet access on smart devices is taken for granted. According to Toffler [1] the third wave rhythm changes the concept and perception of time the X and Y generations' approach to time consumption differs. "But time itself has changed in the "real world," and along with it we have changed the ground rules that once governed us." [1] The value of time is of high importance. As a response, universities and colleges are determined to offer a wide scale of eLearning possibilities being standalone eLearning courses or in the form of blended learning, integrating these courses into existing curriculums thus transforming the structure of teaching and lecturing.

A detailed survey was conducted among university students of Óbuda University and the Budapest Business School to explore the students' eLearning usage motivation, acceptance and attitude [2, 3, 4, 5]. The research has introduced an extended Technology Acceptance model (xTAM) using the Structure Equation Modeling (SEM) with the AMOS program. The research gathered a total of more than 600 responses. After the regular data management processes, data cleansing and transformation, more than 500 questionnaires were used for evaluation [2]. After the determination of the exogenous and the endogenous factors, the ones that are highly significant, regarding the features of the digital environment, i.e. digital learning (DL), as well as, IT security awareness (IT) and smart tools (ST) were analyzed separately. In addition, in case of the questions about digital learning gender was also taken into account.

On the one hand, this paper presents the conclusive findings of the overall research made on the use and acceptance of eLearning systems among the early Z generation-born students in higher education in Hungarian environment. On the other hand, the paper explores two external factors of the model that are related to the digital environment, namely DL and ST. These factors are of crucial

importance in the transformation of the learning process, since digitalization and available internet access revolutionize the practice of studying. The evaluation of these factors reveals the behavior and attitude of students, when using eLearning systems.

2 The Extended TAM Model

The Technology Acceptance Method (TAM) originates from the 80s [6] and throughout time it evolved by separating the external variables to more specialized influencing variables. The model used in the research was compared to other TAM models applied in international and Hungarian environment in [4]. The model is still applicable for technology acceptance and it is used for eLearning usage and acceptance in the research in question. The model is justified by the Structural Equation Modeling (SEM) using the AMOS program.

2.1 Exogenous and Endogenous Variables in xTAM

The extended Technology Acceptance Model (xTAM) includes the following exogenous variables: IT security Awareness (IT), Digital Learning (DL), Smart Tools (ST), System Access (SA), eLearning Anxiety (ANX), Traditional Education (TE) and Social/Cultural Factor (SF) (Figure 1).
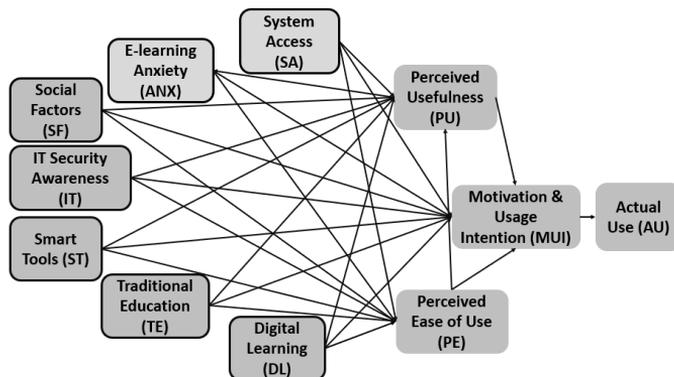
Figure 1
The extended TAM model - xTAM

The use of these external factors for the early Z generation in higher education in Hungary were justified in [2] and [3]. The endogenous variables in the model are the Perceived Usefulness (PU), Perceived Ease of Use (PE), and the Motivation and Usage Intention (MUI) as justified also in [3]. The complex research has allowed the evaluation of the existing influencing effects and the relationships of the xTAM exogenous and the endogenous variables as introduced in [3]. Due to

the criteria for model building, model fit and reliability, some of the questions were excluded from the entire model, however they were included when the factors were analyzed separately. The xTAM has not presumed strong influencing effects between the external factors, which, however, proved to be a significant phenomenon in the research. Previous results [3] showed the relationships of the input factors, however, the strength and nature of the relationships of the exogenous factors in the model seemed to justify the grouping of these factors to *digital* and *human* factors. This paper focuses on the group of digital variables (IT, ST, DL and SA).
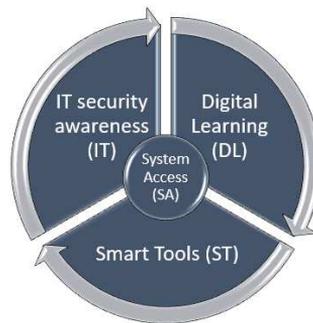
Figure 2

The relationship of the exogenous variables IT, DL, ST and SA of the digital environment

In the course of the evaluation, the exogenous variables related to the digital environment -IT, ST, DL and SA- proved to be in a relatively strong correlation with each other, therefore the analysis of their relationship is also valuable for the research (Figure 2).

3   xTAM Evaluation Results

This chapter presents the results of the evaluation of the extended TAM model that was confirmed with SEM using the AMOS program. The Structure Equation Modeling uses regression-based multivariable technique which is combined with path analysis [3]. The model shows the influencing effect of the external variables and makes direct and indirect effects visible.

3.1   Reliability

Previous studies by the author [3, 4, 5] presented the requirements and criteria for using factorization method and the TAM with the SEM model, the reliability applying the Cronbach's alpha and the Composite Reliability measures. Due to the very low reliability of Traditional Education (TE), it was excluded from the entire model. TE proved to be a poor factor, it turned out that the factor had no impact

on the model itself, which can be explained by the fact that traditional way of education seldom applies eLearning methods and ICT tools. The evaluated model with 56.52% explanatory level had Cronbach's α=0.691, KMO=0.911 and Bartlett test p=0.000. The factor reliabilities were presented in [4]. Excluding the TE slightly worsened model reliability, Cronbach's α did not change at all.

3.2   Evaluation of the Input Factors' Interactions in xTAM

In the course of model verification all the external factors themselves showed correlation with each other, some of them being positive while others negative. Table 1 summarizes the hypotheses and the conclusions about the interactions amid the external variables. The first column lists the factors; the null hypothesis supposes no interaction between the two factors in question while the alternate hypothesis supposes the existence of the interaction.

Table 1
Interaction evaluation of the exogenous factors

Analysis of external factors						
Factor 1 ↔ Factor2	C	S.E.	z test	P	r	hypothesis evaluation
SA ↔ ST	0.948	0.117	8.135	***	0.438	✓
IT ↔ SA	0.642	0.102	6.308	***	0.36	✓
DL ↔ ST	0.726	0.115	6.298	***	0.331	✓
DL ↔ IT	0.563	0.102	5.499	***	0.312	✓
IT ↔ ST	0.529	0.095	5.564	***	0.308	✓
DL ↔ SA	0.623	0.118	5.293	***[a]	0.275	✓
SF ↔ ANX	0.504	0.093	5.414	***	0.283	✓
SF ↔ SA	-0.223	0.11	-2.02	0.043	-0.1	✓
SF ↔ ST	-0.281	0.106	-2.644	0.008	-0.131	✓
ANX ↔ ST	-0.355	0.086	-4.136	***	-0.206	✓
SA ↔ ANX	-0.374	0.09	-4.17	***	-0.209	✓
SF ↔ DL	-0.593	0.119	-4.996	***	-0.262	✓
DL ↔ ANX	0.025	0.09	0.277	0.782	0.014	not significant
SF ↔ IT	-0.023	0.092	-0.248	0.804	-0.013	not significant
IT ↔ ANX	-0.068	0.073	-0.935	0.35	-0.048	not significant

[a] *** p<0.001

Taking the digital factors into account the following can be concluded from the nature of the relationships between the factors. Highlighting the influencing size of the digital factors, all these factors have positive effect on each other, the relationship between ST and SA being the strongest one. This relationship confirms the idea that developing mobile applications for eLearning possibilities on smart device platforms for students in higher education must be in focus. The second strongest relationship in this case is between SA and IT security awareness

(IT), which also justifies the hypothesis that university students are sensitive to IT security issues in case of eLearning system usage and their IT security awareness must be raised. The third strongest relationship exists between DL and ST that proves the concept of providing learning options for the so-called "idle" time that can be used up for learning. Short videos, fast and quickly solvable tasks on mobile platforms help students cover some topics while waiting, travelling or queuing. Students prefer self-study, like eLearning courses and they prefer solving extra tasks and cover extra topics if the application is available on smart tools. IT is also in positive correlation with ST, however, this effect is weaker than in case of SA. System Access questions referred to university computers that form a more robust wired environment. Students usually expect these systems be supervised by university IT staff thus they should rather concentrate on their smart devices' security problems.

In addition, the above table shows that only eLearning anxiety (ANX) has no interaction with DL and IT security awareness, it is in negative correlation with SA and ST while in positive correlation with SF. Obviously, anxiety implies forbearance from technological solutions, some lack of digital literacy, less confidence in internet usage might result in turning to personal face-to-face contact.

One can draw a parallel between eLearning Anxiety and a Technophobes attitude, who think that some technologies are not for them and tend to ignore technological trends. These type of students do not utilize the benefits of digital learning, do not improve their digital skills and, as a result, "negatively affect their own brand" in the digital society. In parallel, SA has a positive interaction with the digital factors (DL, ST, IT) and a negative interaction with personal/human aspects like Social Factor (SF).

3.3   Confirmative Factor Analysis with the AMOS Program

As a first step the factors were created with Principal Axis Extraction with Promax Rotation. The averages of the factor loadings were over 0.6 and none of them were under 0.3. There was no stronger than 0.7 correlation in the correlation matrix. The question DL4 dropped out in the Pattern matrix, but proved to be strong enough in the structure matrix, thus it was considered in the course of evaluation. No strong cross-loadings were allowed during factorization. Figure 3 shows the graph pattern of the model, including the correlation between the exogenous factors and some error terms (exclusively within a factor).

In the course of model building SEM had to be modified several times. In the first step of the fit test, the covariance matrix showed that the exogenous variables are in correlation, thus these relationships were added to the model. These relationships allowed further hypotheses to be formulated, the evaluation of most of these hypotheses are in [3]. The additional correlations in the model improved

on the model's fit test (the degree of freedom increased by 15 and CHI^2 dropped from 3118 to 2788, the difference is 330, which is significantly more than the double of 15 (the difference between the two degrees of freedom). Consequently, the addition of the correlation among the external factors is justified. However, further improvement on the model was still necessary. The factor loadings were under the expected 0.7 in some cases, however, these factor loadings were kept in the course of the original factor analysis, better factor loadings could not be achieved [4].
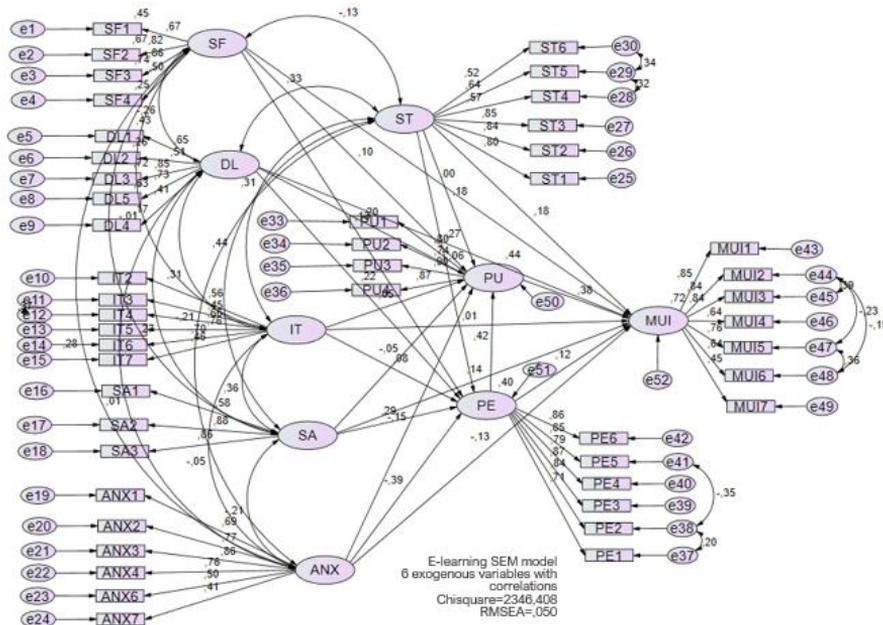
Figure 3
The xTAM model evaluation using SEM with AMOS

The model required even further modifications, which were conducted until the model fitted well to the original data. Certain error terms had to be joined in order to improve the model even more. These are all correlations between individual variances within a single factor, there were no cross-correlations. The result model met the requirement of several fit tests, CHI^2 was still significant (CHI^2=2346.408, and the degree of freedom further dropped by 9), and the other indicators showed a well fitted model. In the end, the model proved to be reliable and fitted to the actual data well. The applied fit test measures with the results and the recommended significance levels are shown in Table 2.

Table 2

Fit measures during SEM

Fit measure	Values	Recommended value
CHI2	2346.408 (p=0.000)	p>0.05
RMSEA	0.05	<0.10
CFI	0.905	>0.9
IFI	0.906	>0.9

3.4   Explanatory Levels and Significance of Exogenous Variables on PU, PE and MUI

The explanatory ratio of the Structural Equation Modeling (SEM) is given by R^2. Considering the three output factors MUI is explained by 71.7% by the factors – which is a strong explanatory ratio - while PE and PU are explained by 40.4% and 43.8%, respectively, which is a relatively strong explanatory ratio. In the two latter cases there exist some other explanatory factors, the exploration of them will be subject to a future research. Conclusively, in case of the three endogenous factors the model has a good explanatory effect.

Upon analyzing the effect size, which shows the predictive capabilities of the model to the MUI factor, it can be seen that if the values determined in [7] is taken (0.02 means small, 0.15 moderate and 0.35 strong effect size), then in almost each case there is a direct moderate or strong effect, precisely, Digital Learning (0.412), Perceived Usage (0.38) and Perceived Ease of Use (0.277) have strong positive effect size while ANX (-0.291) has a strong negative effect on MUI (Table 3). The strongest positive direct effect is PE → PU (0.42) while the strongest negative effect is ANX → PE (-0.388) (Table 3). It implies that the easier students feel the use of an eLearning system, the more often they will use it while the more frustrated they feel about the eLearning system the more cumbersome they will find its usage. The social factor (SF) has a weak direct effect on PE and PU, even the direct effect is not significant on the MUI. It implies that human factor and the cultural norm in "the third wave" [1] is less dominant in the motivation and usage intention than the digital factors.

Table 3

Effect size between the exogenous and the endogenous factors

	ST	ANX	SA	DL	SF	PE	PU
PE		-0.388	0.291	0.22	0.13		
PU	0.211	-0.31		0.293	0.154	0.42	
MUI	0.27	-0.291	0.206	0.412		0.277	0.38

Table 4 summarizes the individual null and alternate hypotheses, the null stating that there is no interaction between the factors, the alternate stating the existence of the relationship. In the model 8 factors were influencing MUI, out of which

Perceived Usefulness (0.324) boosts MUI the most, followed by DL (0.247) and ST (0.172).

In case of PE, System Access has the strongest positive influence (0.254) followed by DL (0.189). PE boosts PU positively the most (0.515), the use of DL and ST moderately, but significantly influence PU (0.211 and 0.204 respectively). ANX proved to worsen all the above three factors – MUI, PE and PU –, which can be explained by the fact that if the use of eLearning systems causes frustration, students will not find it easy to use, will not be motivated and will not have the desire to use the system.

Table 4
Hypothesis evaluation of the interaction of the factors

Evaluation of the model – the strength and significance of the interactions						
Factor1←Factor2	Coefficient	S.E.	z test	P	r	hypothesis evaluation
MUI ← PU	0.324	0.036	8.878	***[a]	0.38	✔
MUI ← DL	0.247	0.037	6.704	***	0.275	✔
MUI ← ST	0.172	0.035	4.882	***	0.183	✔
MUI ← SA	0.13	0.035	3.674	***	0.143	✔
MUI ← PE	0.122	0.045	2.706	0.007	0.117	✔
MUI ← ANX	-0.146	0.044	-3.318	***	-0.128	✔
PE ← SA	0.254	0.042	6.015	***	0.291	✔
PE ← DL	0.189	0.041	4.571	***	0.22	✔
PE ← SF	0.114	0.038	2.963	0.003	0.13	✔
PE ← ANX	-0.424	0.051	-8.321	***	-0.388	✔
PU ← PE	0.515	0.064	8.073	***	0.42	✔
PU ← DL	0.211	0.051	4.117	***	0.201	✔
PU ← ST	0.204	0.051	4.037	***	0.185	✔
PU ← SF	0.107	0.047	2.281	0.023	0.1	✔
PU ← ANX	-0.197	0.064	-3.071	0.002	-0.147	✔
MUI ← IT	0.007	0.042	0.166	0.868	0.006	not significant
MUI ← SF	0.005	0.032	0.143	0.887	0.005	not significant
PE ← IT	0.084	0.051	1.638	0.102	0.077	not significant
PE ← ST	0.057	0.042	1.357	0.175	0.063	not significant
PU ← IT	0.062	0.062	0.997	0.319	0.046	not significant
PU ← SA	-0.05	0.052	-0.962	0.336	-0.047	not significant

[a] *** p<0.001

In summary, it can be stated that the two extra digital factors in xTAM (DL and ST) positively strengthen MUI, PE and PU (except ST → PE), i.e. the more students use the digital form of learning and their smart tools, the more motivated they will be to use such courses and systems, the easier to use and more useful they find the systems. Six hypotheses cannot be supported. IT security awareness

is not in direct relation with MUI, PE and PU. IT security awareness, as proved in the previous chapter has direct relation with DL, ST and SA, the nature of the indirect relation of IT to the endogenous factors will also be explored in a future research. Presumably, IT security awareness should be indirectly related to MUI, PE and PU. At the same time neither SA nor ST influence significantly PU, furthermore, ST does not have an impact on PE. The latter might be surprising, since in the digital age it would be expected to have mobile eLearning apps developed for smart tools that boost the perceived ease of use. Supposedly, eLearning mobile apps are still not up to the expectations, challenging higher educational institutions to develop user friendly, easy access eLearning educational platforms for mobile and smart devices.

The influence of Social Factors on MUI gave a non-significant result, which strengthens the raison d'etre of blended learning, since preferring personal, face-to-face consultations and learning does not strengthen or weaken the motivation to use eLearning systems, especially amid the members of the early Z generation - a transitional generation between digital migrants and digital natives [8], who bear the characteristics of the Y generation as being technology savvy, trend followers, receptive to innovations, having social presence and falling into the cluster of Millennials as defined in [7].

Digital learning intensifies perceived usefulness, perceived ease of use as well as motivation. The early Z generation, being almost digital natives, uses e-learning systems with pleasure. According to the results the social factor – personal, face-to-face learning, learning socialization – also gives a positive impulse to eLearning system usefulness and ease, which also affirms the standing existence of blended learning. In this case the lecturer's persuasive talent could contribute to a great extent to the use of eLearning systems. The significant positive relation between SA, PU and MUI was expected as was proved in a previous study [9].

4   Aspects of Digital Factors – DL and ST

The analysis of the entire xTAM model proved that the digital factors (IT, DL, ST, SA) had a higher influencing effect on MUI, PE and PU than the human factors (ANX, SF). Furthermore, the analysis has shown that there is a direct relationship between these factors as well. Therefore, the separate analysis of the digital factors is justified. The following chapter highlights some aspects of the two digital factors, namely, Digital Learning (DL) and Smart Tools (ST) in terms of the behavior and preference of students in the Hungarian environment.

The classification of the students based on IT Security Awareness and the detailed analysis of the factor IT security awareness was conducted in [4] and [10]. The classification in [4] and [10] showed that students of the early Z generation in higher education in the Hungarian environment could be grouped into

"*Negligents*", "*IT aware*" and "*Sceptics*" requiring different type of training and IT security education. A cardinal point was for students to know whom to turn to in case of hacking or any system problems.

This chapter highlights some aspects of DL and ST, exploring gender differences and clustering in case of DL and a path preference in case of ST.

4.1 Digital Learning

The survey originally included nine questions that dealt with digital type learning (Figure 4). In the course of factorization, some of these questions dropped out because mostly the questions related to eLearning preference had factor loadings over 0.5. However, a separate analysis of the responses focusing on the first two questions in Figure 4 was worthy of consideration. In order to represent the behavior of the students the responses, given use of the 1…7 Likert scale, they were grouped into answer categories: "*non-typical*", "*indifferent*" and "*typical*". Figure 4 shows how students perceive digital learning and how they behave.
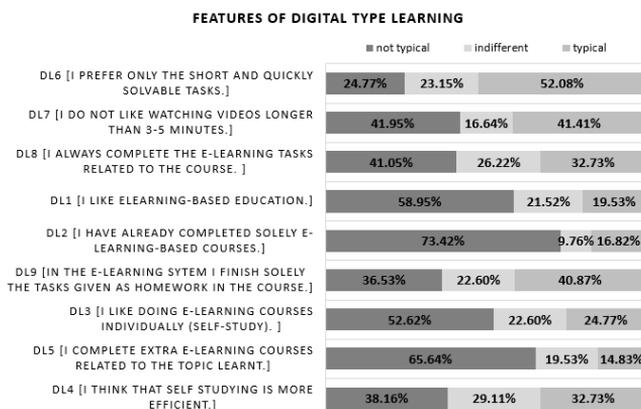
Figure 4

Student behavior regarding Digital Learning

The answers reveal that students of the early Z generation are not fully prepared for digital learning since 73% of them have not finished exclusively eLearning courses yet, 53% of them still do not like doing self-study eLearning courses, 66% of them do not complete extra eLearning courses related to the topic learnt and only 33% of them think that self-paced, individual learning is more efficient. On the other hand, students participating in the research showed that they were in favor of digital learning, they bear the characteristics of digital learners, namely they like short videos and quickly solvable tasks – the basic characteristic of MOOC (Massive Open Online Courses). 52% of them like short videos and quickly solvable tasks and half of them like short educational videos, short exercises, that they can gain with their "scattered brain". According to [11] the

members of the digital natives become shallow-brained, their brain plasticity, cognition and concentration change. Their brain becomes as fragmented as the pieces of information on the Net. The long and deep concentration on one topic ceased and students' focus skips from one bit of information to another within a few seconds. These features appear in the course of eLearning, long videos, long texts and sequential tasks cannot engross the students' attention [12]. Their way of studying and learning is continuously getting scattered, padded with hyperlinks.

4.1.1   Exploring Preferences by Gender

The correlation between DL questions determined their participation in the model. In case of DL6 there seemed to be the strongest and significant relationship (r=0.306, p=.000) with DL7, which could underlie the phenomenon of the "scattered brain" and "scattered attention" in the digital age. The analysis used DL6 as a grouping factor and DL7 as dependent because each measure showed a stronger influence in this order. Using directional and symmetric measures for concordant and discordant pairs in the data set it turns out that all the relevant measures are significant (p=0.000) and show a positive value, i.e. the pairs of answers are rather similar (Table 5).

Table 5
Directional measures of DL6 and DL7

Somer's d	Total	Male	Female	Approx. Sign.
Symmetric	.255	.235	.266	,000
DL7 *Independent*, I prefer only the short and quickly solvable tasks. *Dependent*	.248	.230	.258	.000
DL6 *Independent*, I do not like watching videos longer than 3-5 minutes. *Dependent*	.262	.241	.275	.000

The symmetric Somer's d value equals 0.255, which represents a moderate but significant relationship. The Kendall τ_b is also 0.255 meaning that there are more concordant, similar pairs than discordant ones. At the same time γ=0.306 which underlines that there is a moderate but significant relationship. In summary, the more students prefer short and quickly solvable tasks, the more they prefer short videos. With higher probability, students who prefer short and quickly solvable tasks, will not prefer longer than 3-5 minute-long videos to watch. Pearson's R showed a bit stronger relationship (r=0.324). This phenomenon raises two questions, and challenges educational practices in higher education. On the one hand, it might mean that early Z generation students' concentration on and attention to a topic for a longer time is shallow and superficial, which would hinder deep learning and completing more complex tasks. This phenomenon would support the idea stated by Carr [11] of the existence of the "scattered brain" and "scattered attention". On the other hand, it is a challenge not only for eLearning developments but also for lecturers in the course of face-to-face

teaching to maintain student attention and concentration for the whole lecture or seminar. Students need and require extra impulse, extra activities and "games" during lectures and seminars, newer and newer information impulse must be inserted. A shift to a more informal e-enabled learning environment is emerging. Gamification has become a significant trend in education, the personal ownership of small on hand smart devices with access to online learning apps require all learning-related activities at lectures and seminars to become e-enabled, boosting less tutor-led approach and creating an informal, more social atmosphere [13]. When students watch educational videos it is easy to stop it, forward and rewind it, it is easy to skip and jump on to another short video but a lecture or seminar is not hyperlinked.

The exploration of student digital learning behavior became more differentiated when splitting the responses by gender. The different behavior of male and female students became visible implying that gender could be a well separating factor. As a result, it is proved that there is a difference how male and female students turn to digital learning including e.g. MOOC. In the research 35% of the surveyed were females and 65% were males. In case of female students all the relevant association measures showed stronger relationship than without separating males and females (Somer's d_F=0.266, τ_{bF}=0.266, γ_F=0.318 and r_F=0.310). This can strengthen the trend, that with, the same preference of short and quickly solved tasks female students are getting more and more conscious about their carrier earlier than males, they try to focus more on the content and try to be more success-oriented as well as try to fill in the time spent on learning more and more efficiently. Female students are willing to spend less time on specific educational videos. However, it might mean that male students spend more time on the same topic, deep learning is more characteristic in their case. (Figure 5)
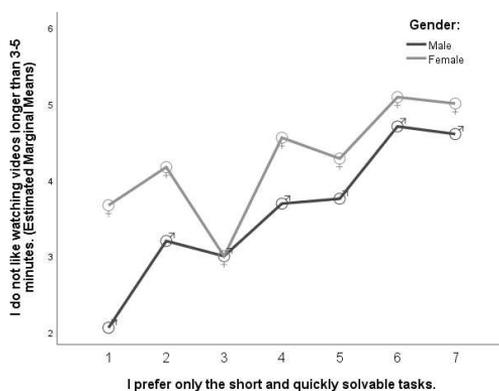
Figure 5

Estimated marginal means of DL7 in respect of DL6 separated by gender

Figure 5 plots that the effect of DL6 on DL7 depends on gender and causes a non-ordinal and non-crossing interaction. At each but level three female students'

resistance to longer educational videos is stronger than of male students. Upon Multiway Analysis of Variance with DL6 and Gender as grouping factors and DL7 as dependent the means are higher for females. Despite the fact that using DL7 as a grouping factor and DL6 as dependent gave a non-ordinal crossing interaction, which is stronger in nature, the above order was evaluated based on figures in Table 5. Despite Levene's test of Equality of Variance being non-significant, the results were evaluated since the Analysis of Variance is robust enough [14]. Table 6 shows that DL6 and gender significantly influence DL7 separately, but their interaction has no effect.

Table 6

Effects of the grouping factors on DL7

Tests of Between-Subjects Effects						
Dependent Variable: DL7 [I do not like watching videos longer than 3-5 minutes.]						
Source	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Corr. Model	297.827[a]	13	22.910	6.571	.000	.137
Intercept	5200.586	1	5200.586	1491.746	.000	.735
Gender	39.040	1	39.040	11.198	.001	.020
DL6	199.911	6	33.319	9.557	.000	.096
Gender * DL7	14.834	6	2.472	.709	.642	.008
Error	1879.084	539	3.486			
Total	10969.000	553				
Corr. Total	2176.911	552				
a. R Squared = .137 (Adjusted R Squared = .116)						

Based on partial η^2 it can be stated that there are several other factors influencing DL7 since gender and DL6 explain the behavior by 2% and 9.6% respectively and the model explains only 13.7%. Other factors could be the level of digital skills, internet knowledge, cognitive behavior, different learning strategies, level of concentration etc., or even the time slot available for the actual task or video. It might occur that the skill of multitasking is also reflected in the learning behavior by gender. Figure 5 also shows a consequent positive relationship, a continuously growing curve for males as well as for females, i.e. the quicker students aim to finish the task, the shorter videos they are willing to watch.

4.1.2   Classification of Students on Digital Learning

For the classification of students regarding digital learning, correlation between the questions were evaluated first. Using the individual questions of the survey relatively weak but significant correlations were found between the individual questions regarding digital learning except DL2 (Table 7).

Table 7
Correlation of DL questions without DL2

	DL6	DL7	DL8	DL1	DL9	DL3	DL5	DL4
DL6	1							
DL7	.306	1						
DL8	.107	.181	1					
DL1	0.080	.132	.350	1				
DL9	.166	.113	.127	.190	1			
DL3	0.026	.102	.357	.531	.157	1		
DL5	0.070	.134	.326	.430	0.063	.646	1	
DL4	.136	0.054	.163	.254	.141	.343	.291	1

The correlation in case of DL2 was non-significant with all the other questions, therefore it was excluded from the entire model. The relatively weak but significant correlations of the other variables allowed a further cluster analysis so as to determine the behavioral pattern of students of the early Z generation studying in higher education. The extreme values determined by the Mahalanobis distance were excluded from the analysis as well (the sample was reduced to 403 responses). Meanwhile in case of some non-significant individual correlations the questions were left in the cluster analysis due to the importance of the question [15].

Two clustering methods were applied (Ward method and K-Means) [10, 16] which was confirmed with discriminant analysis that suggested three separate clusters to be determined. Hierarchical and non-hierarchical methods were applied, but the K-Means cluster gave a better explanatory effect so the clusters determined by the non-hierarchical method was used for further analysis. The linear relationship among the questions proved to be significant (F-test p=0.0000). For the graphical representation of the groups and to determine the discriminant dimensions, discriminant analysis was also carried out. Due to normal distribution and homoscedasticity problems, logistic regression was also used, however, the results were very similar thus the results of the discriminant analysis is presented in the paper.

As a result, students could be clustered into three groups based on the questions on Digital Learning and the questions proved to be good separating variables. Two discriminating functions could be determined (Eigenvalue$_1$=1.893, Canonical Correlation$_1$=0.809, Variance$_1$=59.5%, Eigenvalue$_2$=1.288, Canonical Correlation$_2$=0.750, Variance$_2$=40.5%). The groups significantly differ from each other along the two dimensions (Wilks' $\lambda_{1 \text{ through} 2}$=0.151, Wilks' λ_2=0.437, sig=0.000). The first dimension (DL5, DL3, DL1 and DL4) could be determined as "*e-Self study exploration*" and the second dimension (DL7, DL8, DL6 an DL9) as "*digital minimalism*". Table 8 presents the Group Centroids, Table 9 the structure matrix with the strongest separating variable while Figure 6 displays the groups in the dimensions.

Table 8

Functions at Group Centroids of the clusters

Functions at Group Centroids		
Cluster Number of Case	Function 1	Function 2
1	-.205	-1.566
2	1.697	.665
3	-1.663	.964
Unstandardized canonical discriminant functions evaluated at group means		

Table 9

Effects of the grouping factors on DL7

Structure Matrix Functions	1	2
DL5	.698*	.307
DL3	.661*	.320
DL1	.605*	.348
DL4	.226*	.037
DL7	-.214	.922*
DL8	.268	.374*
DL6	-.108	.292*
DL9	.087	.155*
*Largest absolute correlation between each variable and any discriminant function		
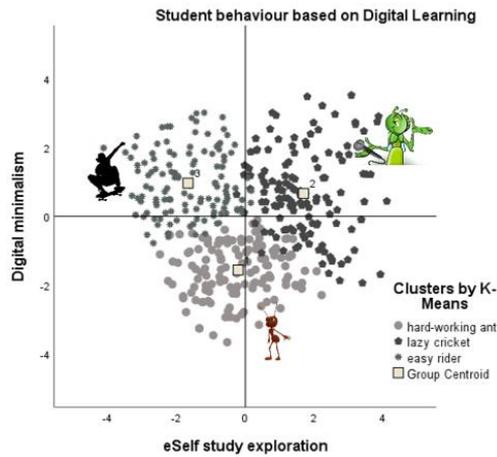
Figure 6

Student behavior based on Digital Learning

The "*hard-working ants*" are willing to explore eLearning possibilities, complete extra eLearning courses and they like more complex tasks and longer educational videos. These students gain knowledge with patience and hard work. The "*lazy crickets*" like playing around extra courses and explore eLearning possibilities but they are not patient enough and are not hard-workers enough to complete more complex tasks and watch longer than 3-5 minute videos. The "*easy riders*" are the students who would like to have all the tasks finished as quickly as possible, they are not willing to add any extra effort to their studies in the digital environment.

4.2   Smart Tools

The digital factor Smart Tools was also analyzed. Smart devices become more and more popular to use for learning processes. The results imply that eLearning system software or online course developers must focus on mobile app developments, because a significant shift can be noticed from desktops and tablets to smart phone usages in case of such courses among Hungarian students in higher education. Based on the survey responses most of the students (79%) reach the eLearning system on their laptops while almost 70% use the eLearning platforms on smart phones, and surprisingly only 17% of the students use tablets for online learning. Students could mark more than one device [2].

A considerable relationship could be detected among the questions on Smart Tools. Based on the answers on the use of eLearning systems on smart phones and devices (independent) there is a direct positive influence on students' belief on mobile learning apps (dependent) in the Hungarian environment. The positive relationship supposes well designed and developed eLearning applications. In Figure 7 the direct relationship is depicted by chords, where a chord connecting to another rating means that rating have changed from one rating to another. The darkness of the chord indicates which rating is more dominate, while if a chord connects back to itself, it shows that there was no change in the rating. It can be seen that in each case there is a better rating in favor of the dependent question.

Figure 7
Influence of 'I use eLearning type…' on "Applications on smart tools …."

Figure 8 plots the relationship by individual rating and it can be seen that in each case the higher rating is more dominant, that is if students use eLearning systems and applications on their smart phones, presumably they experience faster learning, and they actually learn faster.
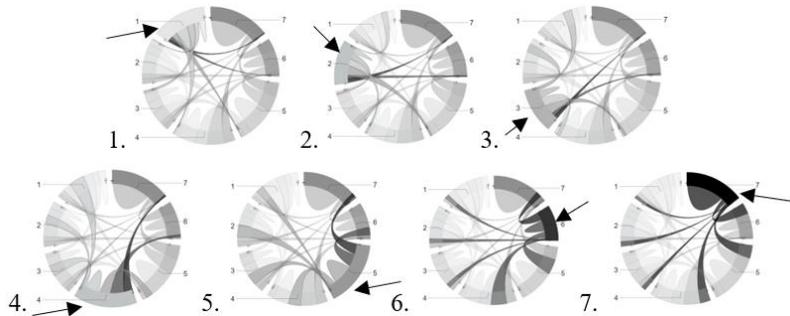
Figure 8

The direct relationship between eLearning usage on smart device and the belief of the learning speeding up with mobile applications

Consequently, mobile apps help students study more efficiently, in an individualized, flexible way, where "Third Wave, brings with it personalized, instead of universal schedules" [1]. Figure 8 also shows that there is no deterioration in the answers, in each case, the dependent variable is the more dominant, i.e. more concordant, than discordant pairs can be found.

Conclusions

This paper focused on the evaluation of an xTAM model, related to eLearning systems, with SEM using the AMOS program while highlighting the digital factors, namely IT security awareness, digital learning and smart tools including their relationship. The extended TAM model included such external variables that might have an impact on the motivation, intention to use, perceived usefulness and perceived ease of use of the eLearning systems due to a high scale digitalization and the boom of smart devices.

The newly introduced external factors, well represent the influence on the usage and motivation behavior of students in the Hungarian environment towards eLearning systems and the rate of influence of the exogenous factors on the endogenous factors in the xTAM could be well quantified.

Upon the evaluation of the xTAM, all newly introduced external factors except TE had a direct positive or negative effect on PE, PU and MUI. DL and ANX, a digital and a personal dimension, became the strongest factors. These factors proved to have a correlation with all three output factors, for instance, ANX i.e. anxiety negatively influences the above factors, so if a student is not confident, then they will prefer not use an eLearning system.

The motivation and intention to use and the perceived usefulness of eLearning systems, do not depend on IT security awareness, which might raise some concerns, provided there are IT negligent and/or IT sceptic sources.

The paper initially analyzed the relationship of the digital factors, then, separately Dl and ST, as these factors are important in the 21st Century, when digitalization is in focus, in business, industry, as well as, in education. IT security awareness as the first digital factor rather has direct influence on the other digital factors like DL, ST and SA but does not influence the personal involvement, ANX and SF. The motivation and intention to use, the perceived usefulness and the perceived ease of use of eLearning systems are all personal dimensions, thus it can be accepted that IT security awareness has no direct relation with them.

IT security awareness boosts DL and ST, meaning that students who are more IT security conscious, will use digital learning and will be more motivated to learn using smart tools. Furthermore, students who prefer DL and use ST realize they have to pay attention to IT security. The IT-ST significant correlation supports that the (online) eLearning systems are used in an IT security conscious way. DL and ST are also in positive correlation, which supports the phenomenon that digital natives – even the members of the early Z generation, who form a transitional generation as stated in [2], are always online, with their smart phones and other mobile devices, they even study with the help of these devices, instead of desktop computers.

Digital Learning also allowed the classification of "hard-working ants", "lazy crickets" and "easy riders" could be identified determined by the preference of eLearning courses and extra tasks and digital minimalism. Furthermore, two questions referred to short and quickly solvable tasks and short videos, a significant difference could be detected depending on gender. The cognitive and learning processes differ by gender, males and females react differently to the digital learning process.

Consequently, it can be stated, that more attention must be paid to IT security awareness trainings and courses for the students of the Z generation within Hungarian higher education. In the course of the elaboration of the University eLearning strategy, it must be noted that these students turn more and more towards digital learning, mainly using smart tools but, at the same time, require an occasional, personal, face-to-face contact.

In summary, digitalization and the usage of smart tools can help students be more motivated to use eLearning systems and the usage of these systems, can be easier when system access is easy and the necessary developments are carried out.

References

[1]   A. Toffler: The Third Wave, William Morrow and company Inc., New York 1980

[2]   A. Tick: "Research on the digital learning and e-learning behavior and habits of the early Z generation", Proceeding of 22nd IEEE International Conference on Intelligent Engineering Systems (INES) June 2018, pp. 33-38

[3]   A. Tick: "IT Security as a Special Awareness at the Analysis of the Digital/E-learning Acceptance Strategies of the Early Z Generation", Proceedings of 22nd IEEE International Conference on Intelligent Engineering Systems (INES) June 2018, pp. 45-50

[4]   A. Tick: "The role of IT security awareness in the TAM model of students' motivation and behavioural intention to use an e-learning environment in Hungary – A structural equation modeling approach", Hadmérnök, Vol. XIII., No. 32, September 2018, pp. 453-471

[5]   A, Tick: "The evaluation of the computer usage and learning habits of the Z generation" in Report of Meeting Researches in Didactics of Mathematics and Computer Sciences (MIDK) Vol. 16, No. 1, p. 145, January 2018

[6]   F. D. Davis: "Perceived usefulness, perceived ease of use, and user acceptance of information technology", MIS Quarterly, Abi/Inform Global, Vol. 13, No. 3, September 1989, pp. 319-339

[7]   W. W. Chin, B. Marcolun, P. R. Newsted: "A Partial Least Squares Latent Variable Modeling Approach for Measuring Interaction Effects: Results from a Monte Carlo Simulation Study and an Electronic mail Emotion/Adoption Study", Information Systems Research. Vol. 14, No. 2, June 2003, pp. 189-217

[8]   A. Kohlhofner-Derecskei, R. Zs. Reichner, Á. Szeghegyi: "The X and Y Generations' Characteristics Comparison", Acta Polytechnica Hungarica, Vol. 14, No. 8, 2017, pp. 107-125

[9]   S. Y. Park: "An analysis of the Technology Acceptance Model in understanding university students' behavioural intention to use e-learning" Educational Technology & Society, Vol. 12, No. 3, 2009, pp. 150-162

[10]   A. Tick: "Evaluating e-learning acceptance and usage motivation including IT Security Awareness amid Z generation Hungarian students with xTAM", Proceedings of IEEE 23nd International Conference on Intelligent Engineering Systems (INES) April 2019, pp. 137-142

[11]   N. Carr: The Shallows: What the Internet is Doing to our Brains, W. W. Norton & Company, Inc., New York, 2011

[12]   P. Luka: "Nicolas Carr's The Shallows. What the internet is Doing to Our Brain and some implications for Net based learning" e-learningeuropa.info, March 2011, pp. 1-4, https://www.openeducationeuropa.eu/sites/default/files/asset/the%20shallows.pdf, retrieved: 12.03.2018

[13] Á. Szeghegyi, V. Szoboszlai, J. Velencei, "Informal Post-Experiential Learning", in Acta Polytechnica Hungarica, 2014, Vol. 11, No. 4, pp. 241-252

[14] L. Sajtos, A. Mitev: SPSS kutatási és adatelemzési kézikönyv, Aliena Kiadó, 2007

[15] A, Jakus, A, Tick: "IT security risks and risk management", Hadmérnök, Vol. XII. No. 1, March 2017, pp. 182-202

[16] IBM Knowledge center, General Procedure, https://www.ibm.com/support/knowledgecenter/en/SSLVMB_22.0.0/com.i bm.spss.statistics.algorithms/alg_cluster_procedure.htm, retrieved: 15. 01. 2018

Handover Process of Autonomous Vehicles – Technology and Application Challenges

Dániel A. Drexler[1], Árpád Takács[1], Tamás D. Nagy[1], and Tamás Haidegger[1]

[1]Óbuda University, Antal Bejczy Center for Intelligent Robotics, University Research, Innovation and Service Center, Bécsi út 96/b, Budapest, H-1034 Hungary, e-mail: {daniel.drexler, arpad.takacs, tamas.daniel.nagy, tamas.haidegger}@irob.uni-obuda.hu

Abstract: Self-driving technologies introduced new challenges to the control engineering community. Autonomous vehicles with limited automation capabilities require constant human supervision, and human drivers have to be able to take back control at any time, which is called handover. This is a critical process in terms of safety, thus appropriate handover modeling is fundamental in design, simulation and education related to self-driving cars. This article reviews the literature of handover processes, situation awareness and control-oriented human driver models. It unifies the psychological and physiological control theory models to create a parameterized engineering tool to quantify the handover processes.

Keywords: autonomous vehicle safety; situation awareness; control-oriented model; takeover; hands-off control

1 Introduction

The versatile autonomous functions of vehicles require different knowledge and control approach from the users (i.e., the human driver). This can be characterized in various ways, broken down to categories from the technical point of view, e.g., Parasuraman et al. provide a well decomposed automation classification with 10 levels of automation [1]. However, the most commonly used automation level classification was created by the Society of Automotive Engineers (SAE), defining five levels of autonomy [2], which has been widely adopted, even by different domains [3, 4]:

L0 no autonomous capability;

L1 driver assistance: specific functions may be under computer control;

L2 partial automation: combined function automation (e.g., Adaptive Cruise Control (ACC));

L3 conditional automation: automation of all critical functions with limitations (limited self-driving), the driver shall be ready to take control all times;

L4 high automation: vehicle can perform all driving tasks under certain conditions; driver may take control;

L5 full automation: vehicle performs all driving tasks under all conditions; driver may not be able to take control.

The safety considerations of cars with partial and conditional automation (L2–L3) are critical, because constant attention of the driver is required due to the limited capabilities of the car; albeit, due to the relatively large portion of fundamental (and comfortable) functions being automated, the driver can easily become distracted and bored, and start to look for other, non-driving related activities. As shown by Stanton et al., this is mainly due to the fact that humans are not efficient in long inactive monitoring tasks, and drivers usually over-trust the system [5]. The problem becomes critical and potentially fatal when the automated system faces a situation that is beyond its functional capabilities, and the human driver has to take back the control from the system, when the driver is not prepared to do so [6].

The situation when the human driver takes back control from the automated system is called both *handover* and *takeover*. In Morgan at al., the term handover is used to define the process when the automated system transfers the control to the human driver, while the term takeover refers to the time instant when the driver had taken full control of the vehicle [7], which has been adopted in many papers. This terminology will be used as well. The time between the handover signal and when the human driver has full control of the vehicle is called *takeover time*. The terminology of handover is reviewed in Section 2.

The safety of autonomous vehicles below L4 is critical in real-life applications. according to Stanton et al, car manufacturers should proceed to L4, or L2 and L3 should be modified such that the driver shall always be responsible for one control input modality, e.g., for handling the steering wheel or the pedals, thus the human would be forced to pay attention during the whole driving process [5], which is a well-established protocol in aviation industry. The first suggestion (i.e., jumping to L4) is not available yet due to technical limitations, while the second suggestion means that the vehicle practically becomes an L1 system. Banks at al. analyzed the fatal Tesla crash happened May 7, 2016, using the Perceptual Cycle Model [6]. Although the investigations showed that the accident was caused by driver error, the authors suggested that "design error" was also part of the cause, which resulted in the over-boosted trust of the driver in the autonomous system. The human trust and situation awareness are critical components in the safety of L2–L3 systems, which are reviewed in Section 3. The connection of handover situations and situation awareness is analyzed in Section 4.

Human driver models and models of the closed-loop system based on a control theory (e.g., [8–10]) approach have been considered in [11]. A human model based on fractional order calculus has also been presented [12]. A recent review

of pilot models based on control theory, physiology and soft computing techniques can be found in [13]. Control and system theoretic models are useful for simulation and analysis purposes, however, they do not provide sufficient insight into the underlying phenomena. The crucial elements in the models are the time delay parts that determine the stability and performance of the closed-loop system. The control oriented models are briefly reviewed in Section 5.

Takeover times in non-critical handover situations are reviewed in [14]. Under noncritical conditions, drivers needed 1.9 to 25.7 seconds to take back control. These data were derived from measurements in non-critical scenarios, however, these takeover times are dangerously high for critical situations (i.e., when the driver has to take back control to possibly avoid an accident). The large takeover time is the main weakness of L2–L3 systems from the safety point of view. The value of the time delay can be approximated by the model of Gold et al., who created an algebraic equation based on regression to calculate the time delay based on selected data (traffic density, time before the accident, age of the driver, the current lane, the number of times the driver has faced similar situations before, and the non-driving related activity of the driver during the handover) [15]. Models for time delays in handover situations are discussed in Section 6. Based on the findings of the literature review, a human driver model is suggested in Section 7, that combines control oriented models with models of situation awareness.

2　Handover Situations

The process of handover, i.e., the process when control is shifted from autonomous to manual, can be a result of various situations; based on the conditions, there are various classifications in the literature. Here, they are considered, the first one is based on the way of handover [16], the other one is categorized by the cause of handover [17].

Based on the way of the handover, four types of handover situations are given in [16]:

- *Immediate handover*, when the control is shifted immediately, e.g., the driver grasps the steering wheel;

- *Step-wise handover*, when the control is shifted step-by-step, e.g., first longitudinal control, then lateral control;

- *Driver monitored handover*, when the driver monitors the system behavior (e.g., force feedback in steering wheel). The control is handed over after a certain period of time (e.g., there is a countdown);

- *System monitored handover*, when the system monitors the inputs of the driver for a certain period of time after the handover, and the system can adjust the inputs if it considers the driver input unsafe.

Based on the cause of the handover, five types of handover situations can be given [17]:

- *Scheduled handover*, when the driver is notified in advance of the handover situation, and has time to prepare;

- *Non-scheduled system initiated handover*, when the driver is not notified in advance, the system realizes that the driver must take control immediately because in the current situation the system would need to operate beyond its functional limits; the driver may not expect this situation;

- *Non-scheduled user initiated handover*: the driver decides to take control while there is no specific need to do so;

- *Non-scheduled user initiated emergency handover*: the user spots a potential risk that was not recognized by the system, and the user takes immediate control;

- *Non-scheduled system initiated emergency*: the system can no longer operate (the cause of this emergency is internal system failure), and notifies the driver.

The handover situations that are non-scheduled and system initiated are also called *self-deactivation* processes. An important difference between L2 and L3 systems is that an L3 system must always be able to realize if a situation is beyond the limits and initiate handover. In this paper, we are interested in immediate handover situations, i.e., the whole control is turned to manual control immediately, caused by self-deactivation, when the handover situations are non-scheduled and initiated by the system. We will also call these handover situations *immediate self-deactivation*. Important to note that handovers could possibly be initiated by cyber-security attacks as well [18].

3   Situation Awareness

Situation Awareness (SA) is used to describe the perception and the understanding of the human driver about the situation. The critical point of L2–L3 systems is when the driver loses SA. Regaining SA during handover is crucial in terms of safety, since SA is indispensable for the driver to find a solution to the problem arose during the handover situation. Thus, designing systems that help drivers regain SA is fundamental in handover management.

3.1   Defining Situation Awareness

Human perception capabilities are modeled by SA, which is a key component in handover processes. SA of the driver is the dynamic understanding of "what is going on" [19]. SA was divided to three levels by Endsley [20]:

- Level 1: perception of the elements in the environment that are relevant to the task;

- Level 2: comprehension of the meaning of these elements relative to the task;

- Level 3: projection of their future states after particular actions.

SA was formally defined as "*the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future*" [21].

Automation of SA was investigated in [22], SA with semi-autonomous agricultural vehicles was analyzed in [23], where they showed that at higher level of automation, the driver has lower SA. The authors used the Situational Awareness Rating Technique (SART) developed by Taylor, which is a self-rating post trial technique [24].

3.2 Measuring Situation Awareness

There are numerous metrics to quantify SA. Stanton at al. compared more than 30 measures of SA [25], which can be categorized into six groups [19, 26]:

1. Freeze probe techniques;

2. Real-time probe techniques;

3. Self-rating techniques;

4. Observer rating techniques;

5. Performance measures;

6. Process indices.

Freeze probe techniques are based on freezing the simulation, and asking questions from the participant right afterwards. Having answered the questions, the simulation continues. The simulation is stopped (frozen) typically randomly, and questions are asked about the tasks performed. The answers are evaluated after the simulation. A popular freeze probe technique measuring the SA along the three levels was proposed by Endsley, and is called Situation Awareness Global Assessment Technique (SAGAT) [27].

3.2.1 Real-time probe techniques

Real-time probe techniques are similar to the above with the difference that during real-time probing, the simulation is not frozen, thus they ask questions from the participants online during the simulation without stopping it. A typical real-time probe technique is the Situation Present Assessment Method (SPAM), developed for air traffic controllers' SA measurement [28].

3.2.2 Self-rating techniques

Self-rating techniques are carried out by the participants, who rate themselves typically after the trial. One such technique is the SART by Taylor [24], which uses ten dimensions to measure the participant's SA. The participant gives a score

to each dimension between 1 and 7, and the result is a subjective measure of the SA.

3.2.3   Observer rating techniques

Observer rating techniques involve experts who observe the participants during task execution, and evaluate their SA. The advantage of this method is that it does not disturb the task execution of the participants, and observer bias is reduced. A typical observer rating technique is the Situation Awareness Behavioral Rating Scale (SABARS), which has been used to asses infantry's SA during field training [29].

3.2.4   Performance measures

Performance measures provide indirect measures of SA by recording some quantities during task performance. For example, Gugerty measured crash avoidance, blocking car detection and hazard detection for driver SA [30]. Process indices involve the recording of certain functions and behaviors that are related to the SA of the participant, e.g., eye-movement is tracked in the study of Smolensky [31].

According to a thorough review that compared these measurement techniques [26], the most typically used are the SAGAT and SART to assess individual or team SA. It was found that the SAGAT technique had the most significant correlation with the task performance [19].

3.3   Losing and Regaining Situation Awareness

During automated cruising, the driver can become inattentive, and start to participate in non-driving related activities, not paying attention to the traffic. This is called Driving Without Attention Mode (DWAM), and was formalized in [32] (also known as Driving Without Awareness (DWA) [17]). In this mode, the driver behaves as a conventional passenger, which is only in line with the SA mode of L4+ cars. For cars under L4, if the driver is in DWAM, wneh a handover request occurs, then the takeover time increases dramatically.

During handover, the driver has to regain SA from DWAM. Assistant systems that help the driver to regain SA may help reducing reaction times and increase safety. In order to understand this process, it is desirable to decompose SA. Matthews et al. describe the following components of SA [33]:

- Spatial awareness: knowledge of the location of all relevant objects in the environment;

- Identity awareness: knowledge of salient items;

- Temporal awareness: knowledge of the change of location of the surroundings;

- Goal awareness: knowledge of the navigational plan, trajectory tracking, maneuvering the vehicle in traffic;

- System awareness: knowing the relevant information about the driving environment.

Regaining full SA means regaining all three SA levels. Driver assistant systems may be characterized and specialized based on the component of SA they help to regain and the level of awareness that can be reached by the assistant system. For example, the car's dashboard can help to regain system awareness, more advanced Human–Machine Interface (HMI) can increase other components of awareness.

Augmented Reality (AR) was used by Lorenz et al. to improve takeover performance of the driver, as described in Section 7 [34]. This experiment showed that an assistance system that helps regaining SA improves takeover performance.

3.4   Critical Performance Assessment

The quantitative assessment of SA, based on the level of autonomy, is crucial for the development of safe and efficient automated driving systems. Until today, there is no widely accepted metrics to quantitatively describe SA indicators, both on global and component levels. Henceforth, new autonomous features are predominantly deployed into driver assistance systems without taking into account the quantitative requirements that the human driver needs to adhere to. In order to address this issue, a systematic assessment method is proposed. Employing this method could enhance the establishment of baseline metrics, and the definition of essential performance for deployment standards.

We call for an assessment method for critical handover performance, to quantitatively define the required level and components of SA with respect to the autonomous functionalities present. To improve system safety, driver assistance systems and automated driving functionalities shall be collected and organized in a hierarchical way, along with the two criteria of SA presented, as a standardized risk assessment protocol:

- Level of SA, based on state of the environment;

- Components of SA, based on knowledge.

Fig. 1 defines SA blocks in autonomous driving, and outlines their hierarchy in accordance to the level of autonomy and SA. As the level increases, i.e., new autonomous features are added incrementally, the required number of SA components decreases for the human driver, as critical driving tasks are temporally or permanently taken over by the system. This representation is in line with the SAE definition of level of autonomy, and can be interpreted as follows:

- L2 ADAS systems require the human driver to remain in control and stay fully aware of the driving situation, possessing all levels and components of SA.

- As a transition from L2 to L3 automated systems, the driver is allowed not to fulfill all the quantitative awareness criteria to the highest level of SA, and an increasing number of components for SA are overseen by the
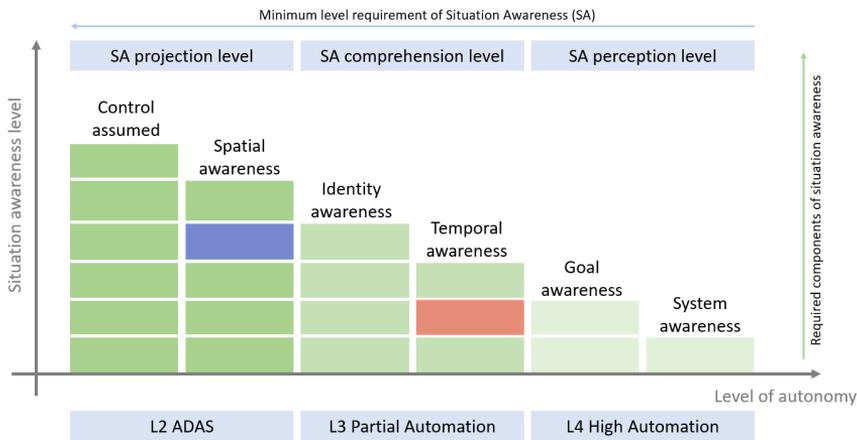
Figure 1
Hierarchical representation of SA blocks in autonomous driving. For each level of autonomy, quantitative requirements shall be defined. E.g., the block highlighted in red corresponds to the SA metrics for L3 autonomy for the comprehension of dynamic states, while the blue block represents the ability of the human driver understanding the spatial structure of the environment, while engaging an L2 driver assistance feature.

system (e.g., state of the traffic participants, expected behavior). However, some components need to stay active on the driver's side, such as handling unexpected behaviour or understanding the driving goals/trajectories.

- Transitioning from L3 to L4 automated driving, the driver is required to perceive the current state of the environment only related to his driving task. However, on the component level, *system knowledge* is interpreted as the knowledge of whether the system can solve critical driving tasks in the current driving environment, i.e., whether the user is educated about the capabilities of the used features.

Each block in Fig. 1 represents a quantitative criteria, which corresponds to the acceptance threshold for the integration of the new functionality into the system. The blocks incorporate metrics in terms of perception (object recognition distance, static and dynamic object state, road topology, actor movement probability and trajectories etc.), time factors (time to collision, takeover time, length of takeover action) and takeover ability (access to driving controls, pose of driver, environmental conditions). The measurement of these quantitative criteria is crucial, however, due to the complexity of the driving task and the human factors of the HMI, it can only be set empirically. The development of the testing framework related to this objective is part of our research, aiming to create a baseline for the definition of upcoming automotive standards.

3.5   Human Trust in Autonomous Systems

A potential safety problem of L2–L3 cars is that human drivers tend to overtrust the system, and as a consequence, they do not pay attention to critical situations [5]. On the other hand, some drivers do not trust autonomous systems at all, and thus do not want to rely on automated functions, even when those would boost their performance [35]. Human automation interaction systems and trust in automation was reviewed recently [36], where the authors pointed out the importance of trust when a human interacts with the autonomous systems. The effect of augmented SA on semi-autonomous car driving is analyzed in [37].

The way the driver treats the autonomous system and reacts to a handover situation can be considered as a problem of Human–Automation Interaction (HAI), which has a rich literature [1, 36, 38, 39]. Trust in Automation (TiA) is found to be a critical component of HAI systems, since TiA effects the decision of the human which leads to the interaction [36]. TiA is usually divided into two domains: compliance and reliance [40]. The advantage of using reliance and compliance is that they can be measured through observable behavior. The disadvantage of using only reliance and compliance is that they can not characterize TiA uniquely.

The tendency of accepting the lack of alarm or a warning is called *reliance*. If the reliance of the driver is large, then he or she believes that there is no problem as long as there is no alarm signal generated by the system, thus the autonomous system needs no supervision. If the driver has low reliance, then he or she believes that there may be errors or critical situation that are neglected by the autonomous system, thus they constantly supervise the functions. In general, the reliance of the driver should be high, however, too high reliance leads to overtrust, while too low reliance renders the autonomous functions idle. The reliance of the driver can change over time, e.g., if the system fails to generate alarms, the reliance of the driver decreases [41]. Since L2–L3 systems need constant supervision of the driver, these systems are unique in the sense that lower reliance is desirable.

The tendency of accepting and carrying out the recommendation from the autonomous system is called *compliance*. Ideally, the compliance of the driver is high, however, too high compliance means overtrust, and accepting all suggestions of the system without checking their validity. False alarms generated by the system decrease compliance, however, if the systems fails to generate an alarm, it has no effect on compliance [40].

Reliance and compliance can not completely characterize trust, since there are other factors that may affect decisions. One such factor is the workload of the driver, i.e., if the driver is kept busy, then they tend to accept the recommendations of the autonomous system, even if their compliance is low. Drnec et al. suggested to model trust as a decision process, since decision making can be objectively measured [36]. However, since decision measurement in their research is done by fMRI (functional magnetic resonance imaging), this measurement can hardly be carried out in a simulated driving environment.

Table 1
The critical SA components of non-scheduled handover situations and their effect on trust

Handover situation	Critical SA component	Effect on trust
non-scheduled system initiated	spatial awareness	reliance and compliance is increased (true positive alarms) or decreased (false positive alarms)
non-scheduled user initiated	spatial awareness	reliance is reduced
non-scheduled user initiated emergency	system awareness	reliance is reduced
non-scheduled system initiated emergency	system awareness	reliance and compliance is increased (true positive alarms) or decreased (false positive alarms)

4   Handover Situations and Situation Awareness

Handover situations are called *automation to human hands-off* in [42], where scheduled handovers are called *structured hands-off*, and non-scheduled handovers are referred to as *unstructured hands-off*. The term *takeover event* is also used to refer to a handover situation. Non-scheduled, system initiated handovers are also called *self-deactivation processes*.

Following the terminology from McCall et al. [17], we collected the non-scheduled handover types, and identified the critical SA components during handover, and the effect of the handover situation on the trust of the driver (Table 1).

4.1   Safety Critical Issues During Handover Process Management

In HAI systems, reliance is considered to be an important component, which should be kept high. However, overtrust can be fatal, since the driver fails to monitor the traffic situation, and may not be able to react in time. Moreover, if the system fails to detect the critical situation or detects the situation too late (e.g., right before the accident), then the driver has no chance to avoid that [43]. As a consequence, for L2–L3 systems, lower reliance is more desirable. Although low reliance implies that the driver has to monitor the system frequently, which is considered to be infeasible for HAI systems, this frequent monitoring is desirable for L2–L3 systems. Based on Table 1, reliance is decreased by non-scheduled user initiated handovers or false positive system initiated alarms. The latter also decreases compliance.

A critical component of handover management systems is the detection system that initiates handover. This system must be able to predict the critical situation as soon as possible, in order to alert the driver in time. If the system fails to alarm the driver in time, and the driver does not pay attention (due to high reliance), the consequences can be fatal. However, detection systems are not perfect, and can make mistakes [44]. Typical question in design is whether false positive or false negative alarms are less desirable. In handover situations, false negative alarms

can be fatal if the driver has large reliance, while false positive alarms decrease reliance as shown in Table 1. Overall, the detection system must be created such that false negative alarms are minimized, while the amount of false positive alarms can be larger.

Too much false positive alarms can lead to significant drop of reliance and compliance, which is good for safety, since it forces the driver to pay attention constantly, however, it is bad for the technology, since drivers will be wary of these systems. In Autonomous Emergency Braking (AEB) systems, false positive detection is avoided by removing stationary objects from radar sensor data, and by treating an object as an obstacle only if it is in the way of the vehicle, which is calculated based on the steering angle [44]. The performance of detection systems will likely improve in the future due to the improvement in artificial intelligence algorithms, like deep neural networks [45] and their training algorithms [46].

Using augmented/virtual reality and advanced HMI can help to improve the performance of the drivers during handover by increasing the SA of the driver, and helping to regain the SA. However, this will only work if the driver trusts the system, and believes that the information given by the HMI is valid, i.e., the driver has high compliance. False positive alarms decrease compliance, and as a result, the trust of the drivers will decrease, and the performance increase due to the advanced HMI may deteriorate as well. To the authors best knowledge, other factors, such as the behavior of drivers when the information of the HMI is not valid has not been researched yet.

5   Control-oriented Driver Models

Control-oriented driver models date back to the '70s. In the work of Kleinman et al., the control-oriented model of the human driver system described human behaviour as a time delay, an equalizer block and a neuro-motor dynamics block, shown in Fig. 2 [47]. The equalizer block contains an observer to estimate the states of the vehicle, and an inverse dynamics block for state estimation. Kleinman and Curry also used a control-oriented approach to predict human operator's performance [48].

Human decision making is modeled as a process based on probabilities in [49, 50]. Gai and Curry modeled human decision making using switches and time delays [51]. Limits of human path tracking capabilities were explored in [52].

Eskandari et al. used a control-oriented framework to model the system under shared control, i.e., the control system with an automated system and the human operator are both presented in the loop [53], shown in Fig. 3. SA is present in the human operator model, along with decision making and acting. The authors modeled SA and regaining SA using dynamical systems in [54]. This model unified the control-oriented approach with the psychological approach characterized by SA [33].

Control-oriented driver modeling was used by Wang et al. to create a control

law for a steering system [55]. Human models were used to evaluate system reliability using simulations in [56].

Driving state recognition is an important component of future autonomous cars. Machine learning was used to learn personalized driving state employing on-board sensor measurements in [57]. Clustering-aided regression is used to predict the driver workload in [58]. Mental workload dynamics was modeled in [59], where linear identification techniques are used to identify the nonlinear model on-line and show robust performance. Workload adaptive cruise control was created in [60], where the adaptive cruise control system is adapted to the current work-load of the human driver in order to tailor the level of assistance to the needs of the driver. Tests in driving simulators showed that this workload adaptive cruise control enables safer driving experience.

6   Critical Components of a Handover Process

Human attention diversion is a critical issue in driving, many studies showed that mental workload has critical effect on the safety of driving [59, 60]. Nevertheless, the study of Gold et al. showed that traffic density has a major effect on takeover performance, while answering questionnaires during the driving process was found to have no significant effect [61]. Identifying large traffic density as a potential danger source in takeover performance leads to the conclusion that for systems under L4 automation, the driver should always pay attention when the traffic is heavy, e.g., by turning automated cruising off. This should not mean that the automated cruising shall be turned off in traffic situations with large density but low velocity, (i.e., traffic jams), which could be safely managed by autonomous vehicles under L4. A possible solution for this situation takes velocity information into account, which can be easily incorporated via on-board sensors. This way, automated cruising can be allowed in large traffic density with low velocity, and remain inaccessible with large traffic density and high velocity.

The U.S. National Highway Traffic Safety Administration (NHTSA) released an

Figure 2
The human driver block, modelled fot the control theory aspect by Kleinman et al., neglecting the noises and disturbances [47].

Figure 3
The block of the closed-loop system under shared control by Eskandari et al. [53].

updated policy *A Vision for Safety* in 2017 [62]: it encourages regularization enti-
ties on the definition and documentation of Operational Design Domains (ODD)
for each automated driving system of the vehicle. An ODD should describe spe-
cific conditions under which the given features are intended to function for au-
tomated vehicles. The minimal information required for the definition of ODD
for a given functionality includes roadway type, geographic area, speed range
and environmental conditions. Pre-defined ODDs could aid the assessment of
the required level of SA in the case of automated systems under L4.

6.1 Time Delay

Time delays are critical components of takeover performance. The takeover time
during highway cruising is modeled by a polynomial in [15] which depends on
the *time budget*, defined as the time between the takeover time and the system
limit (the latest time instant when the driver must take control), the traffic density
measured in cars/kilometer, the lane (right, middle or left), non-driving related
task, repetition (the number of times the driver has faced similar situations be-
fore) and the age of the driver. The *t* takeover time is given as:

$$
\begin{aligned}
t \;=\; & 2.068 + 0.329 TimeBudget - 0.147(Lane - 1.936)^2 \\
& -0.0056(TrafficDensity - 15.667) \\
& -0.571 ln(Repetition) \\
& +2.121 \cdot 10^{-4}(Age - 46.245)^2.
\end{aligned}
\tag{1}
$$

This model implies that traffic density decreases takeover time, and has the least
decreasing effect for medium traffic density, and largest effect for small and large
traffic density. The non-driving related task had no effect, similarly to the study
carried out by Gold et al. [61]. However, it should be emphasized that the same
20-question-long form was used in both experiments. The age and lane did not
affect the results significantly, but the repetition (which is related to the expe-

Figure 4

The model of the human driver included closed-loop control system. The driver block is divided into
3 levels based on SA, representing different decision and action blocks accordingly.

rience of the driver), the time budget (which is related to how early the system
warns the driver) and the traffic density did.

6.2 Transient Quality

Improvement of takeover performance can be achieved through improving tran-
sient quality. Workload-adaptive cruise control does not necessarily reduce reac-
tion time, but it contributes to the improvement of transient quality, e.g., partici-
pants started to break at the same time but the deceleration was lower, as reported
by Hajek et al. [60].

Hence, SA also has an effect on the dynamics of the human model, along with
the time delay. This effect can be incorporated into the human model through the
neuromuscular level, i.e., different transfer functions describing the neuromus-
cular system for different stress levels. As the stress level increases, the settling
time of the transfer function decreases, but other quality factors, such as damping
are most likely to decrease as well.

Creating appropriate warning systems and prediction algorithms do not neces-
sarily improve takeover performance by improving the takeover time, but by im-
proving the reaction quality. This can be modeled through the dynamics of the
human driver, and not the time delay. The importance of this observation lies in
that most of the literature focuses on the time delay effect, and neglects the effect
of dynamics. To incorporate these effects in the model, a combined approach is
presented in the next Section, which is the main contribution of this paper.

7   Human Driver Model with SA

A new model is proposed by combining the model of the classical control theory block diagram of Kleinman et al. [47] with the SA-based block diagram of Es-kandari et al. [53], as shown in Fig. 4. The vehicle block contains the controller block, being responsible for the automation, intelligence of the vehicle, actuators, vehicle model, sensors and finally the handover management block, which, in the trivial case, can be a system that overwrites the decision of the automation with the input signals generated by the human driver.

The human driver block is composed of three levels:

- The first level (Level 1 SA) is comprised of perception, decision and action;

- The second level (Level 2 SA) is responsible for the comprehension of the perceived signal and the corresponding decision and action;

- The third and largest level (Level 3 SA) projects the perceived information on the future, and carries out the corresponding decision and action.

The level of the driver's behavior is specified by the time available for the driver (the time budget by the terminology of Gold et al. [15]). If the time for decision and acting is low, only Level 1 SA is attained, and the driver will use the decision and action corresponding to Level 1 SA. If there is plenty of time, the driver can attain Level 3 SA, and act according to this level, i.e., use the Level 3 decision and action.

The action block contains the neuro-muscular dynamics and the inverse dynamics of the vehicle. The inverse dynamics is the same for all levels, since this block depends on the driver's knowledge of the car dynamics. Note that this statement does not hold if the car is in an extreme situation with unknown dynamics to the driver (e.g., the car slips on ice). The inverse dynamics here is not related to repetition in the model of Gold et al. [15] in (1), since the repetition refers to how many times the driver has faced the critical situation before, and not the knowledge of the car dynamics. While, the possibility of correlation is not excluded, it is not discussed in this work.

The neuro-muscular dynamics can be modeled with the transfer function [13]:

$$W_{\mathrm{NM}} = \frac{e^{-s\tau_{\mathrm{NM}}}}{s^2 T^2 + 2\xi T s + 1},\tag{2}$$

with time constant T, damping coefficient ξ and time delay τ_{NM}. As the level of SA increases, the damping ξ increases, and the time constant T decreases. This way, the quality of the transient improves, as it has been observed [60]. From control theory point of view, decreased time constant would mean decrease in the performance, however, in the current application, decreased time constant results in decreased absolute value of the acceleration. This gives larger comfort to the passenger. This decrease in the acceleration is considered beneficial as long as the value of acceleration is large enough to avoid a possible accident, while it may present some discomfort to the driver and the passengers.

The various levels of SA (perception, comprehension and projection) can be modeled with different time delays with transfer functions:

$$W_{SA} = e^{-s\tau_{\text{SA}}}. \tag{3}$$

As the level of SA increases, so does the time delay τ_{SA}. The modeling of the time delay in the decision block is straightforward.

The model in Fig. 4 gives insight into the process of driver assistance system from a different perspective. For example, Lorenz et al. showed in their study that using augmented reality improves takeover performance [34]. If a green corridor was projected on the path that could be used to avoid the accident, drivers tended to steer the vehicle into that direction, while in the case red corridor was projected onto the path that should have been avoided, the drivers started to brake intensively. This phenomenon could be explained by the decrease in time delays, as shown in [63]. The model presented in Fig. 4 can be used as an explanation, as the augmented reality helps the drivers to attain higher level of SA in a shorter time. Drivers can achieve comprehension through the presented solution (but this comprehension is highly affected by the information shown by the augmented reality), and thus they can achieve Level 2 behavior sooner. This observation can aid the development advanced systems that would improve the safety of autonomous cars.

Conclusions

A complete literature review was provided about the handover processes of autonomous cars. Various terminology can be found in the literature related to handover process, we built on the most common and clarified terms. SA was identified as a fundamental human driver related component in handover situations. We provided a short review about the quantification methods of SA, and established the relationship between SA and handover processes.

Control-oriented human driver modes were reviewed, and the models were extended to incorporate the model of SA. Control-oriented driver models are important to carry out simulations and to specify quantitative measures for human driver performance. Incorporating SA into control-oriented models enforces the fusion of physiological and psychological human models, which have greater modeling power and could enhance the developments aimed at improving handover performance. Out future plan is to build a complete simulator with this knowledge in order to asses SA more efficiently.

Acknowledgment

References

[1] R. Parasuraman, T. B. Sheridan, and C. D. Wickens. A model for types and levels of human interaction with automation. *IEEE Trans. on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 30(3):286–297, May 2000.

[2] J3016b: Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles. Technical report, Society of Automotive Engineers, 2016.

[3] T. Haidegger. Autonomy for surgical robots: Concepts and paradigms. *IEEE Trans. on Medical Robotics and Bionics*, 1(2):65–76, 2019.

[4] A. Takács, D. A. Drexler, P. Galambos, I. J. Rudas, and T. Haidegger. Assessment and standardization of autonomous vehicles. In *Proc. of the 22nd Intl. Conf. on Intelligent Engineering Systems (IEEE INES)*, pages 185–192, 2018.

[5] V. A. Banks, A. Eriksson, J. O'Donoghue, and N. A. Stanton. Is partially automated driving a bad idea? Observations from an on-road study. *Applied Ergonomics*, 68:138–145, 2018.

[6] V. A. Banks, K. L. Plant, and N. A. Stanton. Driver error or designer error: Using the Perceptual Cycle Model to explore the circumstances surrounding the fatal Tesla crash on 7th May 2016. *Safety Science*, 108:278–285, 2018.

[7] P. Morgan, C. Alford, and G. Parkhurst. Handover issues in autonomous driving: A literature review. Technical report, University of the West of England, Bristol, 2016.

[8] J. K. Tar, J. F. Bitó, and I. J. Rudas. Contradiction resolution in the adaptive control of underactuated mechanical systems evading the framework of optimal controllers. *Acta Polytechnica Hungarica*, 13(1):97–121, 2016.

[9] D. A. Drexler. Closed-loop inverse kinematics algorithm with implicit numerical integration. *Acta Polytechnica Hungarica*, 14(1):147–161, 2017.

[10] V. C. da Silva Campos, L. M. S. Vianna, and M. F. Braga. A tensor product model transformation approach to the discretization of uncertain linear systems. *Acta Polytechnica Hungarica*, 15(3):31–43, 2018.

[11] S. L. W. Kleinman, D.; Baron. A control theoretic approach to manned-vehicle systems analysis. *IEEE Trans. on Automatic Control*, 16:824–832, 1971.

[12] Y. L. Z. Huang, Jiacai; Chen. Human operator modeling based on fractional order calculus in the manual control system with second-order controlled element. In *Proc. of the 27th Chinese Control and Decision Conference (CCDC)*, pages 4902–4906, 2015.

[13] S. Xu, W. Tan, A. V. Efremov, L. Sun, and X. Qu. Review of control models for human pilot behavior. *Annual Reviews in Control*, 44:274–291, 2017.

[14] A. Eriksson and N. A. Stanton. Takeover time in highly automated vehicles: Noncritical transitions to and from manual control. *Human factors*, 59(4):689–705, 2017.

[15] C. Gold, R. Happee, and K. Bengler. Modeling take-over performance in level 3 conditionally automated vehicles. *Accident Analysis & Prevention*, 116:3–13, 2018.

[16] M. Walch, K. Lange, M. Baumann, and M. Weber. Autonomous driving: Investigating the feasibility of car-driver handover assistance. In *Proc. of the 7th Intl. Conf. on Automotive User Interfaces and Interactive Vehicular Applications*, pages 11–18, New York, 2015. ACM.

[17] R. McCall, F. McGee, A. Meschtscherjakov, N. Louveton, and T. Engel. Towards a taxonomy of autonomous vehicle handover situations. In *Proc. of the 8th Intl. Conf. on Automotive User Interfaces and Interactive Vehicular Applications*, pages 193–200, New York, 2016. ACM.

[18] J. Contreras-Castillo, S. Zeadally, and J. A. Guerrero-Ibañez. Internet of vehicles: Architecture, protocols, and security. *IEEE internet of things Journal*, 5(5):3701–3709, 2017.

[19] P. M. Salmon, N. A. Stanton, G. H. Walker, D. Jenkins, D. Ladva, L. Rafferty, and M. Young. Measuring situation awareness in complex systems: Comparison of measures study. *International Journal of Industrial Ergonomics*, 39(3):490–500, 2009.

[20] M. R. Endsley. Situation awareness global assessment technique (sagat). In *Proc. of the IEEE 1988 National Aerospace and Electronics Conference*, volume 3, pages 789–795, May 1988.

[21] M. R. Endsley. Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37(1):32–64, 1995.

[22] N. Naikal. Towards autonomous situation awareness. Technical Report UCB/EECS-2014-124, Electrical Engineering and Computer Sciences, University of California at Berkeley, May 2014.

[23] B. Bashiri and D. D. Mann. Automation and the situation awareness of drivers in agricultural semi-autonomous vehicles. *Biosystems Engineering*, 124:8–15, 2014.

[24] R. M. Taylor. Situational awareness rating technique (sart): The development of a tool for aircrew systems design. In E. Salas, editor, *Situational Awareness*, chapter 6, page 18. Taylor & Francis Groups, 1990.

[25] N. A. Stanton, P. M. Salmon, L. A. Rafferty, G. H. Walker, C. Baber, and D. P. Jenkins. *Human Factors Methods: A Practical Guide for Engineering and Design*. CRC Press, 2005.

[26] P. Salmon, N. Stanton, G. Walker, and D. Green. Situation awareness measurement: A review of applicability for c4i environments. *Applied Ergonomics*, 37(2):225–238, 2006.

[27] M. R. Endsley. Measurement of situation awareness in dynamic systems. *Human Factors*, 37(1):65–84, 1995.

[28] F. T. Durso, C. A. Hackworth, T. R. Truitt, J. Crutchfield, D. Nikolic, and C. A. Manning. Situation awareness as a predictor of performance for en route air traffic controllers. *Air Traffic Control Quarterly*, 6, 1998.

[29] M. D. Matthews and S. A. Beal. Assessing situation awareness in field training exercises. Technical Report Research Report 1795, U.S. Army Research Institute for the Behavioral and Social Sciences, September 2002.

[30] L. J. Gugerty. Situation awareness during driving: Explicit and implicit knowledge in dynamic spatial memory. *Journal of Experimental Psychology: Applied*, 3, 1997.

[31] M. Smolensky. Toward the physiological measurement of situation awareness: the case for eye movement measurements. In *Proc. of the Human Factors and Ergonomics Society 37th Annual Meeting. Human Factors and Ergonomics Society*, 1993.

[32] J. S. Kerr. Driving without attention mode (DWAM): A formalisation of inattentive states in driving. In A. G. Gale, editor, *Vision in Vehicles III*, pages 473–479. 1991.

[33] M. L. Matthews, D. J. Bryant, R. D. G. Webb, and J. L. Harbluk. Model for situation awareness and driving: Application to analysis and research for intelligent transportation systems. *Transportation Research Record*, 1779(1):26–32, 2001.

[34] L. Lorenz, P. Kerschbaum, and J. Schumann. Designing take over scenarios for automated driving: How does augmented reality support the driver to get back into the loop? *Proc. of the Human Factors and Ergonomics Society Annual Meeting*, 58, 2014.

[35] B. M. MUIR and N. MORAY. Trust in automation. part ii. experimental studies of trust and human intervention in a process control simulation. *Ergonomics*, 39:492–460, 1996.

[36] K. Drnec, A. R. Marathe, J. R. Lukos, and J. S. Metcalfe. From trust in automation to decision neuroscience: Applying cognitive neuroscience methods to understand and improve interaction decisions involved in human automation interaction. *Frontiers in Human Neuroscience*, 10:1–14, 2016.

[37] L. Petersen, D. Tilbury, L. Robert, and X. J. Yang. Effects of augmented situational awareness on driver trust in semi-autonomous vehicle operation. In *Proc. of the 2017 NDIA GROUND VEHICLE SYSTEMS ENGINEERING AND TECHNOLOGY SYMPOSIUM*. 2017.

[38] R. Parasuraman and D. H. Manzey. Complacency and bias in human use of automation: An attentional integration. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 52(3):381–410, jun 2010.

[39] R. Parasuraman. Designing automation for human use: empirical studies and quantitative models. *Ergonomics*, 43(7):931–951, jul 2000.

[40] J. Meyer, R. Wiczorek, and T. Günzler. Measures of reliance and compliance in aided visual scanning. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 56(5):840–849, nov 2013.

[41] K. Geels-Blair, S. Rice, and J. Schwark. Using system-wide trust theory to reveal the contagion effects of automation false alarms and misses on compliance and reliance in a simulated aviation task. *The International Journal of Aviation Psychology*, 23(3):245–266, jul 2013.

[42] M. Blommer, R. Curry, R. Swaminathan, L. Tijerina, W. Talamonti, and D. Kochhar. Driver brake vs. steer response to sudden forward collision scenario in manual and automated driving modes. *Transportation Research Part F: Traffic Psychology and Behaviour*, 45:93–101, 2017.

[43] Á. Takács, D. A. Drexler, P. Galambos, I. J. Rudas, and T. Haidegger. Assessment and standardization of autonomous vehicles. In *2018 IEEE*

22nd Intl. Conf. on Intelligent Engineering Systems (INES), pages 185–192, 2018.

[44] A. Takács, D. A. Drexler, P. Galambos, I. Rudas, and T. Haidegger. The transition of L2−L3 autonomy through euro NCAP highway assist scenarios. In *Proc. of the 2019 IEEE 17th Intl. Symp. on Applied Machine Intelligence and Informatics*, pages 117–122, 2019.

[45] Z. Fazekas, G. Balázs, and P. Gáspár. ANN-based classification of urban road environments from traffic sign and crossroad data. *Acta Polytechnica Hungarica*, 15(8):29–53, 2018.

[46] A. I. Károly, R. Fullér, and P. Galambos. Unsupervised clustering for deep learning: A tutorial survey. *Acta Polytechnica Hungarica*, 15(8):29–53, 2018.

[47] D. Kleinman, S. Baron, and W. Levison. A control theoretic approach to manned-vehicle systems analysis. *IEEE Trans. on Automatic Control*, 16(6):824–832, December 1971.

[48] D. L. Kleinman and R. E. Curry. Some New Control Theoretic Models for Human Operator Display Monitoring. *IEEE Trans. on Systems, Man, and Cybernetics*, 7(11):778–784, November 1977.

[49] W. B. Rouse. A Theory of Human Decisionmaking in Stochastic Estimation Tasks. *IEEE Trans. on Systems, Man, and Cybernetics*, 7(4):274–283, April 1977.

[50] J. S. Greenstein and W. B. Rouse. A Model of Human Decisionmaking in Multiple Process Monitoring Situations. *IEEE Trans. on Systems, Man, and Cybernetics*, 12(2):182–193, March 1982.

[51] E. G. Gai and R. E. Curry. A Model of the Human Observer in Failure Detection Tasks. *IEEE Trans. on Systems, Man, and Cybernetics*, SMC-6(2):85–94, February 1976.

[52] D. W. Repperger, S. L. Ward, E. J. Hartzell, B. C. Glass, and W. C. Summers. An Algorithm to Ascertain Critical Regions of Human Tracking Ability. *IEEE Trans. on Systems, Man, and Cybernetics*, 9(4):183–196, April 1979.

[53] N. Eskandari, G. A. Dumont, and Z. J. Wang. Delay-incorporating observability and predictability analysis of safety-critical continuous-time systems. *IET Control Theory Applications*, 9(11):1692–1699, 2015.

[54] N. Eskandari, G. A. Dumont, and Z. J. Wang. An Observer/Predictor-Based Model of the User for Attaining Situation Awareness. *IEEE Trans. on Human-Machine Systems*, 46(2):279–290, April 2016.

[55] W. Wang, J. Xi, C. Liu, and X. Li. Human-Centered Feed-Forward Control of a Vehicle Steering System Based on a Driver's Path-Following Characteristics. *IEEE Trans. on Intelligent Transportation Systems*, 18(6):1440–1453, June 2017.

[56] S. B. Bortolami, K. R. Duda, and N. K. Borer. Markov analysis of human-in-the-loop system performance. In *2010 IEEE Aerospace Conference*, pages 1–9, March 2010.

[57] D. Yi, J. Su, C. Liu, and W. Chen. Personalized Driver Workload Inference by Learning From Vehicle Related Measurements. *IEEE Trans. on Systems, Man, and Cybernetics: Systems*, 49(1):159–168, January 2019.

[58] D. Yi, J. Su, C. Liu, and W. Chen. New Driver Workload Prediction Using Clustering-Aided Approaches. *IEEE Trans. on Systems, Man, and Cybernetics: Systems*, 49(1):64–70, January 2019.

[59] W. B. Rouse, S. L. Edwards, and J. M. Hammer. Modeling the dynamics of mental workload and human performance in complex systems. *IEEE Trans. on Systems, Man, and Cybernetics*, 23(6):1662–1671, November 1993.

[60] W. Hajek, I. Gaponova, K. H. Fleischer, and J. Krems. Workload-adaptive cruise control – A new generation of advanced driver assistance systems. *Transportation Research Part F: Traffic Psychology and Behaviour*, 20:108–120, September 2013.

[61] C. Gold, M. Korber, D. Lechner, and K. Bengler. Taking Over Control From Highly Automated Vehicles in Complex Traffic Situations: The Role of Traffic Density. *Human Factors*, 58(4):642–652, 2016.

[62] Automated driving systems 2.0: A vision for safety, October 2017.

[63] D. A. Drexler, A. Takács, P. Galambos, I. J. Rudas, and T. Haidegger. Handover process models of autonomous cars up to level 3 autonomy. In *Proc. of the 18th IEEE Intl. Symp. on Computational Intelligence and Informatics*, pages 307–312, 2018.