



Conference proceedings
Sborník odborného semináře

**Nové metody a postupy v oblasti přístrojové
techniky, automatického řízení a informatiky 2020**
**New Methods and Practices in the Instrumentation,
Automatic Control and Informatics 2020**

14. 9. – 16. 9. 2020, Zámek Lobeč

ISBN 978-80-01-06776-5

OBSAH

<i>Adrian Saldanha, Milan Hofreiter</i> RELAY FEEDBACK IDENTIFICATION USING GUIDING EVOLUTIONARY ALGORITHM	3
<i>Jakub Vaněk, Jan Slabý</i> COMPUTING OF PID PARAMETERS USING TECOMAT FOXTROT AND THE ASYMMETRICAL RELAY SHIFTING.....	24
<i>Jarno Mastomäki, Jan Hošek</i> FINITE ELEMENT ANALYSIS OF AN ENDOSCOPE DURING OPHTHALMIC SURGERY	33
<i>Prem Wongsagoon, Matous Cejnek</i> DISCRETE SIR MODEL FOR SIMULATION OF VARIOUS COVID-19 SCENARIOS	41
<i>Yang Hong-Bin, Matous Cejnek</i> VIDEO_FLAWER: REALISTIC VIDEO DISTURBER FOR ARTIFICIAL VIDEO	44
<i>David Polák, Cyril Oswald</i> ENERGY CONSUMPTION METERING IN PRODUCTION PLANT.....	48
<i>Jana Koubová, Marie Martinásková</i> CODESYS FOR BASIC FESTO MOTION TERMINAL (VTEM) APPLICATION.....	53
<i>Ondrej Baumrt</i> HMI S BIOLOGICKOU ZPETNOU VAZBOU	71
<i>Petr Písarík, Matouš Cejnek</i> KAMEROVÝ SYSTÉM S VLASTNÍ DIAGNOSTIKOU	76
<i>Jan Mareš</i> DETEKCE FYZICKÝCH OBJEKTŮ NA DOTYKOVÉM PANELU	81
<i>Daniel Šebík</i> MOBILNÍ APLIKACE PRO MONITOROVÁNÍ PRŮMYSLOVÝCH PROCESŮ	85
<i>Petr Zvolánek, Šárka Němcová, Daniel Hadraba, David Habart, David Vondrášek, Josef Zicha</i> OPTICKÉ VLASTNOSTI LANGERHANSOVÝCH OSTRŮVKŮ	90
<i>Vladimír Hlaváč</i> UKÁZKA ROZPOZNÁVÁNÍ ŘEČI S VYUŽITÍM KLASIFIKACE NEURONOVOU SÍTÍ.....	94
<i>Michal Kuchař</i> NEURONOVÉ SÍŤE V DETEKCI PORUCH	103
<i>Karel Vošahlík</i> ROBOTIKA V REHABILITACI, OPTIMALIZACE POHONNÝCH JEDNOTEK.....	109
<i>Pavel Blažek</i> AKTIVNÍ MAGNETICKÉ LOŽISKO	114
<i>Marek Ciklamini</i> GRAPH NEURAL NETWORK PREPROCESSING FOR PURPOSE OF DIGITAL TWIN OF MECHANICAL SYSTEM.....	118
<i>David Kovanda, Jan Soukal</i> TOLERANČNÍ ANALÝZA SVĚTELNÉ TECHNIKY V AUTOMOTIVE.....	121
<i>Ahmad Atiyah, Bohumil Šulc</i> NEW MODEL OF ASYNCHRONOUS MOTOR BASED ON NONLINEAR MODEL PREDICTIVE CONTROL IN RAILWAY DRIVE APPLICATION.....	126
<i>Peichl Adam</i> MODELING CORRECTION FACTOR FOR METAL SHEET ROLLING VIA POLYNOMIAL FUZZY SYSTEM.....	131
<i>Martin Procházka</i> PASIVNÍ MAGNETICKÉ LOŽISKO	133
<i>Youssef Alsoufi</i> INTELLIGENT SUSPENSION SYSTEM WITH TWO-DEGREE-OF-FREEDOM HYBRID MASS DAMPERS.....	136

RELAY FEEDBACK IDENTIFICATION USING GUIDING EVOLUTIONARY ALGORITHM

Adrian Saldanha, Milan Hofreiter

Czech Technical University in Prague, Faculty of Mechanical Engineering

Abstract: The Guiding Evolutionary Algorithm (GEA) was proposed by Cao, Xu and Goodman in 2016. The original algorithm was designed to apply the advantages of the previously published Particle Swarm Optimization (PSO), Genetic Algorithm (GA) and the Bat Algorithm (BA). As part of this work, the effectiveness of the GEA algorithm in solving multi-modal optimization problems is analyzed and the results are compared to those obtained using PSO and BA. Once the algorithm is verified and accepted, the same is then applied to a System Identification task using a relay-based feedback method and finally, upon successful identification of a dynamic system, the system is controlled using a PID controller. The work investigates a non-linear approach to identification of dynamic systems using simple Second Order Plus Time-Delay Models (SOPTD).

Keywords: Guiding Evolutionary Algorithm, Relay-feedback identification, Particle Swarm Optimization, Bat Algorithm, Optimization, PID Control

1 Introduction

Since the late 1990's and even more recently in the 2000's, there has been an extensive study of different algorithms for solving optimization problems. The 'No-Free-Lunch' Theorems for Optimization states that there is no universal algorithm for all problems, which means that there is always scope for improvement in the field of optimization [1]. In the book 'Nature-Inspired Optimization Algorithms' [2] Dr. Yang explores a whole new study of metaheuristics using nature inspired techniques. The bat algorithm as well as the Cuckoo-Search algorithm were known to produce promising results compared with those of its predecessors such as Particle Swarm Optimization or Genetic Algorithms. However, it was seen that while BA is a powerful algorithm in exploitation, it may sometimes end up being trapped into some local optima which makes it harder to use for global optimization algorithms [3]. The GEA was introduced to retail some advantages of each algorithm while avoiding some of the disadvantages [4]. With the help of the software MATLAB / Simulink, we aim to visualize the optimization mechanics of the three algorithms, namely GEA, PSO and BA and test and compare the efficiency of each for six pre-defined functions of varying complexity. We shall later test the same for a control application involving system identification and PID control.

The article is arranged as follows: Section 1 provides an introduction to the article as a whole. Section (2) describes optimization approach using metaheuristic algorithms and then explains the three algorithms (GEA, PSO and BA) in brief after which, the results of the optimization are compared for a number of multi-modal functions. Section (3) introduces the concept of relay-based identification of dynamic systems and uses the GEA algorithm for parametric identification of the dynamic process following which, in section (4) a number of PID control algorithms are looked into for appropriate control of the processes using the identified dynamic model of the same. Finally, in section (6), the results of the optimization, identification and control are evaluated for a real dynamic system.

2 Optimization

The two main categories in optimization are called *Deterministic* and *Stochastic* based optimization algorithms. Deterministic algorithms follow a definite procedure for arriving at a solution. Newton's gradient-based algorithm for finding the maxima or minima of a function in finite space, also known as the '*Hill-climbing algorithm*' is one such approach for finding the optima. While deterministic algorithms tend to always find a solution if it does exist, conventional gradient-based approaches are not quite feasible for large-scale problems in which the search space is larger or when the dimensions are much higher and due to this, we need to add a bit of randomness while searching for a solution. With Stochastic algorithms, we use the principle of probability wherein, we use randomization techniques in a well-defined and progressive manner so as to search for the optima in various different sections of the design space. The drawback of this method is that while these algorithms converge very quickly, they may tend to get stuck in a local optimum value instead of finding the true global minimum. That is, we cannot guarantee that we shall arrive at the precise solution, but rather, we look for the best solution which meets our pre-defined criteria which comes under the branch of *Heuristics*. The term '*Metaheuristics*' is a sub-branch of heuristics, but Metaheuristics are inherently problem independent and can be used for a wide range of applications.

As part of this article, the role of Nature-Inspired Metaheuristic algorithms in solving multi-modal problems is looked into. The mechanics governing these algorithms is closely related to natural occurring phenomena and is loosely modelled around natural species like swarm behavior (Particle Swarm Optimization), echolocation of bats (Bat Algorithm), survival of the fittest (Genetic Algorithm). Among all these algorithms, there exist a few main patterns which are similar and are necessary for efficient convergence to the optimum value. These three operators are as follows:

1. **Crossover:** Provides good mixing within the solution space
2. **Mutation:** Provides essential exploratory framework for *diversification* of the solutions to avoid settling at a local optimum solution.
3. **Selection:** Provides the necessary exploitative framework for *intensification* of the search.

Each algorithm essentially uses each of the above operators, but in different fashion. The manner in which it is carried out, thus determines the efficacy of the method. While it is clear that there is no best method in general, the effectiveness is largely dependent on the problem at hand.

To apply the above methods for our problem set, we shall have to use the following terms for building our solution:

1. **Dimensions (d):** The dimensions is the number of variables which influence the objective functions or, in short, this can be given as the number of search parameters to be optimized.
2. **Random variables:** These are variables whose values are randomly determined by a set of random distribution rules such as Gaussian distributions, Uniform distributions, Levy Distributions. For the purpose of the algorithms in this article, we use Gaussian distributed and uniformly distributed random variables.
3. **Random Walk:** A random walk is essentially a step taken randomly from a fixed point. The final value depends on a number of randomly taken steps from the initial position.
4. **Solutions (X):** The solutions is a vector containing a number of points in a d-dimensional search space. In some places, this is called *individuals* especially in relation to nature-inspired metaheuristics, and in some others, like in PSO, these are known as *particles*. Solutions or individuals can be mathematically represented as:

$$\mathbf{x}_N = \sum_{i=1}^N \mathbf{x}_i \quad (1)$$

Where,

$\mathbf{x}_i = [x_{i1}, x_{i2} \dots x_{id}]$, d = number of dimensions of the search space

5. **Cost function (f(x)):** The cost function of any optimization problem is the most important as it describes the problem or the function which needs to be optimized.

With the above definitions, we shall proceed to describe each of the methods.

2.1 Particle Swarm Optimization

Particle Swarm Optimization (PSO) is a meta-heuristic based optimization algorithm which was developed by modelling swarm behavior observed in nature such as bird flocking, fish schooling, swarming theory [5]. The idea behind this technique is based on the principle that each individual in the subsequent generations can benefit from the observations by the other individuals from the previous generations, thus improving the subsequent results.

As described by the original algorithm, with each iteration (generation) each individual particle is attracted toward the position of the current global best \mathbf{g}^* as well as its own best location \mathbf{x}_i^t in history. The positions and velocities of the particles namely, \mathbf{x}_i and \mathbf{v}_i respectively are updated as follows:

$$\mathbf{v}_i^{t+1} = \theta \mathbf{v}_i^t + \alpha \epsilon_1 [\mathbf{g}^* - \mathbf{x}_i^t] + \beta \epsilon_2 [\mathbf{x}_i^* - \mathbf{x}_i^t] \quad (2)$$

$$\mathbf{x}_i^{t+1} = \mathbf{x}_i^t + \mathbf{v}_i^{t+1} \quad (3)$$

The parameters α and β are learning parameters and the values of these can be adjusted based on the problem at hand and depending on how quickly we want the solution to converge. The parameter θ is the inertia weight and while it is not absolutely necessary, it helps to ensure that the particles do not jump around the optimum value. Finally, at the end of every generation, the global best position as well as the individual best is updated using the cost function as follows:

$$\begin{aligned} \mathbf{g}^* &= \min\{f(\mathbf{x}_i)\} \\ \mathbf{x}_i^* &= \mathbf{x}_i, \text{ if } f(\mathbf{x}_i) < f(\mathbf{x}_i^*) \end{aligned} \quad (4)$$

The Pseudo-code of the above algorithm can be given as follows [2]:

Particle Swarm Optimization

1. Objective function $f(\mathbf{x})$, $\mathbf{x} = (x_1, \dots, x_d)^T$
2. Initialize locations \mathbf{x}_i and velocity \mathbf{v}_i of n particles
3. Find \mathbf{g}^* from $\min\{f(\mathbf{x}_1), f(\mathbf{x}_2) \dots f(\mathbf{x}_n)\}$ (at $t = 0$)
4. While (criterion)
 - For (loop over n-particles and d-dimensions)
 - Generate new velocity \mathbf{v}_i^{t+1}
 - Calculate new locations $\mathbf{x}_i^{t+1} = \mathbf{x}_i^t + \mathbf{v}_i^{t+1}$
 - Evaluate objective function at locations \mathbf{x}_i^{t+1}
 - Find the current best for each particle \mathbf{x}_i^*
 - End for
 - Find the current global best \mathbf{g}^*
 - Update $t = t + 1$ (iteration counter)
5. End while
6. Output final results \mathbf{x}_i^* and \mathbf{g}^* .

Figure 1- Pseudo Code - Particle Swarm Optimization

2.2 Bat Algorithm

The inherent principle of BA is quite similar to PSO wherein the BA also uses swarm intelligence for finding the optima, however, in case of BA, the method used is frequency tuning which, is based on the echolocation principle of bats. In essence, PSO is a specific case of BA upon appropriate parameter setting. In case of BA, the position and velocities are updated using the frequency f as follows:

$$f_i = f_{min} + (f_{max} - f_{min})\beta \quad (5)$$

Where,

f_i : frequency of i th particle

f_{max}, f_{min} : max and min search frequency ($f_{min} = 0, f_{max} = 2$)

$$v_i^{t+1} = v_i^t + f_i \cdot (x_i^t - x^*) \quad (6)$$

$$x_i^{t+1} = x_i^t + v_i^{t+1} \quad (7)$$

Once again, the global best position x^* is updated at the end of each iteration similar to PSO. In addition to the above, the BA also provides a necessary mechanism for exploitation using local search as:

$$x_{new} = x_{old} + \sigma \epsilon_t A^t, \quad (8)$$

Where,

σ : Scaling factor (= 0.1 x search range for each dimension)

The intensity of the local search is supposed to increase as the solution moves closer towards the global minimum which is therefore more as the algorithm progresses as compared to the beginning stages. This can be regulated by increasing the *pulse emission rates* as the number of generations rise. Additionally, as the bat moves closer to its prey, the *loudness* decreases, which means that the algorithm moves from exploratory mode to exploitative mode. This can be represented by:

$$A_i^{t+1} = \alpha A_i^t \quad (9)$$

$$r_i^{t+1} = r_i^0 [1 - e^{-\gamma t}] \quad (10)$$

The above equations and search procedure can be seen in the below pseudo-code.

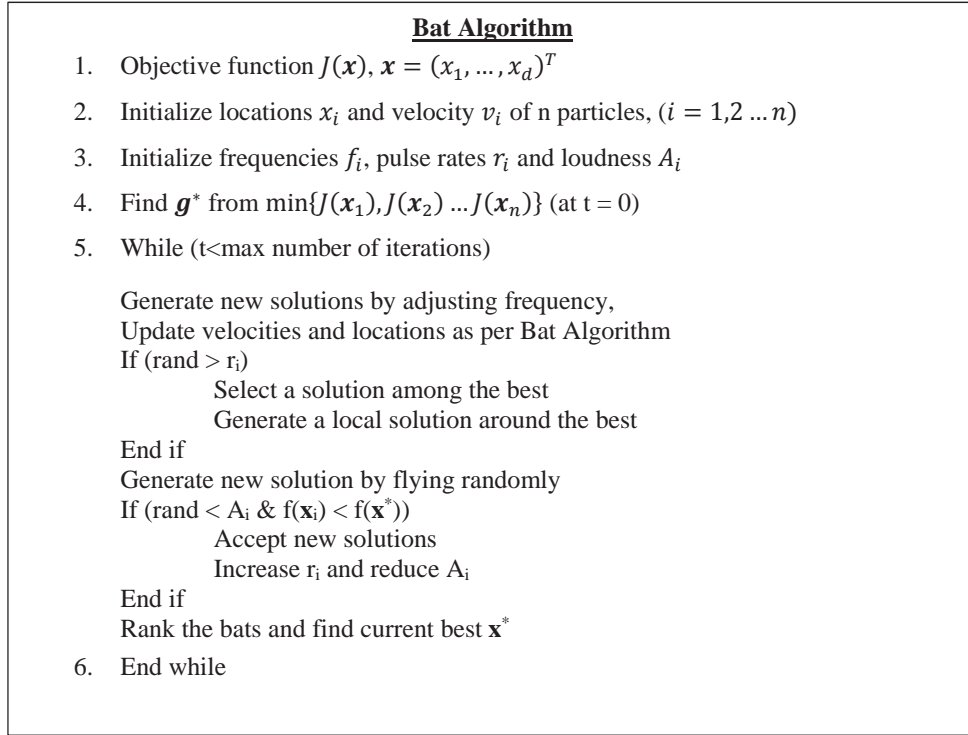


Figure 2 - Pseudo Code - Bat Algorithm

2.3 Guiding Evolutionary Algorithm (GEA)

The Guiding Evolutionary Algorithm, while similar to the above two described algorithms, is essentially simpler to use due to the fewer number of parameters required for tuning the optimization. Like the above, the algorithm consists of Crossover, mutation and local search which can be described by the below equations much simpler as compared to the previous ones.

1. Crossover: The crossover for GEA is given by:

$$x_i^t = x_i^{t-1} + (x_*^{t-1} - x_i^{t-1}) * \beta \quad (11)$$

Where,

β : Step length of position increment, uniformly distributed r.v

The step length β defines the rate of convergence and is generally between 1 and 2. A higher value represents faster convergence.

2. Mutation: The mutation provides the required exploratory mechanics for optimization. It can be given by the equation:

$$x_i^t = x_i^t + \epsilon M \quad (12)$$

Where,

ϵ : Uniform r.v [-1, 1]

M : Mutation vector,

$M_j = (\max(x_{ij}^t - a, b - x_{ij}^t))$; $[a, b]$ = range of j^{th} dimension

According to GEA, the probability of mutation is given by the following:

$$p = c * \ln\left(\frac{T_{max}}{T_{max}-t}\right) \quad (13)$$

Where,

T_{max} : max number of generations

t : current generation

c : 0.2 (constant)

The above equation shows that the probability of mutation increases as the generations pass, thus it helps ensuring that the solution does not settle around the local maxima / minima.

3. Local: As before, the local search provides the necessary exploitative dynamics and can be given as follows:

$$x_i^t = x_*^{t-1} + \epsilon L \quad (14)$$

Where,

L: Local search vector

$L_j = 0.1(b - a)$; [a, b] = range of j^{th} dimension

Similar to the mutation vector, the local search functions similar to the mutation except that it serves to find a solution around the current best unlike mutation, which helps to find a new solution around the unsearched territory.

Again, the probability of local search is given by p defined by equation (13).

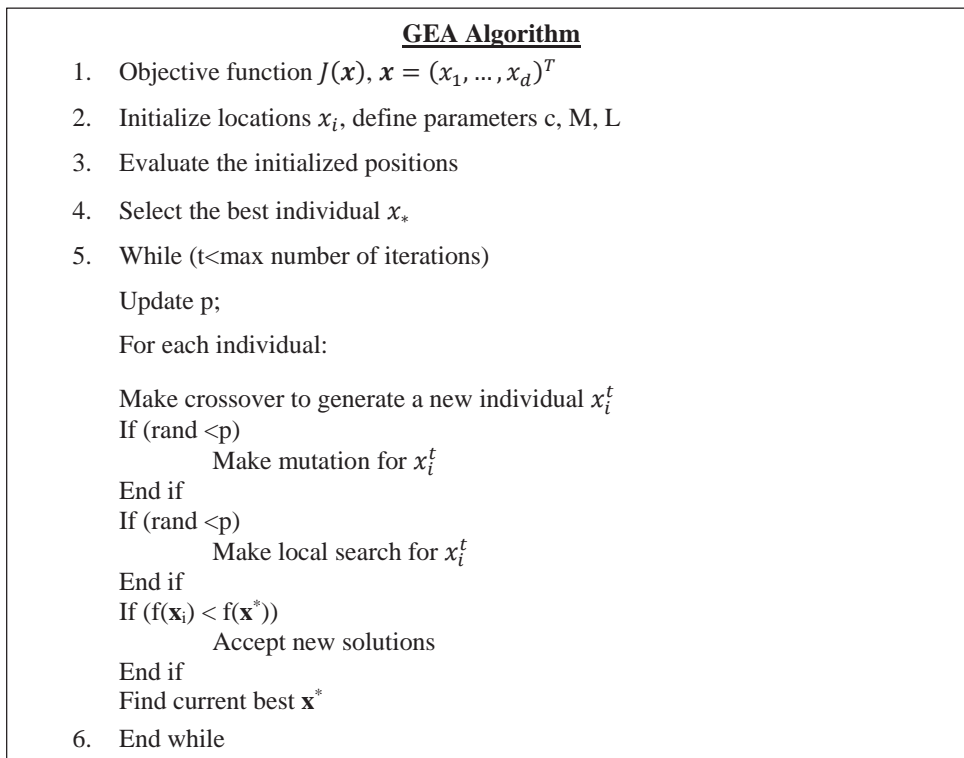
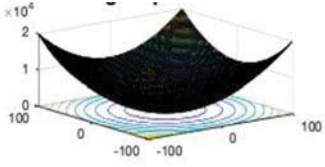
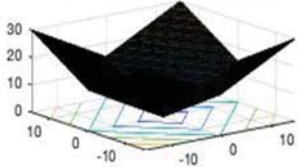
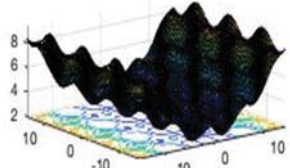
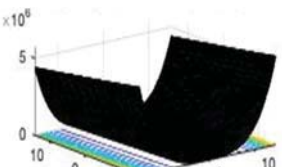
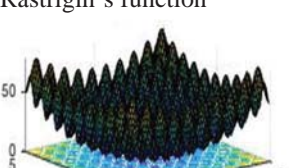



Figure 3 - Pseudocode - GEA Algorithm

2.4 Test functions

To evaluate the above methods, we shall test the above algorithms on six test functions as below and verify how quickly each tends to arrive at the real global optimum value.

Table 1 - Test Functions

Functions	Function Name	Expression	Domain
F1	De-Jong's Sphere function 	$f(x) = \sum_{i=1}^D x_i^2$	[-100, 100]
F2	Schwefel 2.2 function 	$f(x) = \sum_{i=1}^D x_i + \prod_{i=1}^D x_i $	[-15, 15]
F3	Griewangk's function 	$f(x) = -\prod_{i=1}^D \cos\left(\frac{x_i}{\sqrt{i}}\right) + \sum_{i=1}^D \frac{x_i^2}{4000} + 1$	[-15, 15]
F4	Rosenbrock's function 	$f(x) = \sum_{i=1}^{D-1} 100 * (x_{i+1} - x_i^2)^2 + (x_i - 1)^2$	[-15, 15]
F5	Rastrigin's function 	$f(x) = D * 10 + \sum_{i=1}^D (x_i^2 - 10 * \cos(2\pi x_i))$	[-5, 5]
F6	Michalewicz function 	$f(x) = -\left\{ \sin(x) \left[\sin\left(\frac{x^2}{\pi}\right) \right]^{2m} + \sin(y) \left[\sin\left(\frac{2y^2}{\pi}\right) \right]^{2m} \right\}$	[0, 4]

For all of the above functions, our goal is to evaluate the point or points at which, the value of the above functions is at a minimum. We can visualize the above test functions as per the below fig. 4. For the purpose of visualization, we contain the number of dimensions to three at max since, for greater dimensions, we are unable to see the working mechanism of each in detail.

2.5 Parameterization

Before simulating the above functions, we first need to set appropriate parameters to ensure that our algorithm works in an efficient manner. The parameters for each algorithm are taken from literature. For PSO, we consider the parameters as follows [4]:

Parameter	Value
β	1.5
γ	0.9
θ	0.7
α_0	1.5
α	$\alpha_0 * \gamma^t$

Similarly, the parameters for BA are given as [2]:

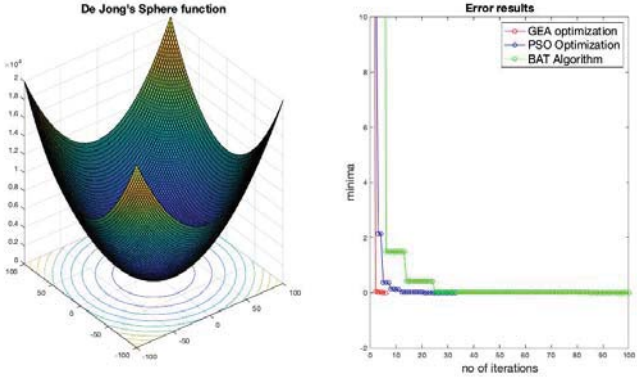
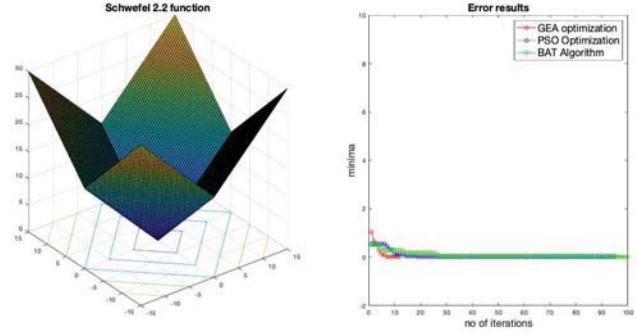
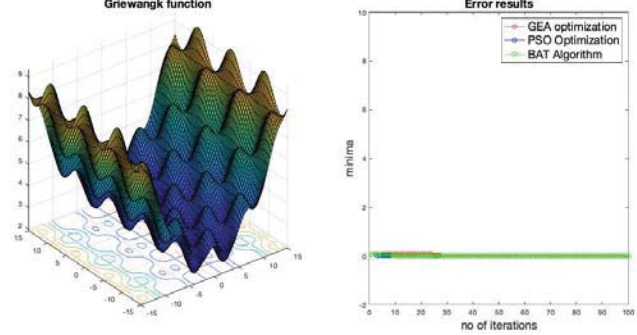
Parameter	Value
Q_{min}	0
Q_{max}	2
α	0.97
γ	0.9
θ	0.7
σ	$0.1 * (ub - lb)$
A	0.95
ρ	0.6

For GEA, we use the parameters as follows [4]:

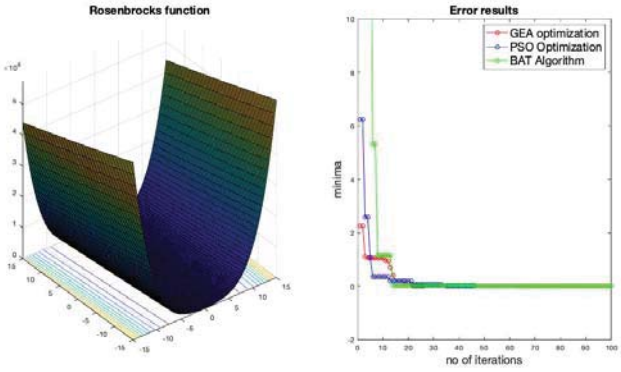
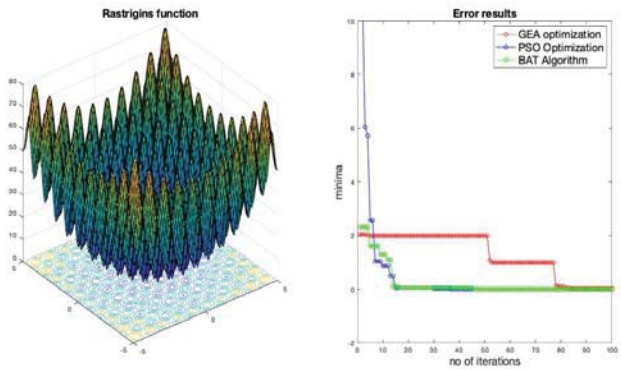
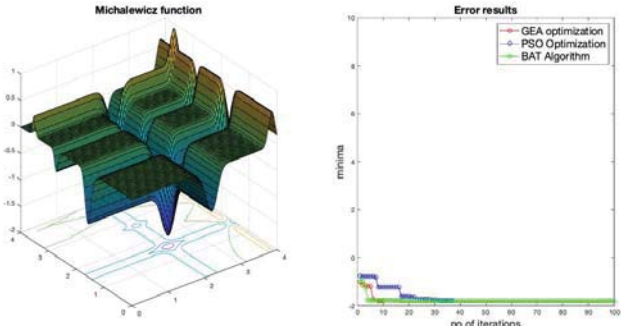
Parameter	Value
c	0.97
β	[0, 2]

2.6 Test Results

Using the above parameters for each algorithm, the functions are simulated in MATLAB from which we obtain the results as per the below table¹.

Functions	Results	Minima
F1	 <p>The figure for F1 consists of two subplots. The left subplot is a 3D surface plot titled "De Jong's Sphere function" showing a smooth, bowl-shaped surface with a central minimum. The axes range from -100 to 100. The right subplot is titled "Error results" and shows the "minima" value on the y-axis (ranging from -2 to 10) against the "no of iterations" on the x-axis (ranging from 0 to 100). Three data series are plotted: GEA optimization (red line with circles), PSO Optimization (blue line with squares), and BAT Algorithm (green line with triangles). All three algorithms show a rapid decrease in error, reaching near-zero values within the first 10 iterations.</p>	(0,0,0)
F2	 <p>The figure for F2 consists of two subplots. The left subplot is a 3D surface plot titled "Schwefel 2.2 function" showing a complex, multi-peaked surface. The axes range from -15 to 15. The right subplot is titled "Error results" and shows the "minima" value on the y-axis (ranging from -2 to 10) against the "no of iterations" on the x-axis (ranging from 0 to 100). Three data series are plotted: GEA optimization (red line with circles), PSO Optimization (blue line with squares), and BAT Algorithm (green line with triangles). All three algorithms show a rapid decrease in error, reaching near-zero values within the first 10 iterations.</p>	(0,0,0)
F3	 <p>The figure for F3 consists of two subplots. The left subplot is a 3D surface plot titled "Griewangk function" showing a highly oscillatory surface with many local minima. The axes range from -15 to 15. The right subplot is titled "Error results" and shows the "minima" value on the y-axis (ranging from -2 to 10) against the "no of iterations" on the x-axis (ranging from 0 to 100). Three data series are plotted: GEA optimization (red line with circles), PSO Optimization (blue line with squares), and BAT Algorithm (green line with triangles). All three algorithms show a rapid decrease in error, reaching near-zero values within the first 10 iterations.</p>	(0,0,0)

¹ It must be noted that the results obtained in the table refers to a single random observation. Nevertheless, the results displayed correctly reflect the observations obtained over a wide number of simulations.

<p>F4</p>	 <p>The figure shows the Rosenbrocks function as a 3D surface plot with a deep valley and a sharp ridge. The error results plot shows the minimum value of the function over 100 iterations for three algorithms: GEA optimization (red line with circles), PSO Optimization (blue line with squares), and BAT Algorithm (green line with triangles). All algorithms converge to a minimum value of approximately 0.1 within the first 10 iterations.</p>	<p>(1,1,0)</p>
<p>F5</p>	 <p>The figure shows the Rastrigins function as a 3D surface plot with many sharp peaks and valleys. The error results plot shows the minimum value of the function over 100 iterations for three algorithms: GEA optimization (red line with circles), PSO Optimization (blue line with squares), and BAT Algorithm (green line with triangles). GEA optimization converges to a minimum value of approximately 0.1, while PSO and BAT algorithms get stuck at a higher minimum value of approximately 1.0.</p>	<p>(0,0,0)</p>
<p>F6</p>	 <p>The figure shows the Michalewicz function as a 3D surface plot with a complex, multi-modal landscape. The error results plot shows the minimum value of the function over 100 iterations for three algorithms: GEA optimization (red line with circles), PSO Optimization (blue line with squares), and BAT Algorithm (green line with triangles). GEA optimization converges to a minimum value of approximately 1.8, while PSO and BAT algorithms get stuck at a higher minimum value of approximately 2.2.</p>	<p>(2.2,1.57,-1.8013)</p>

From the above observations, we see that the algorithms perform similarly for unimodal functions. For multimodal functions F3, F5 and F6, the algorithms GEA performs better than BAT and PSO especially due to the presence of the mutation operator which helps to prevent the algorithm getting stuck at its local optima. Furthermore, to enhance the capability of GEA to not get stuck at local optima, it is necessary to tweak the mutation and local search operator in such a way as to perform the mutation and local search on each dimension individually and not on all parameters as a whole. This can be done by modifying equations (12) and (14) as follows:

$$x_i^t = x_i^t + \epsilon M_j \cdot (\text{rand}(1, j) < p) \tag{15}$$

And,

$$x_i^t = x_i^{t-1} + \epsilon L \cdot (\text{rand}(1, j) < p) \tag{16}$$

The above modification is quite important in the case of multimodal functions wherein, the solution tends to get stuck in a local optima as the combined mutation approach renders it difficult for the solution to move all at once, but, when done individually for each dimension, there is a greater chance of finding a different solution. Without this

modification, the next tasks of system identification are fruitless as the solution rarely reaches its optimum value.

In conclusion, we see that the GEA algorithm works extremely well for unimodal as well as multimodal functions and hence we can confirm its suitability for usage in the next part which involves system identification.

3 Relay-based Feedback Identification

With our optimization algorithm working, the next step is to apply this same algorithm in parametric identification of a dynamic system. In the part the goal is to perform *black box identification* of a dynamic system applicable to a wide range of processes. The main advantage of black box identification is the fact that this approach necessitates no knowledge about the physical system, but rather, uses the experimental data which includes inputs and outputs and a certain defining factor in terms of the cost function to identify the system to a reasonable accuracy. In this case, the ITAE (*Integral Time Absolute Error*) Method will be used as a criterion for the cost function. With system identification, the end goal is to find an accurate model of the process so as to control it correctly. In the later section, we shall use the model identified from this section to tune our PID controller which will be used to control our process.

3.1 System Schematic

The schematic of our system can be seen in the below figure.

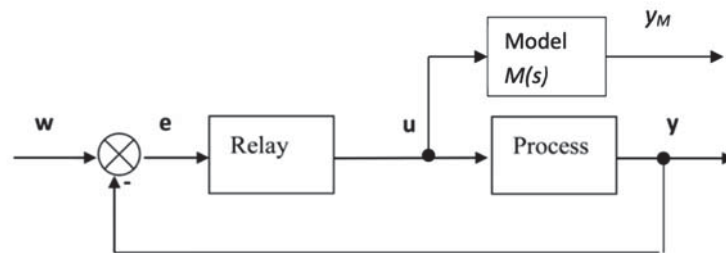


Figure 4 - Experimental Block Diagram Using Simulink

The above figure shows the process $G(s)$ that we wish to identify and then control. The input u is given by the output of the on-off relay which is used to generate sustained oscillations within the closed-loop system as proposed by Astrom and Hagglund [6]. Once the closed-loop system is running automatically, we use the output from the controller 'u' and feed the same as input to our model $G_M(s)$ which, we consider to be a Second Order Plus Time Delay (SOPTD) model which is required to represent the dynamics of the actual process which we are trying to control. The reason for using an SOPTD model is that it can represent almost any linear system. As explained by Ramakrishnan and Chidambaram, the SOPTD model can incorporate various processes such as under-damped and higher order processes in which case, an FOPTD model is not sufficient [7]. Furthermore, SOPTD models can also be used for unstable processes in which case, an FOPTD model is not sufficient.

Thus, by knowing the input 'u', the outputs 'y' and 'y_M', we can proceed with identification of the process.

3.2 Problem Statement

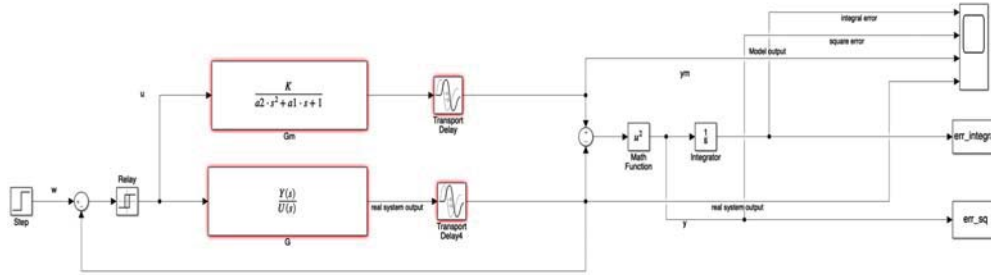


Figure 5 - Relay Identification Schematic

With reference to the schematic from Simulink, we proceed with defining our problem.

1. Model: The SOPTD model describe above can be represented as follows:

$$G_M(s) = \frac{K}{a_2 s^2 + a_1 s + 1} e^{-sT_d} \quad (17)$$

Where, 'K' = Process gain, 'θ' = Time Delay, 'a₁' and 'a₂' are dynamic constants of the transfer function.

2. Parameters: Using the SOPTD model from equation (17), our goal is to identify the parameters as follows:

$$x = [K, a_1, a_2, T_d]^T \quad (18)$$

3. Cost function: The criterion for optimization can be given by the cost function J which can be given by:

$$J = \int_0^{T_k} (y_m(t) - y(t))^2 dt \quad (19)$$

Where,

y_m = Model output

y_t = Process output

T_k = Simulation time

The above cost function uses the ITAE criterion as our optimization constraint for minimizing the error over a pre-set simulation time.

4. Constraints: The constraints restrict the upper and lower limits of each parameter

$$\{K, a_1, a_2, T_d\} > 0 \quad (20)$$

3.3 Pre-defined Processes

For the purpose of identification and verifying our identification procedure, we use a number of pre-defined test functions describing various types of processes. These are given as follows:

Table 2 - Predefined Process Functions

No.	Type of Process	Function
1	Non-oscillatory Process (Lag-dominated)	$P_1(s) = \frac{1}{(s+1)(0.1s+1)(0.01s+1)(0.001s+1)}$
2	Balanced Process	$P_2(s) = \frac{1}{(s+4)^4}$
3	Delay-dominated Process	$P_3(s) = \frac{1}{(0.05s+1)^2} e^{-s}$
4	Oscillatory Process	$P_4(s) = \frac{1}{(0.5s^2+s+1)^2}$
5	Non-oscillatory Process with Time Delay	$P_5(s) = \frac{1}{(s+1)(0.3s+1)^2} e^{-sT_d}$
6	Fifth Order Process with Time Delay	$P_6(s) = \frac{e^{-s}}{(5s+1)^5}$

Each of the above processes represent a different type of process and our goal is to fit the complex process into a lower order model which can closely describe the actual response of the systems.

3.4 Identification Results

The results of the above identification procedure and optimization algorithm can be summarized in the table below:

Table 3 - Identification Results for Pre-defined systems

No.	Function	K	a2	a1	Td (s)	fmin	Elapsed Time (s)
1	$P1 = \frac{1}{(s+1)(0.1s+1)(0.01s+1)(0.001s+1)}$	0.998 ± 0.096	0.0674 ± 0.049	1.172 ± 0.235	0.033 ± 0.039	0.0207 ± 0.023	92.37 ± 33.76
2	$P2 = \frac{1}{(s+1)^4}$	0.966 ± 0.083	3.32 ± 0.369	3.034 ± 0.26	0.891 ± 0.099	0.122 ± 0.049	93,9 ± 45,56
3	$P3 = \frac{e^{-s}}{(0.05s+1)^2}$	1.002 ± 0.009	0.00038 ± 0.001	0.0823 ± 0.008	1.029 ± 0.029	0.324 ± 0.139	135.1 ± 50.1
4	$P4 = \frac{1}{0.5s^2 + s + 1}$	0.943 ± 0.065	1.017 ± 0.122	1.147 ± 0.107	0.364 ± 0.105	0.364 ± 0.105	107.4 ± 39.65
5	$P5 = \frac{e^{-s}}{(s+1)(0.3s+1)^2}$	0.99 ± 0.10	0.515 ± 0.153	1.437 ± 0.114	1.122 ± 0.105	0.155 ± 0.107	107,83 ± 57,30
6	$P6 = \frac{e^{-s}}{(5s+1)^5}$	1.022 ± 0.217	109.9 ± 13.17	17.17 ± 1.33	8.20 ± 0.66	0.222 ± 0.058	83.65 ± 69.09

Each system was simulated ten times for consistency to obtain the above results. We can verify the correctness of the above identified systems by comparing the step response and the Nyquist plot of the modelled system with those of the real processes.

For example, a random simulation of the system 6 (P1) obtained the below step response and Nyquist:

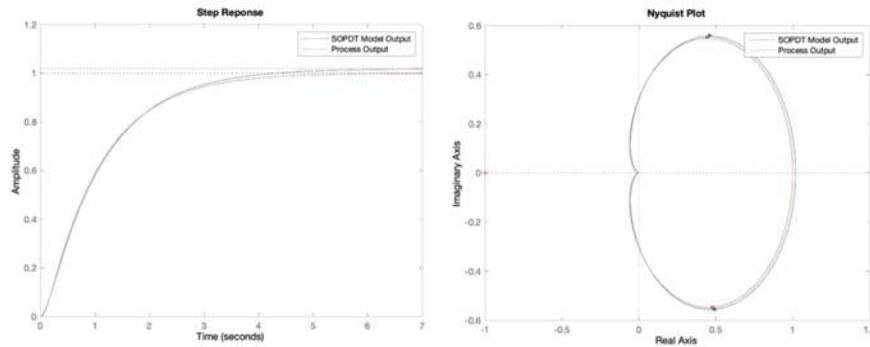


Figure 6 - Step Response, Nyquist Plot - Process P1

We see that both the step response as well as the Nyquist plot, closely resemble the actual system which we are trying to model and therefore, we can say that our identification is correct. Similar results were observed for the remaining systems as well. It must, be noted that although the algorithm manages to converge quickly, it is necessary to set the initial guess in the range selection to reasonable values. In this case the selected range for our initialization was [20, 20, 20, 5].

4 PID Control

As stated earlier, the end goal of system identification of a dynamic process using lower order models is control. Therefore, our goal is to achieve appropriate control of the processes described by the equations in table (2) using the model identified in section (3).

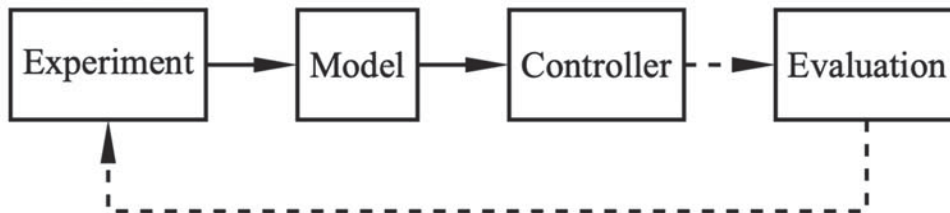


Figure 7 - Tuning Procedure [8]

It must be noted that all the processes described in table (2) are all stable processes. Hence our goal with control is to firstly minimize the settling time and secondly to reduce the overshoot to an acceptable degree. For the purpose of control, we make use of the most commonly used PID controller.

4.1 Closed-Loop System schematic

The block diagram of the controller can be shown in the below figure.

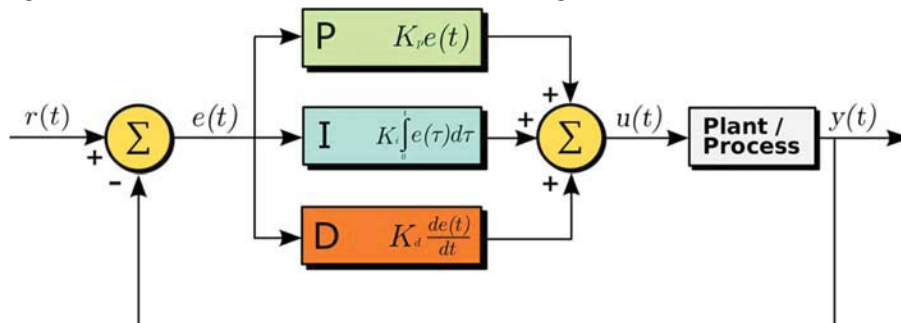


Figure 8 - PID Controller Block Diagram [9]

The equation for the controller in Laplace domain can be given by:

$$R(s) = r_p + \frac{r_I}{s} + r_D s \quad (21)$$

Where,

r_p = Proportional Gain – P-term (K_p)

r_I = Integral Gain – I-term (K_I)

r_D = Derivative Gain – D-term (K_D)

It must be noted that in the above figure (8), we aim to try to control the real unknown process using only the identified model with no more information about the system.

4.2 Tuning Methods

To tune the controller is to find suitable values of the proportional, integral and derivative gain so as to achieve our control objectives. For this, we will have to make use of a few empirical methods for controller tuning, namely:

- Direct Synthesis Method (DS) or Lambda-Tuning Method
- Phase Margin Criterion (PMC) based Tuning Method [10]
- Simple Control (SIMC Tuning Method

For applying the above methods, we shall have to reduce the SOPTD Model $G_M(s) = K \cdot \frac{e^{-sT_d}}{a_2 s^2 + a_1 s + 1}$ obtained in section (3.4) to one of the below forms:

$$G(s) = \frac{(K\omega_0^2)e^{-sT_d}}{s^2 + 2\zeta\omega_0 s + \omega_0^2} \text{ for oscillatory processes} \quad (22)$$

Where,

ω_0 = oscillation frequency

ζ = damping ratio

or

$$G(s) = \frac{K e^{-sT_d}}{(T_1 s + 1)(T_2 s + 1)} \text{ for non-oscillatory processes} \quad (23)$$

Where,

T_1, T_2 = Time constants of modelled system

The relations for tuning the controller by the DS method [11], PMC method and the SIMC methods can be given by the below table:

- Non-oscillatory Processes ($\zeta \geq 1$)

Table 4 - Tuning Equations (Non-Oscillatory Process)

No.	Parameter	DS	PMC	
			$T_d > 0$	$T_d = 0$
1	r_p	$\frac{1}{K} \cdot \frac{T_1 + T_2}{\tau_c + T_d}$	$r_p = a_1 \cdot r_i$	$r_p = a_1 r_i$
2	r_I	$\tau_i = T_1 + T_2;$ $r_i = r_p / \tau_i$	$r_i = \frac{\pi - 2\gamma}{2 \cdot K T_d}$ or $r_i = \frac{\pi}{2 T_d \cdot m_A \cdot K}$	$r_i = \frac{1}{K \tau_c}$
3	r_D	$\tau_d = \frac{T_0}{2\zeta};$ $r_d = r_p \cdot \tau_d$	$r_d = a_2 \cdot r_i$	$r_d = a_2 r_i$

In the above equations, the parameter τ_c is the closed loop time constant and can be estimated using a general rule:

$$\tau_{dom} > \tau_c > T_d \quad (24)$$

Where,

τ_{dom} = dominant time constant of the process

As a thumb rule, we shall estimate the value of τ_c as: $\tau_c = \frac{\tau_{dom}}{3}$.

For PMC Tuning, the parameters γ and m_A refer to the phase margin and gain margin respectively. These values are generally in the range:

$$\frac{\pi}{6} < \gamma < \frac{\pi}{3}; 2 < m_a < 5 \tag{25}$$

It must be noted that the PMC tuning method is suitable for an SOPTD Model only.

- b) Oscillatory Processes ($\zeta < 1$)

Table 5 - Tuning Equations (Oscillatory Process)

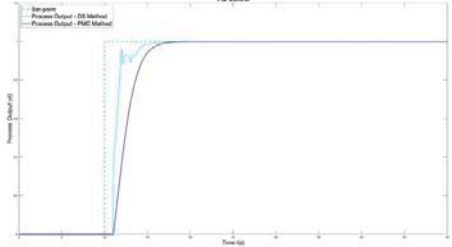
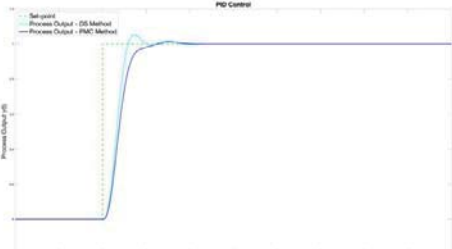
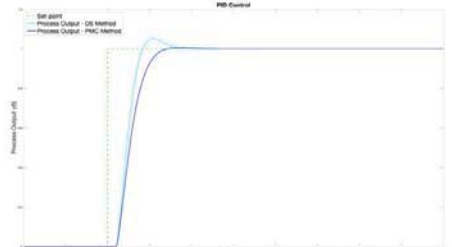
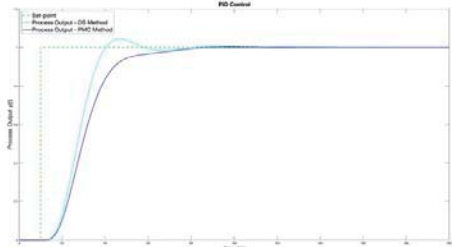
No.	Parameter	DS	PMC	
			$T_d > 0$	$T_d = 0$
1	r_p	$r_p = \frac{2\zeta T_0}{\tau_c + T_d}$	$r_p = a_1 \cdot r_i$	$r_p = a_1 r_i$
2	r_I	$\tau_i = 2\zeta T_0$ $r_i = \frac{r_p}{\tau_i}$	$r_i = \frac{\pi - 2\gamma}{2 \cdot K T_d}$ or $r_i = \frac{\pi}{2T_d \cdot m_A \cdot K}$	$r_i = \frac{1}{K\tau_c}$
3	r_D	$\tau_d = \frac{T_0}{2\zeta}$ $r_d = r_p \cdot \tau_d$	$r_d = a_2 \cdot r_i$	$r_d = a_2 r_i$

4.3 Tuning Results

The results of the tuning using the above methods are displayed in table

Table 6 - PID Tuning Results

No.	Function	Identified Model	PID Control
1	P1	$G_M(s) = \frac{0.9843}{0.1095s^2 + 1.038s + 1}$	
2	P2	$G_M(s) = \frac{0.9446e^{-0.8837s}}{3.3514s^2 + 2.9958s + 1}$	

3	P3	$G_M(s) = 1.0175 \cdot \frac{e^{-1.0285s}}{0.0786s + 1}$	
4	P4	$G_M(s) = 0.9469 \cdot \frac{e^{-0.8172s}}{0.8799s^2 + 1.1987s + 1}$	
5	P5	$G_M(s) = \frac{0.9767 \cdot e^{-1.1509s}}{0.4658s^2 + 1.4062s + 1}$	
6	P6	$G_M(s) = \frac{0.9773 \cdot e^{-8.61s}}{107.21s^2 + 18.07s + 1}$	

From the above output, we can conclude that the DS and PMC algorithms perform satisfactorily when it comes to tuning the processes P1-P6 using the identified SOPTD model.

5 Identification and Control on Physical System

To verify the theory and algorithms presented in the previous sections, we apply the same concepts of identification and PID control on a real system from the Automatic Control Laboratory.

5.1 Experimental Setup

The system which we will be using is a combination of two chambers arranged vertically in a tube with a system of interconnecting valves. The schematic of the setup can be seen in the below figures.

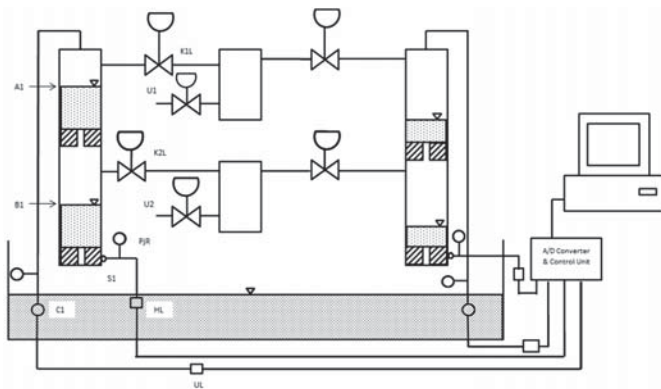


Figure 9 - Schematic Diagram - Two Tanks System

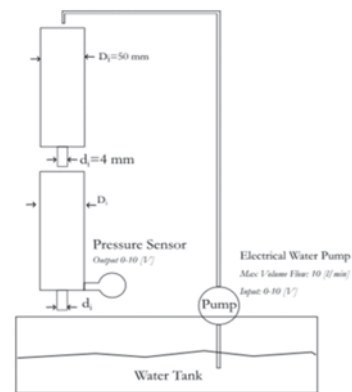


Figure 10 - Functional Diagram - Two Tanks System

The experiment is a combination of two hollow chambers arranged vertically in a tube with a system of interconnecting valves. Pump C1 supplies water to the upper chamber, thus, delivering a pressure head which causes the accumulated water in the upper chamber to trickle down into the lower chamber and eventually back into the tank. A pressure differential sensor is used to map the height of the water in the lower column. Since the upper and lower tanks are connected to each other in series, the system is, by default a second-order system.

To perform the experiment, we need to supply an input signal 0-10V to the pump via the MATLAB / SIMULINK software from the PC which is then transmitted to the pump which in turn pumps up water to the upper chamber A1. From the bottom hole of the upper chamber, water trickles down to the lower tank B1. The system output is measured in terms of the height of water column in the lower chamber B1. This output is measured by a pressure sensor which converts the pressure head to an equivalent height of water column. Before the experiment is performed, it is necessary to test the system for its static characteristics so that we supply the inputs and obtain the readings at around the operating point.

5.2 Simulink Schematic Setup

Similar to the Simulink setup in section (3.2), the experimental Simulink Schema of the physical system is arranged likewise, except that in this case, the input is not fed directly to the theoretical model, but instead is done separately once the readings have been collected. Hence, we will have two separate schematics, one for the physical system and the other, our *Identification Schematic* which uses the values from the physical system to optimize our process model for the purpose of simulation only.

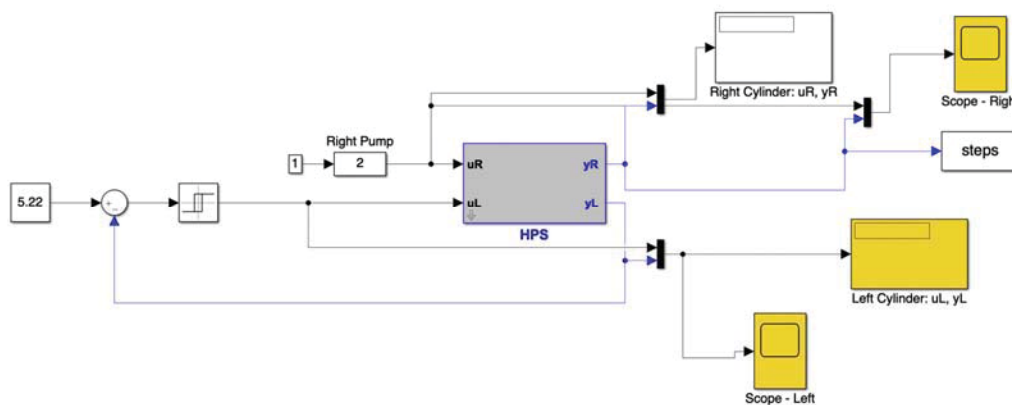


Figure 11 - Simulink Schematic - Physical system

As can be seen from the schematic in Figure (11), we supply the model with an input from our relay controller which in turn causes the system to auto-oscillate. The input set-point 5.22(cm) is supplied to ensure that during experimentation, the process operates around the linear area only which is around the height 5.22 cm and the outputs of the relay are chosen so as be within the linear range. From this setup, the input to the system u_L and the output y_L

are recorded from the practical setup.

Once the values are collected, these are now fed to the simulation setup (*Identification Schematic*) which operates within the MATLAB / Simulink environment only. The advantage of this is that using just one reading from the physical setup, we can simulate the theoretical model a number of times until we obtain the model which nearly describes the unknown process. The schematic of the optimization part can be seen in Figure (12).

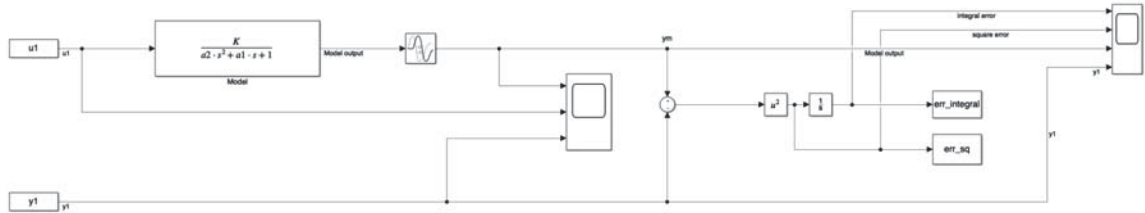


Figure 12 - Identification Schematic

As can be seen from above, the same input u_1 which was supplied to the real process are fed to the theoretical SOPTD model to obtain its output y_m . By comparing the model output y_m with the real output y_1 , we can use our previously described method using the ITAE criterion and the GEA algorithm to identify our system.

5.3 Identification Results

By performing the above experiment and on simulation, we obtain the identified parameters and the model of the system as follows:

$$G_M(s) = \frac{1.10}{376.28s^2 + 49.62s + 1} \tag{26}$$

We shall verify the accuracy of our identified models by plotting the step response and Nyquist plot of the SOPTD model which we have obtained and compare it with the real step response and frequency characteristics.

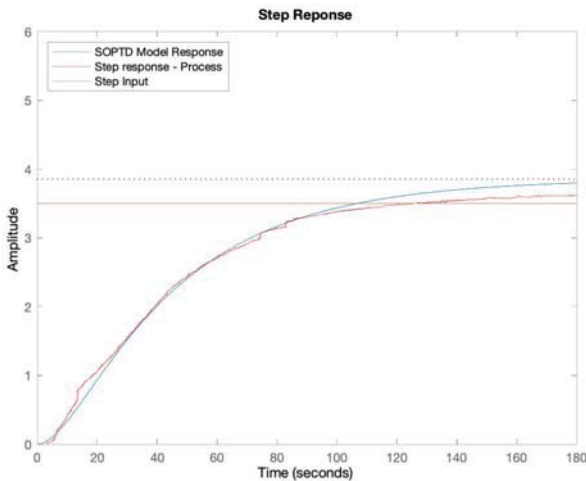


Figure 14 - Step Response - Model vs Real

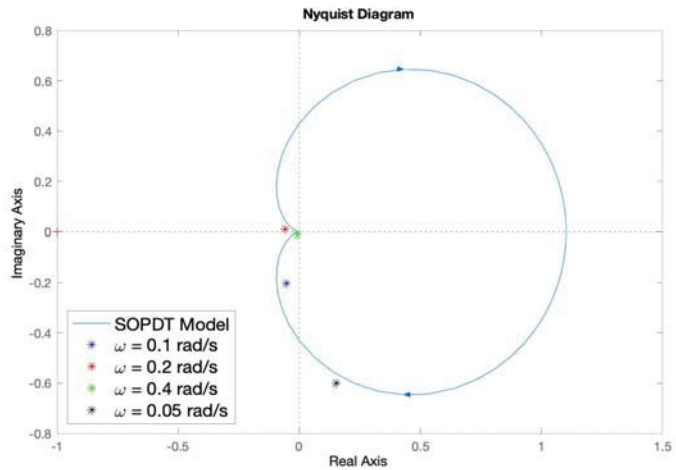


Figure 13 - Nyquist Plot - Model vs Real

We can see from the above step response and the Nyquist plot that the modelled output closely matches the real results process and therefore we can conclude that our identification method works correctly.

5.4 PID Control

Using the above-identified models, we proceed to control the processes using the PID control methods detailed in section (4.2). The tuning parameters for control of the identified process from equation (26) are as follows:

Table 7 - PID Tuning Parameters - Physical System

No.	Parameter	DS	PMC
1	Proportional Gain, r_p	3.36	3.36
2	Integral Gain, r_i	0.07	0.07
3	Derivative Gain, r_d	25.48	25.48

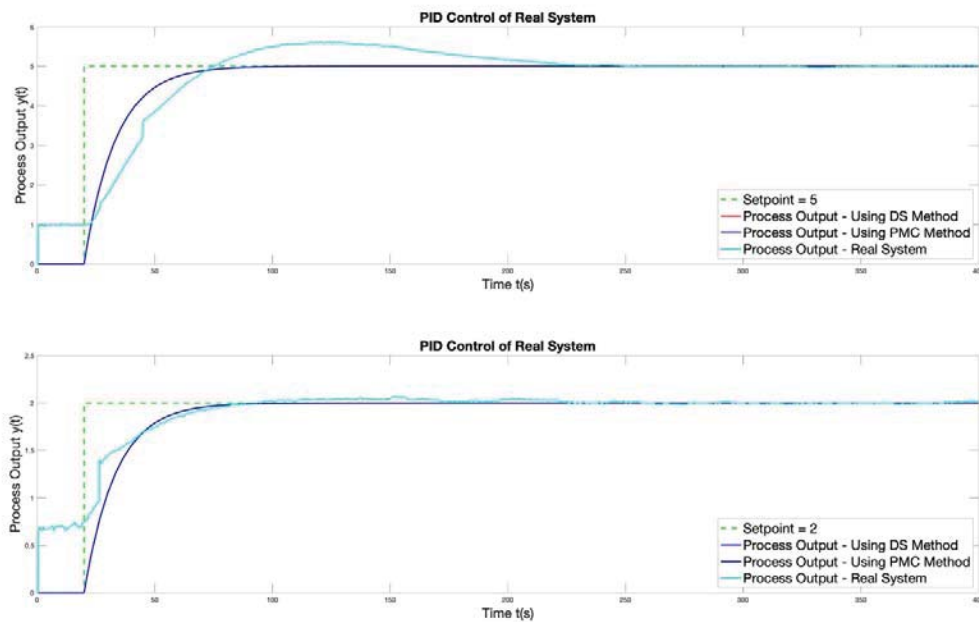


Figure 15 - PID Tuning of Two-tanks system – Simulated vs Real

The figure (15) shows the PID tuning results of the physical setup. As we can see, the control results of the physical system are closer to the simulated results in the second case, when the set-point is 2 (cm) as compared to when the set-point is 5 (cm). This discrepancy can be due to the fact that the theoretical simulation assumes that the system is perfectly linear and doesn't take into account the effects of saturation on the pump output. This results in a large overshoot when the set-point is increased which is a result of integral wind-up effect in the pump. The overshoot can be reduced supplying a setpoint of lower magnitude around a region nearer to the operating point, or by re-simulating the process by taking into account the pump saturation.

6 Conclusion

To sum up the above concepts, we have briefly discussed three important topics namely Optimization, Identification and Control. From section (2), we have evaluated the possibility of using GEA with the afore-mentioned modifications as an effective strategy for non-linear optimization problems. The advantage of this algorithm lies mainly in its inherent simplicity in execution and its versatility in solving multi-modal optimization problems as compared to the other available algorithms which often remain stuck at the local optima rather than finding the global optimum value in a given search-space. The method was later used in section (3) and applied to a task involving relay-based identification of a dynamic system and from the results in section (3.4), it was confirmed to correctly model the

unknown system. Post identification, the process was correctly controlled via simulation using a PID control and with the help of the DS and PMC tuning methods both of which used the identified SOPTD models from identification to effectively control the process. Finally, the results of the complete identification and control using our optimization approach were correctly verified using a real system.

The above methods thus serve as an effective and efficient means of identification and control of unknown processes. Further research into this topic can include application of the GEA algorithm for Model Predictive Control as well as looking into different approaches for optimization like TABU Search [12] or SOMA optimization method or Hybrid Bat Algorithm with Harmony Search [13].

Acknowledgement

This work was supported by the Grant Agency of the Czech Technical University in Prague, grant No. SGS19/158/OHK2/3T/12

REFERENCES

- [1] D. Wolpert and W. Macready, "No free lunch theorems for optimization," *IEEE Transactions on Evolutionary Computation*, vol. 1, no. 1, pp. 67-82, 1997.
- [2] X.-S. Yang, *Nature-Inspired Optimization Algorithms*, Elsevier, 2014.
- [3] G. Wang and L. Guo, "A Novel Hybrid Bat Algorithm with Harmony Search for Global Numerical Optimization," *Journal of Applied Mathematics*, 2013.
- [4] L. Cao, L. Xu and E. Goodman, "A Guiding Evolutionary Algorithm with Greedy Strategy for Global Optimization Problems," *Computational Intelligence and Neuroscience*, pp. 1-10, 2016.
- [5] J. Kennedy and R. Eberhart, "Particle Swarm Optimization," in *Proceedings of ICNN'95 - International Conference on Neural Networks*, Perth, 1995.
- [6] K. Astrom and T. Hagglund, "Automatic Tuning of Simple Regulators with Specifications on Phase and Amplitude Margins," *Automatica*, Volume 20, Issue 5, pp. 645 - 651, 1984.
- [7] V. Ramakrishnan and M. Chidambaram, "Estimation of a SOPTD transfer function model using a single asymmetrical relay feedback test," *Computers & Chemical Engineering*, Volume 27, Issue 12, pp. 1779-1784, 2003.
- [8] J. Berner, T. Hägglund and K. Åström, "Asymmetric Relay Autotuning - Practical features for industrial use," *Control Engineering Practice*, Volume 54, pp. 231-245, 2016.
- [9] "PID Controller," 09 08 2020. [Online]. Available: https://en.wikipedia.org/wiki/PID_controller.
- [10] M. Hofreiter, *Zaklady Automatickeho Rizeni*, Prague: CVUT, 2016.
- [11] D. Chen and D. Seborg, "PI/PID Controller Design Based on Direct Synthesis and Disturbance Rejection," *Industrial & Engineering Chemistry Research*, pp. 4807-4822, 2002.
- [12] F. Glover, "Future Paths for Integer Programming and Links to Artificial Intelligence," *Computers and Operations Research*, pp. 533-549, 1986.
- [13] G. Wang and L. Guo, "A Novel Hybrid Bat Algorithm with Harmony Search for Global Numerical Optimization," *Journal of Applied Mathematics*, 2013.
- [14] L. Ljung, *System Identification: Theory for the User*, Linköping: Prentice Hall, 1987.
- [15] M. Hofreiter, "Alternative Identification Method using Biased Relay Feedback," in *IFAC Papers Online*, Prague, 2018.
- [16] A. Gupta, *Bat Optimization Algorithm*, MATLAB Central File Exchange, 2020.

COMPUTING OF PID PARAMETERS USING TECOMAT FOXTROT AND THE ASYMMETRICAL RELAY SHIFTING

Jakub Vaněk¹, Jan Slabý²

*¹ Department of Instrumentation and Control Engineering, Czech Technical University in Prague
Prague, Czech Republic
Jakub.Vanek@fs.cvut.cz*

*² Department of Instrumentation and Control Engineering, Czech Technical University in Prague
Prague, Czech Republic
Jan.Slaby@fs.cvut.cz*

Abstract: This paper describes the implementation of the relay shifting method for model identification and determining of PID parameters for practical use in process control. For the purpose of this paper, the programmable logic controller Tecomat Foxtrot was used together with the Mosaic development environment. The whole process consists of approximating a controlled system with a second order time delayed model called a SOTD model using the relay shifting method. The PID parameters are then determined from a SOTD model and are used for process control. Emphasis is put on the possibility of simple and effective practical implementation, which enables quick and sufficiently precise determination of the PID parameters for system control. The outcome of this work is the PLC project library together with an easy-to-use function block diagram with a built-in PID controller. The advantage of this method is its versatility, which enables its use even for systems with time delay. A demonstration of this method is given using simulated model as well as real laboratory systems.

Keywords: system control, model identification, PID parameters, autotuning

1. Introduction

Due to rapidly expanding automatization in every industrial sector, great focus is put on fast and sufficiently precise implementation and settings of individual parts of the systems. PID regulators play a large role in industrial automatization. Settings of these regulators might be time consuming and may not offer satisfying results. PID regulators are quite robust, therefore the process can be relatively well controlled, even with nonideally set parameters. However, this might bring extra financial cost, such as higher energy consumption or decreasing lifespan of equipment. For these reasons, it is necessary to conduct further research on the issues of PID controller autotuning. In this paper the application of the relay shifting method for model identification and PID tuning is presented. The method is implemented into PLC Tecomat Foxtrot (Teco a.s, Kolín, Czech Republic).

2. System identification

The first step of the presented method is the identification of the controlled system and its approximation using a SOTD model. Firstly, using the shifting method described in [1], two points of process frequency response $G(j\omega_1)$ and $G(j\omega_2)$ are determined (see Fig. 1). Based on the knowledge of these two points, the second step is the approximation of the system in the form of a SOTD model using the Least Squares Method.

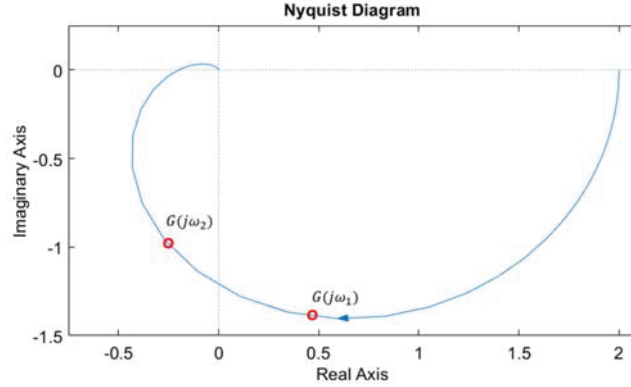


Fig. 1. The Nyquist frequency characteristic with two determined points

2.1 Relay Shifting Method

This method, presented in [1], is based on the use of asymmetrical relay for process control in the proximity of the operating point. A stable oscillation with period T_p after the stabilization time of T_L is prerequisite. Fundamental frequency can be determined by:

$$\omega_1 = 2\pi/T_p \quad (1)$$

and the second harmonic:

$$\omega_2 = 2\omega_1 \quad (2)$$

For the purpose of using this method in real-world systems, T_L is considered to be the time after which 3 subsequent oscillations take place. These meet the following criteria:

- Difference from mean of 3 consequential periods must be lower than 10%.
- Difference from mean of maximal y values from 3 consequential periods must be lower than 10%. Where y is the system output value.
- Difference from mean of minimal y values from 3 consequential periods must be lower than 10%.

The first point is calculated as [1]:

$$G_{(j\omega_1)} = \frac{\int_t^{t+T_p} y(\tau) e^{-j\omega_1 \tau a} d\tau}{\int_t^{t+T_p} u(\tau) e^{-j\omega_1 \tau a} d\tau} \quad (3)$$

The second point is determined as [1]:

$$G_{(j\omega_2)} = \frac{\int_t^{t+T_p} \left(y(\tau) + y\left(\tau - \frac{T_p}{2}\right) \right) e^{-j\omega_2 \tau a} d\tau}{\int_t^{t+T_p} \left(u(\tau) + u\left(\tau - \frac{T_p}{2}\right) \right) e^{-j\omega_2 \tau a} d\tau} \quad (4)$$

Where u is the manipulated variable. The approximating SOTD model is then identified in the form of [1]:

$$M_{(s)} = \frac{K \cdot e^{-s \cdot \tau_d}}{a_2 \cdot s^2 + a_1 \cdot s + 1} \quad (5)$$

Which can be rearranged for non-oscillating systems into the form:

$$M_{(s)} = \frac{K \cdot e^{-s \cdot \tau_d}}{(T_1 \cdot s + 1) \cdot (T_2 \cdot s + 1)} \quad (6)$$

2.2 Computing transition parameters using Least Squares Method

The Least Squares Method is used to determine the parameters K , a_2 and a_1 from the (5) equation. The term $e^{-s \cdot \tau_d}$ element in equation (5) is represented by the Euler formula. By the equation (5) and the points $G(j\omega_1)$ and $G(j\omega_2)$ and the subsequent decomposition into real and imaginary parts, we obtain the equation (7), see [4]:

$$\mathbf{Z} \cdot \boldsymbol{\theta} = \boldsymbol{\zeta} \tag{7}$$

where:

$$\mathbf{Z} = \begin{bmatrix} \cos(\omega_1 \cdot T_d) & R_1 \cdot \omega_1^2 & I_1 \cdot \omega_1 \\ -\sin(\omega_1 \cdot T_d) & I_1 \cdot \omega_1^2 & -R_1 \cdot \omega_1 \\ \cos(\omega_2 \cdot T_d) & R_2 \cdot \omega_2^2 & I_2 \cdot \omega_2 \\ -\sin(\omega_2 \cdot T_d) & I_2 \cdot \omega_2^2 & -R_2 \cdot \omega_2 \end{bmatrix}, \tag{8}$$

$$\boldsymbol{\theta} = \begin{bmatrix} K \\ a_2 \\ a_1 \end{bmatrix}, \tag{9}$$

$$\boldsymbol{\zeta} = \begin{bmatrix} R_1 \\ I_1 \\ R_2 \\ I_2 \end{bmatrix}. \tag{10}$$

R_1, R_2 are real parts of points $G(j\omega_1)$ and $G(j\omega_2)$. And I_1, I_2 are imaginary parts of these points.

The known maximum time delay T_{Dmax} is divided into 500 parts, for each of which a calculation is conducted using the Least Squares Method. T_{Dmax} is estimated as the mean of possible time delays T_{d1}, T_{d2}, T_{d3} (see Fig. 2) increased by 10% to minimize possible error.

$$\boldsymbol{\theta} = (\mathbf{Z}^T \cdot \mathbf{Z})^{-1} \mathbf{Z}^T \cdot \boldsymbol{\zeta} \tag{11}$$

By applying the criterion “Sum of the squared residuals” determined according to:

$$K\tau = \mathbf{E}^T \cdot \mathbf{E}, \tag{12}$$

where:

$$\mathbf{E} = \boldsymbol{\zeta} - \mathbf{Z} \cdot \boldsymbol{\theta}. \tag{13}$$

The desired solution must meet the following conditions: $a_2 > 0$, $a_1 > 0$. These conditions are valid for a stable system.

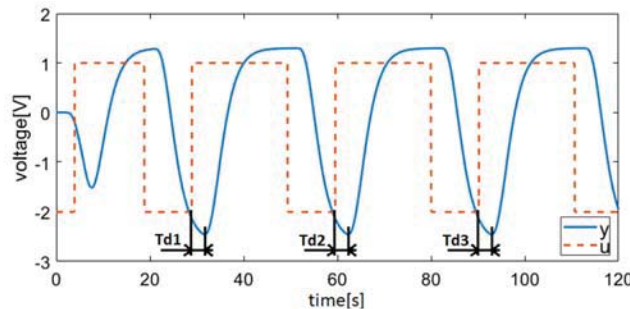


Fig 2. Possible Time delays

2.3 PID Tuning Methods

The identified mathematical model (5) is used to determine the PID parameters of the regulator in use. To ensure versatility, 3 methods of regulator tuning are used. For the purpose of optimization regarding the speed of PLC computing, emphasis is put on simplicity of the calculation. The method described in 3.1 or 3.3 are used for systems

with time delay, based on PID parameter suitability. For systems identified without time delay, the used method is 3.2. The PID controller is described by the formula (25).

2.4 PMC – Phase margin criteria for models with time delay

Based on the identified parameters K, τ_d, a_2 and a_1 , it is possible to determine the parameters of PID regulators for a time delay system in the following manner [2]:

$$r_i = \frac{\pi}{4 \cdot |K| \cdot \tau_d} \quad (14)$$

$$r_0 = a_1 \cdot r_i \quad (15)$$

$$r_d = a_2 \cdot r_i \quad (16)$$

2.5 PMC – Phase margin criteria for models without time delay

This method is suitable for system in form (5) identified without a time delay. Determined parameters are [4]:

$$r_i = \frac{1}{T_w \cdot K'} \quad (16)$$

where T_w is wanted time constant and was chosen as

$$T_w \leq \frac{1}{\left| \operatorname{Re} \left(\frac{-a_1 + \sqrt{a_1^2 - 4 \cdot a_2}}{2 \cdot a_2} \right) \right|} \quad (17)$$

$$r_0 = a_1 \cdot r_i \quad (18)$$

$$r_d = a_2 \cdot r_i \quad (19)$$

2.6 SIMC tuning method

This method is suitable for non-oscillatory time delayed model in form of (6). Determined parameters of PID regulators are [2]:

$$r_0 = \frac{0,5 \cdot (T_1 + T_2)}{K \cdot \tau_d} \quad (20)$$

$$T_i = T_1 + T_2 = a_1 \quad (21)$$

$$T_d = \frac{T_2}{1 + \frac{T_2}{T_1}} \quad (22)$$

where

$$T_i = \frac{r_0}{r_i} \quad (23)$$

$$T_d = \frac{r_d}{r_0} \quad (24)$$

3. Implementation

Implementation is carried out in the Mosaic programming environment for PLC Tecomat Foxtrot. It is divided into three parts. The principles of the first two parts – identification and calculation – are explained in chapters 2 and 3. The third part is feedback control. During programming, emphasis was put on speed of calculation with regard to PLC performance, and also on user-friendliness, specifically on easy implementation into PLC code. The identification and calculation parts are carried out only once. After that, the PLC switches to the control part, where it remains until manual reset.

The user can utilize the PLC program in the form of function block diagram (see Fig. 3), where it is necessary to enter the input parameters.

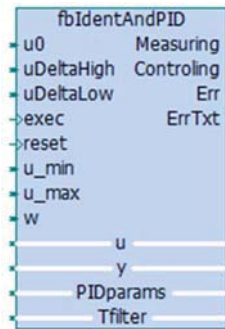


Fig. 3. Function block in Function block diagram

Where:

- u_0 is the input manipulated value for the operating point
- $uDeltaHigh$ and $uDeltaLow$ are relay deviations from the operating point
- u_{min} and u_{max} are limitations for input saturation
- $exec$ starts the program.

The identification part consists of a state machine with three states. In the first phase, resetting of variables is ensured. In the second phase, the stabilized value of the output quantity in the operating point is deducted. The third phase is the measuring itself. The values of the input u and the output y are saved in periods into memory buffers every 20 ms. This sampling period was set as lowest safe time considering the PLC's cycle time. In the event of buffer overflow, the sampling period is automatically adjusted. After the criteria are fulfilled for 3 consecutive periods as described in chapter 2.1, the identification part is complete. With regard to the PLC registry memory, only the last three periods are retained.

In the calculation part, the calculation of points from chapter 2.1, deduction of T_{Dmax} , and subsequently the calculation of the SOTD model using the Least Squares Method from chapter 2.2 take place. After that, the PID parameters for the regulator are determined as described in chapter 3. Since the PLC has a limited computing performance and, at the same time, must control processes in other technologies, the most straightforward calculations were chosen, especially when determining the PID parameters. The constant N for the PID regulator (25) is set to 10 as recommended in [3] and the time constant T_f for the first order filter (26) is then also determined in this part. The filter is put into operation during the start of the control part. It was implemented into the controlling to optimize the course of the output quantity y . After the computing part is completed, buffers and other calculated parameters are uploaded to an SD card for further analysis and the PLC switches to the control part, where inbuilt blocks of the PID regulator and first order filter are utilized.

$$U(s) = r_0 \cdot E(s) + \frac{1}{s \cdot T_i} \cdot E(s) - \frac{r_d \cdot s}{\frac{1}{N} \cdot s + 1} \cdot Y(s) \tag{25}$$

$$M_f(s) = \frac{1}{T_f \cdot s + 1} \tag{26}$$

In Fig. 4, the functional block of the controlling part is depicted using a Simulink scheme.

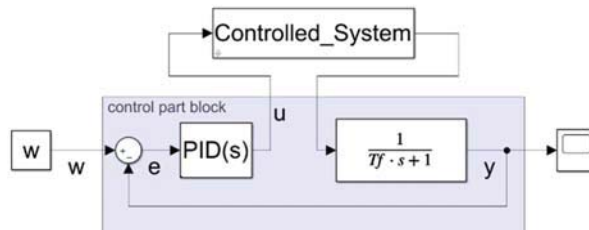


Fig. 4. Controlling part of Function block

4. Experiments

The presented method has been successfully tested on several simulated systems, one of which will be presented here. The tests were also performed using an actual laboratory apparatus.

4.1 Simulated experiment

The test was conducted on a system using the following transfer function:

$$M_{S1} = \frac{2}{(s + 1) \cdot (1.5 \cdot s + 1) \cdot (2 \cdot s + 1)} \quad (27)$$

The system (27) was used in Tecomat Foxtrot simulation mode with results presented in Fig. 5.

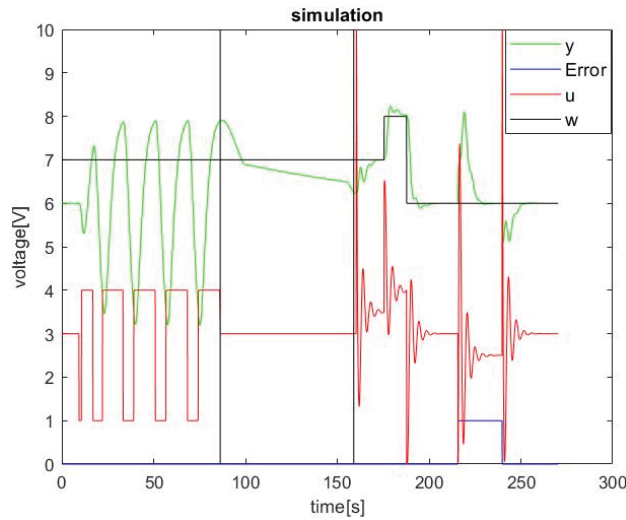


Fig. 5. Course of model identification and process control

Where:

- y is the controlled variable
- $Error$ is the simulated disturbance
- u is the manipulated variable
- w is the desired variable

The process in Fig. 5 is separated into 3 sections by horizontal lines at the times of 86s and 156s. In the first section, system identification using asymmetrical relay takes place. In the second section, the calculation part takes place where two points of frequency response are determined based on the measurement from the first section, then the model in the form of (5) is calculated. Based on this model, the PID parameters are determined. In the third section, feedback process control takes place. After the stabilization of y on the desired variable w at the time of 175s, the desired variable w increases from 7 Volts to 8 Volts and, at the time of 188s, it decreases to 6 Volts. At the time of 216s, the disturbance Err will appear for the following 20 seconds. It is evident that the variable y is controlled successfully as it always stabilizes at the desired variable w .

The transfer function (28) was identified as an approximation of the simulated system.

$$M_{(j\omega)} = \frac{1.91 \cdot e^{-j \cdot \omega \cdot 0.48}}{4.61 \cdot (j \cdot \omega)^2 + 3.76 \cdot j \cdot \omega + 1} \quad (28)$$

The original transfer function (27) is then compared with its identified approximation (28) in Fig. 6.

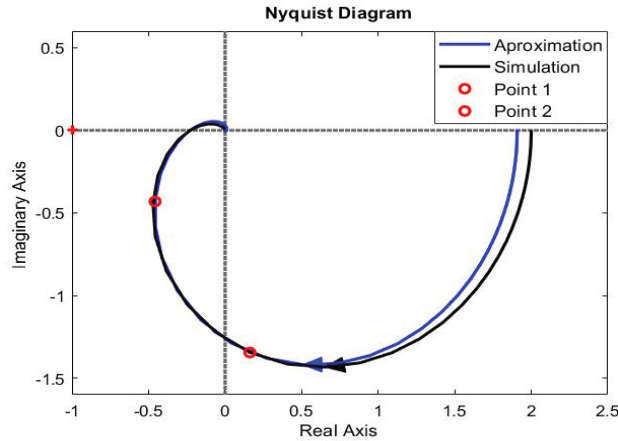


Fig. 6. The Nyquist frequency characteristic comparison for simulated system and its approximation.

According to Fig. 6, it is evident that the approximation (28) describes the simulated system (27) precisely enough and it is possible to use this approximation for PID parameter determination.

The PID parameters were calculated based on calculations described in chapter 3. Based on equations (14), (15) and (16) the results are: $r_0=3.23$ $r_i=0.86s^{-1}$ and $r_d=3.95s$.

4.2 Experiments produced on real systems

For testing purposes, the laboratory apparatus called “air aggregate” (see Fig. 7) was used in two different settings.

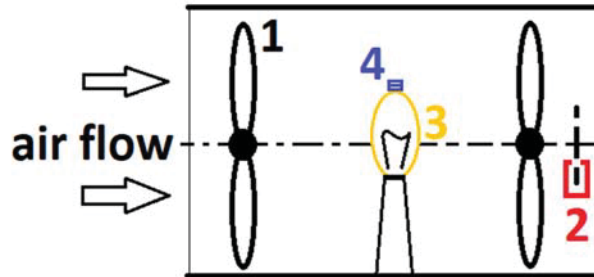


Fig. 7. The laboratory apparatus “Air aggregate”. 1-fan, 2-flow meter, 3-light bulb, 4-thermistor

For the first test, the airflow setting was used where the manipulated variable u is the voltage on the fan and the controlled variable y is the voltage on the flow rate meter. For the second test, the heating setting was used where the manipulated variable u is the voltage on the light bulb and the controlled variable y is the voltage on the thermistor. The fan is used as a disturbance for the second test.

In Fig. 8, the courses of presented method used on air aggregate in airflow mode are presented. The controlled variable y is the voltage on the flow meter, $Error$ is the simulated disturbance on variable y , u is the manipulated variable and w is the desired variable. The process in Fig. 8 is separated into 3 sections by horizontal lines at the times of 236s and 309s. In the first section, system identification using asymmetrical relay takes place. In the second section, the calculation part takes place where two points of frequency response are determined based on the measurements from the first section, then the model in the form of (5) is calculated. Based on this model, the PID parameters are determined. In the third section, feedback process control takes place. The identification section takes longer than during the simulated experiment because of the conditions described in section A of chapter III. were met after 12 periods. It is apparent that in the feedback control section, the system was controlled successfully.

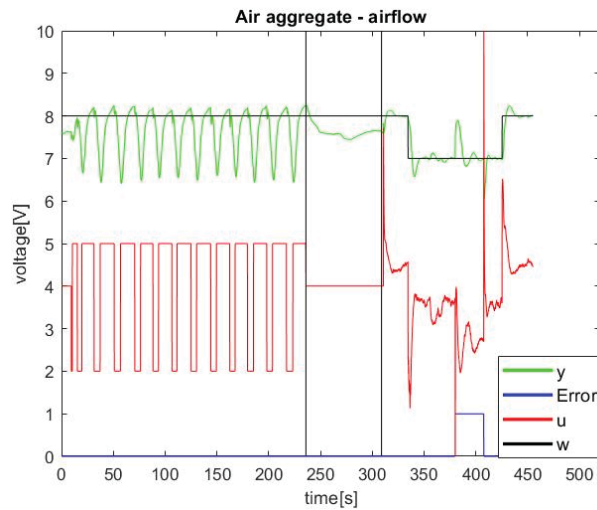


Fig. 8. Course of model identification and process control of the air aggregate with airflow setting.

Results of the test carried out using the Air aggregate laboratory apparatus with heating settings are presented in Fig. 9.

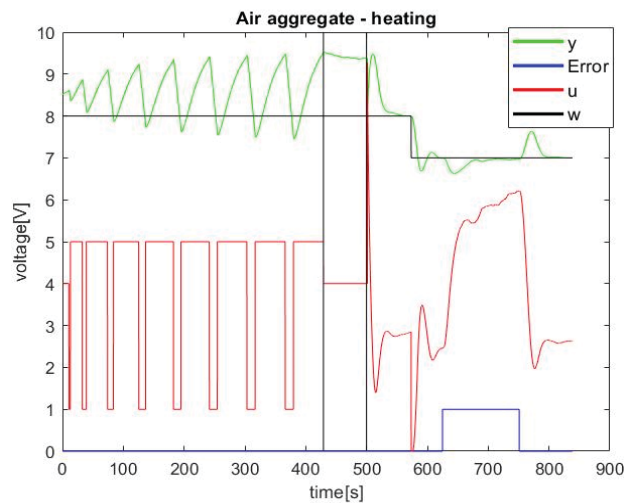


Fig. 9: Course of model identification and process control of the air aggregate with heating setting

The controlled variable y is the voltage on the light bulb, $Error$ is the voltage on the fan creating disturbance, the manipulated variable u is the voltage on the thermistor and w is the desired variable. From Fig. 9 it is evident that the apparatus was successfully identified and controlled.

The duration period is affected by system speed and process stability. The simulated system was identified fastest because the system is not burdened by disturbances, therefore the conditions for measuring termination described in section A of chapter 2. were met immediately.

5. Conclusion

Although it is necessary to further test the procedure before implementation in practice, as well as complete the error handling system, it is possible to state that this procedure has potential for practical use. Tests were conducted on a simulated system of 3rd order and on two actual laboratory apparatuses, where the system was successfully identified, the PID parameters were determined and, subsequently, the system was successfully controlled. The next focus will concentrate on testing and the eventual changes in determining PID parameters to ensure the highest possible versatility of use. Time will also be given to ensure automatic settings of sampling time regarding PLC

cycle time. The next steps for accuracy improvement and higher versatility will be carried out in the identification part.

Acknowledgment

This work was supported by the Grant Agency of the Czech Technical University in Prague, grant No. SGS19/158/OHK2/3T/12.

References

- [1] M. Hofreiter, "Improved relay feedback identification using shifting method," in Proceedings of the 16th International Conference on Informatics in Control, Automation and Robotics (ICINCO 2019). Madeira: SciTePress, 2019. p. 601-608. ISBN 978-989-758-380-3.
- [2] A. Hornychová a M. Hofreiter. Use of the shifting method results for PID controllers parameters estimation. MATEC Web of Conferences. 2019, 292, 01017. DOI: 10.1051/mateconf/201929201017.
- [3] M. Vítečková and A. Víteček, "Vybrané metody seřizování regulátorů," Vysoká Škola Báňská - Technická Univerzita Ostrava, 2011.
- [4] M. Hofreiter. Personal consultation, author of The relay shifting method 2020.

FINITE ELEMENT ANALYSIS OF AN ENDOSCOPE DURING OPHTHALMIC SURGERY

Jarno Mastomäki, Jan Hošek

*Ústav přístrojové a řídicí techniky, ČVUT v Praze, Fakulta strojní, Praha
Czech Technical University in Prague, Faculty of Mechanical Engineering*

Abstract: This thesis work was assigned by Czech Technical University in Prague, further CTU. It is a leading technical University in Czech Republic, located in the country capital, Prague. The purpose of this study is to investigate the stresses and displacements in human eye and instrument during endoscopic procedure. The results of the calculations will be used on developing the instrument for ophthalmic operations. In the process, the eye is penetrated by a needle-like endoscope to visually study the inner parts of the eye before or during the actual surgical operation. The need of further knowledge, especially in the behavior of the eye and the instrument, is still gathered. Because of fairly complicated nature of the interaction between viscoelastic human tissue and metallic material. Recommendations for future research on this field are made.

Keywords: ophthalmology, endoscopy, FEA, biomechanical engineering, optomechanics

1 Introduction

The story of ophthalmic applications of endoscopy can be dated time between the great world wars. In year 1934 procedure was first described by Dr. Harvey E. Thorpe, a native Latvian who immigrated to US in Pittsburgh in 1906. According to Pittsburgh Ophthalmology Society [1], Thorpe received his bachelor degree from electrical engineering and did his post-graduate work in ophthalmology in New York City. As a true interdisciplinary, Thorpe designed an instrument which combines Galilean telescope and an illumination source. The whole instrument shaft (6.5 mm) [2] was inserted to eye thru 8 mm scleral incision.

Thorpe's design was revisited in 1978 with slight modifications by J.L Norris & G.W Cleasby. These two gentlemen, mentioned previously, improved [3] the original design with smaller rigid shaft (1.7 mm). In 1990, V.V Volkov et al. [4] introduced three types of flexible ophthalmic endoscopes which could be used directly to examine the interocular structures. Soon after the laser units were implemented to the endoscopes [5] and therefore, made possible to proceed surgical operations with the same instrument. From that point the prototypes had developed rapidly. All the components had been optimized and the image resolution had improved significantly [6].

Late in the autumn 2019 Dr. Hošek suggested me this topic because of my earlier knowledge and interest for finite element analysis. They had a problem with ophthalmic endoscope in developing. They wanted to improve the endoscope structure by making it more compact and provide better resolution image. During this process they faced a problem with the mechanical endurance of the optical part of the endoscope. Therefore, there was a need for study how much the critical parts can hold a load during the operation. This is done in order to optimize the design without having to compromise between the wanted qualities and strength.

2 Review of the study

Because of the fairly complicated mechanical system of mechanical tool which is interacting with organic tissues is the issue. This makes the study particularly interesting in a simulation point of view. Finite element method had been increasingly applied to biomechanical problems since late nineties. In twenty years the most successful biomechanical applications have been in the field of fracture mechanics of the human bones. These studies are done in the field of fracture mechanics. Mainly because of the solid material knowledge, the method has been successfully applied. These studies have helped researchers to understand the human bone fracture mechanism. On the contrary, the interaction of nonbiological and biomechanical structures has been studied far less. Main reason for this has been the limited computational power. Increasing computational capacity and improved algorithms allow us to study more and more complicated systems without sacrificing the accuracy to the simplifications.

For modeling the operational use of an ophthalmic endoscope, we are using the finite element method. It is a numerical method for solving partial differential equations in boundary value problems. In order to solve extremely complicated systems and the interactions between different bodies, this method subdivides a large system to a finite number of small elements. These elements are solved individually and they finally form a system of algebraic equations. The simple equations that model these finite elements are then assembled into a larger system that covers the entire problem.

The main target of this study is to study the stresses and displacements during ophthalmic surgery. The results of this study could be used for the development of the instrument for ophthalmic operations. The interest is, especially, in the optical part of the instrument and its capability to withstand external loading. The final output is to find the simplest model as possible with the highest possible accuracy.

2.1 Ophthalmic surgery

One of the main problems in this study is that it is controlled directly by the surgeon's hand and therefore the range of operational loading is wide and hard to define. During the ophthalmic operation, the surgeon needs to move the endoscope inside the eye in all directions. Even though the human eye is a very fragile organ and therefore requires extreme care during the surgical procedure, there might occur some forces high enough to damage the weakest parts of the instrument structure.

Although the ophthalmic endoscopes have been around for 1934. According to Marra, K et al, the popularization of ophthalmic endoscopy has been promoted by recent technological advancements that increased the number of indications for endoscopy [6]. Dr. Martina Nemčoková who is a doctor of Ophthalmology from Charles University said that according to their experience the endoscopic examinations are mainly done for infants where, because of natural reasons, every move is done with high care [7]. Even though that the endoscopic view provides two fundamental advantages it is a relatively rarely used tool in ophthalmological operations. These advantages are: firstly, it reaches areas which are normally not so easy to examine because it passes anterior segment and it eases the visualization of anterior structures such as ciliary bodies and sub-iris space. Together these two advantages provide a steady base for surgery. From the following figure (1) we can see the scheme of ophthalmic operation with an endoscope.

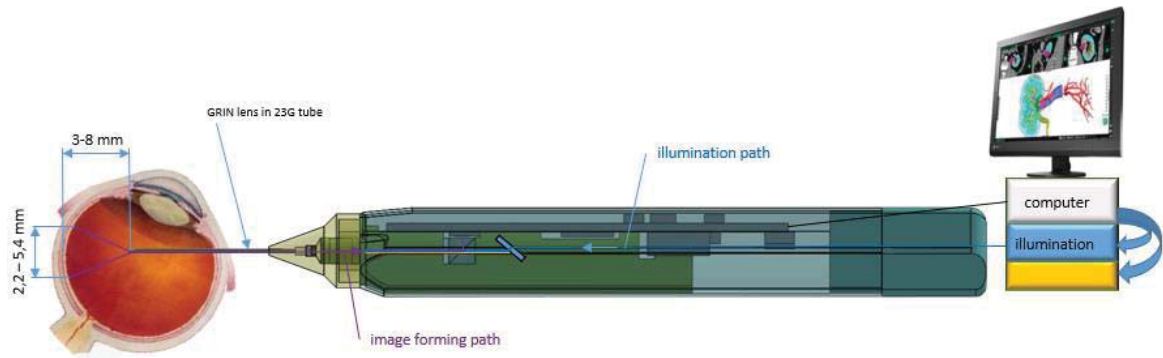


Figure 1. Schematic picture of endoscopic examination [8]. (Němcová, Š, Presentation. Modified)

2.2 Finite Element Model

In this study, we will create a model and we increase the model complexity in phases by increasing the number of non-linearities until we meet the measured values. We will compare the results between several models and subsequently our results will be compared between each other and with the results of similar cases, previously published in the literature. The model will be executed as a multibody model which means that part of the boundary conditions is in the interactions between the different bodies. The definition of these boundary connections will be found by iterative means. The most critical one is the interaction between the endoscope needle and the combined body of sclera, choroid, and retina. In this study, we are simplifying these previously mentioned parts into one and assume that they act as one unit under applied stress.

From the figure (2) we can see the CAD-model of the eye we are going to use in the analysis. From the table (1) we can see the mechanical properties of each body in the model.

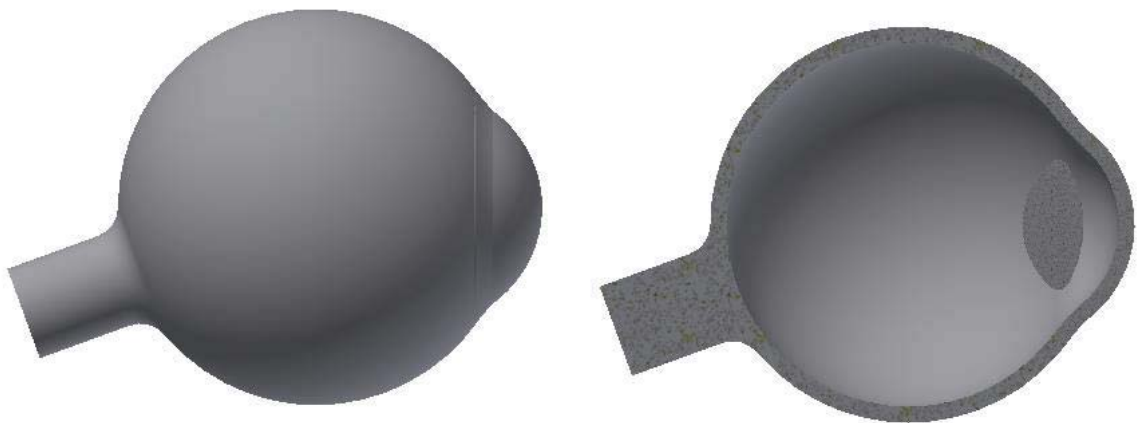


Figure 2. Eye model

Table 1. Mechanical properties of the eye model.

Tissue	E (Pa)	Poisson's ratio	C10 (Pa)	D1 (Pa ⁻¹)	k (m/s)	Source(s)
Choroid	6,00E+05	0,49	100671,1	2,00E-07	5,00E-13	[9,10]
Cornea	2,90E+05	0,42	51056,34	3,31E-06	2,20E-11	[11,12,16]
Lens	8,20E+05	0,47	139455,8	4,40E-07	1,10E-12	[13,14]
Sclera	2,35E+06	0,47	399659,9	1,50E-07	6,37E-11	[9,12,18]
Aqueous humour	4,20E+04	0,49	7047	2,90E-06	1,00E-03	[15]
Vitreous humour	4,20E+04	0,49	7047	2,90E-06	1,00E-03	[15]

In the following figure (3) is named all the major parts of the human eye. Dimensions for annotations presented in the figure can be found in the table (2).

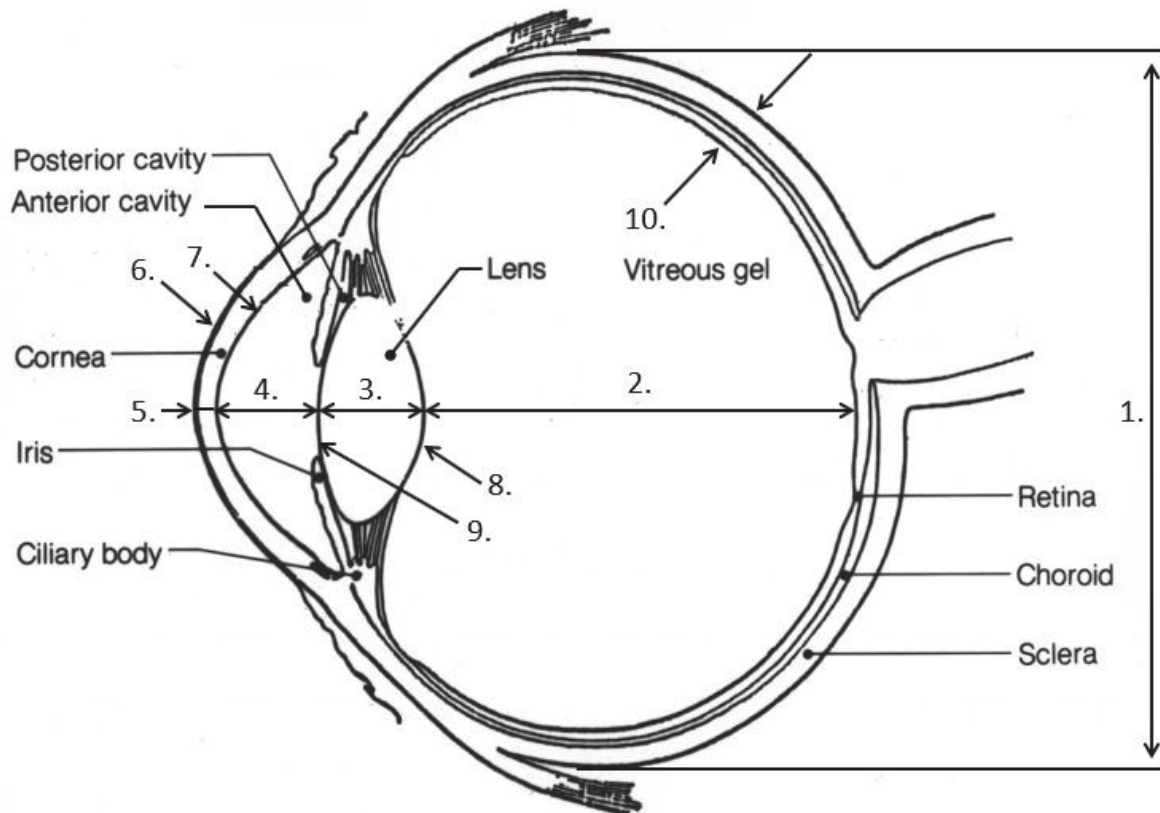


Figure 3. Main parts of human eye [19] (Clusters. E, home page, modified)

In the following table (2) we are presenting the characteristic dimensions of the eye. The CAD-model is based on the following values.

Table 2. Characteristic dimensions of human eye.

Annotation	Dimension	Unit	Source
1.	24	mm	[20]
2.	16	mm	[20]
3.	3,5	mm	[20]
4.	3,3	mm	[20]
5.	0,5	mm	[20]
6.	7,7	mm	[20]
7.	6,8	mm	[20]
8.	6	mm	[20]
9.	11	mm	[20]
10.	0,94	mm	[21]

In the table parameters C_{10} and D_1 are representing the neo-Hookean constants derived from the Young's modulus and Poisson's ratio data. Constant k represents the hydraulic permeability, which is a constant connected to the viscoelastic material model to capture the liquid flow in the porous material. Calculation method for constants C_{10} and D_1 are presented after next chapter.

Mooney-Rivlin models were introduced by Melvin Mooney in 1940 whom theorem was expressed in terms of invariants by Ronald Rivlin in 1948 [22,23]. It is important to realize that the Mooney-Rivlin model does not give any specific insight into the material behavior. Mooney-Rivlin is a curve fits of polynomial to test data. The numerical values of coefficients are results from the curve-fits. These coefficients are given to the FE analysis and they define how stiff the material is under the applied load. Hyperelastic models are defined thru a strain energy potential. Sometimes term strain energy function is used. strain energy potential is commonly denoted as (W) and it can be expressed as a function of the main strain invariants as following,

$$W = W(I_1, I_2, I_3). \tag{1}$$

Equation 1. Strain energy potential as a function of invariants [24]

If we assume that the material is isotropic then the I_3 term always equals 1. This is because when the $\varepsilon = 0$ the $\lambda = 1$. This is caused by the initial assumption of isotropic material. From following equation, we can see why the third invariant is one,

$$I_3 = \lambda_1^2 \cdot \lambda_2^2 \cdot \lambda_3^2 = \left(\frac{V}{V_0}\right)^2. \tag{2}$$

Equation 2. Third invariant [24]

In equation (3) the $\frac{V}{V_0}$ represents the volume ratio. The rest of the invariants are expressed as following (3 & 4),

Jarno Mastomäki

$$I_1 = \lambda_1^2 + \lambda_2^2 + \lambda_3^2, \quad (3)$$

Equation 3. First invariant [24]

$$I_2 = \frac{1}{\lambda_1^2} \cdot \frac{1}{\lambda_2^2} \cdot \frac{1}{\lambda_3^2}. \quad (4)$$

Equation 4. Second invariant [24]

As we can see from the previous equations (2,3 & 4) the main invariants are defined by stretch ratios (λ) which are the ratios of initial lengths in the local principal directions (1, 2, 3). The stretch ratio can be expressed by the help of the relative strain, commonly known as an engineering strain.

$$\lambda = 1 + \varepsilon_{eng} \quad (5)$$

Equation 5. Strech ratio expressed with the help of engineering strain [25]

In the previous equation (6) the sub note (*eng*) indicates the engineering i.e. relative strain. The polynomial form is based on the first and second invariants and its general form can be seen in following,

$$W = \sum_{i+j}^N C_{ij} \cdot (I_1 - 3)^i \cdot (I_2 - 3)^j + \sum_{k=1}^N \frac{1}{D_1} \cdot (J_{el} - 1)^{2k} \quad (6)$$

Equation 6. Strain energy potential in a common form for two term Mooney-Rivlin. [24,25]

Even though the majority of the material models are to be taken from literature, we need to make some material testing to cover our assumptions. Because we are assuming combined bodies in the eye main structure we need to define the combined material properties. This can be done by using nanomechanical testing. This method is mainly used when defining elastic properties for biomaterials and soft tissues. The method is called quasistatic nanoindentation and it is usually used to define the common elastic properties like Young's modulus and Poisson's ratio. From those measured properties we can calculate the visco- and hyperelastic Mooney-Rivlin's two first constants and compare them to values previously published in related literature. The method of calculations is presented in the following formulas (1 & 2).

$$C_{10} = \frac{\mu}{2} \quad (7)$$

Equation 7. constant C_{10} first Mooney-Rivlin term [26]

$$D_1 = \frac{\lambda}{2} \quad (8)$$

Equation 8. constant D_1 Second Mooney-Rivlin term [26]

Jarno Mastomäki

Where λ and μ are called Lamé coefficients. They can be derived from bulk modulus by using following formula.

$$B = \lambda + \frac{2}{3} \cdot \mu \quad (9)$$

Equation 9. Bulk modulus writed in terms of Lamé first and second coefficient [26].

Where B can be written in a terms of Young's modulus and Poisson's ratio as following.

$$B = \frac{E}{3(1 - 2 \cdot \nu)} \quad (10)$$

Equation 10. Bulk modulus written as a function of Young's modulus and Poisson's ratio [26].

Also Lamé's first parameter can be re-written by Young's modulus and Poisson's ratio.

$$\lambda = \frac{E \cdot \nu}{(1 + \nu) \cdot (1 - 2 \cdot \nu)} \quad (11)$$

Equation 11. Lamé's first parameter written as a function of Young's Modulus and Poisson's ratio [26].

By using these previous formulas we can solve the Lamé's second parameter as following,

$$\Rightarrow \mu = \frac{3 \cdot (B - \lambda)}{2}. \quad (12)$$

3 Future development

In the future, we will proceed with the simulation of the endoscope-eye interaction. We are expecting to found that defining the interactions of the different bodies is fairly complicated and time-consuming. There are not much-verified results of similar cases available and therefore we need to rely on specialist dictums and our own perceptions when it comes to determining that boundary conditions are working correctly. The usage of the PHE (porohyperelastic) material model is still a discretionary option. We are expecting to found sufficiently accurate results by using the standard neo-Hookean constitutive approach in material models. For model verification, we are going to use pigs eye. According to Dr. Nemčoková, this is structurally really close to the human eye. According to her, pig eyes are widely used in University teaching when the purpose is to mimic the human eye therefore it shall be suitable in this application too. Model verification will be done by comparing the measured deformation to simulate one.

Acknowledgement

This work was supported by the Grant Agency of the Czech Technical University in Prague, grant No. SGS20/055/OHK2/1T/12.

References

- [1] Harvey E. Thorpe, MD: P.O.Society 2019. Jan 2020. Available: www.pghoph.org.
- [2] Thorpe H. *Ocular endoscope: instrument for removal of intravitreal nonmagnetic foreign bodies*. Trans Am Acad Ophthalmol Otolaryngol. 1934; 39, pp. 422–424
- [3] Norris J. L & Cleasby G.,W. *An endoscope for ophthalmology*. Am J Ophthalmol. 1978; 85, pp.420–422
- [4] Volkov V.V, Danilov AV, Vassin LN, Frolov YA. *Flexible endoscope for intraocular surgery*. Arch Ophthalmol. 1990; 108 pp. 1037–1038.
- [5] Uram M. *Ophthalmic laser microendoscope endophotocoagulation*. Ophthalmology. 1992; 99, pp. 1829–1832.
- [6] Marra K. V., Yonekawa, Y., Papakostas, T. D., & Arroyo, J. G. (2013). *Indications and techniques of endoscope assisted vitrectomy*. Journal of ophthalmic & vision research, 8(3), pp. 282–290.
- [7] Nemčoková M. MD, *Ophthalmological surgery*. 2020. Interview on 24.1.2020. Interviewer Mastomäki, J. Prague
- [8] Němcová Š. Ing. PhD. 2019. CTU in Prague. Faculty of Mechanical Engineering. Presentation. Requires access right.
- [9] Friberg T.R, Lace J.W. *A comparison of the elastic properties of human choroid and sclera*. Exp Eye Res. 1988; 47, pp. 429–436.
- [10] Woo S.L, Kobayashi A.S, Schlegel WA, Lawrence C. *Nonlinear material properties of intact cornea and sclera*. Exp Eye Res. 1972; 14, pp. 29–39.
- [11] Hamilton K.E & Pye D.C. *Young's modulus in normal corneas and the effect on applanation tonometry*. Optom Vis Sci. 2008; 85 pp.445–450.
- [12] Uchio E, Ohno S, Kudoh J, Aoki K & Kisilewicz L.T. *Simulation model of an eyeball based on finite element analysis on a supercomputer*. Clinical Science. 1999; 83 pp. 1106–1111.
- [13] Fisher R.F. *Elastic constants of the human lens capsule*. J Physiology. 1969; 201, pp. 1–19
- [14] Fisher R.F & Wakely J. *The elastic constants and ultrastructural organization of a basement membrane (lens capsule)*. Proc R Soc Lond B Biol Sci. 1976; 193, pp 335–358.
- [15] Cirovic S, Bholra R.M, Hose D.R, Howard I.C, Lawford P.W & Parsons M.A. *A computational study of the optic nerve evulsion*. IUTAM Preceedings on Impact Biomechanics: From Fundamental Insights to Applications. 2005, pp. 469–476.
- [16] Overby D, Ruberti J, Gong H, Freddo T.F & Johnson M. *Specific hydraulic conductivity of corneal stroma as seen by quickfreeze/ deep-etch*. J Biomech Eng. 2001; 123 pp. 154–161.
- [17] Johnson M & Erickson K. *Mechanisms and routes of aqueous humor drainage*.
- [18] Jackson T.L, Hussain A & Hodgetts A. *Human sclera hydraulic conductivity: age-related changes, topographical variation, and potential scleral outflow facility*. Invest Ophthalmol. 2006, 47, pp. 4942–4946.
- [19] Custers E. 2017 *Anatomy of the Eye: Human Eye Anatomy*. Jan 2020 Available: www.owlcation.com/stem/Anatomy-of-the-Eye-Human-Anatomy.
- [20] Atchison, David A. and George Smith. 2000. *Optics of the Human Eye*. 1.th ed. Oxford: Butterworth-Heinemann.
- [21] Vurgese, S., Panda-Jonas, S., & Jonas, J. B. 2012, Jan 2020 *Scleral thickness in human eyes*. PloS one, 7(1), Available: <https://doi.org/10.1371/journal.pone.0029692>
- [22] Mooney, M., 1940, *A theory of large elastic deformation*, Journal of Applied Physics, 11 (9), pp. 582–592.
- [23] Rivlin, R. S., 1948, *Large elastic deformations of isotropic materials. Further developments of the general theory*, Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences, 241(835), pp. 379–397.
- [24] Ansys, Inc. 2005, *Hyperelastic Materials*, Handbook. Jan 2020 Available: http://read.pudn.com/download222/doc/1047742/modelling_simulation/nonlinear-materials-tech+++pdf
- [25] McGinty, B. 2012. Homepage. *Mooney-Rivlin Models*. Jan 2020, Available: <https://www.continuummechanics.org/mooneyrivlin.html>
- [26] Pence T & Gou K. 2014. *On compressible versions of the incompressible neo-Hookean material*. Mathematics and Mechanics of Solids. Available: <https://doi.org/10.1177%2F1081286514544258>

DISCRETE SIR MODEL FOR SIMULATION OF VARIOUS COVID-19 SCENARIOS

*Prem Wongsagoon*¹, *Matous Cejnek*¹

¹ *DICE, FME, CTU, Technicka 4, Praha 6, Czechia, prem.wong@hotmail.com*

Abstract: This project features SIR model to model COVID 19 situation development as a function of various social environment conditions (government restrictions, social distancing, etc.). Such a simulation can help people to understand the strict should be to restrictions to manage the COVID 19 situation development in reasonable way. Simulation are created via Python programming language. The findings of this research shows, that the social distancing and similar concepts could influence the situation in a significant way.

Key words: SIR, COVID 19, Python, social distancing

1 Introduction

There have been a pandemic so called COVID-19 [1] virus since the beginning of 2020. Countries all around the world were decided to lock-down to stop spreading of the virus. In order to understand how behavior of societies affect the infection, we use SIR model to describe.

SIR model [2] is compartmental models simplify the mathematical modelling of infectious diseases. The population is assigned to compartments with labels - for example, susceptible, infective, or remove. People may progress between compartments.

2 SIR Model

The SIR model itself was origin in the early 20th century, with an important work being that of Kermack and McKendrick in 1927. The models are most often run with ordinary differential equations (which are deterministic), but can also be used with a stochastic (random) framework, which is more realistic but much more complicated to analyze. The discrete SIR model can be described by following the equations below. To find susceptibles rate in day t as follows

$$S_{t+1} - S_t = -bS_t I_t, \quad (1)$$

where b is the infection rate. The number of infectives can be obtained as follows

$$I_{t+1} - I_t = bS_t I_t - aI_t, \quad (2)$$

where the a is the recovery rate parameter. And the number of removed for every day is calculated as follows

$$R_{t+1} - R_t = aI_t. \quad (3)$$

2.1 Calculation example

Consider initial population: $S_t = 10000$, $I_t = 1,000$, $R_t = 0$; and model parameters: $a = 0.05$ and $b = 0.0001$. The first steps will look as follows:

t	S_t	I_t	R_t	bS_tI_t	aI_t	S_{t+1}	I_{t+1}	R_{t+1}
0	10,000	1,000	0	1,000	50	9,000	1,950	1,950
1	9,000	1,950	50	1,755	97.5	7,242	3,607.5	147.5
3	7,242	3,607.5	147.5

2.2 Python implementation

Python [3] is an interpreted, high-level, general-purpose programming language. Its language constructs and object-oriented approach aim to help programmers write clear, logical code for small and large-scale projects.

In this work the Python is used to simulate the SIR model with various conditions. To simplify the calculations, the Numpy package [4] is used. The Matplotlib package [5] was used for creation of graphs.

3 Results

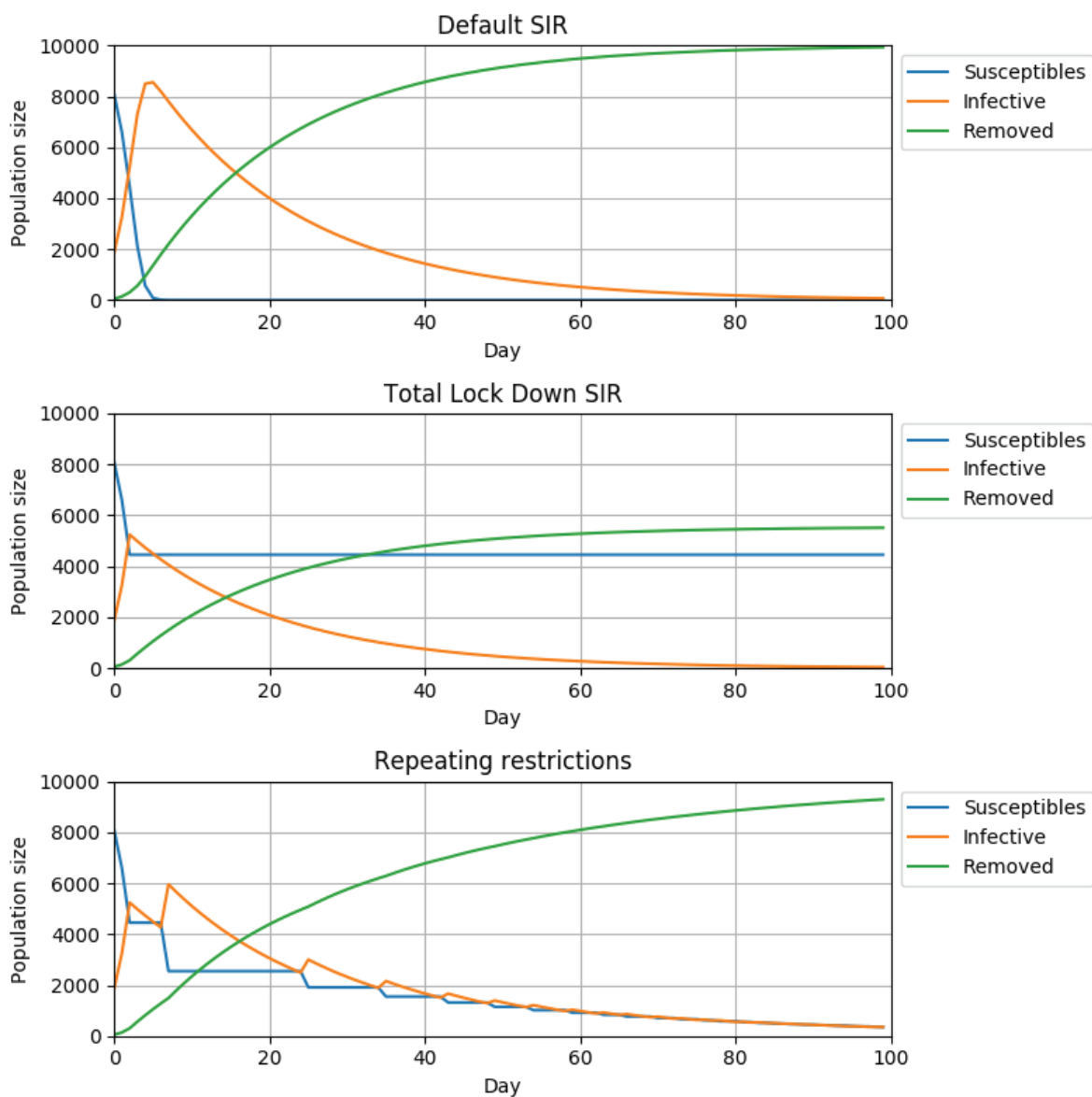


Fig. 1: Simulated development with various settings, from top: a) SIR model with no modification; b) SIR model with total lock-down at day 3; c) Mild restrictions triggered by situation worsening.

For all simulation we used the same following initial conditions: $S_t = 9,000$, $I_t = 1,000$, $R_t = 0$, $a = 0.05$,

$b = 0.0001$. Fig.1 shown results of all simulations. See from top: a) SIR model with no modification; b) SIR model with total lock-down at day 3 - the infection rate is reduced to 0; c) Mild restrictions triggered by situation worsening - if the number of infected increases the infection rate is reduced by factor of 90%.

Conclusion

SIR model was used to describe scenario of diseases infection by using mathematical modelling. While Python was used to simulate graph in different scenarios. By combining these together, it helps to clarify and understand more about how human behaviors trend of infection.

To conclude, It will take several days for this pandemics to decay especially high population countries. In meanwhile, thoughtfulness and awareness behaviors help this pandemic situation getting better such as lock-down, social distancing and wearing mask.

Acknowledgement

This work was supported by the Grant Agency of the Czech Technical University in Prague, grant No. *SGS18/177/OHK2/3T/12*.

Literature

- [1] Zhe Xu, Lei Shi, Yijin Wang, Jiyuan Zhang, Lei Huang, Chao Zhang, Shuhong Liu, Peng Zhao, Hongxia Liu, Li Zhu, et al. Pathological findings of covid-19 associated with acute respiratory distress syndrome. *The Lancet Respiratory Medicine*, 2020.
- [2] Howard Howie Weiss. The sir model and the foundations of public health. *Materials matematics*, pages 0001–17, 2013.
- [3] Guido Van Rossum and Fred L. Drake. *Python 3 Reference Manual*. CreateSpace, Scotts Valley, CA, 2009.
- [4] Travis E Oliphant. *A guide to NumPy*, volume 1. Trelgol Publishing USA, 2006.
- [5] John D Hunter. Matplotlib: A 2d graphics environment. *Computing in science & engineering*, 9(3):90–95, 2007.

VIDEO_FLAWER: REALISTIC VIDEO DISTURBER FOR ARTIFICIAL VIDEO

Yang Hong-Bin¹, Matous Cejnek²

¹ FEL, CTU, Technická 2, Praha 6, Czechia, petingo.yang@gmail.com

² DICE, FME, CTU, Technická 4, Praha 6, Czechia

Abstrakt:

Klíčová slova:

Abstract: Although people try hard to minimize defects in a video most of the time, sometimes it is needed to augment or simulate them for some reasons. We created a program which can add realistic effects to videos, including camera shake, auto focus blur, and noise simulation.

Keywords: noise, synthetic data, digital video processing

1 Introduction

More or less, videos always come with defects. In most cases, people try their best to eliminate them or minimize the affect. Camera and lens manufacturers develop optical image stabilizer to reduce the vibration, continuously updating their noise reduction algorithm in order to attract consumers. Film makers use tripods to shot steady videos, adjusting the parameters in order to decrease the ISO sensitivity. However, in some cases, it is desired to add those defect to the video. For example, if a director wants to make a filmed scene more vivid, adding camera shake effect on it may make audiences feel like they are in the scene personally. Another common example is old video simulation where different kinds of noise is added. Or, as the title of this paper, it is possible to disturb an video generated by computer graphic technology, making it more realistic. This disturber may have further application, which is generating training data for deep learning-based stabilization or denoising method.

In this paper is proposed new python package - *video_flawer* - to create artificial noise-like effects in video records. Three major effects is implemented: camera shake, auto-focus blur, and noise simulation. The program is done by Python with the help of OpenCV and Numpy.

2 Effects

2.1 Camera Shake

The first effect is camera shake. This effect appears in almost every video which is shot by a hand-held camera due to the inevitable minor movement of human arms. That is to say, given a steady video (perhaps a tripod shot), our program can convert it into one with artificial vibrations as if it is recorded by a hand-held camera. To accomplish this, the movement is first broken down into four parts: left-right (x -axis), up-down (y -axis), front-back (z -axis), and rotation (r). Moreover, movement in each axis can be further split into rhythmic movement and random movement. By simulating them separately as a 1-dimensional action, a smooth and realistic camera shake effect is created.

For the four axes, identical formulas with different parameters are used to determine the quantity of the simulated movement. Commencing with rhythmic movement, the following equation is used:

$$qh_i = a \cdot \sin(f * i) \quad (1)$$

where qh_i denotes the quantity of rhythmic movement of frame number i , while a and f controls amplitude and frequency respectively. In implementation, a and f are variables which can be altered by users based on their

preferences. On the other hand, an algorithm which composes a wave with inconstant amplitude and frequency is developed to simulate the quantity of random movement. This algorithm takes five parameters as input, including total frame amount, minimum value of amplitude, maximum value of amplitude, minimum value of wavelength, and maximum value of wavelength. The program would generate waves one by one randomly based on the parameters and concatenate them altogether until every frame is assigned a value qr_i .

With qh_i and qr_i , total movement quantity q_i can be obtained by simply summing them up as:

$$q_i = qh_i + qr_i \quad (2)$$

For each axis, a unique sequence q is generated, and a corresponding affine transformation matrix is then built to express either translation, scale operation, or rotation. In the following equations, $M_{(a,i)}$ and $q_{(a,i)}$ denotes the transformation matrix and movement quantity of axis a , frame number i respectively. For x - and y -axis, the following matrixes are used to express translation:

$$M_{(x,i)} = \begin{bmatrix} 1 & 0 & q_{(x,i)} \\ 0 & 1 & 0 \end{bmatrix} \quad (3)$$

$$M_{(y,i)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & q_{(y,i)} \end{bmatrix} \quad (4)$$

For z -axis, scaling is used to simulate the front-back movement, and the following matrix is used to express the operation:

$$M_{(z,i)} = \begin{bmatrix} s & 0 & 0 \\ 0 & s & 0 \end{bmatrix} \quad (5)$$

where s denotes the scaling factor

$$s = (100 + q_{(z,i)})/100 \quad (6)$$

For rotation, the following matrix is calculated in order to rotate the frame by $q_{(r,i)}$ degrees around the center point:

$$M_{(r,i)} = \begin{bmatrix} \alpha & \beta & (1-\alpha) \cdot c_w - \beta \cdot c_h \\ -\beta & \alpha & \beta \cdot c_w - (1-\alpha) \cdot c_h \end{bmatrix} \quad (7)$$

where

$$\alpha = \cos(q_{(r,i)}) \quad (8)$$

$$\beta = \sin(q_{(r,i)}) \quad (9)$$

$$c_w = (\text{frame width})/2 \quad (10)$$

$$c_h = (\text{frame height})/2 \quad (11)$$

After constructing the four matrixes, they are merged to M_i so that only 1 affine transformation is required to performed instead of 4.

$$M_i = (M_{(x,i)} + M_{(y,i)} + M_{(r,i)} - 2I) \cdot M_{(z,i)} \quad (12)$$

After applying affine transformations on all the frames, camera shake would be successfully added to the video.

2.2 Auto Focus Blur

The second effect is lens blur. If a video is shot with auto-focus function, when the lens try to refocus, the focal point would change and cause image blur. In this program, average blur is used to simulate this effect. An average blur is done by doing a convolution between a frame and a $n \times n$ normalized box filter K .

$$K = \frac{1}{n \times n} \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 1 & 1 & \dots & 1 \\ 1 & 1 & 1 & \dots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \dots & 1 \end{bmatrix} \quad (13)$$

In other words, the value of every pixel in the blurred image would be the average of its neighborhood in the original image. Therefore, the bigger n is, the blurrier the output image will be.

The intensity change is simulated by the combination of two waves: one wave with larger amplitude and longer wavelength following with another smaller one. The idea is, when a lens tries to refocus, it would adjust the focal length until the target object is the clearest. However, it would know that it has reached the desired point after exceeding a little. In implementation, since the intensity should be an integer to perform average blur, the figure would be rounded.

2.3 Noise

It is common to see noise when watching a video, especially in the past when videos are recorded by films. Noise are caused by many reasons and have different form. In our program, three kinds of noise simulation is implemented, including pixel noise, line noise, and pattern noise.

2.3.1 Pixel noise

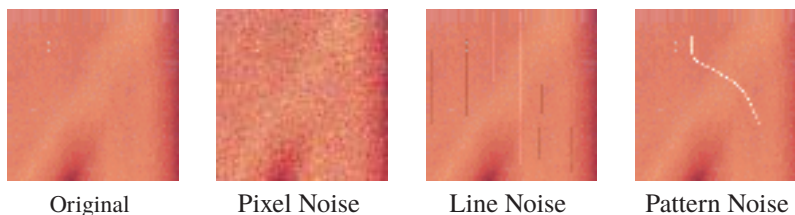
Pixel noise is an inevitable noise, and it is especially obvious if ISO sensitivity is high. This effect makes a frame contain bright spots or randomly-spaced bright pixels. To simulate this effect, two random number is used. The first value controls whether the pixel value should be alter, and the second value controls the quantity of alternation. In order to make it more realistic, the RGB channel are processed separately. The probability and the minimum and maximum of the value can be adjusted by the user.

2.3.2 Line noise

Similar to pixel noise, one random number is used to determine whether a line should be altered. However, in the next step, two random values are required where one controls the length and the other controls the quantity of alternation.

2.3.3 Pattern noise

Pattern noise takes an extra image as input. The image is the reference of the pattern which would be added to the video. It should come with black or transparent background. Otherwise, unnatural edge would appear. The program would randomly scale and rotate the pattern and put it in a randomly selected place.



3 Running Program

In order to run the program, 2 packages are required: `opencv-python` [1] and `numpy` [2].

3.1 Installation

Simplest way to install the `video_flawer` is via pip:

```
pip install video_flawer
```

3.2 Usage as utility

To flaw a video (with path `./test.mp4`) with preset defects call:

```
python -m video_flawer test.mp4
```

The output (for example `/data/out.avi`) file can be setup as:

```
python -m video_flawer test.mp4 /data/out.avi
```

For customization of video defects you should create your custom config file (JSON format). You can provide the file (for example `./my_config.json`) as follows:

```
python -m video_flawer text.mp4 /data/out.avi my_config.json
```

3.3 Usage in script

You can use `video_flower` directly in your script with preset defects as follows

```
import video_flower

INPUT_PATH = "test.mp4"
OUTPUT_PATH = "out.avi"

video_flower.run(INPUT_PATH, OUTPUT_PATH)
```

The defect details can be provided as nested dictionary

```
import video_flower

INPUT_PATH = "test.mp4"
OUTPUT_PATH = "out.avi"

CONFIG = {
    "shift_sin_x": {
        "max_shift": 3,
        "frequency": 0.25
    },
}

video_flower.run(INPUT_PATH, OUTPUT_PATH, config_data=CONFIG)
```

Or the defects can be described in JSON file:

```
import video_flower

INPUT_PATH = "test.mp4"
OUTPUT_PATH = "out.avi"

video_flower.run(INPUT_PATH, OUTPUT_PATH, config_path="config_example.json")
```

4 Conclusion

In this paper the *video_flower* Python package is introduced. The effects implemented in the library are explained. Usage examples are provided. The project repository (all source codes) are accessible under MIT license at https://github.com/matousc89/video_flower.

Acknowledgement

This work was partially supported by the Grant Agency of the Czech Technical University in Prague, grant No. *SGS18/177/OHK2/3T/12*.

Literature

- [1] G. Bradski. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 2000.
- [2] Travis E Oliphant. *A guide to NumPy*, volume 1. Trelgol Publishing USA, 2006.

ENERGY CONSUMPTION METERING IN PRODUCTION PLANT (MĚŘENÍ SPOTŘEBY ENERGIÍ VE VÝROBNÍM ZÁVODU)

*David Polák*¹, *Cyril Oswald*²

¹ ČVUT FS a david.polak@fs.cvut.cz

² ČVUT FS a cyril.oswald@fs.cvut.cz

Abstract: Recherche of energy metering sensors, communication standards and systems for data storage. From acquired knowledge data flow architecture from metering devices to SQL database was designed.

Keywords: PowerMeter, Energy, Communication standards, Ethernet/IP, Modbus, PROFIBUS, PROFINET, Database, SQL, NoSQL, Operational Historian, System architecture, Data flow

Abstrakt: Rešerše sensorů měření energie, komunikačních standardů a systémů pro archivaci dat. Z nabytých znalostí byla navrhuta architektura toku dat z měřících zařízení do SQL databáze.

Klíčová slova: PowerMeter, Energie, Komunikační standardy, Ethernet/IP, Modbus, PROFIBUS, PROFINET, Databáze, SQL, NoSQL, Operační Historian, Architektura systému, Tok dat

1 Introduction

Information technologies are becoming more connected with operation technologies. This opens new windows of opportunities. One of them is energy consumption metering and monitoring. This document serves as a quick overview of what it takes to choose the right metering device and set up a data flow architecture.

2 Energy meters

There are many energy meters producers. In the figure below can be seen the overview of electricity meters series with the most important properties such as output resolution and communication interface.

Measurement instrument: Brand, series/type	Installation	Measurands*	Output resolution*	Communication interface*
IME, NEMO 96	Fixed type	V, A, W, VA, PF, THD	60 s	RS485, Impulse
Siemens, SIMEAS	Fixed type	V, A, W, VAR, VA, PF, THD	< 1s	Profibus, RS485
Schneider, Electrics, PM	Fixed type	V, A, W, VA, PF, THD	60 s	RS485, Impulse
Simpson, GIMA1000	Fixed type	V, A, W, VAR, VA, PF	1 s	RS485, Impulse
Yokogawa,	Fixed type	V, A, W, VAR, VA, PF	<1 s	Ethernet, RS485, Pulse
AccuEnergy, Acuvim	Fixed type	V, A, W, VAR, VA, PF, THD	<1 s	Ethernet, Profibus, RS485, Impulse
Janitza Electronics, UMG 604	Fixed type	V, A, W, VAR, VA, PF, THD	<1 s	Ethernet, Profibus, Impulse
Chauvin Arnoux, C.A.8335	Mobile type	V, A, W, VAR, VA, PF, THD	1 s	USB
Fluke, 434	Mobile type	V, A, W, VAR, VA, PF, THD	0.5 s	USB
Voltech, PM3000	Mobile type	V, A, W, VAR, VA, PF, THD	<10 ms	RS232, IEEE488
Load Controls, PPC	Mobile type	W	15 ms	Analogue 0-10 volts or 4-20 milliampere
National Instruments, cDAQ*	Mobile type	V, A	<1 s	USB

*the listed features are retrieved from the datasheets of the devices and are due to change in future instrument revisions

Figure 1: Overview of electricity meters [1]

In the production plant it was chosen to use PowerMeter series from Schneider Electric brand. These meters satisfy the desired output resolution and provide information about all measurands needed. Schneider Electric offers variety of meters starting from basic one phase energy metering using Acti 9 iEM2000, continuing with three phase multi-function metering using PowerMeter 3000 and ending with high-end meters such as PowerMeter 5000.



Figure 2: PowerMeter 5560 [2]

3 Industrial communication standards

With the rise of Industrial Internet of Things (IIoT) and Industry 4.0, Industrial Ethernet standards are flourishing. In year 2019 Industrial Ethernet solutions were representing 59 % of newly installed nodes compared to Fieldbus solutions being at 35 %. The most used communication protocol over Industrial Ethernet is EtherNet/IP, which managed and developed by ODVA (Open DeviceNet Vendors Association). It is widely used by Rockwell and most of IIoT devices. The second most popular is PROFINET (Process Field Net) utilizing Industrial Ethernet and it even has its counterpart using Fieldbus technology PROFIBUS. These protocols are mainly used by Siemens products. Another example, which is using both Fieldbus and Industrial Ethernet interfaces, was developed by Schneider Electric and it is called Modbus. Modbus protocol can be communicated over RS485, Industrial Ethernet or RS232. [3]

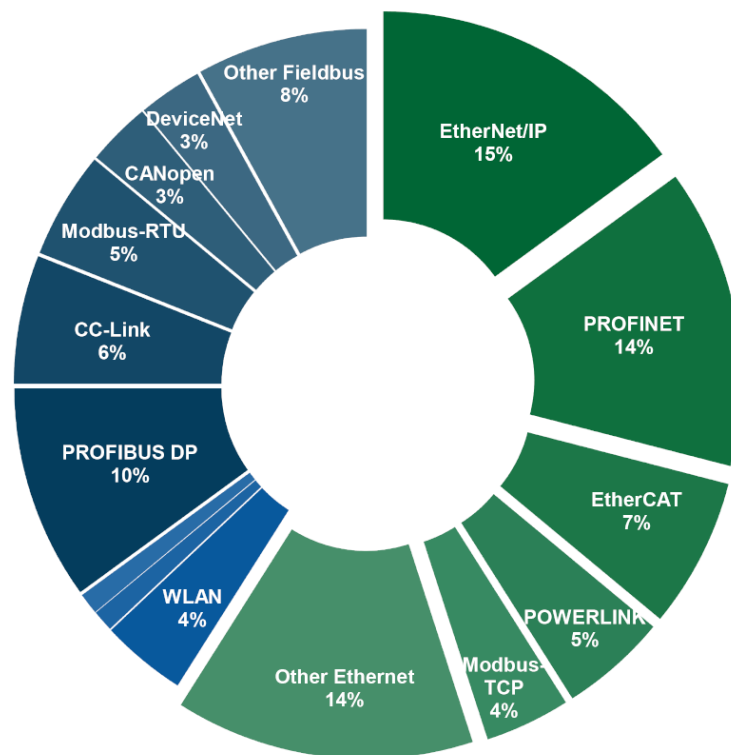


Figure 3: Industrial communication protocols [3]

4 Data storage

Relational databases are popular choice for various systems. It is a set of multiple tables, that are linked through relations. Each table consists of tuples (rows) and attributes (columns). This database type is usually queried and written using Structured Query Language (SQL). It has become so popular, that relation databases are usually referred to as SQL Databases.

The other type is NoSQL database, which refers to Non-SQL or Non-Relational database. The data structure of NoSQL databases is much simpler as it sometimes only stores the key and its value. The simplicity allows some queries to be executed faster than using traditional relational database.

Operational Historian is a time-series database system, which stores industrial process data. In order to do that, it must be able to access Programmable Logic Controllers (PLCs) values. Some Operational Historians have SQL interfaces and therefore they can be accessed via SQL queries (e.g. Historian developed by Wonderware). Others can be accessed via dedicated Application Programming Interfaces (API), which can make it harder to connect it with

SQL databases, but easier for cloud solutions (e.g. Historian developed by GE).

5 Proposed Data Flow Architecture

In order to read the data from energy meters in SQL database, the correct data flow architecture has to be set up. In the factory there are 2 types of meters in terms of communication interface:

1) Impulse

These devices should be connected to PLC Modicon M258, where the pulses shall be counted. This data should be later communicated through Modbus TCP/IP protocol.

2) RS485

Meters, which are communicating over serial bus RS485 should be connected to Gateway EGX100. These meters should act as slaves and the gateway as master using Modbus over RS485 protocol. The data is then translated to Modbus over TCP/IP.



Figure 4: Proposed Data Flow Architecture

Both Modicon and EGX100 should be communicating with PLC Micro850 and it translates the data to Ethernet/IP protocol, which can be accessed by Wonderware Historian. As this type of historian has SQL interface, it can be queried directly from SQL database in order to fill the database.

6 Conclusion

The goal of this project was to do the research of energy metering devices, communication standards and data storage systems. With newly gained knowledge I proposed data flow architecture, which utilizes protocols communicated over both Fieldbus and Industrial Ethernet technology. The data gets first to Wonderware Historian and then through SQL interface it is written into relational database, where it is ready for further analysis.

Bibliography

- [1] Electricity Metering and Monitoring in Manufacturing Systems (2011), S. Kara, G. Bogdanski, W. Li
- [2] PowerLogic™ PM5500 series user manual [online]. Paris: Schneider Electric Industries SAS, 2020 [cit. 2020-08-02].
Dostupné z: https://download.schneider-electric.com/files?p_enDocType=User+guide&p_File_Name=HRB1684301-09-EN.pdf&p_Doc_Ref=HRB1684301
- [3] What is the most popular industrial network protocol? [online]. New York: Engineering 360, May 28, 2019 [cit. 2020-08-03].
Dostupné z: <https://insights.globalspec.com/article/11936/what-is-the-most-popular-industrial-network-protocol>

CODESYS FOR BASIC FESTO MOTION TERMINAL (VTEM) APPLICATION

Jana Koubová¹, Marie Martinásková²

¹*CTU Prague, Jana.Koubova@fs.cvut.cz*

²*CTU Prague, Marie.Martinaskova@fs.cvut.cz*

Abstract:

This paper describes how to use CODESYS software environment for programming of a simple application using Festo Motion Terminal (FMT) - VTEM. This application consists of controlling the movement of one pneumatic actuator with proximity sensors using visualization in the CODESYS software environment. Firstly the main aspects of Festo Motion Terminal - VTEM are described. Secondly the prerequisites for use are described such as installation process of the software, packages and libraries needed for creating an application compatible with VTEM. At last the whole applications composition is described.

Keywords:

valve terminal, pneumatic actuator, CODESYS, Festo Motion Terminal, FMT, VTEM, structured text, , programmable logic controller, proximity switch, electropneumatics, proportional pneumatics, digital pneumatics

1. Introduction

Valve terminals are widely used as a mean to use pneumatics efficiently in automation. So far the valve terminals consisted of specific valves configured together in a cluster for one specific application. Festo Motion Terminal (FMT) introduces a universal valve VEVM, which can switch its inside configuration and thus perform function of different valves and components. [1] Festo Motion Terminal is a valve terminal designed to uphold the standards of Industry 4.0 containing 4 or 8 VEVM valves with a controller module CTMM, to which other controller modules can be connected. In our application the higher control module is CPX-CEC-V1-S3. The application for CPX module is created using CODESYS. CODESYS is development environment for programming of control applications in which all 5 programming languages recommended in IEC/EN 61131-3 for programmable logic controllers are available: Ladder Diagram (LD), Function Block Diagram (FBD), Structured Text (ST), Instruction List (IL), Sequential Function Chart (SFC). Additionally it includes Continuous Function Chart (CFC) which is evolution of FBD, which allows multiple function blocks to be connected. In this application ST and CFC were used.

Application shown in this paper will allow for manual and automatic (sequence) control of one pneumatic actuator with two end position proximity sensors. This document can serve as a guideline for the necessary software installation and creation of code for basic FMT application.

2. Specification of Festo Motion Terminal (FMT)

Structure

FMT design

Conceptually we can divide the FMT into 3 parts: control, pneumatic and input modules. Higher control is provided by CPX module (PLC), which is connected to CTMM controller. The CTMM provides communication directly for FMT terminal (the valves) and also CTMM is where the main air supply with fitting for tubing and main exhaust port with silencer are located. Single universal valves and CTMM are connected pneumatically and electrically via the manifold rail VABM. The manifold also connects CTMM and input modules electrically. On top of the manifold there is certain number of connection ports. Into these ports the valves and input modules are fitted. Leftover ports are covered with cover plates. Unused ports are covered with blanking plugs. [1]

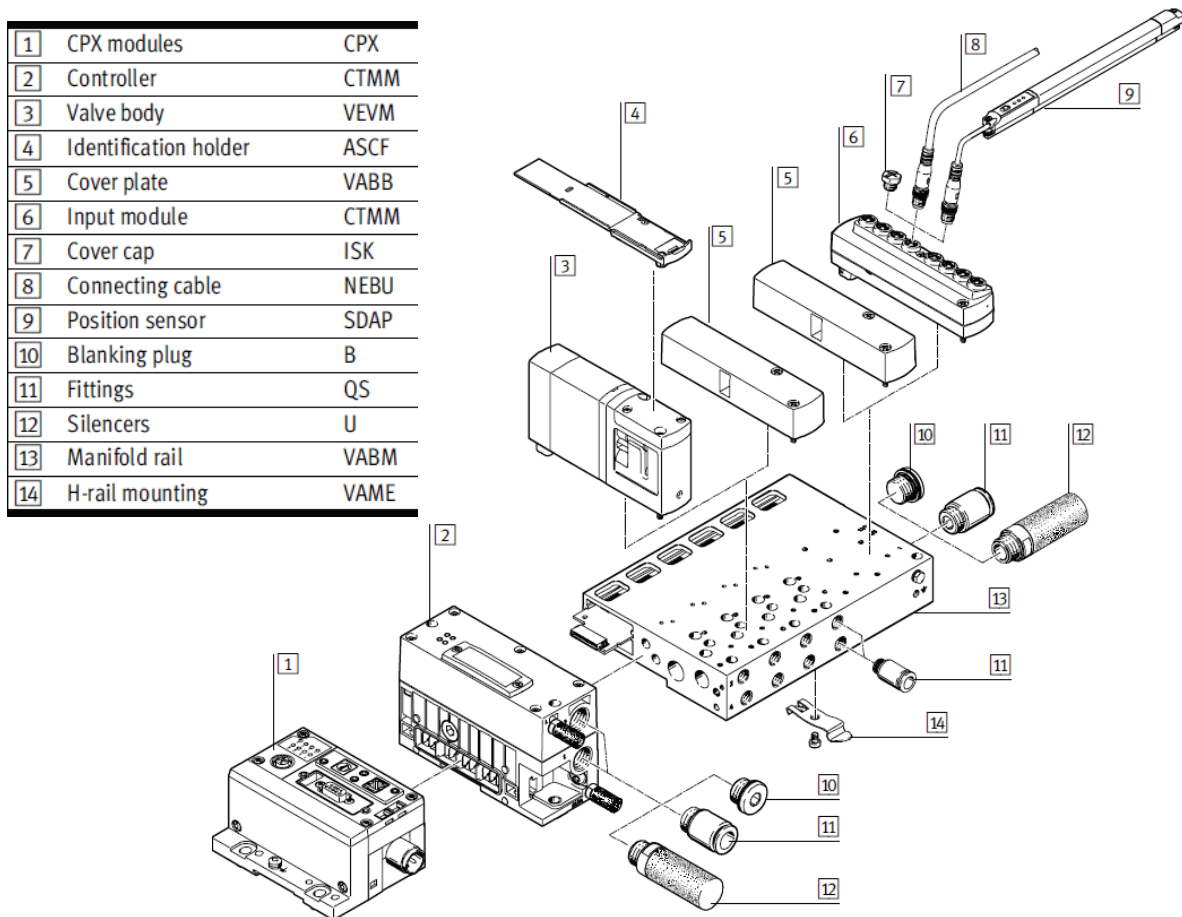


Fig. 1 Festo Motion Terminal – Peripherals overview [1]

VEVM universal valve

The main concept difference between a standard valve terminal and VTEM terminal are the valves. All the universal VEVM valves are the same in build, but they can substitute different components based on their inner parts alignment changing according to the running program. Thus instead of needing to change components for different application, it is only necessary to change the program. It is also not necessary to store multiple spare parts for different types of valves.

Each valve comprises of four 2/2-way proportional valves positioned in bridge circuit (Fig. 2) with integrated sensors. Each 2/2-way valve has 2 piezo pilot valves which are connected to diaphragm poppet valve. The bridge configuration enables for each port to be set in one of following states: open, closed, pressurized, exhausted. By setting the states on port 2 and 4 the valve can simulate function of different components. [2][3]

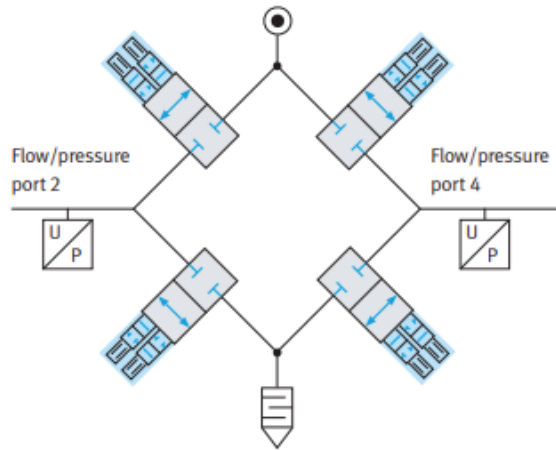


Fig. 2 Bridge circuit in the VEVM valve [3]

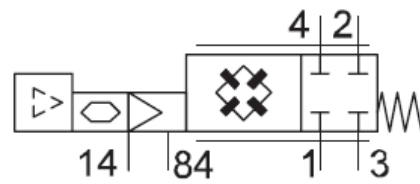


Fig. 3 Circuit symbol for VEVM valve [3]

2.1 Functionality

Motion Apps

The user can assign one Festo Motion App to each valve through the program at one time moment. Through the app, parameters for the function of the valve can be set. Festo Motion Apps are preprogrammed within the VTEM terminal. Through PLC program or in the Web Interface Festo Motion Apps and their parameters are assigned to valves. [2]

Web Interface

There is an option to connect the CTMM module directly to a computer via Ethernet connection. This lets the user interact with CTMM via Web Interface. This option is especially useful for understanding all the Festo Motion Applications and their parameters. In user-friendly environment, apps can be assigned to the valves and their parameters set. For the connection to be successful a specific IP address must be set on the computer to be able to see the terminal. The terminal controller CTMM IP address is then inserted into web browser (Mozilla Firefox is the recommended browser) address line and the interface is loaded – in factory setting the IP address is 192.168.4.2. [4]. Before the terminal can be accessed, user needs to sign in with a password (factory setting password is “vtem”). Working in this environment is very intuitive. For the connection to work only the first two numbers of IP address of the computer are given by the controller address, that means that PC address can be any variation of 192.168.x.x.

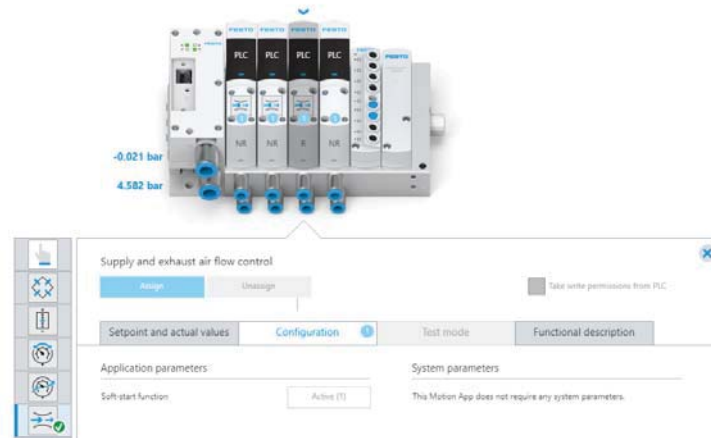


Fig. 4 Web interface for Festo Motion Terminal

PLC module control

Festo Motion Terminal can be controlled by local controller module of the CPX terminal which then communicates with the CTMM controller. In this case it is CPX-CEC-S1-V3. Into CPX the program is downloaded from computer. Programming can be done in software environment CODESYS. There are many CPX communication modules with many possible communication protocols (e.g. CAN, CANopen, Modbus). The technology as a whole has been designed to be in compliance with standards of Industry 4.0. For this project we used Ethernet connection between the CPX module and computer.

2.2 Necessary Software Installation

When installing, it is imperative that the version of the particular software downloaded is compatible with used devices and modules.

For best results installation should be completed in the following order:

a) CODESYS

- 1) Version of CODESYS software environment needs to be downloaded that corresponds to the higher control block device. For this project it is CODESYS V3.5 SP12 Patch6 pbF which supports CPX controller CPX-CEC-S1-V3. This download can be done from relevant webpage, where the used CODESYS is seen as in Fig. 5.

Relevant webpage for download:

https://www.festo.com/net/cs_cz/SupportPortal/default.aspx?q=codesys&tab=4&type=75#result

CODESYS provided by Festo	V3.5 SP12
CODESYS V3.5 SP12 Patch6 pbF	Patch 6 pbF
	22.07.2019
Supported Systems:	
CPX-E-CEC-C1-PN	
CPX-E-CEC-M1-PN	
CPX-E-CEC-C1-EP	
CPX-E-CEC-M1-EP	
CPX-E-CEC-C1	
CPX-E-CEC-M1	
CPX-CEC-S1-V3	
CPX-CEC-M1-V3	
CPX-CEC-C1-V3	

Fig. 5 Description of the used version of CODESYS from the Festo web

- 2) After download the installation is simple using the Setup_CODESYSV35SP12Patch6.exe file located within the downloaded .zip folder, which shall be extracted first prior installation. This installation can be run in default mode.

b) Festo Configuration Tool

Next step is to install the FCT – Festo Configuration Tool – PlugIn.

- 1) Download the FCT – PlugIn from relevant webpage where it is seen as in Fig.6.

Relevant webpage for download:

https://www.festo.com/net/cs_cz/SupportPortal/InternetSearch.aspx?q=codesys&tab=16&type=70#result

FCT - Festo Configuration Tool - PlugIn	1.0
<input type="checkbox"/> configuration and start-up software for the par.kinematic: EXPT-..	16.05.2012
<ul style="list-style-type: none">• FCT PlugIn CMMP-AS v1.4• FW CMMP-AS v3.5• FCT PlugIn CMXR-C1 v1.1.1.6• FCT PlugIn CMXR-C2 v1.1.2.40• Festo Configuration Tool v1.2• CoDeSys provided by Festo v2.3.9.28• TSP (CoDeSys Target Support Package) für CMXR-C2 v01.11• sample projects for EXPT-..	

Fig. 6 Description of the used Configuration Tool from the Festo Web

- 2) After downloading of the .zip file, it is necessary to extract it and then to continue with the installation as described in Readme.txt file. No changes need to be executed in the installation process. When choice is offered chose “Install PlugIns”. Last part of instructions in Readme.txt starting with “To work with CoDeSys...” concerning CoDeSys target needn’t be finished.

c) Target Support Package

- 1) First Target Support Package CODESYS needs to be downloaded that corresponds to the version of CODESYS present on the PC and to the used controller block – for this project CPX-CEC-S1-V3 (3472425) without CPX-CTEL-2-M12-5POL-LK. This shall be downloaded from the relevant webpage. At the webpage this package will be seen as in Fig. 7.

Relevant webpage for download:

https://www.festo.com/net/cs_cz/SupportPortal/InternetSearch.aspx?q=codesys&tab=16&type=73#result

Target Support Package CODESYS	3.5.7.356	→ Balík cílové podpory
<input type="checkbox"/> Supported systems:	06.08.2018	→ Soubory a jazykové verze
<ul style="list-style-type: none">• Control block CPX-CEC-C1-V3 (3473128) without CPX-CTEL-2-M12-5POL-LK• Control block CPX-CEC-M1-V3 (3472765) without CPX-CTEL-2-M12-5POL-LK• Control block CPX-CEC-S1-V3 (3472425) without CPX-CTEL-2-M12-5POL-LK• Motion Terminal VTEM (8047502) without CPX-CTEL-2-M12-5POL-LK		

Fig. 7 Description of the used Target Support Package

- 2) This package needs to be then installed in CODESYS in Package Manager located in Tools tab. In Package Manager it is needed to click on “Install” and select the downloaded package from folder. After the installation the new package should be visible in your Package Manager window.

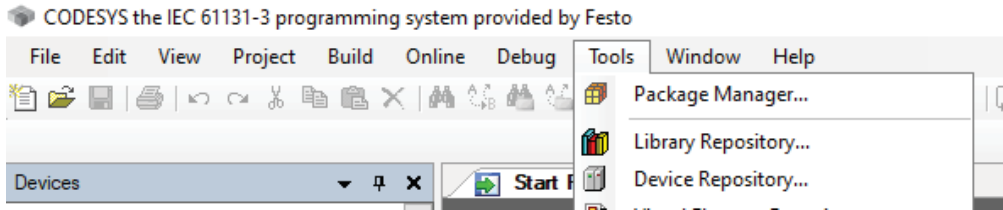


Fig. 8 Location of Package Manager and Library Repository tools in CODESYS software

d) Function Blocks CODESYS

To enable use of VTEM in CODESYS projects, Function blocks CODESYS library needs to be installed. This library has to be installed in CODESYS software in Library Repository.

1) The library needs to be downloaded from relevant webpage where the library is seen as in Fig. 9.

Relevant webpage for download:

https://www.festo.com/net/cs_cz/SupportPortal/InternetSearch.aspx?q=codesys&tab=16&type=74#result

Function blocks CODESYS	3.5.7.221
<input type="checkbox"/> Library valid from FW-Version 4.8.x	06.12.2018
Library for Codesys PLC in version 3.5 (also Beckhoff TwinCAT 3)	
Supported systems:	
• Motion Terminal VTEM (8047502)	

Fig. 9 Description of the used Function blocks CODESYS

2) Next step is to extract the .zip folder. In the folder there is Library subfolder and also Documentation subfolder, which contains additional description of the use of Function blocks. Library Repository is located in Tool same as Package Manager. In Library Repository by clicking “Install”, finding and selecting the library file in the folder it will be installed. After installing the Support Package and VTEM Library restart of CODESYS software is recommended. After that the result should be as seen in Fig. 10.

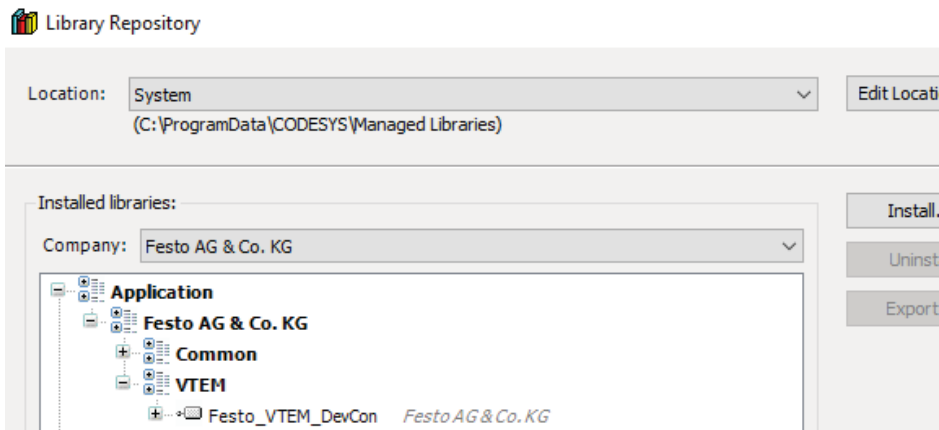


Fig. 10 Festo_VTEM_DevCon library installed in Library Repository

2.3 Connecting the FMT via CPX to computer using Ethernet

Connecting the CPX module Ethernet port with Ethernet port on your computer will allow for applications

to be downloaded to the CPX from CODESYS. Run/stop of the application is possible via position of rotary switch placed on the CPX module or the application can be started and stopped in CODESYS software environment. Proper connection is possible only if computers IP address is set so that it corresponds to the address of the CPX controller. Finding out the CPX IP address in CODESYS using “Scan Festo devices” as shown in Fig.11 is possible. This can be done prior to starting new project as long as the devices are connected in a network.

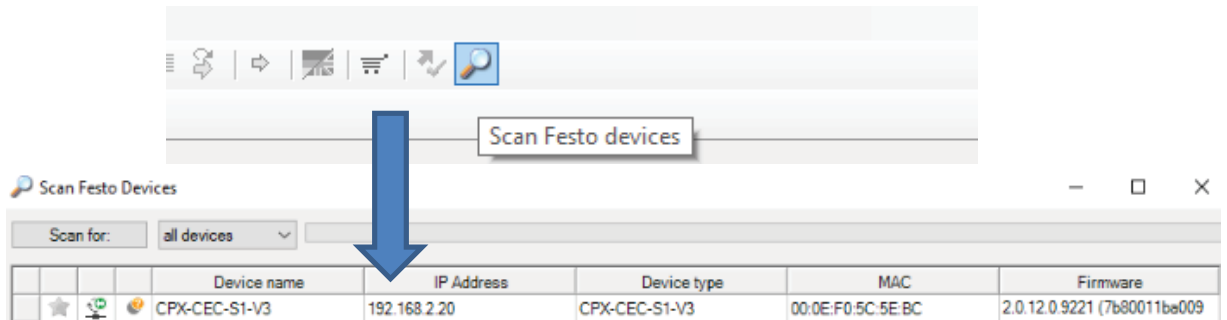


Fig. 11 Finding out the CPX controllers IP address

This means the IP address of the computer should be set to 192.168.2.x and mask to 255.255.0.0. Last number of the IP address should be different than the one of the controller. One way that this can be done on PC run on Windows is to search in START for Network Connections and open Network Connections Control Panel. In this panel Properties of Ethernet should be opened. There in section Internet Protocol Version 4 (TCP/IPv4) is located – IP address and mask can be changed in its Properties.

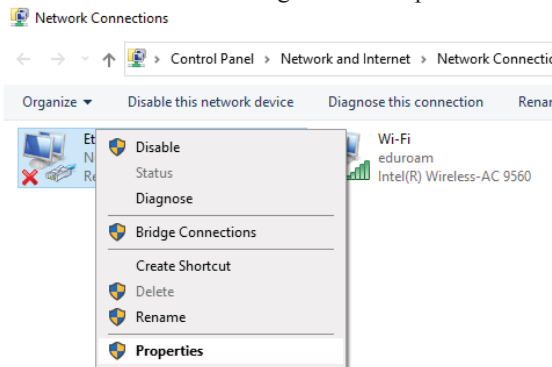


Fig. 12 Changing the IP address of Windows run PC

2.4 Overall project layout

The physical setup of this project is as follows. The FMT is connected to electricity (24V) by special connector cable and to compressed air supply (6 bar input to CTMM module). Components used are from Festo Didactic line. One double acting pneumatic actuator is connected to the first valve position on FMT. This connection is done so that port 2 under VEVN valve is connected to port on the actuator, which when pressurized causes retraction, and port 4 under VEVN valve is connected to port on the actuator, which when pressurized causes advancing as indicated on Fig. 13. This is to the contrary with usual convention, but is used here to be with compliance with logic used in the Web Interface and Apps provided with FMT. There are two 3-pin proximity sensors each at one end position of the piston. These are connected to the first two positions of the input module.

Project layout representation and photo

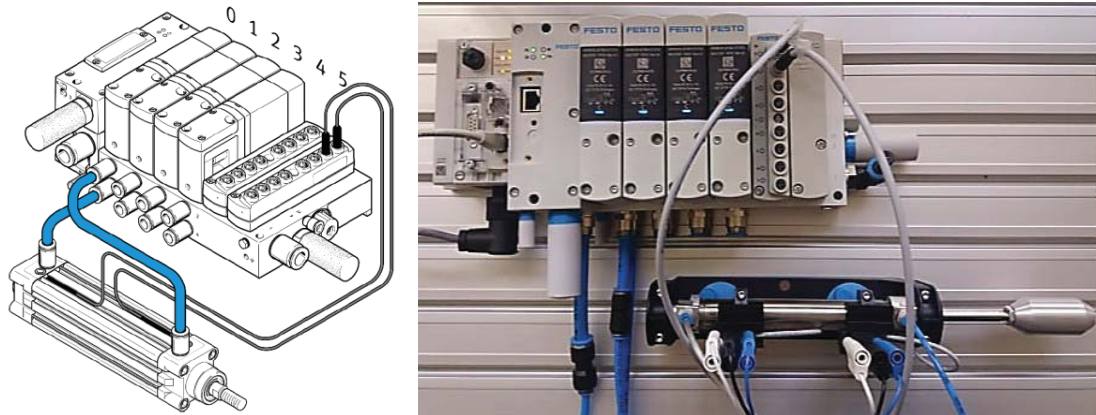


Fig. 13 Left: representation of actuator connection to FMT sourced from [5], right: photo of setup taken in the lab C1-109 CTU FS

3. CODESYS

3.1 Project set-up

New project can be open either from File tab or from the Start Page in CODESYS. After choosing project template (CPX-CEC project) for the type of controller (Fig. 14) it is necessary to make sure that the device set is the same as the one used in the project (Fig.15). In this case CPX-CEC-S1-V3. The language should be set as Structured Text. After that new project with structure as shown on Fig. 16 is formed. Next step is adding the Festo_VTEM_DevCon library to the project library using Library Manager of the project (Fig. 17, Fig. 18).

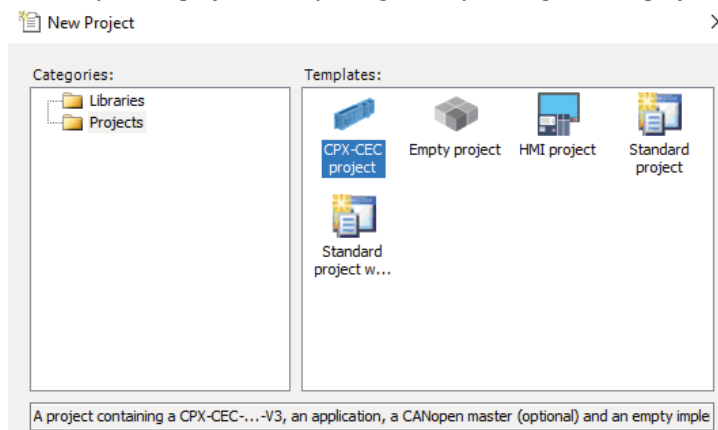


Fig. 14 Choosing the template of the new project

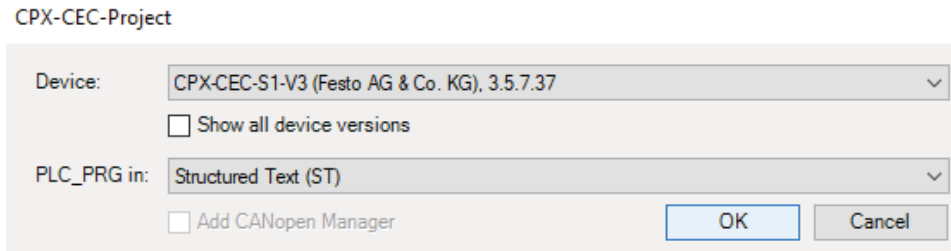


Fig. 15 Choosing device for the project from selection and programming language

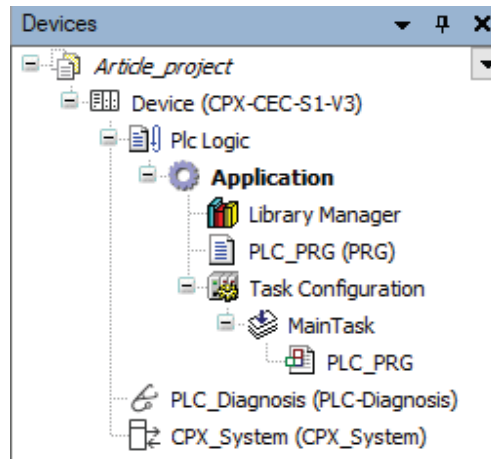


Fig. 16 New projects tree structure

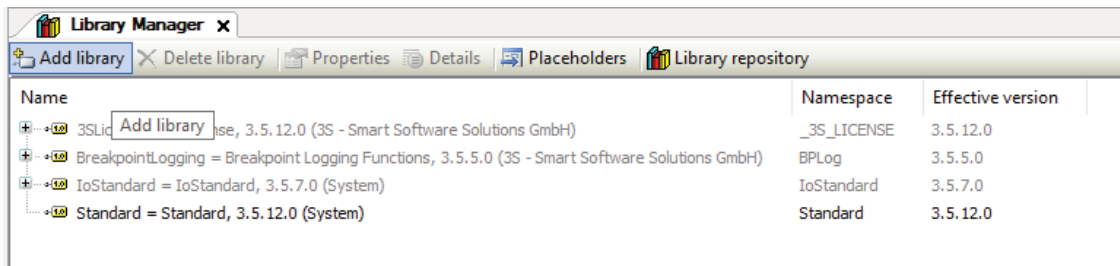


Fig. 17 Library Manager of the project

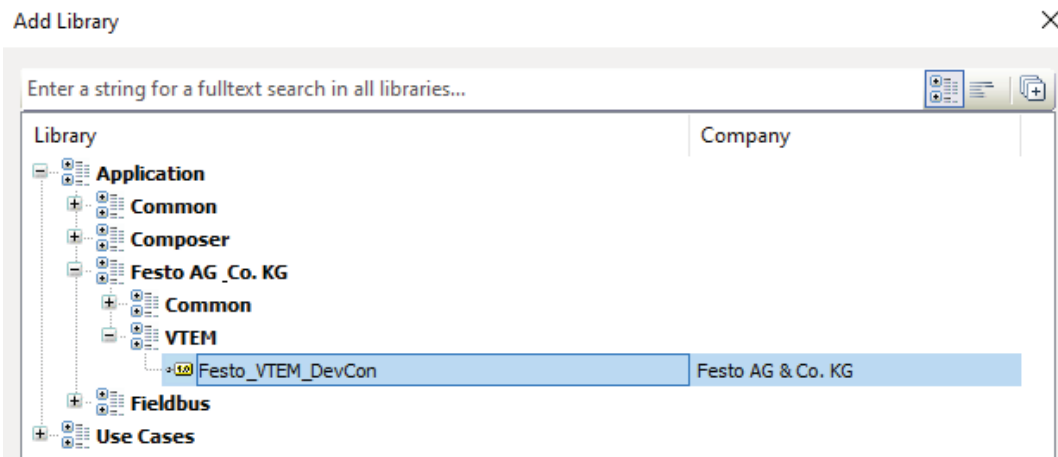


Fig. 18 Adding the Festo_VTEM_DevCon library to the project

3.2 Device set-up and gateway connection

Gateway connection

To establish gateway connection the Device in project tree structure needs to be open. The window that opens is Communication Settings. One way to easily create a gateway is to click on Login in the toolbar (Fig. 19). A following message will appear (Fig. 20). Alternatively the gateways can be added in the Communication Settings window by clicking at “Add gateway...” and then “Scan network” to find the controller block. Fig. 21 shows how active Gateway with device presents. If the PC IP address is set correctly then this automatic Gateway should work in default mode.

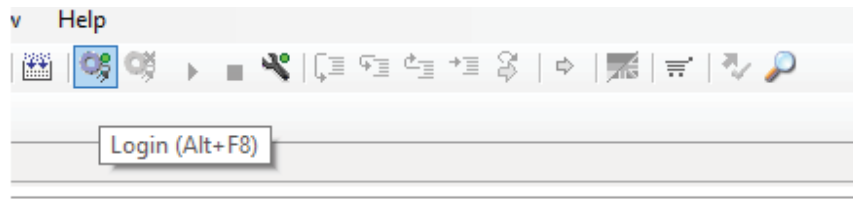


Fig. 19 Location of Login button on upper toolbar in CODESYS

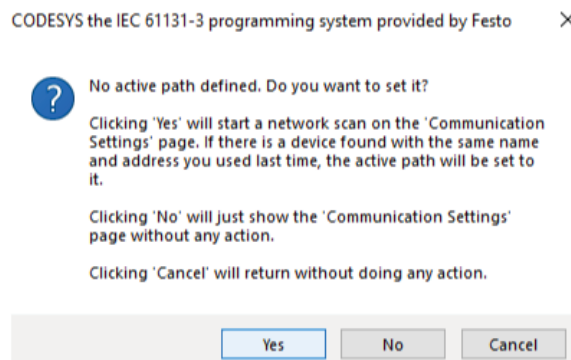


Fig. 20 Message about starting scan and setting pathway

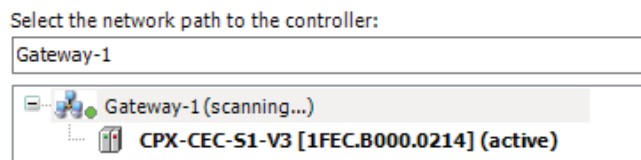


Fig. 21 Active pathway in the Communication Setting window

Configuration scan

In the project tree CPX_System should be opened. In the first folder Module Configuration by clicking on Actual Configuration and performing a Scan (Fig. 22) the actual configuration of modules (Fig. 23) and information about them should appear if your Gateway is working properly. After clicking on the Apply button the specification of the modules will be added to the tree structure of the project (Fig. 24).

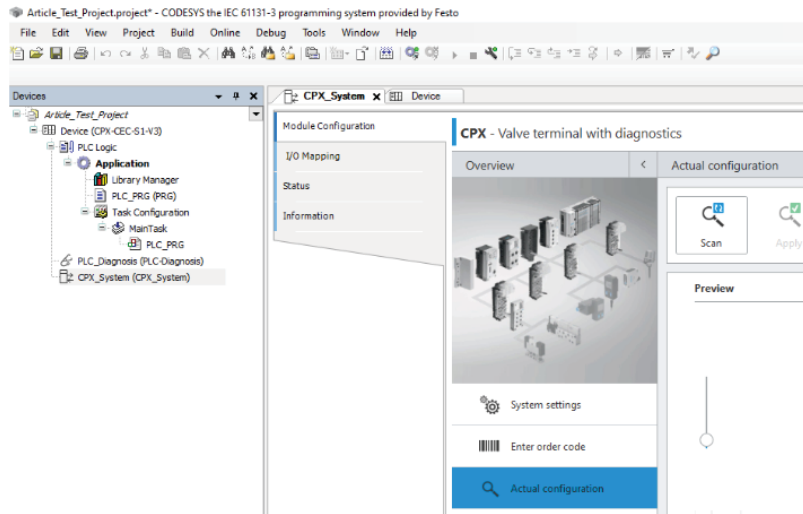


Fig. 22 Scanning for actual configuration of modules

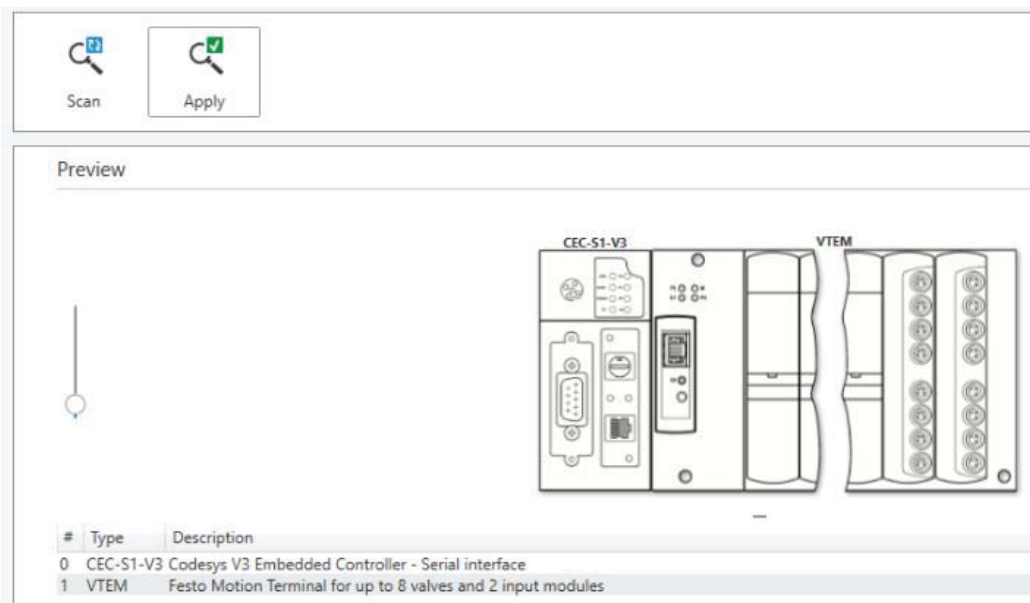


Fig. 23 Actual configuration of modules

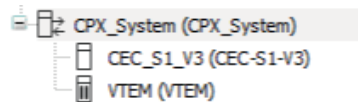
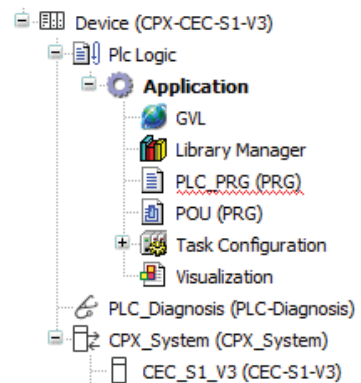


Fig. 24 Modules added to program tree

3.3 Parts of the application

To create a fairly basic application for FMT Application it should contain:

- Global Variable List (GVL)
- Library Manager
- Structured text (ST) POU (Program Organization Unit)



- Continuous Function Chart (CFC) POU
- Task Configuration
- Visualization

Some of them are automatically created when new project is started such as the Library Manager in which the VTEMDevCon library needs to be added as discussed in section Project Set-up.

More information about some of the sections is provided below.

Fig. 25 Finished projects Application tree

Global Variable List - GVL

To define variables that will be used throughout the Application, not only the ST program, Global Variable List has to be created. Variables here should be the valve and terminal variables and also the variables that will be used in Visualization as well as in POU's. To create GVL it is possible to right-click on Application and choose Global Variable List in the Add Object section. How to create GVL and which variables are used in this project is shown in Fig. 26.

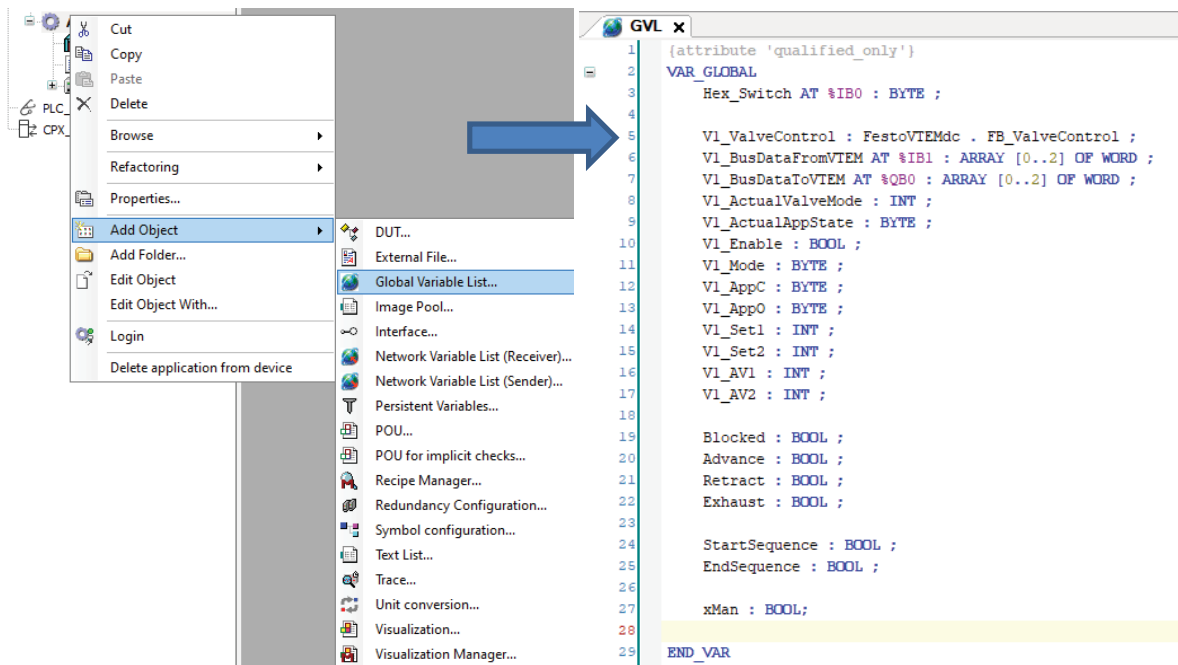


Fig. 26 On the left: creation of GVL, on the right: variables used in the project

Program organization unit in CFC (Continuous Function Chart) - POU

For proper communication between used variables in CPX block and FMT a function block in CFC needs to be used. To the formal parameters of the function block which are prepared in the function block the actual parameters from CPX variables are assigned. This block is prepared specifically for VTEM valve control. CFC POU can be added using same process as above with GVL but choosing POU. Type of POU should be kept as Program and Implementation Language can be chosen. In this case it is CFC which is default when creating new POU. After creating the CFC POU there will be ToolBox window on the right side of the screen. Box is to be chosen and dropped into the down section of the split screen on the left side. To insert the prefabricated setting right-click the box needs to be performed and Input Assistant chosen as shown in Fig. 27. There choice of FB_ValveControl needs to be done in Operations. Note that FestoVTEMdc is only visible when the special library is added previously in Library Manager. In the Input Assistant there is also a Documentation present (in Fig. 28), which explains the input and output variables and also defines their type.

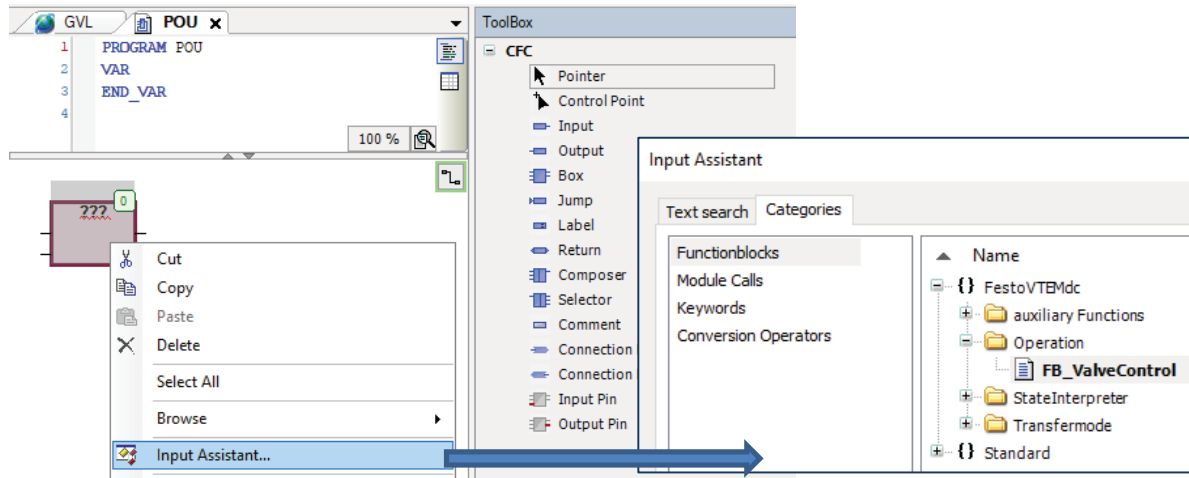


Fig. 27 Input of ValveControl function block

The prefabricated block will appear. To this block inputs and outputs should be connected as in Fig. 29. This function block encompasses only one of the VTEM valves, if more valves are to be used each should have their own function block and their own set of global variables.

Show documentation Insert with arguments Insert with namespace prefix

Documentation:

This is the central function block for the operation of Motion Apps on a valve. The FB serves to operate all Motion Apps with their specific setpoint values. It is able to execute and to stop Motion Apps and to deal with occurring errors. Parameterisation is not possible with this function block.

Variable Name	Data Type	Direction	Description
awBusDataFromVTEM	ARRAY [0..2] OF WORD	VAR_INPUT	6 bytes of process data that is received from the Motion Terminal
xEnable	BOOL	VAR_INPUT	has to be set to TRUE in order to operate a Motion App on the valve, otherwise any motion is stopped
bySetValveMode	BYTE	VAR_INPUT	ID of the Motion App to be executed, 60 for 'commissioning', 61 for 'stop', 62 for 'acknowledge'
bySetAppControl	BYTE	VAR_INPUT	SetAppControl of the chosen bySetMode, see documentation of the valve mode for its meaning
bySetAppOption	BYTE	VAR_INPUT	SetAppOption of the chosen bySetMode, see documentation of the valve mode for its meaning
iSetpointValue1	INT	VAR_INPUT	SetPoint1 of the chosen bySetMode, see documentation of the valve mode for its meaning
iSetpointValue2	INT	VAR_INPUT	SetPoint2 of the chosen bySetMode, see documentation of the valve mode for its meaning
xManualAcknowledge	BOOL	VAR_INPUT	errors can be acknowledged (after they got inactive) with a rising edge at this input
awBusDataToVTEM	ARRAY [0..2] OF WORD	VAR_OUTPUT	6 bytes of process data that is sent to the Motion Terminal
byActualValveMode	BYTE	VAR_OUTPUT	actual valve mode of the Motion Terminal
byActualValveState	BYTE	VAR_OUTPUT	actual valve state of the Motion Terminal
byActualAppState	BYTE	VAR_OUTPUT	Byte1 of the received process data: app state of the desired Motion App, see documentation of the valve mode for its meaning
iActualValue1	INT	VAR_OUTPUT	ActualValue1 of the desired Motion App, see documentation of the valve mode for its meaning
iActualValue2	INT	VAR_OUTPUT	ActualValue2 of the desired Motion App, see documentation of the valve mode for its meaning
iResponseToValveModeSet	eResponseToValveModeSet	VAR_OUTPUT	enumeration that indicates the status of execution of the desired valve mode
iErrorCode	INT	VAR_OUTPUT	contains the malfunction code of the latest error, 0 if no error occurred so far

Fig. 28 Documentation for block variables accessible in Input Assistant

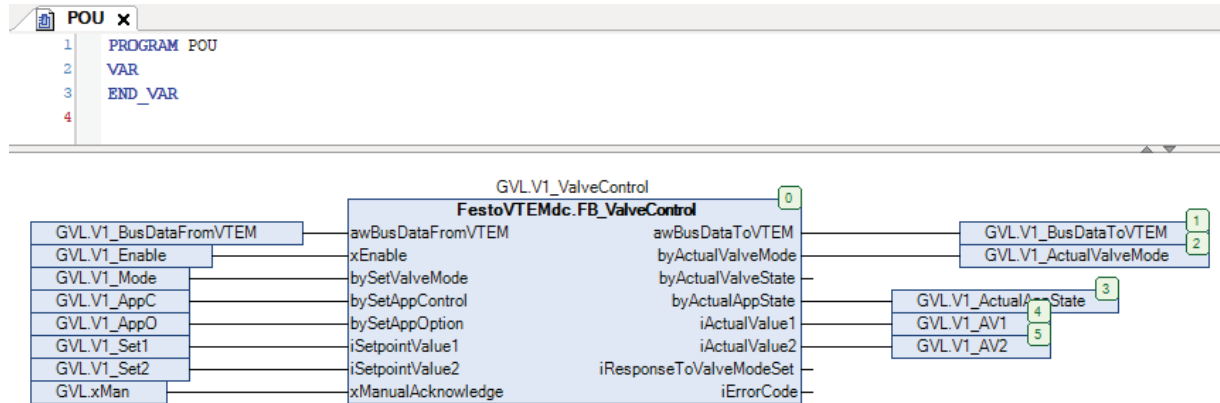


Fig. 29 Finished function block

Task Configuration

In Task Configuration Main Task is already present with the ST program in it. This tool allows for setting of different tasks and their type, when are they executed and so on. This can be left in default setting. The ST program will be executed cyclically every 20 ms. For more complex solutions Task Configuration tool can be used to create tasks that have different priority or are started “on event”. Main Task configuration is shown in Fig. 30

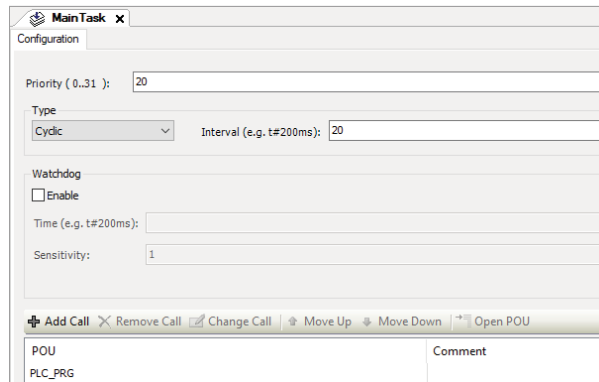


Fig. 30 Main Task Configuration in default mode

Visualization

There is no machine interface panel for the operator therefore the control of movement of the actuator is performed using the Visualization in CODESYS software environment.

Visualization is again created via right-click on Application and choice from the menu Add Object (Fig.31). After the visualization is created there is the visualization window on the left and ToolBox on the right. From the ToolBox in this project the main objects needed are buttons which are used to control BOOL variables in the ST program. These are located in the Common controls tab. Buttons can be labeled by clicking on them once when they are already selected. When button is selected the configurations of the button are displayed on the right. Most important is the Inputconfiguration, which allows the button to be linked with variables.

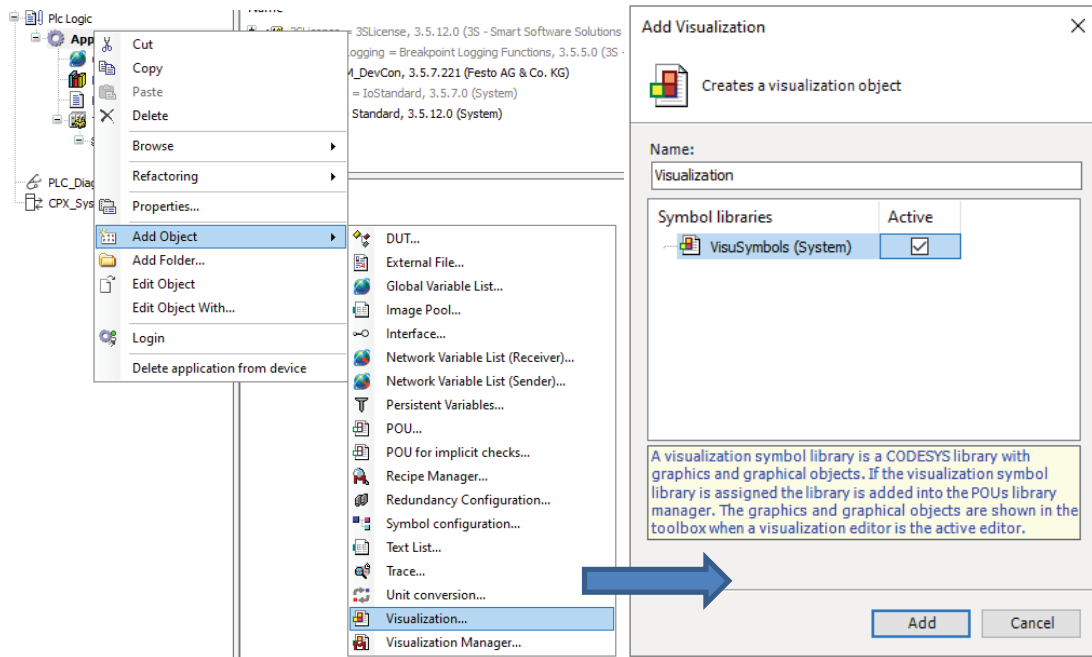


Fig. 31 Input of visualization in the Application

Fig. 32 shows button configuration. Variable that is flipped on mouse tap can be either written in or by clicking on symbol of 3 dots on the right side of the variable name tab you can open Input Assistant. Also it is possible to open Input Assistant by right-clicking the variable name tab, when writing is enabled. By inputting the “Advance” BOOL variable to the Inputconfiguration on Tap of the button labeled “ADVANCE” you link this variable with the use of this button. Same process is followed with other buttons as well. Of course other visualization elements are used such as labels, images and eventually sliders.

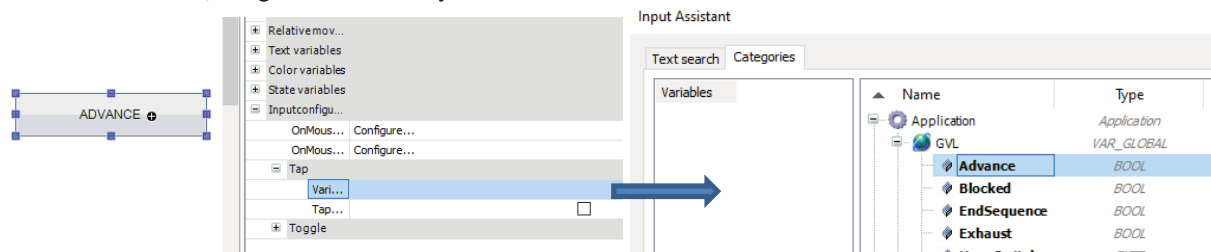


Fig. 32 Linking BOOL variable to the button in button Configuration

In the finished visualization all the buttons are linked to the BOOL variables of the corresponding names. The finished visualization used for this project is shown in Fig. 33. After sequence is started it must be ended before manual control can be resumed. Only button “BLOCKED” working as a safety measure can be used any time and after the piston is blocked the system returns to manual control. In this case sequence means that no matter what position the piston is in it will retract itself and then start cycle of full extension and full retraction.

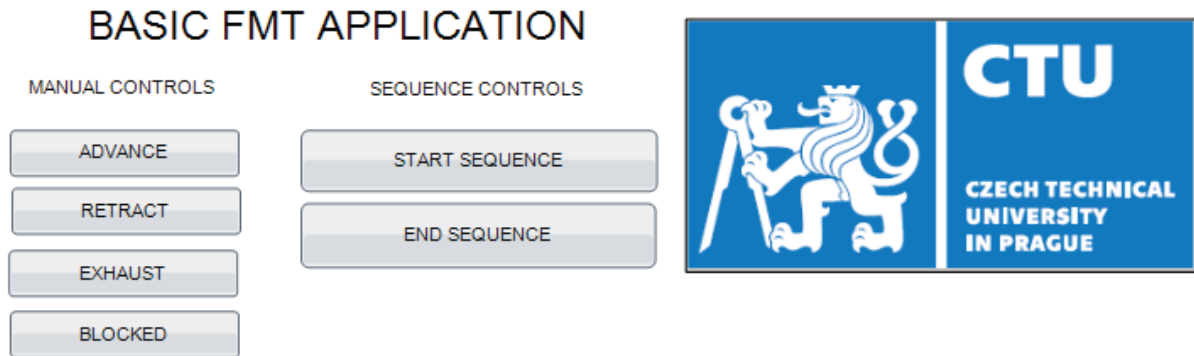


Fig. 33 Finished visualization of the project

Program in structured text (ST)

Finished program is shown here in Fig. 34 as an example of possible FMT application. This program enables manual and sequence control mode of one pneumatic actuator. Manual control allows for advancing, retraction, exhaust and blockage of the actuator. There is also a simple sequence, which uses the proximity switches. Sequence runs after “START SEQUENCE” is activated until either “END SEQUENCE” is activated or button “BLOCKED” is activated as a safety measure. After one of these events the manual control is resumed. The sequence consists of initial retraction of the actuator and then cycles of full extension and full retraction.

Below the logic of the code is described.

This section is already present in project tree. In the upper window local program variables are defined. These are variables that are confined only to this program. In this case it is only the State variable. In the lower window the program in language ST is written. The code used in the program is shown in Fig. 34.

In State 10 the dependence of the running of the program on the position of rotary switch is established. Also in this step the valve opening degree of supply air flow is set, valve is enabled for use and any lingering error statements are acknowledged by setting variable xMan to TRUE.

In state 20 mode of the valve is set to ECO drive and state of the app to Blocked as safe initial setting of the valve.

The program then cycles between states until one of the conditions is met. In this loop manual control is performed as well as the switch to sequence. If the sequence mode is activated via the visualization the program cycles in different states waiting again for one of the conditions that would warrant exiting this loop is met. In all the loops additional control of the state of rotary switch (Hex Switch) is inserted. If the switch has been moved from “run” position the valve will not be enabled.

As was mentioned in visualization “blocked” state as safety element can be activated anytime and sends the program to the manual control loop, if all the conditions for the program to run are met.

This Application, the movements of the actuator, is controlled by the buttons from the Visualisation window.


```

1  PROGRAM PLC_PRG                                48
2  VAR                                             49
3  State : INT;                                  50
4  END VAR                                        51
1  POU( ) ;                                       52
2
3  CASE State OF                                  53
4  0 :                                             54
5      GVL . Vl_Enable := FALSE;                55
6      State := 10;                              56
7
8  10:                                             57
9      IF (GVL . Hex_Switch = 2 ) THEN          58
10     GVL.Vl_Set2:=10000;                       59
11     GVL . Vl_Enable := TRUE;                 60
12     GVL.xMan := TRUE;                        61
13     State := 20;                             62
14     ELSE                                       63
15     State := 0;                               64
16     END_IF                                    65
17
18  20 :                                           66
19     GVL . Vl_Mode := 6 ;                      67
20     GVL . Vl_AppC := 0 ;                      68
21     GVL.xMan := FALSE;                        69
22     State := 100;                             70
23
24  100:                                           71
25     IF GVL . Advance THEN                     72
26     GVL . Vl_AppC := 1 ;                      73
27     State := 120 ;                            74
28     ELSE                                       75
29     State := 120;                             76
30     END_IF                                    77
31
32  120:                                           78
33     IF GVL . Retract THEN                     79
34     GVL . Vl_AppC := 2 ;                      80
35     State := 140 ;                            81
36     ELSE                                       82
37     State :=140;                              83
38     END_IF                                    84
39
40  140:                                           85
41     IF GVL . Exhaust THEN                     86
42     GVL . Vl_AppC := 3 ;                      87
43     State := 160 ;                            88
44     ELSE                                       89
45     State := 160;                              90
46     END_IF                                    91
47
160:                                             92
161     IF NOT ( GVL . Hex_Switch = 2 ) THEN     93
162     GVL . Vl_Enable := FALSE ;              94
163     State := 0;                               95
164     ELSE                                       96
165     State := 180 ;                            97
166     END_IF
167
180:                                             98
181     IF ( GVL.Vl_ActualValveMode = 61 ) THEN  99
182     State := 10;                              100
183     ELSE                                       101
184     State := 190;                              102
185     END_IF
186
190:                                             103
191     IF ( GVL.StartSequence = TRUE) THEN     104
192     State:= 195;                              105
193     ELSE                                       106
194     State :=100;                              107
195     END_IF
196
195:                                             108
197     GVL . Vl_AppC := 2 ;                      109
198     IF (GVL.Vl_ActualAppState = 1) THEN     110
199     State:= 200;                              111
200     ELSE                                       112
201     State:= 195;                              113
202     END IF
203
200:                                             114
204     IF (GVL.Vl_ActualAppState = 4 ) THEN    115
205     GVL . Vl_AppC := 2 ;                      116
206     END_IF
207     IF (GVL.Vl_ActualAppState = 1) THEN     117
208     GVL . Vl_AppC := 1 ;                      118
209     END_IF
210     IF (GVL.EndSequence = TRUE) THEN       119
211     State := 10;                              120
212     ELSE                                       121
213     State := 210;                              122
214     END_IF
215
210:                                             123
216     IF NOT ( GVL . Hex_Switch = 2 ) THEN    124
217     State := 10;                              125
218     ELSE                                       126
219     State := 200;                              127
220     END_IF
221
222     END_CASE
223     IF GVL . Blocked THEN                    128
224     GVL . Vl_AppC := 0 ;                      129
225     State := 0;                               130
226     END_IF

```

Fig. 34 Application POU program in ST

4. Conclusion

In this paper, we have summarized the function of Festo Motion Terminal and its valves. We have also provided a thorough installation and connection guide. Further we demonstrated how to structure a program for FMT application. We have described the components of that program that are mostly given for any such application, thus leaving the reader with a blueprint for designing their own application and developed basis for more complex projects.

Acknowledgements

The authors acknowledge and thank for the initial training provided by Festo, s. r. o. in Prague.

References

- [1] *Motion Terminal VTEM* [online]. Festo, 2018/03, , 33 [cit. 2020-02-18]. Available at: https://www.festo.com/cat/en-gb_gb/data/doc_ENGB/PDF/EN/VTEM_EN.PDF
- [2] Digital pneumatics|Festo Motion Terminal. *Festo* [online]. [cit. 2020-03-30]. Available at: <https://www.festo.com/vtem/en/cms/10169.htm>
- [3] Digital simplicity: World first Festo Motion Terminal VTEM. In: *Festo* [online]. 2017/11, p. 16 [cit. 2020-03-30]. Available at: https://www.festo.com/net/SupportPortal/Files/468005/PSI_Festo_Motion_Terminal_EN.pdf
- [4] FESTO AG & CO. KG. *Motion Terminal VTEM: Instructions for use* [online]. In: . [cit. 2020-07-13]. Dostupné z: https://www.festo.com/net/SupportPortal/Files/469076/VTEM_2017-10_8071722g1.pdf
- [5] FESTO AG & CO. KG. *Festo VTEM DevCon Version 3.5.7.233: Library for the control (Device Control) of the Motion Terminal VTEM from Festo and its elements*. [online]. 2019 [cit. 2020-07-13].

HMI s biologickou zpětnou vazbou

Ondřej Baumrt

Ústav přístrojové a řídicí techniky, FS ČVUT, Ondrej.Baumrt@fs.cvut.cz

Abstrakt: Tento článek se zabývá základním výzkumem k sestavení HMI s biologickou zpětnou vazbou. Jsou zde popsány základní fyziologické parametry člověka, které lze použít ke zjištění aktuálního stavu operátora.

Klíčová slova: HMI, fyziologické parametry, sympatikus, parasympatikus, mydriáza, mióza, EEG, Elektrodermální aktivita

Abstract: This article deals with basic research to create HMI with biological feedback. Basic physiological parameters of a man are described here. They can be used to determine the current operator's status.

Keywords: HMI, physiological parameters, sympathetic, parasympathetic, mydriasis, miosis, EEG, Electrodermal activity

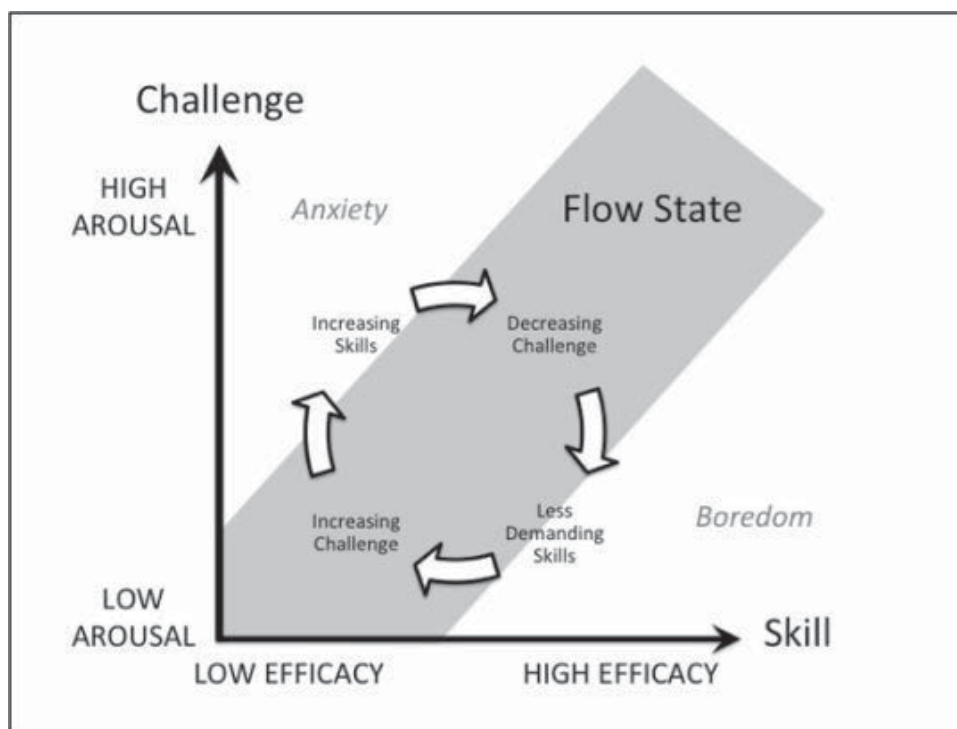
1 Úvod

Fyziologické parametry se v kybernetice dají využít v několika způsobech s tím, že tyto způsoby mohou být i provázány a využívány simultánně. První z těchto oblastí je využití fyziologických parametrů pro zjištění aktuálního stavu operátora.

Druhou možností, jak využít fyziologické parametry, je při návrhu daného zařízení či způsobu jeho využití.

Třetím způsobem využití je vlastní zpětná reakce pro operátora bez jakéhokoli využití pro vlastní proces. Jedná se například o zobrazení stavu jedince jemu samotnému. Kupříkladu u člověka s vyšším krevním tlakem je mu ukázána aktuální hodnota pomocí škály.

Všechny tyto možnosti mají společné to, že se snaží operátora (či jakéhokoli jedince) udržet v co nejideálnějším stavu pro pokračování činnosti, což je popsáno v obrázku 1 níže. [1]



Obr. 1: Graf toku - schopnosti a požadavky [1]

Na obrázku je na horizontální ose vyznačena úroveň schopností (angl. *Skill*) jedince a na vertikální úroveň požadavků (angl. *Demands*) či "výzvy" (angl. *Challenge*). Šedá oblast označuje stav, kdy jsou schopnosti operátora využívány optimálně. Tato oblast se označuje jako oblast toku (angl. *Flow State*). Pokud se jedinec dostane vpravo od této oblasti (čili jeho schopnosti jsou vyšší než požaduje daný úkol), dochází k znužení. Ta může mít za důsledek sníženou koncentraci. Pokud naopak se dostane vlevo mimo oblast "flow", může mít tato situace za následek úzkost. [1]

Existují způsoby, jimiž lze docílit návratu do oblasti "flow", a sice pohybem na grafu:

- **vlevo** - snížením schopností,
- **vpravo** - zvýšením schopností (např. tréninkem)
- **dolů** - snížením požadavků (např. omezení regulovaných parametrů), či
- **nahoru** - zvýšením požadavků (např. přidání regulovaných parametrů).

Je možné tyto "pohyby" vzájemně kombinovat. [1]

2 Fyziologické parametry

Fyziologické parametry, které jsou důležité pro určení duševního stavu souvisí s dvojicí nervových systémů - *Sympatickým* a *Parasympatickým*.

2.1 Sympatikus a parasympatikus

Sympatický a *parasympatický* nervový systém, neboli *sympatikus* a *parasympatikus*, jsou součástí autonomního nervového systému, které se podílejí na řízení cévního oběhu nebo například činnosti určitých orgánů. [2][3]

Obvykle vzájemně působí opačně, přičemž sympatikus je aktivován ve stresových situacích a parasympatikus v útlumových případech (např. jedinec je v klidu, po jídle apod.). Sympatický nervový systém kupříkladu zvyšuje činnost srdce nebo rozšiřuje oční zornice (tzv. *mydriáze*). Jeho efekt na cévy je rozličný dle orgánu, nicméně převažuje zúžení cév (tzv. *vazokonstrikce*) a zvýšení krevního tlaku. Na druhou stranu parasympatikus zvyšuje činnost trávicího ústrojí, oční zornice zužuje (tzv. proces *míózy*) nebo zužuje průdušky. [2][3]

Pokud je lidský organizmus v pořádku, pracují oba systémy přibližně stejnou dobu. [2][3]

2.2 Kontrakce zornic

2.2.1 Mydriáza

Mydriáza je proces rozšiřování oční zornice a jak bylo řečeno v kapitole 2.1, je způsobena aktivací sympatického nervového systému. Zornice se kromě nízkého osvětlení okolí mohou rozšířit kvůli působení stresové situace na jedince. Pochopitelně mydriáza probíhá i z jiných důvodů (nemoc či konzumace léků a drog). [4]

2.2.2 Mióza

V návaznosti na kapitolu 2.2.1, jevem opačným k mydriáze je tzv. *mióza*, proces zúžení zornic. K němu dochází působením parasympatiku, obvykle v klidných situacích či při vysokém osvětlení okolí. Stejně jako v případě mydriázy může být i zúžení zornic způsobeno užíváním určitých léků a drog nebo nemocí. [5]

2.3 Oční pohyby

Pohyby očí lze rozdělit do několika kategorií:

- *sakády*, což jsou rychlé pohyby, které zajišťují, aby byl obsah zájmu zobrazen na střed sítnice co nejrychleji a jsou kontrolovány vědomě;
- *sledovací pohyby*, tedy pomalé pohyby, pomocí kterých je sledování pohybujících se předmětů tak, aby byly zobrazeny na sítnici ostře;
- *vergenční pohyby*, zajišťující zobrazení různě vzdálených objektů na korespondující místa na sítnici; a
- *reflexní pohyby*. [6]

2.4 Mozková aktivita

2.4.1 Základní neurohormonální reakce na stres

Základní reakce organismu na stres má tři fáze:

- poplachovou reakci,
- adaptaci (rezistence) a
- vyčerpání. [7]

Poplachová reakce, označovaná také jako *Cannonův stres* spočívá v aktivaci sympatiku, přičemž se více vyplavují hormony dřeně nadledvinek, konkrétně adrenalin a noradrenalin (zároveň se zvyšuje krevní tlak, jak je popsáno v kapitole 2.1. [7][8])

V této, první, fázi také roste produkce CRH¹-ACTH²-kortizolu, který zvyšuje energetický potenciál jedince. Reakce systému CRH-ACTH-kortizolu je hlavním rysem druhé fáze - adaptace, jinak jmenovanou jako *Selyeho stres*. V této etapě žije člověk stále a je v podstatě základem normálního života. Pokud by totiž z lidského života vymizely veškeré stresující okolnosti, samovolně by stres naopak vznikl. [7][8]

Poslední fází je vyčerpání. Je zapříčiněna zhroucením homeostatických mechanismů, totiž když CRH-ACTH-kortizol má nedostatečnou produkci. Obvykle nastává pouze za extrémních situací (např. nemoc). [7][8]

2.4.2 Elektroencefalografie

Elektroencefalografie sleduje mozkové vlny, které jsou součástí elektromagnetického pole, které vytvářejí svou prací neurony. Obvykle je zmiňováno čtyři až pět druhů, a sice:

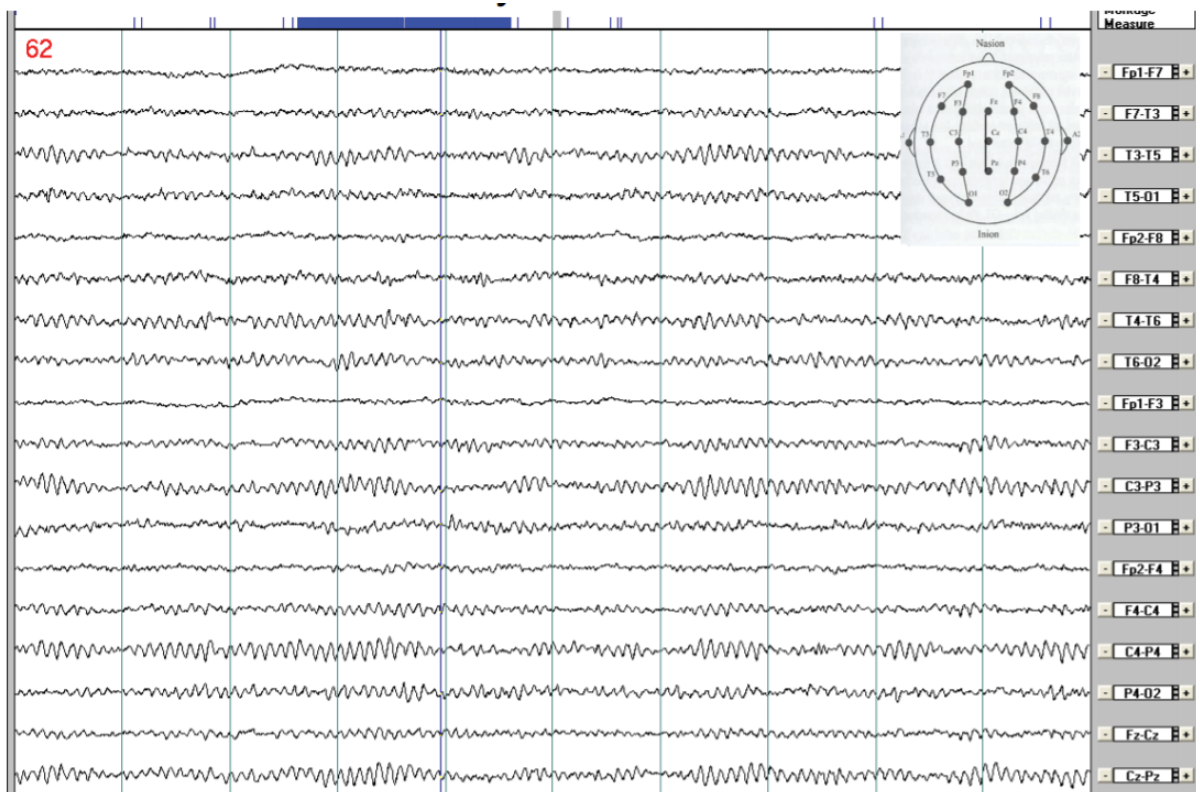
- *alfa vlny*, které mají frekvenci $8 \div 13 \text{ Hz}$ a převládají za stavu uvolnění a odpočinku;
- *beta vlny* s frekvencí $13,5 \div 40 \text{ Hz}$, které jsou aktivní v bdělém stavu, přičemž čím vyšší frekvenci tyto vlny mají, tím je člověk "podrážděnější";
- *gamma vlny* o frekvenci $40 \div 100 \text{ Hz}$ používané v situacích, kdy je mozek vystaven náročným úkolům;

¹Kortikoliberin - peptidový hormon hypotalamu [9]

²Adrenokortikotropní hormon, je produkován adenohypofýzou [9]

- *delta* vlny dosahující frekvencí pouze $1 \div 3,5$ Hz jsou aktivní při hlubokém spánku nebo bezvědomí; a
- *theta* vlny, jejichž frekvence osciluje mezi 4 a 7,5 Hz, převládají během hluboké relaxace či některých fázích spánku. [10][11]

Příklad výsledné křivky je přiložen na obrázku 2. Na obrázku je použito zapojení systémem "10-20". Pro udržení elektrod na správném místě se obvykle používá lepkavá pasta. Obvykle se dané elektrody označují písmenem, které označuje oblast mozku (např. "T" označuje oblast *temporální*) a nejčastěji číslem (lichá pro levou hemisféru, sudá pro pravou) případně malým písmenem "z" či nulou pro elektrody nepárové (tzv. "vertexové"). Umístění elektrod lze vidět na obrázku 2 v pravém horním rohu. [10][12]



Obr. 2: Křivka normálního EEG [10]

2.5 Elektrodermální aktivita

Elektrodermální aktivita (EDA), dříve nazývaná jako *galvanická reakce*, je kožní reakce na elektrický podnět a je velmi často využívána v psychofyziologii. [13][14]

Oproti například *EEG* má velkou výhodu v tom, že jeho použití je technická nenáročnost a snadná aplikace téměř kdekoli, přičemž stále vykazuje velmi dobrou citlivost na zjištění psychického stavu. [13][14]

Elektrickou aktivitu popsali nezávisle na sobě francouzský neurolog *Charles Féré* (v roce 1888) a ruský fyziolog *Tachograff* (v roce 1890). Oba ovšem ke svým závěrům došli odlišnou metodou. Zatímco *Féré* využil stejnosměrný obvod (v podstatě externí stejnosměrný zdroj zapojený na dvě různá místa na kůži), *Tachograff* snímal pouze potenciál mezi dvěma místy na kůži - bez vnějšího zdroje. Veličiny měřené při *EDA* se dělí na:

- endosomatické a
- exosomatické. [13][14]

Jako *endosomatické* označují biofyzikové takové veličiny, které se měří bez zdroje vnějšího napětí. Zde je vyjmenován *kožní potenciál (skin potential - SP)*, což je napětí (rozdíl potenciálů) mezi dvěma místy na pokožce. [13][14]

Za *exosomatické* veličiny lze označit takové, které dle filozofie *Férého* využívají externí zdroj. Teoreticky lze využít i střídavé napětí, avšak prakticky se využívají pouze stejnosměrné zdroje. Mezi exosomatické veličiny patří například *kožní vodivost (skin conductance - SC)* nebo *kožní odpor (skin resistance - SR)*. [13][14]

3 Závěr

Výsledkem tohoto článku je lepší pochopení fyziologických parametrů člověka za účelem jejich využití ve strojírenské praxi pro vytvoření kvalitnějšího HMI, např. z pohledu využití schopností operátora.

Je zřejmé, že velké množství informací o aktuálním stavu lze zjistit ze snímání očí či z mozkové aktivity. Ve velkém množství parametrů se projevuje aktivace sympatického, popř. parasympatického nervového systému.

Článek byl podpořen grantem: SGS20/159/OHK2/3T/12

Literatura

- [1] M. Csikszentmihalyi, "Flow: The Psychology of Optimal Experience", Goodreads, 19-úno-2019. [Online]. Dostupné z: <https://www.goodreads.com/review/show/2714935491>.
- [2] "Sympatický nervový systém (sympatikus)", Velký lékařský slovník. [Online]. Dostupné z: <http://lekarske.slovníky.cz/lexikon-pojem/sympaticky-nervovy-system-sympatikus-5>. [Viděno: 12-čer-2019].
- [3] "Parasympatický nervový systém", Velký lékařský slovník. [Online]. Dostupné z: <http://lekarske.slovníky.cz/lexikon-pojem/parasympticky-nervovy-system-hovor-parasympatikus-6>. [Viděno: 12-čer-2019].
- [4] J. Štefánek, "Rozšířené zornice", Medicína, nemoci, studium na 1. LF UK. [Online]. Dostupné z: <https://www.stefajir.cz/roz sirene-zornice>. [Viděno: 13-čvc-2019].
- [5] J. Štefánek, "Zúžení zornic", Medicína, nemoci, studium na 1. LF UK. [Online]. Dostupné z: <https://www.stefajir.cz/zuzeni-zornic>. [Viděno: 13-čvc-2019].
- [6] K. Skorkovská, "Supranukleární okohybné poruchy", prezentováno v Neurooftalmologie, LF MUNI, 08-lis-2015.
- [7] V. Schreiber, "Současný pohled na stres a endokrinní odpověď", Interní medicína pro praxi, roč. 3/2004, s. 111–112.
- [8] A. Večeřová-Procházková a R. Honzák, "Stres, eustres a distres", Interní medicína pro praxi, roč. 2008, s. 188–192.
- [9] "Adrenokortikotropní hormon", WikiSkripta.eu, 31-říj-2016. [Online]. Dostupné z: https://www.wikiskripta.eu/w/Adrenokortikotropn%C3%AD_hormon. [Viděno: 10-říj-2019].
- [10] H. Krijtová, "Standardní EEG - základy, indikace a základní nálezy", prezentováno v Kurz specializačního vzdělávání, Základy neurologie, Neurologická klinika, 2. LF a FN Motol, 2016.
- [11] T. Procházková, "Mozkové vlny: neviditelné tajemství v naší hlavě", Mentem, 11-dub-2016. [Online]. Dostupné z: <https://www.mentem.cz/blog/mozkove-vlny/>. [Viděno: 16-zář-2019].
- [12] M. Matoušek, "EEG v psychiatrii", in Psychiatrie, 1., Praha: Tigis, 2002, s. 276–282.
- [13] R. Freeman a M. W. Chapleau, "Testing the autonomic nervous system", in Handbook of Clinical Neurology, 1. vyd., roč. 2013, Elsevier, s. 115–136.
- [14] W. Boucsein, Electrodermal Activity, 2. vyd. Springer Science Business Media, 2012.

KAMEROVÝ SYSTÉM S VLASTNÍ DIAGNOSTIKOU

Petr Písařík¹, Matouš Cejnek¹

¹ *Ústav přístrojové a řídicí techniky, Fakulta Strojní, ČVUT v Praze*

Abstrakt: Tato práce popisuje tvorbu kamerového systému, který bude možné použít k monitorování objektů či procesů. Výběr vhodného hardware splňující požadavky. Schéma zapojení kamerového systému. Software kamerového systému je využívá pythonu3 a nástrojů linuxového prostředí Ubuntu Mate 18. Celkový postup implementace a testování.

Klíčová slova: IP-Cam, Python-3, Kamerový systém, Raspberry B3+, Ubuntu, Linux, ffmpeg, cron

Abstract: This work describes the creation of a camera system that can be used to monitor objects or processes. Selection of suitable hardware that meets the requirements. CCTV wiring diagram. The camera system is using Python3 and Linux environment Ubuntu Mate 18. Overall implementation and testing process.

Keywords: IP-Cam, Python-3, Cam system, Raspberry B3+, Ubuntu, Linux, ffmpeg, cron

1 Úvod

Kamerové systémy jsou nedílnou součástí života dnešních lidí. Na trhu je nepřehledná nabídka komerčních řešení již hotových kamerových systémů. V tomto článku jsou popsány některé aspekty návrhu low-cost domácího kamerového systému s důrazem na vlastní monitoring zdraví kamerového systému. V následujících sekcích je popsáno jak a proč byl zvolen použitý hardware a software. Jak byla navržena architektura samotného software řešení. Dále zde jsou také popsány různé klíčové problémy, na které je možné při tvorbě systému narazit, a jak je řešit.

2 Hardware

Použité zařízení, musí splňovat zadané požadavky, tedy být odolné proti venkovnímu prostředí. Umožňovat noční režim. Napájení pomocí PoE (Power over Ethernet) [1]. Poskytovat živý obraz pomocí RTSP (real time stream protokol) [2] a být dosažitelné z internetu. Spolehlivě pracovat i s delšími časovými úseky. Hardware kamerového systému je logicky rozdělen do tří částí: senzorické, síťové a výpočetní.

2.1 Senzorika

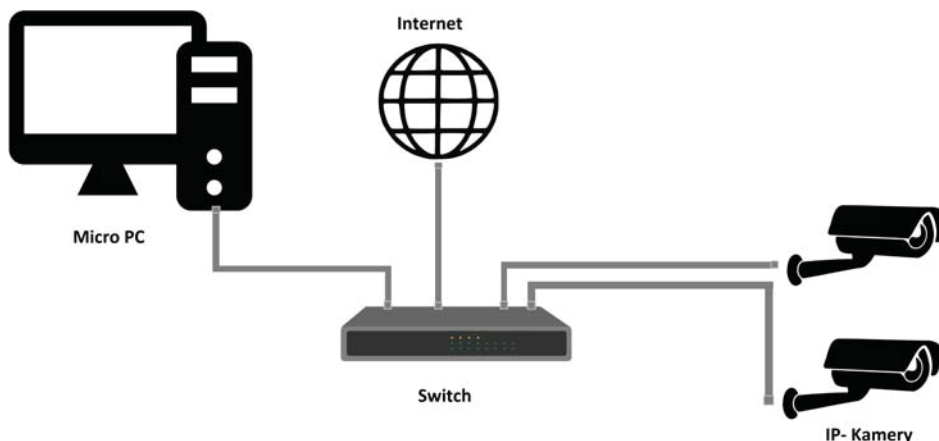
Jako kamera splňující požadavky se ukázala IP Kamera IPSU20HHF200 od firmy Cantonk.

2.2 Síťové připojení

Síťovou část propojující kamery s mikro počítačem a mikro počítač s internetem je realizována pomocí switchu: TL-SF1008P [3]. Poskytující 4 POE porty, kterými je zajištěno napájení i datový přenos kamer.

2.3 Výpočetní zařízení

Jako výpočetní člen je použit mikro počítač Raspberry pi B3+. K mikro-PC je připojený externí HDD disk, jako prostor pro úložiště záznamů.



Obr. 1: Schéma zapojení jednotlivých komponent systému.

3 Software

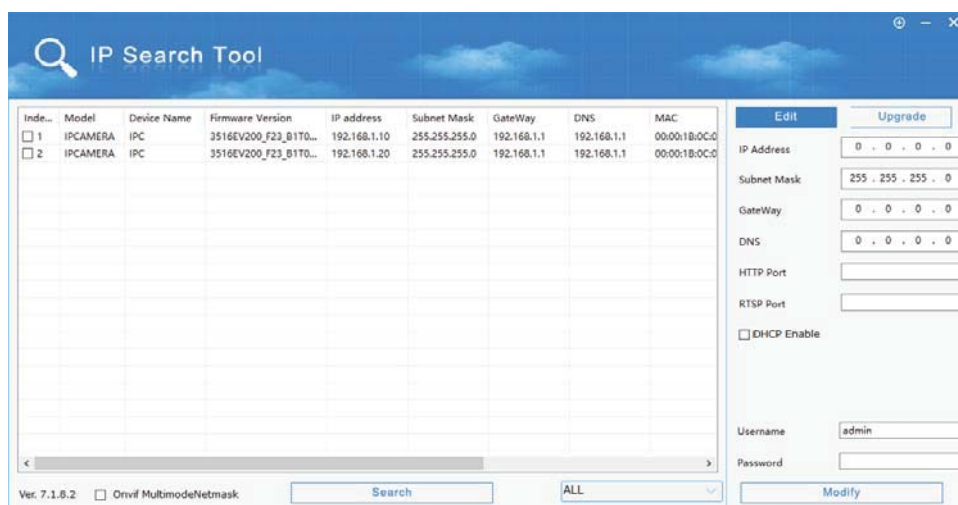
Systém je naprogramovaný v Python 3 [4] a utilitách linuxu. Hlavním cílem je, aby systém umožňoval real-time náhled, ukládal záznam kamer ve smyčce a nejstarší část záznamu automaticky odstraňoval, aby se průběžně uvolňovalo místo na disku. Vše běží v linuxovém prostředí pro Raspberry Pi B3+ Ubuntu Mate verzi 18.04.

4 Implementace

Prvním krokem k zprovoznění systému je instalace operačního systému Raspberry Pi. Na oficiálních stránkách [5] jsem vybral nejnovější verzi operačního systému Ubuntu Mate. Ubuntu Mate jsem zvolil kvůli pohodlnému desktopovému přístupu. Dle instrukcí jsem vytvořil boot-SD card. Po ověření správného nainstalování systému následovala instalace Pythonu 3.8 společně s IDLE ze stránek [6].

4.1 Připojením kamer

Kamerový systém je zapojen podle schématu Obr.1. Raspberry je připojeno pomocí LAN k switch. Kamerové POE LAN zajišťuje zároveň napájení kamer a datový tok od kamer. V takto zapojeném systému jsem vyhledal připojená zařízení pomocí IP Search [7]. Mezi nalezenými zařízeními byly i IP kamery v tovární konfiguraci. Pro budoucí použití bylo nutné kamery vhodně překonfigurovat.



Obr. 2: Nalezení IP kamer dostupných v místní síti a jejich konfigurace

Ke konfiguraci jsem využil výše zmíněný program IP Serch Tool. U kamer bylo nutné nastavit IP adresu tak, aby odpovídala místní síti. Kromě IP adresy jsem nastavil i odpovídající Mask a Gate Way. Zapsal jsem si

označení RTSP portů, na kterých kamera vysílá stream. Dalším krokem bylo připojit se ke kamerám a změnit tovarní přístupové údaje. Hardware kamerového systému byl těmito modifikacemi připraven k provozu.

4.2 Přístup a záznam RTSP streamu

Ve VLC player [8] jsem nejdříve ověřil existenci kamerových streamů a jejich dostupnost. V záložce síťových proudů jsem vyplnil IP adresu, číslo portu RTSP a přihlašovací údaje. Následovala implementace v Pythonu. Využitím knihovny openCV jsem navázal spojení s kamerami a v živém modu zobrazil oba RTSP streamy. Problém nastal až při pokusech o ukládání záznamu. V pythonu jsem si vytvořil skript, který založil video soubor s datem a časem vytvoření. Do tohoto souboru se ukládala přijímaná data z kamery. Skript byl doplněn i o automatické mazání souborů starších než předepsaný čas. Zkušební provoz odhalil, že takto běžící skript, který zobrazuje a ukládá živý obraz z jedné kamery využívá výkon mikro PC zhruba na 90%. Záznam z obou kamer vedl k přehřívání zařízení a ukončení běžících procesů. Laděním skriptu, snižováním kvality (snížení rozlišení a FPS (počet snímků za sekundu)) přijímaného videa se mi podařilo snížit výpočetní náročnost na cca 80%. Problémem byla také časová nespojitost mezi jednotlivými nahranými segmenty. Segmenty byly odděleny časovými prostory v řádech několika sekund (2-13 sec). Pro záznam obou kamer byl výkon mikro PC stále nedostatečný.

Ukládání streamu pomocí pythonu v požadované kvalitě se ukázalo jako neuskutečnitelné. Z těchto důvodů jsem se rozhodl použít alternativní způsob záznamu RTSP streamu. Pro nahrávání jsem se rozhodl otesovat linuxové knihovny a programy umožňující live konverzi RTSP streamu (openRTSP), ffmpeg [9] a vlc [8]. Z testování se nejlépe vyšlo ffmpeg. OpenRTSP se ukázalo příliš náchylné na stabilitu a při drobném kolísání signálu docházelo ke kolapsu nahrávání. Také bylo složité přesně nastavit parametry aby došlo ke správnému záznamu. Automatický záznam pomocí cvlc bylco se týče předepisování parametrů nejjednodušší. Stačilo nastavit pouze adresu streamu, rozdělení na časové segmenty a jména souborů. Kvalita záznamu byla v porovnání s openRTSP horší. Docházelo zde ke krátkým usekům mezi jednotlivými segmenty kdy se nenahrávalo (1-3 sec). Rozlišení samotného záznamu bylo stejné jako v streamu (1920x1080), avšak záznam byl sekaný. Zaznamenaný soubor běžel s nahodilými záseky, kdy se na videu na několik sekund zastaví obraz a pak běží dál. Optimálním nástrojem pro záznam se ukázal ffmpeg.

4.3 FFMPEG

Fmpeg je linuxová utilita pro příkazový řádek pro konverzi video formátů. Umožňuje konverzi streamování a nahrávání digitálního obrazu a zvuku v reálném čase. Neméně důležité je znát i parametry streamu jako je codec, FPS, rozlišení atd. Mnou použitý příkaz pro záznam je následující:

```
>>>ffmpeg -rtsp_transport tcp -i rtsp://login:password@IP_address:RTSP_port -codec:v
copy -codec:a copy -map 0 -f segment -segment_time 28800 -segment_format mkv
-strftime 1 -reset_timestamps 1 /path/name_%Y-%m-%d_%H-%M-%S.mkv
```

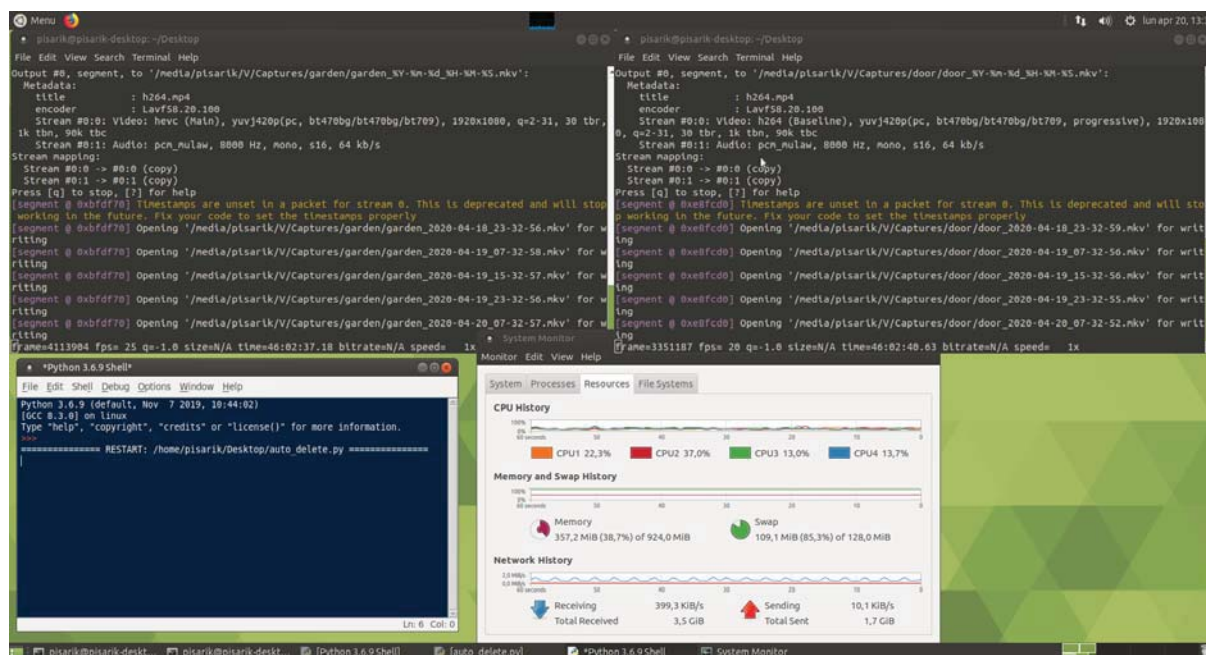
Kde *ffmpeg* je příkaz spouštějící utilitu, *rtsp transport* spouští rtsp transportní protokoly, *tcp* volba TCP jako nižšího transportního protokolu, *i* aktivace nastavení vstupu (input), *rtsp://login:password@IP address:RTSP port* nastavení adresy streamu společně s přihlašovacími údaji a portem rtsp streamu, *-codec:v copy -codec:a copy* volba codecu automaticky podle zdroje, *map 0* mapování jednoho streamu, *f segment -segment time 28800* rozdělení záznamu na 8 hodinové úseky, *segment format mkv* volba výstupního formátu souboru. v našem případě je vhodné použít spíše formáty *mkv* a *mp4*. Použití formátu *avi* vedlo k potížím z důvodů nekonzistence přijímaného signálu. Formát *avi* nedokáže zpracovat dlouhodobě video ve kterém se občas ztratí data. Poslední část *-strftime 1 -reset timestamps 1 /path/name %Y %m %d %H-%M-%S.mkv* nastavuje cestu kam ukládat záznam a jeho jméno tvořeného prefixem a aktuálním časem.

Stream ve formátu h264 je převáděn do mkv po 8 hodinových úsecích. Mkv formát umožňuje ztrátu učitěho množství dat. Mezery mezi jednotlivými záznamy jsou zanedbatelné. Nahrávání je stabilní a může běžet v řádech dní.

Celkové výkonové požadavky zatěžují mikro PC 30% maximálního výkonu, jak je vidět z (3). Takto zavedený systém tedy zaznamenává a automaticky odstraňuje staré soubory.

4.4 Automatické spuštění

Pro spolehlivý chod kamerového systému je také nutné, aby se při výpadku automaticky spustil. Proto jsem nastavil automatické spuštění všech skriptů po rebootování systému. Do operačního systému jsem také zavedl pravidelné restartování mikro PC a kamer jedenkrát za týden. V definovaný čas dojde k uložení záznamu a vypnutí všech zařízení. Po spuštění systému se jako první spustí python skript automatického mazání, který následují



Obr. 3: Nahrávání dvou kamer spolu s python skriptem, který automaticky odstraňuje soubory starší 5 dní.

skripty zaznamenávající kamery. Automatické spuštění je docíleno softwarovým démonem Cron [10], což je vlastně specializovaný systémový proces sloužící jako plánovač dějů. Pro automatické spuštění bylo nutné pomocí příkazů `sudo chmod -x /path/to/my/script.sh` nastavit, aby bylo skript možné spustit. Všechny takto spouštěné skripty bylo třeba doplnit o první řádek `#!/bin/sh`. Bez těchto úprav by nedošlo k jejich aktivaci.

```
@reboot /path/to/my/script/auto_delete.sh
@reboot /path/to/my/script/capture_cam_A.sh
@reboot /path/to/my/script/capture_cam_B.sh
00 03 * * */3 /path/to/my/script/reboot.sh
```



Obr. 4: Syntax zápisu v Cron. `@reboot` znamená aktivaci vždy po zapnutí. Příkaz na poslední řádce znamená spuštění skriptu reboot každou středu ve tři hodiny v noci [11].

5 Vzdálený přístup

Pro kontrolu funkčnosti a přístupu k zánamům jsem na raspberry pi vytvořil FTP (File Transfer Protocol) servr. Kvůli většímu zabezpečení je vhodné použít zabezpečenější verzi SFTP (Secure File Transfer Protocol). Soubory přístupné pro klienty jsou pouze soubory umístěné na externím disku. Přihlásit se jde do tří různých uživatelských úrovní. Administrátorský přístup umožňuje správu a editaci souborů. Uživatelský přístup je pak pouze pro náhled bez možnosti editací souborů. Třetí úroveň povoluje spuštění skriptů a umožňuje nastavování systému.

6 Závěr

Vybral jsem hardware kamerového systému a instaloval jej na budovu. Pomocí raspberry pi jsem vytvořil systém jenž ukládá záznam dvou kamer do 8 hodinových segmentů. K mikro PC jsem připojil externí hard disk, sloužící jako prostor pro data. Jména video segmentů jsou tvořena označením kamery a časem kdy byla stopa založena. Automatické mazání skriptů je uskutečněno skriptem v pythonu, který odstaní soubory starší než předepsaný čas. Vše se spouští automaticky po startu systému Ubuntu mate.

Systém byl extenzivně testován 168 hodinama provozu při podmínkách, v kterých se bude běžně používat.

Poděkování

Hlavní autor článku by rád poděkoval společnosti Alphabet Inc. za provoz služby Google, bez které by tato práce nebyla v této míře uskutečnitelná.

Práce Matouše Cejnika byla podpořena z grantu *SGS18/177/OHK2/3T/12*.

Literatura

- [1] *IEEE standard for Ethernet*. IEEE, New York, 2016. OCLC: 1017937152.
- [2] Henning Schulzrinne, Anup Rao, and Robert Lanphier. Real time streaming protocol (rtsp). 1998.
- [3] Download for TL-SF1008P. Library Catalog: www.tp-link.com.
- [4] Guido Van Rossum and Fred L. Drake. *Python 3 Reference Manual*. CreateSpace, Scotts Valley, CA, 2009.
- [5] Ubuntu MATE Team. Ubuntu MATE | For a retrospective future. Library Catalog: ubuntu-mate.org.
- [6] IDLE — Python 3.8.2 documentation.
- [7] Advanced IP Scanner - stáhněte si bezplatný nástroj na skenování sítě.
- [8] VideoLan. VLC - Features - VideoLAN, 2020.
- [9] Suramya Tomar. Converting video formats with ffmpeg. *Linux Journal*, 2006(146):10, 2006.
- [10] Larry Reznick. Using cron and crontab. *Sys Admin*, 2(4):29–32, 1993.
- [11] Suhesh KS. Crontab tips and tricks, June 2016. Library Catalog: www.eazylinux.com Section: Linux.

DETEKCE FYZICKÝCH OBJEKTŮ NA DOTYKOVÉM PANELU (DETECTION OF PHYSICAL OBJECTS ON A CAPACITIVE TOUCH PANEL)

Jan Mareš

Abstrakt: Tato práce je zaměřena na dotykové obrazovky. V práci je proveden základní výzkum a popis v současnosti používaných dotykových displejů.

Klíčová slova: Kapacitní displej, Dotyková obrazovka, Odporový displej

Abstract: This work aims on touchscreens, fundamental research on how touchscreen work and what types of touchscreen exists.

Keywords: Display, Capacitive Touch, Resistive Touch,

1 Úvod

Tato práce se zabývá tématem dotykových displejů. V první části bylo potřeba udělat rešerši na druhy displejů, jejich zpracování, výhody a nevýhody oproti ostatním, zjistit jejich současnou míru použití na trhu a poté je rozdělit do skupin a kategorizovat. Hlavní a nejvíce používané displeje budou představeny v této práci, ostatní méně používané poté budou popsány až v diplomové práci, ve které bude řešena i problematika, kvůli které tato práce také vznikla. Jedná se o problém firmy EDAG Engineering s.r.o., která má problémy s dotykovým panelem a cílem práce se problém identifikovat a pokusit se vyřešit.

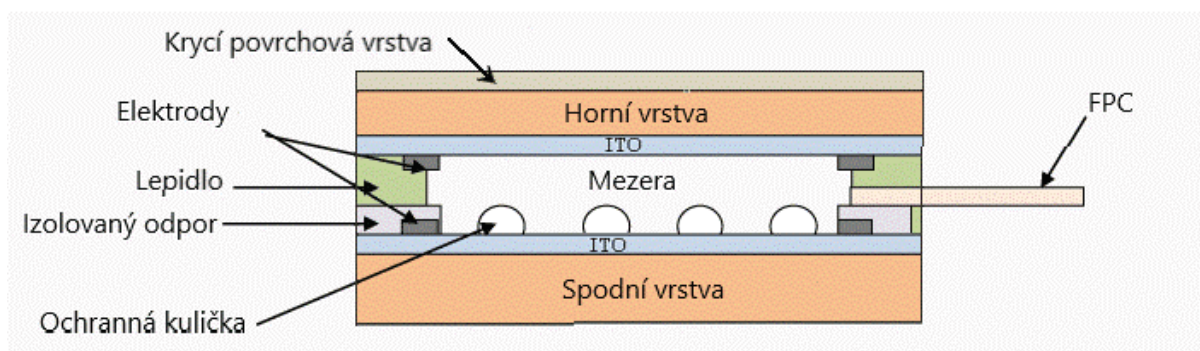
2 Druhy displejů

2.1 Historie

Historicky první dotyková obrazovka, ovládaná pomocí prstu, byla popsána v 1965 americkým vědcem E. A. Johnsonem, který ji v následujících letech ještě podrobněji popisoval. V 70. letech už vědci z CERNu, Ben Stumpe a Frank Beck, kteří vytvořili reálnou transparentní obrazovku. Progres ve vývoji displejů byl poté pozvolný a až v roce 1982 vznikl první displej podporující multi-touch, tzn. detekci více objektů v jeden čas. Za tímto projektem stáli studenti z Univerzity v Torontu. V roce 1993 přichází poprvé na scénu předchůdci telefonů, PDA počítače (Personal Digital Assistant), jejichž představením začíná přebírat vývoj dotykových senzorů a telefonů soukromý sektor, ve zmíněném roce firmy Palm, IBM a Apple. K masivnímu rozšíření dochází ale až po roce 2008, poté co Apple prostřednictvím Steva Jobse představuje iPhone první generace. iPhone disponoval kapacitní multi-touch dotykovou obrazovkou a i když se poté využívali i telefony s rezistivní technologií, kapacitní technologie v současnosti masivně převyšuje druhou zmíněnou technologii. Po telefonech se dotykový displej začal rozšiřovat do všech možných odvětví a velikostí. V dnešní době je masivní produkce tabletů, displeje se objevují už i na počítačích, automobilech nebo výrobních linkách jako operátorské panely. Z výše zmíněného se dá očekávat, že displeje se budou i do budoucna hojně využívat, takže se dá očekávat i jejich postupné vylepšování, jako např. 3D dotyk, ještě odolnější materiály atd. [1]

2.2 Rezistivní displej

Rezistivní, též odporový dotykový panel, se skládá z několika elektricky vodivých vrstev, mezi kterými se nachází tenká membrána. Vrstvy mohou být ze skla, ale dají se použít i další materiály, které jsou průhledné a dají se značknout, například plasty. Membrána je vodivá vrstva, kdy se u většiny displejů využívá nejčastěji slitina oxidu inditého a cínitého (ITO) a tenká vrstva vzduchu. Princip funguje tak, že po stisku ruky nebo jiného předmětu dojde k propojení obou vrstev. V každé z vrstev jsou dvě elektrody, nacházejících se na protilehlých koncích, jsou osy X a Y. Odpor ITO vrstvy vytváří na kontaktním místě dělič napětí. Poměr napětí se následně používá pro výpočet dotykové polohy. Aby se vrstvy nedotýkaly samovolně bez dotyku, je spodní vrstva ještě opatřena ochrannými kuličkami, viz Obrázek 1, kde je i názorně popsán celý displej.



Obrázek 1 – Složení rezistivního displeje **Error! Reference source not found.**

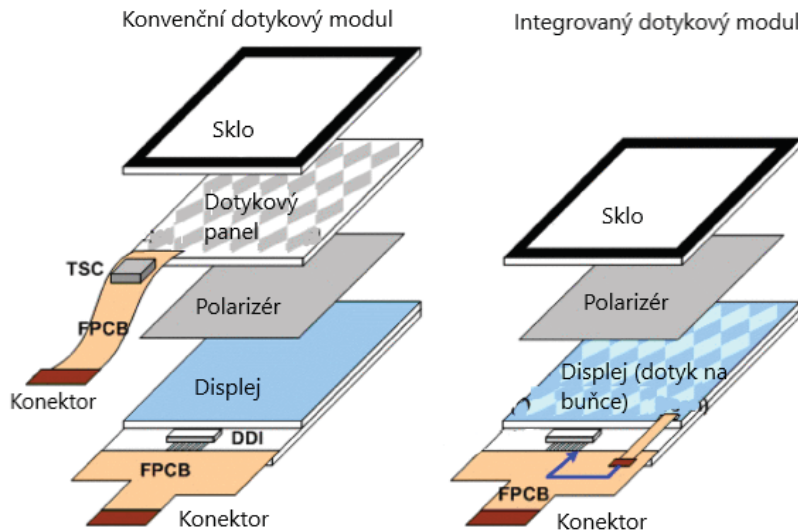
Existují 4 metody určování polohy dotyku: čtyř vodičová metoda, pěti vodičová metoda, osmi vodičová metoda a Digital matrix (Digitální matice). Nejběžnější a nejjednodušší metoda je čtyř drátová, kdy spodní vrstva snímá polohu X a vrchní vrstva snímá polohu Y. Další metody jsou v podstatě variacemi čtyř drátové metody, jen s více elektrodami pro snímání polohy.

Mezi výhody odporových displejů oproti displejům kapacitním patří nízká spotřeba energie, vysoké rozlišení (4096 x 4096 DPI a více), možnost dotyku obrazovky jakýmkoliv předmětem a nízká cena, kvůli jednoduché struktuře čtyř drátových displejů. K nevýhodám naopak patří vzduchová vrstva mezi vrstvami, kdy může snížit průhlednost displeje a může se do vrstvy dostat prach. Nevýhodou je také omezená maximální velikost displeje.

2.3 Kapacitní displej

Oproti rezistivnímu a výše zmíněným displejům pracuje kapacitní displej na principu detekce vodivosti. K doteku displeje tedy je zapotřebí vodivého materiálu, a proto se využívá hlavně vodivosti lidského těla.

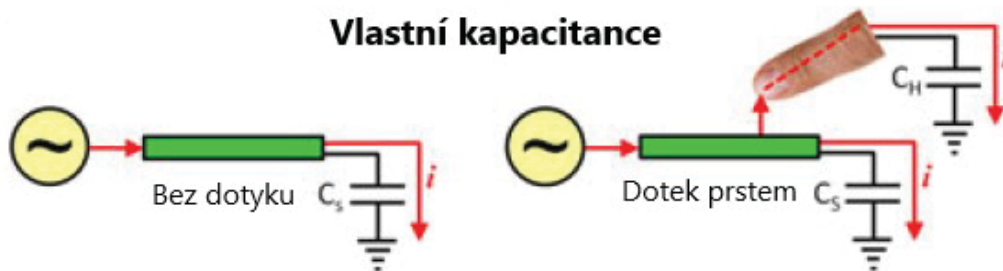
I kapacitní displej má dva druhy zpracování, vlastní a vzájemně kapacitní. Tyto displeje si rozebereme podrobněji v následujících kapitolách. Dva druhy zpracování mohou mít i displeje jako takové, kdy buď může být dotykový panel přímo integrovaný do displeje nebo může být panel oddělený. V dnešní době už se kvůli jednodušší instalaci a výměně používá integrovaný dotykový modul, obě z provedení jsou znázorněny na Obrázku 2.



Obrázek 2 - Typy řešení dotykového displeje [2]

2.4 Vlastní (povrchový) kapacitní displej (Self-capacitance)

Povrchový kapacitní displej se skládá ze tří vrstev. Sklo, vrstva vodivého substrátu a na povrchu je ochranná vrstva. Do 4 rohů vodivé vrstvy je přiváděno napětí stejné fáze, které vytvoří uniformní elektrické pole. Toto napětí je přiváděno průsvitnými elektrodami umístěnými kolem ochranného obalu v maticové formaci. Když se prst dotkne displeje, začne jím téci elektrický proud. Následně se změní proud, které teče z jednotlivých elektrod do prsty. Vzdálenost dotyku je nepřímo úměrná tomuto proudu. Kapacitance může být systémem měřena už od malých výkyvů, řádově 0,5-20 pF.

Obrázek 3 - Vlastní kapacitance **Error! Reference source not found.**

Jedna z elektrod (vertikální) udržuje konstantní hodnotu proudu, když není dotykový displej používán. Další elektrody (horizontální) se spouští v momentě dotyku prstem obrazovky, což inicializuje tok proudu v této části displeje. Formace matic vytváří elektrostatické pole, kde se tyto 2 linie protnou. Díky tomu displej reaguje na dotyk dokonce před tím, než opravdu dojde ke skutečnému dotyku.

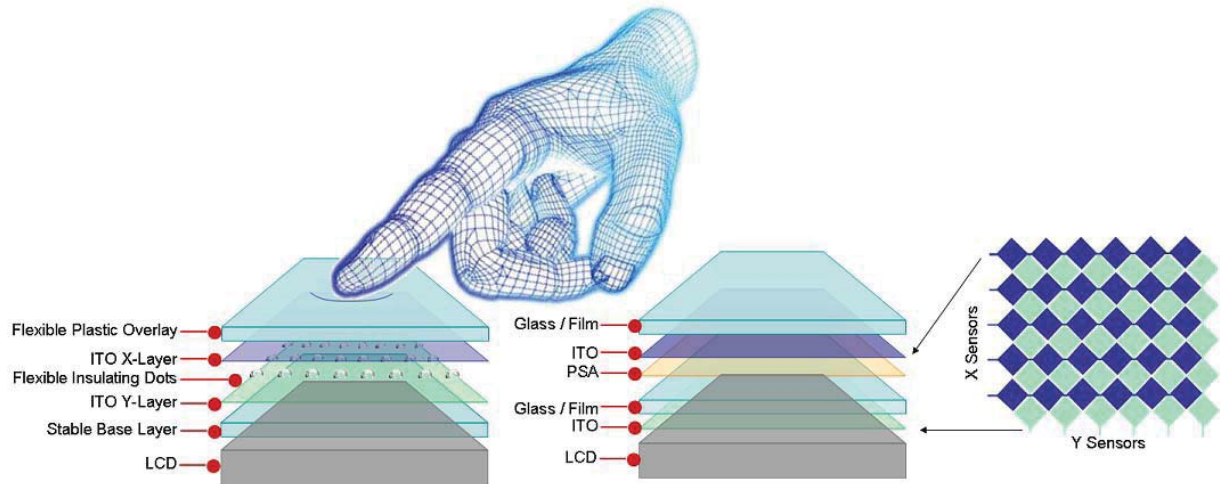
2.5 Porovnání

Odporový displej má výhodu v možnosti dotyku i nevodivými materiály, jako například stylusy, nehtem nebo klidně i v rukavicích, jelikož je nutné k zaznamenání dotyku displej zmáčknout a propojit tím vrstvy. Zatímco u

kapacitního displeje se dotyk pozná už jen lehkým dotykem prstu, ale nefunguje dotyk nehtem ani v rukavicích, pokud nejsou opatřeny dotykovou vrstvou.

Cenově se displeje zlevňují a rozdíly se mezi nimi zmenšují, ale stále je odporový displej levnější. Ovšem o trochu dražší kapacitní umožňuje dotyk na více místech zároveň, tzv. multi touch. K výhodám kapacitního displeje ještě patří všeobecně větší odolnost, ať už proti prachovým částicím nebo mastnotě z prstů.

Jiná je i skladba displeje, která je názorně ukázána na Obrázku 2.



Obrázek 4 - Porovnání technologií **Error! Reference source not found.**

3 Závěr

Z rešerše na hlavní typy displejů jsme zjistili hlavní výhody a nevýhody, které byly uvedeny v předchozí kapitole. Pro dotyk na odporovém displeji může sloužit i jiný předmět, kapacitní displej je spíše přizpůsoben dotyku lidské ruky a má ve výsledku lepší vlastnosti, proto je v dnešní době více využíván.

To, co je jako výhoda normálně, není ovšem výhodou pro aplikaci firmy EDAG Engineering s.r.o. Jelikož je v demonstraci nutno stavět na displej maketu automobilu ovládanou skrze Arduino Mega2560, vhodnější by bylo použít displej rezistivní.

Literatura

- [1] Ion FLORENCE. From touch displays to the Surface: A brief history of touchscreen technology. www.arstechnica.com [online]. 2013. [cit. 2019-12-05]. Dostupné z: <https://arstechnica.com/gadgets/2013/04/from-touch-displays-to-the-surface-a-brief-history-of-touchscreen-technology/>
- [2] Geoff WALKER. A review of technologies for sensing contact location on the surface of a display. onlinelibrary.wiley.com [online]. 2012 [cit. 2019-06-10]. Dostupné z: <https://onlinelibrary.wiley.com/doi/full/10.1002/jsid.100>
- [3] Pavel ŠKOPEK. Techbox: dotykové displeje – čím se liší rezistivní od kapacitního. www.mobilnet.cz [online] 2013. [cit. 2019-01-06] Dostupné z: <https://mobilnet.cz/clanky/techbox-dotykovye-displeje---cim-se-lisi-rezistivni-od-kapacitniho-11566>
- [4] An JAE-SUNG, Hong SEONG-KWAN a Kwan O KYUN. In-cell Capacitive Touch Panel Structures and Their Readout Circuits. ieeexplore.ieee.org [online]. [cit. 2018-06-22]. Dostupné z: <https://ieeexplore.ieee.org/document/7543685>
- [5] Kim HYOUNG-RAE, Choi YOON-KYUNG, Byun SAN-HO, Choi KVANG-HO a Kim SANG-WOO. A mobile-display-driver IC embedding a capacitive-touch-screen controller system. ieeexplore.ieee.org [online]. [cit. 2018-06-24]. Dostupné z: <https://ieeexplore.ieee.org/document/5434080>
- [6]

Seznam obrázků

- [1] <http://noel.feld.cvut.cz/vyu/a2b31hpm/images/thumb/5/51/Rezistivn%C3%AD.gif/180px-Rezistivn%C3%AD.gif>
- [2] https://ieeexplore.ieee.org/mediastore_new/IEEE/content/media/5428240/5433812/5434080/5434080-fig-1-source-small.gif
- [3] https://wol-prod-cdn.literatumonline.com/cms/attachment/ec3b_da35-bfb3-435f-a030-d49ed76056ec/jsid100-fig-0002-m.png
- [4] <https://www.cypress.com/file/95156/download>

MOBILNÍ APLIKACE PRO MONITOROVÁNÍ PRŮMYSLOVÝCH PROCESŮ (INDUSTRIAL PROCESSES MONITORING APP)

Daniel Šebík¹

¹ Ústav přístrojové a řídicí techniky, fakulta strojní, ČVUT v Praze, daniel.sebik@fs.cvut.cz

Abstrakt: V této práci je řešena problematika rychlé orientace na pracovišti s využitím mobilních SCADA systémů. Pro tento účel je navržena aplikace formátu SCADA, ve které je implementováno rozpoznání QR-kódů přiřazených jednotlivým částem sledovaného procesu. Aplikace je psána pro operační systémy iOS v jazyce Swift. Hlavními využitými frameworky jsou Core Image pro rozpoznání obrazu a Macaw pro implementaci vektorově řešené vizualizace procesu.

Klíčová slova: QR-kód, scanování, Swift, mobilní SCADA, SwiftUI, Core Image, AV Foundation

1. Úvod

Monitorování průmyslových procesů je důležitou součástí potřebnou pro řízení komplexních průmyslových procesů. Tato oblast spadá převážně pod SCADA systémy, které mají za účel centralizovat monitorování a řízení těchto procesů. Za pomoci různých operátorských panelů a zobrazovacích zařízení potom může být monitorování dostupné i přímo na daném pracovišti. V některých případech tomu tak ale není. Ať už z důvodu finančních úspor, prostorových omezení nebo z jiných důvodů. Pro technika pohybujícím se na daném pracovišti potom může být v takových případech poněkud složité získávání informací o daném procesu. Tato myšlenka tak vytváří prostor pro návrh aplikace pro různá přenosná zařízení, která může mít pracovník na daném stanovišti stále k dispozici, a pomocí kterých by mohl jednoduše nascanovat identifikační kód dané části procesu, a tím ji identifikovat a získat potřebné informace. Jednalo by se tak o rozšíření mobilních SCADA systémů právě o rychlou identifikaci. Cílem tohoto projektu řešeného v rámci studentského projektu bylo seznámení se s programovacím jazykem Swift za pomoci tvorby mobilní aplikace řešící popsáný problém.

2. Koncept aplikace

Navržená aplikace je rozdělena do tří částí. V první části je vizualizace samotného procesu s interaktivními prvky, pomocí kterých může uživatel přímo přejít na informační stránku dané části procesu, například sebrat data z daného snímače. Obdobným způsobem již přímo SCADA systémy fungují, např. [4]. Ve druhé části si uživatel může navolit seznam několika preferovaných stanovišť, které může bez hledání ve vizualizaci, či scanování identifikačního kódu rychle otevřít. Ve třetí části je implementována inovativní složka. Každá část procesu, snímač či akční člen budou mít vlastní QR kód představující identifikační číslo daného prvku a uživatel jej bude moci přímo na pracovišti nascanovat, čímž se přímo dostane na stránku s

parametry daného prvku v procesu. Nebude tak muset ve vizualizaci hledat, který prvek je ten správný, čímž dojde ke zrychlení hledání. Navíc se zajistí i snížení chybovosti lidského faktoru, protože uživatel nebude moci ve vizualizaci omylem zvolit jiný prvek systému a pracovat se špatnými informacemi.

3. Tvorba UI

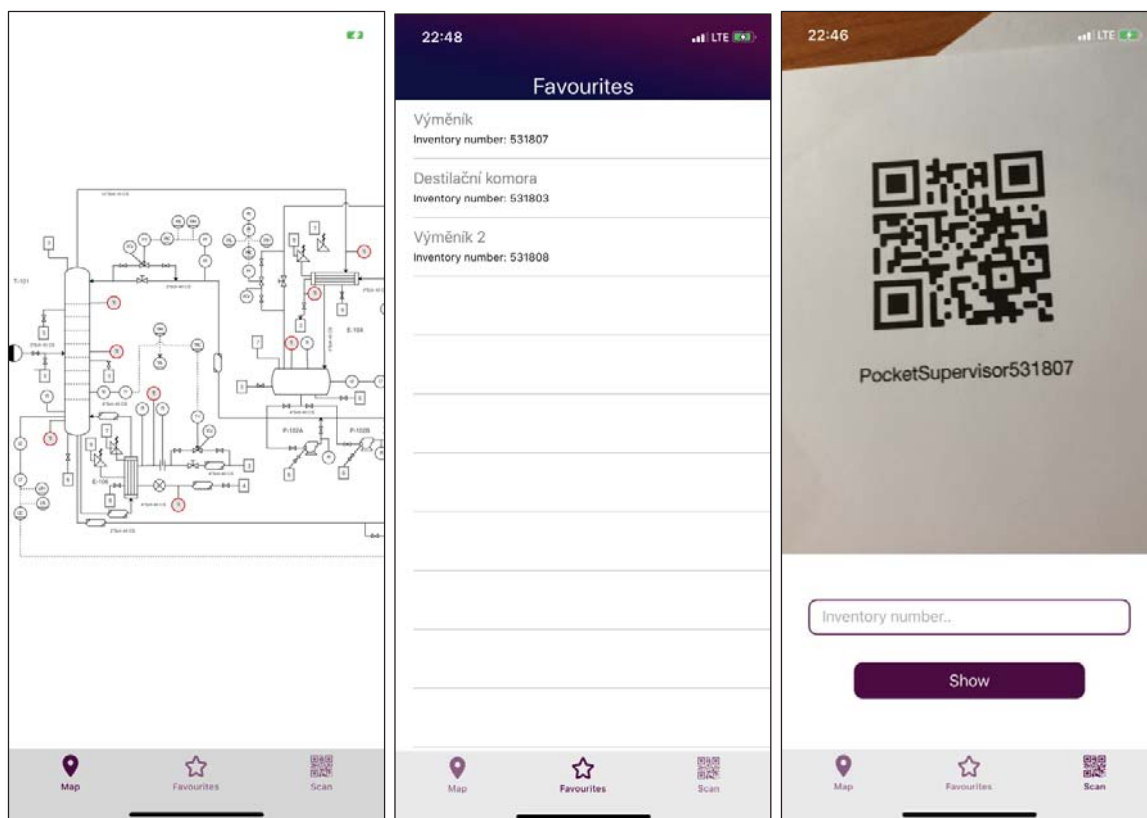
Tvorba uživatelského rozhraní se v nynější době u operačních systémů iOS přesouvá k novému frameworku SwiftUI, nicméně v době tvorby aplikace ještě nebyl tento framework k dispozici a využil jsem frameworku UIKit [2].

Po otevření aplikace se uživateli zobrazí přihlašovací obrazovka. Ta je vytvořena pomocí objektu třídy `UIViewController`. Po zadání správné kombinace přihlašovacího jména a hesla se následně zavře přihlašovací okno a skrze notifikaci objekt třídy `UISceneDelegate`, který spravuje události během životního cyklu aplikace, otevře nové okno. Pro rozdělení aplikace do tří částí je pro toto okno voleno objektu třídy `UITabBarController`, který spravuje tři rozdílné `UITabBarController` jednotlivých částí, mezi nimiž lze přepínat pomocí lišty na spodní straně obrazovky.

První z oken je zmiňovaná vizualizace procesu. Tu jsem řešil pomocí grafického frameworku `Macaw` [1], který mi dovolil využít vektorových formátů grafiky vizualizace ve formátu `SVG`. Zoomování a posouvání ve vizualizaci je řešeno pomocí transformací souřadnic a velikostí v daném vektorovém obrázku. Snímek obrazovky z okna vizualizace je vidět na obr. 1.a.

Druhé z oken je věnováno uloženým prvkům procesu, tak aby se uživatel rychleji dostal k částem procesu, které nejčastěji sleduje. Seznam je řešený pomocí objektu třídy `UITableView`, do kterého se načítají objekty uložených částí procesu podle identifikačních kódů, které jsou uloženy do paměti po přidání na seznam preferovaných objektů. Ukládání do paměti mobilního zařízení je řešeno pomocí frameworku `Core Data`. Okno s uloženými položkami je vidět na obr. 1.b.

V posledním okně `UITabBarController` může uživatel scanovat QR kód přiřazený určité části procesu. Scanování QR kódů je řešeno pomocí frameworků `Core Image` a `AV Foundation`. Okno s nástrojem scanování je vidět na obr. 1.c.



Obr. 1 Okna hlavního UITabViewControllera zleva: a) vizualizace procesu, b) uložené prvky procesu, c) scanování QR kódů.

Posledním oknem (obr. 2) uživatelského rozhraní je okno samotného prvku v procesu, kde se zobrazují potřebné informace, a kde se nachází tlačítka pro přidání do preferovaných položek nebo také jeho zobrazení přímo ve vizualizaci. Data pro testování funkce byla brána z pražských meteostanic [3].



Obr. 2 Stránka s parametry zvoleného prvku procesu.

4. Scanování QR kódů

Pro scanování identifikačních QR kódů jsem využil frameworků Core Image a AV Foundation, jak jsem již zmínil výše. První ze zmiňovaných frameworků je určený ke zpracování a analýze obrazu. V tomto případě jsem využil konkrétně objektu třídy CIDetector, který jsem využil k identifikaci QR kódu z dodaného snímku a jejich přečtení.

Framework AV Foundation potom slouží k přístupu k audiovizuálním prvkům mobilních zařízení, v tomto případě k zadnímu fotoaparátu. Po výběru konkrétního vstupu/výstupu je třeba toto zařízení nakonfigurovat podle požadovaných cílů (počet snímků za sekundu, kvalita snímků a podobně). Po nastavení požadovaného zařízení je možné spustit takzvanou CaptureSession, během které dochází k opakovanému snímání audiovizuálního signálu, v tomto případě tedy ke snímání obrázků. Snímané obrázky se potom za pomoci objektu třídy AVCaptureVideoPreviewLayer zobrazují přímo na obrazovce uživateli, tak aby mohl namířit na identifikovaný kód. Obrázek je po sejmutí předán také objektu třídy CIDetector, který identifikuje, zdali je přítomný nějaký QR-kód a případně přečte zapsaný identifikační kód.

5. Závěr

Během projektu byla v jazyce Swift vytvořena aplikace s prvky vizualizace obdobné SCADA systému pro mobilní zařízení, rozšířená o možnost scanování identifikačního čísla jednotlivých částí průmyslových procesů za pomoci QR kódů pro rychlou orientaci na pracovišti.

Literatura

- [1] P.W.D. Charles, Project Title, (2013), GitHub repository, Dostupné z: <https://github.com/charlespwd/project-title>
- [2] STEELKIWI Inc., “SwiftUI vs UIKit: Benefits and Drawbacks”, 2020, *Medium*.

[3] “Meteostanice”, *Opendata.praha.eu*, Dostupné z: <https://opendata.praha.eu/en/dataset/tsk-meteostanice>

[4] TRIMAX SYSTEMS, “Mobile SCADA”, *Trimaxsystems.com*, Dostupné z: <http://www.trimaxsystems.com/solutions/mobile-scada.html>

OPTICKÉ VLASTNOSTI LANGERHANSOVÝCH OSTRŮVKŮ (OPTICAL PROPERTIES OF ISLETS OF LANGERHANS)

*Petr Zvolánek*¹, *Šárka Němcová*¹, *Daniel Hadraba*^{1,2}, *David Habart*³, *David Vondrášek*²,
*Josef Zicha*¹

¹ *České vysoké učení technické v Praze, Fakulta strojní, Praha, Česká republika, petr.zvolanek@fs.cvut.cz*

² *Akademie věd České republiky, v. v. i., Fyziologický ústav, Praha, Česká republika*

³ *Institut klinické a experimentální medicíny, Praha, Česká republika*

Abstrakt—Transplantace Langerhansových ostrůvků je jedna z metod, která slouží k terapii diabetu mellitu 1. typu. Určení objemu a kvality nebarvených ostrůvků před transplantací je přetrvávající problém. Řada přístupů identifikuje ostrůvky pomocí barvení malého vzorku z materiálu určeného k transplantaci, ale odběr vzoru je provázen významnou chybou. Náš projekt má za cíl umožnit hodnocení celého transplantovaného objemu na základě neinvazivních metod jako světelná reflektance nebo autofluorescence a okamžitou klasifikaci ostrůvků za použití segmentačního přístroje. Tento přístup povede k přesnějšímu stanovení parametrů Langerhansových ostrůvků a následnou větší prediktabilitu výsledků transplantace.

Abstract—Transplantation of the islets of Langerhans is one of the methods used for diabetes mellitus type 1 therapy. Estimation of the volume and quality of unstained islets is a longstanding problem. A number of methods identify islets by staining a small sample from the material for transplantation, but this sampling is accompanied with a significant error. Our project aims to enable the evaluation of the whole volume of the transplanted material based on non-invasive methods such as light reflectance or autofluorescence and the immediate classification of the islets using a sorting device. This approach will lead to a more accurate determination of the parameters of islets of Langerhans and a better predictability of transplantation outcomes.

Klíčová slova—*Langerhansův ostrůvek, zobrazování, velikost, reflektance, autofluorescence*

Keywords—*Islet of Langerhans, imaging, size, reflectance, autofluorescence*

1. Úvod

Langerhansovy ostrůvky jsou shluky buněk o velikosti mezi 0,1 a 0,5 mm nacházející se ve slinivce břišní, které mají mimo jiné na starosti produkci glukagonu a inzulínu – 2 hlavních hormonů regulujících hladinu glukózy v krvi. Glukagon zvyšuje hladinu glukózy a je tvořen v tzv. α -buňkách. Inzulín snižuje hladinu glukózy a je produkován v β -buňkách. Diabetes mellitus 1. typu (DM1T) je autoimunitní onemocnění způsobující úplnou ztrátu β -buněk, čímž znemožňuje přirozenou regulaci glukózy v krvi [1]. Transplantace izolovaných Langerhansových ostrůvků byla vyvíjena od 60. let 20. století. Významně se rozšířila po roce 2000, díky nalezení imunosupresivního protokolu. Je převážně určena nemocným s život ohrožující poruchou rozpoznávání hypoglykemie a tam, kde by transplantace celé slinivky byla příliš riziková [2].

Procedura spočívá v infúzi ostrůvků do vrátnicové žíly vedoucí do jater. Ostrůvky se následně zachytí v jaterních sinusoidách a kontinuálně regulují množství glukózy v krvi [3].

Reprodukovatelnost výsledků transplantace Langerhansových ostrůvků vyžaduje klást větší důraz na jejich kvantitu a kvalitu, které by měly být posuzovány v podobném režimu jako krevní transfúze. Jedná se o bezinfekčnost a životaschopnost ostrůvků a také o jejich funkci, objem jednotlivých ostrůvků a míru příměsí zbytkové tláně slinivky. Ta se od ostrůvků odlišuje například pomocí rychlého barvení dithizonem, který se váže na zinek v molekule inzulínu v β -buňkách ostrůvků a způsobuje jejich zbarvení do červena. Tento proces však

způsobuje znehodnocení vzorku [4]. Velice důležitým parametrem ostrůvku je jeho velikost, protože jeho přežití závisí primárně na zásobování kyslíkem pomocí difuze. Z tohoto důvodu jsou hypoxií více ohroženy větší ostrůvky [5].

Cílem našeho projektu je nalézt takové metody hodnocení viability a čistoty ostrůvků, které nebudou škodlivé pro zkoumané ostrůvky a budou tudíž použitelné pro celý objem materiálu určeného k transplantaci, a odstraní problém variability odběru vzorků. Dalším krokem bude návrh přístroje, který bude schopen vzorky měřit a třídít na základě předem stanovených parametrů a dovolí tím doktorům lépe kontrolovat vlastnosti ostrůvků určených k transplantaci, a tudíž i lépe predikovat výsledky terapie.

2. Parametry Langerhansových ostrůvků

2.1 Velikost, objem a tvar

Jak již bylo řečeno výše, velikost ostrůvku je jedním z jeho nejzásadnějších parametrů při odhadování viability [5]. Pro kvalitní zhodnocení tvarových a rozměrových parametrů se využívá například optická projekční tomografie, při které je vzorek zalit do průhledného materiálu (například agarózový gel), osvětlen a snímán kamerou při synchronizovaném otáčení vzorku kolem podélné osy. Výsledkem je velice přesná 3D rekonstrukce. Nevýhodou je složitá příprava vzorku, destrukce vzorku a časová náročnost metody [6].

Často se tedy přistupuje k odhadu rozměrů ostrůvků pomocí nahrazení jeho tvaru elipsoidem, jehož poloosy jsou odhadovány z 2 na sebe kolmých 2D průřezů [7], nebo pomocí sferoidu kdy se využívá pouze 1 průmět [8]. Současné třídící přístroje fungující na principu průtokové cytometrie často využívají pouze 1D informaci (čas, po který objekt zastíní paprsek světla procházející dráhou objektu) a objem ostrůvku nahrazují koulí [5]. Aproximace složitějšího tvaru ostrůvku sferoidem se ukázala jako relativně nepřesná s velkou systematickou chybou. Tato chyba je větší u ostrůvků větších rozměru, které mají někdy i velice nepravidelný tvar. [6]

2.2 Optické vlastnosti

Výrazná část přístupů, které charakterizují kvalitu Langerhansových ostrůvků, spoléhá na barevné značení sledovaných struktur. Toto barvení probíhá za využití chemických látek, které se naváží do ostrůvků, čímž je další použití ostrůvku znehodnoceno [9].

Pro zobrazení nebo hodnocení ostrůvků však lze využít i velké množství „Label Free“ metod. Autofluorescence využívá fluorescenčních vlastností látek a struktur, které se ve zkoumaném objektu přirozeně vyskytují. Autofluorescence s excitací na 488 nm a detekcí 500–550 nm už byla použita pro zobrazování ostrůvků umístěných do přední komory myšího oka [7]. Autofluorescenci s více fotonovou excitací lze využít ke stanovení tzv. redox poměru – poměru intenzit pro fluorescenci flavinadeninindinukleotidu (FAD) excitovaného světlem okolo 1000 nm a nikotinamidadeninindinukleotidu (NADH) s maximem při excitaci okolo 720nm. Emisní maxima FAD a NADH jsou na 535 a 460 nm [10].

Dále je známo, že intenzita zpětně odraženého signálu při excitaci 633 nm znatelně poklesne, pokud došlo k poškození β -buněk v ostrůvcích [11]. Další metody zahrnují třeba Ramanovu spektroskopii sloužící k detekci inzulinu a glukagonu v ostrůvcích [12] nebo optická koherentní mikroskopie [13].

3. Experiment

Zkoumané ostrůvky byly izolovány z krysí slinivky v laboratoři v IKEMu, následně byly v živném roztoku tkáňové kultury převezeny na oddělení Biomatematické Fyziologického ústavu AV ČR. Po dobu, kdy na nich nebylo prováděno měření byly skladovány v inkubátoru při 37 °C a 5 % CO₂ v prostředí. Měřena byla reflektance pro excitační vlnové délky 470–665 nm s krokem 15 nm, intenzita autofluorescence pro excitační vlnové délky 470–660 nm s krokem 10 nm na konfokálním mikroskopu Leica TCS SP8 (Leica, Německo). Při měření reflektance i autofluorescence byl vždy detekován signál v pásmu vlnových délek o šířce 15 nm. U reflektance bylo toto 15nm

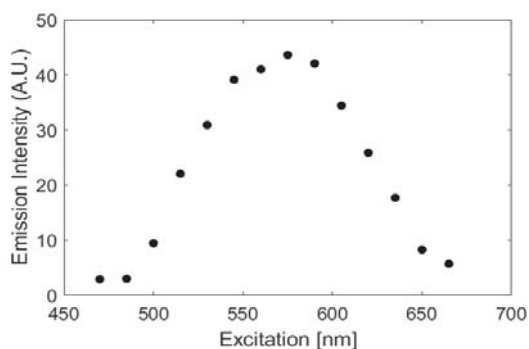
pásmo umístěno symetricky okolo excitační vlnové délky, u autofluorescence šlo o na sebe navazující pásma 500-515 nm až 695-710 nm.

Intenzita emise byla stanovena jako průměrná intenzita všech pixelů ve čtverci 300 x 300 pixelů, který se celý nacházel uvnitř ostrůvku.

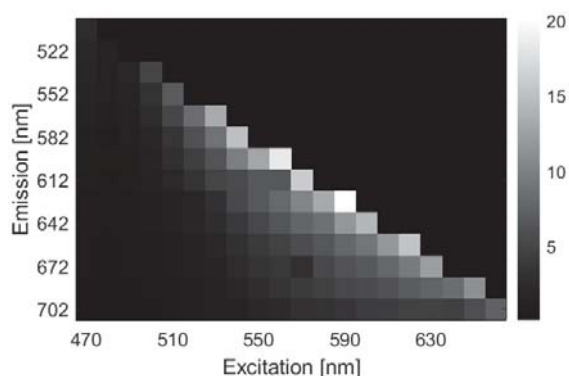
4. Výsledky

Měření zpětně odraženého světla ve viditelném spektru (Obr. 1) mělo maximum mezi vlnovými délkami 545 až 590 včetně (Obr. 2).

Signál při měření autofluorescence (Obr. 3) má 2 výrazná maxima. Jedno pro excitační vlnovou délku 560 nm s emisí 590-605 nm a druhé pro excitaci 590 nm a emisí 620-635 nm. Obecně byl velice slabý.



Obr. 2. Intenzita zpětně odraženého signálu v závislosti na excitační vlnové délce. Intenzita v rozsahu 0-100.



Obr. 3. Intenzita emise autofluorescence v závislosti na excitační a emisní vlnové délce. Intenzita v rozsahu 0-255

5. Diskuze

U měření reflektance byla očekávána větší intenzita signálu pro vlnové délky okolo 633 nm, které se dle literatury používají k hodnocení stavu β -buněk uvnitř ostrůvků [11]. Absence této vlnové délky v naměřeném spektru může znamenat, že měřené ostrůvky již byly poškozené. Poškození mohlo vzniknout během izolace, transportu, nebo následného měření případnou nevhodnou manipulací.

Slabý signál u měření autofluorescence může být zapříčiněn nevhodně zvolenými vlnovými délkami excitace a emise, která mají maxima mírně nad nebo pod viditelným spektrem. Dále třeba úzkými pásmy detekce 15 nm oproti pásmům širokých 50 [7], ale někdy i několik stovek nm [10].

Následně se zaměříme na lepší zreprodukování výsledků již nalezených v literatuře, na jejich využití ke stanovení kvalitativních parametrů ostrůvků a na jejich použitelnost v systému založeném na průtokové cytometrii, který by měl v budoucnosti sloužit k rychlému hodnocení velkého objemu ostrůvků a jejich třídění na základě předem stanovených parametrů.

Poděkování

Autoři článku chtějí poděkovat Studentské grantové soutěži SGS20/055/OHK2/1T/12 a FGU facilitě v rámci MEYS CR (Large RI Project LM2018129 Czech-BioImaging).

Literatura

- [1] HUDÁK, Radovan a David KACHLÍK. *Memorix anatomie*. Praha: Triton, c2013. ISBN978-807-3876-746.
- [2] A. M. J. Shapiro, M. Pokrywczynska, a C. Ricordi, „Clinical pancreatic islet transplantation", *Nat Rev Endocrinol*, roč. 13, č. 5, s. 268–277, kvě. 2017, doi: [10.1038/nrendo.2016.178](https://doi.org/10.1038/nrendo.2016.178).
- [3] K. B. Rodeman a B. Hatipoglu, „Beta-cell therapies for type 1 diabetes: Transplants and bionics", *CCJM*, roč. 85, č. 12, s. 931–937, pro. 2018, doi: [10.3949/ccjm.85a.17088](https://doi.org/10.3949/ccjm.85a.17088).
- [4] T. Yamamoto *et al.*, „Quality control for clinical islet transplantation: organ procurement and preservation, the islet processing facility, isolation, and potency tests", *J Hepatobiliary Pancreat Surg*, roč. 16, č. 2, s. 131–136, bře. 2009, doi: [10.1007/s00534-009-0064-z](https://doi.org/10.1007/s00534-009-0064-z).
- [5] A. Steffen, B. Ludwig, C. Krautz, S. Bornstein, a M. Solimena, „Functional assessment of automatically sorted pancreatic islets using large particle flow cytometry", *Islets*, roč. 3, č. 5, s. 267–270, zář. 2011, doi: [10.4161/isl.3.5.15939](https://doi.org/10.4161/isl.3.5.15939).
- [6] J. Švihlík, J. Kybic, D. Habart, H. Hlushak, J. Dvořák, a B. Radochová, „Langerhans Islet Volume Estimation from 3D Optical Projection Tomography", in *Computer Vision – ACCV 2016 Workshops*, roč. 10117, C.-S. Chen, J. Lu, a K.-K. Ma, Ed. Cham: Springer International Publishing, 2017, s. 583–594.
- [7] E. Ilegems *et al.*, „Light scattering as an intrinsic indicator for pancreatic islet cell mass and secretion", *Sci Rep*, roč. 5, č. 1, s. 10740, zář. 2015, doi: [10.1038/srep10740](https://doi.org/10.1038/srep10740).
- [8] J. Jo, M. Y. Choi, a D.-S. Koh, „Size Distribution of Mouse Langerhans Islets", *Biophysical Journal*, roč. 93, č. 8, s. 2655–2666, říj. 2007, doi: [10.1529/biophysj.107.104125](https://doi.org/10.1529/biophysj.107.104125).
9. A. Pisania. "Development of quantitative methods for quality assessment of islets of Langerhans", Ph.D., Massachusetts Institute of Technology. Dept. of Chemical Engineering., 2007.
- [10] B. Wu, G. Li, M. Hao, a S. Mukherjee, „Non-invasive discrimination between pancreatic islets and exocrine cells using multiphoton microscopy", San Francisco, California, United States, dub. 2015, s. 932935, doi: [10.1117/12.2083555](https://doi.org/10.1117/12.2083555).
- [11] H. Chmelova *et al.*, „Distinct Roles of β -Cell Mass and Function During Type 1 Diabetes Onset and Remission", *Diabetes*, roč. 64, č. 6, s. 2148–2160, čer. 2015, doi: [10.2337/db14-1055](https://doi.org/10.2337/db14-1055).
- [12] J. Hilderink *et al.*, „Label-Free Detection of Insulin and Glucagon within Human Islets of Langerhans Using Raman Spectroscopy", *PLoS ONE*, roč. 8, č. 10, s. e78148, říj. 2013, doi: [10.1371/journal.pone.0078148](https://doi.org/10.1371/journal.pone.0078148).
- [13] M. Villiger *et al.*, „In vivo imaging of murine endocrine islets of Langerhans with extended-focus optical coherence microscopy", *Diabetologia*, roč. 52, č. 8, s. 1599–1607, srp. 2009, doi: [10.1007/s00125-009-1383-y](https://doi.org/10.1007/s00125-009-1383-y).

UKÁZKA ROZPOZNÁVÁNÍ ŘEČI S VYUŽITÍM KLASIFIKACE NEURONOVOU SÍTÍ

Vladimír Hlaváč

ČVUT v Praze, Fakulta strojní, hlavac@fs.cvut.cz

*Abstrakt: Článek demonstruje použití klasifikace neuronovou sítí na data, získaná FFT transformací a pásmovým rozdělením do složek vektoru ze zvukového záznamu několika samohlásek, zaznamenaných zvukovou kartou. Experiment byl proveden za účelem porovnání s metodou SVM pro klasifikaci dat.
Klíčová slova: Rozpoznávání zvuku, FFT, Neuronová síť, Klasifikace dat.*

1. Úvod

Rozpoznávání mluveného slova pomocí analýzy frekvenčního spektra bylo zkoumáno již od sedmdesátých let, dlouho před rozšířením počítačů. Použití výpočetních metod popisuje již například profesor Josef Psutka v roce 1995 [1]. Jednou ze základních je použití rychlé Fourierovy transformace a nerovnoměrné (logaritmické) rozdělení pásem, kdy při sečtení amplitud v těchto pásmech vznikne vektor, který je pro danou hlásku charakteristický. Tímto způsobem lze s minimální chybovostí rozlišit přinejmenším vyslovované samohlásky za podmínky jednoho mluvčího (pro souhlásky a více mluvčích je nutné použít další metody, popsané v [1]).

Tento článek popisuje, jak získat tyto vektory, a demonstruje jejich rozpoznávání neuronovou sítí. Byl napsán jako podklad pro další práce, kdy studenti řeší obdobný problém metodou SVM, za účelem jejich porovnání.

2. Získávání dat pro rozpoznávání

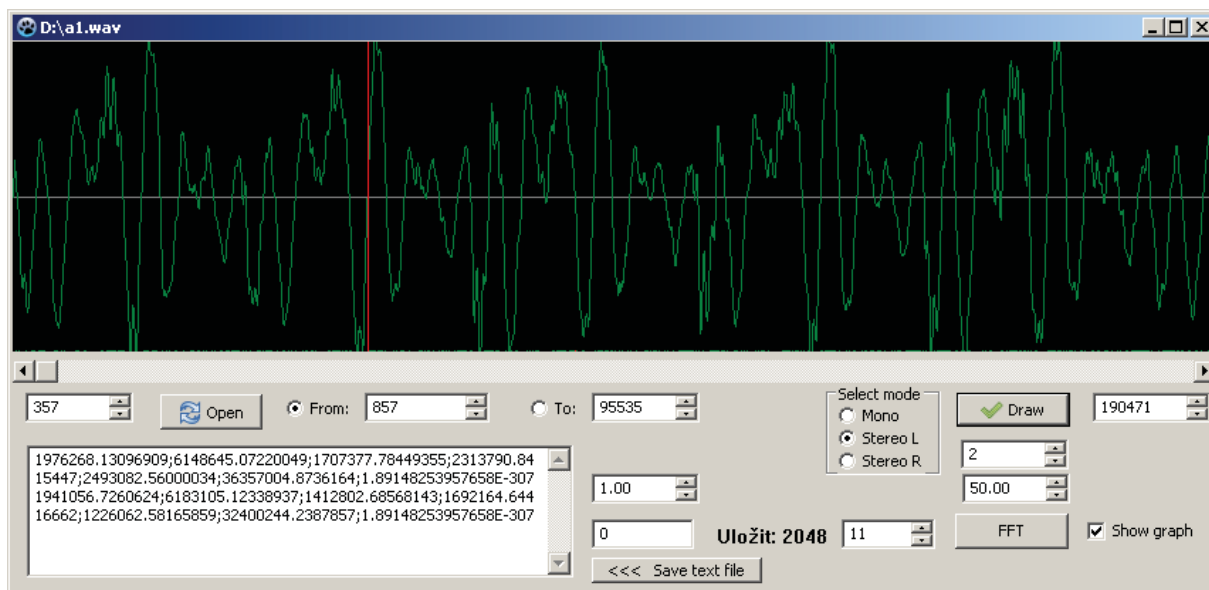
2.1 Data v časové oblasti

Data jsou načtena pomocí zvukové karty a programu Záznam zvuku, který je součástí Windows. Data byla zaznamenána ve formátu wav, který není pakovaný. Pro záznam byly vybrány hlásky, které lze vyslovovat v délce sekund, po jisté subjektivní úvaze samohlásky a, e, i, o, u a souhlásky s. Záznam neobsahuje začátek ani konec vyslovování.

2.2 Převod do frekvenční oblasti a rozdělení do pásem

Funkce FFT [2] by byla nejrychlejší v Matlabu, ale je třeba nejdříve načíst data, a stejně jako v jazyce Python, musela by být použita řada knihoven. Protože se nejedná o prestižní projekt, bylo použito jednodušší řešení a vše bylo napsáno v Pascalu, kde je k dispozici výkonné RAD prostředí Lazarus. Ukázkové řešení FFT pro Delphi lze nalézt na internetu, ale bylo použito z archivu autora.

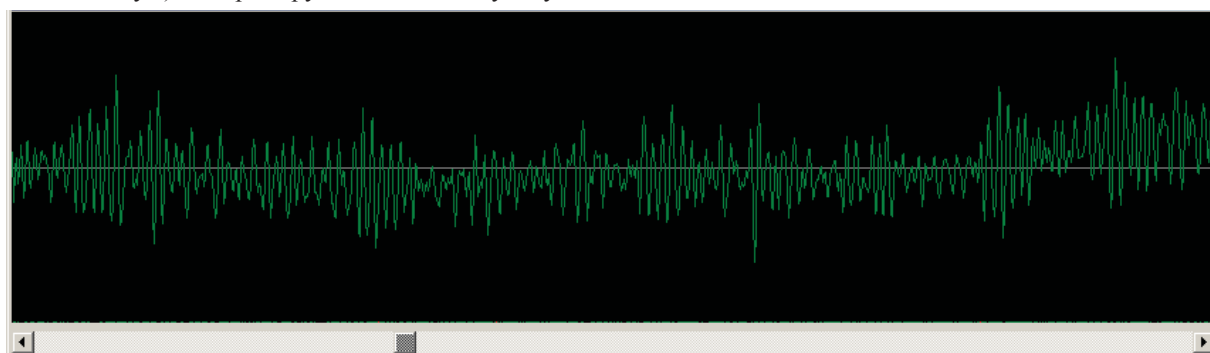
Program umožňuje načíst přímo soubory formátu wav v mono nebo stereo verzi pro 16bitové vzorky. Po načtení je třeba nastavit, zda se jedná o záznam mono nebo stereo a v druhém případě, který kanál číst. Všechny analýzy byly provedeny pro stereo záznam a levý kanál.



Obr. 1 Program pro výpočet FFT, načtena hláska "a". Po kliknutí na graf se doplní orientační červená čára, v daném bodě se zahájí čtení dat pro FFT (počet vzorků je dán nastavením exponentu základu „2“ v okně nalevo od tlačítka FFT, zde 11 – odpovídá 2048, ale používalo se 4096), výsledné frekvence se po intervalech sečtou (viz text) a výsledek se zapíše do okna textového editoru vlevo dole. Odtud je lze kopírovat přes schránku, nebo tlačítkem všechna najednou zaznamenat do souboru. Výměnou zpracovávaných dat (zelený graf) se toto okno nemaže, a lze tak snadno vybírat tutéž hlásku z různých záznamů zvuku.

Po načtení dat je třeba zvolit parametry transformace. Pro dále vyhodnocované vzorky bylo použito okno o šířce 4096 vzorků (počet iterací FFT je třeba přenastavit z 11 na 12). Data na obr. 1 jsou 50× zmenšená a vynesena jako 2 vzorky na pixel.

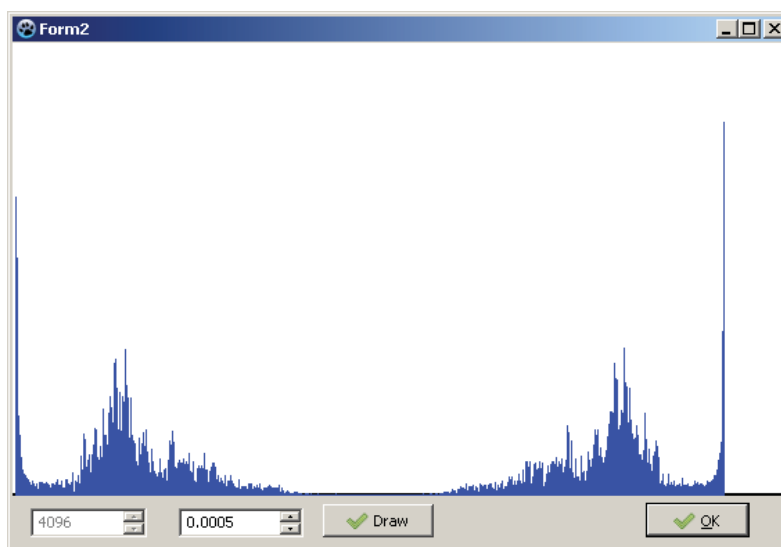
Pro nízké frekvence FFT má velký vliv poloha okna. I když funkce je zjevně periodická, FFT ve skutečnosti vyhodnocuje spektrum z dat, která vzniknou tak, že načtených (zde) 4096 vzorků se periodicky opakuje, což je velmi ovlivněno místem, kde se přesně začnou data přebírat. Tomuto efektu se nechá zamezit aplikací Hammingova okna. Ve zde popisovaném případě byl ale použit manuální postup, kdy vzorky jsou vždy načítány od místa průchodu nulou, bezprostředně předcházející lokálnímu maximu funkce (na obr. 1 červená čára v grafu, označí se kliknutím myši). Oba postupy vedou k obdobným výsledkům.



Obr. 2 Záznam hlásky "s"

Pro samohlásky „a“ a „e“ vychází převaha nízkých frekvencí, pro „o“, „u“ a „i“ jsou pak vyšší. Všeobecně se vyšší frekvence objevují hlavně u souhlásek, ale zde jsou stabilní průběhy jen u „s“, „š“, „z“ a „ž“. Některé souhlásky mají hodně krátký průběh („k“), u jiných se frekvence v průběhu hlásky mění („l“).

Po stisknutí tlačítka FFT je zobrazena funkce (obr. 3). Samotný vektor je do okna výsledků zapsán po kliknutí do grafu záznamu v časové oblasti, aby šlo rychle vybrat dostatečný počet dat.



Obr. 3 Průběh spektra hlásky "s" (levá polovina grafu). Výpočtem rychlé Fourierovy transformace vyjdou komplexní hodnoty, které pro každou dílčí frekvenci reprezentují vektor otočený podle fázového zpoždění dílčí frekvence oproti začátku skenovaného intervalu. FFT z reálných hodnot je vždy osově symetrická, kromě první hodnoty, která má vazbu na nevyváženost stejnosměrné hodnoty signálu (program ji v dalším ignoruje). Zobrazeny jsou ale velikosti dílčích frekvencí (magnitudy), získané jako odmocnina ze součtu druhých mocnin reálné a imaginární hodnoty.

Vektor (jednorozměrná matice) hodnot se ze spektra získá tak, že do první složky se sečte prvních 10 hodnot, do další následujících 20 (frekvence č. 11 až 30), pak dalších 40, atd. Způsob rozdělení na frekvenční pásma má velký vliv na výsledné rozpoznávání.

Program umožňuje rychle zaznamenat velké množství vzorků a následně je uložit do textového souboru. Byl uložen různý počet vzorků, od 63 pro „e“ po 140 pro „i“. Jednotlivé textové soubory byly předzpracovány programem MS Excel. Poslední prvek vektoru, který obsahoval pravou polovinu spektra, byl odstraněn, a ostatní složky byly poděleny takovým číslem, aby součet byl 10000. Takto upravená data se hodí na porovnání různých způsobů rozpoznávání.

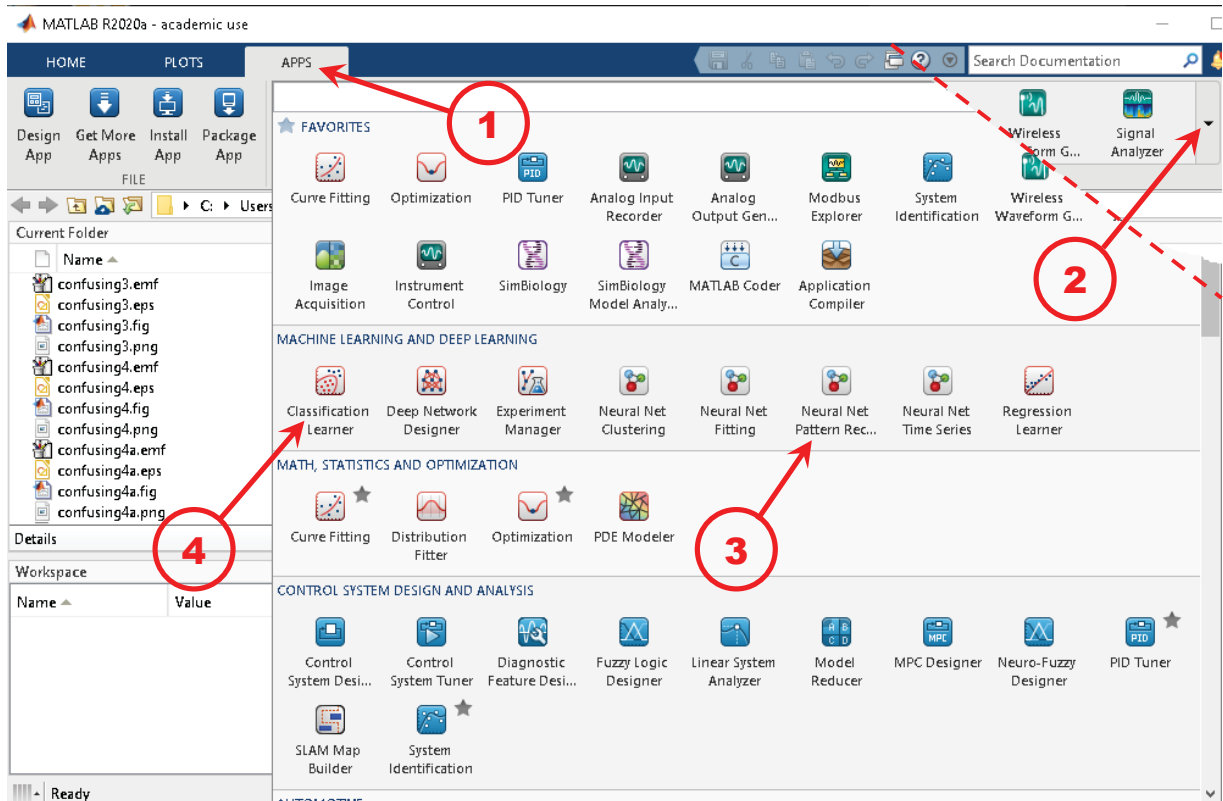
Pro zpracování v Matlabu byly doplněny sloupcečky s příslušností dat (pro daný vzorek se doplní do řádku nuly kromě místa, které odpovídá dané hláске, kde se doplní jednička), všechny tabulky spojeny a zaznamenány do textového souboru. Pro načtení Matlabem řádky nemají záhlaví.

3. Rozpoznávání neuronovou sítí

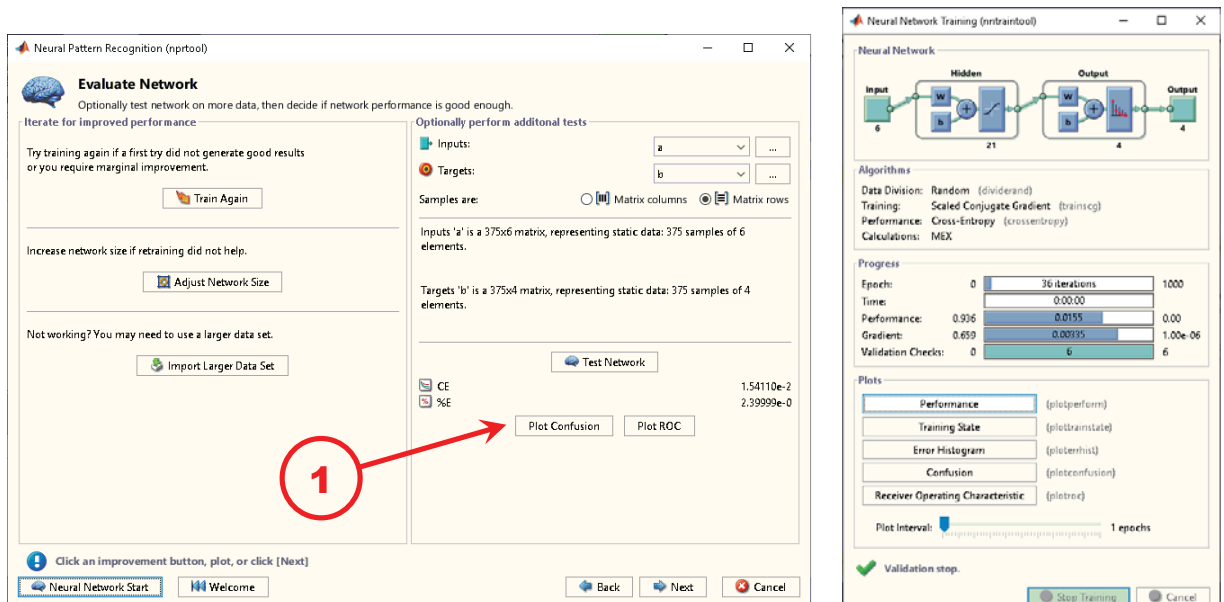
Aplikace pro pokročilé zpracování dat lze v Matlabu nalézt pod záložkou APPS. Seznam aplikací otevřeme a vybereme Neural Net Pattern Recognition, viz obr. 4. Také je možné ji přímo spustit příkazem nprtool.

Aplikace nprtool připomíná vzhledem aplikaci nftool. V několika krocích zvolíme zdrojová data, jejich rozdělení na trénovací, validační a testovací sadu, metodu trénování, počet neuronů ve skryté vrstvě a můžeme zkusit první pokus o trénování. Po natrénování aplikace umožňuje přenastavit parametry nebo změnit data a zkusit trénování opakovat (to lze i beze změn, metoda je stochastická a získáme jiné výsledky). Na závěr je možné výslednou síť vyexportovat, například jako funkci v Matlabu, a tu pak používat pro klasifikaci dalších dat.

Data (popis vyhodnocovaných dat viz předchozí kapitola) byla tvořena šesti sloupci součtu magnitud frekvencí v jednotlivých pásmech, ke kterým bylo doplněno šest sloupců správné klasifikace. Nejprve byla vyhodnocována data pouze pro čtyři písmena, „a-e-i-s“, kdy byly samozřejmě nadbytečné sloupce vynechány. Tato zkrácená tabulka měla 375 vzorků. Pro trénování byly rozděleny mezi trénovací, validační a testovací sadu v poměru 80% – 10% – 10%, tedy 299 – 38 – 38 (default je 70% – 15% – 15%). Bylo nastaveno 21 neuronů ve skryté vrstvě. Matlab provedl 26 iterací. Výsledky trénování jsou na obr. 6, obr. 7 a obr. 8.



Obr. 4 Vyhledání aplikace pro klasifikaci pomocí Neuronové sítě. Kliknutím na záložku APPS (1) se přepneme na kartu aplikací. Vpravo (2) otevřeme nabídku (čárkovaná čára odděluje vzhled okne před kliknutím, vpravo nahoře, a po rozevření nabídky, většina obrázků). Většina nabídky není zobrazena (posuvník vpravo). Kroužek (3) označuje použitý modul pro klasifikaci neuronovou sítí. Kroužek (4) pak aplikaci pro metodu podpůrných vektorů, SVM, příkaz Matlabu classificationLearner.



Obr. 5 Předposlední okno aplikace nprtool. Stisknutím označeného tlačítka Plot Confusion získáme tabulky, které jsou na obr. 6. (vpravo okno z průběhu trénování sítě; byla použita metoda Scaled Conjugate Gradient).



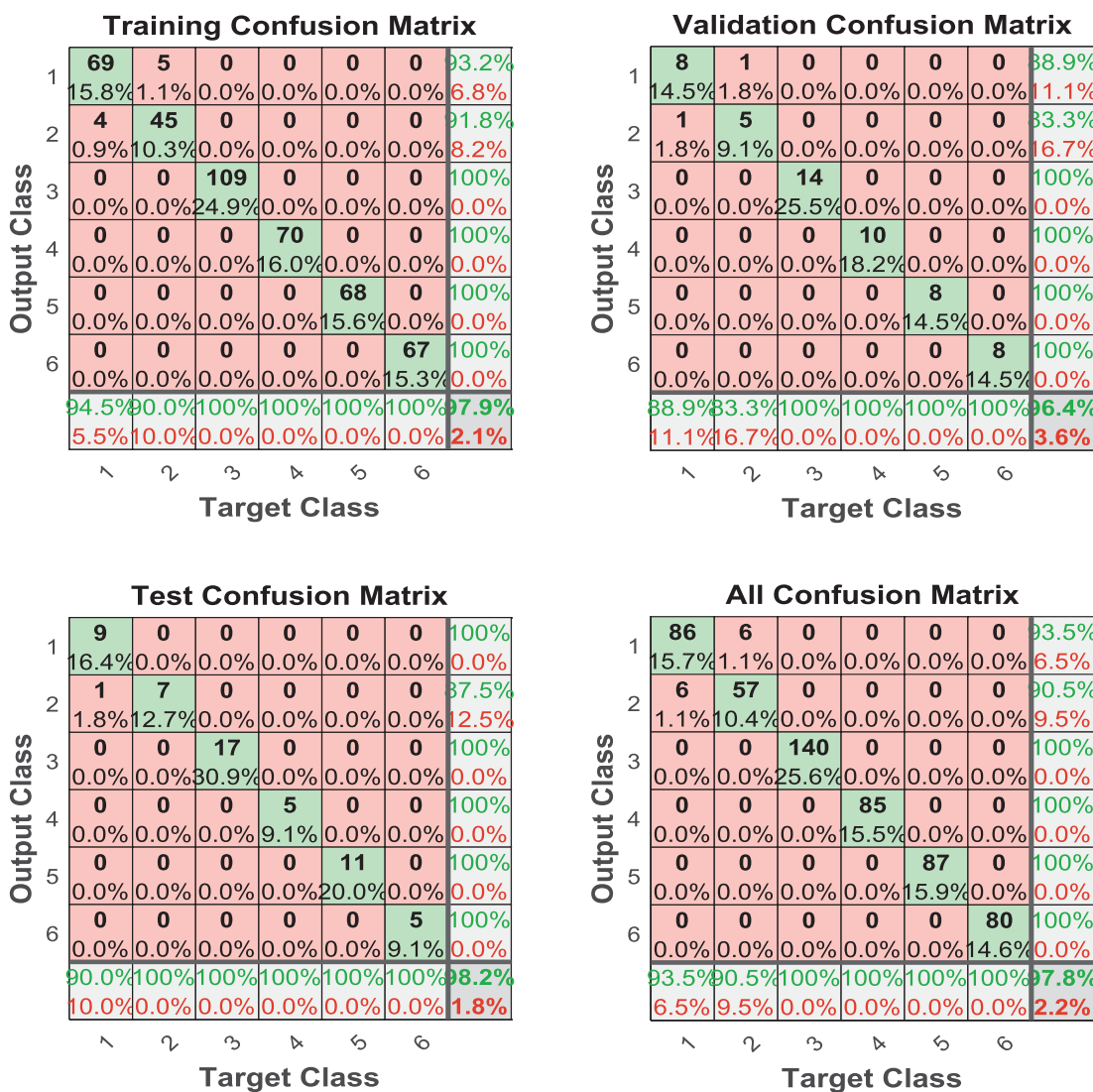
Obr. 6 Výsledky trénování. Svisle výsledek klasifikace, vodorovně správná hodnota. Dochází jen k záměnám hlásek „a“ a „e“, což může být poměrně hrubým rozdělením začátku frekvencí.



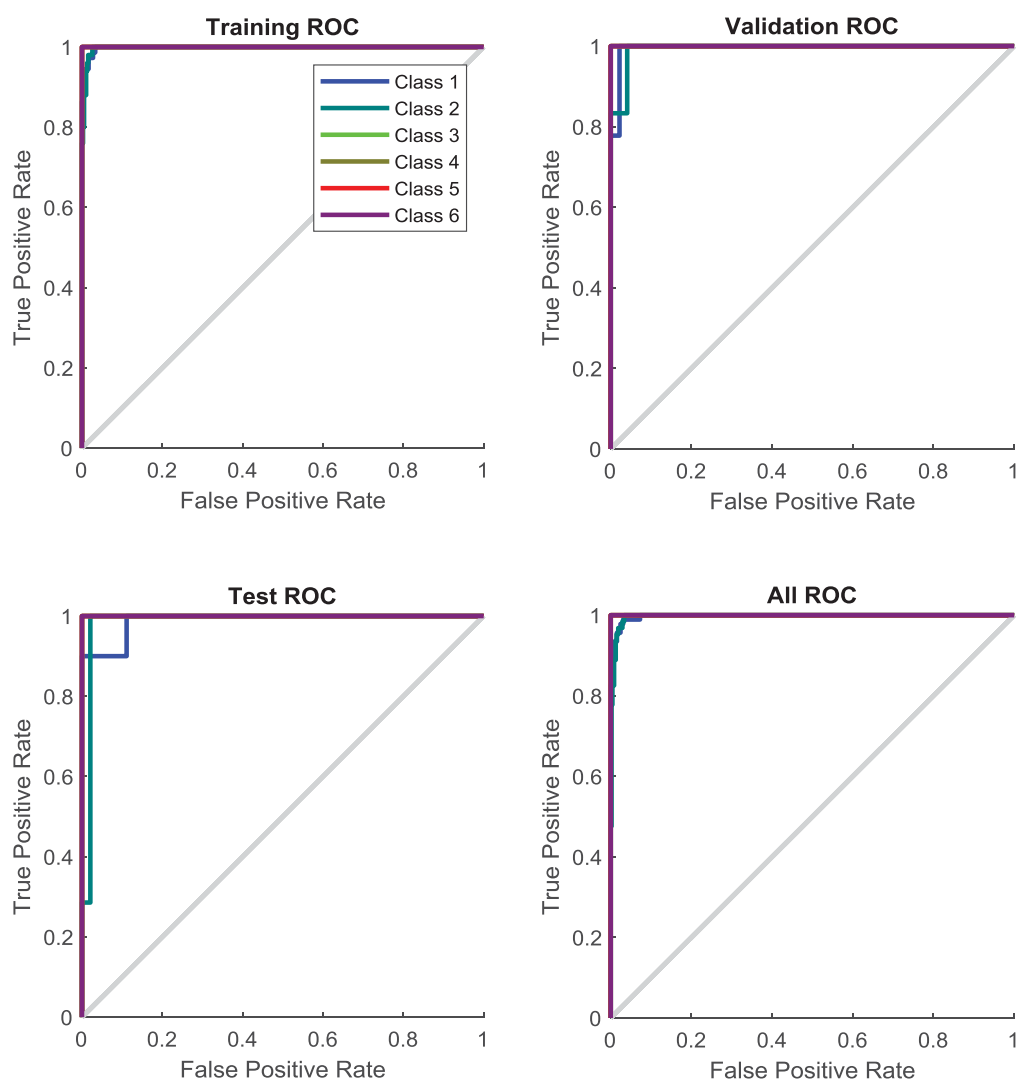
Obr. 7 Po sloučení hlásek "a" a "e" do jedné skupiny máme naprosto bezchybné rozpoznávání.



Obr. 8 Původní data, původní nastavení Matlabu (263 trénovacích, 56 validačních i testovacích, 10 neuronů ve skryté vrstvě, Matlab provedl 20 iterací). Výsledek je prakticky stejný, resp. rozdíl může být dán stochastickým charakterem metody.



Obr. 9 Úplná sada dat (1=a, 2=e, 3=i, 4=o, 5=u, 6=s). Výsledky jsou obdobné. Celý soubor má 576 řádek, ty se dělily na 437 trénovacích, 55 validačních a stejný počet testovacích. Skrytá vrstva byla nastavena na 21 neuronů. Proběhlo 46 iterací. I v tomto případě nejsou rozpoznány hlásky „a“ a „e“, zatímco ostatní jsou bezchybně rozlišeny.



Obr. 10 Plot recognition, vnesení přiřazení čarově. Tatáž data jako na obr. 9. Tento druh zobrazení má smysl, pokud je příliš mnoho tříd objektů na rozpoznávání.

4. Závěr

Byla připravena data pro testování modulu nprtool, která by se měla využít i pro metodu podpůrných vektorů (SVM), v Matlabu funkce fitcecoc (error-correcting output codes, ECOC). Data se ukázala po drobné úpravě (sloučení hlásek „a“ a „e“) plně rozpoznatelná neuronovou sítí, měla by tedy být rozpoznatelná i pomocí SVM, pravděpodobně i ve zjednodušené lineární formě.

5. Reference

- [1] Pšutka, Josef: Komunikace s počítačem mluvenou řečí. Academia, Praha, 1995. ISBN 80-200-0203-0
- [2] Erickson, Jeff: Algorithms. Amazon, 2019. ISBN 1792644833

NEURONOVÉ SÍTĚ V DETEKCI PORUCH (NEURAL NETWORKS IN FAULT DETECTION)

Michal Kuchař¹

¹ Ústav přístrojové a řídicí techniky, Fakulta strojní ČVUT v Praze. Michal.Kuchar@fs.cvut.cz

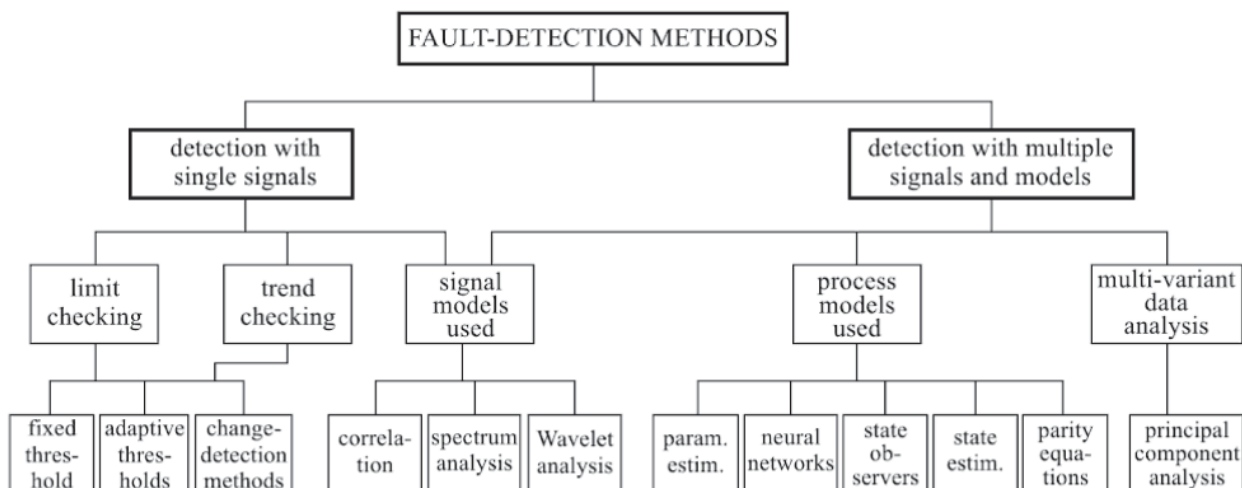
Abstrakt: Příspěvek pojednává o základním rozdělení metod detekce poruch a následně o možných aplikacích neuronových sítí v této oblasti. Ukazuje metody učení s učitelem (vícevrstvá neuronová síť) i bez učitele (samo-organizační mapa).

Klíčová slova: Detekce poruch, umělá inteligence, neuronové sítě, analýza dat, detekce novosti

1. Úvod

Diagnostika poruch je stále se rozvíjející obor, který je aplikován čím dál více i na systémy, které nejsou kritické z hlediska bezpečnosti (jako například jaderné elektrárny, letadla apod.) [1]. V dnešní době internetu věcí a rostoucího výpočetního výkonu je umožněn sběr velkého množství dat z různých zařízení a aplikací, proto je čím dál snazší aplikovat statistické metody a metody umělé inteligence.

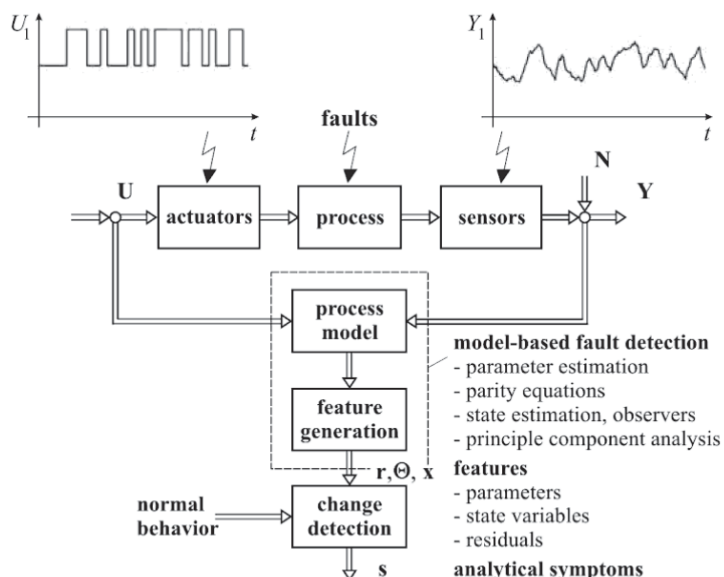
Pro přiblížení je na obr. 1 znázorněno základní rozdělení detekce poruch. Jednou skupinou je detekce poruch založena na porovnání modelu s měřením (obr. 2), druhou skupinou jsou metody detekce poruch na základě analýzy signálu bez znalosti modelu (obr. 3).



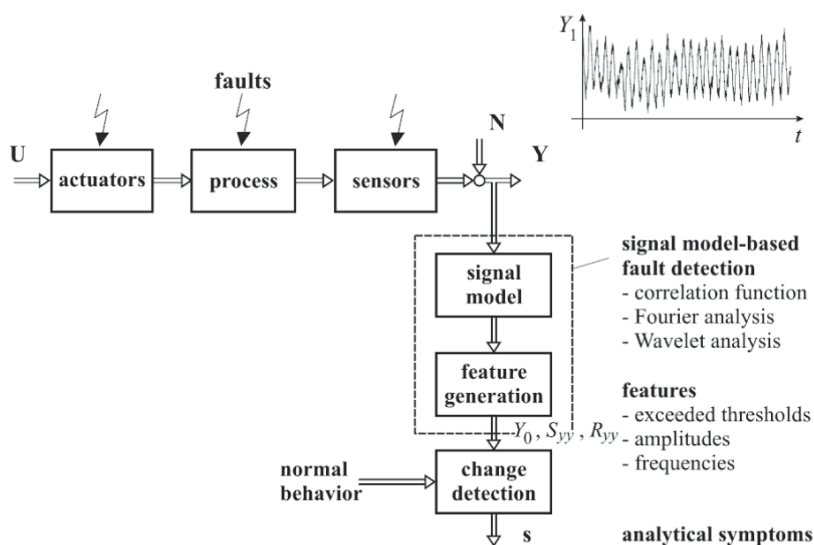
Obr. 1 Základní rozdělení metod detekce poruch [2].

Jak již z názvu plyne, tak pro detekce poruch na základě porovnání modelu s měřením je potřeba mít model, což je výhodné, pokud je model vůbec dostupný a pokud dostatečně popisuje realitu. V praxi se objevuje velké

množství systémů, pro které je velice obtížné odvodit uspokojivý model. Již v tomto bodě nastává prostor pro aplikace některých metod umělé inteligence. Analytický model může být nahrazen například neuronovým modelem, fuzzy modelem nebo jejich kombinací [3].



Obr. 2 Detekce poruch na základě porovnání modelu s měřením [2].



Obr. 3 Detekce poruch na základě analýzy signálu bez znalosti modelu [2].

Co se týče metod analýzy signálu, je možné aplikovat metody umělé inteligence na samotný signál, nebo na jeho vlastnosti. Výčet vybraných vlastností signálu je uveden v tabulce 1. Sledováním těchto vlastností lze nastavit prahy, které určují, zdali se v signálu neobjevují neobvyklé události. Prahly lze nastavit analyticky, pokud jsou známy limitní hodnoty pozorovaného systému, nebo je lze nastavit na základě statistické analýzy, nebo je možné použít například neuronovou síť.

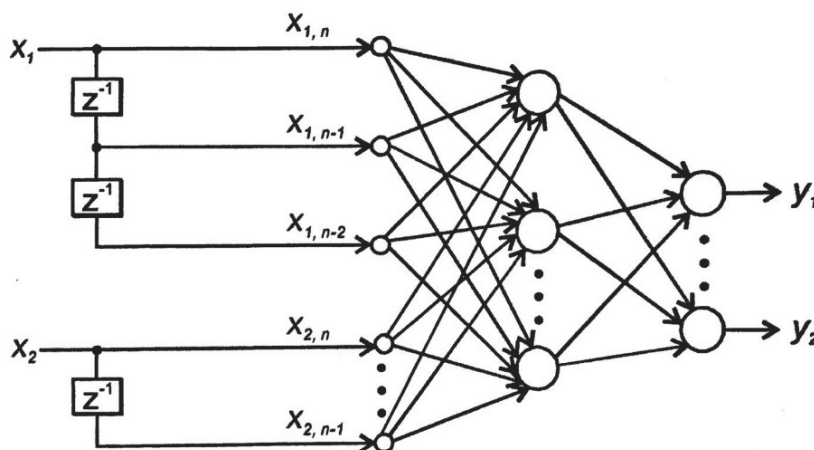
V případě multidimenzionální analýzy dat je vhodné nějakým způsobem redukovat dimenze pro snazší vizualizaci. Mezi takové metody se řadí metoda hlavních komponent nebo Kohonenova samo-organizační mapa (dále SOM). SOM se řadí mezi neuronové sítě, které je možné učit bez učitele, což je výhoda, jelikož není zapotřebí mít k dispozici anotovaná data.

Tab. 1. Vybrané vlastnosti signálu [3].

Feature name	Feature definition	Value for the harmonic signal
Mean value	$\bar{x} = \frac{1}{T} \int_t^{t+T} x(u) du$	$\bar{x} = 0$
Absolute mean value	$x_{AVE} = \frac{1}{T} \int_t^{t+T} x(u) du$	$x_{AVE} = \frac{2}{\pi} X \cong 0,637X$
Root mean square value	$x_{RMS} = \sqrt{\frac{1}{T} \int_t^{t+T} x^2(u) du}$	$x_{RMS} = \frac{X}{\sqrt{2}} \cong 0,707X$
Absolute peak value	$x_{PEAK} = \max_{t < u \leq t+T} x(u) $	$x_{PEAK} = X$
Positive peak value	$x_{PEAK+} = \max_{t < u \leq t+T} x(u)$	$x_{PEAK+} = X$
Negative peak value	$x_{PEAK-} = \min_{t < u \leq t+T} x(u)$	$x_{PEAK-} = -X$
Peak to peak value	$x_{P-P} = x_{PEAK+} - x_{PEAK-}$	$x_{P-P} = 2X$
Form factor	$K = \frac{x_{RMS}}{x_{AVE}}$	$K = \frac{\pi}{2\sqrt{2}} \cong 1,111$
Crest factor	$C = \frac{x_{PEAK}}{x_{RMS}}$	$C = \sqrt{2} \cong 1,414$

2. Neuronové sítě

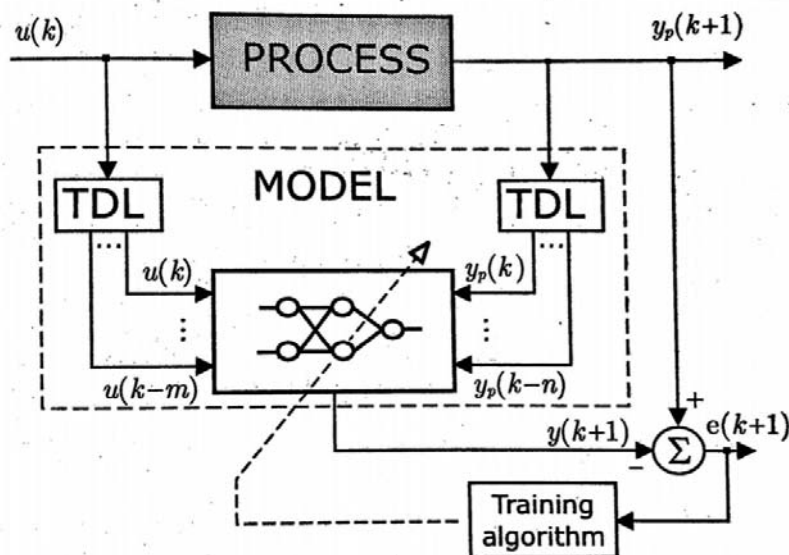
Neuronové sítě se dají použít pro tvorbu neuronového modelu, ale i pro samotnou klasifikaci a detekci poruch. Na obr. 4 je znázorněna vícevrstvá neuronová síť jejíž vstupy obsahují i určitou historii vstupních veličin. Tato síť může reprezentovat model systému. Celková reprezentace systému je znázorněna na obr. 5.



Obr. 4 Neuronový model dynamického systému [3].

Je nutno podotknout, že architektury neuronových sítí je velice mnoho, jejichž rozebrání je na celou knihu. To samé se dá říci i u učících algoritmů. Výchět nejčastějších architektur pro učení s učitelem je následovný:

- Vícevrstvá neuronová síť (MLP)
- Konvoluční neuronová síť (CNN)
- Síť s kruhovými funkcemi (RBF)
- Rekurentní neuronová síť (RNN)

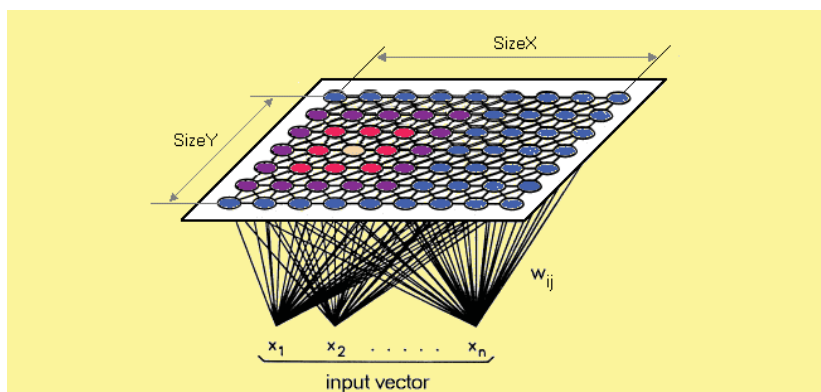


Obr. 5 Schéma použití neuronového modelu [3].

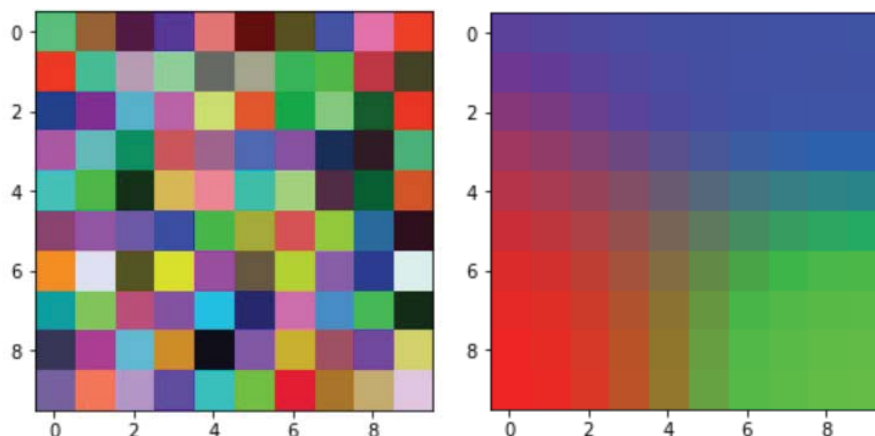
2.1 Samo-organizační mapy

Jak bylo řečeno, SOM se řadí mezi neuronové sítě, které se učí bez učitele. Podobně jako u metody hlavních komponent je tedy cílem promítnutí vícerozměrných dat obvykle do roviny. Body v SOM, které leží u sebe, leží u sebe i v původním prostoru dat (což se obecně u metody hlavních komponent nemusí vždy povést). Další výhodou je i adaptivita tj. pokud chceme mapu obohatit o další pozorování, je doučení mapy mnohem méně náročné, než například přepočítat celou metodu hlavních komponent (to je důležité především u opravdu nadměrně velkém počtu dat). [4]

Základní architektura dvourozměrné SOM je znázorněna na obr. 6. Jednotlivé váhy neuronů se snaží přibližovat (dle zvolené metriky – např. Euklidovská) vstupním datům, čímž se tvoří shluky pozorování, které k sobě mají v prostoru blízko.



Obr. 6 Architektura SOM [5].



Obr. 7 Ukázka SOM. Obrázek vlevo ukazuje náhodné rozřazení jednotlivých vah reprezentovaných barvami (RGB). Obrázek vpravo ukazuje váhy po určité době učení, z čehož je názorné samo-organizování.

Vlastnosti SOM se dají použít jak při průzkumové analýze dat, tak při detekci i samotné diagnostice poruch. SOM tvoří během učení shluky a pokud se některé nové pozorování vyskytne mimo ně, je možné z tohoto pozorování vyvozovat poruchy. Vzhledem k tvorbě shluků je možné i jednotlivé poruchové shluky klasifikovat a vyvozovat z nich diagnózu.

2.2 Adaptivní filtry

Další z přístupů detekce poruch je použití adaptivních filtrů. Klasickým reprezentantem adaptivního filtru je LMS (least mean squares) filtr jehož přepis je vyjádřen dle rovnice (1). Jednotlivé váhy jsou vypočteny tak, aby střední hodnota kvadratických odchylek byla minimální. K tomu lze použít například metodu největšího spádu, z čehož se odvodí rovnice (2).

$$\tilde{y}(k) = w_1(k) \cdot x_1(k) + \dots + w_n(k) \cdot x_n(k) = \mathbf{w}^T(k) \cdot \mathbf{x}(k) \quad (1)$$

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \Delta \mathbf{w} = \mathbf{w}(k) + \eta(k) \mathbf{w}(k) e(k) \quad (2)$$

Jedním ze způsobů využití adaptivního filtru k detekci poruch jsou metody error and learning based novelty detection (ELBND) a learning entropy (LE) [6], které vyvozují novost na základě adaptace vah. Neobvyklá adaptace vah nad určité meze může být indikátorem poruchy v pozorovaném systému.

3. Závěr

Byly shrnuty základní metody detekce poruch a možné prostory pro využití neuronových sítí. Bylo uvedeno základní schéma neuronového modelu. Dále byla představena role samo-organizační mapy jako nástroje pro detekci poruch i klasifikaci symptomů. Posledním nástrojem byly uvedené metody detekce novosti (ELBND a LE) jako možné nástroje detekce poruch.

Poděkování

Práce byla podpořena grantem SGS18/177/OHK2/3T/12

Literatura

- [1] R. Patton, P. Frank a R. Clark, Issues of Fault Diagnosis for Dynamic Systems, London: Springer, 2000.

- [2] R. Isermann, *Fault-Diagnosis Systems, An introduction from Fault Detection to Fault Tolerance*, Berlin: Springer, 2006.
- [3] J. Korbicz, J. Kościelny, Z. Kowalczyk and W. Cholewa, *Fault Diagnosis, Models, Artificial Intelligence, Applications*, Berlin: Springer, 2004.
- [4] P. Hebák, *Statistické myšlení a nástroje analýzy dat*, Praha: Informatorium, 2013.
- [5] A. Ralhan, „Self Organizing Maps,“ 2018. [Online]. Available: <https://medium.com/@abhinavr8/self-organizing-maps-ff5853a118d4>. [Přístup získán 30 Srpen 2020].
- [6] M. Cejnek a I. Bukovský, „Concept drift robust adaptive novelty detection for data streams,“ *Neurocomputing*, pp. 46-53, 2018.

ROBOTIKA V REHABILITACI, OPTIMALIZACE POHONNÝCH JEDNOTEK

Karel Vošahlík¹

¹ Ústav přístrojové a řídicí techniky, Fakulta strojní, ČVUT v Praze, karel.vosahlik@fs.cvut.cz

Abstrakt: Robotizace je neustále rozšiřována napříč všemi obory. Jedním z oborů, ve kterém se robotika využívá, je zdravotnictví. V tomto oboru je robotika využívána například při zpracování vzorků a provádění rozborů, chirurgii či rehabilitaci. V tomto článku budou uvedeny příklady využití robotiky ve zdravotnictví, zejména v oblasti rehabilitace. Dále budou uvedena některá kritéria pro optimalizaci pohonných jednotek při navrhování robotického zařízení.

Klíčová slova: rehabilitace, robotizace, konstrukce, optimalizace

1. Úvod

Lidé jsou v průběhu života vystavováni různému onemocnění či úrazům. Z neurologického či ortopedického onemocnění lze jmenovat například cévní mozkovou příhodu, dětskou mozkovou obrnu, roztroušenou sklerózu, paraplegii, centrální i periferní parézy, ataxii, skoliózu, svalové a kloubní poruchy. Rehabilitační a nemocniční zařízení navštěvují pacienti také ambulantně, například při léčbě pooperačních stavů či chronických potíží.

Některé způsoby léčby jsou nenáročné, jiné způsoby jsou velmi fyzicky náročné jak pro fyzioterapeuty, tak pro samotné pacienty. Jeden z méně náročných způsobů léčby je léčebná tělesná výchova. Při tomto způsobu léčby cvičí pacienti cviky, které jim fyzioterapeut ukáže, nebo cvičí společně s fyzioterapeutem. Dalším méně náročným způsobem léčby je například elektroterapie. Mezi náročnější způsoby léčby je nácvik chůze pacienta. Při této léčbě jsou potřeba až čtyři fyzioterapeuti na jednoho pacienta. Dva fyzioterapeuti vedou pacientovi dolní končetiny, jeden fyzioterapeut zajišťuje stabilitu pacienta a čtvrtý fyzioterapeut říká pacientovi pokyny.

Vzhledem k náročnosti léčby a nedostatku fyzioterapeutů je postupně zaváděna robotika i do oboru rehabilitace. Jednoduché léčebné metody, jako je léčebná tělesná výchova, vykonává pacient zcela sám v rehabilitačním zařízení či doma. Jiné jednodušší léčebné metody mohou být vykonávány přímo na lůžku bez přítomnosti fyzioterapeutů za použití robotického zařízení. Náročnější léčebné metody jsou vykonávány v nemocniční rehabilitačním oddělení či rehabilitačním zařízení.

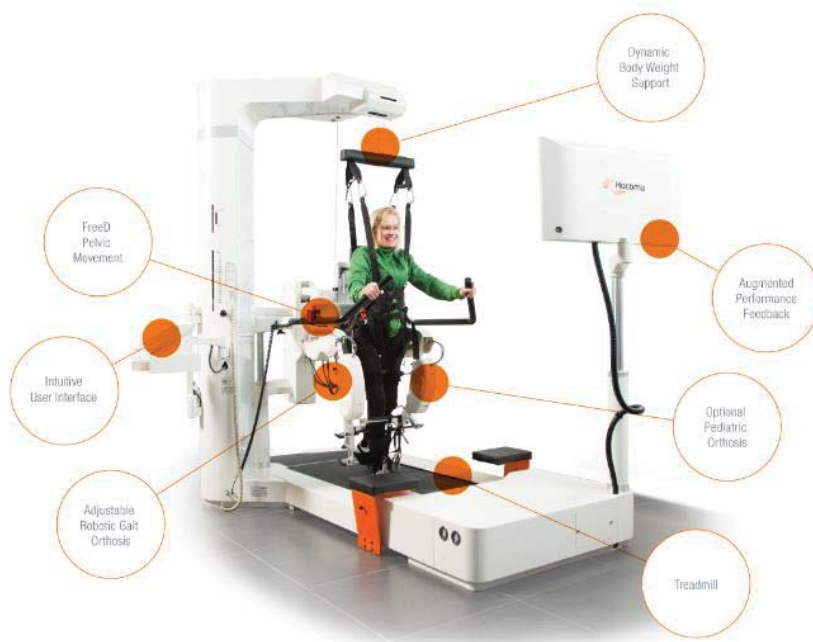
2. Robotika v rehabilitaci

Robotika je v rehabilitaci stále více používána při léčebných metodách, u kterých je to možné. Konstrukčně jednodušší robotická zařízení jsou motodlahy. Motodlahy jsou používány pro léčbu v oblasti ramene, loktu, zápěstí, prstů, kyčelního kloubu, kolene a hlezenního kloubu. Na obr. 1 je zobrazena motodlaha BTL-CPMotion K Elite od firmy BTL [1] pro dolní končetiny, která slouží pro rehabilitaci kyčelního, kolenního a hlezenního kloubu. Tato motodlaha má 13 automatických přednastavených protokolů, 50 uživatelských protokolů, unikátní click systém pro jednodušší nastavení, dálkové ovládání s dotykovým displejem.



Obr. 1 Motodlaha BTL-CPMotion K Elite od firmy BTL. Tato profesionální motodlaha je určena pro všechny tři klouby dolní končetiny

Dalším příkladem robotického rehabilitačního zařízení je například Lokomat od firmy Hocoma [2], který je zobrazen na obr. 2. Lokomat je simulátor chůze, ve kterém je pacient dynamicky zavěšen pomocí nosného postroje. Dolní končetiny jsou uchyceny do robotických ortéz. Pacient chodí po pohyblivém běžeckém pásu pomocí robotických ortéz. Rychlost běžeckého pásu je nastavitelná v rozmezí 0-10 km/h. Jako zpětná vazba pro pacienta je 32" palcový displej, na kterém pacient často vidí jdoucí postavu. Pro obsluhu simulátoru je určen 15" palcový displej.



Obr. 2 LokomatPro od firmy Hocoma. LokomatPro je robotický simulátor chůze.

Posledním uvedeným příkladem v tomto příspěvku je nemocniční robotické lůžko Anymov od firmy BTS BioEngineering [3] zobrazené na obr. 3. Toto lůžko je vybaveno 13 přesnými pohonnými jednotkami. Zdvihací sloupky umožňují výškové nastavení ložné plochy a její laterální náklony. Ložná plocha je v oblasti dolních končetin rozdělena, čímž lze provádět abdukci a addukci dolních končetin. Pomocí robotických ortéz lze provádět simulaci chůze. Ložnou plochu lze vertikalizovat, čímž lze postupně zatěžovat pacienta při simulaci chůze. Celé lůžko lze ovládat pomocí displeje. K lůžku je možné dokoupit další volitelné příslušenství.



Obr. 3 Robotické lůžko Anymov od firmy BTS Bioengineering.

Nanorobotika si našla své uplatnění i v oboru zdravotnictví. Nanoroboti jsou robotická zařízení, jejichž rozměry jsou v nanometrech. Cíle využití nanorobotiky je například v onkologii-chirurgii, kdy nanoroboti budou cíleně ničit postiženou tkáň, další možností je dodávání léčiv k cílové oblasti, odstraňování nevhodných částic z krevního řečiště. Nanorobotika se neustále vyvíjí a její uplatnění se bude neustále rozšiřovat

3. Pohonné jednotky a jejich optimalizace

Robotická zařízení jsou poháněna pohonnými jednotkami. Obecně se v robotice používají pneumatické pohonné jednotky, hydraulické pohonné jednotky či elektrické pohony. Pro zdravotnická zařízení jsou však jistá omezení. K důležitým omezením patří zejména hygienické požadavky a bezpečnostní požadavky. Příkladem může být vyhláška č. 306/2012 Sb. ministerstva zdravotnictví České republiky o podmínkách předcházení vzniku a šíření infekčních onemocnění a o hygienických požadavcích na provoz zdravotnických zařízení a ústavů sociální péče. Pneumatické pohonné jednotky potřebují pro svou funkci stlačený vzduch. Stlačený vzduch je poměrně nebezpečné medium, neboť je velmi stlačitelné a v případě poškození zařízení dochází k velké expanzi stlačeného media. Při své činnosti dochází k vypouštění expandovaného media, což do vnitřních prostor není z hygienických důvodů možné, neboť stlačený vzduch je znečištěný od oleje. Jedná se tak o rizikové faktory pro zdravotnická zařízení.

Hydraulické pohonné jednotky vyžadují kapalinu o potřebném tlaku. Hydraulická kapalina (olej) je však téměř nestlačitelná a při poškození pohonné jednotky či rozvodů kapaliny nedochází k expanzi jako u stlačeného vzduchu. Unik hydraulické kapaliny není taktéž z hygienického důvodu přípustný. Rozvody stlačeného vzduchu či hydraulické kapaliny nejsou v nemocničních či rehabilitačních zařízeních standardně zhotovené. Výjimkou jsou technická oddělení.

Z uvedených důvodů se nejčastěji používají elektrické pohonné jednotky. Podle požadovaného pohybu se používají jednotky s rotačním pohybem nebo s posuvným pohybem. Nejčastěji používané pohonné jednotky s rotačním pohybem jsou stejnosměrné elektromotory s napětím 12 V, 24 V či 36 V. Je-li potřeba větší výkon či krouticí moment, jsou tyto elektromotory doplněny o převodovku. Pro lineární pohyb se převážně používají lineární aktuátory. Lineární aktuátor je tvořen elektromotorem s převodovkou a pohybovým šroubem s maticí. Lineární aktuátor od firmy Timotion [4] je na obr. 4. Použití lineárních aktuátorů je však konstrukčně náročnější, neboť mají větší zástavbové rozměry, omezený krok-vysunutí a náročnější způsob uchycení.

Před volbou pohonné jednotky je nutné vyřešit zatížení celého mechanismu pomocí mechaniky statiky, kinematiky a dynamiky. Zatížení lze určit výpočtově, graficky či v simulačních programech. Po předběžném zjištění zatížení pohonných jednotek, lze vybrat konkrétní pohonnou jednotku. Dle této jednotky se upraví rozměry konstrukce mechanismu. A následně se provede znovu výpočet zatížení. Pro optimalizaci výběru pohonných jednotek lze použít například metodu vícekritériálního rozhodování. Tato metoda je založena na volbě vhodných kritérií. Poté je oslovena skupina expertů, kteří ohodnotí jednotlivá kritéria bodově či stanový jejich pořadí dle důležitosti. Kritéria mohou být například zástavbové rozměry pohonné jednotky, hmotnost, cena, servis a dostupnost, výkonové a silové parametry.



Obr. 4 Lineární aktuátor od firmy Timotion.

Zástavbové rozměry značně ovlivňují rozměry navrhovaného zařízení. Hmotnost pohonných jednotek ovlivňuje hmotnost celého zařízení, ale také zatížení ostatních pohonných jednotek, které jsou v navrhovaném mechanismu této jednotce předřazené. Výkonové a silové parametry jsou při výběru typu pohonné jednotky důležité. Existuje řada výrobců pohonných jednotek. Jednotky se však vzájemně liší nejen výkonovými a silovými parametry, ale i tvarově, rozměry či typem vykonávaného pohybu. Důležitým parametrem je i cena, neboť je-li na navrhovaném zdravotnickém zařízení použito více pohonných jednotek, je také cena celého zařízení vyšší.

4. Diskuze

Uvedené typy pohonných jednotek jsou jen výběrem z mnoha, které jsou standardně k dispozici. Pokud standardní pohonné jednotky nevyhovují požadavkům, je nutné vyrobit pohonné jednotky dle požadavků. Při návrhu robotického zařízení je však vhodné posoudit, zda optimalizovat návrh konstrukce za účelem volby konkrétního typu pohonné jednotky, nebo zda optimalizovat vybraný typ pohonné jednotky. Pro optimalizaci výběru typu pohonných jednotek lze použít i jiné optimalizační metody. Výsledek zmíněné metody vícekriteriálního rozhodování je závislý na názoru expertů, kteří jednotlivá kritéria hodnotí.

5. Závěr

V tomto článku jsou uvedeny některé příklady robotických zařízení, která jsou využívána v léčebné rehabilitaci. Jedná se pouze o vybrané příklady využití robotiky v rehabilitaci. Robotika má samozřejmě širší využití ve zdravotnictví a stále se obor robotiky vyvíjí. Robotická zařízení usnadňují práci zdravotnickému personálu a zároveň urychluje léčbu pacientů. Nejnovější odvětvím robotiky je nanorobotika. Co se týče výsledků zmíněné vícekriteriální metody rozhodování výběru typů pohonných jednotek, tak výsledky se budou pravděpodobně lišit dle hodnotících expertů. Někteří experti mohou více upřednostňovat konstrukční parametry, jiní experti mohou upřednostňovat parametry ekonomického charakteru.

Poděkování

Práce byla podpořena grantem SGS20/055/OHK2/1T/12

Literatura

- [1] BTL. [online]. [cit. 15.9.2020]. Dostupné z: <https://www.btl.cz/produkty-motodlaha-cpmotion-btl-cpmotion-elite>
- [2] Hocoma. [online]. [cit. 15.9.2020]. Dostupné z: <https://www.hocoma.com/solutions/lokomat-2/?variation=LokomatPro#why>
- [3] BTS Bioengineering. [online]. [cit. 15.9.2020]. Dostupné z: <https://www.btsbioengineering.com/products/anymov/>
- [4] Timotion. [online]. [cit. 15.9.2020]. Dostupné z: <https://www.timotion.com/en/product/detail/linear-actuators/ta23-series?upcls=1481269298&guid=1494581213>

DOKTORANDSKÁ SEKCE

Aktivní magnetické ložisko

Bc. Pavel Blažek

České vysoké učení technické

Fakulta strojní, Ústav přístrojové a řídicí techniky

Praha, Česká republika

Pavel.Blazek@fs.cvut.cz

Abstrakt – V článku je záznam pokračování vývoje školního demonstrátoru aktivního magnetického ložiska. V úvodních kapitolách je popsána funkce takového zařízení a seznámení s postupem práce na školním prototypu. Další částí textu jsou věnovány popisu konstrukčních úprav celkové sestavy, vytvoření zjednodušeného matematického modelu odpovídající nové sestavě a návrh řízení. Součástí jsou i dosažené výsledky simulací v podobě grafů a fotografie konstrukčních úprav.

Klíčová slova – aktivní magnetické ložisko, systém, stabilizace, stavové řízení, PID regulace

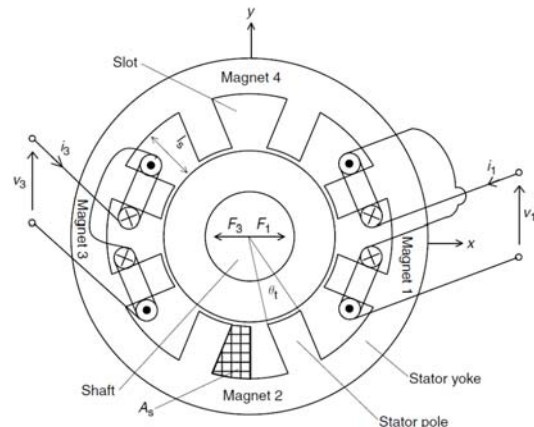
I. ÚVOD

Na tomto tématu pracuji v rámci studentské grantové soutěže s názvem *Vysokorychlostní elektrické stroje pro automobilové aplikace – řízení, výkonová elektronika*. Mým úkolem je pokračovat na vývoji demonstrátoru aktivního magnetického ložiska, na kterém by se dalo ukázat, jak takový systém funguje, a který by mohl do budoucna sloužit pro testování různých řídicích algoritmů.

II. AKTIVNÍ MAGNETICKÉ LOŽISKO

Aktivní magnetické ložisko je takový typ ložiska, které umožňuje uložení rotujícího hřídele bez přítomnosti pasivních odporů. Rychlost rotace hřídele je tedy teoreticky neomezená, pokud vynecháme vliv odstředivé síly na hmotu hřídele. Možnost dosahovat vysokých otáček bez použití maziva přináší řadu výhod, čímž tato technologie nachází uplatnění v řadě aplikací, například ve sterilních nebo potravinářských procesech nebo právě v aplikacích, kde potřebujeme realizovat vysoké otáčky.

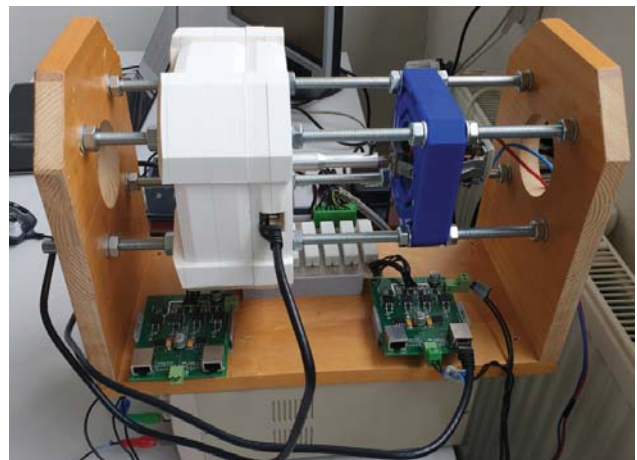
Magnetické ložisko se skládá ze statoru, které obsahuje elektromagnety. Magnetickými silami je rotující hřídel centrován do osy magnetického ložiska, čímž je zachována konstantní vzduchová mezera mezi hřídelem a ložiskem. Slovo *aktivní* znamená, že magnetické síly jsou regulovány řídicím systémem. Na obr. 1. je průřez radiálním magnetickým ložiskem. Stator je rozdělen do osmi částí. Každé dvě části patří k jednomu elektromagnetu. V ose x a y jsou tedy vždy dva elektromagnety působící proti sobě. Senzory měří aktuální polohu hřídele v obou osách a řídicí systém reguluje proud v elektromagnetech tak, aby byl hřídel neustále v ose ložiska. Při vychýlení ze své referenční polohy dojde k akčnímu zásahu a jeho zpětnému návratu do osy ložiska. [1]



Obr. 1. Řez ložiskem [1]

III. KONSTRUKCE ŠKOLNÍHO PROTOTYPU

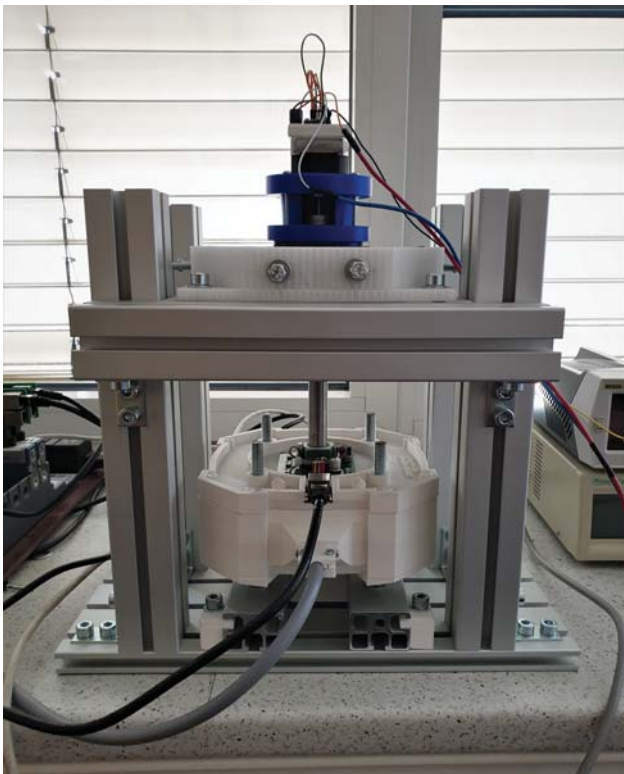
V počátku mého působení na projektu, byl školní prototyp radiálního aktivního magnetického ložiska ve stavu, kdy byla dokončena celková konstrukce ložiska, příslušná elektronika i funkční program v LabVIEW FPGA. Regulace polohy hřídele však fungovala pouze pokud se hřídel netočil. Po roztočení došlo k destabilizaci systému. Na obr. 2. je k vidění původní konstrukce.



Obr. 2. Původní model - ložisko (bíle), kardanův kloub (modře) [2].

V dřevěném rámu jsou na společných závitových tyčích nasazeny ložisko i kardanův kloub, v němž je uložen DC motor. Kloub umožňuje naklápění motoru s hřídelem. Hřídel je nasezen na motoru bez použití spojky a veškeré navržené komponenty, tedy hlavní pouzdro ložiska a kardanův kloub, jsou vytištěny na 3D tiskárně z plastu, což je jednou z příčin vzniku konstrukčních nepřesností.

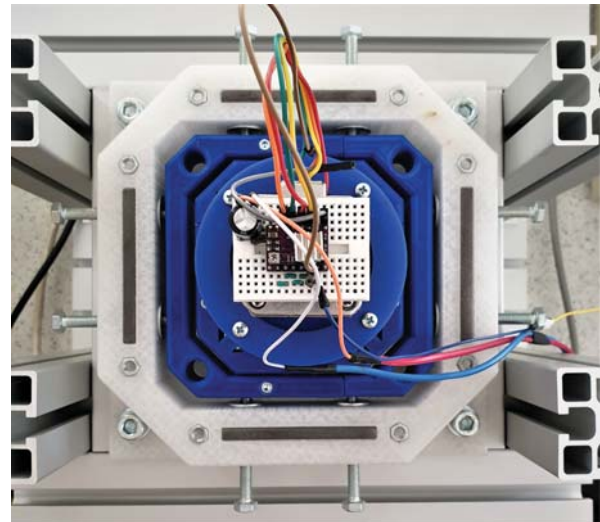
IV. KONSTRUKČNÍ ÚPRAVY



Obr. 3. Nová sestava

Abychom byli schopni řídit aplikaci, ve které pracujeme se vzduchovou mezerou mezi ložiskem a hřídelem 1.5 mm, bylo zapotřebí zajistit určitou mechanickou přesnost konstrukce. Navrhl jsem tedy novou konstrukci. Celou sestavu jsem uspořádal do vertikální polohy, přičemž hřídel směřuje směrem dolů, podobně jako u kyvadla. Takové uspořádání umožňuje získat rovnoběžnost hřídele s osou ložiska. Samotné magnetické ložisko a kardanův kloub jsem zachoval. Dřevěný rám jsem nahradil konstrukcí z hliníkových profilů, které výrazně přispěly ke zvýšení tuhosti konstrukce a přidal jsem centrovací hlavu, pomocí níž lze jemně posouvat kardanův kloub s motorem vůči magnetickému ložisku tak, aby bylo možné srovnat jejich osy. V původní konstrukci byla osa motoru pevně svázána s osou ložiska pomocí společných závitových tyčí. Mírné vyosení tak mohlo způsobovat další problémy s řízením. Hlava obsahuje osm šroubů, které slouží pro jemný posuv kardanu. Další úpravou bylo vyrobení nové delší hřídele z hlazené ocele. Na jednom konci je nalisováno kuličkové ložisko, které je uloženo uvnitř plastového domku viz obr. 5.. Hřídel pohání krokový

motor, který nahradil původní DC motor. Ten nám umožnil i velice pomalé otáčení hřídele s možností snadného řízení otáček. Motor je s hřídelem spojen zubovou spojkou, která je uvnitř plastového domku.



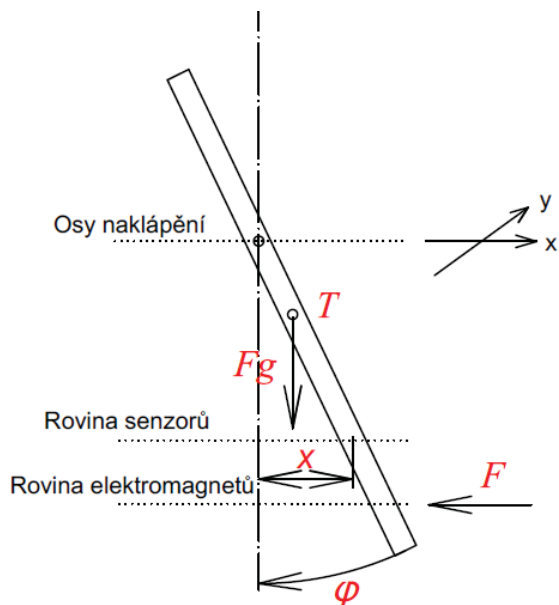
Obr. 4. Centrovací hlava



Obr. 5. Detail na uložení hřídele

V. MATEMATICKÝ MODEL

Aktivní magnetické ložisko je nestabilní systém. Na obr. 6. je schéma nového systému, který je velice zjednodušený. V rovině, ve které leží osy naklápění, je již zmíněný kardanův kloub. Těžiště se podle výpočtů nachází 21 mm od této roviny. Rovina senzorů je pak vzdálena 140 mm a elektromagnety 180 mm od osy naklápění. Hmotnost v těžišti je 800 g. Odvození síly F zde neuvádím z důvodu náročnosti. Čerpal jsem z odborné literatury [1] a při výpočtu přenosu už pracuji s výslednými vztahy. Nový matematický model jsem realizoval v prostředí MATLAB a



Obr. 6. Schéma - nový model

Simulink. Samotný systém jehož vstupem je proud je druhého řádu. Model je vytvořen pouze pro jednu osu, protože druhá osa je teoreticky stejná.

Síla elektromagnetů F působící na hřídel v jedné ose má tvar

$$F = k_x x + k_i i, \quad (1)$$

kde x je výchylka v ose x a i je proud v cívce. Konstanty k_i a k_x pak mají tvar

$$k_i = 2L_o \left(\frac{I_b}{g} \right) \cos \left(\frac{\pi}{8} \right), k_x = 2L_o \left(\frac{I_b}{g} \right)^2 \cos \left(\frac{\pi}{8} \right), \quad (2)$$

kde g je tíhové zrychlení, I_b je stálá složka proudu a indukčnost L_o je

$$L_o = \frac{N^2 \mu S}{2g}, \quad (3)$$

kde N je počet závitů cívky, μ je permeabilita vakua a S je průřez kotvy [1].

Přenos systému se vyjádří jako

$$G(s) = \frac{X(s)}{I(s)} = \frac{l_s l_f k_i}{I_x s^2 - (l_s l_f k_x + F_g l_t)}, \quad (4)$$

kde I_x je moment setrvačnosti závěsu, který jsem spočítal v programu SolidWorks, l_s je vzdálenost senzorů od osy naklápění, l_f je vzdálenost elektromagnetů od osy naklápění a F_g je tíhová síla. Výchylky hřídele jsou tak malé, že jsem si mohl dovolit využít zjednodušení $\varphi \approx x/l_s$.

V reálné aplikaci však není řízen proud v cívkách ale napětí. Jedná se o RL obvod, kde musíme brát v úvahu přechodové jevy ze vztahu

$$I(s) = \frac{U(s)}{sL + R}, \quad (5)$$

kde R je odpor vodiče a L je indukčnost cívky [2]. Tím vneseme do systému další dynamiku a přenos výsledného systému je

$$G(s) = \frac{X(s)}{U(s)} = \frac{l_s l_f k_i}{I_x L s^3 + I_x R s^2 - L(l_s l_f k_x + F_g l_t) s - R(l_s l_f k_x + F_g l_t)}. \quad (6)$$

VI. STABILIZACE SYSTÉMU

Stabilizace takového systému je možné dosáhnout použitím PD regulátoru. Po vyjádření charakteristického polynomu uzavřeného regulačního obvodu

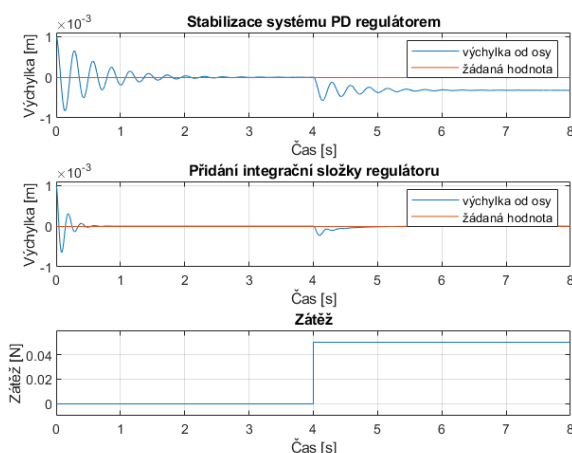
$$Ls^3 + Rs^2 + (-ML + Nr_d)s - MR + Nr_p = 0, \quad (7)$$

kde

$$M = \frac{l_s l_f k_x + F_g l_t}{I_x}, N = \frac{l_s l_f k_i}{I_x} \quad (8)$$

je patrné, že lze pomocí parametrů regulátoru ovlivnit některé členy charakteristického polynomu. Aby platila Stodolova podmínka stability musí být splněno, že všechny koeficienty charakteristického polynomu existují a mají stejné znaménko. Musí platit, že

$$r_p \geq \frac{(l_s l_f k_x + F_g l_t)R}{l_s l_f k_i}, r_d \geq \frac{(l_s l_f k_x + F_g l_t)L}{l_s l_f k_i}$$



Obr. 7. Stabilizace systému.

Výpočet jsem ověřil v prostředí MATLAB a Simulink. V grafu na obr. 7. je zobrazeno chování systému s nastavenou nenulovou počáteční podmínkou (s vychýlením 1 mm od

osy). V čase 4 sekundy začne působit konstantní zatížení 0,05 N kolmo na hřídel. Je vidět, že systém je stabilní, ale při silovém působení na hřídel dojde ke vzniku trvalé regulační odchylky. Ta je odstraněna přidáním integrační složky regulátoru. Regulační pochod je poté výrazně rychlejší a nezanechává trvalou regulační odchylku.

VII. STAVOVÉ ŘÍZENÍ

Další možností jak řídit takový systém by mohla být stavová regulace. Ta umožňuje pracovat se zpětnou vazbou od všech stavových veličin. V mém případě pracuji se třemi stavy - polohou (aktuální výchylkou x), rychlostí \dot{x} a proudem i . Matice stavového popisu mají tvar

$$A = \begin{pmatrix} 0 & 1 & 0 \\ M & 0 & N \\ 0 & 0 & -\frac{R}{L} \end{pmatrix}, B = \begin{pmatrix} 0 \\ 0 \\ \frac{1}{L} \end{pmatrix}, C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (9)$$

Abych mohl použít stavový regulátor, musím měřit všechny stavové veličiny. V mém případě mohu měřit aktuální výchylku a velikost proudu, což je patrné z matice C. Rychlost by se měřila obtížně, proto jsem použil pozorovatele stavu, který slouží k odhadování stavu. Dynamika odhadu je popsána vztahem

$$\dot{\hat{x}}(t) = \mathbf{A}\hat{x}(t) + \mathbf{B}u(t) + \mathbf{L}(y(t) - \hat{y}(t)), \quad (10)$$

kde veličiny se stříškou jsou odhadované, y je vektor výstupu (poloha a proud) a vektor L modifikuje dynamické vlastnosti, se kterými se budou eliminovat odchylky odhadu.

Pro řízení jsem navrhl rozšířený stavový regulátor pro možnost systém řídit a ne ho pouze stabilizovat. Zavedl jsem nový stav pro vnesení astatismu do soustavy

$$\dot{I}(t) = x(t) - w(t), \quad (11)$$

kde w je žádaná hodnota a x je výchylka hřídele. Rozšířený stavový regulátor má tvar

$$u(t) = -\mathbf{K}_p \hat{x}(t) - k_i I(t) \quad (12)$$

a celková zpětná vazba je $\mathbf{K} = (\mathbf{K}_p \quad k_i)$. Rozšířený stavový popis se vyjádří jako

$$\begin{pmatrix} \dot{\hat{x}}(t) \\ \dot{I}(t) \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{C} & 0 \end{pmatrix} \begin{pmatrix} \hat{x}(t) \\ I(t) \end{pmatrix} + \begin{pmatrix} \mathbf{B} \\ 0 \end{pmatrix} u(t) + \begin{pmatrix} \mathbf{0} \\ -1 \end{pmatrix} w(t). \quad (13)$$

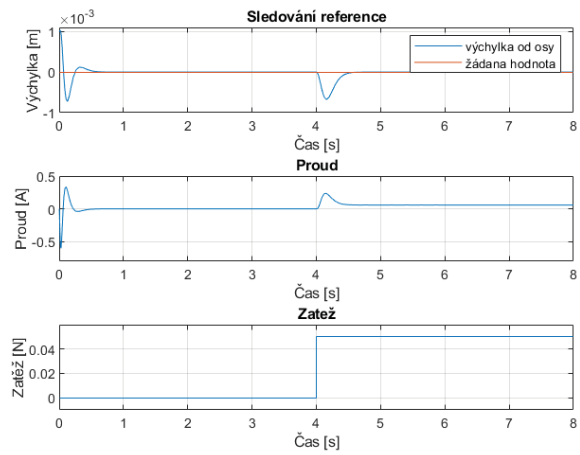
Pozorovatele stavu a stavový regulátor navrhuji odděleně. K výpočtu parametrů stavového regulátoru a pozorovatele se využívá v programu MATLAB příkaz *acker* popřípadě *place*, který však neumožňuje definovat vícenásobný kořen [3]. Příkaz se zapisuje jako $\mathbf{K} = \text{acker}(\mathbf{A}_{\text{hat}}, \mathbf{B}_{\text{hat}}, p)$, kde se matice nové dynamiky zapíšou jako $\mathbf{A}_{\text{hat}} = [\mathbf{A} \text{ zeros}(3,1); -[1 \ 0 \ 0] \ 0]$; $\mathbf{B}_{\text{hat}} = [\mathbf{B}; 0]$; a

pro pozorovatele stavu se příkaz zapíše jako $\mathbf{L} = (\text{place}(\mathbf{A}', \mathbf{C}', p_{\text{poz}}))'$.

Pro můj systém jsem zvolil čtyřnásobný stabilní pól $p = -20$ (1 1 1 1). Póly pozorovatele stavu se volí 2 - 6 x pomalejší [3]. Zvolil jsem $P_{\text{poz}} = -60$ (1 1.2 1.3 1). Z výpočtů vychází

$$\mathbf{K}_p = (54.7 \quad 2.2 \quad -2.6 \quad -118), k_i = -118,04 \quad (14)$$

$$\mathbf{L} = \begin{pmatrix} 150 & 0 \\ 6140 & 2.5 \\ 0 & -1400 \end{pmatrix}. \quad (15)$$



Obr. 8. Stavová regulace.

Na obr. 8. je znázorněn průběh výchylky od osy a proud. Navržené stavové řízení má lepší odezvu než v případě použití PID regulátoru.

VIII. ZÁVĚR

Rotující hřídel se podařilo stabilizovat ve vertikální poloze s použitím dvou PD regulátorů, ale s jinými parametry než vychází z výpočtů. Po přidání integrační složky se systém destabilizuje, a protože model neodpovídá realitě je obtížné správně nastavit parametry regulátoru. S PD regulátorem je takový systém schopen rotovat v ose s mírnou deviací, způsobenou nevyvážením hřídele. Po vychýlení se hřídel vrací zpět do osy, kompenzovat silové působení bez trvalé regulační odchylky však není možné bez integrační složky. Dalším postupem je tedy pokus o zpřesnění matematického modelu a zdokonalení implementace řízení v prostředí LabVIEW FPGA.

REFERENCES

- [1] Akira Chiba. Magnetic bearings and bearingless drives. English. Oxford: Newnes, 2005. isbn: 9780750657273;0750657278;
- [2] Jan Krofta. Řízení aktivního radiálního magnetického ložiska. Praha, 2019. Diplomová práce. České vysoké učení technické v Praze.
- [3] Tomáš Vyhliďal. Řízení systémů a procesů. Téma 9: Aplikace stavové regulace s pozorovatelem stavu [přednáška]. Praha, 2020. České vysoké učení technické v Praze.

Graph neural network preprocessing for purpose of digital twin of mechanical system

Marek Ciklamini

*dept. of Instrumentation and Control Engineering
Czech Technical University
Prague, Czech Republic
Marek.Ciklamini@fs.cvut.cz*

Abstract—Graph neural networks might be brilliant tool how to effectively build digital twin of mechanical system and more.

The suggested pre-step technique of novel hybrid modelling is introduced as subset of the main hypothesis, where is assumed synergy between positives of hybrid modelling. Those are especially descriptiveness of model, which is characteristic for physical based modelling. Fast execution and results delivering is than underlying approach of data driven modelling.

Therefore is explored possibility into building empirical model based on extraction of FE model structure with aim to reinforcing whole model by new by information, f.g. measurement data.

Index Terms—digital twin, graph neural networks, finite element method, Fibonacci's spring

I. INTRODUCTION

Digital twin (DT) is currently well known words (one can say - nearly cliché) whose cannot be such easily overlooked. This term is carrying a simple idea in imitation of a real system in a virtual environment. DT has been usually connected in context with artificial intelligence and that is fact for most scientific fields. Since it is not absolutely clear what DT should have covered, it is obvious that there exist various approaches on how to build a DT of a certain system. This tool might have a big potential from a whole perspective to solving specific tasks, where different types of information are available. Those information might be measured data, but also models with certain accuracy describing particular phenomena.

Following lines are devoted to the topic of designing DT by usage of graph neural networks (GNN) with combination of finite element method (FEM) in order to store knowledge of some kind of mechanical or multi-physical system. For this purpose, it has to be initially defined the whole mechanical system in FEM software and then, this representation can be transferred to the graph network, which will be further used for the purpose of GNN in order to closely imitate behaviour of a real system.

First three subsections are aiming to briefly describe basic topics DT, GNN and FEM. Those are necessary ingredients for final compilation to a hybrid model of mechanical system. Next section is then describing how to pre-process and create GNS based on the geometry extracted from FEM model representation.

A. Digital twin

The concept of DT might be defined as an adaptive model of a physical system combining various approaches of mathematical modelling with connection of current technologies [1]. Those technologies can be multi-physical solvers, cybernetic of big data, artificial intelligence, augmented and virtual reality. DT can be also defined as the next milestone of simulations connected to the real world. This milestone was aptly depicted in [1] as a successor of product life cycle management approach (PLM), which is in exaggeration simple tool managing database of models and measured data concerned to a specific product.

B. Hybrid modelling

Term Hybrid modelling can be described as approach, which is combining two different types of model building - data driven modelling (DDM) and physical based modelling (PBM). DDM is not more, than building empirical model based on experimental data with usage of statistical methods, or with help of AI and machine learning (ML) methods. On other hand PBM is more transparent approach, with contrary to DDM. Usually, geometrical description of system with combination of FEM is helping to analyse system very deeply and inaccuracy can be easily found, since the partials equation of specific tasks are known. Main cons of PBM are traditionally very high computational cost, which should be reduced when the DDM is used.

C. Graph neural networks

GNN is used for model creation by nodes and edges [2]. Interaction between neighbours are defined by edges, which can (not) have specified direction concerning information exchange. A graph can describe various systems, from social networks in order to investigate voting preferences, or for purpose weather forecasts based on data from weather stations. Most Likely, GNN can be used for compiling information from FEM, where the model already has defined nodes and between them are specified behaviours by element information.

Architecture of probably the latest neural network is by author understood as a more corresponding architecture of the biological brain. The view is coming from naive idea, that the brain has no really defined input and output layer of neurons as

it is usual in standard feed forward neural networks (FFNN), but has specific areas, which are activated in time based on different actions. In other words, specific groups of neurons are excited by certain subjects. Another comparison GNN to FFNN is the fact that the biological brain is not built by a neat structure of layers of neurons, yet the chaotic graph is more representative, when close neurons are connected.

II. ILLUSTRATIVE MECHANICAL SYSTEM - FIBONACCI'S SPRING

For demonstration of main idea were designed simple geometry of Fibonacci's spring (FBS). FBS is not really following the standard golden spiral and geometry has a slightly different growing ratio of logarithmic spiral. The purpose of designed part is to have mechanical system with visually helpful geometry in order to briefly describe main topic of article.

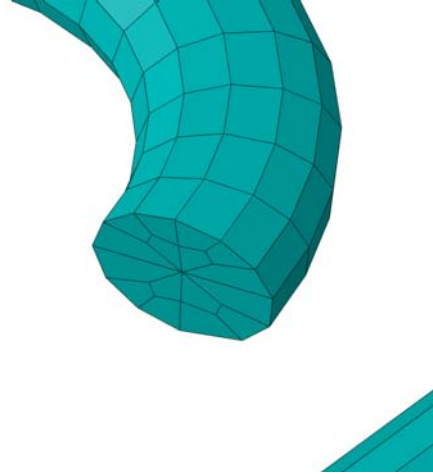


Fig. 2. Detail of swept pattern along the inner edge of FBS



Fig. 1. Representative example of mechanical system - Fibonacci spring

A. Mesh of FBS

Continuously growing mesh seed applied on the part is established only by hexahedral elements. The path defining swept pattern on the cross section is located on the inner edge of FBS. This pattern is chosen due to expectations of not sharp bricks in mesh structure, which brings at least usually good convergence behaviour of the FEM model important for potential further investigation.

Linear Hexahedral elements of the FBS part were chosen for the aspect of following simple numbering convention and therefore can be the shape of the part imitated back, once a GNN is built. Mentioned element is in Abaqus software well known abbreviation C3D8, which is not really recommended for tasks, where accurate model is needed, but for our purpose to initiate GNN is more than sufficient.

B. Input information for GNN

Input file (INP) generated in computer aided engineering (CAE) software is containing information regarding the whole FEM model needed for solver performing than calculation of task. INP is not more than a text file representing by its keywords the most important behaviour of a model, such as the

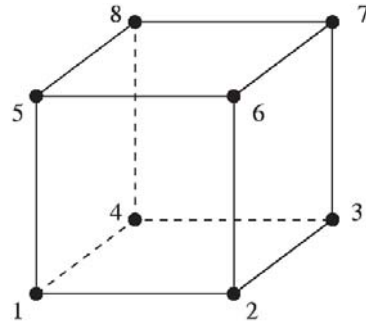


Fig. 3. Element C3D8 and convention of nodes numbering

geometry of the model, materials used, boundary conditions, load, etc. In this INP we will for the next step need information concerning mesh representation. Our essential keywords are:

- *Node ... specification of node label and its space position
- *Element ... with type definition (f.e.: type=C3D8) with element label and nodes shaping certain element

C. Building GNN

Once the information regarding FEM model mesh is available the GNN architecture can be initiated. For this purpose it is selected to use programming language Python and especially Deep Graph Library (DGL) which is mainly developed for deep learning applied on a graph structure. By DGL tool can then be defined graph G based on previously extracted nodes information N and edges E .

$$G = (N, E) \quad (1)$$

One interesting aspect about graphs is that not usually euclidean dependency is required and for their function, position of nodes in space are not really important. Therefore, following possible visualisation of FBS in GNN is not a big

surprise. On 4 can be seen detail of small area focusing on connection made between neighbouring nodes.

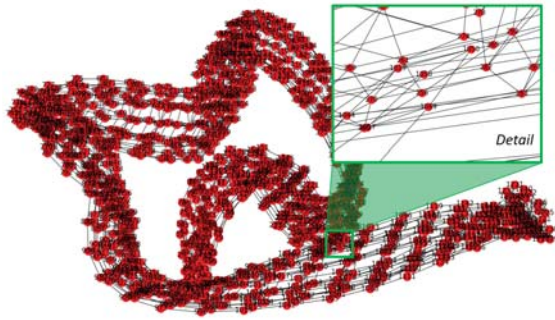


Fig. 4. Possible visualisation of PBS by GNN

Human eye needs to see the initial model and this need can be fulfilled, of course. Since, the INP is still available we can to G put information of N position and then visualise back the full part accordingly to specifications needed.

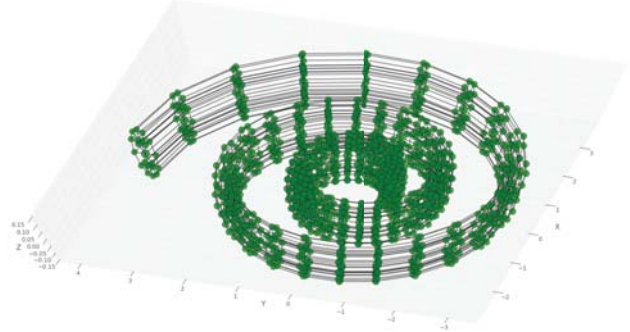


Fig. 5. Visualisation of PBS GNN with positions of nodes in order to validate correctness of built GNN

III. FURTHER WORKS IN GNN AND FEM HYBRID MODELLING

Previous lines has briefly introduced the main idea of author hypothesis, that graph neural networks as representative of data driven modelling, might be an elegant tool for building a digital twin with help of physical based modelling such as finite element method and was shown as initial step into exploration of topic.

Since, there are many hidden steps, the next one is to distinguish between hierarchical graph structures focusing on idea of loads definition present in system.

Another obvious milestone will be into extracting results data from FEM model and pre-process them (deflections, stresses, etc.) in order to have correctly prepared data-set for learning developed architecture of GNN.

IV. CONCLUSIONS

In the article were introduced 3 essential topics which are necessary for compilation of presented pre-processing of GNN.

Next section was devoted to a brief description of workflow dealing with conversion of the FEM model - FBS into the final requested graph.

Last but not least point of article was dealing with few lines introducing next author's intentions in the area focusing to develop theory concerning of building DT by a hybrid modelling with mentioned combination of mathematical approaches.

REFERENCES

- [1] Adil Rasheed, Omer San, Trond Kvamsdal, Digital Twin: Values, Challenges and Enablers, arxiv, 2019
- [2] Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, Philip S. Yu, A Comprehensive Survey on Graph Neural Networks, arxiv, 2019

Toleranční analýza světelné techniky v automotive

David Kovanda
fakulta strojní
Česká technická univerzita
Praha, Česká republika
david.kovanda@fs.cvut.cz

Jan Soukal
fakulta strojní
Česká technická univerzita
Praha, Česká republika
johny.soukal@gmail.com

Abstrakt—Článek se zabývá identifikací a analýzou tolerančních řetězců v konstrukci a výrobě předního světlomety automobilu a její vliv na výslednou fotometrii, která podléhá předpisům ECE. Hlavní naplní je vliv rozložení světelné stopy potkávací funkce ECE 112, která je tvořena indirektním reflektorovým systémem nebo direktním modulovým systémem na výrobních tolerancích.

Klíčová slova—toleranční analýza, automobilové osvětlení, ECE 112

I. ÚVOD

Automobilový průmysl je velice progresivní obor, který musí reagovat na požadavky uživatelů a nové technické trendy a nabízet atraktivní řešení a držet krok s konkurencí. Spojuje téměř celou paletu strojírenských oborů, které se v dnešní době neobejdou bez mezioborové spolupráce. Přední světlomety, které procházejí za posledních pár let masivní revolucí, stále zůstávají nezbytnou a zajímavou částí, která dává automobilu ojedinělý vzhled a duši. Často bývají základním poznávacím znamením dané značky a to díky různým tvarům denní signatury. Není proto divu, že se této části vnějšího vzhledu vozu věnuje taková pozornost. Také je snaha propojovat funkce jiných dílů právě se světlomety nebo do nich tyto funkce úplně implementovat. Současný trend osvětlení směřuje k nahrazení klasických žárovek či výbojek výkonnější LED technologií. To s sebou přináší možnosti a potenciál, o kterých se při použití žárovek ani nezdálo. Nejen že se zvyšuje vyzařovaný světelný tok, a tím vyšší intenzita osvětlení vozovky, lepší homogenita, možnost nových designových prvků, jako například použití světlovodů, nastavitelná chromatičnost zdroje, ale také se může uvažovat o animaci některých funkcí, nebo dokonce o

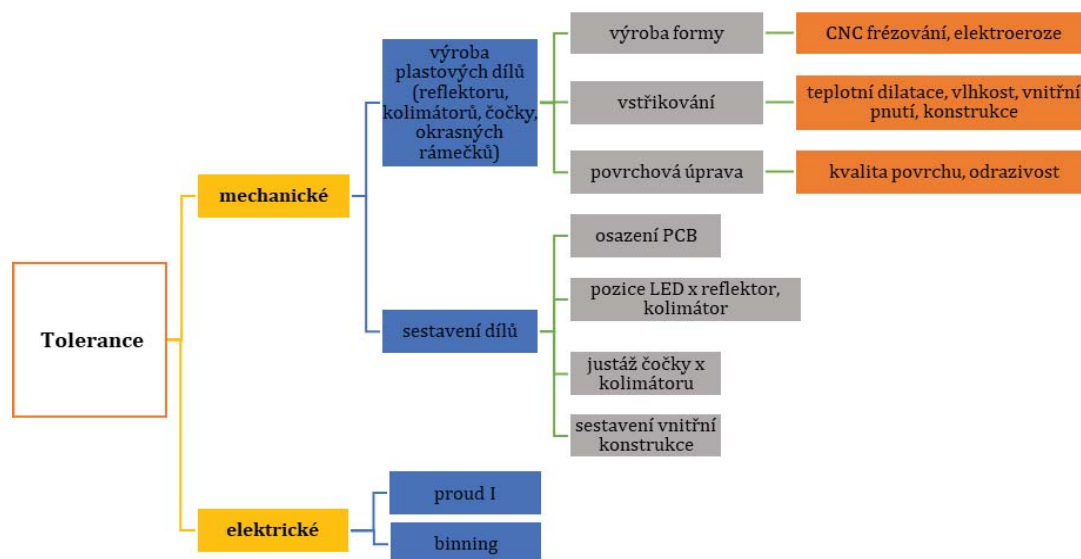
pokročilých systémech aktivních dálkových funkcí Matrix-LED s několika desítkami segmentů nebo „HD“ technologie, které mají tisíce segmentů a svým technickým řešením připomínají spíše než automobilová světla dalekohled. Cenou za tyto výhody je komplikovanost řešení oproti reflektorovým soustavám se žárovkami a jiné aspekty, které se dříve řešit nemusely.

II. TOLERANČNÍ ŘETĚZEC

Určení tolerančního řetězce (OBRÁZEK 1) výroby automobilového světlomety je velmi široká úloha, protože zdroje chyb se skládají jak z vlastní tolerance výroby jednotlivých dílů, tak z tolerance celkové sestavy. V rámci stanovení tolerančního řetězce byly rozděleny chyby do dvou oblastí – mechanické a elektrické. Do elektrických byly zahrnuty vlivy svázané se zdrojem světla, tedy s LED technologií. Jako první jsou zde uvedeny odchylky zdroje proudu, kterým jsou LED napájeny. Velikost světelného toku LED totiž přímo závisí na velikosti proudu. Do této kategorie byl také zařazen vliv binningu, tedy rozdílných vlastností dané skupiny LED (chromatičnost, světelný tok,...).

Mechanické chyby jsou odvozeny z výrobních technologií a z technologií které vymezují, které vymezují vzájemnou polohu ve světlomety. Výrobních technologií se používá velké množství a závisí od vyráběného dílu. Protože se jako konstrukční materiál používá plast, základní výrobní technologií je vstřikování do formy. Na samotném plastu také záleží, jakého je druhu, čistoty nebo vlhkosti. Během tohoto procesu dochází k vysokým teplotním rozdílům, se kterými souvisí teplotní rozpínatelnost materiálů. To zahrnuje velikost technologických parametrů během vstřikování jako například teplota taveniny, velikost tlaku, teplota a chlazení

OBRÁZEK 1- Určení tolerančního řetězce



formy. Po výstřiku na produkt působí děje, nazývané obecně smrštění, které mění jeho rozměry. Přesnost výroby formy pro vstřikování se odráží do výsledného dílu. Touto technologií se vyrábějí veškeré optické členy, jak indirektní reflektorové systémy, tak direktní systém skládající se z TIR kolimátorů a projekční čočky. Při použití indirektní reflektorové technologie se opatřují odrazné části reflektoru tenkou hliníkovou vrstvou, která se ještě překrývá ochranným průhledným lakem. Pórovitost a tloušťka tenkých vrstev ovlivňují odražené paprsky. Druhým druhem mechanických vlastností bylo zvoleno sestavení dílů do jejich vzájemné pozice avymezení vůlí. To zahrnuje jednak přesnost umístění LED na PCB osazovacími stroji a samotné ustavení PCB vůči reflektoru nebo PCB vůči držáku primární optiky (TIR kolimátorů). Oba aspekty vstupují do optických zákonů reflektoru nebo direktní optické soustavy a určují výsledné světelné rozložení potkávací funkce.

Téměř každý zdroj chyb působí na jinou část soustavy optické soustavy a má různě veliký vliv na výslednou světelnou stopu. Nelze je proto počítat. Rovněž nejde hodnotit stejným způsobem dva reflektorové systémy odlišných konstrukcí nebo stejné modulové řešení. Protože nemáme k dispozici ostrá data konstrukce, držíme se obecných předpokladů a zjednodušení. Proto se omezíme pouze na vychýlení zdroje světla vůči optickým prvků dané soustavě. V tomto jsme se pokusily shrnout veškeré technologické vlivy a určit velikost chyby.

Pro reflektorový systém budou uvažovány dvě skupiny chyb. Do první skupiny jsme zahrnuli vlivy, které budou mít za následek pohyb zdroje světla v rovně PCB. Druhá skupina ovlivní pozici LED v ose kolmé na PCB.

V úvahu bereme tyto zdroje: nepřesnosti výsledného reflektoru, teplotní roztažnost materiálů, nepřesnost osazení LED, vymezení vůle mezi reflektorem a PCB. Číselně jsme tyto vlivy vyjádřili následovně:

- celková nepřesnost reflektoru: 0,2 mm;
- teplotní zatížení materiálu: 0,2 mm;
- osazení LED: 0,05 mm;
- vůle mezi PCB a reflektorem: 0,05 mm.

Do první skupiny jsme zahrnuli vliv nepřesného osazení LED, vůli mezi PCB a reflektorem před její vymezení a nepřesnost reflektoru. Výsledná hodnota tedy je 0,3 mm. V druhé skupině uvažujeme teplotní zatížení materiálu – teplotní roztažnost. Velikost uvažované chyby druhé skupiny je 0,2 mm.

Pro projekční systém budeme uvažovat tyto zdroje nepřesnosti: vymezení vůle mezi TIR kolimátorem a PCB, nepřesnost osazení PCB a justáže projekční čočky. Číselně jsme tyto vlivy vyjádřili následovně:

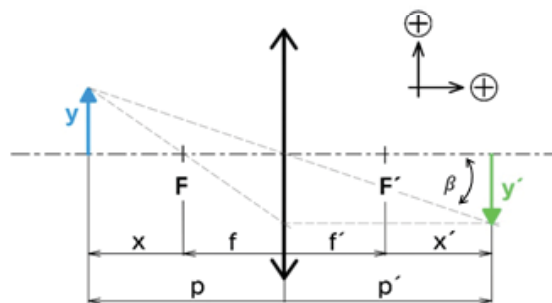
- justáž projekční čočky: 1 mm
 - osazení LED: 0,05 mm
 - vůle mezi PCB a TIR kolimátorem: 0,05 mm
- Výsledná chyba je tedy 2 mm.

III. CITLIVOST OPTICKÉHO SYSTÉMU NA VELIKOST OHNISKOVÉ VZDÁLENOSTI

Než se pustíme do samotné simulace těchto systémů, musíme ještě zmínit, že jeden z trendů vývoje světelné

techniky je neustále vše zmenšovat a přitom zachovat nebo ještě vylepšit světelné vlastnosti osvětlovacích soustav. Mezi nejlepší taktiky, jak toho dosáhnout je zkracovat ohniskovou vzdálenost. Tato změna znamená, že můžeme optický zdroj přiblížit blíže k optickému členu a tudíž dokážeme usměrnit větší část světelného toku LED zdroje. Tento trend má ale za následek, že se soustava stane mnohem citlivější na nepřesnosti při výrobě. Tento trend se dá matematicky odvodit, viz OBRÁZEK 2.

OBRÁZEK 2- Znázornění znaménkové konvence



Zdroj: Vlastní zpracování

Protože se během návrhu světelné soustavy používá výhradně úhlový popis, uvádíme zde odvozený vzorec:

$$\operatorname{tg} \beta = \frac{x^2 - f'x}{-f'^2 + f'x} \cdot \operatorname{tg} \alpha$$

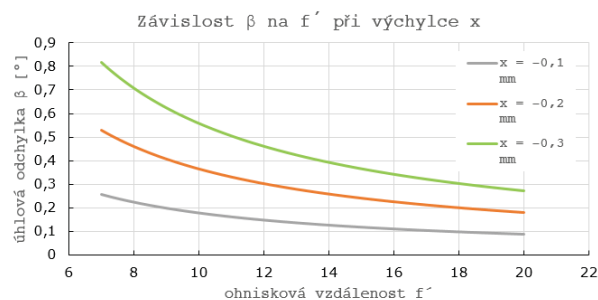
x ...vzdálenost předmětu od předmětového ohniska

f' ...obrazová ohnisková vzdálenost

α ...úhel mezi optickou osou a krajním paprskem vycházejícím z bodového zdroje světla

Úhel α jsme zvolili na základě modelových soustav. Na GRAFU 1 je vidět, že se zmenšujícím ohniskem roste úhlová odchylka. Tento případ jsme provedli pro 3 modelové chyby 0,1 mm; 0,2mm a 0,3 mm.

GRAF 1- Závislost β na f' při výchylce x



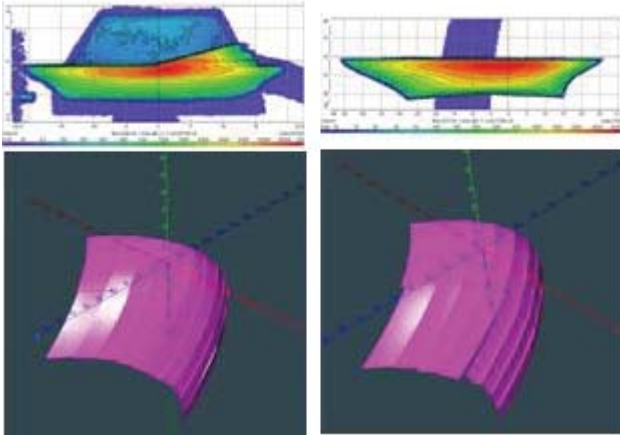
Zdroj: Vlastní zpracování

IV. SIMULACE

A. Reflektorový systém

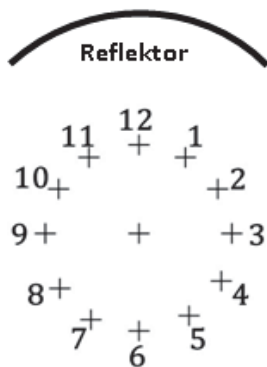
Pro simulaci byly použity dva multifokální reflektory (OBRÁZEK 3) s ohniskovými vzdálenostmi 20 mm pro reflektor tvořící spot a 18 mm pro reflektor předpolí. Postupně bylo se zdroji pohybováno do dvanácti bodů ve třech rovinách, viz OBRÁZEK 4.

OBRÁZEK 3- Světelný příspěvek jednotlivých reflektorů



Zdroj: Vlastní zpracování

OBRÁZEK 4- Označení vychýlených bodů



Zdroj: Vlastní zpracování

První rovina leží v rovině výchozího světelného zdroje – označujeme ji jako „0“. Další dvě roviny jsou s touto rovnoběžné a od ní vzdálené o 0,2 mm. Tu, která leží blíže k reflektoru, označujeme „+“ a rovinu vzdálenější „-“. Znázorněné body jsou od výchozí – středové polohy vzdáleny o 0,3 mm.

Pro každé vychýlení zdroje jsme po simulaci získali světelnou mapu. Popis jejích os je v úhlech. Na barevném měřítku pod mapou je odstupňována hodnota svítivosti jednotlivých bodů mapy. Protože potkávací funkce je realizována dvěma reflektory, součtem map z obou reflektorů získáme výslednou celkovou mapu. Na ni lze aplikovat přednastavenou analýzu předpisu ECE 112 13,2V. Podle této analýzy se vyhodnocují body a oblasti předepsané touto normou, zda plní nebo neplní předpis. Simulace byly provedeny metodou accurate s 1.mld paprsků.

Protože vyhodnocujících parametrů je 23, byly určeny 3, z nich, které bývají nejvíce problematické, viz OBRÁZEK 5. U některých parametrů je předepsána maximální hodnota svítivosti, která se nesmí překročit, u některých hodnota minimální, která nesmí být podkročena.

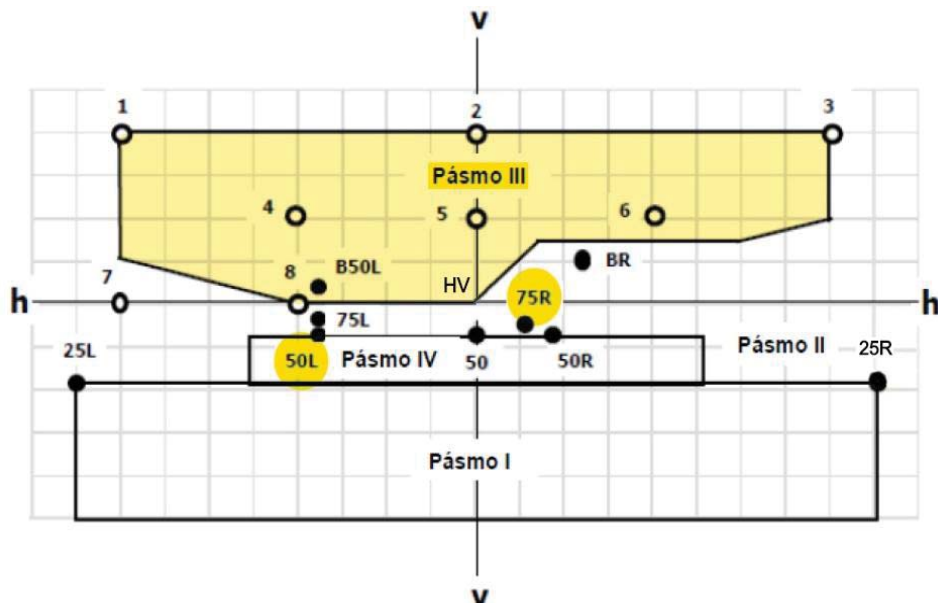
1) Vyhodnocení bodu 75R

V měřícím Bodě 75R norma předepisuje minimální hodnotu svítivosti 10100 cd. Na jednotlivá ramena paprskovitého grafu (GRAF 2) jsou zaneseny získané hodnoty svítivosti při odpovídajícím vychýlení zdroje světla. Tyto stavy spojuje modrá křivka – pro rovinu „-“, oranžová křivka – pro rovinu „0“ a fialová křivka – pro body v rovině „+“. Černá čárkovaná čára znázorňuje výchozí stav, tedy hodnotu svítivosti pro nevychýlený světelný zdroj. Zelená vyjadřuje předepsanou minimální svítivost. Z grafu můžeme vyčíst, že pro polohy od bodu 11 do bodu 4 neplní některé z nich předpis. Svítivost klesá až na 30 % požadované hodnoty. Nejhorší případy jsou pro body v rovině „-“. Svítivost ale také pro některé body roste, a to až o 160 %.

2) Vyhodnocení bodu 50L

Bod 50L se nachází pod bodem 75L a vyjadřuje hodnotu svítivosti na levé krajině 50 m od automobilu. Je umístěn do světelného vertikálního gradientu. V této oblasti je požadováno velké osvětlení bez oslnění protijedoucích vozidel. To právě zajišťuje tento gradient svítivosti, jehož oblast největší změny musí ležet pod horizontálou (h-h). Předpis pro tento bod stanovuje maximální přípustnou hodnotu svítivosti. Křivka změřených svítivostí musí tedy

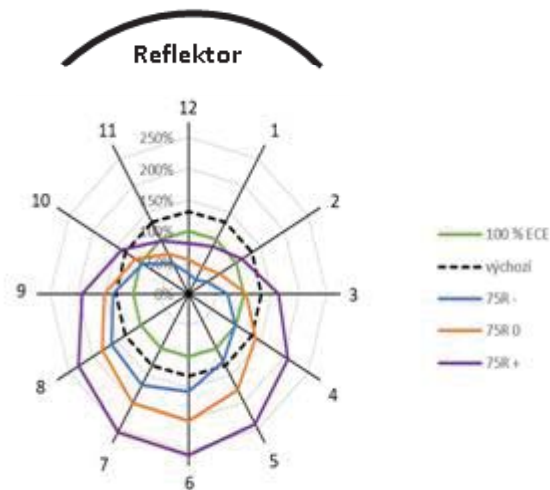
OBRÁZEK 5- Zvýraznění hodnotících parametrů na světelné mapě



Zdroj: Regulation ECE 112

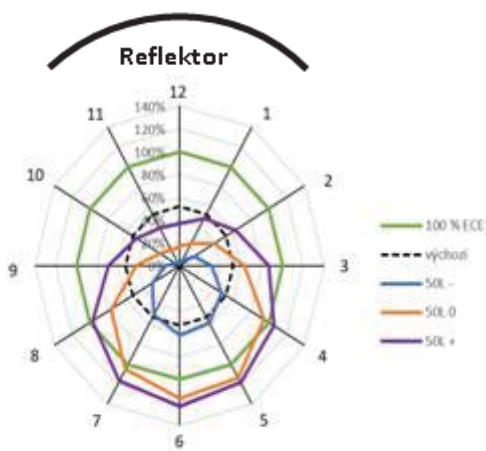
ležet uvnitř zelené křivky (GRAF 3), která vyjadřuje maximální povolenou svítivost. Černá čárkovaná křivka vyjadřuje hodnotu svítivosti nevychýleného zdroje. Z grafu je patrný nárůst svítivosti pro pozice 3 až 8 ve všech rovinách. Z toho lze usoudit, že se bude pro tyto polohy zvyšovat poloha gradientu až dojde k oslňování. Svítivosti s vychýlením do bodů 4 až 7 v rovinách „0“ a „+“ už jsou mimo předepsanou hodnotu.

GRAF 2- Paprskový graf pro bod 75R



Zdroj: Vlastní zpracování

GRAF 3- Paprskový graf pro bod 50L



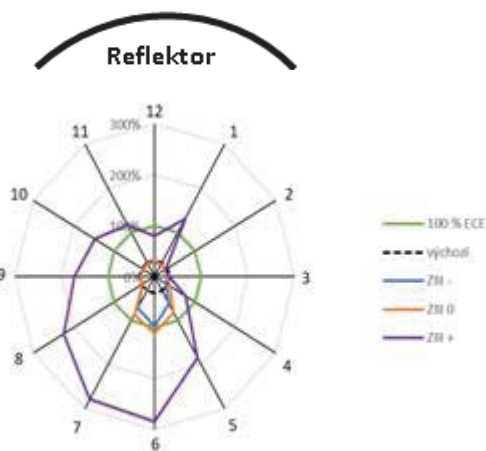
Zdroj: Vlastní zpracování

3) Vyhodnocení oblasti ZIII

Oblast ZIII má tvar mnohoúhelníku a nachází se nad horizontálou. Od bodu HV se její strana zvedá směrem vpravo pod úhlem 45° , kde je kink, a horní část je omezena úsečkou umístěnou 4° nad křivku h-h. ZIII obsahuje dalších 8 bodů, které se hodnotí vlastním kritériem. Mezi ně patří i HV. Od zmiňovaných bodů liší principem vyhodnocování. V žádném bodě oblasti nesmí být překročena hodnota 625 cd, aby nedošlo k oslňování. Hodnoty maximální svítivosti pro dané vychýlky zdroje jsou uvedeny v GRAFU 4. Body v rovinách „-“ a „0“ splňují předpis kromě polohy zdroje v bodě 6 v rovině „0“. Zdroje umístěné v rovině „+“ ve většině případů překračují maximální přípustnou hodnotu. Největší svítivost jsme zaznamenali v bodě 6 a zvedla se o 180 % proti normě. Mezi polohami 4 až 9 bylo místo největší

svítivosti v těsné blízkosti bodu HV. U ostatních případů se toto místo vyskytovalo spíše kolem dalších částí spodního ohraničení oblasti ZIII. Pro zdroje v této rovině tedy dochází k výraznějšímu zvedání světelné stopy, než je tomu u ostatních dvou rovin.

GRAF 4- Paprskový graf pro oblast ZIII



Zdroj: Vlastní zpracování

4) Shrnutí

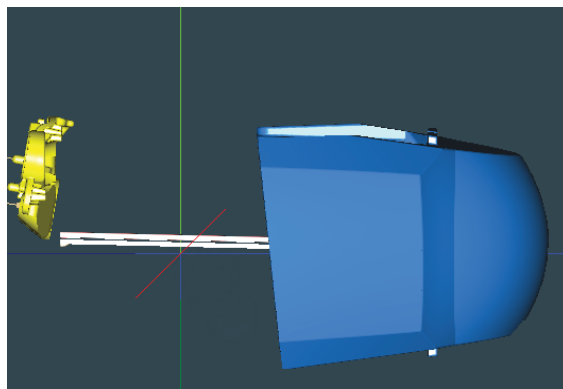
Vychýlení zdroje světla ze své výpočtové polohy deformuje rozložení osvětlení. Z provedených simulací je patrné, že každý vyhodnocovaný bod předpisem pro potkávací funkci je jinak citlivý na tyto vychýlení. Lze ale vypořádat určité trendy. První z nich bereme z pohledu vychýlení zdroje světla od výchozí polohy v rámci této roviny. Pro body v polohách 3 až 9 obecně dochází ke zvyšování svítivosti. Protože vyhodnocované body 75R, 50L a HV leží v oblasti blízké světelného rozrání, dochází obecně k vertikálnímu posunu světelné stopy.

Druhý pohled srovnává stejné body z jednotlivých rovin. Obecně lze z provedených simulací vyvodit, že změna polohy zdroje světla směrem od reflektoru (rovina „-“) snižuje svítivost ve sledovaném bodě. Naopak posunem zdroje světla směrem k reflektoru (rovina „+“) dochází ke zvyšování svítivosti. U některých bodů se zvyšuje rychleji, u některých pomaleji.

B. Projekční optický systém

Pro simulaci byla použita optická koncepce Bi – modul,

OBRÁZEK 6- Bi-modul



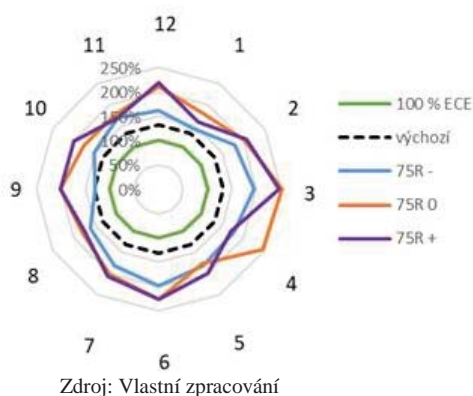
Zdroj: Vlastní zpracování

kteřý má v sobě implementované dvě světelné funkce, viz OBRÁZEK 6.

Jedna je dálková funkce a druhá je potkávací funkce. Simulovat se bude ale pouze optická soustava, která se stará o potkávací funkci. Tato soustava se skládá z 9 LED zdrojů, které jsou umístěna v ohniscích TIR kolimátorů, které přes projekční čočku tvoří stopu světelné funkce. V tomto konceptu se pro tvarování ostrého rozhraní a tvaru kinku použita kovová clonka. Stejně jako u reflektorového systému budeme pohybovat se zdroji ve dvanácti bodech a ve 3 rovinách, které jsou kolmé na osu šíření světla.

1) Vyhodnocení bodu 75R

GRAF 5- Paprskový graf bodu 75R



Z GRAFU 5 můžeme vyčíst, že pro všechny polohy jsou minimální požadavky na bod 75R splněny, spíše se defocusem a decentralizací významně zvětšují a to až 200% ECE.

2) Vyhodnocení bodu 50L

GRAF 6- Paprskový graf bodu 50L



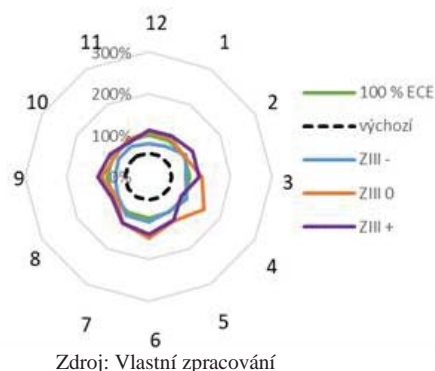
Fotometrický bod 50L už tak dobře nevychází. Z GRAFU 6 lze vyčíst, že ve všech rovinách jsme mimo zelenou křivku, tudíž nesplňujeme fotometrické maximum, které je povolené. Z toho vyplývá, že se všech vychýlených polohách, se zvýšila pozice gradientu a dochází k oslňování protijedoucích vozidel a tudíž by se světlo stal nelegální.

3) Vyhodnocení oblasti ZIII

Z GRAFU 7 lze vyčíst, že se ve všech rovinách se zvětšil příspěvek od nominálního stavu svítivosti téměř o 100% a tudíž se nachází mimo dovolenou oblast zelené křivky, která

označuje 100% ECE, což je maximální povolená svítivost. V tomto případě dochází k přímému oslňení zpětných zrcátek jedoucích vozidel a taktéž oslňení protijedoucích vozidel.

GRAF 7- Paprskový graf oblasti ZIII



4) Shrnutí

Z provedených simulací, lze vyvodit, že projekční systém, je oproti reflektorovému systému mnohem citlivější a že mnohem více deformuje celkové rozložení světelné stopy. Tento fakt je způsoben složitostí optického systému, který se oproti reflektoru, který se skládá pouze z LED zdroje a samotného reflektoru. Nýbrž projekční systém se skládá z 9 LED zdrojů, 9 TIR kolimátorů, kovové clonky a projekční čočky. Právě poloha kovové clonky je v tomto systému klíčová, protože vytváří ostré rozhraní a odclouňuje krajní paprsky, které míří do oblasti nad HV tudíž do oblasti ZIII, které způsobuje oslňování protijedoucích vozidel. Při vychýlení LED zdrojů v TIR kolimátoru dochází k vychýlování světelného kužele a následně způsobuje překročení povolených hodnot nad HV rozhraním.

V. ZÁVĚR

Cílem textu bylo seznámit vliv konstrukčních a výrobních tolerancí předních světlometů na podobu výsledné světelné stopy potkávací funkce, která je předepsána předpisem ECE 112. V textu jsou vyhodnoceny dva optické systémy. První řešení realizováno pomocí reflektorů, které jsou osvětleny LED zdrojem a druhé řešení je direktní řešení, které je realizováno TIR kolimátory a projekční čočkou. Z provedených simulací je zřetelné, že tolerance, které vznikají při výrobě, zasahují do výsledné stopy a mají vliv na legalitu celé potkávací funkce. Proto je nutné při konstrukci optických ploch dbát v úvahu i všechny tolerance, které v systému vznikají.

VI. PODĚKOVÁNÍ

Autoři článku chtějí poděkovat Studentské grantové soutěži SGS20/055/OHK2/1T/12.

ZDROJE

- [1] SOUKAL, JAN. Toleranční analýza potkávací funkce předního světlometu automobile, 2020, Diplomová práce . ČVUT v Praze, Fakulta strojní. Vedoucí práce Šárka Němcová

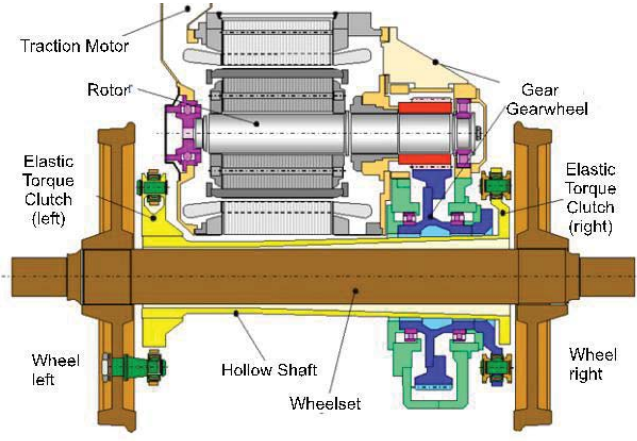


Fig. 2. Cross section of a wheel-set drive

The whole torque transmission path from the ASM to wheels, including the gear, right clutch, hollow shaft, left clutch, left wheel, axle, right wheel, we can depict it as if it were composed of such fictive substitutes. One of the possible ways how to do this is depicted in Fig. 3.

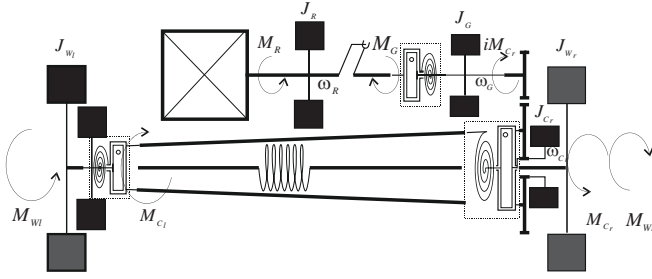


Fig. 3. Substitute concept of modelling dynamic behaviour of mechanical wheelset components using symbolic functional elements

B. Used Asynchronous Motor Modelling in dq0 frame and its disadvantages in control

The AC motor is considered as the main component of the train's wheelset drive unit, as it generates the required power to the wheels which push the train forward on the rails. Here comes the importance of creating an accurate model of ASM that takes into account changing dynamic characteristics when a comprehensive model of the wheelset unit is established for the purpose of simulation

When modelling an ASM with torque control, a variety of modelling methods used to represent the mathematical model of the motor. One of the favourite modelling methods known as direct quadratic frame dq0, where it converts the measured physical quantities such as voltage, current, and torque in the natural reference frame to abstract quantities (from three to two quantities) within a rotating reference frame. dq0 transformations or Park transformation [6] are based on the space phasor theory to study the dynamic behaviour of an ASM. These converted quantities cannot be measured directly on the ASM, and they are used as controlled variables when designing the ASM control system, they provide the ability to control torque and magnetic flux independently of each other.

In Fig. 4, a block scheme of the control structure with a drive motor cross section is shown, where the natural measurable physical quantities, i.e. voltages u_{SA}, u_{SB}, u_{SC} of the three-phase motor supply, inverter DC supply voltage

u_{DC} together with the current, the natural quantities are illustrated in blue color arrows. Other arrows and blocks, except torque setpoint representation, are in black because of their purely abstract character.

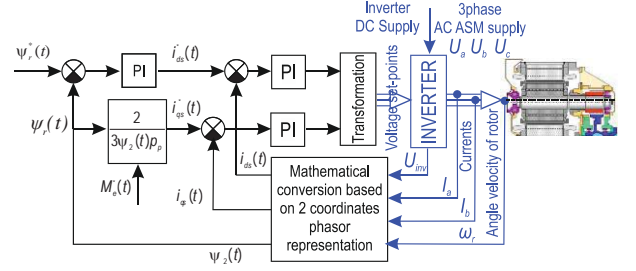


Fig. 4. A torque control scheme using abstract mathematical model based on dq0 frame representation and PI controllers.

The torque control set-point is defined, but the controlled variable is not directly measured, which results in disability in evaluating the topical control error. Also, the ASM should be equipped with an Inverter in order to generate control actions. The dq0 stand as converter between the physical reality, and abstract (mathematical) image of it, where the problems of controller design, optimal parameter setting are not easy to solve by standard time-domain methods. This has evoked our attention to study them in more detail, using the developed Simulink model.

Basically, indices dq express affiliation to orthogonal coordinate system, added index s linkage to the stator, r then to the rotor. For variables that are introduced only on the basis of the mathematically formulated definitions without any direct link to natural physical quantities common symbols of their physical counterparts are used, i.e. for currents i [A], voltages u [V]. Magnetic fluxes Ψ [Wb]. Also, the same physical units are assigned them because their values are computed according to the same formulas used for calculations in electrical circuit schemes with lumped parameters such as the resistance R in $[\Omega]$, the (mutual) inductance L in [H].

The Simulink model which we developed was based on the dq model equations as follows:

a. Stator voltage equations:

$$u_{qs}(t) = R_s i_{qs}(t) + \frac{d}{dt} \Psi_{qs}(t) - \omega_s(t) \Psi_{ds}(t) \quad (2)$$

$$u_{ds}(t) = R_s i_{ds}(t) + \frac{d}{dt} \Psi_{ds}(t) - \omega_s(t) \Psi_{qs}(t) \quad (3)$$

b. Rotor voltage equations:

$$u_{dr}(t) = R_r i_{dr}(t) + \frac{d}{dt} \Psi_{dr}(t) - (\omega_s(t) - \omega(t)) \Psi_{qr}(t) = 0 \quad (4)$$

$$u_{qr}(t) = R_r i_{qr}(t) + \frac{d}{dt} \Psi_{qr}(t) - (\omega_s(t) - \omega(t)) \Psi_{dr}(t) = 0 \quad (5)$$

c. Flux linkages of the stator and rotor are expressed in terms of the stator and rotor current in dq axis:

$$\Psi_{ds}(t) = L_s i_{ds}(t) + L_m i_{dr}(t) \quad (6)$$

$$\Psi_{qs}(t) = L_s i_{qs}(t) + L_m i_{qr}(t) \quad (7)$$

$$\Psi_{dr}(t) = L_r i_{dr}(t) + L_m i_{ds}(t) \quad (8)$$

$$\Psi_{qr}(t) = L_r i_{qr}(t) + L_m i_{qs}(t) \quad (9)$$

c. Magnetizing fluxes are defined as:

$$\Psi_{dm}(t) = L_m i_{ds}(t) + L_m i_{dr}(t) \quad (10)$$

$$\Psi_{qm}(t) = L_m i_{qs}(t) + L_m i_{qr}(t) \quad (11)$$

$$L_m = 2/3 L_{ms} \quad (12)$$

d. Electromagnetic torque expressed by:

$$M_e(t) = 2/3 P_p \left(\Psi_{ds}(t) i_{qs}(t) - \Psi_{qs}(t) i_{ds}(t) \right) \quad (13)$$

The stator and rotor winding are mutually coupled. Thus, when rotor angle changes, the mutual induction also changes i.e. as it is dependent on the rotor position.

The dq0 frame allows us to remove the time varying inductances position dependencies between the stator and the rotor, making it easier to model the motor, also, it decouples the magnetic flux linkage, and thus easier to control the motor torque. These advantages of dq0 representation may not be valid for many cases in which it is necessary to create a motor model in abc natural coordinates frame which takes into account transient dynamic changes and nonlinear components. This prompted us to propose a new ASM model based on the natural frame discussed in the next chapter.

III. NEW MODELLING METHOD OF ASYNCHRONOUS MOTOR IN ABC FRME

The process of delivering the optimal torque value from the ASM to the wheelset is subject to a variety of parameters and constraints such as degree of ASM model accuracy, control system stability and response, in addition to dynamical conditions influence the wheelset slip and adhesion such as weather conditions, and unpredicted changes in operation conditions.

When adhesion of the wheelset is lost, a significant transient changes in the motor voltage could be noticed [7] due to the fluctuation in the motor traction torque in order to reestablish the adhesion of the wheels to the rails. This unwanted situation affect the torque transmitted to the wheels. In transient conditions, a number of harmonic components which has a nonlinearity nature are produced [4] which results in the inability to directly obtain real values of the currents since those linearity components are located in space harmonic vector. These quantities affecting the relationship between the original current in abc natural frame and the transformed current in dq0 frame are difficult to be represented precisely based on dq analysis since it has a linear nature. In addition, dq transformation model deals with id current as a function of iq, which eliminates the possibility to do simulation independently from each other, and also the dq model is not able to identify the differences of the phase currents. Considering the above-mentioned reasons, attention is switched to developing a new ASM model that can accurately and precisely reflect the transient conditions and steady-state operation of the ASM.

Some of the literature work focused on modeling ASM in abc frame using Matrix partitioning was done by Dimitrovski and Luther [5]. We propose an ABC natural coordinate model Fig. 5 of which should improve the accuracy of the transient process and controlling the response of the parameters of the analysis more precisely knowing that in practice, the currents flow through phases are not equal.

The stator voltage differential equations are defined as follows

$$u_{sA}(t) = R_{sA} i_{sA}(t) + \frac{d}{dt} \Psi_{sA}(t) \quad (14)$$

$$u_{sB}(t) = R_{sB} i_{sB}(t) + \frac{d}{dt} \Psi_{sB}(t) \quad (15)$$

$$u_{sC}(t) = R_{sC} i_{sC}(t) + \frac{d}{dt} \Psi_{sC}(t) \quad (16)$$

Similarly, the rotor voltage differential equations are

$$u_{rA}(t) = R_{rA} i_{rA}(t) + \frac{d}{dt} \Psi_{rA}(t) \quad (17)$$

$$u_{rB}(t) = R_{rB} i_{rB}(t) + \frac{d}{dt} \Psi_{rB}(t) \quad (18)$$

$$u_{rC}(t) = R_{rC} i_{rC}(t) + \frac{d}{dt} \Psi_{rC}(t) \quad (19)$$

In real operation of the ASM, the magnetic flux of the stator and rotor are mutually connected, and the magnetic flux of the one winding depends on the magnetic fluxes of the rest of the windings. The magnetic flux equation of the new ASM model can be stated in matrix form as follows:

$$\begin{bmatrix} \Psi_A(t) \\ \Psi_B(t) \\ \Psi_C(t) \end{bmatrix} = [L_{ms}] \cdot [I_{sABC}] + [L_{mr}] \cdot [I_{rABC}] \quad (21)$$

$$= \begin{bmatrix} L_{sA} & L_{sAsB} & L_{sAsC} \\ L_{sAsB} & L_{sB} & L_{sBsC} \\ 0 & L_{sBsC} & L_{sAsC} + L_s \end{bmatrix}$$

Where is

$$[L_{mr}] = \begin{bmatrix} L_{sArA} & L_{sArB} & L_{sArC} \\ L_{sBrA} & L_{sBrB} & L_{sBrC} \\ L_{sCrA} & L_{sCrB} & L_{sCrC} \end{bmatrix} \quad (23)$$

$$[I_{sABC}] = \begin{bmatrix} I_{sA} \\ I_{sB} \\ I_{sC} \end{bmatrix} \quad (24)$$

$$[I_{rABC}] = \begin{bmatrix} I_{rA} \\ I_{rB} \\ I_{rC} \end{bmatrix} \quad (25)$$

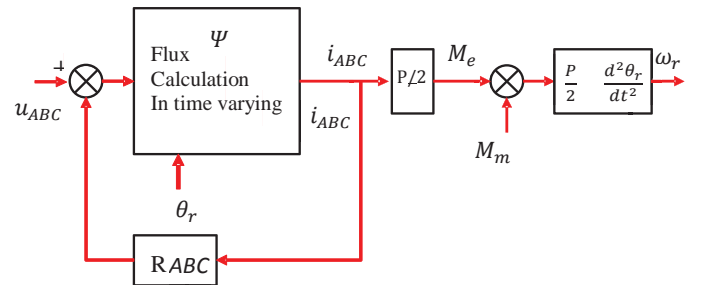


Fig. 5. ABC natural coordinate model proposal for ASM

Where in (14) – (25) symbols $i_{sA}, i_{sB}, i_{sC}, i_{rA}, i_{rC}$ are the natural 3-phase currents. $u_{sA}, u_{sB}, u_{sC}, u_{rA}, u_{rB}, u_{rC}$ are the natural 3-phase stator and rotor voltages as follows.

IV. PROPOSAL OF MODEL PREDICTIVE CONTROL FOR ASM

If an inadequate tractive torque is detected, it may lead to an inadequate wheel slip and a subsequent limiting vehicle dynamics. So far, a variety of vector modelling method like Field Oriented Control (FOC) and Direct Torque Control (DTC) [8] has been used in practice to control the ASM. These methods depend on two reference frames, stationary which is based on alpha and beta abstract quantities and known as Clark transformation [9], and arbitrary which based on direct quadratic abstract quantities that were used to simulate our ASM Simulink model.

FOC emulates the behaviour of ASM as if it is a DC machine by decomposed the stator currents into two-axis, d-axis correspond to (Flux) and q axis correspond to (Torque) which results in the decoupling of torque and flux control. Those separated quantities are then regulated using a two tuned PI controller, and good steady-state stability and fast dynamic response are obtained. On the other hand, DTC directly selects the best voltage vector using a predefined table based on stator flux position and error signals of the torque and stator flux. Thus, by eliminating the use of current control of the ASM, a faster dynamic response is achieved.

The features which FOC and DTC present, come with limitation represented in low-speed operation and computational effort and inability to handle constraints which influence the stability, in addition to difficulties in considering the nonlinear components in the ASM. It is also good to mention that based on some previous experience, the good tuning of the PI Controller in an untraditional control structure standardly used in the ASM torque is not a straightforward process, since the PI control loops work with signals computed in this model, not directly with physical quantities.

In order to approach a precise ASM model based on abc frame, we need to consider the nonlinearities nature of the stator flux which is not decoupled in abc frame due to harmonic components such as saturation. These nonlinear components are difficult to be represented in the dq0 frame.

Based on what previously mentioned, it is good to obtain a new control model based on Nonlinear Model Predictive Control (NMPC) [10] and replacing the implemented PI controllers in order to facilitate the adjustment. The benefits of the necessary state-of-the-art in-system order reduction, linearization for a computer simulation that is quite dependent on step size may be indispensable. In general, The MPC controller has a quick response, multivariable control, ability to handle various nonlinear constraints, also, MPC [11] can eliminate the use of the inner and outer loops and result in an increment of steady-state error. On the other hand, the computational effort, the requirements of the high sampling frequency, and the difference in the torque and stator flux amplitudes and units restrict the implementation speed rate of the traditional MPC in real application.

We propose a new control strategy based on the NMPC controller for voltage vector control and disturbance observing which caused by load torque variation and parameter deviation. The new control strategy will use an explicit nonlinear MPC, where an additional loop of current

control will be neglected. Increasing the predicted instances of both torque and stator flux will help us in obtaining the best voltage vector come through minimizing MPC cost function J.

Developing such a new control method will result in a better stability and higher performance since the physical quantities will be directly measured in abc frame without using dq0 frame. The proposed MPC controller is depicted in Fig. 6.

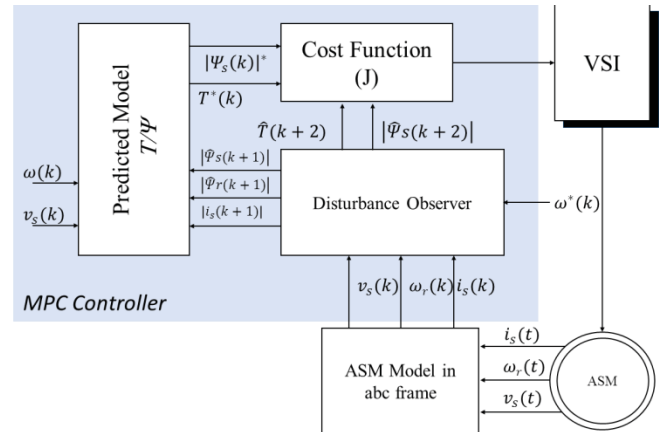


Fig. 6. Block scheme of proposed NMPC controller for ASM drive

The expected results of the MPC controller comparing with PI controller may include a better transient response, better robustness and improved stability in face of uncertainties including load disturbance, and torque fluctuating transmitted to the locomotive wheelset.

CONCLUSIONS

The urgent need, and the growing interest in obtaining an accurate model of ASM which consider the transient changes that affect the torque transmitted to the wheels under undesirable operating conditions, prompted us to propose a new model of an ASM based on its representation in the natural coordinate frame instead of the commonly used two-axis dq0 reference frame. In this paper, we have emphasized the importance of the proposed model in its ability to take into account nonlinear components that appear during motor operation. The conventional control method such as FOC and DTC cannot be employed with the proposed motor model, as it deals with torque and magnetic flux within a two-axis dq0 frame, so we proposed a new control system based on a nonlinear predictive controller. By predicting the motor variable within a predefined prediction horizon, taking into account all possible voltage vectors, it is possible to select the optimal voltage vector by minimizing the cost function of the controller.

ACKNOWLEDGMENT

The creation of this paper was supported by The Czech Technical University and the related grants – No. SGS18/177/OHK2/3T/12..

REFERENCES

- [1] P. Kaderavek, J. Pernicka, "Torsion oscillations of powered wheelsets," Railvolution, 2013, pp. 34-37.
- [2] J. Kolář, "Modern Trends in the Drive Wheelsets of Rail Vehicles", Springer International Publishing, 2016, pp. 27-35.

- [3] V. Dybala, B. Šulc, "Simulation of influence of transverse wheel-set movement on torsion oscillation" Technical University of Ostrava, , 2019, p.1-47.
- [4] R. H. Park, "Two-reaction theory of synchronous machines generalized method of analysis-part I," Trans. Am. Inst. Electr. Eng., vol. 48, no. 3, pp. 716–727, 1929.
- [5] T. Fridrichovský, B. Šulc, "Investigation of Torsional Oscillations in Railway Vehicles", MATEC Web of Conferences 76, 2016.
- [6] J. M. Aller , D. Delgado, A . Bueno, J. C. Viola, J. A. Restrepo, "Model of the Induction Machine including Saturation," 15th European Conference on Power Electronics and Applications (EPE). 2013.
- [7] R. Dimitrovski, M. Luther, "Modeling and Simulation of an induction machine in the abc-reference frame using inversion of a matrix by partitioning" ICREPQ 16, No.14 ,May 2016.
- [8] D. Casadei, F. Profumo, G. Serra, A. Tani "FOC and DTC: two viable schemes for induction motors torque control" IEEE Trans. Power Electron., vol. 17, no. 5, pp. 779-787.
- [9] W. C. Duesterhoeft; M. W. Schulz; E. Clarke. "Determination of Instantaneous Currents and Voltages by Means of Alpha, Beta, and Zero Components". Transactions of the American Institute of Electrical Engineers, July 1951. 70 (2): 1248–1255.
- [10] M. Diehl, R. Findeisen, S. Schwarzkopf, I. Uslu, F. Allgöwer, H. G. Bock, J. P. Schlöder, "An efficient algorithm for nonlinear model predictive control of large-scale systems. Part I: Description of the Method. At Automatisierungstechnik". 2002., 50 (12), 557–567.
- [11] W. Liuping, "Model Predictive Control System Design and Implementation Using MATLAB ®". Springer Science & Business Media. pp. xii, 2009

Modeling correction factor for metal sheet rolling via polynomial fuzzy system

1st Adam Peichl

Department of Instrumentation and Control Engineering
Faculty of Mechanical Engineering
Czech Technical University in Prague
Prague, Czech Republic
adam.peichl@fs.cvut.cz

Abstract—This paper proposes a fuzzy based polynomial model for prediction of force correction factors for metal sheet rolling. The goal of this work is to design, test and implement set of robust algorithms in C++.

Index Terms—mathematical modeling, metal sheets rolling, fuzzy, polynomial, machine learning

I. INTRODUCTION

Correct design of metal sheet rolling production process needs to determine forces and moments with adequate precision. Unfortunately the process is so complicated that need for some kind of numerical adaptive model emerges.

In [1] we designed model based on *polynomial neural units* (HONUs). This model proved to be robust, simple and reliable enough to be deployed and successfully tested in real conditions.

Unfortunately production line for metal sheet rolling usually needs to process different kinds of metals and model proposed in [1] assumed that for every single metal there will be dedicated polynomial unit, which proved to be a little bit unpractical at best.

Therefore we propose usage of *polynomial fuzzy system* which preserves desirable properties of original model to some degree but also removes the need for multiple separately trained and stored polynomial units.

II. METHODOLOGY

The methods and approaches used to accomplish the goal of this project are described in this section.

A. Preprocessing

Since our model works with various different inputs and those inputs also have huge scale differences¹ it is important to put all our inputs into similar scale. Therefore we use (same as in [1]) *z-scoring* which is defined as follows:

$$x_z = \frac{x - \bar{x}}{\sigma_x} \quad (1)$$

Where x is original input, \bar{x} is mean of original input, σ_x is standart deviation of original input and x_z is z-scored input.

This work was supported by grant SGS18/177/OHK2/3T/12. Algorithms were developed in collaboration with PT Solutions Worldwide s.r.o.

¹For instance *temperature* of metal sheet is in hundreds of °C and *aspect ratio* is usually somewhere between 0.8 and 2.0.

From now on, if there is input somewhere (in text, equation etc), we think of that as it is already z-scored.

B. Polynomial Fuzzy System

Proposed fuzzy system is basically the same as the one proposed by Kim and Vachtsevanos in [2].

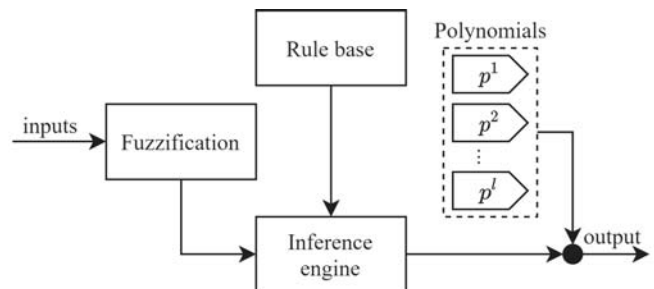


Fig. 1. Scheme of a polynomial fuzzy system

Fuzzy IF-THEN rules are defined as follows:

$$IF \bigwedge_{x_i \in \mathbf{x}} (x_i \text{ is } A_i^l) \text{ THEN } y \text{ is } p^l(\mathbf{x}) \quad (2)$$

The fuzzy system can be described via following equation:

$$f(\mathbf{x}) = \frac{\sum_{l=1}^M p^l(\mathbf{x}) \prod_{i=1}^n \mu_{A_i^l}(x_i)}{\sum_{l=1}^M \prod_{i=1}^n \mu_{A_i^l}(x_i)} \quad (3)$$

Where $p^l(\mathbf{x})$ is a polynomial connected to fuzzy IF-THEN rule l and $\mu_{A_i^l}(x_i)$ is a membership function of a fuzzy set A_i^l defined as follows:

$$\mu_{A_i^l}(x_i) = \exp\left(-\left(\frac{x_i - c_{A_i^l}}{\sigma_{A_i^l}}\right)^2\right) \quad (4)$$

Where $c_{A_i^l}$ is center and $\sigma_{A_i^l}$ is a parameter affecting the width of gaussian membership function.

C. Training Algorithm

In [2] they actually trained their model via complicated *genetic algorithm*. Since our system is much smaller and

simpler, we will use *stochastic gradient descent* as described in [3]:

$$\theta_{k+1} = \theta_k - \alpha \nabla_{\theta} Q(\mathbf{x}_r, y_r; \theta_k) \quad (5)$$

Where θ is a vector of *trainable parameters*², α is *learning rate* and $Q(\mathbf{x}_r, y_r; \theta_k)$ is standard SE criterion, defined as follows:

$$Q(\mathbf{x}_r, y_r; \theta_k) = (f(\mathbf{x}_r; \theta) - y_r)^2 \quad (6)$$

Where input-output pair \mathbf{x}_r, y_r is randomly selected from our training set.

III. MODEL VALIDATION AND TESTING

Model was validated on small number of input-output pairs not used for training process.

As you can see on figures 2 and 3, there is significant improvement in absolute force and torque correction factor errors.

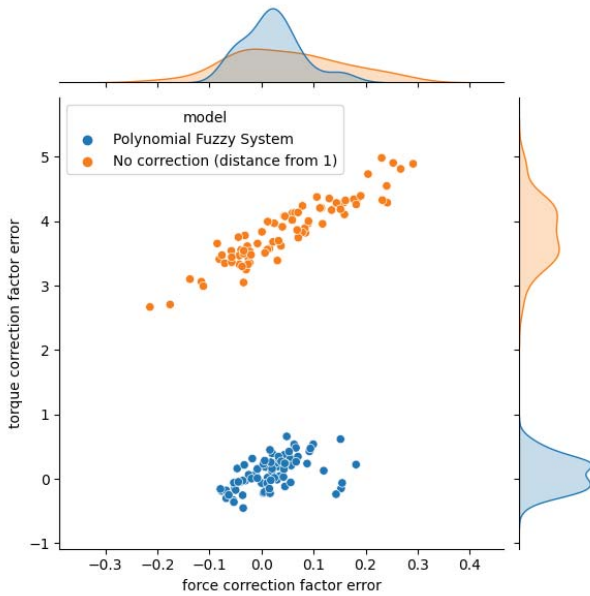


Fig. 2. Scatter plot of errors for roller FM2 for 80 input-output pairs left for validation (not used for training)

Model will be tested on another huge set of measured data in 2021.

IV. IMPLEMENTATION

The whole set of classes and functions was implemented in C++. The result of this project is C++ header library.

V. CONCLUSION

The proposed model solves problem with multiple separated polynomial units.

Model was validated on small set of real measured data and results shown are very promising.

²So in our case all centers and sigmas of gaussian membership functions and coefficients of polynomials

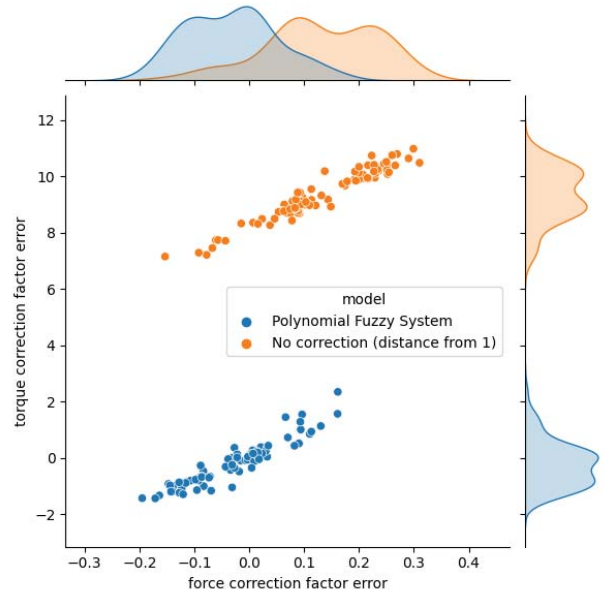


Fig. 3. Scatter plot of errors for roller FM1 for 80 input-output pairs left for validation (not used for training), you can see, that results for this roller are not that great (but still good in comparison to no correction) as in figure 2

Model was not yet fully tested in real application. On the other hand, when properly trained, it should not be worse than single polynomial unit.

ACKNOWLEDGMENT

Special thanks to PT Solutions Worldwide s.r.o., all algorithms were developed in collaboration with them. Also, this work was supported by grant SGS18/177/OHK2/3T/12.

REFERENCES

- [1] Pechl, Adam, Oswald, Cyril. Predikce silových a momentových korekčních faktorů pro válcování kovů pomocí AI, Studentská tvůrčí činnost 2019. Praha: České vysoké učení technické v Praze, Fakulta strojní, 2019. ISBN 978-80-01-06564-8.
- [2] Kim, Sungshin and Vachtsevanos, George J, A polynomial fuzzy neural network for identification and control, Proceedings of North American Fuzzy Information Processing, 1996
- [3] Bottou, Léon. Large-scale machine learning with stochastic gradient descent. In: Proceedings of COMPSTAT'2010. Physica-Verlag HD, 2010

Pasivní magnetické ložisko

Příprava a test pro hybridní magnetické ložisko

Martin Procházka

Fakulta strojní

České vysoké učení technické v Praze

Praha, Česká republika

martin.prochazka@fs.cvut.cz

Abstrakt—Tento článek se zabývá pasivním magnetickým ložiskem, testovacím přípravkem pro ložisko a statickými silovými testy ložiska v rámci přípravy přeměny pasivního magnetického ložiska na hybridní magnetické ložisko (pasivní magnetické ložisko s aktivním tlumením kmitání).

Klíčová slova—pasivní magnetické ložisko, test magnetického ložiska, hybridní magnetické ložisko

I. ÚVOD

Magnetické ložisko je ložisko, kde rotující část je držena ve středu ložiska pomocí magnetického pole. Vnitřní a vnější kroužek ložiska (nebo vnější část a hřídel) se navzájem nedotýkají, je mezi nimi vzduchová mezera. [1]

Magnetická ložiska se dělí podle principu funkce na pasivní a aktivní. U pasivních magnetických ložisek je hřídel držena ve středu ložiska pomocí odpudivých magnetických sil mezi vnitřním a vnějším kroužkem – permanentním magnetem. U aktivního magnetického ložiska se vystředění hřídele v ložisku děje pomocí přitažlivé magnetické síly mezi aktivně řízenými elektromagnety ve vnější části ložiska a feromagnetickou hřídelí. [1]

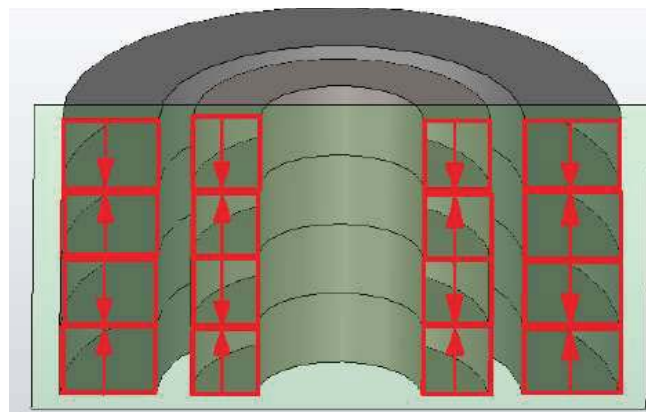
Tato ložiska mají několik výhod, převážně kvůli absenci tření [1]:

- Nemusejí se mazat
- Nedochozí zde k opotřebení a vzniku prachových částic a mají tedy vysokou životnost
- Mají zanedbatelný mechanický odpor, tudíž jsou zde zanedbatelné ztráty mechanické energie.
- Jsou vhodné pro velmi vysoké otáčky

Mají ale také své nevýhody [1] [2]:

- Nízká nosnost vzhledem k velikosti
- Omezená provozní teplota, pokud jsou použity permanentní magnety (např. některé neodymové magnety ztrácí své vlastnosti už při 80°C)
- Spotřeba elektrické energie u aktivních magnetických ložisek
- Relativně složité řízení aktivního magnetického ložiska
- Náhylnost ke kmitům a vibracím u pasivního ložiska

Je také možné tyto způsoby zkombinovat a vytvořit tzv. hybridní magnetické ložisko, což bude cílem tohoto projektu.



Obr. 1: Magnety v pasivním magnetickém ložisku [2]

Hybridní magnetické ložisko je pasivní magnetické ložisko doplněné o aktivní tlumení kmitů a vibrací. Oproti aktivnímu ložisku se předpokládá menší zástavbový prostor, jednodušší řízení a nižší spotřeba elektrické energie.

II. PASIVNÍ MAGNETICKÉ LOŽISKO

A. Důvod měření pasivního magnetického ložiska

Pro vytvoření hybridního magnetického ložiska bylo potřeba zkonstruovat nejdříve pasivní magnetické ložisko a na testovacím/měřicím přípravku je proměřit. Podle změřených dat následně proběhne vytvoření matematického modelu ložiska, návrh řízení a výběr a dimenzování hardwarových prvků aktivního tlumení (např. dimenzování tlumících elektromagnetů).

B. Pasivní magnetické ložisko - popis

Pro tento projekt bylo vybráno provedení magnetického ložiska s osmi neodymovými axiálně polarizovanými prstencovými magnety. Magnety jsou po čtyřech ve vnějším i vnitřním ložiskovém kroužku a jsou otočeny póly proti sobě, čímž tvoří tzv. Halbachovo uspořádání (obr. 1), pomocí kterého lze zesílit magnetické pole v místě kde potřebujeme (viz zdroj [2]). V praxi se toto projeví vyšší nosností ložiska, ale také vyšší citlivostí ložiska na axiální posunutí ložiskových kroužků. Zástavbové rozměry ložiska jsou vnitřní průměr $r=10$ mm, vnější průměr $R=40$ mm a hloubka $h=24$ mm. [2]

III. TESTOVACÍ PŘÍPRAVEK

A. Požadavky

Na testovací přípravek byly kladeny tyto požadavky:

- Vytvoření ložiskového domečku tak, aby se magnety nemuseli lepit a ložisko bylo rozebíratelné. Použité magnety jsou velice křehké a při sestavování, či měřicích pokusech se

může stát, že bude potřeba vyměnit poškozený magnet.

- Upevnění testovací hřídele zajišťující nastavitelnou vzájemnou axiální polohu ložiskových kroužků a zároveň umožňující jejich vzájemný volný radiální pohyb.
- Volný zástavbový prostor pro následné připojení vysokootáčkového motoru.
- Variabilní sestavení pro budoucí aplikaci různých prvků aktivního tlumení vibrací.

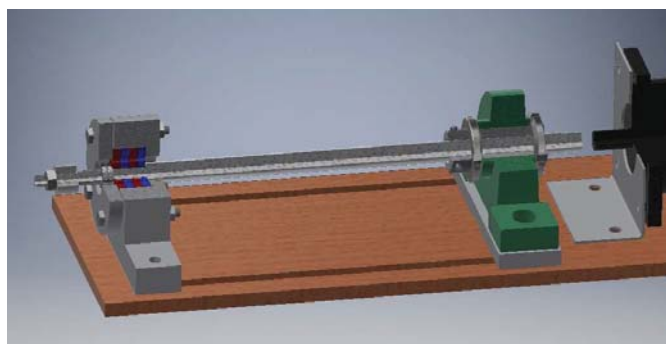
B. Realizace

Provedení přípravku bylo realizováno na základě výše zmíněných bodů. Domeček ložiska je sešroubován z plastových dílů výtisků na 3D tiskárně. Magnety vnitřního kroužku ložiska jsou navlečeny přímo na hřídeli a zajištěny z jedné strany osazením a z druhé strany dvěma KM maticemi. Závit pro KM matice je delší z důvodu snadnější montáže odpuzujících se magnetů. Hřídel je na druhém konci zajištěna v naklápěcím valivém ložisku s vložkou umožňující přesné axiální nastavení hřídele pomocí dvou KM matic.

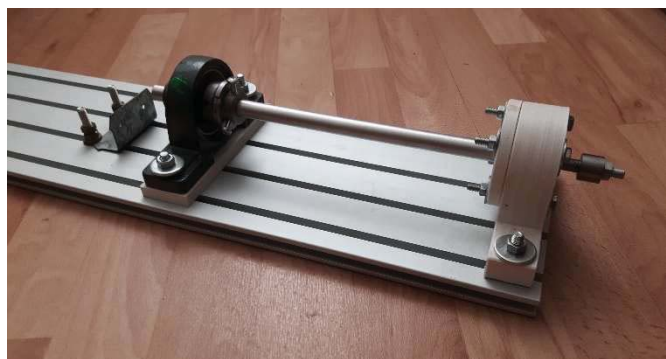
Hřídel má konec na straně valivého ložiska upraven pro připojení pružné spojky a vysokootáčkového motoru. Na konci bližším magnetickému ložisku je připevněn feromagnetický váleček, který bude přitahován magnety aktivního tlumení (samotná hřídel je z hliníkové slitiny).

Na obr. 2 je předběžný návrh přípravku s viditelnými ložiskovými magnety (modře a červeně). Na obr. 3 je už realizace přípravku s bílým domečkem magnetického ložiska a zeleným valivým ložiskem.

V průběhu měření byl přípravek dle potřeby doplněn dorazy pro měření polohy hřídele.



Obr. 2: Návrh testovacího přípravku

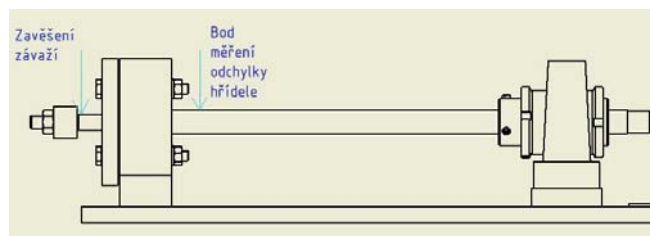


Obr. 3: Realizace testovacího přípravku

IV. STATICKÉ MĚŘENÍ SILOVÉ CHARAKTERISTIKY LOŽISKA

A. Radiální měření

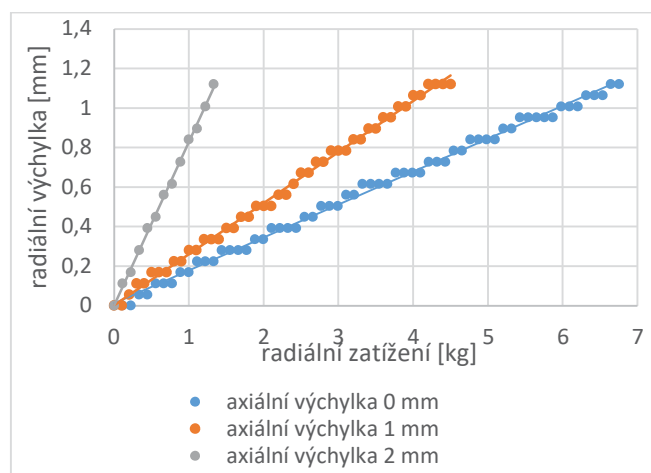
Nejdůležitějším měřením pro následnou realizaci hybridního ložiska je změření průběhu radiální výchylky hřídele v závislosti na jejím zatížení. V průběhu měření bylo na konec hřídele u magnetického ložiska zavěšováno závaží po 100 gramech a při tom byla měřena radiální výchylka hřídele (obr. 4). Jelikož není možné závaží ani měření umístit do středu ložiska, byly naměřené hodnoty i hmotnost závaží přepočítána do místa středu ložiska.



Obr. 4: Schéma radiálního měření

Měření proběhlo třikrát, pokaždé pro jiné axiální nastavení hřídele. Poprvé pro přesné nastavení ložiskových kroužků proti sobě a následně s axiální výchylkou 1 a 2 mm. Měření bylo provedeno od nulové zátěže až do zatížení, kdy se vnitřní a vnější ložiskové kroužky vzájemně dotýkaly.

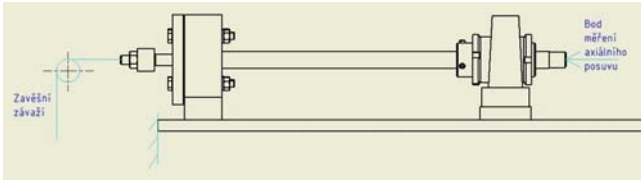
Z grafu (obr. 5), kde jsou taktéž zobrazeny porovnávací přímky pro zhodnocení linearity, vyplývá, že průběh výchylky je jen mírně odlišný od lineárního průběhu a že axiální výchylka má značný vliv na nosnost ložiska.



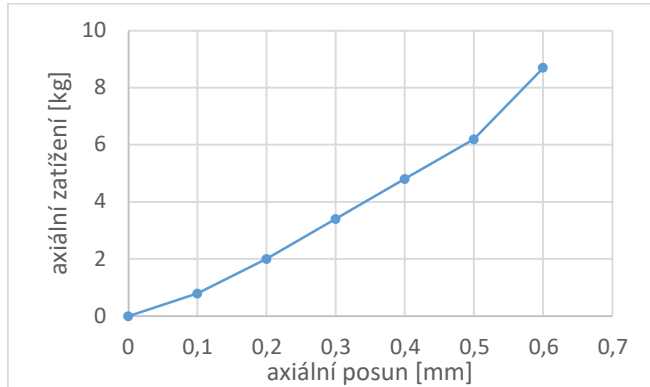
Obr. 5: Graf závislosti radiální výchylky hřídele na radiálním zatížení

B. Axiální měření

Vzhledem k relativně velikým axiálním silám v ložisku při axiální výchylce ložiskových kroužků bylo shledáno jako vhodné změřit i průběh síly v závislosti na této výchylce (obr. 6). Změřené výsledky mohou být potřebné při sestavování a úpravách ložiska a při dimenzování axiálního zajištění hřídele, zvláště pokud nejsme schopni zaručit přesnou axiální pozici ložiskových kroužků. Výsledný graf je na obr. 7. Kvůli mechanickým omezením dodatečného měřicího přípravku nebylo možno změřit sílu při větší axiální odchylce jak 0,6 mm, i přesto je zde však potvrzena nutnost přesného nastavení ložiskových kroužků, pokud nechceme příliš axiálně zatěžovat pomocné ložisko.



Obr. 6: Schéma axiálního měření

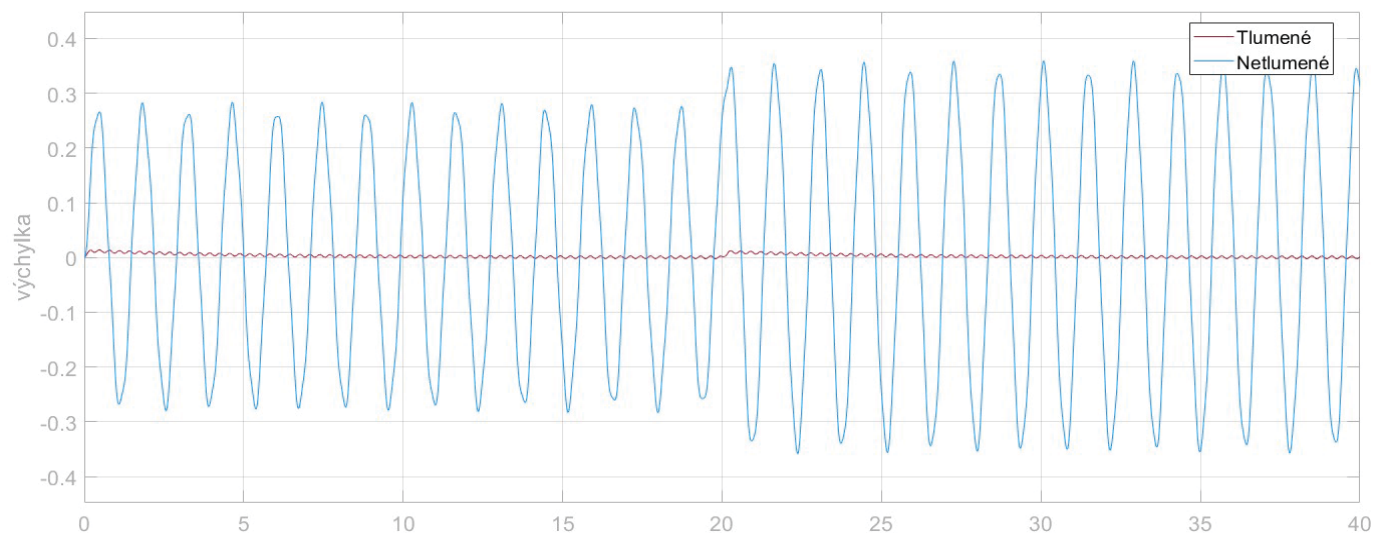


Obr. 7: Graf závislosti axiálního zatížení na axiálním posuvu

V. MOŽNOST TLUMENÍ VIBRACÍ S VYUŽITÍM DERIVAČNÍ SLOŽKY PID REGULÁTORU

Vzhledem k lineárnímu průběhu radiální výchylky při radiálně působící síle byla zvažována možnost tlumení vibrací pouze pomocí derivační složky PID regulátoru, což by umožňovalo relativně jednoduchou implementaci. Celý systém (pasivní ložisko + aktivní tlumení) představoval PD regulátor, kde derivační složku by zajišťovali elektromagnety aktivního tlumení a proporcionální složku by vzhledem k změřenému lineárnímu průběhu zajišťovalo pasivní magnetické ložisko. Předpokládá se ovšem, že nevádí trvalá regulační odchylka v případě působení statické síly (např. gravitační síla).

Vzhledem k podobnosti s lineární pružinou bylo tlumení vibrací otestováno na jednoduchém modelu lineární pružiny. Z výsledných grafů bylo zjištěno, že silový pulz toto tlumení dokáže rychle ztlumit bez jakýchkoli zbytkových kmitů, ale při testu kmitání hřídele při periodické harmonické působící síle (simulace nevyvážky na hřídeli) je v grafu vidět značné



Obr. 8: Graf reakce hřídele na silové pulzy (0 a 20 s) a periodickou sílu

ztlumení kmitů, ale nelze je tímto způsobem zcela odstranit. V grafu na obr. 8 je zobrazeno chování hřídele při kombinaci dvou pulzů (v 0. a 20. sekundě) a periodické harmonické síly. Při dalším testování bude třeba také vzít v úvahu možné nelinearity senzorů a elektromagnetů aktivního tlumení, které bude třeba kompenzovat.

VI. ZÁVĚR

V průběhu práce na projektu byly stanoveny požadavky na testovací přípravek s pasivním magnetickým ložiskem, podle kterých byl přípravek s ložiskem následně zkonstruován. Na ložisku proběhlo statické měření silových vlastností, které budou následně využity při návrhu aktivního tlumení. Předběžně byla zvažena možnost tlumení vibrací pomocí D složky PID regulátoru.

Pro vytvoření hybridního ložiska bude třeba splnit ještě tyto cíle:

- Výběr a dimenzování senzorů a aktivních prvků – elektromagnetů.
- Vytvoření matematického modelu včetně modelu aktivních prvků a senzorů.
- Výběr způsobu řízení aktivního tlumení.
- Vlastní realizace a test celého hybridního magnetického ložiska.

VII. ZDROJE

- [1] ŠAMÁNEK, O. Magnetická Ložiska. VUT v Brně, FSI. Copyright © [cit. 18.09.2020]. Dostupné z: http://dl.uk.fme.vutbr.cz/zobraz_soubor.php?id=531
- [2] MYSTKOWSKI Arkadiusz, AMBROZIAK Leszek, INVESTIGATION OF PASSIVE MAGNETIC BEARING WITH HALBACH-ARRAY, Bialystok University of Technology, Faculty of Mechanical Engineering Wiejska 45 C, 15-351 Bialystok.
- [3] Bezpečnostní tipy - Unimagnet.cz. Neodymové magnety - Unimagnet.cz [online]. Copyright © 2020 UNIMAGNET [cit. 21.09.2020]. Dostupné z: <https://www.unimagnet.cz/bezpecnost/>

Intelligent suspension system with two-degree-of-freedom hybrid mass dampers

Youssef Alsoufi

*Department of Instrumentation and Control Engineering
Faculty of Mechanical Engineering
Czech Technical University in Prague , Czech Republic
Youssef.Alsoufi@fs.cvut.cz*

Abstract—The presented model discusses the advantages and drawbacks of combining three concepts in vibration suspension. Dual-Loop Controller (DLC) with two-degree-of-freedom Hybrid Mass Damper (HMD) attached to a single degree of freedom primary system is studied. The benefits of using the adaptive actuators instead of classical ones in terms of reducing the amount of energy required in active part are mentioned in this design by adding feed-forward Deep Neural Network (DNN) to the primary system. The proposed model aims to increase the performance of damping, reduce energy consumption with achieving the fail-safe behavior of dampers.

Index Terms—Dynamic vibration absorber, active suspension system, hybrid vibration absorber, deep learning, deep neural network, dual loop controller.

I. INTRODUCTION

In general, problems caused by vibration are basically related to structure resonances or important harmonic loads. When the frequency of a harmonic load matches the resonance frequency of a structure, huge amplitude vibration will be created by dynamic amplification which linked to the internal damping of the structure. To minimize these vibrations, structural damping can be increased by adding the vibration absorbers. Passive and active vibration dampers are used widely to reduce the effects of structure oscillations. The Tuned Mass Damper (TMD) as a traditional Passive device is an auxiliary mass–spring–damper system which correctly tuned to a target mode, or a well-known harmonic perturbation, then the device is called Dynamic Vibration Absorber (DVA) [1]. The main benefits of the passive approach are the inherent stability of the overall set-up and no energy needs to be exerted to damp the oscillations. The design of a single-degree-of-freedom (SDOF) tuned-mass-damper (TMD) to attenuate the vibration of a single mode of the main structure under various conditions is studied e.g., [2]–[6]. Multiple SDOF TMDs with frequencies tuned in the neighborhood of a mode of the main structure is proposed in [7] to enhance the robustness and effectiveness of TMDs performance. Using a multiple SDOF TMDs to absorb more than one mode of main structure is studied in [8] and [9] by tuning each TMD to an individual mode of interest in the primary system. The problem with DVA is that in practice, the resonant frequency of the primary structure or the disturbance frequency may change during the time but the traditional DVA has no ability to automatically adjust its passive parameters or absorption

frequency .So, to overcome this problem, a lot of attempts were made to improve the vibration suppression performance of a TMDs and create an adaptive DVA, which has a self-adjusting ability on its absorption frequency. The adaptive DVA can adjust its absorption frequency by modifying its passive parameters. However, the effect of adaptive DVA can only perform well on vibration control with slow time-varying disturbances frequency. Adding an actuator controls the force that the reaction mass applies to the structure gives more efficient method to treat several resonances and improves the efficiency of the vibration absorber on controlling the structural vibration excited by forces of random disturbance. The resulted proposed device is called Active Mass Damper (AMD) which can be efficient on all the controllable modes and usually an absolute velocity (or acceleration) measurement is required as an input of AMD [10]. Active Mass Damper uses the active element to generate a force to reduce the effects of the exciting forces acting on a vibrating structure in order to minimize the vibration of the controlled structure [11]. However, the drawback of traditional AMD design is that all counter forces are created by the active element and therefore the power consumption and the size of the active actuator become very large in case dealing with large structural vibration [12]–[14]. Consequently, some suggestions are proposed. An electromagnetic actuator or piezoelectric actuator is integrated with a passive DVA as an active element to reduce the large counter-acting force requirement from the actuator [15], [16]. The absorber in this design uses the inertial force generated by the passive components as part of the counter-acting force to the excitations of the structure of the vibration. Therefore, with this feature, the control effort and the power requirement of the counterforce actuator can be reduced [14]. In the same context, the experimental results of some researches showed that the developed Neural Network (NN) as controller also gives a great performance and better results than the optimal controller under uncertainties in terms of reducing actuator energy consumption [17]. Attaching a lightweight mass damper to minimizing oscillating motions of structures like ships or buildings has been discussed in the late 1980s and early 1990s. In [18], different types of controller designs are reported. The authors in the research proposed a new concept of active vibration absorption, “Delayed Resonator” where a controlled time delay has been used in a feedback loop. In fact, the main

difference between this new generation damper and TMD is that this device is mainly developed for harmonic perturbation rejection and thus cannot be considered as a fail-safe, i.e. that the damper is not still efficient when the controller is turned off. Also, some attempts to enhance the suspension technology by transition traditional passive suspension systems to active suspension systems are reported in [19]. Combination of an actuator with spring and dampers is discussed to generate a force that controls suspensions' dynamics behavior in order to achieve a higher level of performance [20]. Improving the classical resonator setting with lumped delay introducing a distributed delay structure into the control law with modification of delayed resonator concept for vibration suppression using both delayed and non-delayed acceleration feedback control laws together taking into account both the stability and implementation constraints is discussed in [?]. A robust alternative of the delayed resonator using a spectral approach is proposed in [?]. Because of the nonlinearities and time-varying characteristics of the suspension system, it is difficult to design and tune the parameters of Active Suspension System (ASS) or to establish an appropriate dynamic model for model-based controller design, many researchers focused on ASS. Sliding mode is one of the common examples of control systems [21]. Also with advances in technology, more complex controllers are developed. Some researches discussed simple solutions to design controllers able to deal with uncertainties such as the use of a self-tuning pole assignment and fuzzy logic to adjust the gains of the controller [22]. Using linear-quadratic (LQ) control and pole placement techniques for a linearized model is developed in [23]; however, the results were not satisfying with respect to sprung acceleration suspension.

Recently, a new definition appeared to refer to the new generation of dampers which combining the behavior of an optimal TMD and the active damping devices. These dampers are named Hybrid Mass Damper (HMD), or Hybrid Vibration Absorber (HVA). The HMD over the last decade is used widely and many control laws have been proposed. The researches attempted to prove that using HMD will enhance the dynamic systems by increasing the performance, reducing the stroke of the moving mass, decreasing the energy consumption, reducing the embedded mass and to ensure fail-safe behavior. Optimal control is used to combine structural damping with a restricted stroke of the actuator to create a HMD in [24]. In summary, the hybrid controllers attempt to combine the features of optimal TMD with AMD with robust and simple control law, while maintaining the basic goals of increasing the performance and/or decreasing the control effort of the actuator.

More recently, Neural Network (NN) is used in vibration suspension to enhance the controller performance. In [25], the author proposed a novel neural network-based sliding mode control by combining the advantages of the adaptive, radial basis function neural network and sliding mode control strategies to decide the dynamical model requirement. So, by tuning the neural network weightings and/or radial basis function parameters, NN can learn online to deal with the system time-

varying and non-linear uncertainty behaviors. Development of a Recurrent Neural Network (RNN) is discussed in [26]. RNN here is based on inverse dynamics derived from the Bouc–Wen model of a semi-ASS with magnetic damper as well as an actual damper fitted in a hardware-in-the-loop simulation (HILS) showed that the RNN offered improved actuator voltage signal which ensures extended service life and lower power requirement. In [27], back propagation neural network (BPN) has been used to determine the gain parameters of a PID controller by using NN trained with a Levenberg–Marquardt algorithm and the outputs were compared with the results of other algorithms. In fact, almost all the previous NN which used in the control of ASSs were combined with sliding mode, LQR, PID, neuro-fuzzy or similar model to adjust their parameters either offline or in real-time. In [17], [26] RNN and DNN are reported where RNN is used with a semi-ASS utilizing a magnetic damper in [26], while the author substituted the conventional controller with a DNN controller in [17] which is the closest to our proposed design. In the present model, we will discuss using DNN controller added to Dual-Loop Controller (DLC) with Two-Degree-Of-Freedom hyperstable HMD to:

- reduce the actuator energy consumption and the sprung mass acceleration.
- provide stable and efficient control.
- increase the damping performance.
- achieve fail-safe behavior based on optimally tuned passive device.

II. REVIEW OF PASSIVE TUNED MASS DAMPER

The passive devices do not require an external power supply to control its operation and use the internal motion of the structure to enhance the absorption forces and dissipate the vibration energy of a specific resonance. For the ideal passive devices, the control forces applied to the structure are only dependent to the structural motion [28]. Vibration absorber systems such as Tuned Mass Damper (TMD) has been widely used for vibration control in mechanical systems. Basically, a TMD is a device consisting of a mass attached to a structure through a spring-dashpot system in parallel [29], see the Fig. 1 where a simple mechanical model for TMD is shown. In Fig. 1(a), Mass M represents the main structural mass and m represents the added mass of the TMD which is usually about 10% of the main structure mass and mounted on a spring and a damper. The equations for stiffness and damping can be found in [30]. The feature of this kind of dampers is that the same design rules can be used to damp nanometer vibrations or meter vibrations. The most popular is the so-called Den Hartog's method [1], [30]. The optimal design of a dynamic absorber can be classified into time-domain optimization and frequency domain optimization. In [31], four optimum design methods for a dynamic absorber are compared when they are applied to a single-degree-of-freedom system with primary damping. Furthermore, they can be used on another complex systems, including continuous systems. The undamped natural frequency of the TMD depends on the mass and stiffness of the

TMD and has to be tuned close to the natural frequency of the main structure, therefore, reducing the amplitude of the main structure is achieved at the resonance frequency by adjusting TMD parameters damping correctly [32]–[34]. In Fig. 1(b), TMD oscillates at the same frequency of the structure, but with a phase-shift. The challenge in designing TMD is hard tuning correctly the parameters of stiffness and damping of the absorber because of the large sensitivity of the suppression amplitude to parameter variations of stiffness and damping of TMD which consider a drawback in passive devices, in addition to another drawback in DVAs work which is that they are tuned to damp only one specific resonance. This problem was addressed by adding an actuator which controls the force that the reaction mass applies to the structure, we mention it in section III and section IV. In fact, overall performance of designed systems is affected by parameter variations of the dynamics and natural frequency variations of the main structure. Some methods designed to deal with this sensitivity. One of them is represented in [35]. TMD devices have proven as a useful concept in a lot of fields and were used in the widespread field of applications. For example, In buildings and structures, automotive industry, airplanes, space structures and so on. in Fig. 2(a) you can notice many types of passive devices with the vibration performance according to each dynamical components. One thing should be noticed that during the work of TMD, the main advantage is that these structures do not make reaction forces in the dynamical systems. A semi-active tuned mass damper (STMD) with variable damping under harmonic excitation is studied in [36], [37]. A study about the comparison between passive and active dampers is presented in [38]. In [39], the designed model is closed to our proposed model in this research where a complex TMD tuned optimally with two degrees of freedom (displacement and rotational movements) is achieved to increase the efficiency of the damper.

III. REVIEW OF SEMI-ACTIVE MASS DAMPERS

The controllers must be designed so as to achieve an acceptable trade-off between control effectiveness and energy consumption. From this point of view, the control strategies can be grouped into two main categories: active and semi-active. When the controller requires a small external power source for its operation by utilizing the motion of the structure to develop control force, it is called Semi-active controller. Where the magnitude of the force can be adjusted by an external power source [40]. Since its introduction [41] these systems somewhat trying to use the advantages of both passive and active devices by filling the gap between purely passive and purely active suspensions. However, the benefits of semi-active devices over active devices are their fewer power requirements. And moreover, that semi-active devices cannot inject the mechanical energy into the controlled structural system, but has properties that can be controlled to optimally reduce the response of the system. Therefore, in contrast to active control devices, semi-active control devices do not have the potential to destabilize the structural system [42].

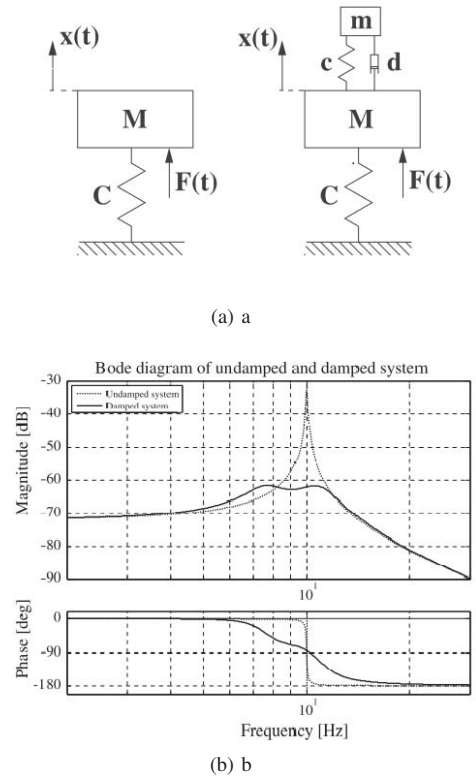


Fig. 1. (a) Main structural mass with TMD added to reduce the vibration amplitude with force $F(t)$ and positions measurements $X(t)$ (b) Bode diagrams of the undamped and the damped system.

The design of semi-active damper with its performance is shown in Fig. 2.b . In fact, the semi-active concept has been applied to many applications in vibration isolation problems. One of these applications used the concept to high-speed ground transportation vehicles [43]. In [44], a two-degree-of-freedom model of a semi-actively suspended vehicle is used as a starting point in the design of an optimal suspension. Using semi-active dampers for structural control application is proposed in [45]. Semi-active controllable fluid damper is studied in [46]. These devices have some special fluid, where applying external energy gives the ability to adjust their property. The electric and magnetic fields are mainly used to control these devices. Electro Rheological (ER) and Magneto-Rheological (MR) dampers are proposed in [40]. In [47], design and analysis of a control strategy, for semi-active suspensions in road vehicles, based on model-predictive control (MPC) strategies is studied.

IV. REVIEW OF ACTIVE MASS DAMPERS

The Active Mass Damper (AMD) is a mechanical device that attenuates the vibrations of a structural system. It uses reaction forces generated by moving the auxiliary mass with an actuator connected between the structure and the auxiliary mass. An AMD concept is shown in Fig. 2.c and the equations of motion can be found in [48]. In fact, many applications used the AMD concept in the last decades. For

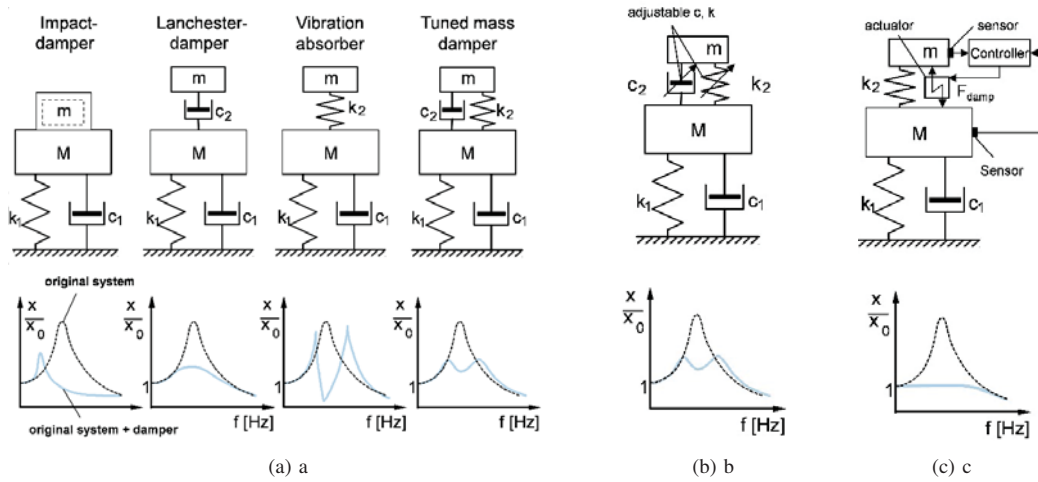


Fig. 2. (a) Passive Auxiliary System (b) Semi-active system (c) Active system

examples, in buildings, Moving a relatively small mass with a limited amplitude is required to suppress the vibrations of the structures. This is the reason why AMDs do not have a great ability to absorb the large earthquake excitations. However, the goal of absorption energy of structural vibrations under small earthquake excitations or under strong wind is achieved. Usually, The auxiliary mass of an AMD for a building structure is less than 0.4% of the total mass of the structure. In [49], some switching between passive and active vibration control methods in buildings are proposed. Where behaves as an AMD to suppress the vibrations caused by small earthquakes, and as a tuned mass damper to suppress the vibrations of a targeted mode excited by a big earthquake. The response motion of the auxiliary system is ordinarily increased as a result of active control. Hence, it is desired not only to increase the control efficiency but also to restrict the auxiliary mass motion as much as possible. Generally, the size of AMD or the installations space decides the allowable amplitude of the AMD with strong restrictions. These restrictions of the amplitude of the auxiliary mass are one of the main reasons for the limited performance of AMDs. In [24], a control law for AMD that effectively suppress the vibrations of a one-degree-of-freedom structural system under the amplitude constraint of the auxiliary mass is studied. Optimal control is used to combine structural damping with a restricted stroke of the actuator. Active Suspension Systems (ASS) also have used widely in the automotive industry [50], [51]. The research [52] discussed the advantages of ASS compared with Passive Suspension Systems (PSS) where the parameters of PSS are normally chosen based on the vehicle's design requirements and are fixed generally, therefore do not have the ability to improve both the passenger comfort and the vehicle handling at the same time. Force actuators, such as pneumatic, linear motor, hydraulic actuators, or so on can provide the required forces between the sprung mass and the main structure to control the attitude of the vehicles in [53]. Many attempts

and methods have been used and proposed to improve the performance of ASS from simple on-off control to highly advanced linear and non-linear control techniques, including linear- quadratic-Gaussian control [54], linear parameter-varying control (LPV) control [55], adaptive control [56], a multi-objective control method with wheelbase preview for active vehicle suspension. A four-degree-of-freedom half-car model with active suspension is studied [57], and quantitative feedback theory [58].

Recently with self-driving vehicles, the suspension of the vertical vibration amplitude of the cameras which mounted on the vehicles is needed to reduce the unwanted motion effects. In this context, a new assisted technologies started to emerge. A novel adaptive neural network based on sliding mode control strategy to stabilize the image captured area of the camera, therefore, suppress vertical displacement of sprung mass with the application of active suspension system is proposed in [59]. In [60], an adaptive fuzzy optimal control design is addressed for a class of unknown. The control objective is to design a controller not only to guarantee the stability of the systems but also to achieve the optimal control performance as well.

In fact, with the many uses of ASS in the automotive industry, a problem has appeared. It is the parametric uncertainties and changing the road disturbances during the time which in turn will adversely affect vibration mitigation. To solve this problem, many methods have been used. For instance, sliding mode control [61]. a velocity-dependent multi-objective control method to solve the problem of preview control with velocity uncertainty is presented in [62]. Neural network showed great potential for solving non-linear state observation problems when the system model has uncertain parameters. Wherein practical terms, the state variables are usually not available for direct online measurement so, an observer is needed. A stable neural network NN-based observer for general multi-variable non-linear systems is proposed in [63] which uses a back-propagation algorithm with a modification

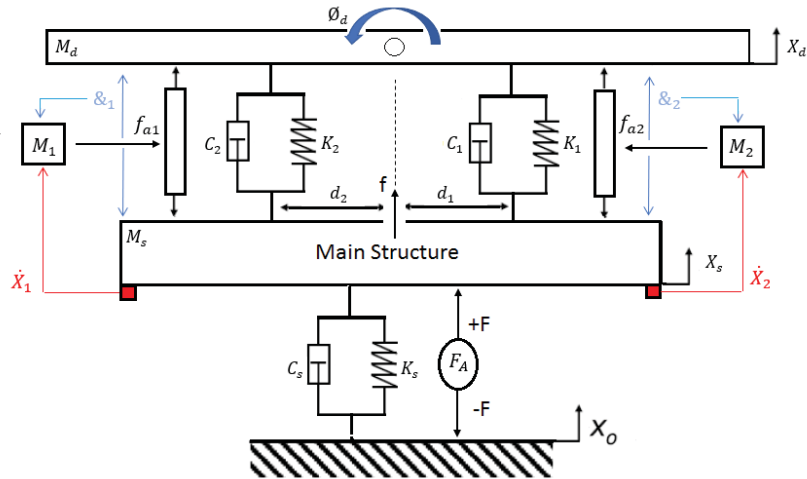


Fig. 3. intelligent suspension system with two degree of freedom hybrid mass damper

term. Adaptive neural network (NN) state feedback control and robust observation for an active suspension system that considers parametric uncertainties, road disturbances and actuator saturation is studied in [64]. Since an ASS actuator only provides a limited control force, input saturation with a non-linear factor may adversely affect system performance and can destabilize the system [65]–[67].

V. REVIEW OF HYBRID MASS DAMPER

After many studies in the last decades, a novel class of dampers has appeared that are trying to combine several objectives and features at the same time. These devices are gathered under the common name of Hybrid Mass Dampers (HMDs), or Hybrid Vibration Absorbers (HVAs) which combine passive and active vibration control. Combining passive and active elements the system is fail-safe that the damper will behave as a passive device even when the feedback control is turned off. The goal of using HMD may differ from one to other. For example, in [68], H_∞ optimal design of HMD is used for the minimization of the resonant vibration amplitude of a single-degree-of-freedom (sdf) vibrating structure.

In [69], [70], a pole placement technique is proposed to ensure performance and stability where the proposed hybrid absorber design could provide a simple alternative to adjust the conventional TMD as a higher performance HVA [69]. And a special pole placement controller is designed such that all vibration modes of the flexible structures become critically damped in [70]. In [71], a dual loop approach is preferred to increase the stability margins.

Improving the performance and stability of hybrid mass dampers by creating a hyperstable controller is studied in [72]–[74]. Where two zeros are added to interact with the poles of the system so as to reduce the resonant vibration amplitude of a multiple degree-of-freedom structure [72]. In [74], a compensator is introduced in the control loop to correct the phase of the actuator in order to be stable at low frequency. A

hybrid fuzzy logic approach which combines fuzzy logic and PID controllers is designed in the automotive industry [75]. In [76], a hybrid self-organizing fuzzy radial basis-function neural-network controller has been proposed by manipulating an ASS to reduce the power spectral density of the vehicle body acceleration. A fail-safe and unconditionally stable controller with a simple control law which is extremely efficient under harmonic excitation and required a low consumption are proposed in [72]. In [72], Small active forces are required when hypersatble controller is applied to three degrees of freedom.

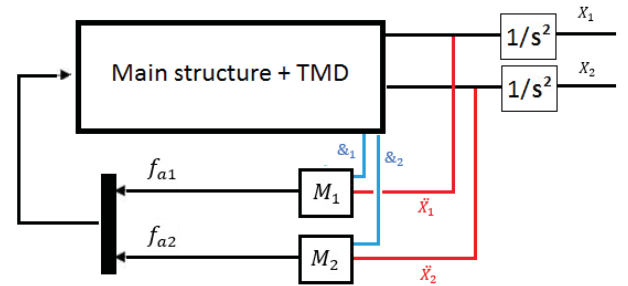


Fig. 4. Controller loops.

VI. PROPOSED MODEL

More than one mode of vibration of an absorber body relative to a primary system be tuned to suppress single-mode vibration of a primary system with developing an algorithm to enhance the connection of multi-degree-of-freedom between the absorber body and primary structure, therefore, mitigation the response to random and harmonic excitations is proposed in the present model. The model simulates two-degree of freedom HMD attached to a single degree of freedom primary system Fig. 3. Vertical displacement movement and

rotational movement around the center of the auxiliary mass are proposed where the results proved that an optimal two-degree of freedom TMD performs better than a traditional single degree of freedom or two separate TMDs with optimal mass distribution, even for the case where the rotary inertia of the absorber tends to zero. [39]. We will use this prove and continue to convert the TMDs to HMDs.

When we use a HMD to attenuate the vibration; the passive part has a positive impact where mitigate the amplitudes of the vibrations i.e. damping the oscillations is occurring without the need for any external power. On the other hand, the frequency band where the absorber suppresses the vibrations efficiently is relatively narrow, being centered at the natural frequency of the absorber. Practically, the damper cannot absorb the vibration entirely even if the vibration frequency is identical with the natural frequency of the absorber because of fact that the physical absorber is never ideal, i.e. it features nonzero damping. Solving this problem is presented in section II. The proposed model tries to achieve three main objectives: (1) increase the performance of the suspension system, (2) reduce energy consumption, and (3) ensure fail-safe behavior. We consider passive absorbers are supplemented by dual transducer loops in the model Fig.3. The passive parts are a TMDs optimally tuned using Den Hartog's law [1], and the active control forces (f_{a1}) and (f_{a2}) are introduced between the two masses. The auxiliary mass is the moving part of the actuator. In fact, the general concept of dual-loop controller is started in [10] and is used precisely as parallel dual loops and act on the same transducer in [77] which is the closest to our model with the difference that in the previous research a dual loop controller (DLC) is used with single degree of freedom. The proposed controller combines two control laws using two inputs (i) the relative displacement between the inertial mass and the main structure, and (ii) the absolute velocity of the main structure. Subsequently, dual loops in each side of the damper are used. One to detune the HMD (negative stiffness feedback), and the second one to damp the main structure (direct velocity feedback). Since the relative velocity or absolute acceleration will be measured, so integration has to be added to the control laws M1 and M2 Fig 4. The proposed controller will be designed to increase the margins and reduce the stiffness of the TMD by develop the controller to hyperstable controller because it is known that direct velocity feedback associated with a TMD results in very poor stability margins [72]–[74]. The rotational movement of the damper gives ability to absorb not only vertical disturbances but also disturbances with angles. Also, as long as we use an active impact in our design, we have to keep in mind the necessity of restriction the amount of energy required for absorption operation. Where as we mentioned in section I, one of the drawback of AMD design that all counter forces are created by the active element, therefore, required forces of the actuator become very large in a case dealing with large structural vibration. In this context, we proposed using feed-forward DNN as a controller to decrease the actuator energy consumption. The network will be trained through supervised

learning using the back-propagation algorithm. We relied on this type of network because experimental results showed that the developed DNN controller outperforms the optimal controller under uncertainties in terms of reducing the sprung mass acceleration and actuator energy consumption, with a 4% and 14% reduction, respectively [17].

VII. CONCLUSION

Combining many concepts in vibration absorption will be achieved in this study to attain an optimal absorber over a relatively wide frequency band. A fail-safe behaviour will be guaranteed by taking advantage of the features of combining between active and passive elements in one absorber which called a hybrid mass damper. We cast the problem of optimization of the multi-degree-of-freedom connection between the absorber body and primary structure. Developing optimization algorithms based on the H2 and H-infinity norms will be discussed to reduce the response to random and harmonic excitations with attaining better performance in vibration by using an optimal 2DOF HMD instead of traditional SDOF HMD. Finally, actuator energy consumption was taken into consideration by using a back-propagation algorithm with deep neural network DNN to reduce the sprung mass acceleration.

ACKNOWLEDGEMENT

The author has been supported by the Grant Agency of the Czech Technical University in Prague, grant No. SGS20/159/OHK2/3T/12.

TABLE I
NOMENCLATURE

Abbreviation	Definition
AMD	Active Mass Damper
ASS	Active Suspension System
BPN	Back-Propagation Neural Network
DLC	Dual-Loop-Controller
DNN	Deep Neural Network
DVA	Dynamic Vibration Absorber
HILS	Hardware-In-the-Loop-Simulation
HMD	Hybrid Mass Damper
HVA	Hybrid Vibration Absorber
NN	Neural Network
PSS	Passive Suspension System
RNN	Recurrent Neural Network
STMD	Semi active Tuned Mass Damper
SDOF	Single-Degree-Of-Freedom
TMD	Tuned Mass Damper

REFERENCES

- [1] DenHartog, J., "Mechanical vibrations, (1956), 373."
- [2] Den Hartog, J. and Vibrations, M., "Mcgraw-hill book co," *New York*, 1947.
- [3] Brock, J. E., "A note on the damped vibration absorber," *Trans. ASME, J. Appl. Mech.*, vol. 13, no. 4, pp. A-284, 1946.
- [4] Snowdon, J. C., *Vibration and shock in damped mechanical systems*. J. Wiley, 1968.
- [5] Warburton, G., "Optimum absorber parameters for various combinations of response and excitation parameters," *Earthquake Engineering & Structural Dynamics*, vol. 10, no. 3, pp. 381–401, 1982.
- [6] Asami, T., Nishihara, O., and Baz, A. M., "Analytical solutions to h_{∞} and h_2 optimization of dynamic vibration absorbers attached to damped linear systems," *J. Vib. Acoust.*, vol. 124, no. 2, pp. 284–295, 2002.

- [7] Xu, K. and Igusa, T., "Dynamic characteristics of multiple substructures with closely spaced frequencies," *Earthquake engineering & structural dynamics*, vol. 21, no. 12, pp. 1059–1070, 1992.
- [8] Kitis, L., Wang, B., and Pilkey, W., "Vibration reduction over a frequency range," *Journal of Sound and Vibration*, vol. 89, no. 4, pp. 559–569, 1983.
- [9] Rice, H., "Design of multiple vibration absorber systems using model data," *Journal of sound and Vibration*, vol. 160, no. 2, pp. 378–385, 1993.
- [10] Preumont, A. and Seto, K., *Active control of structures*. Wiley Online Library, 2008.
- [11] Williams, K., Chiu, G., and Bernhard, R., "Adaptive-passive absorbers using shape-memory alloys," *Journal of Sound and Vibration*, vol. 249, no. 5, pp. 835–848, 2002.
- [12] Vonflotow, A. H., Beard, A., and Bailey, D., "Adaptive tuned vibration absorbers: tuning laws, tracking agility, sizing, and physical implementations," *nce*, pp. 437–454, 1994.
- [13] Franchek, M., Ryan, M., and Bernhard, R., "Adaptive passive vibration control," *Journal of Sound and Vibration*, vol. 189, no. 5, pp. 565–585, 1996.
- [14] Wu, S.-T., Chiu, Y.-Y., and Yeh, Y.-C., "Hybrid vibration absorber with virtual passive devices," *Journal of sound and vibration*, vol. 299, no. 1-2, pp. 247–260, 2007.
- [15] Ikai, S., Ohsawa, K., Nagaya, K., and Kashimoto, H., "Electromagnetic actuator and stacked piezoelectric sensors for controlling vibrations of a motor on a flexible structure," *Journal of sound and vibration*, vol. 231, no. 2, pp. 393–409, 2000.
- [16] Morgan, R. and Wang, K., "Active-passive piezoelectric absorbers for systems under multiple non-stationary harmonic excitations," *Journal of sound and vibration*, vol. 255, no. 4, pp. 685–700, 2002.
- [17] Konoiko, A., Kadhem, A., Saiful, I., Ghorbanian, N., Zweiri, Y., and Sahinkaya, M. N., "Deep learning framework for controlling an active suspension system," *JVC/Journal of Vibration and Control*, vol. 25, no. 17, pp. 2316–2329, 2019.
- [18] Olgac, N. and Holm-Hansen, B., "A novel active vibration absorption technique: delayed resonator," *Journal of Sound and Vibration*, vol. 176, no. 1, pp. 93–104, 1994.
- [19] Alexandru, C. and Alexandru, P., "A comparative analysis between the vehicles' passive and active suspensions," *International Journal of Mechanics*, vol. 5, no. 4, pp. 371–378, 2011.
- [20] Fischer, D. and Isermann, R., "Mechatronic semi-active and active vehicle suspensions," *Control Engineering Practice*, vol. 12, no. 11, pp. 1353–1367, 2004.
- [21] Kim, C. and Ro, P., "A sliding mode controller for vehicle active suspension systems with non-linearities," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 212, no. 2, pp. 79–92, 1998.
- [22] Ramsbottom, M., Crolla, D., and Plummer, A. R., "Robust adaptive control of an active vehicle suspension system," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 213, no. 1, pp. 1–17, 1999.
- [23] Watton, J., Holford, K. M., and Surawattanawan, P., "The application of a programmable servo controller to state control of an electrohydraulic active suspension," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 218, no. 12, pp. 1367–1377, 2004.
- [24] Yoshida, K. and Matsumoto, I., "Vibration suppression control for a multi-degree-of-freedom structural system using an amd with restricted stroke," in *2009 International Conference on Networking, Sensing and Control*. IEEE, 2009, pp. 912–917.
- [25] Huang, S. and Lin, W., "A neural network based sliding mode controller for active vehicle suspension," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 221, no. 11, pp. 1381–1397, 2007.
- [26] Metered, H., Bonello, P., and Oyadiji, S., "An investigation into the use of neural networks for the semi-active control of a magnetorheologically damped vehicle suspension," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 224, no. 7, pp. 829–848, 2010.
- [27] Heidari, M. and Homaei, H., "Design a pid controller for suspension system by back propagation neural network," *Journal of Engineering*, vol. 2013, 2013.
- [28] Yi, F. and Dyke, S. J., "Structural control systems: performance assessment," in *Proceedings of the 2000 American Control Conference. ACC (IEEE Cat. No. 00CH36334)*, vol. 1, no. 6. IEEE, 2000, pp. 14–18.
- [29] Acar, M. and Yilmaz, C., "Design of an adaptive-passive dynamic vibration absorber composed of a string-mass system equipped with negative stiffness tension adjusting mechanism," *Journal of Sound and Vibration*, vol. 332, no. 2, pp. 231–245, 2013.
- [30] Den Hartog, J. P., *Mechanical vibrations*. Courier Corporation, 1985.
- [31] Wang, Y. and Cheng, S., "The optimal design of dynamic absorber in the time domain and the frequency domain," *Applied Acoustics*, vol. 28, no. 1, pp. 67–78, 1989.
- [32] Rana, R. and Soong, T., "Parametric study and simplified design of tuned mass dampers," *Engineering structures*, vol. 20, no. 3, pp. 193–204, 1998.
- [33] Ozer, M. B. and Royston, T. J., "Extending den hartog's vibration absorber technique to multi-degree-of-freedom systems," 2005.
- [34] Lin, G.-L., Lin, C.-C., Chen, B.-C., and Soong, T.-T., "Vibration control performance of tuned mass dampers with resettable variable stiffness," *Engineering Structures*, vol. 83, pp. 187–197, 2015.
- [35] Zhang, H., Shi, Y., and Mehr, A. S., "Robust non-fragile dynamic vibration absorbers with uncertain factors," *Journal of Sound and Vibration*, vol. 330, no. 4, pp. 559–566, 2011.
- [36] Pinkaew, T. and Fujino, Y., "Effectiveness of semi-active tuned mass dampers under harmonic excitation," *Engineering Structures*, vol. 23, no. 7, pp. 850–856, 2001.
- [37] Lallart, M., Yan, L., Wu, Y.-C., and Guyomar, D., "Electromechanical semi-passive nonlinear tuned mass damper for efficient vibration damping," *Journal of Sound and Vibration*, vol. 332, no. 22, pp. 5696–5709, 2013.
- [38] Kim, S.-M., Wang, S., and Brennan, M. J., "Dynamic analysis and optimal design of a passive and an active piezo-electrical dynamic vibration absorber," *Journal of sound and vibration*, vol. 330, no. 4, pp. 603–614, 2011.
- [39] Zuo, L. and Nayfeh, S. A., "The two-degree-of-freedom tuned-mass damper for suppression of single-mode vibration under random and harmonic excitation," 2006.
- [40] Symans, M. D. and Constantinou, M. C., "Semi-active control systems for seismic protection of structures: a state-of-the-art review," *Engineering structures*, vol. 21, no. 6, pp. 469–487, 1999.
- [41] Karnopp, D., Crosby, M. J., and Harwood, R., "Vibration control using semi-active force generators," 1974.
- [42] Forrai, A., Hashimoto, S., Funato, H., and Kamiyama, K., "Structural control technology: system identification and control of flexible structures," *Computing & Control Engineering Journal*, vol. 12, no. 6, pp. 257–262, 2001.
- [43] Margolis, D., Tylee, J., and Hrovat, D., "Heave mode dynamics of a tracked air cushion vehicle with semiactive airbag secondary suspension," 1975.
- [44] Hrovat, D., Margolis, D., and Hubbard, M., "An approach toward the optimal semi-active suspension," 1988.
- [45] Hrovat, D., Barak, P., and Rabins, M., "Semi-active versus passive or active tuned mass dampers for structural control," *Journal of Engineering Mechanics*, vol. 109, no. 3, pp. 691–705, 1983.
- [46] Soong, T. and Spencer Jr, B., "Supplemental energy dissipation: state-of-the-art and state-of-the-practice," *Engineering structures*, vol. 24, no. 3, pp. 243–259, 2002.
- [47] Canale, M., Milanese, M., and Novara, C., "Semi-active suspension control using "fast" model-predictive techniques," *IEEE Transactions on control systems technology*, vol. 14, no. 6, pp. 1034–1046, 2006.
- [48] Nishimura, I., Kobori, T., Sakamoto, M., Koshika, N., Sasaki, K., and Ohru, S., "Active tuned mass damper," *Smart Materials and Structures*, vol. 1, no. 4, p. 306, 1992.
- [49] Abe, N., "Passive and active switching vibration control with pendulum type damper," in *Proceedings of the 2004 IEEE International Conference on Control Applications, 2004.*, vol. 2. IEEE, 2004, pp. 1037–1042.
- [50] Cao, D., Song, X., and Ahmadian, M., "Editors' perspectives: road vehicle suspension design, dynamics, and control," *Vehicle system dynamics*, vol. 49, no. 1-2, pp. 3–28, 2011.
- [51] Appleyard, M. and Wellstead, P., "Active suspensions: some background," *IEE Proceedings-Control Theory and Applications*, vol. 142, no. 2, pp. 123–128, 1995.
- [52] Van der Sande, T., Gysen, B., Besselink, I., Paulides, J., Lomonova, E., and Nijmeijer, H., "Robust control of an electromagnetic active suspen-

- sion system: Simulations and measurements,” *Mechatronics*, vol. 23, no. 2, pp. 204–212, 2013.
- [53] Cao, J., Liu, H., Li, P., and Brown, D. J., “State of the art in vehicle active suspension adaptive control systems based on intelligent methodologies,” *IEEE transactions on intelligent transportation systems*, vol. 9, no. 3, pp. 392–405, 2008.
- [54] Chen, S., He, R., Liu, H., and Yao, M., “Probe into necessity of active suspension based on lqg control,” *Physics Procedia*, vol. 25, pp. 932–938, 2012.
- [55] Wang, X., Li, Y. W., Blaabjerg, F., and Loh, P. C., “Virtual-impedance-based control for voltage-source and current-source converters,” *IEEE Transactions on Power Electronics*, vol. 30, no. 12, pp. 7019–7037, 2014.
- [56] Alleyne, A. and Hedrick, J. K., “Nonlinear adaptive control of active suspensions,” *IEEE transactions on control systems technology*, vol. 3, no. 1, pp. 94–101, 1995.
- [57] Li, P., Lam, J., and Cheung, K. C., “Multi-objective control for active vehicle suspension with wheelbase preview,” *Journal of Sound and Vibration*, vol. 333, no. 21, pp. 5269–5282, 2014.
- [58] Niksefat, N. and Sepehri, N., “Design and experimental evaluation of a robust force controller for an electro-hydraulic actuator via quantitative feedback theory,” *Control Engineering Practice*, vol. 8, no. 12, pp. 1335–1345, 2000.
- [59] Zhao, F., Dong, M., Qin, Y., Gu, L., and Guan, J., “Adaptive neural-sliding mode control of active suspension system for camera stabilization,” *Shock and Vibration*, vol. 2015, 2015.
- [60] Liu, Y.-J., Gao, Y., Tong, S., and Li, Y., “Fuzzy approximation-based adaptive backstepping optimal control for a class of nonlinear discrete-time systems with dead-zone,” *IEEE Transactions on Fuzzy Systems*, vol. 24, no. 1, pp. 16–28, 2015.
- [61] Deshpande, V. S., Mohan, B., Shendge, P., and Phadke, S., “Disturbance observer based sliding mode control of active suspension systems,” *Journal of Sound and Vibration*, vol. 333, no. 11, pp. 2281–2296, 2014.
- [62] Li, P., Lam, J., and Cheung, K. C., “Velocity-dependent multi-objective control of vehicle suspension with preview measurements,” *Mechatronics*, vol. 24, no. 5, pp. 464–475, 2014.
- [63] Abdollahi, F., Talebi, H. A., and Patel, R. V., “A stable neural network-based observer with application to flexible-joint manipulators,” *IEEE Transactions on Neural Networks*, vol. 17, no. 1, pp. 118–129, 2006.
- [64] Zhao, F., Ge, S. S., Tu, F., Qin, Y., and Dong, M., “Adaptive neural network control for active suspension system with actuator saturation,” *IET control theory & applications*, vol. 10, no. 14, pp. 1696–1705, 2016.
- [65] Gayaka, S., Lu, L., and Yao, B., “Global stabilization of a chain of integrators with input saturation and disturbances: A new approach,” *Automatica*, vol. 48, no. 7, pp. 1389–1396, 2012.
- [66] Zhou, B., Lam, J., and Duan, G.-R., “An are approach to semi-global stabilization of discrete-time descriptor linear systems with input saturation,” *Systems & Control Letters*, vol. 58, no. 8, pp. 609–616, 2009.
- [67] Chen, M., Ren, B., Wu, Q., and Jiang, C., “Anti-disturbance control of hypersonic flight vehicles with input saturation using disturbance observer,” *Science China Information Sciences*, vol. 58, no. 7, pp. 1–12, 2015.
- [68] Cheung, Y., Wong, W., and Cheng, L., “Design optimization of a damped hybrid vibration absorber,” *Journal of sound and vibration*, vol. 331, no. 4, pp. 750–766, 2012.
- [69] Tso, M., Yuan, J., and Wong, W., “Design and experimental study of a hybrid vibration absorber for global vibration control,” *Engineering structures*, vol. 56, pp. 1058–1069, 2013.
- [70] —, “Suppression of random vibration in flexible structures using a hybrid vibration absorber,” *Journal of sound and vibration*, vol. 331, no. 5, pp. 974–986, 2012.
- [71] Hagedorn, P. and Spelsberg-Korspeter, G., *Active and passive vibration control of structures*. Springer, 2014.
- [72] Collette, C. and Chesne, S., “Robust hybrid mass damper,” *Journal of sound and vibration*, vol. 375, pp. 19–27, 2016.
- [73] Chesné, S. and Collette, C., “Experimental validation of fail-safe hybrid mass damper,” *Journal of Vibration and Control*, vol. 24, no. 19, pp. 4395–4406, 2018.
- [74] —, “Phase compensator for hyperstable hybrid mass,” 2017.
- [75] Demir, O., Keskin, I., and Cetin, S., “Modeling and control of a nonlinear half-vehicle suspension system: a hybrid fuzzy logic approach,” *Nonlinear Dynamics*, vol. 67, no. 3, pp. 2139–2151, 2012.
- [76] Lin, J. and Lian, R.-J., “Intelligent control of active suspension systems,” *IEEE Transactions on industrial electronics*, vol. 58, no. 2, pp. 618–628, 2010.
- [77] Chesne, S., Inquieté, G., Cranga, P., Legrand, F., and Petitjean, B., “Innovative hybrid mass damper for dual-loop controller,” *Mechanical Systems and Signal Processing*, vol. 115, pp. 514–523, 2019.



Editor: Ing. Vladimír Hlaváč, Ph.D.
Název díla: Nové metody a postupy v oblasti přístrojové techniky,
automatického řízení a informatiky 2020
Vydalo: České vysoké učení technické v Praze
Zpracoval: Ústav přístrojové a řídicí techniky Fakulty strojní ČVUT v Praze
Kontaktní adresa: Ústav přístrojové a řídicí techniky, FS ČVUT v Praze, Technická 4, Praha 6
Tel.: +420 2 2435 9750
Počet stran: 144 Pořadí vydání: 1
Vydáno dne 7. 10. 2020 jako pdf soubor na CD.
Dostupné na webových stránkách Ústavu přístrojové a řídicí techniky: <http://iat.fs.cvut.cz/nmp>
ISBN: 978-80-01-06776-5