



Research in Cybernetics at FEI STU in Bratislava

reviewed slovak professional magazine for scientific and engineering issues

Výskum v kybernetike na FEI STU v Bratislave

recenzované periodikum vedeckých a inžinierskych publikácií

Research in Cybernetics at FEI STU in Bratislava

Odborný garant

prof. Ing. Ján Murgaš, PhD. Slovenská technická univerzita v Bratislave Fakulta elektrotechniky a informatiky Ústav robotiky a kybernetiky Oddelenie kybernetiky

Technical guarantee

prof. Ing. Ján Murgaš, PhD. Slovak University of Technology in Bratislava Faculty of Electrical Engineering and Information Technology Institute of Robotics and Cybernetics Department of Cybernetics

Vydavateľ

HMH s.r.o. Tavarikova osada 39 841 02 Bratislava 42 IČO: 31356273

Spoluzakladateľ

Katedra ASR, EF STU Katedra automatizácie a regulácie, EF STU Katedra automatizácie, ChtF STU PPA CONTROLL, a.s. **Co-founder**

Publisher



www.atpjournal.sk

ATP Journal PLUS 2/2019

Redakčná rada

Draft committee

prof. Ing. Alexík Mikuláš, PhD., FRI UNIZA, Žilina Ing. Balogh Richard, PhD., FEI STU, Bratislava prof. Ing. Belavý Cyril, CSc., SjF STU, Bratislava prof. Ing. Duchoň František, PhD., FEI STU – NCR, Bratislava prof. Ing. Fikar Miroslav, DrSc., FCHPT STU, Bratislava prof. Ing. Hulkó Gabriel, DrSc., SjF STU, Bratislava prof. Ing. Janíček František, PhD., FEI STU, Bratislava prof. Ing. Krokavec Dušan, CSc., FEI TU Košice doc. Ing. Kvasnica Michal, PhD., FCHPT STU, Bratislava prof. Ing. Malindžák Dušan, CSc., BERG TU, Košice prof. Ing. Mészáros Alajos, CSc., FCHPT STU, Bratislava prof. Ing. Murgaš Ján, PhD., FEI STU, Bratislava prof. Ing. Rástočný Karol, PhD., FEIT UNIZA, Žilina doc. Ing. Schreiber Peter, CSc., MTF STU, Trnava prof. Ing. Smieško Viktor, PhD., FEI STU, Bratislava prof. Ing. Taufer Ivan, DrSc., FEI Univerzita Pardubice prof. Ing. Veselý Vojtech, DrSc., FEI STU, Bratislava prof. Ing. Zolotová Iveta, CSc., FEI TU, Košice prof. Ing. Žalman Milan, PhD., FEI STU, Bratislava doc. Ing. Ždánsky Juraj, PhD., FEIT UNIZA, Žilina

Babic Branislav,

výkonný riaditeľ ProCS, s.r.o.

Ing. Horváth Tomáš, riaditeľ HMH, s.r.o.

Ing. Hrica Marián, riaditeľ divízie A & D, Siemens, s.r.o.

Kroupa Jiří,

riaditeľ kancelárie pre SK, DEHN+SÖHNE Ing. Lásik Vladimír,

PPA CONTROLL, a.s.

Ing. Mašláni Marek,

riaditeľ B+R automatizace, s.r.o. – o. z.

Mík Pavel,

obchodný riaditeľ ABB, s.r.o.

Ing. Petergáč Štefan,

predseda predstavenstva Datalan, a.s.

Ing. Széplaky Ladislav,

riaditeľ Emerson Process Management, s.r.o.

Redakcia

Editors office

ATP Journal Galvaniho 7/D 821 04 Bratislava tel.: +421 2 32 332 182 fax: +421 2 32 332 109 vydavatelstvo@hmh.sk www.atpjournal.sk

Ing. Anton Gérer, šéfredaktor – editor in chief gerer@hmh.sk

Zuzana Pettingerová, DTP grafik – DTP graphic designer dtp@hmh.sk

Dagmar Votavová, obchod a marketing – sales and Marketing mediamarketing@hmh.sk

Mgr. Bronislava Chocholová jazyková redaktorka – text corrector

> Datum vydania: november 2019 Autorské práva HMH, s.r.o., Všetky práva vyhradené Copyright HMH, Ltd., 2019, All Rights Reserved ISSN 1336-5010 ATP Journal PLUS 2/2019

Obsah

60 Rokov kybernetiky na FEI STU v Bratislave
Modelovanie a riadenie procesov
Systémovo založená analýza podmienok možnej náhrady in-vivo štúdií modelovaním in-silico
Identifikácia nelineárnych kaskádových systémov so statickými vstupnými a dynamickými výstupnými nelinearitami s pásmami necitlivosti
Kybernetická podpora pri kompenzácii diabetu
Kybernetické modelovanie v energetike
Biokybernetika v diabetickej edukácii
Robotika
Systém lokalizácie a mapovania pre drony
Inteligentný menič so vstavanými PLC pre BVDC motory
SLAM na platforme mobilného robota
Umelá inteligencia
Rozpoznávanie gest ruky pomocou konvolučných neurónových sietí
Neuro-evolúcia riadenia dynamických systémov
Možnosti riadenia inteligentných budov
Detekcia rovinných útvarov v obraze

Content

60 Years of Cybernetics at FEI STU in Bratislava
Process control and modeling
System – based analysis of conditions for using in-silico modelling as surrogate for in-vivo studies
Identification on nonlinear cascade systems with static input and dynamic output nonlinearities including dead-zones
Rehabilitation equipment for diabetes patienst
Computing moddeling in power systems
Biocybernetics in diabetes education
Robotics
Localization and mapping system for drones
Intelligent drive with embedded PLC for BLDC motors
SLAM on Ackernann mobile robot platform
Atificial Intelligence
Hand gesture recognition using convolutional neural networks

Slavomír Kajan, Jozef Goga

 Neuro-evolution of dynamic system control
 56

 Ivan Sekaj, Dominika Banásová, Martin Komák, Milan Hvozdík
 62

 Control options for inteligent buildings
 62

 Juraj Slačka, Marek Budaj
 65

 Juraj Slačka, Filip Zúbek
 65

60 rokov kybernetiky na FEI STU v Bratislave

Ján Murgaš

V roku 2019 si pripomíname 60. výročie začatia výučby v odbore Kybernetika na Elektrotechnickej fakulte SVŠT, ktorej pokračovateľom je Fakulta elektrotechniky a informatiky STU v Bratislave. Hneď na úvod treba podotknúť, že kybernetika bola len rok pred tým oficiálne zaradená medzi svetové vedy. Naši predchodcovia veľmi rýchlo zareagovali na vznik novej vedy aj napriek nepriaznivým podmienkam v bývalej republike. Žiaľ, politickým vplyvom sa kybernetika nevyhla ani za oceánom, kde sa v tejto oblasti začali používať iné pojmy.

Pripomeňme si len základné prínosy kybernetiky:

- zaviedla k dovtedy uznávaným pojmom "hmota" a "energia" nový pojem "informácia"
- zaviedla modelovanie procesov a nový pohľad na dynamiku systémov
- zaviedla automatické riadenie procesov

Tak, ako sa kybernetika najskôr rozvíjala v oblasti riadenia procesov, kde bola najväčšia požiadavka spoločnosti, aj na našej fakulte boli študijné odbory, resp. programy, pomenované v tomto zmysle, napr. Automatizácia, Technická kybernetika a pod. Podobne aj výskum na katedrách, ktoré boli predchodcami Ústavu robotiky a kybernetiky, sa orientoval jednak na spojité a tiež na diskrétne procesy. Hlavné tézy výskumnej činnosti boli orientované na výskum modelovania a riadenia procesov a systémov a tiež na metódy a algoritmy riadenia. V aplikačnej oblasti v spojitých procesoch to bolo modelovanie a riadenie elektrizačných sústav a v diskrétnej oblasti modelovanie a riadenie servosystémov. Tento smer výskumu bol dobrým základom pre vznik nového smeru vo výskume robotika. V spojitých procesoch popri tradičných smeroch sa začala rozvíjať biokybernetika (niekde lekárska kybernetika) a bola podporená významným projektom (pozri projekty). Veľmi významný je aj rozvoj výskumu v oblasti umelej inteligencie.

60. výročie vzniku kybernetiky si pripomíname aj určitým pohľadom na súčasný výskum na Ústave robotiky a kybernetiky FEI STU v Bratislave.

V súčasnosti výskum v oblasti kybernetiky nadväzuje na oblasti rozvíjané v minulosti, ako je teória automatického riadenia, kde najvýznamnejšie výsledky boli dosiahnuté v oblasti riadenia nelineárnych systémov a v oblasti riadenia zložitých systémov. Aplikačná oblasť pri riadení systémov bola najmä trieda zložitých elektrizačných systémov. V posledných rokoch bol začatý výskum v oblasti biokybernetiky (lekárskej kybernetiky), konkrétne sa jedná o závažné civilizačné ochorenia, ako je diabetes mellitus a hypertenzia. Výskum v oblasti robotiky sa aktuálne zameriava na inovačné potreby priemyslu, konkrétne silovo poddajné roboty, mobilné manipulátory a drony. S tým úzko súvisia riešené témy spoľahlivého určenia polohy robota v prostredí, navigácia robotov v 3D reprezentácii prostredia, optimalizácia údajových reprezentácií o prostredí - tzv. mapy prostredia, fúzia informácií z viacerých snímačov, interakcia robotčlovek a pod. Cieľom je dosiahnuť čo najväčšiu autonómnosť robota nielen z pohľadu riadenia jeho pohybu, ale aj z pohľadu komunikácie s človekom vo forme prirodzenej pre ľudí, tzv. human-oriented collaboration.

Výskum v oblasti umelej inteligencie sa prioritne venuje dvom smerom. Prvý je orientovaný na výskum bio-inšpirovaných optimalizačných metód, predovšetkým evolučných a ich algoritmov implementáciám v riadení procesov a v robotike. Aktuálny výskum sa sústreďuje na metódy neuroevolúcie a jej implementácie v riadení procesov. Druhý smer nášho výskumu je orientovaný na metódy strojového videnia a rozpoznávania obrazu na báze hlbokých neurónových sietí. Medzi hlavné aplikačné domény tu patrí lekárska diagnostika a komunikácia človeka s počítačom statickými a dynamickými gestami a podobné aplikácie.

Výsledky ústavu vo výskume boli publikované v mnohých významných publikáciách. Vzhľadom na ich veľké množstvo uvádzame len výber publikácií z vedeckých časopisov prvého kvartilu (Q1) za posledných šesť rokov. Pracovníci ústavu prezentovali svoje výsledky na mnohých vedeckých konferenciách, vrátane svetových kongresov IFAC. Bolo vydaných viac ako desať vedeckých monografií.

Niektoré publikácie v kategórii Q1 v posledných šiestich rokoch:

VESELÝ, Vojtech - ILKA, Adrian. Gain-Scheduled PID Controller Design. In Journal of Process Control. Vol. 23 (2013), s.1141-1148. ISSN 0959-1524 (2.179 - 2013).

VÖRÖS, Jozef. Identification of nonlinear dynamic systems with input saturation and output backlash using three-block cascade models. In Journal of the Franklin Institute. Vol. 351, No. 12 (2014), s. 5455-5466. ISSN 0016-0032 (2.395 - 2014).

BABINEC, Andrej - DUCHOŇ, František - DEKAN, Martin - PÁSZTÓ, Peter - KELEMEN, Michal. VFH*TDT(VFH* with Time Dependent Tree): A new laser rangefinder based obstacle avoidance method designed for environment with non-static obstacles. In Robotics and Autonomous Systems. Vol. 62 (2014), s. 1098-1115. ISSN 0921-8890 (1.256 - 2014). V databáze: CC: 000338405000002.

VESELÝ, Vojtech - ILKA, Adrian. Design of robust gain-scheduled PI controllers. In Journal of the Franklin Institute. Vol. 352, No. 4 (2015), s. 1476-1494. ISSN 0016-0032 (2.327 - 2015). V databáze: CC: 000353084900008.

VÖRÖS, Jozef. Iterative identification of nonlinear dynamic systems with output backlash using threeblock cascade models. In Nonlinear Dynamics. Vol. 79, No. 3 (2015), s. 2187-2195. ISSN 0924-090X (3.000 - 2015). V databáze: WOS: 000349244900037

DEKAN, Martin - DUCHOŇ, František - BABINEC, Andrej - HUBINSKÝ, Peter - KAJAN, Martin -SZABOVÁ, Martina. Versatile approach to probabilistic modeling of Hokuyo UTM-30LX. In IEEE Sensors Journal. Vol. 16, Iss. 6 (2016), pp. 1814-1828. ISSN 1530-437X

KONIAR, Dušan - HARGAŠ, Libor - LONCOVÁ, Zuzana - DUCHOŇ, František - BEŇO, Peter. Machine vision application in animal trajectory tracking. In Computer Methods and Programs in Biomedicine. Vol. 127, (2016), pp. 258-272. ISSN 0169-2607

CHOVANCOVÁ, Anežka - FICO, Tomáš -HUBINSKÝ, Peter - DUCHOŇ, František. Comparison of various quaternion-based control methods applied to quadrotor with disturbance observer and position estimator. In Robotics and Autonomous Systems. Vol. 79, (2016), s. 87-98. ISSN 0921-8890 (1.950 - 2016). V databáze: SCOPUS: 2-s2.0-84958581863 ; CC: 000374358500007.

VESELÝ, Vojtech - ILKA, Adrian. Generalized robust gain-scheduled PID controller design for affine LPV systems with polytopic uncertainty. In Systems and Control Letters. Vol. 105, (2017), s. 6-13. ISSN 0167-6911 (2.656 - 2017). V databáze: CC: 000405764200002.

VÖRÖS, Jozef. Recursive identification of discretetime nonlinear cascade systems with time-varying output hysteresis. In Nonlinear Dynamics. Vol. 87, Iss. 2 (2017), s. 1427-1434. ISSN 0924-090X (4.339 - 2017). V databáze: CC: 000392293200049

TÖLGYESSY, Michal - DEKAN, Martin - DUCHOŇ, František - RODINA, Jozef - HUBINSKÝ, Peter -CHOVANEC, Ľuboš. Foundations of visual linear human-robot interaction via pointing gesture navigation. In International Journal of Social Robotics. Vol. 9, Iss. 4 (2017), pp. 509-523. ISSN 1875-4791.

BERNÁTOVÁ, Iveta - BALIŠ, Peter - GOGA, Rudolf - BEHULIAK, Michal - ZICHA, Josef - SEKAJ, Ivan. Lack of reactive oxygen species deteriorates blood pressure regulation in acute stress. In Physiological research. Vol. 65, Suppl. 3 (2016), s. S381-S390. ISSN 0862-8408 (1.461 - 2016). V databáze: SCOPUS: 2-s2.0-84997542933 ; CC: 000386689100008. Monografie:_

MURGAŠ, Ján - POLIAČIK, Michal. Priame adaptívne riadenie spojitých procesov. 1. vyd. Bratislava : Nakladateľstvo STU, 2013. 137 s. ISBN 978-80-227-4109-5

ŠKULTÉTY, Jozef - MIKLOVIČOVÁ, Eva. Riadenie systémov s využitím Laguerreových sietí. 1. vyd. Bratislava : Felia, 2014. 97 s. Edícia: Kybernetika. ISBN 978-80-971512-3-2.

DUCHOŇ, František - BABINEC, Andrej. Lokalizácia mobilných robotov. 1. vyd. Bratislava : Nakladateľstvo STU, 2015. 121 s. ISBN 978-80-227-4461-4.

LUDWIG, Tomáš - TÁRNÍK, Marián. Biokybernetické modely diabetu. 1. vyd. Bratislava : Felia, s.r.o., 2015. 119 s. Edícia: Kybernetika. ISBN 978-80-971512-8-7.

BABINEC, Andrej - VITKO, Anton. Navigačné metódy VFH v nestatickom prostredí. 1. vyd. Bratislava : Vydavateľstvo STU, 2016. 106 s. ISBN 978-80-227-4548-2.

LODERER, Marek - PAVLOVIČOVÁ, Jarmila. Biometria. Rozpoznávanie ľudských tvárí [elektronický zdroj]. 1. vyd. Bratislava : Felia, 2016. CD-ROM, 113 s. ISBN 978-80-89824-06-9.

TÖLGYESSY, Michal - HUBINSKÝ, Peter. Metódy na riadenie mobilných robotických systémov pomocou gest. 1. vyd. Bratislava : Spektrum STU, 2018. 105 s. Dostupné na internete: . ISBN 978-80-227-4789-9.

Na podporu výskumu sme získali desiatky vedeckých projektov od agentúr VEGA a APVV v aplikovanom výskume aj veľkých projektov zo štrukturálnych fondov s významnou finančnou podporou. Uvedieme len tie s najvyššou podporou::

Murgaš, Ján a kol.: Kompetenčné centrum inteligentných technológií pre elektronizáciu a informatizáciu systémov a služieb, ITMS 26240220072, 2011-2015

Murgaš, Ján a kol.: Centrum výskumu závažných ochorení a ich komplikácií, 2014-2015, ITMS 26240120038

Prof. Ing. Ján Murgaš, PhD.

Slovenská technická univerzita v Bratislave, Ilkovičova 3 812 19 Bratislava

Ústav robotiky a kybernetiky

e-mail: jan.murgas@stuba.sk

SYSTEM – BASED ANALYSIS OF CONDITIONS FOR USING IN-SILICO MODELLING AS SURROGATE FOR IN-VIVO STUDIES

Zuzana Vitková, Ján Murgaš, Anton Vitko, Andrej Babinec, Eva Miklovičová and Marián Tárník

Abstract

The contemporary cybernetics has a decisive impact on the successful solutions of many problems in biology and medicine. There is the strong need for analyses of the macro and micro processes running inside the body or molecules. Most of them are initiated just by inserting a drug into a living organism. The extraction of explicit knowledge that is hidden in the results of the *in-vivo* experiments performed on the animals or human volunteers is one of the current key topics of bio-cybernetics. The problem covered in this paper is related to the system-based study of conditions under which the analysis of a model of a specific structure may be used as a surrogate of the *in-vivo* experiment. An incentive for tackling the problem is the fact that the model-based prediction of the fate of drugs in the body can significantly reduce the cost of the *in-vivo* experiment as a whole.

Keywords: compartment model; identifiability; sensitivity; surrogate, in-vivo and insilico experiments

Introduction

Due to growing amount of biologic and medicine data, which are stored in databases around the word raises the need for development of the well-founded reliable models based on experimental data and give them sound biological meanings. The need is similar on the level of whole body, organs and even molecules. To fill the gap, the scientists call for integration of experimental results with modelling formalism in order to get deeper understanding of the leaving systems. The dominant formalisms used on the level of the whole body (or body organs) is linked with the differential/difference equations while on inter/intra molecular level the discrete means, like automata and Petri nets are used [1].

The system biology intends to describe bioprocesses in a holistic way and to understand them by integration of analytic experimental approaches with synthetic computational models. The same approach is adopted in this paper. The paper deals with analyzing the fate of a drug delivered into a body and tries to find conditions under which it would be possible to replace the *in-vivo* experiments by the experiments *in-silico*.

The drug delivery systems transport the drug either into the site of absorption or directly into the blood circulation. In both cases the aim is to evoke a desirable therapeutic response. Many principles and technologies of the drug delivery were developed, and many novel approaches are under development. This naturally calls for effective methods of assessing the influence of various technological modifications of dosage forms on *in-vivo* performance measures, like the therapeutic effect or the drug bioavailability. In what follows the term "drug" denotes an active therapeutic agent (medicine), while the "dosage form" means a manufactured product like tablets, drageés, ointments, powders, etc. In cases where any ambiguity is impossible we simply use the term "drug"

The holistic system-based approach conceives all presystemic and systemic processes induced by the application of a drug as a unique process, which may or may not incorporate some sub-processes into which the drug enters before reaching the blood circulation. These are mainly the disintegration of a solid dosage form into small-size particles, followed by the drug dissolution (release) of drug particles, which finally gives rise to the solution. These processes are pre-absorption processes and may be analyzed in-vitro i.e. in a test tube. The knowledge about these processes may be obtained separately by the in-vitro experiment and then used for evaluation of in-vitro/in-vivo correlation (IVIVC). Pre problems of IVIVC are beyond the subject of this paper. Only the dissolved drug can be absorbed into the blood circulation as illustrated in Fig. 1. A possible assignment of the dynamics equations to the particular processes is shown in Fig.1 below. For illustration of the dynamics of the disintegration, dissolution and absorption are in Fig.1 represented by transfer functions. Depending on the particular drug and/or dosage form, these processes may be strongly nonlinear. The only nonlinear process shown in Fig.1 is the drug dissolution, which is frequently modelled by the Higuchi's square root law [2]. The block denoted as a "reservoir of the dissoluted drug" is also depicted, but it make sense only when a depot drug form (that of with very slow drug release) was applicate. This case will not be considered here. The serial (catenary) structure the block scheme as shown in Fig.1 is typical for pre-absorption processes and very common for post-absorption ones.

Though the entire process of the *pre* and *post* absorption sub-processes may be analyzed as a unique process, the pre-absorption processes are not analyzed in this paper. The first author has analyzed them in advance and using special methods she determined both the character of absorption dynamics and the absorption rate constant K_a [2].

Subject of this paper is restricted to the analysis of the process running in between the blood circulation and the rest of the body. The process starts at the instant of entering the first molecules of the drug into blood circulations and ends at the instant of elimination of the last molecule from the body.

The absorption is the first process that is controlled by the body. All pre-absorption processes may be controlled by the both "construction" and composition of the dosage form, for instance by using of a layered or in other way structured tablet, drageé and the like. The special drug delivery dosage forms, for instance a pump for delivery of insulin are much easily controlled on the closed loop and create a specific category.



Fig.1 Chain of pre-absorption processes (above) and corresponding block scheme (below)

General issues of compartmental modelling

From the view of pharmacokinetics the body is divided into a few compartments i.e. fictitious volumes in which are homogenously distributed drugs. The compartments are supposed to be separated by a semipermeable membrane through the drug can penetrate. Every compartment is characterized by its own distribution volume V_d. Due to concentration gradient the drug flows form one compartment to another. The flow is oriented from the donor compartment to the acceptor compartment. The flow rates are proportional to the current drug concentration in the donor compartment. Note that in the drug development parlance the term "pharmacokinetic model", which actually is an umbrella term, is sometimes erroneously put on a par with the "compartment model" though there exist also non-compartment, statistic and other kinds of pharmacokinetic models. [3-5]

The simplest model is a first order dynamic model, which represents the whole body as a single compartment. Here it is used the model shown in Fig. 2, which consists of two compartments, each exhibiting the first order dynamics. Such a model better matches the experimental data.

One very important question is related to the number of compartments to be used. The answer is rather complex from the aspects of the both system theory and physiology. The thing is that the model's structure must be physiologically substantiated, clearly interpretable and structurally identifiable (see next paragraph). Another problem is related to the system realization, as it is known in the realm of the system theory. Namely, from the *in–vivo* measurements

one can only obtain a series of output samples as a response to series of input samples, but they do not bear any information about the "structure" of the process running inside of the body. In case when the processes are linear the input/output model may take the form of the transfer function. Then, having the transfer function one may derive a specific realization (the state space model), which must be not only physiologically sensible, but, it must represent a positive realization. Let us note that the drug amounts, concentrations or volumes are always positive quantities. The compartment models belong to category of internally positive systems. It means that not only inputs and outputs but also internal states take exclusively positive values. It can be shown [6,7] that state trajectory x(t) of the linear compartment model that starts from the nonnegative initial state $\dot{x}(0)$ will remain nonnegative, if and only if the control vector **b** is nonnegative and the matrix A is a Metzler matrix. (a nonzero matrix with nonnegative off-diagonal entries).

In this study was used the linear SISO compartment model described by the state space model (1)

$$\frac{d\mathbf{x}(t)}{dt} = A\mathbf{x}(t) + bu(t)$$
$$y(t) = \mathbf{c}^{\mathsf{T}}\mathbf{x}(t)$$
(1)

The constraints laid on model parameters follow from the fact that compartment models respect validity of the law of mass conservation and the system positivity. The constrains may be summarized as follows:

*a*_{i,j} ≥0, *i*≠j,

$$a_{ii} + \sum_{j \neq i} a_{j,i} \leq 0$$
 $i, j = 1, 2, ...n$ (2)

The model used in this study is described by the Eqs. (4)

Structural identifiability

The need for identification of model parameters opens the questions on existence of the non-unique sets of values of model parameters, which correspond to the I/O measurements or even principal impossibility to obtain such a unique set. When for the model parameters is not requested to have physical meanings the input-output representation may be sufficient. But if one wants to understand the observed phenomena and wants to integrate *a prior* knowledge, the both model structure and parameters should be biologically interpretable.

Contrary to the canonical structures known from the theory of linear systems, parameters of compartment models are different from those of the transfer function. On other hand, identifiability analysis has strong impact on the design of invivo experiments. For instance it provides information on the allowable locations of inputs (the sites where drugs may be administered) and also on the inputs (the sites where the measurements may be performed) to allow unique parametric identifiability. It is fundamental knowledge the model designer should have on disposal before setting up the invivo experiment. Obviously, such knowledge is also a basic prerequisite for the process of parameter estimation to be well posed. Unfortunately, the model designer has on its disposal only input/output samples; hence he is only able to uniquely identify the parameters of the transfer function. The only thing he is able to do, is to find relations between parameters of the transfer function and parameters of the compartment model. In this way he can obtain a set of algebraic equations. The identifiability of pharmacokinetic parameters than depends on the unambiguous solvability of these algebraic equations. The problem is known as *structural identifiability*. [6-9] Saying this in a slightly formal way, the problem of structural identifiability is considered resolved if a choice of the vectors **b** and **c** exists that parameters of the system matrix A can be unambiguously determine from I/O data. Only then the estimation of the model parameters can be carried out by an appropriate regression method.

The analysis of a chosen compartment model for structural identifiability is always an important task of the experimental design. Therefor the structural identifiability test is always an *a-priori* test, which should be performed prior to the experimentation with the model.

The transfer function H(s) of the model shown in Fig.2 between the input $u(t) = K_a M_0 \exp(-K_a t)$ and the output $y(t) = M_1(t)$ takes the form:

$$H(s) = \frac{s + k_{21}}{s^2 + (k_{21} + k_{12} + k_{e}) + k_{e}k_{21}} = \frac{\beta_2 s + \beta_1}{s^2 + \alpha_2 s + \alpha_1}$$
(3)

where

 $\beta_2 = 1$ $\beta_1 = k_{21}$

 $\alpha_2 = k_{11} + k_{12} + k_e$

 $\alpha_1 = k_e k_{21}$

 $\alpha_2 + \beta_1 = k_{12} + k_e$

From this follows:

$$k_e = \frac{\beta_1}{\alpha_1}$$

 $k_{21} = \beta_1$

$$k_{12} = \alpha_2 - \beta_1 - \frac{\beta_1}{\alpha_1}$$

Hence, the compartmental model is structurally identifiable.

See also for instance [6-9] where the reader may find other methods and operations that preserve structural identifiability. of solving problem of structural identifiability.

The following is worth mentioning: The structure of the compartment model shown in Fig. 2 belongs to the category of so called *catenary* structures. On other hand it is known that the single input/output catenary structures in which the input/output measurements are performed on the most left or the most right compartment, (in this case on either M₁ or M₂ compartment) is structurally identifiable [8]. Actually, in the Fig 2 the drug is inputting and outputting from the most right compartment M₁. Let us note again that the site of application is not considered as a part of the compartment model. In view of (3) and [8] the structure shown in Fig.2 is structurally identifiable

Structures of compartment models follow from preliminary analyses of in-*vivo* measurements, which are performed on the experimental animals or human volunteers. Leaving alone all pre-absorption processes, which were separately analyzed *in-vitro*, the compartment model shown in Fig. 2 illustrates movement of the absorbed drug into the *central compartment* M₁ (the systemic blood circulation). A measure of the rate of absorption is represented by the *absorption rate* constant K_a. From the central compartment it is partially eliminated by the *elimination rate* constant K_e and partially enters the *peripheral* compartment M₂ (the rest of the body). While not completely eliminated the drug circulates between compartments in the rates given by the *hybrid* rate constant K₁₂ and K₂₁. Symbols M₁, M₂ represent instantaneous

amounts of the drug in compartments. The site of application is not considered to be a compartment.



Fig.2 Two compartment model of an extravascular application with elimination from the central compartment

The meaning of the symbols is the following:

M₀ is an administered dose

M₁ is an amount of drug in central compartment

M₂ is an amount of drug in peripheral compartment

Ka is a rate constant of drug absorption

 K_{12} , K_{21} are rate constants of the drug exchange between compartments

Ke is a rate constant of the drug elimination

The compartment model is described by the state space equations (5). The site of application is supposed to exhibit the first order absorption, which is represented by the exponential term in the end of the first equation. Because the absorption rate is commonly much greate*r* than the release rate, the absorption rate equals to the release rate. In other words, what is released that is instantly absorbed. That is why we set $K_a = K_r$, where K_r (not mentioned in Fig.2) is the release rate constant that was determined by the separate *in-vitro* experiment and it takes value $K_r = 4 h^{-1}$. This imply that the input to the system is $u(t) = K_a M_0 exp(-K_a t)$. As can be seen in (4) it is directly inserted into the first equation.

$$\frac{dM_1}{dt} = -(K_{12} + K_e)M_1 + K_{21}M_2 + K_aM_0 \exp(-K_at)$$

$$\frac{dM_2}{dt} = K_{12}M_1 - K_{21}M_2$$
(4)

 $y(t) = M_1(t)$

Parameters of the system matrix A were estimated by the MLS algorithm as K_e = 0.98 h⁻¹, K_{12} = 2.06 h⁻¹ and K_{21} = 0.61 h-1.

The criterial function used in the parameters estimation is given in (4)

$$J = \sqrt{\frac{\mathbf{e}^{\mathsf{T}}\mathbf{e}}{\mathsf{N}}} \tag{5}$$

Vector **e** stands for differences between samples of the blood concentrations or masses measured in the central compartment (denoted as M1) and the regressing values and N is a number of measurements. For the purposes of the simulation the distribution volumes V_d were set to one and in-vivo values were also normalized by their maximum values. Due to that the time courses of the drug amounts M₁, M₂ and the corresponding drug concentrations coincide. The minimum of the criterial function was found by the simplex algorithm (Nelder-Mead algorithm).

Concept of sensitivity functions

A general form of a linear n-compartment pharmacokinetic model is given by the set of the first order differential equations (6)

$$\frac{dM_{i}(t)}{dt} = F_{i}(M_{1}, M_{2}, \dots, M_{n}, q_{1}, q_{2}, \dots, q_{m}) \quad i=1, 2, \dots n$$
(6)

where F_{i} is a (possibly nonlinear) function describing the i-th compartment.

 $M_{i}(t)$ is an instantaneous drug amount in the i-th compartment.

- n is a number of compartments.
- m is a number of model parameters qi.

The sensitivity function (in short "sensitivity") of the amount M_i to the deviations of the parameter q_j is defined as partial derivation of M_i w.r.t parameter q_j .

$$S_{M_i,q_j} = \frac{\partial M_i(t)}{\partial q_i} \tag{7}$$

For instance, the sensitivity of M_2 to a (small) deviation of the absorption rate constant K_a is given by the partial derivation

$$S_{M_2,K_a} = \frac{\partial M_2}{\partial K_a} \tag{8}$$

All (n x m) sensitivities may be obtained as solutions of the set of (n x m) differential equations (9) that are known as a sensitivity equations.

$$\frac{dS_{M_i,q_j}}{dt} = \sum_{i=1}^{n} \frac{\partial F_K}{\partial M_i} S_{M_i,q_j} + \frac{\partial F_K}{\partial q_i} \qquad k=1,2,\dots,n; \quad j=1,2,\dots m (9)$$

If a parameter does not represent an initial condition then initial conditions of the sensitivity equations are equal to zeros. Having the sensitivities known one can calculate deviations ΔM_i (t), which were caused by deviation Δq_i of the parameter q_i from its nominal value as it is indicated in (10)

$$\Delta M_i(t) = S_{M_i,q_j}(t) \Delta q_j \quad i=1.2..., n; \quad j=1.2...m$$
(10)

The new (deviated) values of Mi will be given by (11)

$$M_i(t)_{new} = M_i(t) + \Delta M_i(t) \tag{11}$$

As it was envisaged, the main aim of this study is to analyze the influence of variations of the absorption rate constant K_a on the time course of the drug amount M_1 in the blood circulation. Such knowledge is inevitable for making decision about the extent to which the model (4) may be considered as a surrogate of *in-vivo* studies. This is done on the basis of the sensitivity analysis. The sensitivity of the amounts M_1 and M_2 to (small) deviations of K_a are obtained as solutions of the sensitivity equations (12)

$$\frac{dS_{M_1,K_a}}{dt} = -(K_{12} + K_a) + K_{21}S_{M_2,K_a} + M_0(1 - t \cdot K_a) \cdot exp(-K_a \cdot t)$$

$$\frac{dS_{M_2,K_a}}{dt} = K_{12}S_{M_1,K_a} - K_{21}S_{M_2,K_a}$$
(12)

Results of simulation

Equations (5) and (12) were solved by a computer and the obtained results are shown in Fig 3. The vertical axis is scaled by the normalized values of M_1 and $S_{M1,Ka}$. The dashed piecewise linear curves with asterisks and squares represent in-*vivo* samples. In particular, the asterisks represent *in-vivo* samples of the amount M_1 for the increased absorption rate K_a , while the squares represent nominal samples. The *in-vivo* increase of the absorption rate constant Ka was induced by addition of the surface active substance Tween 80 to the initial dose M_0 . One can see that the asterisks are located at the beginning above the squares but after approximately 0.65 hour they fall below them. The same tendency can be observed in the simulated trajectories of $M_1(t)$ that are drawn by the full curves.

In relation to the subject of this paper it is important to note that the same behavior is also predicted by the time course of the sensitivity $S_{Ka}(t)$, (in Fig.3 denoted as SM1,Ka). Its

sign is changing just at the time instant 0.65 h. Hence, the real *in-vivo* process virtually exactly follows the one predicted by the augmented compartment model. On the top of that, the sensitivity $S_{M1,Ka}(t)$ correctly predicts not only magnitude but also the direction of the deviated time profile of the drug amount as it is indicated in (11).

Conclusions

The increase of the time profile of the drug amount M_1 by ΔM_1 = $S_{M1.Ka}$. ΔKa caused by the increased value of the absorption rate constant Ka was predicted by the augmented two-compartment pharmacokinetic model (4) and (12). The time profiles of the nominal $M_1(t)$ and increased drug amounts $M_1(t)$ + $\Delta M_1(t)$ in the central compartment were generated by the validated compartment model and subsequently compared with the results obtained by the *in-vivo* experiment.

The increase of the *in-vivo* absorption rate constant K_a was observed in the following *in-vivo* experiment: The drug Natrium p-aminosalicylicum (NaPAS) was administered into the rats' gastrointestinal tract. After some washout period that is necessary for total elimination of the drug from the body, the experiment was performed again with the same drug but with addition of the surface active agent Tween 80, which increased the rate of absorption K_a . The *in–vivo* time profile and the simulated time profiles were drown and compared.

The rate constant K_a was determined from the *in-vivo* data by the "peeling off "method [2]. For the sake of simplicity were the volumes of distribution set to one. In case when the actual value of V_d is needed it may be approximated by the expression V_d= 0.06 x BW +0.77 where BW means the body weight [4]. Pharmacokinetic parameters K_e, K₁₂ and K₂₁ were estimated by the Nelder–Mead simplex algorithm with the quadratic criterial function.

The number of compartments n=2 was determined by "the rule of thumb" [2] form the semi-logarithmic plot of *in-vivo* data.

From the system point of view it is important to note that the choice of the model's structure is closely related to the structural identifiability. Those issues must be resolved prior to practical experimentations. Therefore; the structural identifiability of the compartment parameters was also verified before performing the experiments.. To this end the transfer function was first derived and its (still symbolic) coefficients of the numeration and denominator were used for formulation of a set of algebraic equations for unknown parameters K_{e} , K_{12} , K_{21} . The unique solutions of these equations indicate that the compartment model (5) is structurally identifiable.

It was stressed that the pharmacokinetic compartment models are commonly derived on the basis of physical laws. In this case they are Fick's law of simple diffusion and the law of mass conversion. Besides, the mode's structure has to allow for the clear physiological interpretations. The qualitative and quantitative behaviors of both the simulated and the in-vivo processes were shown to be virtually the same what certifies model validity. Sensitivities of the drug amount M1 in the blood circulation (the central compartment) to deviations of the absorption rate constant Ka observed in the both in-vivo and in-silico experiments were also very close. These facts, together with fulfillment of the constraint imposed on the structure and parameters, proves validity of the compartment model, what has important consequences: The identifiable and validated compartment model can be used as a surrogate of the *in-vivo* experiment, due to which



Fig.3 Time courses of $S_{M1,Ka}$ and M_1 before and after increase of K_a

the in-vivo experiment may be significantly reduced.

Leaving alone the mentioned advantages the presented approach allows also for generalization of the traditionally conceived problem of the *in-vitro* versus *in–vivo* correlation (IVIVC). In particular, the criteria for evaluation of the similarity between *in-vitro* dissolution time profiles and *in-vivo* absorption time profiles may be significantly generalized [10]. It is worth noting that question of this correlation play a decisive role in the way of the drug development], in particular in the decisions on the drug bio-equivalency.

To this end we formulated and resolved an inverse problem, namely determination of an optimal absorption rate for a given (desired) time course of the therapeutic effect. As a byproduct it is shown that the *in-silico* generated sensitivity of plasmatic concentration to changes of the absorption rate (caused arbitrary post-approval changes of auxiliary substances) provides invaluable information about a degree of belief that an *in-silico* testing can serve as a waiver of *invivo* bioavailability and bioequivalence studies [11]

Acknowledgment

The paper is one of the outcomes of the research work for the project entitled "Research center for severe diseases and related complications", "ITMS: 26240120038". "This project is being co-financed by the European Union. We support research activities in Slovakia".

References

[1] BALDAN, P., COCO, N, MARTIN, A.: Petri nets for modelling metabolic pathways: A survey. Nat. Comput. (2010), 955-989. DOI: 10.1007/s11047-010-9180_6

[2] VITKOVÁ, Z.: Úvod do farmakokinetickej analýzy. STU, Bratislava 2001, ISBN 80-227-1766-5

[3] VITKOVÁ, Z., VITKO, A., ŽABKA, M. MANDÁK. M.: Bezkompartmentová – modelovo nezávislá farmakokinetická analýza I. - Základné princípy štatistickej farmakokinetickej analýzy. Čs. farmácia, 38(1989), 64-68 ISSN 0009-0530

[4] VITKOVÁ, Z., VITKO, A., ŽABKA, M. MANDAK. M.: Bezkompartmentová – modelovo nezávislá farmakokinetická analýza II. – Využitie Laplaceovej transformácie v bezkompartmentovej farmakokinetickej analýze. Čs. farmácia, 38(1989), 141-144, ISSN 0009-0530

[5] VITKOVÁ, Z., VITKO, A., ŽABKA, M. MANDAK. M.: Bezkompartmentová – modelovo nezávislá farmakokinetická analýza III. – Dekonvolúcia a štatistické momenty v analýze procesov absorpcie liečiva. Čs. farmácia, 38(1989), 370-373, ISSN 0009-0530

[6] MESHKAT, N., ER-ZEN KUO, CH., DISTEFANO, J.: On finding and using identifiable parameter combinations in nonlinear dynamic systems biology models and COMBOS. A novel web implementation. PloS ONE 9(10): e110261, 9(2014), Issue 10, DOI: 10.1371/journal.pone0110261

[7] GREISZ, W. BOURIN,NM: Identification of pharmacokinetic two compartment model application to the case of salicylates administered. Mathematical modelling 7(1986), 1245 -1253

[8] COBELLI C., A.: A priori identifiability analysis in pharmacokinetic experiment design. IN: Endrenyi (ed) Kinetic drug analysis, Plenum press, New York 1981, 181-208

[9] GROSS E., HARRINGTON, H., MESHKAT. N., SHIU A.: Linear compartmental models: Input-output equations and

operations that preserve identifiability. SIAM J. Appl. Math, 79(4), 2423-1447, https://arxiv.org/pdf/1810.05575

[10] VITKOVÁ, Z., VITKO, A., ŽABKA, M. ČIŽMÁRIK, J.: Alternative and generalized approach to *in vitro-in vivo* correlation. Acta Poloniae Pharmaceutica - Drug research, 68(2011), 417-421, ISSN 0001-6837

[11] VITKOVÁ Z., TÁRNÍK, M, MIKLOVIČOVÁ, E., MURGAŠ, J., OREMUSOVÁ, J., AND VITKO, A.: Surfactants' Influences on Pharmacokinetics and Pharmacodynamics of Drugs. Tensides Surfactants Detergents, 2019, *DOI:* org/10.3139/113.014042 (accepted for publication)

Doc. RNDr. Zuzana Vitková, PhD. Prof. Ing. Ján Murgaš, PhD. Prof. Ing. Anton Vitko, PhD Doc. Ing. Andrej Babinec, PhD. Doc. Ing. Eva Miklovičová, PhD. Ing. Marián Tárník, PhD.

Slovak University if Technology, Faculty of Electrical Engineering and Information Technology Institute of Robotics and Cybernetics Ilkovičova 3, 812 19 Bratislava, Slovakia e-mails: zuzana.vitková@stuba.sk jan.murgas@stuba.sk anton.vitko@stuba.sk andrej.babinec@stuba.sk

marian.tarnik@stuba.sk

IDENTIFICATION OF NONLINEAR CASCADE SYSTEMS WITH STATIC INPUT AND DYNAMIC OUTPUT NONLINEARITIES INCLUDING DEAD- ZONES

Jozef Vörös

Abstract

Parameter identification of nonlinear dynamic systems with static input and dynamic output nonlinearities including dead-zones is presented. Multiple application of a decomposition technique provides special expressions for the corresponding nonlinear model description that are linear in parameters. A least-squares-based iterative technique allows estimation of all the model parameters based on measured input/output data. An illustrative example of three-block cascade system identification with dead-zone and backlash characteristics is included.

Keywords: Nonlinear systems; identification; cascade models; dead-zone; backlash

Introduction

Cascade models are a popular type of block-oriented models and the simplest types of cascade nonlinear models consist of two blocks. The so-called Hammerstein model consists of a static nonlinear block followed by a linear dynamic block and the so-called Wiener model consists of a linear dynamic block followed by a static nonlinear block. Generally, the Wiener model is supposed to represent sensor nonlinearities, while the Hammerstein model is supposed to represent actuator nonlinearities. These models appear in many engineering applications not only for their relative simplicity but also for their ability to approximate more general nonlinear systems, which are not necessarily of this form. Therefore, the identification of nonlinear dynamic systems using Hammerstein and Wiener models has been an active research area for many years. Some recent works dealing with the Hammerstein models can be found in [1 - 9], while the Wiener models are used in e.g., [10 - 21].

If the systems to be identified contain nonlinearities with memory such as backlash or hysteresis [22], the choice of Hammerstein or Wiener models is not appropriate, because the nonlinear static block cannot characterize these nonlinearities. Therefore, a special case of two-block models consisting of the cascade of linear dynamic and nonlinear dynamic blocks have to be used for modeling and identification of nonlinear dynamic systems with backlash or hysteresis.

In analogy with the Wiener model, the cascade model structure consisting of a linear dynamic block followed by a nonlinear dynamic block was often used for the identification of nonlinear dynamic systems with output dynamic nonlinearities, e.g., in [23 - 25]. However, in some cases of more complex nonlinear dynamic systems with actuator nonlinearities such as dead-zones, the two-block cascade model may be not precise enough and it is appropriate to choose a threeblock cascade model with combination of nonlinear static, linear dynamic, and nonlinear dynamic blocks. It means, compared to the well-known structure of Hammerstein-Wiener model (see e.g. [26 - 30]), that the input block contains static nonlinearities while the output block contains dynamic nonlinearities, e.g., backlash or hysteresis. Actually, this form of three-block cascade model can be considered as a cascade of Hammerstein model and a dynamic nonlinearity block and significantly extends the applicability for systems with both actuator and sensor nonlinearities.

In this paper, the three-block cascade model is used to the parameter identification of nonlinear dynamic systems with static input and dynamic output nonlinearities including deadzones (Fig. 1). The previous results on the decomposition of compound operators [8, 18] are effectively applied to simplify the mathematical description of this complex system. The resulting input/output equation is without cross-multiplication of parameters; nevertheless, it contains more internal variables, which are generally unmeasurable. Application of a least-squares-based iterative algorithm enables estimation of all the model parameters on the basis of measured input/output data. An illustrative example of nonlinear dynamic system identification with input dead-zone and output backlash characteristic is included.

Three-block cascade model

Let a three-block cascade model be given by the cascade connection of a nonlinear static block followed by a linear dynamic block, which is followed by a nonlinear dynamic block according to Fig. 1.



Fig.1 Three-block cascade system with input dead-zone and output backlash

The output of nonlinear static block with dead-zone v(t) can be written as

$$v(t) = \begin{cases} 0 & \text{if } 0 \le u(t) \le d_1 \\ m_1[u(t) - d_1] & \text{if } u(t) > d_1 \end{cases}$$
(1)

$$v(t) = \begin{cases} 0 & \text{if } d_2 \le u(t) < 0\\ m_2[u(t) - d_2] & \text{if } u(t) < d_2 \end{cases}$$
(2)





After introducing internal variables $g_1(t)$ and $g_2(t)$ as follows:

$$g_1(t) = m_1 h[d_1 - u(t)]$$
(3)

$$g_2(t) = m_2 h[u(t) - d_2]$$
(4)

where h(.) is a switching function defined as

$$h(\sigma) = \begin{cases} 0 & \text{if } \sigma \ge 0\\ 1 & \text{if } \sigma < 0 \end{cases}$$
(5)

the input/output relation of nonlinear block can be given by

$$v(t) = m_1 h[d_1 - u(t)]u(t) - d_1 g_1(t) + m_2 h[u(t) - d_2]u(t) - d_2 g_2(t)$$
(6)

The linear dynamic block can be described by the following difference equation

$$x(t) = \sum_{i=1}^{r} a_i v(t-i) - \sum_{j=1}^{p} b_j x(t-j)$$
(7)

where x(t) and y(t) are the inputs and outputs of LD, respectively, and we assume that *r* and *p* are known.

Let the nonlinear dynamic block be a backlash with inputs x(t) and outputs y(t). The backlash is a dynamic nonlinearity (Fig. 3) and can be described by the following first order nonlinear difference equation [25]

$$y(t) = m_L(x(t) + c_L)f_1(t) + m_R(x(t) - c_R)f_2(t) + y(t-1)(1 - f_1(t))(1 - f_2(t))$$
(8)

where $f_1(t)$ and $f_2(t)$ are auxiliary internal variables defined as:

$$f_1(t) = h((m_L x(t) + m_L c_L - y(t-1))/m_L)$$
(9)

$$f_2(t) = h((y(t-1) - m_R x(t) + m_R c_R) / m_R)$$
(10)

and h(.) is the above-defined switching function.



Fig.3 Backlash

The input-output equation of the three-block cascade system resulting from direct substitutions of the corresponding variables from (6) into (7) and then into (8) would be a very complex equation that is strongly nonlinear both in the variables and in the parameters, hence not very suitable for the parameter estimation. To find a simpler form of this description, the so-called key-term separation principle will be applied [8, 18]. Because of the cascade connection of three blocks, the parameterization of the blocks is not unique, as many combinations of parameters can be found. Therefore, in at least two blocks, one parameter has to be fixed. Choosing $m_L = 1$, we rewrite (8) as follows:

$$y(t) = x(t)f_1(t) + c_L f_1(t) + m_R x(t)f_2(t) - m_R c_R f_2(t) + y(t-1)(1-f_1(t))(1-f_2(t))$$
(11)

and half-substitute x(t) from (7) into (11), i.e. only for x(t) in the first term on the right-hand side of (11) obtaining

$$y(t) = \sum_{i=1}^{r} a_i v(t-i) f_1(t) - \sum_{j=1}^{p} b_j x(t-j) f_1(t) + c_L f_1(t) + m_R x(t) f_2(t) - c f_2(t) + (12) y(t-1)(1-f_1(t))(1-f_2(t))$$

where

$$c = c_R m_R \tag{13}$$

Then we can choose $a_1 = 1$, and half-substitute (6) into (12), i.e. only for the term with variable v(t-1). This will lead to the three-block cascade model description

$$y(t) = m_1 h[d_1 - u(t-1)]u(t-1)f_1(t) - d_1 g_1(t-1)f_1(t) + + m_2 h[u(t-1) - d_2]u(t-1)f_1(t) - d_2 g_2(t-1)f_1(t) + \sum_{i=2}^r a_i v(t-i)f_1(t) - \sum_{j=1}^p b_j x(t-j)f_1(t) + c_L f_1(t) + m_R x(t) f_2(t) - cf_2(t) + y(t-1)(1 - f_1(t))(1 - f_2(t))$$
(14)

This equation is linear in all the model parameters, but nonlinear in some variables. The model inputs u(t) and outputs y(t) are measurable, while the internal variables v(t), x(t), $g_1(t)$ $g_2(t)$, $f_1(t)$ and $f_2(t)$ are not.

Parameter estimation

The three-block cascade model equation can be written in the following concise form

$$y(t) = \varphi^{T}(t) \theta + y(t-1)(1 - f_{1}(t))(1 - f_{2}(t))$$
(15)

where

$$\varphi(t) = [h[d_1 - u(t-1)]u(t-1)f_1(t), -g_1(t-1)f_1(t),
h[u(t-1) - d_2]u(t-1)f_1(t), -g_2(t-1)f_1(t),
v(t-2)f_1(t), ..., v(t-r)f_1(t), -x(t-1)f_1(t), ...,
-x(t-p)f_1(t), f_1(t), x(t)f_2(t), -f_2(t)]$$
(16)

is the vector of data and

$$\theta = [m_1, d_1, m_2, d_2, a_2, \cdots, a_r, b_1, \cdots, b_p, c_L, m_R, c]^T$$
(17)

is the vector of parameters. As more variables in (15) are unmeasurable and must be estimated, an iterative parameter estimation process has to be considered. The techniques presented in [1, 8, 18], which are based on the use of the preceding estimates of model parameters for the estimation of internal variables and vice-versa, can be applied to this three-block model. We replace the internal variables in (15) by their estimates defined as follows:

$${}^{s}g_{1}(t) = {}^{s}m_{1}h[{}^{s}d_{1} - u(t)]$$
(18)

$${}^{s}g_{2}(t) = {}^{s}m_{2}h[u(t) - {}^{s}d_{2}]$$
(19)

$${}^{s}v(t) = {}^{s}m_{1}h[{}^{s}d_{1} - u(t)]u(t) - {}^{s}d_{1} {}^{s}g_{1}(t)$$
(20)

$$+^{s}m_{2}h[u(t) - ^{s}d_{2}]u(t) - ^{s}d_{2} ^{s}g_{2}(t)$$

$${}^{s}x(t) = {}^{s}v(t-i) + \sum_{i=2}^{r} {}^{s}a_{i}{}^{s}v(t-i) - \sum_{j=1}^{p} {}^{s}b_{j}{}^{s}x(t-j)$$
(21)

$${}^{s}f_{1}(t) = h({}^{s}x(t) + {}^{s}c_{L} - y(t-1))$$
(22)

$${}^{s}f_{2}(t) = h((y(t-1) - {}^{s}m_{R}{}^{s}x(t) + {}^{s}c)/{}^{s}m_{R})$$
(23)

Then the error to be minimized in the estimation procedure will be

$${}^{s+1}\varepsilon(t) = y(t) - {}^{s} \varphi^{T}(t)^{s+1} \theta - y(t-1)(1 - {}^{s} f_{1}(t))(1 - {}^{s} f_{2}(t))$$
(24)

where ${}^{s}\varphi(t)$ is the data vector with the corresponding estimates of variables according to (18) – (23) and ${}^{s+1}\theta$ is the (s+1)-st estimate of the parameter vector.

Assume the input is persistently excited with respect to the dead-zone. Then the iterative parameter estimation algorithm based on N input/output data consists of the following steps:

a) the s-th estimates of internal variables, i.e., ${}^{s}g_{1}(t)$, ${}^{s}g_{2}(t)$, ${}^{s}v(t)$, ${}^{s}x(t)$, ${}^{s}f_{1}(t)$, ${}^{s}f_{2}(t)$ are computed using (18) – (23);

b) the s-th estimate of ${}^{s}\varphi(t)$ is generated using the variables computed in the step a);

c) the (s+1)-st estimates of the model parameters are computed minimizing (24);

d) if the estimation criterion is met (e.g. the mean squares error is less than a predetermined value), the procedure ends, else it continues by repeating steps a) - c).

In the first iteration, only the parameters of nonlinear static block and the linear dynamic block are estimated where ${}^{1}x(t-j)$ is approximated by y(t-j). Nonzero initial values of the parameters ${}^{1}d_{1}$ and ${}^{1}d_{2}$ have to be considered for evaluation of ${}^{1}g_{1}(t)$, ${}^{1}g_{2}(t)$. Also, nonzero initial values of the backlash parameters ${}^{2}m_{R}$, ${}^{2}c_{L}$ and ${}^{2}c_{R}$ have to be considered for evaluation of ${}^{2}f_{1}(t)$, ${}^{2}f_{2}(t)$.

Example

The following example of simulated three-block cascade systems with input dead-zone and output backlash illustrate the feasibility of proposed identification method. The input dead-zone of the three-block cascade system (Fig. 4) was given by $m_1 = 0.9$, $m_2 = 0.5$, $d_1 = 0.4$ and $d_2 = -0.3$. The linear dynamic system was given by the difference equation

$$x(t) = v(t-1) + 0.5v(t-2) + 0.5x(t-1) - 0.4x(t-2)$$

followed by the output backlash (Fig. 5) characterized by the parameters m_L = 1.0, c_L = 0.2, m_R = 1.0, c_R = 0.3.





Fig.5 Output backlash

The identification was performed on the basis of 4000 samples of uniformly distributed random inputs with |u(t)| < 2.0and simulated outputs. Normally distributed random noise with zero mean and signal-to-noise ratio - SNR = 25 (the square root of the ratio of output and noise variances) was added to the outputs to make the simulation more realistic. The iterative estimation algorithm was applied with initial values ${}^{1}d_{1} = 0.001$, ${}^{1}d_{2} = -0.001$ for the first estimates of $g_{1}(t)$ and $g_2(t)$ and $m_R = 0.5$ and $c_L = c_R = 0.001$ for the second estimates of $f_1(t)$ and $f_2(t)$. The process of parameter estimation is shown in Fig. 6 for the dead-zone (the top-down order of parameters is: m_1 , m_2 , d_1 , d_2), in Fig. 7 for the linear block (the top-down order of parameters is: a2, b2, b1) and in Fig. 8 for the backlash (the top-down order of parameters is: mR, cR, c_L). The estimates meet the values of real parameters after about 9 iterations.





for the linear block



for the backlash

Conclusions

The presented approach to the identification of nonlinear dynamic systems with static input and dynamic output nonlinearities including dead-zones is based on a special form of three-block cascade model description resulting from more consecutive decompositions of compound mappings describing this block-oriented model. These models are appropriate for systems with both actuator and sensor nonlinearities.

An iterative least squares-based parameter estimation algorithm with internal variables estimations has been proposed and illustrated on examples of simulated nonlinear dynamic systems with dead-zone in the input block and with the backlash in the output block.

Finally, note that the proposed three-block cascade model can also be applied for on-line identification of nonlinear dynamic systems with input dead-zone and output backlash using the known recursive least-squares algorithm [31, 32]. The presented identification method can be extended also to the systems with the so-called general backlash [33 - 35].

Acknowledgment

The author gratefully acknowledges financial support from the Slovak Scientific Grant Agency (VEGA).

References

- [1] CHEN, H., DING, F., XIAO, Y.: Decomposition-based least squares parameter estimation algorithm for input nonlinear systems using the key term separation technique. Nonlinear Dynamics, Vol. 79, No. 3, pp. 2027-2035, 2015.
- [2] CHEN, J., WANG, X.: Identification of Hammerstein systems with continuous nonlinearity. Information Processing Letters, Vol. 115, No. 11, pp. 822-827, 2015.
- [3] LI, J., DING, F., HUA, L.: Maximum likelihood Newton recursive and the Newton iterative estimation algorithms for Hammerstein CARAR systems. Nonlinear Dynamics, Vol. 75, No. 1-2, pp. 235-245, 2014.

- [4] MA, J., XIONG, W., DING, F., ALSAEDI, A., HAYAT, T.: Data filtering based forgetting factor stochastic gradient algorithm for Hammerstein systems with saturation and preload nonlinearities. Journal of the Franklin Institute, Vol. 353, No. 16, pp. 4280-4299, 2016.
- [5] MAO, Y., DING, F.: A novel data filtering based multiinnovation stochastic gradient algorithm for Hammerstein nonlinear systems. Digital Signal Processing, Vol. 46, November 2015, pp. 215-225, 2015.
- [6] SHEN, Q., DING, F.: Least Squares Identification for Hammerstein Multi-input Multi-output Systems Based on the Key-Term Separation Technique. Circuits, Systems, and Signal Processing, Vol. 35, No. 10, pp. 3745-3758, 2016.
- [7] SHI, Z., WANG, Y., JI, Z.: A Multi-innovation Recursive Least Squares Algorithm with a Forgetting Factor for Hammerstein CAR Systems with Backlash. Circuits, Systems, and Signal Processing, Vol. 35, No. 12, pp. 4271-4289, 2016.
- [8] VÖRÖS, J.: Iterative algorithm for parameter identification of Hammerstein systems with two-segment nonlinearities. IEEE Trans. Automatic Control, Vol. 44, pp. 2145-2149, 1999.
- [9] WANG, X., DING, F., HAYAT, T., ALSAEDI, A.: Combined state and multi-innovation parameter estimation for an input non-linear state-space system using the key term separation. IET Control Theory & Applications, Vol. 10, No. 13, pp. 1503-1512, 2016.
- [10] CHEN, J., LU, X., DING, R.: Gradient-based iterative algorithm for Wiener systems with saturation and deadzone nonlinearities. Journal of Vibration and Control, Vol. 20, No. 4, pp. 634-640, 2014.
- [11] GIRI, F., ROCHDI, Y., CHAOUI, F.Z.: An analytic geometry approach to Wiener system frequency identification. IEEE Trans. Automatic Control, Vol. 54, pp. 683-696, 2009.
- [12] JANCZAK, A.: Instrumental variables approach to identification of a class of MIMO Wiener systems. Nonlinear Dynamics, Vol. 48, pp. 275-284, 2007.
- [13] JIN, Q., DOU, J., DING, F., CAO, L.: A novel identification method for Wiener systems with the limited information. Mathematical and Computer Modelling, Vol. 58, No. 7-8, pp. 1531-1539, 2013.
- [14] KAZLAUSKAS, K., PUPEIKIS, R.: On intelligent extraction of an internal signal in a Wiener system consisting of a linear block followed by hard-nonlinearity. Informatica, Vol. 24, pp. 35-58, 2013.
- [15] LIU, M.M., XIAO, Y.S., DING, R.F.: Iterative identification algorithm for Wiener nonlinear systems using the Newton method. Applied Mathematical Modelling, Vol. 37, pp. 6584-6591, 2013.
- [16] PAWLAK, M., HASIEWICZ, Z., WACHEL, P.: On nonparametric identification of Wiener systems. IEEE Trans. Signal Processing, Vol. 55, pp. 482-492, 2007.
- [17] PUPEIKIS, R.: On recursive parametric identification of Wiener systems. Information Technology and Control, Vol. 40, pp. 21-28, 2011.
- [18] VÖRÖS, J.: Parameter identification of Wiener systems with multisegment piecewise-linear nonlinearities. Systems and Control Letters, Vol. 56, pp. 99-105, 2007.
- [19] WACHEL, P., MZYK, G.: Direct identification of the linear block in Wiener system. Int. Journal of Adaptive

Control and Signal Processing, Vol. 30, No. 1, pp. 93–105, 2016.

- [20] WANG, D., DING, F.: Least squares based and gradient based iterative identification for Wiener nonlinear systems. Signal Processing, Vol. 91, pp. 1182-1189, 2011.
- [21] ZHOU, L., LI, X., PAN, F.: Gradient based iterative parameter identification for Wiener nonlinear systems. Applied Mathematical Modelling, Vol. 37, No. 16–17, pp. 8203–8209, 2013.
- [22] KALAŠ, V., JURIŠICA, L., ŽALMAN, M., ALMÁSSY, S., SIVIČEK, P., VARGA, A., KALAŠ, D.: Nonlinear and Numerical Servosystems. Alfa/SNTL, Bratislava, Slovakia, 1985 (in Slovak).
- [23] DONG, R., TAN, Q., TAN, Y.: Recursive identification algorithm for dynamic systems with output backlash and its convergence. Int. J. Appl. Math. Comput. Sci, Vol. 19, No. 4, pp. 631-638, 2009.
- [24] GIRI, F., ROCHDI, Y., BROURI, A., RADOUANE, A., CHAOUI, F.Z.: Frequency identification of nonparametric Wiener systems containing backlash nonlinearities. Automatica, Vol. 49, pp.124–137, 2013.
- [25] VÖRÖS, J.: Identification of cascade systems with backlash. International Journal of Control, Vol. 83, pp. 1117-1124, 2010.
- [26] BAI, E.W.: A blind approach to the Hammerstein-Wiener model identification. Automatica, Vol. 38, pp. 967-979, 2002.
- [27] BAI, E.W.: An optimal two-stage identification algorithm for Hammerstein-Wiener nonlinear systems. Automatica, Vol. 34, pp. 333-338, 1998.
- [28] CRAMA, P., SCHOUKENS, J.: Hammerstein-Wiener system estimator initialization. Automatica, Vol. 40, pp. 1543–1550, 2004.
- [29] WANG, D., DING, F.: Hierarchical Least Squares Estimation Algorithm for Hammerstein–Wiener Systems.

IEEE Signal Processing Letters, Vol. 19, No. 12, pp. 825-828, 2012.

- [30] YU, F., MAO, Z., JIA, M.: Recursive identification for Hammerstein–Wiener systems with dead-zone input nonlinearity. Journal of Process Control, Vol. 23, No. 8, pp. 1108-1115, 2013.
- [31] LJUNG, L., SÖDERSTRÖM, T.: Theory and Practice of Recursive Identification, MIT Press, Massachusetts, 1983.
- [32] CHIDAMBARAM, M.: Computer Control of Processes, CRC Press, New York, 2001.
- [33] GIRI, F., RADOUANE, A., BROURI, A., CHAOUI, F-Z.: Combined frequency-prediction error identification approach for Wiener systems with backlash and backlashinverse operators. Automatica, Vol. 50, No. 3, pp. 768– 783, 2014.
- [34] REYLAND, J., BAI, E.W.: Generalized Wiener system identification: General backlash nonlinearity and finite impulse response linear part. Int. Journal of Adaptive Control and Signal Processing, Vol. 28, No. 11, pp. 1174–1188, 2014.
- [35] VÖRÖS, J.: Modeling and identification of nonlinear cascade and sandwich systems with general backlash. Journal of Electrical Engineering, Vol. 65, No. 2, pp. 104–110, 2014.

Ing. Jozef Vörös, PhD.

Slovak Technical University Faculty of Electrical Engineering and Information Technology Institute of Robotics and Cybernetics Ilkovicova 3, 812 19 Bratislava, Slovakia

e-mail: jozef.voros@stuba.sk

KYBERNETICKÁ PODPORA PRI KOMPENZÁCII DIABETU

Martin Ernek, Tomáš Murgaš, Ján Murgaš

Abstrakt

Závažné ochorenia, niekedy nazývané aj civilizačné, majú v súčasnosti vysoký nárast a tak je to aj pri ochorení Diabetes mellitus, kde je veľmi dôležité udržiavať koncentráciu glukózy v krvi človeka v stanovených medziach. Pre túto činnosť existujú medicínske riešenia, ktoré je vhodné doplniť riadenou telesnou aktivitou pacienta. Aby táto aktivita bola čo najúčinnejšia musí mať predpísané parametre. Pre riešenie takéhoto problému je možné použiť rehabilitačné zariadenie s predprogramovaným potrebným profilom zaťaženia v čase. Príspevok opisuje návrh rehabilitačného zariadenia, ktoré je využiteľné aj pri stabilizácii glukózy v krvi človeka.

Kľúčové slová: Diabetes mellitus, kompenzácia glukózy v krvi, ergometer

Úvod

Civilizačné ochorenia sú v súčasnosti veľkou hrozbou a spôsobujú znižovanie kvality života ľudí trpiacich týmito ochoreniami. Diabetes mellitus (DM) je jedným z najzávažnejších civilizačných ochorení a postihuje stále širšie masy ľudí. Liečba pacientov s týmto ochorením je stanovovaná diabetológom, pričom neustále prebieha výskum príčin tohto ochorenia a teda aj forma liečby. Dôležité je udržiavanie koncentrácie glukózy v krvi pacientov v stanovených hraniciach. Základom kompenzácie DM (cukrovka) sú medicínske predpisy, ktoré sú kombinované odporúčaniami v oblasti stravovania a pohybovej aktivity pacienta. Dostupné pohybové aktivity stanoví lekár avšak sofistikované formy je výhodné podporiť technickými zariadeniami. Tieto zariadenia sú podobné ako moderné tréningové športovcov rehabilitačné zariadenia alebo zariadenia pacientov [3].

V súčasnosti nie je dostatočne analyzovaný vplyv pohybových aktivít pre kompenzáciu diabetu. Väčšinou ide o všeobecné odporúčania o význame pohybu pri liečbe diabetu. Budúcnosť tejto časti kompenzácie diabetických pacientov bude vo vytváraní kompenzačných centier, kde pod dozorom riadiacich systémov a lekára bude možné pre pacientov zabezpečiť účinnú kompenzáciu. Výber vhodnej pohybovej aktivity pri liečbe cukrovky môže mať zásadný význam, k čomu môže prispieť aj inteligentné rehabilitačné zariadenie. Inteligentné rehabilitačné zariadenie spolupracuje s pacientom tak aby čas venovaný tréningu bol maximálne efektívny a zameraný na konkrétne ochorenie. Návrh inteligentného rehabilitačného zariadenia bol súčasťou projektu "Centrum výskumu závažných ochorení a ich komplikácií" [1], ktorý bol zameraný aj na riešenie problematiky modelovania a kompenzácie diabetu [4], [5].

Návrh zariadenia na kompenzáciu DM

V uvedenom projekte boli skúmané dve koncepcie inteligentného rehabilitačného zariadenia pre efektívnu reguláciu glukózy v krvi. V tomto príspevku bude opísaná alternatíva s rotačným pohybom pacienta a zariadenie budeme uvádzať pod názvom bicyklový ergometer, skrátene len ergometer. Predstavuje jeden zo základných rehabilitačných strojov v celom rade medicínských odborov. Konštrukčne ide o zariadenie podobné stacionárnemu bicyklu. Rotačným pohybom dolných končatín sa vykonáva práca úmerná odporu ktorý je možné na stacionárnom bicykli zvoliť. Odpor je spravidla vytvorený jednoduchým konštrukčným systémom, ktorý brzdí hnacie koleso pomocou mechanického alebo magnetického prítlaku. Navrhnutý ergometer je vybavený snímačmi a pokročilým riadiacim systémom na platforme TwinCAT 3 spoločnosti Beckhoff. Zaťaženie pre rehabilitáciu je vyvíjané magneto-práškovou brzdou, ktorá je mechanicky spojená s pedálmi ergometra pomocou remenice. Modulárny riadiaci systém umožňuje snímanie sily, otáčok, teploty brzdy a tepu monitorovaného pacienta. Vybrané veličiny sú potom zobrazované počas rehabilitácie na dotykovom paneli umiestnenom na riadidlách ergometra. Komunikácia s dotykovým panelom s informáciami pre rehabilitovaného pacienta je zabezpečená bezdrôtovo. Počas rehabilitačného cvičenia sú zbierané údaje do databázy. Brzdenie

zotrvačníka zabezpečuje magneto-prášková brzda o výkone 65 Nm. Magneto-práškové brzdy sú jednoduchej konštrukcie a vzhľadom na svoju veľkosť produkujú vysoký brzdný moment v porovnaní napr. s hysteréznou brzdou. Princíp brzdy spočíva v elektrickom prúde prechádzajúcim cievkou, ktorá vytvára magnetické pole a priťahuje magnetické častice. Magnetická brzda sa skladá z dvoch rotorov rozdelených na vonkajší a vnútorný. Medzi rotormi je vzduchová medzera, kde sú rozptýlené magnetické častice. Magnetické pole prúdom vytvorené prechádzajúcim cievkou zarovnáva magnetický prášok do medzery. Čím väčší prúd prechádza cievkou, tým je pevnejšie spojenie medzi vnútorným a vonkajším rotorom. Výhody magnetickej brzdy:

brzdný moment nezávislý od otáčok za minútu

· vysoký rozsah brzdného momentu

Pre snímanie momentu na hnacom hriadeli je použitý rotačný snímač, ktorý odmeriava natočenie hriadeľa medzi vstupom a výstupom. Na vstupnej strane hriadeľa sa nachádza remenica a na výstupnej sa nachádza elektromagnetická brzda FAT 650. Rotačné snímače krútiaceho momentu so zbernými krúžkami posielajú signál v analógovej forme. Výstupný signál je v jednotkách mV/V a je prenášaný cez vysoko kvalitné, nízko opotrebovateľné zberné krúžky. Tieto senzory sú navrhnuté pre momenty od 1 Nm do 500 Nm.

Matematický model

Nakoľko ide o stacionárny bicykel budeme uvažovať, že naklonenie bicykla do strán a riadenia bude nulové. Stacionárny bicykel je náhrada bicykla, ktorý robí pohyb priamočiaro dopredu a nakláňa sa okolo osi *y*. Pre odvodenie modelu použijeme rovnicu opisujúcu priamočiary pohyb po naklonenej rovine [2]

$$m\frac{\mathrm{d}^2 \mathrm{s}}{\mathrm{d} \mathrm{t}} = \frac{\mathrm{P}}{\mathrm{v}} - \left[\mathrm{mgC} + \frac{1}{2}\mathrm{A}\rho\mathrm{C}_{\mathrm{D}}\mathrm{v}^2 + \mathrm{mg\sin\alpha}\right]$$

alebo po použití Laplaceovej transformácii:

$$m\ddot{s} = \frac{P}{v} - [mgC + \frac{1}{2}A\rho C_D \dot{s}^2 + mg\sin\alpha]$$

kde:

 α – uhol naklonenia, m – hmotnosť, C – koeficient valivého odporu, P – výkon hnacej sily, C_D – koeficient valivého odporu, A – čelný prierez telesa, ρ – hustota vzduchu, v- rýchlosť.

Z tejto rovnice je možné po menších úpravách namodelovať a vytvoriť model bicykla. Stavová veličina systému je dráha, z ktorej po derivácii dostaneme aktuálnu rýchlosť a uhol α akým je bicykel naklonený okolo osi *y*.

Porovnaním priebehov rýchlosti pre definovanej trati bola odskúšaná adekvátnosť modelu. Výsledky ukázali dostatočnú adekvátnosť pre ďalšie využívanie modelu.

Prototyp ergometra

Robotický rehabilitačný systém riadi priemyselný počítač Advantech UNO-2272G. Na meranie momentu a otáčok sú použité systémy v podobe rotačných členov medzi elektromagnetickou brzdou a menšou prevodovou remenicou. Na odmeranie presnej polohy pedálov je použitý inkrementálny snímač polohy. UNO-2272G je zabudovaná platforma ARP priemyselného počítača, ktorá dokáže skrátiť vývojový čas a ponúka širokú škálu sieťových rozhraní pre rozsiahle potreby rôznych projektov. UNO-2272G obsahuje technológie spoločnosti Intel Atom a poskytujú bohaté rozhranie



Obr.1 Simulačný model ergometra



Obr.2 Prototyp ergometra

Prototyp ergometra bol realizovaný na princípe regulovania brzdnej sily na poháňanej náprave. Týmto je možné docieliť, že osoba cvičiaca na tomto zariadení bude pociťovať silu pri rotačnom pohybe pedálov. Napodobní to pocit, akoby sa zariadenie pohybovalo na reálnej trati. Odpor, ktorý je cítiť je tvorený elektromagnetickým členom, ktorý sa nachádza na osi predného kolesa. Ide o magnetickú brzdu, ktorá vyvíja maximálny moment 65Nm. Brzda sa nachádza na prednom kolese, ktoré je spojené cez väčšie prevodové koleso pomocou hnacieho remeňa s menším prevodovým kolesom, ktorý je spojený s pedálmi. Medzi magnetickou brzdou a kolesom je umiestnený rotačný senzor momentu, ktorý meria veľkosť pôsobiaceho momentu na koleso. Na opačnej strane osi predného kolesa sa nachádza inkrementálny snímač polohy, pomocou ktorého je možné merať rýchlosť kolesa. Tieto snímače sú pripojené k riadiacej skrini, kde sa nachádzajú moduly riadiaceho systému značky Beckhoff a taktiež komunikačná jednotka EK1814 vstupno-výstupných modulov. Táto jednotka slúži pre komunikáciu vstupných a výstupných zariadení s riadiacim počítačom pomocou zbernice EtherCAT. Ďalej sa v riadiacej skrini nachádzajú moduly EL3104 a EL4104 pre analógové vstupy, ktoré zbierajú a transformujú signály do digitálnej podoby.



Obr.3 Riadiaci počítač prototypu

Brzdenie zotrvačníka zabezpečuje magnetoprášková brzda o výkone 65 Nm. Pre snímanie momentu na hnacom hriadeli je použitý rotačný snímač, ktorý odmeriava natočenie hriadeľa medzi vstupom a výstupom.



Obr.4 Rotačný senzor momentu [6]

Na meranie sily hnacieho remeňa bol použitý tenzonometrický snímač RMGZ223, ktorý sa vyznačuje požadovanými vlastnosťami: odstredivé sily nemajú vplyv na meranie, má široký rozsah od 6 do 300 N a je postavený na bezúdržbových ložiskách bez potreby ich mazania.

Riadiaci systém je cez sieťový kábel spojený s priemyselným počítačom Advantech UNO-2272G. Na tomto počítači je nainštalovaný systém TwinCAT, v ktorom je možné naprogramovať raidenie, diagnostikovať alebo sledovať všetky merané veličiny. Na naprogramovanie riadenia elektromechanického člena bol vybraný jazyk ST, čiže štruktúrovaný text, pretože tento jazyk sa najviac hodí pre zložitejšiu implementáciu programu.



Obr.5 Tenzometrický snímač sily RMGZ223

V časti okna Solutions Explorer sa v sekcii PLC a pod sekcii POUs nachádza program s názvom MAIN. Tento program je spúšťaný v riadiacom PLC ako hlavný a z neho sa ďalej spúšťajú ďalšie súčasti programu. Ďalej sa v tejto podsekcii nachádzajú aj funkcie modelBicykla a interpolacia a Sklon. Tieto funkcie sú vedľajšie a môžu byť spúšťané viackrát v iných programoch.

Funkcia Šklon slúži na implementáciu testovaného zaťaženia simulovaného sklonom trate pohybu bicykla.. Výstupom z tejto funkcie je vektor hodnôt výšky, v ktorej sa trať nachádza. K nemu prislúcha vektor dĺžky trate. To znamená, že každému bodu v určitom poradí vo vektore dĺžky existuje aj bod v tom istom poradí vo vektore výšky trate.

Režimy cvičení

Režim s konštantným odporom

V tomto režime užívateľ nastaví východiskovú polohu (spravidla v dolnej časti rozsahu pohybu plánovaného cviku) a požadovaný odpor. Brzda kladie konštantný odpor, a to bez ohľadu na silu, ktorú naň cvičiaca osoba prostredníctvom pedálov vyvíja. Znamená to, že kým vyvíjaná sila nedosiahne úroveň odporu, pedál zostáva v pôvodnej polohe. K ich pohybu dochádza až v okamihu, keď vyvíjaná sila presiahne úroveň nastaveného odporu. Zrýchlenie ako aj rýchlosť pedála a pohybu brzdy závisia okrem ich hmotnosti najmä od toho, o koľko presahuje vyvíjaná sila nastavený odpor. Zvýšené úsilie sa na tomto režime prejaví zvýšením rýchlosti. Keďže súčin rýchlosti a sily predstavuje výkon, závisí aj tento parameter silového cvičenia od vyvíjaného úsilia [1].

Režim s konštantnou rýchlosťou (izokinetický)

Po nastavení rozsahu pohybu definovaním východiskovej a konečnej polohy sa nastaví požadovaná rýchlosť pohybu, osobitne pre koncentrickú a excentrickú fázu. Okrem toho sa nastavuje trvanie fáz zrýchlenia na začiatku a spomaľovania na konci koncentrickej a excentrickej fázy ako aj štartovacia poloha (buď bližšia k telu pri flexii v kolenovom kĺbe, alebo ďalej od tela pri vystretých dolných končatinách). Ak sa zvolí východisková poloha bližšie k telu, po aktivácii režimu začne koncentrická fáza. Pedál zrýchli z východiskovej polohy na nastavenú rýchlosť a bez ohľadu na vyvíjanú silu sa pohybujú rovnomerne

smerom ku konečnej polohe. Tesne pred ňou vo vopred definovaných úsekoch spomalí na nulu, čím sa ukončí koncentrická fáza. Pedál prakticky okamžite začne excentrickú fázu. Opäť zrýchli, tentoraz však opačným smerom na definovanú rýchlosť a pohybuje sa konštantnou rýchlosťou k východiskovej polohe. Tesne pred jej dosiahnutím opäť vo vopred definovaných úsekoch spomalí na nulu a celý cyklus sa začne odznova, t.j. pedál opäť zrýchli na definovanú rýchlosť a v koncentrickej fáze smeruje ku konečnej polohe. Systém vykoná nastavený počet cyklov a pedál zostane stáť vo východiskovej polohe. Konštantnú rýchlosť, a to tak v koncentrickej ako aj excentrickej fáze pohybu zabezpečuje veľmi tesná spätnoväzobná regulácia brzdnej resp. hnacej sily, ktorá úplne kompenzuje sily, ktorými cvičiaca osoba pôsobí na pedále. Zvýšené úsilie sa v tomto režime neprejaví zrýchlením, ako to je v režime s konštantným odporom, ale prejaví sa zvýšením sily, ktorú registrujú pedálmi tenzometre medzi a elektromagnetickou brzdou. Vyššia sila pri konštantnej rýchlosti aj v tomto režime znamená vyšší výkon (súčin sily a rýchlosti). Keďže sval je schopný generovať vo fáze excentrickej kontrakcie vyššiu silu ako vo fáze koncentrickej, prejavia sa takéto krátke a rýchle fázy proti pohybu opakovanými výraznými zvýšeniami sily svalovej kontrakcie. V excentrickej fáze vedie k podobným silovým špičkám zaradenie krátkych úsekov s vyššou rýchlosťou brzdeného pohybu. Možno predpokladať, že opakované silové špičky výraznejšie aktivujú fyziologické mechanizmy svalovej kontrakcie ako pôsobenie proti klasickému pohybu. jednosmernému izokinetickému Za takýchto podmienok možno tiež očakávať, že adaptačné zmeny, vznikajúce pri systematickom vykonávaní takéhoto cvičenia budú mať či už kvantitatívne alebo kvalitatívne odlišný charakter ako pri cvičení v klasickom izokinetickom režime. Pedále sa v izokinetickom režime môžu pohybovať zhodne, opačne, prípadne ich pohyby možno naprogramovať úplne nezávisle [1].

Izometrický režim

Z hľadiska koncepcie zariadenia môže byť tento režim osobitnou formou buď režim s konštantným odporom alebo režimu izokinetického. V režime s konštantným odporom ho možno dosiahnuť tak, že sa vzvolenej polohe nastaví odpor, ktorý presahuje maximálnu silu, ktorú je cvičiaca osoba schopná vyvinúť. Za takýchto podmienok zostáva pedál v nastavenej polohe a systém indikuje vyvíjanú silu pri nulovej rýchlosti. Inou alternatívou, ako dosiahnuť izometrického režimu je nastavenie nulovej rýchlosti v izokinetickom režime. Tak ako pri pohybe, aj v tomto špecifickom prípade sa udržiava konštantná, t.j. nulová rýchlosť. Pokiaľ vyvíjaná sila nepresiahne brzdnú kapacitu elektromagnetickej brzdy, poloha pedálov zostane stála a režim svalovej práce si zachová izometrický charakter. Bez ohľadu na to, akým spôsobom sa zabezpečí stabilná poloha pedálov, zariadenie sa mení na izometrický dynamometer. Elektrické signály zo zabudovaných senzorov sa privádzajú do riadiaceho systému, čo umožňuje registráciu maximálnej izometrickej sily. Vzhľadom na vysokú

frekvenciu merania je možné deriváciou silovej krivky podľa času určiť aj tzv. maximálny silový gradient, ktorý sa považuje za dôležitý parameter silových schopností. Opakovaným meraním v rôznych polohách možno získať informácie o maximálnej izometrickej sile pri rôznych stupňoch flexie v kolenovom kĺbe [1].

Záver

V aktívnej rehabilitácii ako aj pri kompenzácii DM sú navrhované koncepcie zariadení s čoraz vyššou inteligenciou tak aby sa zvyšovala účinnosť týchto zariadení. Problémom je fakt. že tieto medziodborové výskumné aktivity nemajú dostatočnú podporu v spoločnosti. Viac podpory si nachádzajú v medicínskych vyšetrovaniach a pri zvyšovaní kvality života pacientov. Pokiaľ však v tejto oblasti nebude synergia technikov, kybernetikov, biológov a lekárov, zariadenia nezískajú potrebnú podporu v spoločnosti. Zatiaľ však v SR takéto kolektívy sú skôr výnimkou ako zavedenou praxou, čo je na škodu tých čo potrebujú podporu pri zvyšovaní kvality ich života.

Poďakovanie:

Článok je jedným z výstupov výskumnej práce projektu s názvom Centrum výskumu závažných ochorení a ich komplikácií, ITMS projektu: 26240120038. "Projekt je spolufinancovaný zo zdrojov EÚ. Podporujeme výskumné aktivity na Slovensku."

Literatúra

[1] Záverečná správa projektu: "Centrum výskumu závažných ochorení a ich komplikácií" ITMS 26240120038,2016, FEI STU v Bratislave.

[2] PARKS, J. Simulation of riding a bicycle using Simulink. [online]. Sacramento: California State University, 2010. [cit. 2016-11-10].

[3] HAMAR, D,ZEMKOVÁ. E,SCHICKHOFER. P.,GAŽOVIČ, O,BOHMEROVÁ., 2007. Alternatívne metódy rozvoja a posudzovania nervovosvalových funkcií., 2007. 106 s. ISBN 978-80-89197-79-7

[4] BÁTORA, Vladimír - TÁRNÍK, Marián -MURGAŠ, Ján - SCHMIDT, Singe - NØRGAARD, Kirsten - POULSEN, Niels Kjolstad - MADSEN, Henrik - JØRGENSEN, John Bagterp. Bihormonal model predictive control of blood glucose in people with type 1 diabetes. In *Proceedings of IEEE MSC 2014* : IEEE, 2014, s. 1693-1698. ISBN 978-1-4799-7409-2.

[5] TÁRNÍK, Marián - BÁTORA, Vladimír -JORGENSEN, John Bagterp - MIKLOVIČOVÁ, Eva - LUDWIG, Tomáš - OTTINGER, Ivan - MURGAŠ, Ján. Estimation of the carbohydrate to insulin ratio using empirical models. In Diabetes technology and therapeutics : 8th international conference on advanced technologies and treatments for diabetes. Paris, France, February 18-25, 2015. Vol. 17, No. 1 (2015), s. A-99. ISSN 1520-9156 (2015: 3 - JCR Best Q).

[6] https://www.fmstechnology.com/en/for e-sensors.php

Abstract

Serious diseases, sometimes called civilization diseases, are currently affecting more people, and so is Diabetes mellitus, where it is very important to keep the blood glucose concentration within the established limits. There are many medical solutions to keep the glucose concentration in limits and they should be complemented with physical activity of the patient. Physical activity has to fulfill required parameters to make it more effective. It is possible to solve this problem using a rehabilitation device with a preprogrammed necessary time dependent load profile. This paper describes the design of rehabilitation device, which is also useful in the stabilization of glucose concentration level in human blood.

Prof. Ing. Ján Murgaš, PhD Ing. Martin Ernek, PhD.

Slovenská technická univerzita v Bratislave Fakulta elektrotechniky a informatiky Ústav robotiky a kybernetiky Ilkovičova 3, 812 19 Bratislava Tel: +421 2 602 91731

Ing. Tomáš Murgaš

– Systémy priemyselnej informatiky, s.r.o. Kopčianska 18, 85101 Bratislava

KYBERNETICKÉ MODELOVANIE V ENERGETIKE

Tomáš Murgaš, Ivan Sekaj, Ján Murgaš

Abstrakt

Energetické procesy patria z pohľadu kybernetiky do kategórie spojitých a priestorovo rozľahlých systémov. Elektrizačná sústava (ES) vytvorená z viacerých národných sústav, ako je napríklad ES kontinentálnej Európy sa považuje za najrozsiahlejší umelo vytvorený systém so zložitou štruktúrou. Kybernetika, veda o modelovaní a riadení systémov dáva riešenia aj pre vytváranie simulačných modelov ES v spolupráci s odborníkmi energetikmi a elektrotechnikmi. Dôležitosť vytvárania simulačných dynamických modelov ES vzrástla po prijatí európskej legislatívy v roku 2016 [1]. Cieľom tohto príspevku je poukázať na význam modelovania pre riešenie problémov v energetike.

Kľúčové slová: Elektrizačná sústava, kybernetický model, počítačové modelovanie

Úvod

Z histórie kybernetiky vyplýva, že okrem toho, že zaviedla pojem informácia popri predtým zavedených pojmoch ako energia a hmota, zameranie orientovala od začiatku na modelovanie a riadenie dynamických systémov. V prvých fázach išlo najmä o modelovanie jednoduchších procesov v elektrotechnike a priemysle, ale neskôr sa modelovanie prenieslo aj do roviny čoraz zložitejších procesov. Veľkým impulzom pre rozvoj modelovania bol aj vývoj autopilotov a kozmický výskum. Najmodernejšie informačné technológie priniesli nové možnosti aj v oblasti modelovania najrôznejších procesov. Nové kvalitné modely spolu so superpočítačmi, ktoré majú veľký výpočtový výkon, prispeli aj k náhrade skutočných jadrových skúšok počítačovými simuláciami, čo prinieslo ohromný príspevok ku kvalite životného prostredia. Prognostické modely v rôznych oblastiach ekonómie, politiky, sociálnych procesov ale aj predpovede počasia našli svoje uplatnenie v každodennom živote. Modelovanie postupne preniká aj do vied o živej prírody a medicíny. Navrhnuté sú modely diabetes mellitus ako aj závažných civilizačných ďalších ochorení. V technických vedách, kde sa modely používajú už niekoľko desaťročí tiež nastal pokrok nielen prístupnosťou výpočtovej techniky vyšších výkonov, prístupnosťou stále dokonalejších ale aj simulačných prostredí. Aj v elektroenergetike sa požívajú stále dokonalejšie modely elektrizačných sústav, ktoré poskytujú možnosti analyzovať javy v sústave, ktoré vznikajú nezávisle alebo vplyvom zmien vyvolaných človekom.

V súčasnej energetike sa okrem moderných "klasických" elektrární objavuje stále viac malých lokálnych energetických jednotiek pracujúcich na rôznych fyzikálnych princípoch. Tým vznikajú veľké technické problémy pri riadení zložitej elektrizačnej sústavy, ktorá je rozložená v celom priestore kontinentálnej Európy. Zvyšuje sa riziko vzniku globálneho výpadky elektrickej energie typu blackoutu, spojeného s ohromnými ekonomickými stratami a technickými poruchami zariadení. Úloha riadenia ES je globálny problém a nielen z pohľadu systémových porúch ale aj z pohľadu obchodovania na trhu s elektrickou energiu.

Matematické modelovanie elektrizačných sústav

Riešenie otázok analýzy prechodných procesov v elektrizačných sústavách sa s rozvojom výpočtovej techniky prenieslo do oblasti použitia simulačných modelov a priame riešenie rovníc opisujúcich ES prešlo do úzadia. Hlavným dôvodom boli dva faktory a to zložitosť riešenia nelineárnych diferenciálnych rovníc ES ako aj vyššia dôvera výsledkov získaných zo simulačných systémov. Kybernetické modelovanie sa stalo základným predpokladom pre realizáciu výskumných zámerov v oblasti riadenia elektrizačných sústav

Medzi prvé modely vyžívané pri analýze vlastnosti turbogenerátorov ako aj pri návrhu riadenia týchto agregátov bol tzv. model synchrónneho generátora, ktorý spolupracuje so sieťou veľkého výkonu (obr.1).



Obr. 1 Zjednodušený model ES

SG	 – synchrónny generátor
Т	– turbína
BS	– budiaci systém
х	– reaktancia vedenia
Ug,Us,	- Napätie generátora, napätie siete
fs – fre	kvencia siete

Tento jednoduchý model bol používaný v zmysle zásad používaných v kybernetickom modelovaní o modelovaní len skúmaných vlastností systémov pri zanedbaní alebo zjednodušení ostatných vlastností modelovaných procesov – systémov.

Tieto modely boli v minulosti použité v stovkách vedeckých príspevkoch z oblasti riadenia ES. Ozvlášť boli používané pri návrhu riadiacich štruktúr SG ako regulátor statorového napätia, ale najmä pri overovaní účinnosti tlmenia systémových stabilizátorov (PSS). V minulosti boli navrhnuté mnohé algoritmy PSS, ktoré využívali nové metodiky z oblasti teórie automatického riadenia. Tieto algoritmy našli uplatnenie aj v reálnej praxi ako príklad možno uviesť najčastejšie využívané PSS, ktoré vychádzajú zo štruktúry 3B podľa IEEE (obr.2).



Obr. 2 Model PSS 3B

Model zložitej elektrizačnej sústavy

Zložité ES vznikajú najmä prepojením menších sústav národného charakteru. Význam prepojených sústav vzrástol najmä v súčasnosti vplyvom zabezpečenia vyššej miery stability ES, zlepšenia kvality elektriny ako aj zabezpečením trhu s elektrinou vo vnútri zložitej ES. Príkladom je ES kontinentálnej Európy, ktorá vznikla spojením národných sústav štátov a samostatných regiónov.

Dynamika zložitej ES je z pohľadu matematického modelovania systém nelineárnych diferenciálnych rovníc doplnený mnohými nelineárnymi algebrickými rovnicami. Riešenie úloh v takomto systéme metódou vyčlenenia objektu, ktorý je predmetom riešenia a zvyšok ES nahradiť sieťou veľkého výkonu má nižšiu kvalitu. Vyššiu kvalitu riešení poskytujú simulačné modely zložitejších ES, ktoré obsahujú viac zdrojov, silových prvkov, vedení aj záťaží. V týchto prípadoch už nie je možné použiť jednoduché modelovanie ES ale je potrebné vytvoriť dynamický model, kde aspoň určitá oblasť, napr. národná ES, je modelovaná podrobne a zvyšok zložitej sústavy môže byť modelovaný zjednodušene. Zjednodušenie sa najčastejšie realizuje tak, že je úplné v ustálených stavoch ES, čo znamená, že sú zabezpečené toky (činný a jalový výkon) a napätia v uzloch. Avšak aj v takomto prípade ide o model s veľmi vysokým rádom systému diferenciálnych rovníc opisujúcich model. Preto pri výstavbe simulačného modelu sa využívajú špecializované simulačné prostredia.

Simulačné prostredia pre modelovanie zložitých ES

Dôležitým faktorom pri modelovaní zložitých ES je správny výber simulačného prostredia, čo predurčuje aplikačnú formu, a teda aj mieru využitia v praxi. Európska snaha o vybudovanie globálnej Elektrizačnej sústavy EÚ si vyžaduje na úrovni národných elektrizačných sústav (ES) zabezpečiť výskum problémov zabezpečenia stability, kvality riadenia a bezpečnosti jednotlivých národných ES. Na splnenie tejto veľkej úlohy je potrebné vyriešiť viacero problémov, ktoré sú náplňou mnohých projektov. K riešeniu je nevyhnutné realizovať výskumné činnosti vedúce k získaniu špičkových nástrojov na analýzu prechodných procesov a ustálených stavov ES, ktorými sú napr. simulačné - počítačové modely ES, pričom okrem globálneho modelu je nevyhnutné mať k dispozície kvalitné počítačové modely základných stavebných prvkov týchto rozsiahlych ES.

Jednotlivé simulačné prostredia môžu byť realizované v jednofázovom prostredí ako je PSLF, PSS-E, Neplan, Powerfactory a mnohé iné. Moderné simulačné prostredie v oblasti elektrizačných sústav musí poskytovať:

- Modelovanie elektrických sietí
- Analýzy dynamických javov
- Simulácie prechodných javov.
- Výpočty ustálených stavov, zaťaženia, elektrických tokov a skratov
- Optimalizáciu a identifikáciu parametrov, modelovanie uzlov sústavy
- Dátový manažment a jednoduchý import a export aplikácií.
- Moderné grafické užívateľské prostredie
- Podporu veľkého množstva uzlov siete
- Optimalizáciu užívateľského prostredia pre špecifické úlohy.
- Podporu simulácií v reálnom čase.
- Podporu viacero programovacích jazykov

Realizácia výpočtov v trojfázových sústavách je menej častá vzhľadom na problémy s objemom výpočtov. V tejto oblasti sa môže požívať všeobecný simulačný systém Matlab od fy MatWorks.

Proti tvorbe simulačného modelu zložitej elektrizačnej sústavy sa teda najčastejšie využívajú vlastnosti špecializovaných simulačných prostredí. Tieto prostredia možno všeobecne charakterizovať:

- Umožňujú navrhovať modely, ktoré obsahujú tisíce prvkov aktívnych aj pasívnych,
- Výpočty prechodných procesov s veľkou rýchlosťou bez nutnosti prevádzkovania na superpočítačoch
- Poskytujú bohaté knižnice modelov aktívnych prvkov (turbogenerátory všetkých typov), pasívne prvky rôznych typov (záťaže, vedenia, transformátory a pod.)
- Poskytujú knižnice riadiacich a regulačných zariadení
- Sú ľahko osvojiteľné používateľom.

Dôležitou podmienkou pri výbere simulačného prostredia pre výstavbu modelu zložitej ES, napr. model ES SR so sieťami kontinentálnej Európy vrátane dodávaných a odoberaných výkonov v uzloch, je široká knižnica programových modulov. Hlavné skupiny programových modulov sú

 Moduly synchrónnych generátorov (SG) so štruktúrami podľa odporúčaní IEEE, ktoré boli vypracovaní odbornými komisiami].

V prílohe 1 je uvedený príklad knižničného modulu pre modelovanie synchrónneho generátora s hladkým rotorom.

- Moduly budičov SG tiež podľa odporúčaní IEEE
 [4] a v rozdelení na:
 - o jednosmerné budiče
 - o striedavé budiče nezávislé
 - o statické budiče

V prílohe 2 je uvedený príklad modelu statického budiča s regulátorom svorkového napätia tak aby bol modifikovateľný pre viac verzií tohto typu budiča. Dôležitou súčasťou budiaceho systému je aj systémový stabilizátor (PSS), ktorý môže mať veľmi rozmanitú štruktúru. Rozmanitosť je taká veľká, že v oblasti teoretických návrhov existujú stovky štruktúr, ale v praxi sa často implementuje len niekoľko typov. V ES SR sú najrozšírenejšie PSS 3B, ktorého základná štruktúra je uvedená na obr.2.

 Moduly turbín v knižniciach sú veľmi dôležité a mali by mať návrh vzorových parametrov pre turbíny rôznych výkonov, nakoľko praktické meranie parametrov turbín je takmer nemožné.

V prílohe 3 je uvedený odporúčaný model od IEEE pre tepelné turbíny [5], kde je zrejmé, že časové konštanty nie možné priamo merať v prevádzke.

 Knižnica simulačného prostredia by mala obsahovať tiež možnosti vytvárania modelov obmedzovacích systémov jednoduchou parametrizáciou knižničných modulov.

V prílohe 4 je uvedený odporúčaný model pre modelovanie fotovoltaickej elektrárne. Štruktúry riadiacich systémov na vyšších stupňoch riadenia možno tiež často nájsť ako knižničné moduly, ale vzhľadom na rozmanitosť sú pri vytváraní modelu zložitej ES programovaní užívateľom.

Záver

Význam dynamických modelov komponentov ES je zvýraznený vyhláškou EÚ [1], ktorá túto povinnosť stanovuje nielen pre prevádzkovateľov prenosových sústav ale aj pre všetkých účastníkov ES. Dôvodom je najmä potreba vytvárania reálnych plánov obrany a obnovy ES po veľkých systémových poruchách ako je napr. blackout, kedy sa veľké celky, prípadne aj celá ES kontinentálnej Európy môže dostať do beznapäťového stavu, čo by znamenalo bezprecedentné ekonomické straty a s veľkou pravdepodobnosťou aj ľudské a materiálne straty.

Poďakovanie:

Príspevok vznikol s podporou projektu VEGA

Literatúra

[1] vyhláška 631

[2] Ernek, M., Tárník, M.: Dynamika elektrizačných sústav. FEI STU v Bratislave, elektronické vydanie.

[3] J. Machowski, J. W. Bialek, and J. R. Bumby. Power system dynamics: stability and control. John Wiley & Sons, Ltd, 422:.

[4] IEEE Power and Energy Society. IEEE recommended practice for excitation system models for power system stability studies. *IEEE Std 421.5 -2016(Revision of IEEE Std 421.5-2005)*, pages 1– 207, Aug. 2016

[5] IEEE IEEE Power and Energy Society Task Force on Turbine-Governor Modeling. Dynamic models for turbinegovernors in power system

Abstract

From the cybernetics point of view, energy processes belong to the category of continuous and large systems. The power system network constructed from several national systems, such as the continental Europe power system network, is considered to be the largest artificially created system with a complex structure. Cybernetics, system modeling and control science also provides solutions for the creation of power system simulation models in collaboration with energy and electrical engineering experts. The importance of developing power system network simulation dynamic models has increased since the adoption of European legislation in 2016 [1]. The aim of this paper is to point out the importance of modeling for solving energy problems.

Ing. Tomáš Murgaš

Systémy priemyselnej informatiky, s.r.o. Kopčianska 14, 85101 Bratislava e-mail: tomas.murgas@syprin.sk

Prof. Ing. Ivan Sekaj, PhD Prof. Ing. Ján Murgaš, PhD

Slovenská technická univerzita v Bratislave Fakulta elektrotechniky a informatiky Ústav robotiky a kybernetiky Ilkovičova 3, 812 19 Bratislava





Para- meter	Popis
Sn	Nominálny zdanlivý výkon generátora [MVA]
Un	Nominálne napätie generátora [kv]
cos φ _n	Nominálny účinník generátora [-]
U _{max}	Maximálne dovolené statorové napätie generátora [kv]
Umin	Minimálne dovolené statorové napätie generátora [kv]
Rst	Rezistancia jednej fázy statorového vinutia [Ω]
X _{2n}	Spätná reaktancia v nenasýtenom stave [%]
Xdn	Synchrónna reaktancia v pozdĺžnej osi v nenasýtenom stave [%]
Xds	Synchrónna reaktancia v pozdĺžnej osi v nasýtenom stave [%]
X _{qn}	Synchrónna reaktancia v priečnej osi v nenasýtenom stave [%]
Xqs	Synchrónna reaktancia v priečnej osi v nasýtenom stave [%]
X [´] dn	Prechodná reaktancia v pozdĺžnej osi v nenasýtenom stave [%]
X [´] ds	Prechodná reaktancia v pozdĺžnej osi v nasýtenom stave [%]

X´qn	Prechodná reaktancia v priečnej osi v nenasýtenom stave [%]
X´qs	Prechodná reaktancia v priečnej osi v nasýtenom stave [%]
X´´dn	Rázová reaktancia v pozdĺžnej osi v nenasýtenom stave [%]
X´´ds	Rázová reaktancia v pozdĺžnej osi v nasýtenom stave [%]
X ^{′′′} qn	Rázová reaktancia v priečnej osi v nenasýtenom stave [%]
X´´qs	Rázová reaktancia v priečnej osi v nasýtenom stave [%]
X 0	Netočivá reaktancia [%]
X ₁	Rozptylová reaktancia rotora [%]
Ta	Časová konštanta jednosmernej zložky [s]
Τ́d	Prechodná časová konštanta pre pozdĺžnu os [s]
T´´d	Rázová časová konštanta pre pozdĺžnu os [s]
Τ́q	Prechodná časová konštanta pre priečnu os [s]
Τ΄́q	Rázová časová konštanta pre priečnu os [s]
Н	Zotrvačnosť
D	TImenie

Príloha č.2: Štruktúra knižného modulu pre modelovanie budiaceho systému



Parameter	Popis
T _R	Časová konštanta filtra napätia
VIMAX	Maximálna regulačná odchýlka
V _{IMIN}	Minimálna regulačná odchýlka
T _A	Časová konštanta Lag člena
Тв	Časová konštanta Lag člena
K _A	Zosilnenie regulátora
T _A	Časová konštanta Lead člena
V _{RMAX}	Horné obmedzenie budiaceho napätia
V _{RMIN}	Dolné obmedzenie budiaceho napätia
Kc	Konštanta nasýtenia budiča od budiaceho prúdu
K _F	Spätná väzba od budiaceho napätia - zosilnenie
TF	Spätná väzba od budiaceho napätia – časová konštanta filtra

Príloha č.3: Štruktúra knižného modulu pre modelovanie parnej turbíny



Par.	Popis					
P _{m1}	Mechanický výkon turbíny pre generátor 1					
P _{m2}	Mechanický výkon turbíny pre generátor 2					
CV, IV	Riadiace, záchytné ventily					
Pt	Tlak pary z kotla					
T ₄ , T ₅ , T ₆ , T ₇	Časové konštanty reprezentujúce časti turbíny					
k₁, k₃, k₅, k⁊	Zosilnenia uvádzajúce podiel na výkone príslušnej časti turbíny pre generátor 1					
k2, k4, k6, k8	Zosilnenia uvádzajúce podiel na výkone príslušnej časti turbíny pre generátor 2					
m _s , m _r	Prietok pary z kotla, prehrievača					
HP, LP	Prepojenie vysokotlakej, nízkotlakej časti s generátorom 1,2					



Príloha č.4: Štruktúra knižného modulu pre modelovanie fotovoltaickej elektrárne



Param	Popis					
pqflag	Priorita pre jalový výkon 0 alebo činný výkon					
Xc	Kompenzačná reaktancia					
qmin	Dolné obmedzenie želanej hodnoty jalového výkonu					
qmax	Horné obmedzenie želanej hodnoty jalového výkonu					
V ₀	Dolné obmedzenie necitlivosti napäťovej charakteristiky					
V1	Horné obmedzenie necitlivosti napäťovej charakteristiky					
dqdv	Napäťová charakteristika					
fbdb	Necitlivosť regulátora pre nadfrekvenciu					
ddn	Zosilnenie regulátora výkonu , väzba od frekvencie					
I _{max}	Obmedzenie prúdu					
Vt0	Krivka pre odpínanie podľa napätia bod 0					
Vt1	Krivka pre odpínanie podľa napätia bod 1					
V _{t2}	Krivka pre odpínanie podľa napätia bod 2					
Vt3	Krivka pre odpínanie podľa napätia bod 3					
vrflag	Nastavenie ochrany od napätia 0 – vypínanie, 1 samoresetovanie					
fto	Krivka pre odpínanie podľa frekvencie bod 0					
f _{t1}	Krivka pre odpínanie podľa frekvencie bod 1					

BIOCYBERNETICS IN DIABETES EDUCATION

Marián Tárník, Eva Miklovičová

Abstract

Paper presents an empirical model of glucose metabolism in type 1 diabetes. Model can be used as a diabetes simulator for diabetes education. Model is parameterized using clinical characteristics and parameters such as an insulin sensitivity, carbohydrate ratio and insulin on board curves. In comparison with other relevant empirical models, proposed model allows to reflect the variance of the clinical parameters independently on meal or insulin bolus events.

Keywords: empirical model, diabetes, education, T1DM

Introduction

Type 1 diabetes mellitus (T1DM) is associated with insulin deficiency in the body caused by defects in the insulin secretion by the pancreatic beta cells. The result is an inability of the body to autonomously regulate the blood glucose concentration.

People with T1DM have to replace an endogenous insulin secretion by external insulin administration. The idea is to mimic the physiological insulin secretion pattern of a non-diabetic person. The most common therapy is the so-called basal-bolus therapy which consist of two types of insulin infusion rates (two types of insulin doses).

Basal insulin is the low-rate infusion which mimics the slow basal insulin secretion of a non-diabetic person throughout the day. Bolus (prandial) insulin, i.e. insulin administered at mealtimes, provides an increased glucose uptake during and after meals. Both the basal insulin dose and the bolus insulin dose have to be properly adjusted to maintain normoglycemia.

Outlined task can be seen as a control problem where the blood glucose concentration is the output of the controlled system. Insulin infusion is the controlled system input and the meals act as disturbance on the controlled system. A significant research effort is being directed towards the Artificial Pancreas (AP), which is device capable to cope with this control problem, see paper [1]. One of the main issues is an control algorithm, for example see paper [2] and references therein.

Nevertheless, the human factor is still crucial in the issue of an appropriate insulin bolus size calculation for a given meal and it does not matter if it is in the case of manual dosing or the case of configuring the automated closed loop system (AP). Therefore, the diabetes education always plays an important role.

This paper addresses the use of the theoretical tools associated with the Artificial Pancreas research for the purpose of diabetes education. The mathematical models of T1DM, which are suitable for AP design, can be also used as an T1DM simulator with focus on diabetes education as is shown in this paper. Particularly, empirical models, which are typically control-oriented, can have a structure which directly includes clinical parameters such as carbohydrate ratio (CR) (also often referred as carbohydrate to insulin ratio or insulin to carbohydrate ratio) and an insulin sensitivity (IS). It can be shown that the unit of insulin sensitivity is [mmol/l/U] and the carb. ratio unit is [g/U].

Empirical Model

T1DM models can be divided into several groups as have been discussed in paper [3]. The models based on physiology, are focused at giving insight into the metabolism of the average T1DM subject. An interpretation of the model parameters in terms of clinical parameters, such as CR and IS, is often not straightforward.

The black-box models, see [4], are usually designed using data-based methods. The main application is a short-term prediction of glycemia. The main drawback is that the model output is given predominantly by an auto-regressive part of the model thus the prediction of glycemia hardly takes the insulin and carbohydrates into account.

The individualised empirical models have been proposed to describe the glycemia dynamics for T1DM subjects, see [4]. These models have several common features. The model structure is simple. It is not derived from a detailed physiological analysis. The model is continuous-time system with only few adjustable parameters which allow the parameter, i.e. the static gain, time constants or impulse responses to have a clear physiologic interpretation. The resulting model has clinician friendly structure and is control oriented.

Authors have proposed an empirical model suitable for simple diabetes simulator as follows:

$$\dot{x}_{i1}(t) = -\frac{1}{T_i} x_{i1}(t) + u_i(t) \tag{1}$$

$$\dot{x}_{i2}(t) = -\frac{1}{T_i} x_{i2}(t) + \frac{1}{T_i} K_i x_{i1}(t)$$
(2)

$$y_i(t) = e x_{i2}(t)$$
 (3)

$$\dot{x}_{c1}(t) = -\frac{1}{T_c} x_{c1}(t) + u_c(t)$$
(4)

$$\dot{x}_{c2}(t) = -\frac{1}{T_c} x_{c2}(t) + \frac{1}{T_c} K_c x_{c1}(t)$$
(5)

$$y_c(t) = e x_{c2}(t) \tag{6}$$

$$G(t) = G_{b} + y_{c}(t) - y_{i}(t)$$
 (7)

where G(t) [mmol/I] is an output glycemia. Signals $x_{i1}(t)$, $x_{i2}(t)$, $x_{c1}(t)$, $x_{c2}(t)$, $y_i(t)$ and $y_c(t)$ are model internal signals and e is the Euler constant. Signals $u_i(t)$ [U/min] and $u_c(t)$ [g/min] are insulin and carbohydrate input rates respectively, T_i [min] and T_c [min] are time constants. Consequently K_i [mmol/I/U] corresponds to IS, K_c [mmol/I/g] is the carbohydrate sensitivity and K_i/K_c [g/U] corresponds to CR.

Two subsystems of the model can be identified. First an insulin subsystem and second a carbohydrate subsystem. This overall structure of the empirical model is depicted in fig. 1.





For the proposed model the effect of the model parameters can be illustrated as follows. In fig. 2 the effect of insulin subsystem parameters, i.e. K_i [mmol/l/U] and T_i [min], is shown. Simulated subject is kept at basal glycemia 7 [mmol/l] with corresponding non-zero basal insulin infusion rate. At the time t=1 [h] an insulin bolus is administered. Resulting glycemia response is given by the insulin subsystem parameters.



Fig.2 Effect of the insulin subsystem parameters.





Fig.3 Effect of the carbohydrate subsystem parameters.

As can be seen in the model equations the parameter K_i does have effect on model output even if the input signal $u_i(t)$ is zero. This is illustrated in fig. 4.



Fig.4 Example of time variant paremeter K_i (step change at time t=10 [h]). Simulation with insulin bolus given at t=9 [h] (5 min. impulse of the signal u_i(t)).

It is clear, that the two effects, the effect of insulin and the effect of carbohydrates, can cancel each other, which will result in no changes in simulated glycemia. In general, the peaks of each effects are given by the model parameters as is depicted in the figures 2 and 3. These parameters can be determined using a priori clinical patient characteristics.

Parameters K_i and K_c , as has been mentioned, are related to the clinical parameters IS and CR. The time constants T_i and T_c can be also related to the known information from everyday practice. Time constant T_c can be related to the type of meal, i.e. to the rate of carbohydrate absorption from the particular type of meal.

The time constant T_i can be based on the known IOB characteristics. IOB stands for insulin on board, which is an information on the remaining amount of bolus insulin which has not been actively used in the body. Clearly, IOB as used in everyday practice is related to the pharmacodynamics (PD) of the particular insulin. For example, IOB curve of insulin Aspart based on available PD data [5] is shown in fig. 5. As can be seen in fig. 5 this IOB curve corresponds to the real IOB curve (pump data) as implemented in the bolus calculator of particular insulin pump.



Fig.5 IOB curves (active insulin) for bolus given at t = 0. Curve based on pharmacodynamics (insulin Aspart) is label as *PD*. IOB curve as implemented in real insulin pump is labeled as *pump data*. Solid line is the curve given by the proposed model.

In order to relate the IOB curve to the parameter T_i of the proposed empirical model, we propose the IOB curve approximation (model) in the form

$$IOB(t) = e^{-\frac{1}{T_i}t} \left(1 + \frac{1}{T_i}t\right) \times 100\%$$
(8)

where IOB(t) [%] is remaining amount of bolus insulin and T_i is used to parametrize this model. We propose formula

$$T_i = \frac{I_{IOB}}{5} \tag{9}$$

to determine the value of T_i where T_{IOB} is time when IOB from a given bolus decreases below 5 %. Resulting IOB model approximation is shown in fig. 5.

Simulator for Diabetes Education

In this section we present T1DM simulator which uses the empirical model presented in the previous section. Purpose of the simulator is to illustrate the basic principles of bolus calculators. As have been mentioned the parameters K_i and K_c are related to the clinical parameters IS and CR, which are typically used in bolus calculators, see [6].

Simple bolus calculation rules can be summarized as follows. For example, glycemia is 7.5 [mmol/l] which is within the target interval 5.6 -- 7.5 [mmol/l]. Therefore, the total bolus estimate is based only on the carbohydrate amount as follows:

bolus=
$$\frac{\text{carb.}}{\text{CR}} = \frac{45 \text{ [g]}}{12 \text{ [g]}} \approx 3.7 \text{ [U]}$$
 (10)

where the CR value is 12 [g/U]. For the case when the blood glucose (BG) level is greater than maximum of target interval (max.BG), the calculation rule includes so called correction bolus

$$bolus = \frac{carb.}{CR} + \frac{BG-max.BG}{IS}$$
(11)

For example, if BG = 9.7 [mmol/l] and IS = 4 [mmol/l/U] and carb. intake is 50 [g] we have

$$bolus = \frac{50}{12} + \frac{9.7 - 7.5}{4} \approx 4.7 [U]$$
(12)

In this case the first step is to identify the parameters of the empirical model. For this purpose a CGM (continuous glucose monitoring) data of the particular person with diabetes are used. The data used in this work have been collected in everyday conditions from one T1DM subject (male, age 14, body weight BW = 64.6 [kg], height h = 1.85 [m]). The subject is on insulin pump therapy and uses the bolus calculator which is a feature of the pump to calculate the bolus. The bolus calculator uses following parameters. The insulin sensitivity is IS = 4 [mmol/I/U] for whole 24 hour period. The carbohydrate ratio CR = 6 [g/U] for the period 5:30 -- 9:00 and CR = 12 [g/U] for the rest of the day. All these parameters have been determined by the healthcare professional.

Apart from the insulin pump, the patient uses a glucose sensor with CGM system. The patient has estimated the carbohydrate content of the meals used in the bolus calculator. The pump has logged all information related to the insulin bolus.

The data from one day are used in this case. These data are shown in fig. 6.



Fig.6 Result of the identification of empirical model parameters - graphical comparison of simulated glycemia G(t) and measured CGM (continuous glucose monitoring) data. Time t = 0 [min] corresponds to 05:00 AM.



Fig.7 Result of the identification of empirical model parameters - graphical comparison of simulated glycemia G(t) and measured CGM (continuous glucose monitoring) data. Time t = 0 [min] corresponds to 05:00 AM.

The unknown parameters of the model are identified so that the estimation error between the measured glycemia data vector y_{CGM} and the simulated glycemia vector y is minimized in a least-squares sense. A derivative free optimization method, namely the Nelder-Mead method or downhill simplex method, is used to minimize the objective function. Results are summarized in the table 1 and in fig. 6. A time variant parameter K_i had to be used in order to achieve satisfactory results which corresponds to the a priori information available.

Begin time [min]	$ \begin{bmatrix} K_i \\ \frac{\text{mmol}}{\text{l/U}} \end{bmatrix} $	$\frac{K_i/K_c}{[g/U]}$	$ \begin{bmatrix} K_c \\ \frac{\text{mmol}}{\text{l/g}} \end{bmatrix} $	$\frac{T_i}{[\min]}$	T_c [min]
0	2.748	10.890	0.252	17.2	26.9
30	1.097	4.346	0.252	17.2	26.9
240	2.648	10.496	0.252	17.2	26.9
870	2.748	10.890	0.252	17.2	26.9

Tab.1 Identified parameters of empirical model used in T1DM simulator. Glycemia basal value: $G_b = 5.6$ [mmol/l].

Once the simulator which corresponds to the real T1DM data is obtained, as shown in fig. 6, it can be provided with any inputs, i.e. insulin boluses and CHO ingestion, and the resulting simulated glycemia can be used for diabetes education purposes.

In this case simulator is used to provide an information on glycemia for the simulated bolus calculator with same settings as in the case of real T1DM subject. Same CHO ingestion is considered as well. Simulation results are shown in fig 7.

Simulated bolus insulin amounts are comparable with the original T1DM subject data, see bottom panels in fig. 6 and in fig. 7. Resulting simulated glycemia is also comparable with original data. Simulator would be able to simulate other sequence of CHO ingestions so that the resulting simulated glycemia can be used as an example in diabetes education.

At time t = 635 [min] in fig. 7 the insulin bolus is deliberately omitted. Resulting glycemia related to corresponding meal event reaches, as expected, higher levels, i.e. simulator illustrates the situation with missing insulin bolus.

Conclusions

The paper briefly introduces relevant empirical model for simple diabetes simulator.

The need for changes in initial model structures in order to allow effective time variant parameters is presented. A modified empirical model structure with emphasis on time variant parameters is proposed. Model maintains an advantages of T1DM empirical models, i.e. small number of parameters, control oriented, and clinician friendly and simple structure. Proposed model also maintains all properties required in applications such as estimation of clinical parameters IS and CR as presented in our previous work, see [3].

Presented T1DM simulator with proposed empirical model has been implemented in a form of web application. It has been developed in Python programming language using primarily the Python SciPy libraries. Django has been used as web framework. The main goal is the support of diabetes education.

Acknowledgment

The paper is one of the outcomes of the research work for the project entitled "Research center for severe diseases and related complications", "ITMS: 26240120038". "This project is being co-financed by the European Union. We support research activities in Slovakia".

References

[1] Russell, S.J., Hillard, M.A., Balliro, C., Magyar, K.L., Selagamsetty, R., Sinha, M., Grennan, K., Mondesir, D., Ehklaspour, L., Zheng, H., Damiano, E.R., and El-Khatib, F.H. (2016). Day and night glycaemic control with a bionic pancreas versus conventional insulin pump therapy in preadolescent children with type 1 diabetes: a randomised crossover trial. The Lancet Diabetes & Endocrinology, 3(4).

[2] Batora, V., Tarnik, M., Murgas, J., Schmidt, S., Norgaard, K., Poulsen, N.K., Madsen, H., and Jorgensen, J. (2015). Bihormonal control of blood glucose in people with type 1 diabetes. In ECC 2015 : European Control Conference. Linz, Austria. July 15-17, 2015.

[3] Tarnik, M., Batora, V., Jorgensen, J.B., Boiroux, D., Miklovicova, E., Ludwig, T., Ottinger, I., and Murgas, J. (2015). Remarks on models for estimating the carbohydrate to insulin ratio and insulin sensitivity in t1dm. In 14th European Control Conference, Linz, Austria, July 15-17, 2015.

[4] Cescon, M. and Johansson, R. (2009). Glycemic trend prediction using empirical model identification. In Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference on, 3501–3506. doi: 10.1109/CDC.2009.5400219.

[5] Mudaliar, S., Lindberg, F.A., Joyce, M., Beerdsen, P., Strange, P., Lin, A., and Henry, R.R. (1999). Insulin aspart (b28 asp-insulin): A fast-acting analog of human insulin. Diabetes Care, 22(9), 1501–1506.

[6] Schmidt, S. and Nørgaard, K. (2014). Bolus calculators. Journal of Diabetes Science and Technology, 8(5), 1035–1041. doi:10.1177/1932296814532906.

Ing. Marián Tárník, PhD.

Slovenská technická univerzita v Bratislave Fakulta elektrotechniky a informatiky Ústav robotiky a kybernetiky Ilkovičova 3 812 19 Bratislava marian.tarnik@stuba.sk

doc. Ing. Eva Miklovičová, PhD.

Slovenská technická univerzita v Bratislave Fakulta elektrotechniky a informatiky Ústav robotiky a kybernetiky Ilkovičova 3 812 19 Bratislava eva.miklovicova@stuba.sk

SYSTÉM LOKALIZÁCIE A MAPOVANIA PRE DRONY

Andrej Babinec, Jozef Rodina, Eduard Mráz

Abstrakt

Tento článok sa zaoberá problematikou, ktorá tvorí jeden z hlavných pilierov robotiky - lokalizáciou. Konkrétne predstavuje lokalizačný systém, ktorý je možné použiť pri lietajúcich robotoch. Špecifikom, ktorým tento systém disponuje, je fakt, že je určený do vnútorných priestorov. V praxi to predstavuje vylúčenie použitia klasických GNSS systémov (akým je napríklad GPS) pri lokalizácii. Cieľom článku nie je len vysvetlenie princípu fungovania jadra tohto systému, či ukážka výsledkov. Článok totiž obsahuje aj vzorovú aplikáciu lokalizačného systému - mapovanie, alebo inak - 3D rekonštrukciu obrazu. Je to síce iná problematika ako lokalizácia, no v konečnom dôsledku spolu veľmi úzko súvisia.

Kľúčové slová: SLAM, LOKALIZÁCIA, UAV, DRONY

Úvod

Drony (pozn. v kontexte tohto článku drony chápme ako bezpilotné lietajúce roboty) sú už dnes väčšine svetovej populácie známe. Sú globálne rozšírené, a čo je dôležité, dostupné aj pre bežných ľudí. Drony spadajú do podkategórie robotiky - mobilná robotika. A teda, ak má ktorýkoľvek mobilný robot vykonávať úlohy, ktoré vyžadujú istú mieru autonómnosti, bude musieť s veľmi veľkou pravdepodobnosťou disponovať lokalizačným systémom. Pod autonómnou úlohou si môžeme predstaviť napríklad pravidelný prelet po vopred stanovenej trajektórii a vyhotovenie fotografií v určitých bodoch tejto trajektórie. Práve takéto aplikácie sú pre nás smerodajné pri vývoji nášho lokalizačného systému. Predpokladáme, že náš lokalizačný systém bude tvoriť dobrý základ pre celkový, komplexný riadiaci systém drona, ktorý umožní kompletné plánovanie autonómnych úloh. V robotickej terminológii - tento systém bude poskytovať základ pri riešení ďalšej problematiky mobilnej robotiky, ktorou je navigácia.

Dnes už poznáme mnoho aplikácií, kedy sú drony používané na rôzne druhy inšpekcií. Treba podotknúť, že aktuálne ponúkané riešenia sú často neautonómne a vyžadujú priame ovládanie človekom. Pre predstavu uvádzame možné aplikácie firmy AIROBOTICS[1], keďže celkom dobre pokrývajú spektrum možných úloh pre drony: inšpekcia, mapovanie, bezpečnosť a iné.

Treba povedať, že riešenia firmy AIROBOTICS sú automatizované, teda autonómne. Avšak z nám dostupných informácií o ich službách je zrejmé, že autonómne fungujú len vo vonkajších priestoroch - za použitia GNSS systémov. Pri tejto informácii sa dostávame k podstate našej práce, a teda prečo vidíme význam vo vývoji takéhoto indoor lokalizačného systému (absencia GNSS). Indoor lokalizácia je pomerne obšírna problematika v zmysle, že existuje niekoľko metód, ako ju riešiť. Napríklad: systém POZYX[9], vizuálna odometria, lokalizácia pomocou WIFI, iné typy odometrie (na základe IMU jednotky, alebo enkóderov z kolies pri robotoch s kolesami), magnetická lokalizácia a iné. Pre náš systém sme si ako stavebný prvok vybrali vizuálnu odometriu pomocou hĺbkovej kamery. Na záver úvodnej časti si len priblížime filozofiu fungovania tohto systému, k vysvetleniu jednotlivých pojmov a metód sa dostaneme v jadre článku. Pri GNSS systémoch poznáme pojem absolútna poloha, to znamená, že kdekoľvek na planéte Zem máme svoju unikátnu, jednoznačne určenú polohu. Pri indoor lokalizácii predstavuje našu "Zem" budova, alebo ktorýkoľvek iný vnútorný priestor, v ktorom sa pokúšame lokalizovať. Aby sme jednoznačne a unikátne vedeli určiť polohu v takomto priestore, je dobré urobiť si jeho mapu a určiť si počiatok súradnej sústavy, v ktorej sa chceme lokalizovať (v rámci mapy). Týmto však nastáva nový problém - ako sa lokalizovať a mapovať súčasne? Tento problém je v súčasnosti intenzívne skúmaný v rámci robotickej komunity a dostal pomenovanie SLAM (simultánna lokalizácia a mapovanie, z ang. simultanneous localization and mapping). Jadro nášho systému je postavené na metóde ORB-SLAM2[8]. Mapovanie vnútorného priestoru bolo spomenuté preto, aby bol jasný nasledovný fakt: ORB-SLAM2 dosahuje nulový drift odometrie pri už zmapovaných priestoroch[8].

1. Návrh systému

Táto kapitola predstavuje jadro článku. Tu odhalíme všetky nástroje a metódy používané v našom systéme. Taktiež pre ilustráciu uvádzame architektúru celého systému (obr. 1).

1.1 Použitý hardvér

Je všeobecne známe, že systémy, ktoré sú postavené na báze počítačového videnia potrebujú pomerne veľký výpočtový výkon. Samozrejme bavíme sa o aplikáciách v mobilnej robotike, čo znamená, že sme vo veľkej miere obmedzení aj rozmermi. Firma NVIDIA v tomto poskytuje riešenie v podobe výpočtového modulu s označením Jetson TX2[2]. A teda ako hlavnú výpočtovú jednotku sme zvolili modul NVIDIA JET-SON TX2.

Senzorické vybavenie lokalizačného systému v našom prípade zatiaľ pozostáva z RGB-D kamery Intel Realsense D435, ktorú vidíme na obr. 1 vľavo hore. Takýto (RGB-D) typ kamery poskytuje okrem klasickej farebnej snímky



Obr. 1 Architektúra lokalizačného systému

taktiež informáciu o hĺbke každého pixelu (pozn. RGB-D, D na konci označuje depth - hĺbku). Túto informáciu vieme využiť pri samotnej lokalizácii, alebo pri 3D rekonštrukcii priestoru. Význam tejto voľby vysvetlíme neskôr v tomto článku, konkrétne v sekcii 1.3.

Samotné hardvérové vybavenie drona nebudeme uvádzať, pretože toto nie je súčasťou našej práce. Jediná časť, ktorá je z tohto celku pre nás podstatná je zariadenie Pixhawk, ktoré predstavuje autopilot pre dron. Autopilot v tomto kontexte znamená, že zabezpečuje riadenie drona, obsahuje základné senzorické vybavenie nutné pre let (IMU jednotka, kompas atď.), taktiež disponuje rôznymi perifériami pre komunikáciu s inými zariadeniami.

Ďalej už toto fungovanie rozvinieme v nasledujúcej časti, kde celú architektúru popíšeme z hľadiska softvéru. Veríme, že takýto opis poskytne dobrú predstavu o celkovej funkcionalite nášho systému.

1.2 Architektúra systému

V tejto časti článku uvádzame jednotlivé moduly nášho systému. Vysvetľujeme ich význam v celkovom kontexte nášho systému.

Ako hlavný stavebný prvok sme zvolili Robot Operating System (skr. ROS). ROS je flexibilný framework pre písanie robotického softvéru. Je to kolekcia nástrojov, knižníc a konvencií[10]. Okrem iného nám ROS ponúka medziprocesnú komunikačnú vrstvu tzv. middleware. Pomocou nej môžeme prepojiť jednotlivé balíčky, ktoré ROS ponúka, a ktoré konkrétne uvádzame v ďalšej časti článku.

1.2.1 Opis jednotlivých modulov lokalizačného systému

Táto časť sa vo svojej podstate zaoberá opisom všetkých častí (okrem hardvéru spomínaného v časti 2.1), ktoré môžeme vidieť na obr. 1.

- realsense2_camera, balíček, ktorý komunikuje s kamerou. Stará sa o zber údajov z kamery a taktiež s jeho pomocou môžeme nastaviť parametre kamery.
- rtab_map, balíček, ktorý poskytuje nástroje na 3D rekonštrukciu priestoru a podporuje priamu integráciu s metódou orb-slam2, ktorú spomíname v úvode.
- mavros, stará sa o komunikáciu s Pixhawk autopilotom.
 Zberá množstvo dát z jeho senzorov. Dokážeme

pomocou neho priamo posielať riadiace príkazy autopilotu, ktorý potom dokáže riadiť samotný hardvér drona. V prípade autonómnej misie by sme doňho posielali polohu, ktorá je určená naším lokalizačným systémom.

 robot_localization, momentálne nevyužívaný balíček. Je to však predpríprava na vývoj robustného lokalizačného systému, ktorý pozostáva z fúzie niekoľkých lokalizačných zdrojov, či senzorov. Tento balíček poskytuje nástroje, ako napríklad Kalmanov filter, pomocou ktorého vieme takúto fúziu viacerých zdrojov uskutočniť.

1.2.2 Sekvenčný opis fungovania systému

Písmená v tomto zozname prislúchajú písmenám v kruhu na obr. 1:

- A Sú zozbierané surové dáta z kamery, teda hlavne RGBD snímky
- B Dáta z bodu A sú distribuované do balíčka rtabmap ros, kde môžu nastať dve situácie:
 - a) v prípade SLAM sa spájajú jednotlivé snímky do jedného 3D objektu (mapovanie) a zároveň sa určuje poloha (lokalizácia)
 - b) v prípade lokalizácie je určovaná poloha z predtým vytvorenej mapy
- C informácia o polohe je distribuovaná do balíčka mavros
- D mavros podáva informáciu o polohe z bodu C priamo autopilotu, ktorý ju zohľadňuje pri nízkoúrovňovom riadení drona.

1.3 Metóda SLAM

Keďže funkčné jadro nášho systému tvorí metóda Orb-Slam2, je vhodné opísať princíp jej fungovania trochu bližšie. Ako poznámku uvedieme fakt, že Orb-Slam2 v našom systéme nevystupuje v tzv. čistej podobe. Implementácia, aj princíp metódy Orb-Slam2 pozostáva z niekoľkých (troch) paralelne vykonávajúcich sa vlákien (procesov)[8]. Jedným z týchto procesov je detekcia uzavretej slučky. Aby systém dokázal uzavrieť slučku, musí detegovať miesto, na ktoré sa vrátil, resp. ktoré bolo už predtým navštívené. Keď sa táto slučka deteguje, globálna hierarchická mapa je aktualizovaná úpravou transformácií medzi jednotlivými snímkami, alebo podmapami[11]. A práve táto časť metódy Orb-Slam2 je pri jej integrácii do balíčka Rtab-Map deaktivovaná [4] a používa sa detekcia uzavretej slučky z balíčka Rtab-Map. Túto modifikovanú verziu autori M. Labbé a F. Michaud pomenovali ako ORB2-RTAB[4]. Práve ORB2-RTAB používame pri 3D rekonštrukcii priestoru, keďže jeho výstup predstavuje husté mračno bodov (z ang. termínu Dense Point Cloud). Natívny výstup metódy Orb-Slam2 naopak poskytuje iba riedke



Obr. 2 Rozdiel medzi riedkym a hustým mračnom bodov. Vľavo riedke mračno bodov - riedka rekonštrukcia mestského prostredia s niekoľkými uzavretiami slučiek. Vpravo kľúčové snímky a hustá 3D rekonštrukcia izby[7]. mračno bodov (voľný preklad, z ang. Sparse Point Cloud). Pri lokalizácii však plne postačuje základný Orb-Slam2.

1.3.1 Výber verzie Orb-Slam2

Prvá vec, ktorú uvedieme na pravú mieru je vysvetlenie rozhodnutia z časti 1.1. Implementácia metódy Orb-Slam2 ponúka celkovo 3 verzie, a to:

- Verzia pre stereo kamery
- Verzia pre **RGB-D** kamery
- Verzia pre mono kamery

Vybrali sme si teda verziu pre RGBD kamery, keďže kamera použitá v našom systéme je práve RGBD. Je hneď niekoľko dôvodov pre toto rozhodnutie. Napriek tomu, že verzia pre obyčajnú mono kameru je zrejme najlacnejším riešením, no obsahuje niekoľko úskalí. Napríklad už pri evaulácii výsledkov na známom datasete KITTI spozorovali Raúl Mur-Artal and Juan D. Tardós[8], že sekvenciu, kde mono verzia zlyhala, stereo verzia dokázala spracovať. Mono verzia taktiež mala problém s veľkou odchýlkou mierky (voľný preklad z ang. scale drift), čo môžeme vidieť na obr. 3. Mono verzia taktiež spotrebuje viac výpočtového výkonu, pretože hĺbka bodov nemôže byť odhadnutá z jednej snímky, a teda vyžaduje inicializáciu mapy[6]. Oproti tomu, kamera Intel Real-Sense D435 už poskytuje výstup, ktorý obsahuje hĺbkovú mapu, a teda nie je potrebné žiadne ďalšie spracovanie dát z kamery - pokiaľ nechceme napríklad určitým spôsobom filtrovať dáta z tejto kamery. Toto je vo všeobecnosti podstata RGBD kamier.

Treba ešte podotknúť, že stereo a RGBD kamery sú technologicky podobné. Nami zvolená kamera RealSense D435 používa aktívny stereo pár IR (infračervených) kamier na vytvorenie hĺbkovej mapy[3], k čomu je potom zarovnaná RGB snímka, aby vznikla textúra hĺbkovej mapy.



Obr. 4 Odhadnutá (čierna) a skutočná (červená) trajektória známeho datasetu KITTI 08. Vľavo mono Orb-Slam[6], vpravo Orb-Slam2 (stereo)[7].

1.3.2 Princíp Orb-Slam2

Ako už bolo spomínané v hlavnej sekcii 1.3, metóda Orb-Slam2 pozostáva z troch hlavných, paralelne vykonávajúcich sa vlákien (procesov), konkrétne sú ich funkcie nasledovné:

- Sledovanie lokalizácia kamery pri každej novej snímke - hľadanie zhody čŕt (z ang. feature matching) s lokálnou mapou. Minimalizácia chyby pri spätnej projekcii aplikáciou balíkových úprav (voľne preložené z ang. termínu Bundle Adjustments[5], ďalej len skrátene BA).
- Lokálne mapovanie spracovanie lokálnej mapy a jej optimalizácia, vykonávanie lokálnych BA
- Uzatváranie slučiek detegovanie veľkej slučky a korekcia naakumulovanej odchýlky



Obr. 5 Ukážka významu uzatvorenia slučky[11]. Vľavo lokálna mapa bez detegovaného uzatvorenia slučky. Vpravo s uzatvorením slučky.

2. Dosiahnuté výsledky

Pri overovaní presnosti nášho lokalizačného systému sme nepoužívali samotný dron. Slam sme realizovali ručne. Ako referenčný systém je použitý indoor lokalizačný systém Pozyx. Tento vyžaduje rozmiestnenie tzv. kotiev (z ang. anchors) vo vnútornom priestore. Systém je potom schopný lokalizovať zariadenia, tzv. značky (z ang. tags), ktoré sa v tomto vnútornom priestore pohybujú. Systém pracuje na báze UWB (ultra-široké pásmo, z ang. ultra-wideband) a k určeniu polohy sa využíva metóda s názvom multilaterácia[9].



Obr. 3 Porovnanie trajektórií nášho lokalizačného systému (v legende Orb-Slam2, keďže tento algoritmus predstavuje jadro nášho systému) a trajektórie zo systému Pozyx. Obrázok obsahuje dve merania.

Orb-Slam2 bol pri získavaní dát v režime Slam. A teda pri získavaní dát zo systému prebiehal slam, čo znamená, že súčasne prebiehala lokalizácia aj mapovanie. Keďže sme systém nespúšťali v počiatku súradného systému, ktorý je zhodný s tým, ktorý využíva Pozyx, bolo potrebné dáta posunú tak, aby mali rovnaký počiatok. Porovnanie je dostupné na obr. 4.

Na obr. 6 vidíme robustnosť systému, kedy sa systém pomerne pohotovo vrátil na správnu polohu po očividne chybnom určení polohy.

Na nasledujúcom obr. 7 uvádzame ukážku už spomínanej 3D rekonštrukcie priestoru. Pri zhotovení tejto 3D snímky bola použitá Slam metóda ORB2-RTAB (spomínaná v časti 1.3). Metóda optimalizácie v tomto prípade bola nastavená na metódu TORO[5].

2.1 Zhodnotenie výsledkov práce

V práci sa podarilo integrovať všetky časti lokalizačného systému, ktoré boli stanovené v počiatkoch tvorenia tohto systému a uviesť ich od fungujúcej podoby. Taktiež môžeme zodpovedne prehlásiť, že tento systém počas testovania fungoval stabilne z hľadiska jeho programového chodu. To znamená, že systém bol schopný pracovať bez prerušenia. Neznamená to však, že jeho výstup chybový nebol.

Na obr. 5 je porovnanie dvoch meraní. S meraniami sme spokojní, keďže trajektórie sú v podstate do veľkej miery zhodné. Ukazuje sa však nepresnosť v pravej časti trajektórie, kedy náš systém je voči Pozyx-u nepresný o cca 0.3m. Pozyx je v tomto prípade presný, čo sme overili klasickým pásmovým meradlom. Domnievame sa, že práve na tejto strane obdĺžnika trajektórie má táto chyba opodstatnenie. Je to podľa nás spôsobené charakteristikou kamery.

Na obr. 8 vidíme, že práve za stranou trajektórie, ktorá je odchýlená od referenčného merania nie je v blízkosti stena, ani žiadna iná súvislá, rozmerovo veľká prekážka.

Kamera má charakteristický prejav, ktorý vidno na obr. 9. Vidíme, že vzdialené plochy sú zo značnej časti zvlnené. Takto isto bola pri meraní zvlnená v hĺbkovej mape kamery vzdialená stena za pravou časťou trajektórie. Toto je naše vysvetlenie pre fakt, že práve táto strana meranej trajektórie bola odchýlená od referenčnej.

Na obr. 6 vidíme, že v meraní nastala pomerne veľká chyba. Podstatný je pre nás ale fakt, že systém bol po chvíli schopný re-lokalizácie na správnu polohu.



Obr. 8 Ukážka schopnosti systému vrátiť sa do správnej polohy pri náhlej veľkej chybe.

Systém pozyx, ktorý používame ako referenčný, poskytuje pomerne zašumený výstup. Avšak programové vybavenie, ktoré je k dispozícii na prácu s týmto systémom, obsahuje niekoľko filtrov, ktoré pomerne dobre signál vyhladia. Použili sme filter pohyblivého priemeru (z ang. moving average filter).



Obr. 9 3D rekonštrukcia priestoru, ukážka z dvoch perspektív. Na obrázku sú 3D tlačiarne.

Podľa obr. 7 môžeme prehlásiť, že optimalizácia grafu a celkovo spájanie hĺbkových snímok do jedného veľkého mračna bodov funguje spoľahlivo. Napríklad, keď sa pozrieme na stôl, na ktorom stoja tlačiarne (oranžovo-čierne objekty), vidíme, že jeho hrana je takmer dokonale rovná. Tento obrázok však ukazuje aj niektoré nedostatky, ktoré vyplývajú hlavne zo samotného výstupu použitej kamery. Kamera má tendenciu mierne skresľovať obraz, hlavne na okraji objektov. Pod skreslením v tomto prípade myslíme "roztrúsenie" hrán objektov. Domnievame sa, že pri použití napríklad v navigácii, kedy sa vytvorí napríklad mriežka obsadenosti, by tento defekt nemal veľký vplyv.



Obr. 6 Náčrt meranej trajektórie vrámci miestnosti. Modrá farba predstavuje samotnú trajektóriu, čierna farba predstavuje steny.



Obr. 7 Ukážka hĺbkovej snímky zarovnanej s RGB snímkou. Priamo z kamery D435.

Poďakovanie

Článok vznikol vďaka podpore projektov VEGA 1/0752/17 a APVV-17-0116.

Literatúra

[1] AI ROBOTICS.: Automated Drone (UAV) Applications. [Online] [Dostupné na internete: 24.03.2019] https://www. airoboticsdrones.com/applications/.

[2] NVIDIA.: Hardware For Every Situation. [Online] [Dostupné na internete: 25.03.2019]. https://developer.nvidia.com/embedded/develop/hardware.

[3] INTEL.: Intel RealSense[™] Depth Camera D435 – Intel RealSense[™] Depth Cameras. [Online] [Dostupné na internete: 25.03.2019]. https://click.intel.com/intelr- realsensetm-depthcamera-d435.html. [4] Mathieu Labbé a François Michaud.: RTAB-Map as an open-source lidar and visual simultaneous localization and mapping library for large-scale and long-term online operation. In: Journal of Field Robotics (2018). ISSN: 15564967. DOI: 10.1002/rob.21831.

[5] N. Mazouni et al.: SYSCOLAG: A transdisciplinary and multi-stakeholder approach towards integrated co-astal area management. An experiment in Languedoc-Roussillion (France). In: Vie et Milieu 56.4 (2006), s. 265–274. ISSN: 02408759.

[6] Raul Mur-Artal, J. M.M. Montiel a Juan D. Tardos.: ORB-SLAM: A Versatile and Accurate Monocular SLAM System. In: IEEE Transactions on Robotics 31.5 (2015), s. 1147-1163. ISSN: 15523098. DOI: 10.1109/TRO.2015.2463671.

[7] Raul Mur-Artal a Juan D. Tardos.: ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras. In: IEEE Transactions on Robotics 33.5 (2017), s. 1255–1262. ISSN: 15523098. DOI: 10.1109/TRO.2017.2705103.

[8] Raúl Mur-Artal a Juan D. Tardós.: ORB-SLAM2: an Open-Source SLAM System for Monocular, Stereo and RGB-D Cameras. In: IEEE Transactions on Robotics 33.5 (2017), s. 1255–1262. DOI:10.1109/TRO.2017.2705103.

[9] POZYX: Pozyx. [Online] [Dostupné na internete: 24.03.2019]. https://www.pozyx.io/.

[10] ROS.: About ROS. [Online] [Dostupné na internete: 25.03.2019]. http://www.ros.org/about-ros/.

[11] Brian Williams et al.: A comparison of loop closing techniques in monocular SLAM. In: Robotics and Autonomous Systems 57.12 (2009), s. 1188–1197. ISSN: 09218890. DOI: 10.1016/j.robot.2009.06.010.

Abstract

This article deals with field, which forms one of the main pillars of robotics - localization. Specifically it introduces localization system, which is dedicated to flying robots. Particularity, which this system disposes of, is the fact, that it is designated to be used in indoor environments and in 3D space. In practice, it means exclusion of any GNSS systems (i.e. GPS) used for localization. Aim of the article is not only the explanation of core principles of our system, or to show the results. It also contains one of the sample applications of such system - 3D reconstruction. However it is different field than localization, but in the end, these two are closely related.

doc. Ing. Andrej Babinec, PhD. Ing. Jozef Rodina, PhD. Ing. Eduard Mráz

Slovenská technická univerzita v Bratislave Fakulta elektrotechniky a informatiky Ústav robotiky a kybernetiky Ilkovičova 3 812 19 Bratislava 1 E-mail: eduard.mraz@stuba.sk

INTELIGENTNÝ MENIČ SO VSTAVANÝM PLC PRE BLDC MOTORY

Jaromír Stanko, Jozef Rodina, Peter Hubinský

Abstrakt

V dnešnej dobe existuje na trhu veľký výber meničov pre BLDC motory. Niektoré meniče využívajú pre riadenie FOC (field oriented control – vektorové riadenie) a z nich len malá časť navyše ponúka funkcionality ako konektivita s inými zariadeniami a možnosť programovať správanie meniča priamo používateľom zariadenia. Článok sa zaoberá meničom so senzorovým a bez senzorovým vektorovým riadením pre BLDC motory, ktorý má implementovaný modul pre vytváranie programov používateľom a zároveň dokáže komunikovať s inými zariadeniami pomocou komunikačných zberníc CAN, UART

Kľúčové slová: UAVCAN, BLDC, FOC, PLC, InstaSPIN-MOTION

Úvod

Využitie BLDC motorov v priemyselných a robotických aplikáciách sa stáva čoraz bežnejším. S polohovým senzorom dokážu BLDC motory vďaka dostatočnej presnosti polohovania nahradiť krokové motory. V určitých aplikáciách sú BLDC motory schopné nahradiť aj asynchrónne motory.

V roku 2013 predstavila spoločnosť Texas Instruments (TI) ich technológiu pre riadenie motorov InstaSPIN-FOC. Táto technológia umožnila vývojárom v oblasti riadenia elektromotorov sústrediť sa na vývoj samotnej aplikácie namiesto implementácie vlastných algoritmov riadenia čo je časovo a finančne neefektívne. InstaSPIN-FOC využíva pozorovateľa stavu FAST (Flux, Angle, Speed, Torque), ktorý nepotrebuje žiaden senzor. Okrem iného taktiež obsahuje nástroj na identifikáciu parametrov motora, ktorý má byť riadený. TI neskoršie prišla s nadstavbou - InstaSPIN-MOTION, ktorá rozšírila možnosti InstaSPIN-FOC. InstaSPIN-MOTION bol navrhnutý pre optimalizovanie sekvencii pohybov motora, sledovanie trajektórii s vysokou presnosťou. Obsahuje modul SpinTAC, ktorý dokáže vytvárať optimalizované trajektórie, ktoré motor bude sledovať, ladenie regulátora len s jedným parametrom a ADRC (Active Disturbance Rejection Control) riadenie.

Vytvorený menič by bol určený pre využitie v komplexnejších robotických alebo priemyselných aplikáciách využívajúcich viacero pohonov. To je dôvod, prečo má menič implementovanú komunikáciu prostredníctvom CAN zbernice. Hlavnou myšlienkou je využiť viacero meničov, ktoré spolu komunikujú priamo bez potreby PC. Šetrí sa tým čas keďže dáta nemusia byť najprv posielané do PC a až potom odoslané k cieľovému zariadeniu. Toto je užitočné v prípade ak musí riadený systém čo najrýchlejšie reagovať na poruchu, ktorá by zhoršila kvalitu riadenia ak by nebola dostatočne rýchlo vykompenzovaná.

Vytváranie a vykonávania programov, ktoré môže naprogramovať používateľ umožňuje vnorený "PLC" modul. Prostredníctvom API (aplication programming interface) má program vytvorený používateľom prístup k analógovým a digitálnym vstupom, výstupom, obmedzený prístup k polohovej, rýchlostnej, prúdovej riadiacej slučke. Po nahratí programu dokáže menič samostatne alebo v spolupráci s ďalšími meničmi vykonať naprogramované úlohy bez potreby nadradeného PC alebo PLC.

1. BLDC motory

BLDC motory sa od bežných jednosmerných motorov líšia tým, že nepoužívajú mechanický komutátor a klzný kontakt medzi komutátorom a kefovým zberačom. BLDC motory využívajú elektronický komutátor. Svojou konštrukciou pripomínajú synchrónny motor a preto nie je možné na BLDC motory priamo pripojiť jednosmerné napätie. Napájacie napätie musí byť postupným spínaním privádzane na vinutia motora.



Obr.1 Ukážka prierezu BLDC motora

Fig.1 BLDC motor cross section

Pokiaľ preteká vinutiami motora prúd, vytvára sa magnetický tok. Obr. 2 znázorňuje, čo sa deje ak vinutiami U a W preteká prúd. Veľká šípka reprezentuje výsledný magnetický tok, súčet magnetických tokoch od jednotlivých vinutí. Výsledný magnetický tok spôsobí rotáciu rotora s permanentným magnetom pretože magnetické pole od permanentného magnetu sa snaží zarovnať sa s výsledným magnetickým tokom (N pól

magnetu bude bližšie k hrotu šípky reprezentujúcej výsledný magnetický tok).



Obr.2 Výsledný magnetický tok vytvorený vinutiami U a W

Fig.2 Resultant magnetic flux produced by phase U and W

Rotácia rotora je zabezpečená neustálim pripájaním a odpájaním napájacieho napätia na vinutia statora. Ak sa jednotlivé vinutia prepínajú len medzi dvomi stavmi, preteká nimi prúd alebo nepreteká, a prúd preteká dvomi vinutiami, tak sa výsledný magnetický tok skokovo otočí o 60°. Takýto spôsob rotácie môže spôsobovať nežiadúce vibrácia a zvuky.

Ak by prúd pretekal všetkými vinutiami a jeho veľkosť by bola presne riadená, dosiahla by sa tým plynula zmena otočenia výsledného magnetického toku. Ten by bol v tomto prípade súčtom magnetických tokov od všetkých vinutí. Veľkosť prúdu pretekajúceho cez vinutia sa riadi napájacím napätím, ktoré je pripájané na vinutia. Zmena napätia je dosiahnutá pomocou PWM. Tento spôsob riadenia otáčania rotora sa nazýva vektorové riadenie.

Vektorové riadenie je komplikovanejšie ako "6-step" pretože je potrebné riadiť veľkosť prúdu pretekajúceho jednotlivými vinutiami. Riadenou veličinou je vektor statorového prúdu **i**s.



Obr.3 Rozloženie statorového prúdu na iq a id zložku [4]

Fig.3 Reduction of stator current is into components of d, q rotor coordinate frame (id and iq) [4]

Fázové prúdy (prúdy pretekajúce vinutiami) sú transformované Clarkovej a Parkovou transformáciou. Po transformáciách je statorový prúd i_s reprezentovaný dvoma zložkami – i_d a i_q . Žiadaná hodnota prúdu i_d ovplyvňuje úroveň magnetického toku v smere rotorového magnetického toku. i_q zložka sa podieľa na tvorbe momentu motora.

$$i_q = i_s * \cos\theta_i \tag{1}$$

$$i_d = i_s * \sin\theta_i \tag{2}$$

Rovnica (1) a (2) [4] popisuje rozdelenie statorového prúdu. θi predstavuje záťažový uhol. Pre riadenie BLDC motora je potrebné poznať aktuálnu polohu rotora (aktuálne natočenie). Z tohto dôvodu je BLDC motor doplnený o senzor ako napr. optický enkodér alebo Hallove sondy. Existuje však aj bez senzorová metóda, ktorá prepočítava polohu rotora na základe fázových prúdov.

2. InstaSPIN-MOTION – FOC riešenie pre riadenie elektromotorov

Texas Instruments InstaSPIN-MOTION je komplexné riešenie pre robustné riadenie momentu, rýchlosti a polohy elektromotorov. Celá knižnica je obsiahnutá v ROM 32-bit mikrokontrolera z rady TI C-2000 Piccolo.

InstaSPIN-MOTION využíva FAST pozorovateľa, ktorý meria magnetický tok rotora (veľkosť, uhol. rýchlosť) a moment hriadeľa pri bez senzorovom vektorovom riadení. Ďalej poskytuje identifikáciu parametrov motora, ktoré sú následné použité pri nastavení parametrov pozorovateľa FAST a inicializovaní prúdových PI regulátorov pre id a iq. Modul Spin-TAC, ktorý je súčasťou InstaSPIN-MOTION, poskytuje plánovanie trajektórii (lichobežník, s-krivka, st-krivka). Generuje profil sledovaného signálu na základe počiatočnej rýchlosti alebo pozície, požadovanej rýchlosti alebo pozície a na základe obmedzení pre akceleráciu a trh. SpinTAC tiež obsahuje ADRC algoritmus riadenia, ktorý proaktívne prepočítava a kompenzuje poruchy systému v reálnom čase.







2. Konektivita

Možnosť komunikácie s inými zariadeniami je dôležitá ak chce byť menič využitý v komplexnejších aplikáciách. Menič využíva komunikáciu cez CAN zbernicu a komunikáciu prostredníctvom UART.

CAN štandard definuje fyzickú a linkovú vrstvu OSI modelu a preto je potrebné využiť niektorý z protokolov vyššej vrstvy ako napr. CANopen, MilCAN, UAVCAN Protokol ako CA-Nopen je v priemysle rozšírený a dobre známy avšak jeho implementácia na nový hardvér (HW) je zložitejšia v porovnaní s inými protokolmi a zároveň je aj náročnejšia na pamäť a celkovo výkon HW na ktorom je implementovaný. Z tohto dôvodu bol ako protokol vyššej vrstvy vybraný protokol UAV-CAN, ktorý má odľahčenú implementáciu pre "hlboko" vnorené systému, ktoré sú obmedzené z pohľadu pamäti a výkonu. Aktuálne existuje jeho implementácia v C, C++, Python a Rust. UAVCAN využíva dva spôsoby komunikácie po zbernici: vysielanie (message broadcasting) a vyvolanie služby (service invocation). Pri vysielaní je po zbernici vyslaná správa, ktorú môže prijať každé zariadenie pripojené ku zbernici. Pri vyvolaní služby sa jedná o komunikáciu medzi dvoma zariadeniami, ktorá prebieha na báze požiadavka – odpoveď. Každé zariadenie pripojené k zbernici má vlastné unikátne ID.



Obr.5 Základný koncept UAVCAN protokolu [8]

Fig.5 Basic concept of UAVCAN protocol [8]

Pre oba spôsoby komunikácie existuje preddefinovaný súbor dátových štruktúr. Každá táto štruktúra má vlastný identifikátor (DTID - data type ID). Niektoré sú definované protokolom, ostatné môžu byť definované programátorom a prispôsobené potrebám konkrétnej aplikácie. Posielané dátové štruktúry sú definované s použitím DSDL (data structure description language). DSDL popis je použitý na serializáciu a deserializáciu dátovej štruktúry. Tento krok pomáha kompilátoru určiť koľko každá dátová štruktúra zaberá pamäte čo vedie k lepšej optimalizácii využívania pamäte a rýchlosti programu. Okrem štandardných dátových typov sú aj protokolom definované niektoré funkcie ako napr. aktualizácia firmware, monitorovanie stavu zariadenia, časová synchronizácia





Fig.6 Communication over CAN with UAVCAN protocol

3. Vstavaný "PLC" modul

"Soft" PLC alebo vstavané PLC je softvérový balíček, ktorý emuluje správanie klasického PLC na PC či mikropočítači, na

ktorom je nainštalovaný. V prípade riadenia malých zariadení ako môže byť napr. mobilný robot je kvôli rozmerom praktickejšie a zároveň lacnejšie využitie takéhoto riešenia.

PLC implementované v mikrokontrolery meniča umožňuje používateľovi vytvoriť vlastný program v jazyku C, s ktorým vie vďaka preddefinovaným funkciám riadiť BLDC motor. Užívateľský program neovplyvňuje bez firmvéru meniča a priamo nezasahuje do riadiacich slučiek a nepristupuje priamo k perifériám a registrom. Prístup k perifériám je zabezpečený prostredníctvom predkompilovaných funkcii, ktoré vytvárajú API. Výhodou využitia jazyka C je, že pre vykonanie užívateľského programu už nie je potrebný interpreter, ktorý by zaberal drahocenné miesto v pamäti mikrokontrolera. Užívateľský program využívajúci predkompilované funkcie je vykonávaný v reálnom čase a môžu byť využité pri vykonávaní časovo kritických úloh na všetkých úrovniach riadenia. Ďalšou výhodou je, že nie je potrebné sa zaoberať návrhom vlastného riadiaceho algoritmu pretože ten už je implementovaný v rámci InstaSPIN-MOTION a zároveň je vďaka predkompilovaným funkciám zjednodušená práca so samotným InstaSPIN-MO-TION.

Užívateľský program je kompilovaný, linkovaný a nahrávaný do pamäte mikrokontrolera nezávisle od firmware. Je cyklicky volaný z hlavnej slučky firmware meniča. K predkompilovaným funkciám sa pristupuje cez štruktúru, v ktorej sú uložene smerníky na predkompilované funkcii. Množstvo funkcii, ktoré môže byť preddefinované v štruktúre je obmedzené len množstvom periférii, pamäťou a fantáziou vývojára. Vďaka takto pripraveným a predkompilovaným funkciám, sa používateľ nemusí trápiť s konfiguráciou periférii, navrhovaním riadiaceho algoritmu a podobne. O všetky nízkoúrovňové záležitosti sa stará firmware meniča. Užívateľ už len tvorí aplikačnú časť programu.

4. Modelová aplikácia



Obr.5 Základný koncept UAVCAN protokolu

Fig.5 Hardware configuration of model of simplified industrial application

Vývoj a overovanie funkčnosti firmware meniča a užívateľských programov prebiehali na vytvorenom zjednodušenom modely priemyselnej aplikácie.

Pre každý z motorov je k dispozícii vlastná vývojová doska s nahratým firmware a užívateľským programom. Pre kontrolu komunikácie je k dispozícii USB-CAN adaptér Zubax Babel. BLDC motor 2 má k dispozícii inkrementálny snímač a je riadená jeho poloha. BLDC motor 1 nemá k dispozícii žiaden senzor, využíva bez senzorové vektorové riadenie a je riadená jeho rýchlosť. Meniče navzájom komunikujú prostredníctvom CAN zbernice.

Priebeh aplikácie je nasledujúci:

- motor 2 vykoná požadovaný počet otáčok
- po dosiahnutí žiadaného počtu otáčok odošle menič 2 správu meniču 1
- motor 1 sa roztočí na žiadanú hodnotu otáčok za minútu
- po určitom čase motor 1 plynule zmení svoju rýchlosť otáčania na novú žiadanú hodnotu otáčok za minútu
- po uplynutí ďalšieho času motor 1 plynulo zastaví
- menič 1 odošle správu meniču 2, že môže zmeniť polohu hriadeľa (aplikácia sa takto dostane do bodu 1 a cyklicky sa bude vykonávať)

Zariadenia využité v modelovej aplikácii:

- LAUNCHXL-F28069M vývojová doska s výkonovou časťou BOOSTXL-DRV8301 revB.
- BLDC motor 1 DJI 2212/920KV
- BLDC motor 2 EC-max 30 (60W) s inkrementálnym snímačom MR typ ML 128-1000CPT a planétovou prevodovkou
- CAN USART prevodník Zubax Babel

Záver

Vytvorený BLDC menič so vstavaným PLC modulom prezentovaný v článku umožňuje vytváranie užívateľských programov vykonávaných priamo v riadiacej slučke firmware meniča. Vykonávajú sa v reálnom čase. Užívateľský program písaný v jazyku C nepotrebuje žiaden interpreter. Šetrí sa tak pamäť mikrokontrolera a zároveň je takýto program vykonaný rýchlejšie ako keby bol využitý skriptovací jazyk s interpreterom. Používateľ sa nemusí zaoberať návrhom a implementáciou vlastného algoritmu pre riadenie motora pričom sa môže sústrediť na vývoj samotnej aplikácie. Ovládače, prístup k registrom mikrokontrolera, vektorové riadenie motora zabezpečuje firmware meniča obsahujúci InstaSPIN-MO-TION.

Komunikácia cez CAN zbernicu umožňuje prepojenie viacerých meničov a ich spoluprácu pri vykonávaní komplexnejších úloh.

Poďakovanie

Článok vznikol vďaka podpore Agentúry na podporu výskumu a vývoja z výskumného projektu APVV-17-0214: Kolaboratívny robot pre použitie v laboratóriu.

Literatúra

[1] Keeping, S.: Controlling sensorless BLDC motors via back EMF, 2013, https://www.digikey.com/en/articles/techzone/2013/jun/controlling-sensorless-bldc-motorsvia-back-emf

[2] Suslov, A.: Řízení soustavy bezkartáčových motorů: bakalárska práca, Praha: České vysoké učení technické v Praze, 2014.

[3] Kozáček, P.: Riadenie bldc motora v oblasti nízkych otáčiek: diplomová práca, Brno: Vysoké učení technické, 2015.

[4] Rau, D.: Pokročilé riadenie synchrónnych motorov s permanentnými magnetmi: písomný referát k dizertačnej skúške, Bratislava: Slovenská technická univerzita v Bratislave, 2016.

[5] Carlos, J.: Position and Speed Control of Brushless DC Motors Using Sensorless Techniques and Application Trends, Sensors (Basel, Switzerland), Molecular Diversity Preservation International (MDPI), 2010, www.ncbi.nlm.nih.gov/pmc/articles/PMC3231115/

- [6] Texas Instruments: LAUNCHXL-F28069M overview,
- 2015, http://www.ti.com/lit/ug/sprui11/sprui11.pdf

[7] Texas Instruments: InstaSPIN-FOC [™] and InstaSPIN-MOTION [™]; 2017,

http://www.ti.com/lit/ug/spruhj1g/spruhj1g.pdf

[8] UAVCAN Development Team: UAVCAN specification v1.0, 2019, https://new.uavcan.org/specification/UAV-CAN_Specification_v1.0.pdf

[9] Šimon, J.: BLDC aneb DC motor s nulovými náklady na údržbu, IN Elektro, 2011, http://www.odbornecasopisy.cz/res/pdf/44507.pdf

[10] Jovankovič, J., Žalman, M.: Mechatronické pohybové systémy (6), AT&P journal, 2006, http://robotika.yweb.sk/skola/jednosmerne%20mo-

tory /snimace%20polohy%20natocenia.pdf

[11] Balogh, T., Fedák, V., Ďurovský, F.: 2011 IEEE International Symposium on Industrial Electronics. pp.1403-1407, 2011.

[12] Texas Instruments: Controller area network physical layer requirements, 2008,

http://www.ti.com/lit/an/slla270/slla270.pdf

[13] Texas Instruments: Introduction to the Controller Area Network (CAN), 2016,

http://www.ti.com/lit/an/sloa101b/sloa101b.pdf

[14] CAN in Automation: CAN data link layer, http://affon.na-rod.ru/CAN/CANdII.pdf

[15] Texas Instruments: Introduction to the Controller Area Network (CAN), 2016,

http://www.ti.com/lit/an/sloa101b/sloa101b.pdf

[16] Texas Instruments: TMS320C28x optimizing C/C++ compiler v18.1.0.LTS, 2018.

http://www.ti.com/lit/ug/spru514p/spru514p.pdf

[17] Texas Instruments: TMS320C28x assembly language tools v18.1.0.LTS, 2018,

http://www.ti.com/lit/ug/spru513p/spru513p.pdf

Abstract

This paper describes an intelligent BLDC motor controller using sensored or sensor-less FOC solution InstaSPIN-MO-TION for torque, speed and position control. Enabled CAN communication allows controller to communicate with other controllers or devices supporting communication via CAN. Embedded "PLC like" module allows to create small usermade programs executed by controller's firmware. Programs written in C language are executed faster than if they were written in interpreted programming language. It allows them to be used in time critical applications in all levels of motor control. Created hardware and software were tested on simplified model of industrial application.

Prof. Ing. Peter Hubinský, PhD. Ing. Jaromír Stanko Ing. Jozef Rodina, PhD.

Slovenská technická univerzita v Bratislave Fakulta elektrotechniky a informatiky Ústav Robotiky a Kybernetiky Ilkovičova 3 812 19 Bratislava +421 2 654 20 4

SLAM NA PLATFORME MOBILNÉHO ROBOTA

Martin Lučan, František Duchoň, Martin Gulan

Abstrakt

Simultánna lokalizácia a mapovanie je veľmi aktuálny problém v rámci robotiky, vývojom nových SLAM algoritmov sa zaoberajú popredné výskumné inštitúcie. Text prezentuje implementáciu vybraných SLAM algoritmov na vlastnú hardvérovú platformu v podobe mobilného robota postaveného na podvozku RC modelu a porovnal ich charakteristiku v zvolených praktických experimentoch.. Prestavba pôvodného modelu vyžadovala mnoho hardvérových zmien, samostatná úloha bola tiež voľba vhodných snímačov a výpočtovej jednotky.

Kľúčové slová: SLAM, lokalizácia, mapovanie, odometria, snímače

Úvod

V priebehu posledných desaťročí rapídne stúpa nasadenie robotov prakticky vo všetkých odvetviach priemyslu. Stále častejšie sme však svedkami nasadenia robotov v domácnostiach a ostatných sférach, kde robot prichádza do blízkeho kontaktu s človekom. V súčasnosti od mobilných robotov očakávame vykonávanie zložitých úloh, ktoré pravidelne vyžadujú pohyb vo vonkajšom a vnútornom prostredí bez akýchkoľvek zásahov človeka. V prípade, že poloha robota v prostredí je známa, napr. pomocou GPS, tvorba mapy je zjednodušená. Rovnako to platí v prípade, ak máme vopred k dispozícii presnú mapu prostredia, lokalizácia v takomto prípade nie je príliš zložitým problémom. Najkomplikovanejšia situácia nastáva v prípade, keď ide o pohyb v neznámom prostredí, bez možnosti externej lokalizácie. Nutnou schopnosťou takéhoto robota je v reálnom čase tvoriť mapu prostredia a súčasne určiť svoju aktuálnu polohu v takto vytvorenej mape prostredia. Túto problematiku označujeme pojmom "Simultánna lokalizácia a mapovanie" (z angl. SLAM - Simultaneous Localization and Mapping).

1. Simultánna lokalizácia a mapovanie

Autonómny pohyb robota v neznámom prostredí je komplexný a zložitý proces, pri ktorom musia byť úspešne zvládnuté nasledovné procesy:

- mapovanie získavanie informácie o prostredí prostredníctvom snímačov;
- lokalizácia určenie polohy robota v lokálnom alebo globálnom prostredí;
- voľba trajektórie výber optimálnej trasy pohybu;
- pohyb nasledovanie vypočítanej trajektórie prostredníctvom akčných členov.

Simultánna lokalizácia a Mapovanie (SLAM) je prienikom mapovania a lokalizácie. Nemáme na mysli konkrétny algoritmus, tento pojem odkazuje na komplexnú problematiku tvorby mapy okolitého prostredia pomocou dostupného senzorového vybavenia a súčasnú lokalizáciu robota v takto vytvorenej mape. Významným prvkov tohto procesu je použitie orientačných bodov v mape na spresnenie odhadnutej polohy robota. Pre prvotný odhad polohy robota sa štandardne používa odometria - odhad relatívnej polohy robota vzhľadom na východiskovú (nulovú) polohu. SLAM je považovaný za tzv. "chicken-egg" problém, keďže pre tvorbu mapy prostredia je nutné poznať z ktorých pozícii boli merania vykonané a naopak, pri lokalizácii je dôležitá informácia o polohe orientačných bodov.

Matematicky je tento vzťah formulovaný výrazom

$p(x_k, m | z_{1:k}, u_{1:k})$

Pravdepodobnostné rozdelenie polohy robota x_k a mapy prostredia *m* preto závisi od sady meraní $z_{1:k}$ a vektora akčných zásahov $u_{1:k}$ predstavujúcich pohyb robota. Samotné merania aj pohyb podliehajú nepresnostiam, SLAM je preto fúziou dvoch nepresných meraní, ako je znázornené na Obr.1.



Obr.1 Ilustrácia problému SLAM. Súčasne je odhadovaná poloha robota aj orientačných bodov. [1]

1.1 Rôzne prístupy k riešeniu SLAM

V rámci teórie SLAM existuje mnoho prístupov a spôsobov, ktorým problém riešime. Vo všeobecnosti sa líšia typom a kombináciou použitých snímačov, vhodnosťou pre rozdielne aplikácie a v neposlednom rade aj výpočtovou náročnosťou. Viaceré z prístupov sú dokonca špecifické pre konkrétnu robotickú platformu a oblasť nasadenia. Väčšinu z nich však vieme zaradiť do nasledujúcich skupín:

- Filtračné metódy. Najpoužívanejší prístup využivajúci implementáciu Bayesovho filtra v rámci ktorého rekurzívne prebiehajú krok predikcie a merania. Informácia o prostredí aj poloha robota je reprezentovaná hustotou pravdepodobnosti. Do tejto skupiny patria používané metódy ako rozšírený Kalmánov filter a časticový filter (angl. Particle filter). Túto metódu používa napríklad EKF-SLAM, Gmapping, FastSLAM a ďalšie.
- 2) Grafické metódy (Graph-based approach). V tejto metóde je prostredie aj každá poloha robota na mape prezentovaná uzlami. Hrana medzi uzlami predstavuje priestorové obmedzenie, ktoré je zvyčajne dané transformáciou medzi uzlami. Cieľom optimalizácie je nájsť konfiguráciu uzlov, ktorá najlepšie vyhovuje obmedzeniam. Príkladom tejto metódy je GraphSLAM.
- Neurónové siete. Zatiaľ málo rozšírený prístup, za istých podmienok však dosahuje výborné výsledky. Príkladom je RatSLAM.

1.2 Snímače pre SLAM

Väčšina mobilných robotov využiva pre svoj pohyb kolesový podvozok. Prvotným predpokladom pre odhad polohy robota je správna voľba snímačov vnútorných veličín, ako kolesové enkódery, IMU jednotky a iné. Pre správnu funkciu SLAM algoritmu je nevyhnutné snímanie dát z okolitého prostredia, z ktorých sú následne extrahované kľučové orientačné body, medzi ktoré najčastejšie patria hrany, steny, rohy a iné. Medzi snímače schopné takto snímať okolité prostredie radíme v prvom rade laserový snímač, mono a stereovízne kamery, ultrazvukové snímače a množstvo ďalších špecifických snímačov. Keďže úspech SLAM-u závisí vo veľkej miere na senzorovom systéme, sú kladené náročné požiadavky najmä na:

- funkčnosť v reálnom čase;
- redundantnosť systém musí zachovať svoju funkčnosť aj pri poruche niektorého zo snímačov;

 presnosť systému — zvyšujeme ju použitím viacerých senzorov a následnou fúziou meraných dát.

1.2.1 Odometria

Odometria je proces, ktorý na základe poskytnutých údajov zo snímača udáva relatívnu polohu robota vzhľadom na nulový bod merania. Bez definovania nulového bodu nám údaj z odometrie nehovorí nič o absolútnej polohe. V praxi, pri väčšine robotických podvozkov na tento účel slúžia najčastejšie rotačné enkódery snímajúce počet impulzov, ktoré priamo súvisia s pootočením kolesa. Spolu so známou hodnotou polomeru kolesa tento údaj umožňuje odhadnúť prejdenú vzdialenosť. Každý nameraný údaj z enkóderu zahŕňa malú chybu merania, ktorá je sama o sebe zanedbateľná. Problém však nastáva pri narastajúcom počte takýchto meraní. Údaje sú snímané inkrementálne, v každej vzorke sa teda k súčtu starých meraní pripočíta nové meranie. Rovnako funguje súčet chýb merania, ktorý v čase týmto spôsobom narastá a nadobúda podobu Gaussového rozloženia pravdepodobnosti. Nezriedkavé je aj použitie tzv. vizuálnej odometrie, ktorá spočíva v porovnávaní za sebou idúcich snímkov, pričom odhaduje posun a pootočenie význačných bodov medzi jednotlivými snímkami. Táto metóda odometrie je vhodná hlavne pre nekolesové podvozky, podvozky s vysokou vôľou a pre použitie na šmykľavých a hrboľatých povrchoch, kde predpokladáme vysokú chybu odometrie.

1.2.2 Laserový snímač

Aktuálne najpresnejším spôsobom merania vzdialeností v priestore je použitie laserového snímaču (LIDAR, z angl. Light Detection and Ranging). Na základe údajov z laserového snímača je možné určiť vzdialenosť a relatívnu polohu orientačných bodov v prostredí. Táto informácia je kľúčová pri ďalšej tvorbe mapy a určovaní pozície robota. Silnými stránkami LIDAR-ov sú výsoká presnosť, bezpečnosť, rozlíšenie a v neposlednom rade aj jednoduchá interpretácia nameraných údajov. Narozdiel od ultrazvukových snímačov, ktoré merajú odozvu zvukového signálu, laserové snímače ako médium používajú svetelný lúč, ktorý sa vyznačuje výrazne dlhším dosahom. Údaje z laserového snímača vychádzajú vo formáte LaserScan alebo Point Cloud, ktorý namerané údaje prezentuje v bodoch umiestnených v dvojrozmernom alebo trojrozmernom priestore.

1.2.3 Kamera

Počítačové videnie je nevyhnutnou funkciou pre mnohé aplikácie, kamera je preto v robotike veľmi rozšíreným snímačom. Narozdiel od LIDAR-u, ktorý je výhodný pre určenie vzialenosti a rozloženie objektov v priestore, pomocou kamerového systému a algoritmov počítačového videnia dokážeme podrobnejšie určiť povahu samotných objektov v prostredí a určiť doplňujúce vlastnosti. Častým použitím kamery je práve počítačové rozlišovanie objektov, ktoré rozlíšenému objektu v zornom poli priradí niektorú z vopred vytvorených tried. V rámci SLAM sú pre kameru vytvorené osobitné algoritmy, ktoré umožňujú lokalizáciu a mapovanie iba pomocou kamery, prípadne kamery a IMU jednotky.

1.2.3 Inerciálny snímač (IMU)

IMU(z angl. Inertial measurement unit) je snímač, ktorý v kompaktnej jednotke zväčša združuje nasledovné snímače:

akcelerometer — translačné zrýchlenie v osi x, y, z

- gyroskop rotačné zrýchlenie okolo osi x, y, z, resp. klonenie, klopenie, zatáčanie (angl. yaw, pitch, roll);
- magnetometer intenzita magnetického pola.

Pomocou uvedených meraných veličín, dokážeme určiť relatívnu polohu vzhľadom na nulový bod. Avšak nevýhodou IMU je, podobne ako u rotačného enkódera, akumulácia chyby v čase. Nezriedka je takáto jednotka doplnená aj o korekciu pomocou GPS, ktorá eliminuje akumuláciu chyby v čase.

2. Platforma mobilného robota

2.1 Hardvérová konfigurácia



Obr.2 Mobilný robot postavený na platforme RC modelu.

Základom robota je šasi z RC modelu Traxxas SLASH 2WD, ktorý je vďaka svojim rozmerom a typu podvozku vhodnou platformou pre dodatočné hardvérové úpravy. Štandardná platforma bola upravená v nasledujúcich bodoch:

- V prvom kroku boli z modelu odstránené všetky nadbytočné komponenty. Konkrétne išlo o odstránenie plastovej karosérie a príjimaču signálu z ručného ovládača. Vybratý bol tiež pôvodný motor a kontróler, ktoré budú v ďalších krokoch úprav vhodne nahradené.
- Robot s kompletným hardvérovým vybavením má výrazne vyššiu hmotnosť ako pôvodný model, pôvodné pruženie podvozku bolo preto zamenené za pružiny s vyššou tuhosťou a dodatočne spevnené vymedzovacími vložkami.
- Testovanie SLAM algoritmov si vyžaduje pomalé a precízne riadenie pohonu. Štandardný motor bol preto zamenený za Brushless Sensored DC motor.
- 4. Z dôvodu uvedeného v predchádzajúcom bode došlo tiež k výmene základného ESC za VESC schopné uriadiť silný DC motor. Takáto zostava motoru a VESC je vďaka spätnej väzbe zo senzoru v motore schopná vykonávať pomalý a precízny pohyb.
- Pre uchytenie výpočtovej jednotky a senzorov bola navrhnutá a osadená platforma vyrobená technológiou 3D tlače.
- 6. K pôvodnej baterii ktorá napája VESC, ktoré delí napájacie napätie medzi motor a servo pribudla batéria pre napájanie výpočtovej jednotky Jetson TX2.



Obr.3 Schéma znázorňujúca fyzické prepojenie hardvérových častí upraveného RC modelu.

Výber snímačov použitých na robote zohladňuje účel, pre ktorý bol robot postavený. Vybrané snímače sú modely štandardne dostupné na trhu. Pre použitie na tento druh robota ich predurčuje nízka hmotnosť, kompaktná veľkosť a vhodné zbernice ktorými snímače komunikujú s výpočtovou jednotkou.

Laserový snímač	Slamtec RpLidar A2M6		
Stereokamera	Stereolabs ZED		
Inerciálna jednotka	Xsens Mti-G710		
Výpočtová jednotka	Nvidia Jetson TX2		

Tab.1 Hardvérové komponenty mobilného robota.

2.2 Softvérová architektúra

Softvérový systém zabezpečujúci potrebné funkcie funguje na báze ROS. Jednotlivé funkcie systému vykonávajú ROS uzly organizované do balíkov. Globálne je architektúra rozdelená na dve hlavné vetvy. Prvá z nich zabezpečuje zber údajov zo snímačov, ktoré vstupujú do SLAM bloku. Finálnym výstupom SLAM-u je mriežková mapa a relatívna poloha voči nulovému bodu. Druhá vetva pokrýva navigáciu robota prostredníctvom povelov z joysticku, ktoré transformuje do priameho pohonu motora a serva. Obe vrstvy prepája odometria, ktorá je ďalším vstupom do SLAM-u.



Obr.4 Zjednodušená softvérová schéma

3. Mapovanie robotického bludiska



Obr.5 Robot v bludisku Národného centra robotiky.

Cieľom experimentu je tvorba mapy pomocou niekoľkých vybraných SLAM balíkov s cieľom vizuálneho porovnania kvality vytvorených máp. Pre účely tohto experimentu som použil robotické bludisko v priestoroch Národného centra robotiky. Každý z testovaných mapovacích algoritmov má svoje špecifiká a dosahuje rôzne výsledky. V záujme kvalitnej výstupnej mapy a pozície robota som testoval zmenu väčšiny dôležitých parametrov. Mobilný robot je opakovane navigovaný bludiskom a dopĺňa tak chýbajúce časti, ktoré pri jednorazovom prejazde neboli zmapované.

Mapy vytvorené pomocou testovaných algoritmov sú reprezentované mriežkou obsadenosti (mriežková mapa). Celá mapovaná plocha je rozdelená do siete štvorcov s definovaným rozmerom, pričom každému štvorcu je priradená hodnota pravdepodobnosti výskytu objektu. Akonáhle hodnota pravdepodobnosti prekročí definovanú hodnotu, mriežka je považovaná za obsadenú (na mape zobrazená čiernou farbou).



Obr.6 Rozmery a usporiadanie stien robotického bludiska Národného centra robotiky. Farebne sú označené steny, ktorých nameraný rozmer ďalej vyhodnocujem.

3.2.1 Gmapping

Gmapping patrí medzi najpoužívanejšie SLAM algoritmy. Je založený na tzv. Rao-Blackwellized časticovom filtri. Z hľadiska hardvéru je nutné aby mobilný robot disponoval horizontálne umiestneným laserovým snímačom. Nutnosťou je tiež niektorá z foriem odometrie. Veľký vplyv na kvalitu výslednej mapy má kvalita odometrie, ktorá zvyšuje efektivitu algoritmu aj pri použití lidar-ov nižšej kategórie.

Pri teste Gmapping-u v prostredí robotického bludiska som dosiahol pri nastavených parametroch relatívne presné výsledky. Testoval som Gmapping s oboma formami odometrie: VESC odometria a vizuálna odometria. VESC odometria založená na motorových enkóderoch podlieha výraznej nepresnosti, táto chyba je však v rámci SLAM algoritmu eliminovaná.



Obr.7 Výsledná mapa bludiska v NCR vytvorená pomocou Gmapping s vizuálnou odometriou

3.2.2 Hector SLAM

Základným princípom Hector SLAM-u je metóda tzv. spájania skenov (z angl. scan matching), ktorá s vysokou frekvenciou porovnáva za sebou idúce skeny s cieľom vypočítať translačnú a rotačnú transformáciu medzi nimi, pomocou ktorej odhadnutá pozícia robota. Významným znakom ie HectorSLAM-u je, že pre svoju funkcionalitu nevyžaduje údaje z odometrie. Tento algoritmus je preto vhodný pre roboty pohybujúce sa v prostredí, kde je problematické dosiahnuť presnú odometriu, ako napríklad pohyb po hrboľatom alebo členitom teréne. Vhodný je tiež pre roboty so značnou vôľou v podvozku, ktorá ovplyvňuje meranie pomocou štandardných rotačných enkóderov. Nakoľko Hector SLAM vie fungovať iba s jediným snímačom, laserovým skenerom, sú na tento snímač preto kladené vyššie nároky ako v prípade Gmapping-u.

Vďaka členitosti a množstvu orientačných bodov pre laserový snímač, je mapa vyhotovená v bludisku relatívne presná. Nakoľko pri meraní bol použitý RPLidar A2M6 s frekvenciou merania maximálne 10Hz, v prípade Hector SLAM-u bolo nutné robotom vykonávať čo najpomalší pohyb. Problémom pre Hector SLAM môžu byť dlhé rovné plochy, ako napríklad chodby, bez dostatočného množstva význačných orientačných bodov, ktoré nie sú priaznivým prostredím pre metódu scan matching.



Obr.8 Výsledná mapa bludiska v NCR vytvorená pomocou Hector SLAM

3.2.3 CRSM SLAM

Skratka CRSM reprezentuje Critical Rays Scan Match SLAM. Základný princíp spočíva v spájaní skenov s vytváranou mapou (z angl.scan-to-map matching), namiesto štandardného (z angl. scan-to-scan matching), ktorý používa napríklad Hector SLAM. Vzťah medzi skenmi je v tomto prípade riešený pomocou tzv. Random Restart Hill Climbing algoritmu, ktorý do procesu spracovania vyberá iba kritické skeny, čo eliminuje akumuláciu šumu. Za kritické sú považované skeny, ktoré vynikajú medzi zvyšnými z hľadiska priestorovej informácie. Tento algoritmus je nenáročný na senzoriku, nakoľko vyžaduje iba použitie laserového snímača. Mapa vytvorená CRSM SLAM je už na prvý pohľad vizuálne odlišná. V porovnaní s predchádzajúcimi prístupmi reprezentuje rovinu priečky vernejšie a prakticky bez viditeľných skokov.



Obr.9 Výsledná mapa bludiska v NCR vytvorená pomocou CRSM SLAM

3.2 Reprezentácia výsledkov

Obsahom tejto state je porovnanie skutočných hodnôt vybraných rozmerov robotického bludiska (Obr.6) s rozmermi, ktoré boli odčítané z výsledných máp vytvorených uvedenými spôsobmi získavania mapy. Z týchto hodnôt sú následne vypočítané stredné hodnoty μ a smerodajné odchýlky σ pre každý z prípadov (Tab. 2). Pre každý z porovnávaných rozmerov vykresľujem graf s cieľom porovnať vypočítané parametre každého zo SLAM balíkov.

	μ_A	μ_B	μ _C	σ_A	σ_B	σ_{C}
GMAPPING + VESC	576.52	151.36	56.33	3.015	1.038	1.084
gmapping + Vo	573.10	151.43	54.37	1.313	0.731	0.841
HECTOR SLAM	579.41	151.76	57.05	1.557	0.865	0.489
CRSM SLAM	572.44	152.60	56.17	0.829	0.784	0.522

Tab. 2 Porovnanie strednej hodnoty μ a smerodajnej odchýlky σ pre vybrané rozmery bludiska.



Obr.10 Graf zobrazujúci Gaussovo rozdelenie pravdepodobnosti hodnôt rozmeru A.



Obr.11 Graf zobrazujúci Gaussovo rozdelenie pravdepodobnosti hodnôt rozmeru B.



Obr.12 Graf zobrazujúci Gaussovo rozdelenie pravdepodobnosti hodnôt rozmeru C.

V grafoch porovnávajúcich referenčnú hodnotu s rozdelením pravdepodobnosti príslušného rozmeru je viditeľný relatívne široký rozsah nameraných hodnôt.

Z hľadiska opakovateľnosti, t.j. konzistentnosti výsledku pri opakovaných merianiach, prakticky pri každom z rozmerov zaostáva Gmapping v kombinácii s VESC odometriou, následkom spomínej vôle v podvozku. Na výsledku sa nesporne podpísal aj fakt, že pri niektorých rozbehoch robota zo statickej polohy dochádza k skokovej zmene rýchlosti, ktorú odometria nevie vhodne kompenzovať. Naopak najlepšiu opakovateľnosť, naopriek tomu, že rozdiely sú tesné, vykazuje CRSM SLAM.

Záver

Stavba samotného robota vyžadovala hardvérovú úpravu v podobe nahradenia pôvodných komponentov vhodnými náhradami pre pomalý a plynulý pohyb. Následne bola nainštalovaná senzorová a PC nadstavba potrebná pre zbieranie dát a beh programu. Softvérové moduly implementujem do ROS, ktorý sa vďaka svojej univerzálnosti, usporiadanej architektúre a prehľadnej komunikácii medzi uzlami ukázal ako efektívny nástroj pre riešenie mnohých úloh v robotike. Z hľadiska opakovateľnosti nameraných údajov dosahujú podobné výsledky Gmapping v kombinácii s VO, HectorSLAM a CRSM SLAM. Výraznejšie za nimi zaostáva Gmapping s VESC odometriou, čo je spôsobené najmä vôľami a charakterom samotného podvozku robota. Výber vhodnej varianty z testovaných spôsobov preto závisí najmä od prostredia v ktorom plánujeme SLAM využívať a hardvéru, ktorý máme k dispozícii. Spoliehanie sa na samostatný LIDAR môže zlyhávať v prípade prostredia s výskytom predmetov s nízkou reflexivitou, prípadne sklom, ktoré laser nezaznamená.

Poďakovanie

Uvedený robot bol zostavený pre potreby Driverless divízie STUBA Green Team, ktorý pracuje na vývoji autonómneho študentského monopostu. Poďakovanie patrí členom tímu, rovnako ako aj konzultantom Ing. František Duchoň, PhD., doc. Ing. Martin Gulan, PhD. za odborné usmernenie pri práci. Vďaka patrí tiež Národnému centru robotiky za poskytnutie robotického bludiska, ktoré bolo vhodným priestorom pre praktické experimenty.

Tento článok vznikol vďaka podpore projektov APVV-16-0006, DIH² a VEGA 1/0754/19.

Literatúra

[1] BAILEY, T., DURRANT-WHYTE, H.: Simultaneous localization and mapping: Part 1.

[2] THRUN, S., BURGARD, W., FOX, D.: Probabilistic robotics

[3] STACHNISS, C., BURGARD, W.: Particle filters for robot navigation

Abstract

The main aim of the thesis is to implement selected SLAM algorithms on self-assembled hardware platform represented by a mobile robot built on the chassis of RC model. Rebuilding the original model required a number of hardware adjustments. A separate task was also an appropriate selection of sensors setup and computer unit. Simultaneous localization and mapping is a very topical issue in robotics, with the development of new SLAM algorithms being addressed by leading research institutions. In current thesis I have implemented several commonly used approaches to the SLAM solution and compared their performance and characteristics in selected practical experiments.

Ing. Martin Lučan prof. Ing. František Duchoň, PhD. doc. Ing. Martin Gulan, PhD.

Slovenská technická univerzita v Bratislave Fakulta elektrotechniky a informatiky Ústav robotiky a kybernetiky Oddelenie robotiky Ilkovičova 2961/3 84104 Bratislava Iucan.martin@gmail.com

ROZPOZNÁVANIE GEST RUKY POMOCOU KONVOLUČNÝCH NEURÓNOVÝCH SIETÍ

Slavomír Kajan, Jozef Goga

Abstrakt

Príspevok sa zaoberá problematikou rozpoznávania statických a dynamických gest ruky v oblasti hlbokého učenia. Využitím vhodných vlastností konvolučných neurónových sietí (CNN) boli navrhnuté postupy na trénovanie týchto modelov v oblasti rozpoznávania statických a dynamických gest ruky. Za účelom vyhodnotenia dosiahnutých výsledkov boli vytvorené dve databázy gest v aplikáciách rozpoznávania abecedy amerického znakového jazyka (ASL) a dynamických gest rukou kreslených číslic. Na snímanie obrazu bol použitý senzor Kinect vo verzii 2. Metodika rozpoznávania gest bola verifikovaná pre rôzne štruktúry CNN a boli realizované demonštračné aplikácie v prostredí Python.

Kľúčové slová : rozpoznávanie gest, konvolučné neurónové siete, statické gestá ruky, dynamické gestá ruky, senzor Kinect

Úvod

Interakcia medzi človekom a počítačom prináša každým dňom do oblasti výpočtovej techniky nové výzvy a problémy, ktoré je potrebné riešiť. Jedným z kľúčových problémov tejto interakcie je práve vizuálne rozpoznávanie gest. Efektívna detekcia gestikulácie pri komunikácii ako aj jej klasifikácia je náročná úloha z viacerých hľadísk. Ak neberieme v úvahu náročnosť na kvalitu snímania a spracovania vizuálnych snímok pri gestikulácii, stále sa musíme vysporiadiať s kultúrnymi, osobnostnými a fyziologickými zmenami parametrov komunikácie [1].

Technický pokrok v oblasti hĺbkových senzorov a umelej inteligencie priniesol za posledné roky nové nástroje, ako napríklad konvolučné neurónové siete (CNN) s pomocou ktorých sa dosahujú čoraz lepšie výsledky rozpoznávania tejto prirodzenej interakcie. Najčastejšie používaným typom neverbálnej l'udskej komunikácie sú gestá rúk, na ktoré sa momentálne upriamuje najväčšia pozornosť výskumu [2]. Existuje niekoľko možností rozdelenia gest rúk na základe rôznych charakteristických vlastností a interpretácie. Medzi jedno zo základných rozdelení patrí klasifikácia do dvoch skupín na statické a dynamické gestá.

Na základe rôznych prístupov k problematike samotného rozpoznávania gest rúk sa v priebehu rokov podarilo úspešne aplikovať viacero metód. Vo všeobecnosti metódy rozpoznávania gest pracujú na princípe extrakcie príznakov zo vstupného obrazu a následnej klasifikácii do príslušnej triedy [2, 3]. Použitie konvolučných neurónových sietí na rozpoznávanie gest [2, 4, 5, 6, 7], ktoré popisujeme v príspevku dokáže obe metodiky skĺbiť do jednej neurónovej siete.

1. Klasické metódy rozpoznávania statických gest ruky

Na obrázku č.1 sa nachádza všeobecná bloková schéma pre rozpoznávanie statických gest [3]. Nasnímaný obraz zo senzorov sa spracuje a uloží vo vhodnej forme. Následne sa môže dané gesto v snímku segmentovať, alebo sa príznaky extrahujú z celého vstupného obrazu. Pod extrakciou príznakov rozumieme vyhodnotenie kvantitatívnych a štatistických ukazovateľov na základe nami zvolenej metriky, ktoré reprezentujú dané gesto. Extrahované príznaky musia byť v ideálnom prípade invariantné voči posunutiu, pootočeniu, zmene mierky a prípadnému poškodeniu získaného obrazu.



Obr.1 Bloková schéma pre rozpoznávanie gest Fig.1 Block scheme for gesture recognition

Príklady rôznych extrahovaných príznakov sú zobrazené na obrázkoch 2, 3 a 4, kde príznaky sú napríklad získavané pomocou detekcie koncov prstov a stredu dlane [4], Hougovou transformáciou alebo Voronoiovými diagramami [2].



- Obr.2 Extrakcia príznakov pomocou detekcie koncov prstov a stredu dlane
- Fig.2 Features extraction using end fingers and center palm detection



- Obr.3 Extrakcia príznakov pomocou Houghovej transformácie
- Fig.3 Features extraction using Hough transformation



Obr.4 Extrakcia príznakov pomocou Voronoiových diagramov (skeletonizácia gesta)

Fig.4 Features extraction using Voronoi diagrams (skeletonizations of gestures)

Tieto extrahované príznaky sú vstupom do výpočtového klasifikačného modelu, ktorý má za úlohu správne určiť rozpoznávané gesto. Klasifikačný model je najčastejšie realizovaný neurónovou sieťou alebo metódou strojového učenia, ako sú napríklad metódy podporných vektorových strojov (support vector machines), rozhodovacie stromy (decision trees), náhodné lesy (random forests).

Ako príklad najčastejšie používaných neurónových sietí určených na klasifikáciu statických gest uvádzame viacvrstvové perceptrónové siete (MLP) (Obr. 5)



Obr.5 MLP sieť na rozpoznávanie statických gest Fig.5 MLP network for static gestures recognition

2. Konvolučné neurónové siete

Konvolučné neurónové siete je možné využiť na rozpoznávanie gest ruky z dôvodu, že sú špeciálne navrhnuté na rozpoznávanie vzorov v dvojrozmernom priestore dát s veľkou mierou invariantnosti voči posunutiu, zmene mierky alebo iným formám skreslenia. Tieto vlastnosti sieť získava prostredníctvom učenia s učiteľom (supervised learning).

Už pomenovanie konvolučná neurónová sieť naznačuje, že táto sieť obsahuje matematickú operáciu nazývanú konvolúcia. Konvolúcia je matematická operácia, ktorej vstupom sú dve funkcie s reálnym argumentom. Vo všeobecnosti je definovaná vzťahom, ktorý vyjadruje integrál násobenia dvoch funkcií, pričom jedna z nich je symetricky prevrátená a posunutá.

$$s(t) = \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau$$
(1)

najčastejšie sa označuje operátorom * .

$$s(t) = (f * g)(t) \tag{2}$$

V princípe je táto matematická operácia veľmi podobná krížovej korelácii, ktorá vyjadruje mieru podobnosti funkcií vzhľadom na ich vzájomné posunutie. Ak uvažujeme jednu zo vstupujúcich funkcií, ako funkciu dát obsahujúcu šum a druhú funkciu ako vzor, ktorý obsahuje informáciu o tom, čo prenášajú alebo obsahujú naše dáta. Potom má konvolúcia význam kvantitatívneho ukazovateľa pravdepodobnosti, ako sa v každom vzájomnom posunutí zhodujú naše dáta so zvoleným vzorom.



Obr.6 Konvolúcia v jednorozmernom priestore Fig.6 Convolution in one-dimensional space

Konvolúciu môžeme vnímať aj ako vyhladzovanie, prípadne rozmazávanie funkcie zvoleným filtrom. To akým spôsobom ovplyvní výsledok konvolúcie vstupnú funkciu záleží od konkrétnej voľby filtra (Obr. 7). V terminológii strojového učenia sa zvyčajne funkcia dát vstupujúca do operácie konvolúcie nazýva vstup *(input)*. Funkcia obsahujúca vzor, alebo informáciu o funkcii dát sa nazýva jadro *(kernel)*. Výstupom konvolúcie je funkcia označovaná najčastie ako mapa príznakov *(feature map)*. V reálnych aplikáciách strojového učenia pracujeme s diskrétnymi signálmi a preto sa častejšie používa diskrétna forma konvolúcie:

$$s[n] = (f * g)[n] = \sum_{m=-\infty}^{\infty} f[m]g[n-m]$$
(3)

Operácia konvolúcie dvojrozmernom diskrétnom priestore je definovaná nasledovne:

$$S(i, j) = (I * K)(i, j) = \sum_{m} \sum_{n} I(m, n) K(i - m, i - n)$$
(4)

Výsledkom konvolúcie je v tomto prípade dvojrozmerná mapa príznakov (Obr. 7).





Architektúra konvolučnej neurónovej siete pozostáva z rôznych vrstiev. Konvolučná a zhlukovacia vrstva sa správajú ako extraktory príznakov a plneprepojená vrstva ako klasifikátor. V konvolučnej vrstve sa vykonáva výpočet 2rozmernej diskrétnej konvolúcie jedného farebného kanála vstupného obrazu s konvolučným jadrom. Z konvolučnej vstvy dostávame tzv. mapy príznakov, ktoré sa aktivujú pri detekcii rôznych obrazových vzorov ako sú napríklad hrany pod určitým uhlom, farebné zhluky a podobne. Zhlukovacia vrstva vykonáva podvzorkovanie vstupnej mapy príznakov, čím zmenšuje veľkosť mapy príznakov, ale zachováva najdôležitejšie informácie v nej obsiahnuté. Týmto výrazne zmenšuje počet parametrov, výpočtovú náročnosť siete a tiež aj možnosť predtrénovania siete. Konvolučná neurónová sieť môže obsahovať aj ďalšie typy vrstiev ako normalizačná vstva, nelinearita typu ReLu alebo Dropout vrstva, ktoré prispievajú k rýchlejšej konvergencii učenia siete a zabraňujú pretrénovaniu.

Tieto vrstvy môžeme rôznym spôsobom radiť za sebou a takto vytvárať rôznu architektúru siete (Obr. 8).



Obr.8 Príklad architektúry konvolučnej neurónovej siete s dvoma konvolučnými vrstvami

Fig.8 Example of convolutional neural network architecture with two convolutional layers

3. Databázy gest ruky

3.1 Databáza gest americkej znakovej reči - ASL

Keďže verejne dostupné databázy obsahujú malý počet gest alebo malý počet vzoriek, vytvorili sme vlastnú databázu gest americkej znakovej abecedy (ASL) [5]. Ako základné gestá sme si zvolili znaky americkej znakovej abecedy (ASL), ktorú sme pozmenili kvôli dynamickým znakom "J" a "Z". Týmto sme vybrali 35 znakov, ktoré tvoria základ našej databázy. Pomocou hĺbkového senzora Microsoft Kinect v2 sme následne uložili každé z 35 gest vo forme farebnej, infračervenej a hĺbkovej snímky. Pri snímaní gest sme pre jednoduchosť použili interné prostredie s bielym pozadím. Na vytvorení tejto databázy sa podiel'alo 65 dobrovol'níkov z radov študentov a akademického zboru ÚRK FEI STU. Z toho 60 mužov a 5 žien. Pri snímaní jednotlivých gest sme rukou hýbali v malom rozsahu, aby sme vytvorili nové pootočené a posunuté vzorky. Jedno snímanie teda pozostávalo z 5 jedinečných nasnímaných vzoriek pre každé gesto. Týmto sa nám podarilo vytvoriť 175 snímkov (35 gest po 5 snímkov) pre farebné, infračervené a hĺbkové snímky. Celkovo teda 525 snímkov na osobu. Výsledná databáza má veľkosť 34 125 snímkov (65 osôb po 525 snímkov). Tieto snímky sme následne upravili do formy vhodnej na trénovanie konvolučnej siete.

V originálnej farebnej snímke zo senzora *Kinect* s rozlíšením 1920x1080 pixelov sme segmentovali len obrazovú časť zobrazujúcu dané gesto. Takýmto spôsobom sme vytvorili štvorcový obraz vhodný na trénovanie s rozlíšením 640x640 pixelov. Segmentáciu gest sme vykonali taktiež pre originálne hĺbkové a infračervené snímky. Pôvodné rozlíšenie 512x424 pixelov, sme taktiež upravili na štvorcový obraz vhodný na trénovanie s rozlíšením 220x220 pixelov.



Obr.9 Ukážka nasnímaného statického gesta v databáze (zľava doprava farebný, infračervený a hĺbkový snímok) Fig.9 Previewed static gesture in database (left to right color, infrared and depth frame)

3.2 Databáza rukou kreslených číslic

Pomocou senzora Kinect v2 sme vytvorili databázu rukou kreslených číslic [8]. V databáze je uložený pohyb stredu dlane ruky, t.j. priebehy x-ovej a y-ovej súradnice v čase (obr. 10). Databázu na trénovanie vytváralo 6 študentov, pričom snímanie sa opakovalo 3 krát. Ako začiatok a koniec gesta sme použili otvorenie a zatvorenie ruky.



Obr.10 Ukážka dynamického gesta v databáze Fig.10 Previewed dynamic gesture in database

4. Dosiahnuté výsledky

4.1 Rozpoznávanie gest americkej znakovej reči

Praktickú realizáciu konvolučných neurónových sietí sme vytvorili v jazyku Python s využitím frameworkov Keras a TenzorFlow. Databázu gest sme najprv náhodne rozdelili na dve časti, kde 60% tvorili trénovacie dáta, 40% testovacie dáta a v oboch častiasť boli triedy percentuálne rovnako zastúpené. Ako počiatočný experiment sme na vstup siete aplikovali farebné (rgb) snímky databázy. Celkovo bolo v trénovacej množine 6825 farebných snímkov a v testovacej množine 4550 snímkov. Porovnávali sme viacero architektúr konvolučných neurónových sietí z hľadiska rýchlosti konvergencie a úspešnosti klasifikácie pre prvých 500 trénovacích epoch pri rovnakom rozdelení dát. Najlepšia architektúra je zobrazená na obr. 11. Ako optimalizačný algoritmus pri trénovaní siete sme zvolili algoritmus Adam (adaptive moment estimation), ktorý má z najčastejšie používaných algoritmov pri trénovaní neurónových sietí jednu z najrýchlejších konvergencií [5].



Obr.11 Architektúra CNN na rozpoznávanie statických gest

Fig.11 CNN architecture for static gestures recognition

V tabuľke 1 je vyčíslená chyba siete a dosiahnutá presnosť klasifikácie. Uvedené hodnoty sú priemerom posledných 20tich hodnôt učiacich kriviek na trénovacej a testovacej množine. Hodnota chybovej funkcie je priemerná krížová entropia (ACE - average cross-entropy error). Priebehy trénovacieho cyklu architektúry siete (obr.11) sa nachádzajú na obrázkoch 12 a 13.

Priemerná o	hyba [ACE]	Priemerná úspešnosť [%]		
Train	Test	Train Test		
0.5249	0.4634	80.13	82.44	

Tab.1 Vyčíslené hodnoty chyby a úspešnosti siete

Na učiacich krivkách (Obr. 12, 13) je vidieť, že úspešnosť na testovacej množine je vyššia ako úspešnosť na trénovacej množine dát. Je to spôsobené použitím takzvaného dropout-u, teda náhodného vypnutia určitého počtu neurónov v priebehu trénovacieho procesu.



Obr.12 Priebeh chyby siete v trénovacom procese Fig.12 Time response of network error for training



Obr.13 Priebeh úspešnosti v trénovacom procese Fig.13 Time response of classification score for training

Následne sme v architektúre sieti zvýšili počet jadier na dvojnásobok a trénovanie sme predĺžili o ďalších 500 epoch. Trénovanie sme realizovali na farebných, hĺbkových a infračervených snímkoch našej databázy, pričom zakaždým sme trénovanie začínali s natrénovanými váhami z predchádzajúceho experimentu. V tomto prístupe sme zvolili krížovú validáciu dát, kde sme jednotlivé osoby náhodne zaradili do trénovacej (60 %) alebo testovacej skupiny (40 %). Dosiahnuté výsledky sme následne vyhodnotili v tabuľke 2.

Typ snímok	Priemerná chyba [ACE]		Priemerná úspeš- nosť [%]	
	Train	Test	Train	Test
Farebné	0.5831	0.3894	77.69	84.87
Infračervené	0,9178	0,6287	64,71	75,02
Hĺbkové	1,0102	0,6546	62,22	74,52

Tab.2 Vyčíslené hodnoty chyby a úspešnosti siete

Keďže maximálna dosiahnutá úspešnosť v tomto porovnaní (Tab. 2) bola menej ako 85% pri plnofarebných snímkach, rozhodli sme sa v ďalších experimentoch zamerať hlavne na zvýšenie úspešnosti samotnej klasifikácie. V predchádzajúcich porovnaniach sme rozdelili databázu na trénovaciu a testovaciu množinu, pričom sme náhodne nerozdeľovali jednotlivé vzorky, ale ľudí zaradených do databázy. Týmto sme krížovou validáciou overili úspešnosť siete na testovacích osobách, ktorých vzorky neboli obsiahnuté v trénovacej množine počas trénovania. Trénovacia množina pritom obsahovala 60 % všetkých dát, čo predstavovalo v našom prípade 39 osôb. Naproti tomu testovacia množina obsahovala zvyšných 40 % dát, teda zostávajúcich 26 osôb. Keďže našou snahou bolo zvýšiť úspešnosť klasifikácie, rozhodli sme sa predložiť sieti viac vzoriek a následne vyhodnotiť dosiahnuté výsledky. V tabuľke 3 je porovnanie priemerných dosiahnutých výsledkov najúspešnejšej architektúry pri krížovej validácii s rôznym percentuálnym rozdelením dát.

Rozdelenie dát	Priemerná chyba [ACE]		Priemerr nost	iá úspeš- ː [%]
	Train	Test	Train	Test
60%/40%	0.5831	0.3894	77.69	84.87
80%/20%	0,5676	0,3384	80,42	89,79

Tah 3 V	včíslené	hodnoty	chyby a	a úsi	nešnosti	siete
1 4 8 1 9 1	,	nounory			000110011	0.000

Testovaním sme zistili, že úspešnosť siete bola veľmi dobrá, s výnimkou niektorých špecifických gest (obr.14)



Obr.14 Najčastejšie nesprávne klasifikované gestá Fig.14 The most commonly misclassified gestures

Z pôvodnej 35 znakovej abecedy sme odstránili znaky 'D','H','M','S','V' a 'W', čím nám vznikla upravená znaková sada s počtom 29 znakov. Na tejto znakovej sade sme znova natrénovali našu najúspešnejšiu neurónovú sieť" pri náhodnom 80% rozdelení dát do trénovacej množiny. Na upravenej znakovej sade farebných testovacích snímkov, ktorá obsahovala 1885 vzoriek sa nám podarilo dosiahnuť úspešnosť klasifikácie 94.48% na testovacej množine. Najúspešnejšiu konvolučnú sieť natrénovanú na upravenej znakovej sade sme následne použili v aplikácii na rozpoznávanie v reálnom čase (Obr.15).



Obr.15 Aplikácia rozpoznávania gest ruky Fig.15 Application for hand gesture recognition

4.2 Rozpoznávanie rukou kreslených číslic

V súčastnosti existuje viacero algoritmov, pomocou ktorých je možné rozpoznávať dynamické gestá. Patrí sem napríklad algoritmus dynamického zarovnania časovej mierky (DTW) [8], skryté Markovove modely (HMM) [8] alebo konvolučné neurónové siete (CNN) [2].





- Obr.16 Číslo vykreslené zo súradníc x, y (vľavo) a predspracované pre CNN (vpravo)
- Fig.16 Number paints from coordinates x, y (left) and pretreated for CNN (right)

Štandardným vstupom do konvolučných neurónových sietí je informácia vo forme obrazu. Z vytvorenej databázy časových priebehov pohybu ruky sme vykreslenú trajektóriu transformovali do obrazovej informácie vhodnej pre vstup do CNN (obr. 16).

Na obrázku 17 sa nachádza architektúra CNN pre vstupný obraz 28x28x1, ktorá bola použitá na rozpoznávanie rukou kreslených číslic od 0 do 9 [8].



Obr.17 Architektúra CNN na rozpoznávanie rukou kreslených číslic

Fig.17 CNN architecture for recognition of hand painting numbers

V procese trénovania CNN sme rovnako zvolili vzájomnú entropiu ako chybovú funkciu (ACE), a optimalizačnú metódu Adam. V tabul'ke 4 sú zobrazené úspešnosti klasifikácie jednotlivých dynamických gest pri testovaní tromi nezávislími použivateľmi. Každé gesto bolo nezávisle testované 5krát, pričom výsledná priemerná dosiahnutá úspešnosť klasifikácie jednotlivých gest bola 96,67 %.

Úspešnosť rozpoznávania dynamických gest [%]									
0 1 2 3 4 5 6 7 8 9									
100	100	100	100	100	93,3	100	86,7	100	93,3

Tab.4 Úspešnosť klasifikácie jednotlivých gest pomocou CNN

Záver

V tomto príspevku sme sa zaoberali problémom rozpoznávania statických a dynamických gest rúk s využitím modelov hlbokého učenia. Na nami vytvorenej databáze gest sme testovali schopnosť konvolučných neurónových sietí rozpoznávať statické a dynamické vzory. Sledovali sme vplyv parametrov siete na rýchlosť konvergencie a úspešnosť klasifikácie. Na overenie správnosti riešenia sme následne naprogramovali aplikácie v prostredí Python pre rozpoznávanie statických a dynamických gest v reálnom čase využívajúce senzor Kinect v2. Tieto aplikácie sme overili na testovacej množine, ktorá nebola súčasťou vytvorených databáz. Z celkových dosiahnutých výsledkov tejto práce vyplýva, že konvolučné neurónové siete sú vhodný nástroj na rozpoznávanie statických a dynamických gest vďaka ich vysokej úspešnosti klasifikácie a schopnosti rýchleho rozpoznávania po natrénovaní.

Poďakovanie

Výskum opísaný v tomto príspevku sa uskutočnil v rámci projektu Slovenskej grantovej agentúry VEGA č. 1/0867/17

Literatúra

[1] KELLY, S. D., MANNING, S. M., RODAK, S.: Gesture gives a hand to language and learning: Perspectives from cognitive neuroscience, developmental psychology and

education. In: Language and Linguistics Compass, 2008, 2.4: 569-588.

[2] PISHARADY, P. K., SAERBECK, M.: Recent methods and databases in vision-based hand gesture recognition: A review. In: Computer Vision and Image Understanding, 2015, 141: 152-165.

[3] KAJAN, S., PERNECKÝ, D.: Gesture recognition of American sing language using kinect. In IN-TECH 2016 : International conference on innovative technologies. Proceedings. 1. vyd. Rijeka 2016, 213-216.

[4] KAJAN, S., GOGA, J.: Simulation-based model control using static hand gestures in Matlab. In Technical computing Prague 2017

[5] GOGA, J., KAJAN, S.: Hand gesture recognition using 3D sensors. In Proceedings ELMAR-2017 : 59th International symposium. Zadar, Croatia. 2017, S. 181-184.

[6] TANG, A., et al.: A real-time hand posture recognition system using deep neural networks. In: ACM Transactions on Intelligent Systems and Technology (TIST), 2015, 6.2: 21.

[7] STREZOSKI, G., et al.: Hand Gesture Recognition using Deep Convolutional Neural Networks. ICT Innovations, 2016

[8] ZSIROS, O.: Rozpoznávanie dynamických gest ruky senzorom Kinect, diplomová práca FEI STU Bratislava, 2019

Abstract

The paper deals with the issue of recognizing static and dynamic hand gestures in the area of deep learning. Using suitable properties of convolutional neural networks (CNN), procedures were designed to train these models in the task of recognizing static and dynamic hand gestures. In order to evaluate the results, two gesture databases have been created in applications of the American Sign Language Alphabet (ASL) and dynamic hand drawn digit gesture recognition. The Kinect sensor in version 2 was used to capture the image. The gesture recognition methodology was verified for various CNN structures and demonstration applications were implemented in the Python environment.

Ing. Slavomír Kajan, PhD.

Slovenská technická univerzita v Bratislave Fakulta elektrotechniky a informatiky Ústav robotiky a kybernetiky Ilkovičova 3 81219 Bratislava E-mail: slavomir.kajan@stuba.sk

Ing. Jozef Goga

Slovenská technická univerzita v Bratislave Fakulta elektrotechniky a informatiky Ústav robotiky a kybernetiky Ilkovičova 3 81219 Bratislava E-mail: jozef.goga@stuba.sk

NEURO-EVOLÚCIA RIADENIA DYNAMICKÝCH SYSTÉMOV

Ivan Sekaj, Dominika Banásová, Martin Komák, Milan Hvozdík

Abstrakt

V predkladanom projekte predstavujeme metodiku návrhu riadenia dynamických systémov na báze umelých neurónových sietí (UNS), pričom proces ich učenia (parametrizácie) je realizovaný evolučným algoritmom. Takýto neuro-evolučný prístup dovoľuje implementáciu UNS v širokom spektre aplikácií. Spomedzi rôznych možných v článku demonštrujeme 2 typické prípady riadenia. Prvý je riadenie spojitého, nelineárneho dynamického systému a druhý prípad je riadenie mobilného robota.

Kľúčové slová: umelá neurónová sieť, evolučný algoritmus, spojitý dynamický systém, mobilný robot, riadenie

Úvod

Umelé neurónové siete (UNS) predstavujú prelomovú technológiu v metódach strojového učenia. Ich hlavná aplikačná doména sú problémy klasifikácie objektov, rozpoznávania vzorov (obrazu, zvuku a iných signálov), tiež modelovanie systémov a aproximácia funkčných závislostí. V poslednej dobe ale UNS zažívajú rozkvet vďaka novým mnohovrstvovým architektúram (tzv. hlboké siete) a novým typom metód spracovania vstupných dát, ako sú konvolučné metódy a metódy hlbokého učenia. Tu sa jedná predovšetkým o cieľ naučiť UNS reprezentovať a využívať závislosti medzi vstupno-výstupnými dátami získané z praktických problémov a týmto spôsobom nahradiť, automatizovať, resp. ešte viac zdokonaliť schopnosti a funkcionalitu stroja pri nahrádzaní človeka pri riešení problémov. Pri tomto type úloh sa jedná o učenie s učiteľom. Iná sféra implementácie UNS, kde spravidla nemáme dispozícii vstupom zodpovedajúce známe výstupné dáta, čiže a-priori nepoznáme "správne odpovede", je učenie bez učiteľa. Tu vieme definovať aktuálne vstupy do predmetného objektu, ale poznáme iba cieľ, ktorý chceme splniť vo forme vhodného kritéria či pokutovej funkcie. Medzi úlohy tohto typu patria aj mnohé úlohy riadenia. Vieme, čo sú vstupy do systému (informácie zo snímačov atď.) a hľadáme optimálnu stratégiu na dosiahnutie cieľa.

Predkladaný článok predstavuje metodiku, ktorá umožňuje naučiť UNS riadiť systém na báze používateľom definovaného cieľa. Inými slovami, máme k dispozícii riadiaci algoritmus vo forme UNS, ktorý riadi systém tak, aby bolo splnené určené kritérium. Pre účely učenia UNS (resp. jej parametrizácie) je použitý evolučný algoritmus (EA), v našom prípade genetický algoritmus (GA). Pre proces učenia UNS za pomoci EA sa začal používať pojem neuroevolúcia. Proces neuro-evolúcie riadenia demonštrujeme na dvoch principiálne odlišných aplikáciách. Prvou je riadenie (regulácia) spojitého nelineárneho dynamického systému a druhou je riadenie pohybu mobilného robota v 2D prostredí s prekážkami.

V prvej kapitole predstavíme princíp neuro-evolúcie (NE). V druhej vysvetlíme použite NE pri regulácii spojitého nelineárneho dynamického systému, prezentujeme výsledky experimentov a porovnáme ich s konvenčnou PID reguláciou. Riadenie mobilného robota na báze NE aj s experimentálnymi výsledkami bude opísané v kapitole 3.

1. Neuro-evolúcia

Pri pojme neuro-evolúcia riadenia máme na mysli takú architektúru riadenia, kde objekt, ktorý realizuje riadenie (počíta a generuje riadiacu veličinu) je umelá neurónová sieť (UNS), pre učenie ktorej sa využíva evolučný algoritmus (EA). Bez ujmy na všeobecnosti v našom prípade uvažujeme genetický algoritmus (GA). Ako je známe, EA aj GA sú efektívne optimalizačné prístupy, ktoré sú schopné nachádzať (sub)optimálne riešenia komplexných problémov v mnohodimenzionálnych priestoroch (v problémoch s mnohými hľadanými parametrami). EA aj GA sú v dnešnej dobe už známe a často publikované prístupy [1-5]. Pracujú nad množinou (populáciou) potenciálnych riešení súčasne a pomocou operátorov: výber, mutácia, kríženie sú schopné nájsť obyčajne lepšie výsledky, než je to možné bežnými optimalizačnými metódami. UNS je univerzálny aproximátor, ktorý dokáže reprezentovať ľubovoľné nelineárne aj nespojité zobrazenia z $\mathbb{R}^n \rightarrow \mathbb{R}^m$ (vstupno/výstupné závislosti s n vstupmi a m výstupmi). Problémom je iba existencia vhodnej metódy na parametrizáciu UNS. Pri úlohách klasifikácie resp. rozpoznávania sú vstupné aj výstupné dáta známe (učenie s učiteľom). Naproti tomu existujú úlohy, kde výstupné dáta nie sú známe (učenie bez učiteľa). Medzi tieto úlohy patrí aj riadenie. Ideálny riadiaci zásah nie je a-priori známy, známy je iba želaný cieľ vo forme účelovej funkcie, ktorú chceme minimalizovať (maximalizovať).

Predpokladajme teraz UNS vo forme doprednej viacvrstvovej perceptrónovej siete (MLP) ako na Obr.1. Ako aktivačná funkcia vo vnútri neurónov je použitý hyperbolický tangens. Na jej vstupe je *n* premenných $x_1...x_n$, na výstupe *m* premenných $u_1...u_m$, obsahuje jednu alebo viac vnútorných (skrytých) vrstiev neurónov, pričom predpokladáme úplné prepojenie neurónov medzi jednotlivými vrstvami (každý neurón vrstvy *k* je prepojený s každým neurónom vrstvy *k*+1). Prepojenia medzi neurónmi (synaptické spoje) multiplikujú prechádzajúci signál a tieto zosilnenia predstavujú hľadané parametre UNS, ktoré sa modifikujú v procese učenia s cieľom minimalizovať zvolenú účelovú funkciu. Ľubovoľná množina všetkých týchto parametrov UNS predstavuje potenciálne riešenie zvoleného problému.

Predpokladajme teraz štandardný genetický algoritmus, ktorého veľkosť populácie je *N*. Každý prvok populácie (jedinec, chromozóm) reprezentuje vektor všetkých parametrov jednej UNS. Tvar i-teho chromozómu populácie je

$$X_{i}=\{W_{1,1,1}, \dots, W_{P,Q,R}\}$$
(1)

kde $w_{1,1,1}$ je prvá synapsa medzi vstupnou a prvou vnútornou vrstvou neurónov a $w_{P,Q,R}$ je posledná synapsa medzi poslednou vnútornou a výstupnou vrstvou neurónov.

GA prostredníctvom operácií výber, mutácia a kríženie v priebehu jednotlivých výpočtových cyklov (generácií) minimalizuje účelovú funkciu (fitness) a výsledkom je (sub)optimálne riešenie definovaného problému. Cieľom neuro-evolúcie teda je hľadanie takého optimálneho chromozómu, ktorý minimalizuje zvolenú účelovú funkciu *F*.

2. Riadenie spojitého nelineárneho dynamického systému

V prvom prípade implementujeme vyššie uvedený princíp neuro-evolúcie pri návrhu diskrétneho regulátora nelineárneho dynamického systému, ktorý je definovaný diferenčnou rovnicou

 $\Delta^2 y + \Delta y + 0.2y + y^3 = 0.01u$

Predpokladajme diskrétny spätnoväzobný regulačný obvod podľa obr.2, kde regulátor je vo forme UNS (Obr.1) so 7 vstupmi a jedným výstupom. Vstupné veličiny sú e_k , se_k , y_k , Δy_k , $\Delta^2 y$, u_{k-1} , w_k , keď $e_k = w_k - y_k$ je regulačná odchýlka, $se_k = e_k + e_{k-1}$ je suma regulačných odchýlok, $\Delta y_k = y_k - y_{k-1}$ je diferencia regulovanej veličiny, $\Delta^2 y_k = dy_k - dy_{k-1}$ je druhá diferencia regulovanej veličiny, u_{k-1} je riadiaca veličina v predchádzajúcom kroku riadenia, w_k je aktuálna žiadaná hodnota regulácie a k je krok riadenia. Výstupom UNS regulátora je aktuálna riadiaca veličina u_k .

UNS má v našom prípade dve vnútorné (skryté) vrstvy po 100 neurónov. Celkový počet synáps v UNS, ktorých hodnoty hľadáme je 7.100+100.100+100.1=10800. Veľký počet neurónov v skrytých vrstvách UNS bol zvolený vzhľadom na veľký pracovný rozsah regulácie a nelineárny charakter systému. Účelová funkcia, ktorá je minimalizovaná je vo forme

$$I_{AE} = \sum_{k=1}^{N} \left| e_k \right| \tag{2}$$

čo reprezentuje absolútnu regulačnú plochu [3-5].

Na Obr.3 je graf evolúcie účelovej funkcie (2) (fitness funkcie) v priebehu 500 generácií výpočtu GA. Na Obr.4 je priebeh regulácie pri trénovacom scenári pri zmenách žiadanej hodnoty v rozsahu od -60 po 60. Na Obr.5 je priebeh regulácie pre testovací scenár. Pre porovnanie uvádzame priebehy regulácie s PID regulátorom, ktorého parametre boli navrhnuté tiež na báze GA. Na Obr.6 je trénovací scenár a na Obr.7 je testovací scenár. Z porovnaní priebehov na Obr.4 a Obr.6 na trénovacích priebehoch, ako aj porovnaním Obr.5 a Obr.7 na testovacích priebehoch je zrejmé, že jednoduchý a lineárny PID regulátor nie je schopný zabezpečiť požadovanú kvalitu (Obr.6) a ani stabilitu (Obr.7) regulácie nelineárneho systému, na rozdiel od neuro-regulátora, ktorého kvalita regulácie je veľmi dobrá, hoci dochádza k istému preregulovaniu.



Obr.1 Dopredná umelá neurónová sieť s *n* vstupmi a *m* výstupmi.

Fig.1 Feedforward artificial neural network with *n* inputs and *m* outputs.



Obr.2 Spätnoväzobný regulačný obvod Fig.2 Feedback control loop



Obr.3 Graf evolúcie fitness funkcie Fig.3 Graph of the fitness function evolution



Obr.4 Priebeh regulácie s UNS regulátorom pre trénovací scenár, ref=w

Fig.4 System control with neurocontroller for train data, /ref=w



Obr.5 Priebeh regulácie s UNS regulátorom pre testovací scenár





Obr.6 Priebeh regulácie s PID regulátorom pre trénovací scenár





Obr.7 Priebeh regulácie s PID regulátorom pre testovací scenár

Fig.7 System control with PID controller for test data

3. Riadenie mobilného robota

Nami uvažovaný mobilný robot pozostáva z konvenčnej štvorkolesovej konfigurácie s otočnými prednými kolesami, ako je znázornené na Obr. 8 [6-9]. Informácia o prostredí pred robotom je poskytnutá snímačmi, ktoré merajú vzdialenosť medzi prekážkami a čelom robota, rovnako ako aj typ detegovanej prekážky (stena, cieľ, atď.) v niekoľkých lúčoch (Obr. 8a). Riadiacu veličinu reprezentuje uhol natočenia φ , ktorý je odchýlkou medzi pozdĺžnou osou robota a osou natočenia kolies (Obr. 8b). Bez ujmy na všeobecnosti rýchlosť považujeme za konštantnú.



Obr.8 Mobilný robot so štyrmi kolesami, a) lúče snímača merajúce vzdialenosti od prekážok,
 b) uhol natočenia φ robota

Fig.8 Mobile robot with 4 wheels, a) sensor beams measuring distances from obstacles, b) robot orientation angle ϕ

Riadiaci algoritmus obsahuje nasledovné tri kroky v každej perióde riadenia: 1. meranie vzdialeností prekážok a identifikácia ich typu (senzory S1-Sn), 2. výpočet nového uhlu natočenia φ, 3. vykonanie jedného kroku pohybu v smere φ. Výpočet aktuálneho uhla natočenia je zabezpečený pomocou UNS. Pre tento účel bola použitá UNS ako na Obr.1. Vstupné veličiny x1 ... xn, sú aktuálne hodnoty vzdialeností získaných zo senzorov S1 ... Sn. Ďalších 5 vstupov x₆ ... x₁₀ poskytuje informáciu o type prekážky (hranica, cieľ). Dve skryté vrstvy neurónov boli použité s maticami váh W_{H1} obsahujúce váhy $w_{H1,1,1}$, ..., $w_{H1,N2,N1}$ a W_{H2} s váhami WH2,1,1, ..., WH2,0,N2. Vektor váh Wo výstupnej vrstvy zahŕňa váhy wo,1, ..., wo,N2. Použitou aktivačnou funkciou f(a) je hyperbolický tangens. N_1 je počet neurónov v prvej skrytej vrstve, N₂ je počet neurónov v druhej skrytej vrstve, obe hodnoty boli stanovené na 20. Počet výstupov O je v našom prípade 1 (uhol natočenia kolies), ak by sme uvažovali aj premenlivú rýchlosť robota, pribudla by aj druhá veličina rýchlosť.

Neznáme parametre UNS riadiaceho algoritmu, ktoré sú predmetom optimalizačného procesu sú synaptické váhy $w_{i,j,k}$, kde i je poradové číslo vrstvy (1, 2 sú skryté vrstvy, 3 je výstupná vrstva), k je počet neurónov danej vrstvy a j predstavuje počet prepojených neurónov v nasledujúcej vrstve. K nájdeniu ich optimálnych hodnôt je využitý znova GA. Každý jedinec (chromozóm) populácie GA je vektor všetkých váh UNS v tvare

$$ch = \{W_{H1}, W_{H2}, W_{O}\} = \{w_{H1,1,1}, ..., w_{H1,N2,N1}, w_{H2,1,1}..., w_{H2,ON2}, w_{O1,1}, ..., w_{O1,N2}\}$$
(3)

Cieľom úlohy je nájsť také optimálne hodnoty chromozómu *chopt*, ktoré minimalizujú účelovú funkciu *fitness(ch)*

$$ch_{opt} = \arg\min fitness(ch)$$
 (4)

Účelová funkcia reprezentuje mieru úspešnosti správania mobilného robota vzhľadom na vzdialenosť k cieľu, dĺžku trajektórie, počet kolízií s prekážkami a iné. Použitý genetický algoritmus pozostáva z týchto krokov:

- 1. Náhodná inicializácia populácie o veľkosti Npop chromozómov.
- 2. Ohodnotenie každého chromozómu populácie použitím *fitness* funkcie – simulácia trajektórie robota a výpočet kriteriálnej funkcie.
- 3. Koniec, ak je dosiahnutý vopred určený počet generácií, alebo nastala ukončovacia podmienka, inak prechod na krok 4.

- 4. Výber 40% rodičov1 na kríženie, výber 40% rodičov2 na mutáciu. V oboch prípadoch je použitý ruletový výber [1-3]. Výber 2% najlepších chromozómov a výber 18% náhodných chromozómov, ktoré prežijú bez modifikácie.
- 5. Modifikácia rodičov1 jednobodovým krížením → potomkovia1. Modifikácia rodičov2 mutáciou → potomkovia2. Miera mutácie je 0.1 (10% génov populácie je mutovaných).
- 7. Kompletizácia novej populácie: potomkovia + najlepšie + nezmenené chromozómy.
- 8. Späť na krok 2.

V experimentálnom nastavení robota je použitých päť lúčov snímača s uhlami {-30°, -15°, 0°, +15°, +30°} voči pozdĺžnej osi. Všetky príslušné vstupy UNS sú normalizované do rozsahu $\langle 0,1 \rangle$. Výstup UNS (uhol natočenia φ) je v intervale

 $-\pi/2 < \phi < \pi/2$. Ďalších päť vstupov je použitých na identifikáciu typu prekážky pre každý lúč: stena, cieľ. Architektúra UNS obsahuje 10 neurónov vo vstupnej vrstve, dve skryté vrstvy po 20 neurónov a jeden neurón vo výstupnej vrstve.

Účelová funkcia (fitness) má nasledovnú formu

$$fitness = c_1.DTT + c_2.LT + c_3.C + B \tag{5}$$

DTT je euklidovská vzdialenosť medzi robotom a cieľom, *LT* je dĺžka trajektórie robota, *C* je počet kolízií medzi robotom a prekážkami a *B* je bonus, ak robot dosiahol cieľ. Váhové konštanty boli nastavené experimentálne na hodnoty c_1 =100, c_2 =100, c_3 =150, *B*=5000, aby proces evolúcie konvergoval čo najrýchlejšie. Veľkosť použitej populácie sa pohybovala medzi 80 a 160 jednotlivcov v závislosti od náročnosti problému.

Bolo uvažovaných niekoľko experimentálnych scenárov. Najjednoduchším prípadom bolo učenie robota pohybovať sa od počiatočnej do cieľovej pozície v prostredí bez prekážok (Obr. 9). Na zabezpečenie dobrej schopnosti zovšeobecňovania robota bolo počas procesu učenia použitých päť rôznych konfigurácii súradníc štart/cieľ pre výpočet jednej hodnoty účelovej funkcie pre každého jedinca populácie. Robot bol po evolúcii schopný pohybu z ľubovoľného počiatočného do ľubovoľného cieľového bodu. Je možné pozorovať, že v prípade použitia iba jedinej konfigurácie štart/cieľ pri vyhodnotení fitness sa robot naučí iba túto špecifickú trajektóriu a nebude sa schopný vysporiadať s neznámymi situáciami. Náročnejší príklad je ukázaný na Obr. 10, kde prostredie obsahuje prekážky. Na Obr. 11 je znázornený graf fitness funkcie pre dva nezávislé behy GA. Na Obr. 12 sú samostatne zobrazené grafy štyroch čiastočných kritérií fitness funkcie (5): 1. počet kolízií s prekážkami, 2. vzdialenosť od cieľa, 3. bonus za dosiahnutie cieľa, 4. dĺžka trajektórie. Na naučenie robota pohybovať sa v prostredí s prekážkami bolo potrebných v priemere šesťdesiat generácií evolúcie. Najnáročnejší scenár je ukázaný na Obr. 13. Miera úspešnosti dosiahnutia cieľa po 200 generáciách evolúcie bola 75%. Príklad neúspešného tréningového behu je ukázaný na Obr. 14. Miera úspešnosti môže byť zvýšená navýšením počtu neurónov v skrytých vrstvách, použitím rekurentných spojení v UNS, alebo použitím väčšieho množstva generácii behu GA. Experimenty boli realizované v programovacom prostredí Matlab [10].

Záver

Našim cieľom bolo ukázať, že umelá neurónová sieť môže byť úspešne implementovaná vo funkcii riadiaceho algoritmu pri rôznych prípadoch riadenia, či už ako regulátor nelineárnych dynamických systémov, v robotických aplikáciách ako je riadenie pohybu mobilného robota, ale aj v mnohých iných prípadoch, keď riadený objekt je nelineárny, má zložitú a rozvetvenú štruktúru, netriviálne správanie, keď konvenčné metódy riadenia nedávajú uspokojivé výsledky. V príspevku bola demonštrovaná metodika učenia UNS na báze evolučného algoritmu (neuro-evolúcia), ktorá nevyžaduje a-priórnu existenciu vstupno-výstupných dát, iba používateľom zvolené vhodné kritérium. Metodika neuroevolúcie vyžaduje veľký počet simulácií riešeného problému, ktorý zodpovedá súčinu veľkosti populácie a počtu generácií, čo kladie zvýšené nároky na výpočtový výkon.



Obr.9 Dráha robota zo štartu (modrá značka) do cieľa (zelená značka)





Obr.10 Dráha robota v prostredí s prekážkami Fig.10 The robot path in environment with obstacles



Obr.11 Graf evolúcie fitness funkcie pre dve spustenia GA Fig.11 Evolution of fitness function

for two GA runs



Obr.12 Graf evolúcie parciálnych kritérií fitness funkcie pre dve spustenia GA

Fig.12 Evolution of partial criteria for two GA runs



Obr.13 Dráha robota pre netriviálny prípad prostredia s prekážkami

Fig.13 The robot path for a non-trivial environment with obstacles



Obr.14 Neúspešný výsledok evolúcie Fig.14 Failed result of evolution

Poďakovanie

Článok vznikol vďaka podpore projektu s názvom: Centrum výskumu závažných ochorení a ich komplikácií, ITMS projektu: 26240120038. "Projekt je spolufinancovaný zo zdrojov EÚ. Podporujeme výskumné aktivity na Slovensku."

Literatúra

[1] GOLDBERG, D. E.: Genetic Algorithms in Search, Optimisation and Machine Learning, Addisson-Wesley, 1989.

[2] KVASNIČKA V., POSPÍCHAL J., TIŇO P.: Evolučné algoritmy, Vydavateľstvo STU, Bratislava, 2000

[3] SEKAJ I.: Evolučné výpočty, IRIS, Bratislava, 2005

[4] SEKAJ, I. Evolutionary Controller design. In: Wellington Pinheiro dos Santos. Evolutionary Computation, In-Teh, Vukovar, Chroatia, 2009 (www.intechopen.com)

[5] SEKAJ, I. Control algorithm design based on evolutionary algorithms. In: Chugo, D., Yokota, S. Introduction to Modern Robotics. iConcept Press, Hong Kong, 2011, ISBN 978-0980733068

(http://www.iconceptpress.com/books/introduction-tomodern-robotics)

[6] SICILIANO, B., KHATIB, O.: Springer handbook of robotics, Springer (2008)

[7] DUCHOŇ, F., BABINEC, A., KAJAN, M., BEŇO, P., FLOREK, M., FICO, T., JURIŠICA, L.: Path planning with modified a star algorithm for a mobile robot. Procedia Engineering, 96, 59-69, (2014)

[8] BABINEC, A., DUCHOŇ, F., DEKAN, M., PÁSZTÓ, P., & Kelemen, M.: VFH* TDT (VFH* with Time Dependent Tree): A new laser rangefinder based obstacle avoidance method designed for environment with non-static obstacles. Robotics and autonomous systems, 62(8), 1098-1115, (2014)

[9] DUCHOŇ, F., HUŇADY, D., DEKAN, M., BABINEC, A.: Optimal navigation for mobile robot in known environment. In Applied Mechanics and Materials (Vol. 282, pp. 33-38). Trans Tech Publications, (2013)

[10] Matlab R2017b, $\ensuremath{\mathbb{C}}$ 1994-2018 The MathWorks, Inc. (2017). www.mathworks.com

Abstract

In this project, we present a methodology for the control of dynamic systems design based on artificial neural networks (ANN), while the process of their learning (parameterization) is based on evolutionary algorithms. Such a neuroevolutionary approach allows the implementation of ANN in a wide range of applications. In the article, we demonstrate 2 representative implementations. The first is the control of a continuous-time, non-linear dynamic system and the second is the control of a mobile robot.

prof. Ing. Ivan Sekaj, PhD.

Slovenská technická univerzita v Bratislave Fakulta elektrotechniky a informatiky Ústav robotiky a kybernetiky Ilkovičova 3 812 19 Bratislava 1 ivan.sekaj@stuba.sk

Ing. Dominika Banásová

Slovenská technická univerzita v Bratislave Fakulta elektrotechniky a informatiky Ústav robotiky a kybernetiky Ilkovičova 3 812 19 Bratislava 1 dominika.banasova@stuba.sk

Ing. Martin Komák, PhD.

Slovenská technická univerzita v Bratislave Fakulta elektrotechniky a informatiky Ústav robotiky a kybernetiky Ilkovičova 3 812 19 Bratislava 1 martin.komak@stuba.sk

Ing. Milan Hvozdík

Slovenská technická univerzita v Bratislave Fakulta elektrotechniky a informatiky Ústav robotiky a kybernetiky Ilkovičova 3 812 19 Bratislava 1 milan.hvozdik@stuba.sk

MOŽNOSTI RIADENIA INTELIGENTNÝCH BUDOV

Juraj Slačka, Marek Buday

Abstrakt

Tento príspevok sa venuje problematike riadenia inteligentných budov. Čitateľovi spresní predstavu čo vlastne inteligentná budova znamená, aké sú jej výhody a nevýhody. Príspevok ďalej popisuje jednotlivé riadiace systémy, ktoré sa používajú v inteligentných budovách, porovnáva jednotlivých výrobcov riadiacich systémov ich možnosti, výhody a nevýhody.

Kľúčové slová: Inteligené budovy, Automatizácia, Loxone, Mervis

Úvod

Dnešná moderná doba dáva "zelenú" inteligentným budovám. Inteligentné budovy sa čoraz viac stávajú súčasťou našich miest a obydlí. Inteligentná budova je pojem , ktorý vyvoláva diskusie v mnohých skupinách ľudí. Či už ide o mladých ľudí , ktorí plánujú nový domov, cez veľkopodnikateľov ,ktorí chcú ušetriť svoje výdavky práve touto cestou až po inžinierov , ktorí takéto budovy navrhujú a riadia.

Aby sme mohli budovu označiť za inteligentnú musí spĺňať určité štandardy. A tu sa názory odborníkov rozchádzajú. Pre niektorých je budova dostatočne inteligentná , keď má inteligentné riadenie, pre iných to nie je postačujúca podmienka. Inteligenciu budovy vidia napríklad aj v tom , ako je budova navrhnutá , z akých materiálov je postavená a mnohých ďalších faktoroch.

Budova môže byť inteligentná z rôznych pohľadov a to záleží od toho , kto budovu hodnotí. Pre architekta, ktorý budovu projektoval pojem inteligencia budovy znamená niečo iné ako pre informatika , ktorý programoval jej riadenie. Z pohľadu informatika je práve riadenie tou najdôležitejšou súčasťou celej inteligentnej budovy. Nemá zmysel mať inteligentnú budovu , ktorá nebude inteligentne riadená a tým nebude naplno využitý jej potenciál.

1. Riadiace systémy pre inteligentné budovy

Výber vhodného riadiaceho systému je najdôležitejšou súčasťou návrhu riadenia inteligentnej budovy. Od výberu riadiacej jednotky sa ďalej odvíja celková hardvérová konfigurácia. V dnešnej dobre je mnoho firiem , ktoré takéto riadiace systémy vyrábajú. Avšak neexistuje monopol, ktorý by prevyšoval ostatných a mal neohrozené postavenie na trhu. To spôsobuje aj fakt, ,že medzi výrobcami je veľká konkurencia a preto sa firmy snažia eliminovať konkurentov, napríklad aj tak ,že ich riadiace systémy nie sú kompatibilné s rozširujúcimi modulmi inej firmy a tak je odberateľ nútený zakúpiť celú hardvérovú konfiguráciu práve od jednej firmy. všetky Samozrejme nie firmy robia. to Medzi najznámejších výrobcov a dodávateľov patrí Loxone. Táto spoločnosť so sídlom v Rakúskom Kollerschlagu ponúka kompletné riešenia inteligentných budov pomocou ich riadiaceho systému Minisever Obr.1. . Ponúkajú celkovú hardvérovú konfiguráciu , vlastné vývojové prostredie pre programovanie riadenia , vlastné HMI (Human Machine Interface) na ovládanie domu užívateľov a to vo forme webu ako aj mobilnej aplikácie. Mať všetky komponenty od jedného výrobcu je veľkou výhodou, keďže určite nenastane problém pri komunikácií jednotlivých modulov alebo iné podobné problémy. Avšak na druhej strane sa táto výhoda otočí vo forme ceny , ktorú je treba investovať. Preto je na mieste prehodnotiť aj iných výrobcov poprípade poskladať si celý systém od rôznych výrobcov s čo najvýhodnejším pomerom ceny , kvality a výkonu.



Obr.1 Loxone miniserver Fig.1 Loxone miniserver

Novinkou na trhu riadiacich systémov je Unipi Neuron od firmy Faster Obr. 2. Spoločnosť so sídlom v Brne ponúka viacero riadiacich systémov. Práve Unipi Neuron je porovnateľný so systémom MiniServer od firmy Loxone. Keďže firma Faster neponúka ďalšie rozširovacie moduly pre Unipi je potrebné nájsť vhodné rozširovacie moduly ,aby bola zabezpečená potrebná hardvérová konfigurácia. Vhodnou voľbou môže byť napríklad rozširovací modul Railduino od firmy SEDtronic. Pre pohodlnosť je samozrejme výhodnejšie zakúpiť celú hardvérovú konfiguráciu od jednej firmy, avšak pri vyskladaní komponentov od viacerých firiem je možné ušetriť stovky eur.



Obr.2 Unipi Neuron Fig.2 Unipi Neuron

2. Porovnanie ceny hardvérovej konfigurácie pre 2-poschodvý rodinný dom

Pre riadenie 2 poschodového rodinného domu je potrebné zakúpiť riadiacu jednotku a aspoň 3 rozširovacie moduly. Celková suma bude však vyššia , keďže treba počítať s veľkým množstvom senzorov. V tabuľke Tab.1 je uvedená cena hardvérovej konfigurácie od firmy Loxone a v tabuľke Tab. 2 je uvedená cena hardvérovej konfigurácie poskladanej od rôznych výrobcov. Cena medzi jednotlivými konfiguráciami je značne rozdielna a prevyšuje hodnotu 1000€.

Zariadenie	Cena	
Miniserver	498,00 €	
Multi Extension Air	474,00 €	
Multi Extension Air	474,00 €	
Multi Extension Air	474,00 €	
Cena spolu:	1 920,00 €	

Tab.1 Ceny Loxone komponentov

Zariadenie	Cena
Unipi Neuron M203	369,00 €
Railduino v1.3	136,81 €
Railduino v1.3	136,81 €
Railduino v1.3	136,81 €
Cena spolu:	779,43€

	Tab.2 Cen	v komponen	itov českých	výrobcov
--	-----------	------------	--------------	----------

3. Vývojové prostredia a HMI

Pre riadiacu jednotku Miniserver od firmy Loxone nie je veľmi na výber pri voľbe vývojového prostredia. Pre svoje riadiace jednotky majú vlastné vývojové prostredie Loxone Config. Keďže toto prostredie je priamo vytvorené pre programovanie logiky inteligentných budov obsahuje viac ako 100 vytvorených funkcií pre inteligentné budovy. Firma Loxone ho prezentuje ako najpokročilejší softvér pre inteligentné bývanie.

Unipi Neuron má na výber viacero vývojových prostredí. Medzi najpoužívanejšie patria Mervis , Evok, Rex, Codesys, Fhem. Práve Mervis je veľmi dobrou voľbou. Rovnako ako Loxone Config obsahuje mnoho preddefinovaných funkcií na riadenie inteligentných budov.

Human machine interface alebo rozhranie stroj - človek , v tomto prípade riadiaci systém - človek je taktiež nevyhnutnou súčasťou inteligentných budov. Keďže celá budova je riadená centrálne je vhodné mať centrálny nadriadený systém , ktorý bude riadiť človek. Existuje mnoho spôsobov ako vytvoriť HMI. Či už ide o vytvorenie webovej aplikácie, mobilnej aplikácie alebo zabudovaného riadiaceho panelu priamo v dome. Pri porovnaní týchto možností je najlepšou voľbou mobilná aplikácia pre jej flexibilnosť , keďže je možné systém ovládať takmer z celého sveta za predpokladu internetového pripojenia.

Pre spomínané riadiace systémy Unipi Neuron a Loxone Miniserver nie je problém vytvoriť respektíve zakúpiť HMI. Firma Loxone ponúka svojím zákazníkom či už webovú formu HMI, ako aj mobilnú aplikáciu. Pri riadiacom systéme Unipi Neuron je to o niečo zložitejšie. Je potrebné vybrať si vhodné vývojové prostredie, ktoré podporuje tvorbu HMI alebo vytvoriť si vlastnú mobilnú aplikáciu. Pre Unipi Neuron je z tohto ohľadu vhodné spomínané vývojové prostredie Mervis od firmy ENERGOCENTRUM.

4. Nevyhnutné súčasti inteligentných budov a komunikačné protokoly

Okrem riadiacej jednotky a rozširovacích modulov sú dôležitou súčasťou inteligentnej budovy aj moderné zariadenia ako napríklad rekuperačná jednotka , tepelné čerpadlo , kotol na vykurovanie , alarm a samozrejme množstvo snímačov. Keďže celá budova je riadená centrálne je potrebné aby informácie s týchto zariadení boli k dispozícií pri inteligentnom riadení. Inteligentné riadenie v tomto ohľade predstavuje napríklad to, že rekuperačná jednotka alebo kotol sa zapína na základe informácií zo snímačov v oknách. Je zbytočné zapínať vykurovanie ak sú otvorené okná. Informácie zo zariadení získavame pomocou komunikácie medzi zariadením a riadiacou jednotkou pomocou rôznych komunikačných protokolov.

Medzi najrozšírenejšie komunikačné protokoly v systémoch riadenia inteligentných budov patria Modbus RTU (sériová linka RS485), ktorý je používaný napríklad na komunikáciu medzi riadiacou jednotkou a rozširujúcim modulom. Na komunikáciu riadiacej jednotky s rekuperačnou jednotkou alebo vykurovacím kotlom sa zväčša používa Modbus TCP (Ethernet). Snímače sa najčastejšie pripájajú na zbernicu 1Wire alebo I2C.

Záver

Možností riadenia inteligentných budov je skutočne veľa. Pri výbere vhodného riadiaceho systému je dôležitým aspektom finančná situácia objednávateľa. Obe spomínané možnosti majú svoje výhody aj nevýhody. Najväčšou nevýhodou Miniserveru od Loxonu je jeho privysoká cena hardvéru. Nehovoriac o tom ,že je potrebné zaplatiť aj za softvér riadenia , HMI a mnoho ďalších zariadení. Riadenie pomocou Unipi Neuron od Faster je o niečo zložitejšie, avšak podstatne lacnejšie.

Abstract

This paper deals with the issues of intelligent building management. It gives the reader an idea of what an intelligent building means, what are its advantages and disadvantages. The paper further describes the individual control systems that are used in intelligent buildings, compares individual control system manufacturers and their possibilities, advantages and disadvantages.

Ing. Juraj Slačka, PhD.

Slovenská Technická Univerzita Fakulta Elektrotechniky a Informatiky Ústav Robotiky a Kybernetiky Ilkovičova 3 81219 Bratislava E-mail: juraj.slacka@stuba.sk

DETEKCIA ROVINNÝCH ÚTVAROV V OBRAZE

Juraj Slačka, Filip Zúbek

Abstrakt

Tento článok popisuje možnosti detegovania rovinných útvarov v obraze za účelom získavania dát z obrazu pre systém na lúštenie rubikovej kocky. Tento systém nie je predmetom tohto článku. Príspevok hovorí o spôsoboch ako riešiť problematiku detekcie rovinných útvarov v obraze a venuje sa aj možnostiam spracovania obrazu za účelom spresnenia a zvýšenia kvality výsledkov.

Kľúčové slová: detekcia v obraze, Cannyho detektor, Rubiková kocka, Houghová transformácia, detekcia farby, ROI, RGB, HSV, HSL

Úvod

Rovinné útvary sú základné objekty vyskytujúce sa v dvojrozmernom priestore. Tieto dvojrozmerné útvary sú pravidelného typu a dostatočne opisujú hrany rubikovho hlavolamu. Avšak ich samotná detekcia v obraze nie je triviálnou. Pre robotický systém bola použitá statická metóda detekcie hrán. Takáto metóda pozostáva vo vytvorení regiónov záujmu (ROI, z ang. Region Of Interest) [10], ktoré sú staticky umiestnené v obraze a následne z nich extrahujeme potrebné dáta (farba). Statická metóda detekcie je nenáročná na výpočtový výkon a algoritmicky je (pre detekciu) jednoduchšia, ale v prípade vychýlenia kamery môže spôsobiť problém a vyžadovať zásah operátora.

Riešením takéhoto problému je dynamická metóda detekcie, pričom narastá zložitosť a komplexnosť problematiky. Pre ešte väčšiu rozmanitosť zadania sme sa rozhodli riešiť hlavolam rubikovej kocky, typu "Molecul", ktorá je vyobrazená na obrázku č.1. Rozmer takejto kocky je rovnaký ako pri rubikovej kocke t.j. 3x3x3, pričom hrany hlavolamu sú v tvare gule. V takomto prípade uvažujeme detekciu kruhových útvarov v obraze. Pre detegovanú stranu natočenú oproti kamere sú takéto hrany detegované ako kruh. Pre detekciu kruhu bola použitá Houghová transformácia pre kruh (Sekcia 1.4.2). Riešenie pomocou dynamickej metódy detekcie je ďaleko robustnejšie, avšak narastajú nároky na výpočtový výkon a zložitosť algoritmov pre detekciu. Nároky sa zvyšujú v závislosti od natočenia a osvetlenia kocky, pričom vzniká chyba



1. Dynamická metóda detekcie

Dynamická metóda detekcie je používaná pre hľadanie objektov v obraze. V obraze rozlišujeme objekty pomocou hrán objektu. Hrany objektu dokážeme určiť na základe zmeny intenzity bodu v obraze oproti okolitým bodom. Body obrazu reprezentujú pixely, pričom obraz tvoria pixely v 2D (dvoj-rozmernom) priestore. Každý pixel je definovaný polohou a farbou. Hľadáme teda zmenu intenzity pixelu obrazu v závislosti od okolitých pixelov v smere osi x a v smere osi y. Takúto zmenu intenzity reprezentuje gradient obrazu. [7],[4]

1.1 Gradient obrazu

Gradient obrazu je zmena intenzity bodu v smere, vzhľadom na okolie bodu. [7],[4] Hovorí nám o tom ako sa zmenila intenzita v smere osi x a v smere osi y oproti okolitým bodom ako môžeme vidieť v rovnici (1).

$$\Delta f = \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$
(1)

V obraze zloženého z pixelov sa vieme pohybovať len diskrétne a teda sa používa diskrétna forma gradientu. Pomocou rovnice (2) dokážeme vypočítať veľkosť vektora gradientu a pomocou rovnice (3) smer vektora gradientu.

$$|G| = \sqrt{g_x^2 + g_y^2} \tag{2}$$

$$\theta = tan^{-1} \begin{bmatrix} g_y \\ g_x \end{bmatrix}$$
(3)

Gradient obrazu tvorí základ pre detekciu hrán v obraze. Výpočet gradientu v obraze pre každý pixel vyžaduje množstvo výpočtov a pre výpočet v reálnom čase vzniká obrovský nárok na výpočtový výkon. Túto problematiku rieši Sobel-Feldmanov operátor alebo Sobelov filter ktorý je reprezentovaný ako isotropický 3x3 operátor pre gradient. [8] Sobelov filter je založený na konvolenčnom súčine obrazu a gradientu v smere osi x a v smere osi y.

Obr.1 Rubikov hlavolam typu "molecube"

1.2 Cannyho detektor

Cannyho detektor je algoritmus na zistenie hrán v obraze ktorý využíva práve Sobelov filter vo vertikálnom a horizontálnom smere. [5],[8] Výstupom Sobelovho filtra sú hrany v obraze, ktoré stále nemusia byť skutočnou hranou hľadaného objektu. Práve tieto hrany sa snaží algoritmus potlačiť.



Obr.2 Určovanie skutočnej hrany [5]

Algoritmus prechádza všetky pixely a v okolí pixelu hľadá lokálne maximum v smere gradientu. Na Obr. 2. môžeme vidieť práve takýto prípad. Bod A leží na hrane a body B a C ležia v smere gradientu, ktorý je kolmý na hranu. Algoritmus hľadá v bode A lokálne maximum. Pokiaľ takéto maximum nájde body nechá v obraze, inak ich potlačí. [5]

Následne sa hľadajú väzby medzi gradientom a bodom ktorý skutočne tvorí hranu, a ostatné algoritmus podtlačí v závislosti od thresholdu. Na Obr. 3. je zobrazené rozhodovanie algoritmu či sa skutočne jedná o hranu. Hodnoty minVal a maxVal tvoria hranice thresholdu na základe ktorých sa algoritmus rozhoduje. Pokiaľ je hodnota gradientu hrany väčšia ako maxVal, hrana je prehlásená za skutočnú tzv. "sure-edge"(bod A) a je ponechaná. Pokiaľ sa bod nachádza medzi týmito hodnotami thresholdu a hrana nadväzuje na skutočnú hranu, hrana je ponechaná (bod C). V prípade že sa hrana nachádza medzi hranicami thresholdu a nemá nadväznosť na skutočnú hranu alebo sa nachádza pod spodnou hranicou (maxVal), hrana je potlačená (bod B). Výsledkom sú hrany objektu, bez tých hrán, ktoré by mohli byť detegované vplyvom šumu alebo inými nežiadúcimi javmi. [5]



Obr.3 Určovanie väzby so skutočnou hranou [5]

1.3 Predspracovanie obrazu

Pred použitím Cannyho detektora hrán je vhodné obraz spracovať do podoby kde pixely závisia od intenzity obrazu, pomocou čierno-bieleho filtra. Keďže detekcia hrán vyplýva z diferencii v obraze (detekcia hrán je citlivá na šum), je potrebné obraz zbaviť šumu čo sa prevažne aplikuje pomocou Gaussovej filtrácie (Obr.4.).



Obr.4 Gaussov filter

1.4 Houghová transformácia

Houghová transformácia (skratka HT) sa môže použiť na zistenie ľubovoľného tvaru objektu v obraze, ktorý sa dá opísať pomocou matematickej funkcie.[2],[3] Pri detekcii klasickej rubikovej kocky sú hrany takéhoto hlavolamu reprezentované ako štvorec. V prípade, že uvažujeme hlavolam typu "molecube" a detegovaná strana je umiestnená oproti snímaciemu zariadeniu, detegujeme takúto hranu ako kruh.

1.4.1 Houghová transformácia pre čiaru

Pre získanie štvorca z obrazu sme uvažovali metódu hľadania strán štvorcového útvar. Každá strana štvorca je reprezentovaná ako čiara, ktorej matematický opis je

$$y = ax + b \tag{4}$$

$$b = x\cos(\theta) + y\sin(\theta)$$
(5)

kde x a y reprezentujú súradnice bodu v obraze (rovnica č.4) a kde b je vzdialenosť od počiatku a θ je uhol pod ktorým sa čiara nachádza. Z uvedených rovníc číslo č.4 a č.5 vieme. Rovnica číslo 5 vyjadruje vzťah medzi súradnicovým priestorom obrazu a tzv. Houghovím priestorom do ktorého obraz transformujme. V takomto priestore vieme opísať čiaru pomocou dvojice údajov (b, θ). Pre prácu s takýmto priestorom bolo vytvorené dvojrozmerné pole ktoré je aj inak nazývané akumulátor.

Akumulátor slúži pre ukladanie hodnotenia bodov v transformácii kde hodnoty b tvoria stĺpce a hodnoty θ tvoria riadky. Každý bod ktorý bol detegovaný ako hrana je následne preložený čiarou ktorú rotujeme okolo osi bodu. Hodnotenie pre generovanú čiaru sa zvyšuje vždy keď generovaná čiara prechádza niektorým z ďalších bodov detegovaných hrán. Pomocou takéhoto priestoru vieme každý bod opísať čiarou. Takto opíšeme každý bod hrany a dostávame transformáciu pôvodného obrazu. Vo výsledku sa vyhodnocuje bod s najväčším hodnotením t.j. bod kde sa po transformácii čiary pretli najviac krát, a dostávame dvojicu (b, θ), ktorá reprezentuje čiaru v obraze. [2],[3],[14]

Vykreslením výslednej čiary Houghovej transformácie do obrazu s hranami rubikovej kocky, získame hrany rubikovej kocky preložené čiarami (tak ako môžeme vidieť na Obr. 5.).



Obr.5 Hrany preložené čiarami z Houghovej transformácie

Všimnime si na Obr. 6. uhol θ ktorý zviera čiara so súradnicovým systémom ide v protismere hodinových ručičiek. To znamená, že čiara pod uhlom 0° je orientovaná vertikálne a v prípade že uhol θ je 90° čiara je orientovaná horizontálne. Smer osí závisí od implementácie.



Obr.6 Opis súradníc čiary v Houghovom priestore[3]

1.4.2 Houghová transformácia pre kruh

Detekcia kruhu využíva tak isto vlastnosti Houghovho priestoru, pričom sa líši transformačný vzťah v dôsledku zmeny matematického zápisu hľadaného útvaru z ktorého transformácia vyplýva.

$$x = a + R\cos\theta \tag{6}$$

$$y = b + Rsin\theta \tag{7}$$

V rovniciach č.6 a č.7 [17],[18] sa nachádza matematický opis kruhu kde x a y sú súradnice bodu v obraze. Hľadaná trojica parametrov (a, b, R) označuje súradnice stredu kružnice a polomer kružnice. Vzhľadom na počet neznámych parametrov bol vytvorený trojrozmerný akumulátor hodnotenia možných výsledkov. Postup detekcie je podobný ako pri detekcii čiary (Sekcia 1.4.1). Pre každý bod z Cannyho detektora (x, y) bola generovaná kružnica pomocou rovníc č.6 a č.7 pre rôzne R a θ . Pre každú časť tejto kružnice sme priradili hodnotu do akumulátora. Hodnota akumulátora s najväčším hodnotením (obr.č.7) je prehlásená za stred kružnice a vráti nám (a, b, R) práve tejto kružnice. [17],[18]



Obr.7 Vykreslenie akumulátora detekcie dvoch kružníc pomocou HT

2. Detekcia farby rubikovej kocky

Pre lúštenie hlavolamu rubikovej kocky je potrebné poznať farebné rozloženie kocky, ktoré nám hovorí o stave, v ktorom sa kocka nachádza. Farebné rozloženie kocky určuje plocha každej strany rubikovej kocky. Po aplikácii Cannyho detektora sme získali detegovanú plochu rubikovej kocky v obraze. V prípade klasickej rubikovej kocky má táto detegovaná plocha svoje hrany preložené čiarami, ktoré sme získali pomocou Houghovej transformácie. Tieto čiary nám rozdeľujú obraz kocky na ďalších deväť podpriestorov, ktorých hlavnou vlastnosťou je farba. Pre hlavolam typu "molecube" sme takto detegovanú plochu opísali kružnicami získaných z Houghovej transformácie a tým získali deväť podpriestorov pre extrakciu farby z obrazu.

Každá farba je reprezentovaná pomocou pixelov v príslušnom formáte, kde každý pixel je definovaný farbou a polohou. Formát farby je priestor ktorý opisujú súradnice v závislosti od modelu farebného priestoru a na základe takýchto súradníc, nám určuje farbu daného bodu. Najbežnejším výstupným formátom snímacích zariadení je RGB.

2.1 RGB model

RGB (Red,Green,Blue) model je aditívny farebný model pri ktorom daná farba vzniká zmiešaním červenej, zelenej a modrej farby príslušnej intenzity. Charakteristikou pre tieto farby je citlivosť ľudského oka práve na vlnové dĺžky týchto troch farieb. Všetky viditeľné farby vnímané ľudským okom sú kombináciou týchto troch zložiek. Vytvárané kombinácie sú založené na aditívnom spôsobe, t.j. sčítavanie intenzity každej z príslušných zložiek. V najbežnejšej forme sa RGB formát prezentuje ako osem bitové číslo, pre každú zložku v diskrétnej forme, kde čierna farba je reprezentovaná ako trojica čísel (0,0,0), pričom nie je žiadna zo zložiek prítomná a biela farba je reprezentovaná ako trojica čísel (255,255,255), kde prítomnosť každej z trojice farieb je v plnej intenzite. [11]



Obr.8 Kubický model priestoru RGB [12]

Na Obr. 8. môžeme vidieť opísaný kubický model farebného priestoru RGB, pričom intenzita jednotlivých farieb narastá v smere šípok popísaných v obrázku.

2.2 HSV/HSL model

Detekcia farby zosnímaného obrazu je citlivá na osvetlenie detegovanej časti (pri snímaní lesklých povrchov ako sú napríklad nálepky rubikovej kocky). Pri veľkom osvetlení sa môže snímaná farba javiť pod úplne iným odtieňom prípadne pod inou farbou a farba nemusí byť správne detegovaná. Aby sme takýto jav eliminovali čo najviac, volíme vhodný threshold, ktorý nám určuje prípustné vychýlenie v odtieňoch detegovanej farby. Spomínaný RGB formát ťažko opisuje odtiene farieb a z tohoto dôvodu sa používa niektorý z alternatívnych modelov farebných priestorov (ako sú HSV alebo HSL modely).

HSV a HSL sú alternatívne cylindrické modely navrhnuté pre bližšie opísanie vnímania farieb ľudským okom. V modeloch HSV a HSL reprezentujú farbu tri zložky (viď Obr. 9. a Obr. 10.). Zložka "Hue" reprezentuje farby a odtiene farieb, a v cylindrickom modeli je rozdelená radiálne po obvode valcového modelu. Saturácia (sýtosť) farby je určená polomerom v takomto modeli a to od stredu po okraj, pričom na okraji má farba najväčšiu sýtosť. Poslednou zložkou HSV modelu je hodnota jasu (označovaná z ang. "Value"), ktorá určuje výšku jasu, pričom spodná hodnota reprezentuje čiernu farbu (nízky jas) a vrchná hodnota predstavuje bielu farbu (vysoký jas).[16]



Obr.9 Cylindrický model priestoru HSV [12]

HSV a HSL modely sú si veľmi podobné, avšak HSL model vníma svetlosť (z ang. "Ligtness") a saturáciu ako dve nezávislé veličiny, čo si môžeme všimnúť na Obr. 10. Tu nastáva prípad kde na rozdiel od HSV modelu môže mať skoro biela farba 100% sýtosť. [16]



Obr.10 Cylindrický model priestoru HSL [12]

2.3 Maskovanie farby

Maskovanie farby nám umožňuje odfiltrovať všetky farby v obraze okrem hľadaného spektra farieb. Hľadanú farbu potom vieme určiť pomocou masky v príslušnom farebnom formáte (pre výhody HSV/HSL formátu je vhodnejšie pri maskovaní spraviť konverziu z RGB do HSV/HSL modelu.). Masku vytvoríme pomocou hľadanej farby vo zvolenom farebnom modeli, kde aplikujeme threshold na niektoré z parametrov určujúce farbu, podľa zvoleného farebného modelu (pre HSV/HSL model viď odsek 2.2), pričom dostávame interval hľadaných farieb a odtieňov. Zlúčením takejto masky a obrazu v ktorom hľadáme danú farbu, dostávame bitový obraz (mapu), na základe ktorého vieme prehlásiť či sa hľadaná farba nachádza v obraze (pixel má hodnotu logickei 1) alebo či sa farba z rozsahu nenachádza v obraze (pixel má hodnotu logickej 0). V takomto obraze sa poloha pixelov zhoduje s pôvodným obrazom. [13],[16]



Obr.11 ROI pred a po maskovaní červenej farby

Na Obr.11 môžeme vidieť príklad maskovania pre klasickú rubikovú kocku, kde bol obraz orezaný na veľkosť jedného z regiónov záujmu (vľavo) a následne bola maskovaná červená farba vo formáte HSV pomocou knižníc OpenCV (vpravo).

Záver

Tento príspevok sa zaoberal problematikou detekcie v obraze pre robotický systém na lúštenie rubikovej kocky. Poukázal na možnosti riešenia tejto problematiky pomocou statickej a dynamickej metódy. Podrobnejšie sme sa venovali dynamickej metóde pre rubikovú kocku a prispôsobili aj pre iný typ rubikovho hlavolamu konkrétne "molecube". Dynamická metóda odstraňuje nedostatky statickej metódy za cenu potreby vyššieho výpočtového výkonu a zložitosti algoritmu pre detekciu. Dynamická metóda je robustnejšia v prípade výskytu problémov, ako je napríklad natočenie kocky, a nežiada si prítomnosť operátora tak často ako to je v prípade statickej metódy, ale dokáže si s niektorými defektami poradiť sama. V tomto príspevku sme ozrejmili aj výhody farebného priestoru HSV, ktorý dokáže lepšie pracovať s jasom farby a teda je vhodnejší pri detekcii farby z obrazu ako farebný priestor RGB. Vytvorili sme masku farby, ktorá poslúžila ako filter a na základe tohoto filtra sme dokázali detegovať farbu v obraze (viď Obr.11.). Výsledkom tohoto príspevku je detekcia hrán rubikovej kocky (Obr. 5.) a detekcia farieb (Obr. 11.).

Literatúra

[1] Wang Mingang ,Wang Chao ,Fan Yingping, Research for Target Recognition of Infrared Bridge Based on Morphological operator and Bridge Template, 2012, 978-0-7695-4719-0, 2012 International Conference on Computer Science and Service System.

[2] Nebojsa M. Ralevic, Slobodan Drazic, Radovan Obradovic, The Hough transformation of rectangle, 2008, 2008 6th International Symposium on Intelligent Systems and Informatics

[3] Hough Line Transform, dostupné online: https://docs.opencv.org/3.0-beta/doc/py_tutorials/py_imgproc/py_houghlines/py_houghlines.html

[4] Image Gradient, dostupné online: https://docs.opencv.org/3.0-beta/doc/py_tutorials/py_imgproc/py_gradients/py_gradients.html#gradients

[5] Canny Edge Detection, dostupné online: https://docs.opencv.org/3.0-beta/doc/py_tutorials/py_imgproc/py_canny/py_canny.html#canny

[6] Oravec Jakub, Houghova transformácia a jej použitie v oblasti spracovania obrazu, 2017, ISSN 1338-0087
[7] Jacobs, David, Image graddients, 2005, Class Notes for CMSC 426,

[8] Mouna Afif, Yahia Said, Haythem Bahri, Mohamed Atri, Efficient implementation of sobel filter baed on GPUs cards, 2016, 978-1-5090-1645-7, 2016 International Image Processing, Applications and Systems (IPAS)

[9] Tianzheng Wang, Qingang An, Jie Li, Yujia Zhang, Junyu Han, Shuai Wang, Shiying Sun, Xiaoguang Zhao, Vision-based illegal human ladder climbing action recognition in substation, 2017, 978-1-5090-4726-0, 2017 Ninth International Conference on Advanced Computational Intelligence (ICACI) [10] Basic Operations on Images, dostupné online: https://docs.opencv.org/3.4.2/de/d06/tutorial_js_basic_ops.html

[11] Balkrishan Ahirwalk, Mahesh Khadtare, Rakesh Mehta, FPGA based system for color space transformation RGB to YIQ and YcbCr, 2007, 978-1-4244-1355-3, 2007 International Conference on Intelligent and Advanced Systems

[12] Thresholding Operations using inRange, dostupné online: https://docs.opencv.org/3.4.3/da/d97/tutorial_threshold_inRange.html

[13] Image Segmenatation Using Color Spaces in OpenCV + Python, Rebecca Stone, dostupné online: https://realpython.com/python-opencv-color-spaces/

[14] Oliver Clark, The development of the UAV image detection software and control, 2015, Development of the University of Liverpool Entry in the Proposed ImechE UAS Challenge.

[15] Baohua Jin, Yong Gan, Yongquan Xia, An Object Region Extraction Approach Based on Interest Pixel Detection and Inner Filling Strategy, 2010, 978-1-4244-6596-5, 2010 International Conference on Machine Vision and Human-machine Interface

[16] Rafael C. Gonzalez, Richard Eugene Woods, Digital Image Processing, 2008, 0-13-168728-X , 3rd ed.Upper Saddle River, Prentice Hall

[17] Janak Trivedi, Mandalapu Sarada Devi, Dave Dhara, OpenCV and Matlab based car parking system module for smart city using circle hough transform, 2018, 978-1-5386-1887-5, 2017 International Conference on Energy, Comunication, Data Analytics and Soft Computing (ICECDS)

[18] Nova Hadi Lestriandoko, Rifki Sadikin, Circle detection based on Hough transform and Mexican Hat filter, 2017, 978-1-5090-2323-3, 2016 International Conference on Computer, Control, Informatics and its Aplications (IC3INA)

Abstract

This paper describes the possibilities of detecting planar formations in an image to extract image data for a rubic cube solving system. The paper talks about possibilities of solving problems of detection of planar objects in the image and also deals with the possibilities of image processing in order to increase accuracy of detection and to improve the results.

Ing. Juraj Slačka PhD. Bc. Filip Zúbek

Slovenská Technická Univerzita v Bratislave Fakulta Elektrotechniky a Informatiky Ústav Robotiky a Kybernetiky Ilkovičova 2961/3, 841 04 Karlova Ves xzubek@stuba.sk juraj.slacka@stuba.sk

Archív

Archive

2001 AT&P journal PLUS 1: Adaptívne a nelineárne riadenie systémov (tlačená verzia) Adaptive and nonlinear control systems (printed version) AT&P journal PLUS 2: Robotika, mechatronika, diskrétne výrobné systémy (tlačená verzia) Robotics, mechatronics, discrete manufacturing systems (printed version) 2002 AT&P journal PLUS 3: Robustné systémy riadenia (tlačená verzia) Robust control systems (printed version) 2003 AT&P journal PLUS 4: Samonastavujúce sa systémy v riadení procesov (tlačená verzia) Selftuning systems in process control (printed version) 2004 AT&P journal PLUS 5: Robotické systémy (elektronická – CD verzia) Robotics systems (electronic - CD version) 2005 AT&P journal PLUS 6: Mechatronika (elektronická – CD verzia) Mechatronics (electronic – CD version) AT&P journal PLUS 7: Umelá inteligencia v praxi (elektronická – CD verzia) Artifical intelligence in Practise (electronic - CD version) 2006 AT&P journal PLUS 1: Mechatronické systémy (elektronická – CD verzia) Mechatronic systems (electronic - CD version) AT&P journal PLUS 2: Inteligentné meracie systémy (elektronická – CD verzia) Intelligent measurement systems (electronic - CD version) 2007 AT&P journal PLUS 1: MMaMS'2007 (elektronická – CD verzia) MMaMS'2007 (electronic - CD version) AT&P journal PLUS 2: Riadenie procesov (elektronická - CD verzia) Process Control (electronic - CD version) 2008 AT&P journal PLUS 1: Mobilné robotické systémy (elektronická – CD verzia) Mobile robotic systems (electronic - CD version) AT&P journal PLUS 2: Riadenie v energetike (elektronická – CD verzia) Control of Power Systems (electronic – CD version) 2009 AT&P journal PLUS 1: Inteligentné pohybové systémy (elektronická – CD verzia) Intelligent motion control systems (electronic - CD version) AT&P journal PLUS 2: Riadenie procesov (elektronická – CD verzia) Process control (electronic - CD version) 2010 AT&P journal PLUS 1: Systémy automatického riadenia (elektronická - CD verzia) Systems of automatic control (electronic - CD version) AT&P journal PLUS 2: Robotika vo vzdelávaní (elektronická – CD verzia) Robotics in education (electronic – CD version) 2011 ATP Journal PLUS 1: Systémy automatického riadenia (elektronická - CD verzia) Systems of automatic control (electronic - CD version) ATP Journal PLUS 2: Riadenie procesov (elektronická - CD verzia) Process Control (electronic – CD version) 2012 ATP Journal PLUS 1: Modelovanie mechanických a mechatronických sústav (elektronická - CD verzia) Modelling of Mechanical and Mechatronic Systems (electronic – CD version) 2013 ATP Journal PLUS 1: Robtep 2012 (elektronická – CD verzia) Robtep 2012 (electronic - CD version) ATP Journal PLUS 2: Riadenie dopravných a priemyselných procesov (elektronická – CD verzia) Control of transport and industrial processes (electronic - CD version) 2019 ATP Journal PLUS 1: Aktivity Katedry riadiacich a informačných systémov Žilinskej univerzity v Žiline v oblasti riadenia priemyselných a dopravných procesov (elektronická – CD verzia) Activities of the Department of Control and Information Systems University of Zilina in the field of industrial and transport processes (electronic - CD version)

> www.atpjournal.sk ATP Journal PLUS 2/2019