

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Workshop Perspektivní projekty vývoje řídicích a senzorických technologií

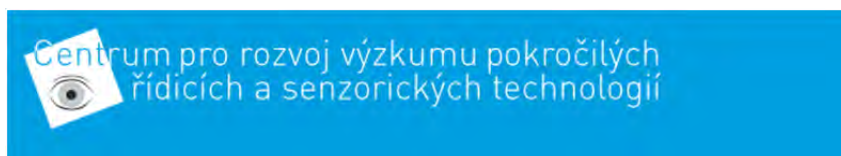
Sborník příspěvků

Místo konání:

hotel Ráztoka, Trojanovice

Datum:

27.-29. června.2011



Centrum pro rozvoj výzkumu pokročilých řídicích a senzorických technologií CZ.1.07/2.3.00/09.0031

TENTO SBORNÍK JE SPOLUFINANCOVÁN EVROPSKÝM SOCIÁLNÍM FONDEM
A STÁTNÍM ROZPOČTEM ČESKÉ REPUBLIKY



INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Workshop Perspektivní projekty vývoje řídicích a senzorických technologií

Sborník příspěvků

Místo konání:

hotel Ráztoka, Trojanovice

Datum:

27.-29. června.2011

Centrum pro rozvoj výzkumu pokročilých řídicích a senzorických technologií CZ.1.07/2.3.00/09.0031

TENTO SBORNÍK JE SPOLUFINANCOVÁN EVROPSKÝM SOCIÁLNÍM FONDEM
A STÁTNÍM ROZPOČTEM ČESKÉ REPUBLIKY

Název: Workshop Perspektivní projekty vývoje řídicích a senzorických
technologií
Editoři: Doc. Ing. Pavel Václavek, Ph.D.
Doc. Ing. Petr Beneš, Ph.D.
Vydavatel: Vysoké učení technické v Brně,
Fakulta elektrotechniky a komunikačních technologií
Odpovědný redaktor: doc. Ing. Petr Blaha, Ph.D.
Tiskárna: LITERA Brno
Tábor 43a
612 00 Brno
Počet stránek: 240
Vydáno v roce: 2011
Vydání: první
Počet výtisků: 50

Za obsahovou a jazykovou správnost odpovídají autoři.

© Vysoké učení technické v Brně

ISBN 978-80-214-4297-9

OBSAH

Vyžádané příspěvky	1
Řídicí systém pro sklářský stroj ISS firmy Sklostroj	
<i>Karel Dočkal</i>	3
Capacitive sensors with precalculable capacitance	
Stanislav Ďaďo	9
Prodej a ochrana duševního vlastnictví	
<i>Tomáš Halva</i>	19
Modelování a prediktivní řízení spalovacího motoru	
<i>Daniel Pachner</i>	22
Přiřazení Jordanovy formy zpětnou vazbou: metoda, aplikace, podpůrný software	
<i>Miloš Schlegel</i>	27
Řádová analýza signálů z točivých strojů s proměnlivými nebo neustálenými otáčkami	
<i>Jiří Tůma</i>	41
Nejnovější trendy výzkumu inteligentních řídicích systémů	
<i>Pavel Vrba</i>	59
Řídicí technika	67
Některé aplikace řídicího systému REX	
<i>Pavel Balda, Miloš Schlegel</i>	69
Rekuperace energie mobilních robotů	
<i>František Burian, Luděk Žalud</i>	76
Jednotná platforma pro sdílení dat a dokumentů při řešení projektů výzkumnými konsorciemi	
<i>Filip Buršík, Martin Sýkora, Martin Heneš</i>	81
Realizace gravírovacího zařízení	
<i>Pavel Číp, Tomáš Babinec</i>	89
Adaptivní učení modelu ARMAX rekurzivní klouzavou metodou Levenberg-Marquardt	
<i>Jakub Dokoupil, Vladimír Burlak</i>	95

Vylepšení vlastností samočinně se nastavujících regulátorů s identifikací založenou na neuronových sítích	
<i>Vlastimil Lorenc</i>	102
Akcelerace hashovacího algoritmu MD5 za pomoci FPGA	
<i>Daniel Piši</i>	107
Syntéza regulátoru asynchronního motoru metodou smíšených citlivostních funkcí	
<i>Lukáš Pohl</i>	112
Různé způsoby výpočtu příčné indukčnosti u online metody vylepšené frekvenční analýzy	
<i>Ivo Veselý, Dušan Zámečník</i>	118
Možnosti bezsnímačové řízení synchronního motoru s permanentními magnety v oblasti nízkých otáček	
<i>Libor Veselý</i>	124
Prediktivní proudový regulátor v řízení asynchronního motoru	
<i>Dušan Zámečník, Ivo Veselý</i>	130
Senzory a měření	137
Zařízení pro optickou detekci a kontrolu předložené obrazové informace	
<i>Tomáš Babinec, Pavel Číp</i>	139
Stabilizace frekvenční charakteristiky piezoelektrické keramiky částečnou depolarizací tepelnou vlnou	
<i>Jiří Fialka</i>	145
Malý průzkumný robot – snímače, elektronika a komponenty	
<i>Tomáš Florián</i>	151
Měřicí soustava pro sledování chybových artefaktů fotovoltaických solárních článků metodou elektroluminiscence	
<i>Kristýna Friedrischková, Bohumil Horák</i>	156
Measurements of multi-copter propulsion systems	
<i>Jan Gebauer</i>	162
Využití mems akcelerometru pro měření jízdního pohodlí na sedadle	
<i>Martin Juránek</i>	170

Head tracking - sensors and test bed requirements	
<i>Lukáš Kopečný, Luděk Žalud</i>	176
Vestavěný senzorový systém pro sběr a vyhodnocování biomedicínských dat v reálném čase	
<i>Tereza Otáhalová, Zdeněk Slanina, David Vala</i>	181
Vizuální inspekce v klinické praxi	
<i>Radka Pustkova, David Vala, Zdeněk Slanina</i>	186
Time measurement synchronization using gps units	
<i>Zdeněk Slanina, Vladimír Kašík, Karel Musil</i>	192
Photogrammetry in mining hole inspection using MCU controled LED flash	
<i>Dagmar Šeligová, David Vala</i>	198
Methodology for assessing efficiency of photovoltaic power plant	
<i>Michal Tutsch, Petr Vojčínák</i>	202
Srovnání metod analýzy signálu akustické emise v časové a frekvenční oblasti při měření průtoku dvoufázových médií	
<i>Miroslav Uher</i>	209
Malý datalogger s flash kartou	
<i>Martin Vágner</i>	215
Using software applications for modelling of temperature conditions inside geothermal resources	
<i>Petr Vojčínák, Radovan Hájovský, Jiří Koziorek, Martin Pieš</i>	220
Index autorů	233

Vyžádané příspěvky

ŘÍDICÍ SYSTÉM PRO SKLÁŘSKÝ STROJ ISS FIRMY SKLOSTROJ

Karel DOČKAL

Siemens s.r.o. divize Industry, oddělení Motion Control Production Machines

Olomoucká 7/9, Brno, 618 00

E-mail: karel.dockal@siemens.com

Abstrakt/Abstract:

The article describe Individual section (IS) machine for production of container glass, especially oriented on IS machines of company Sklostroj Turnov CZ s.r.o. It also describe basic principles, which was required on Siemens Motion Control system Simotion during development of control system for new Sklostroj machine ISS. This totally new Individual Section machine has most of mechanisms driven by synchronous servomotors, instead of old fashion pneumatic drives. The IS machine required high level of modularity, extremely precise synchronization of the mechanisms, modern and powerful communication concept and high level of safety. All this requirements has been fulfilled by Simotion with communication over IRT (Isochronous Real Time) PROFINET

Klíčová slova/Keywords: IS machine, container glass, Sklostroj, ISS, Simotion, Motion Control, PROFINET IRT.

1 Úvod

Firma Sklostroj Turnov CZ s.r.o. je renomovaným výrobcem řadových sklářských strojů (IS – Individual section machine) pro výrobu obalového skla. Zabývá se jak dodávkou samostatných IS strojů, tak kompletními dodávkami sklářských linek. Firma patří k těm několika českým firmám, které sázejí na inovativní přístupy, moderní technologie a vývoj.

Spolupráce firmy Sklostroj s firmou Siemens začala již mnoha lety a v cca posledních 5 letech je velmi intenzivní. Firma Siemens spolupracovala na vývoji nových mechanismů pro standardní sklářské stroje, které dříve byly ovládané pneumaticky. Kvůli požadavkům na vysokou přesnost, dynamiku a synchronizaci s jinými mechanismy se začaly vyvíjet řešení s elektrickými pohony. Takto jsme například spolupracovali při vývoji mechanismů obracečů, odnímačů, nůžek atd. Na základě zkušeností s firmou Siemens, nás firma Sklostroj Turnov CZ s.r.o. v roce 2009 požádala o pomoc při vývoji nové generace IS strojů.

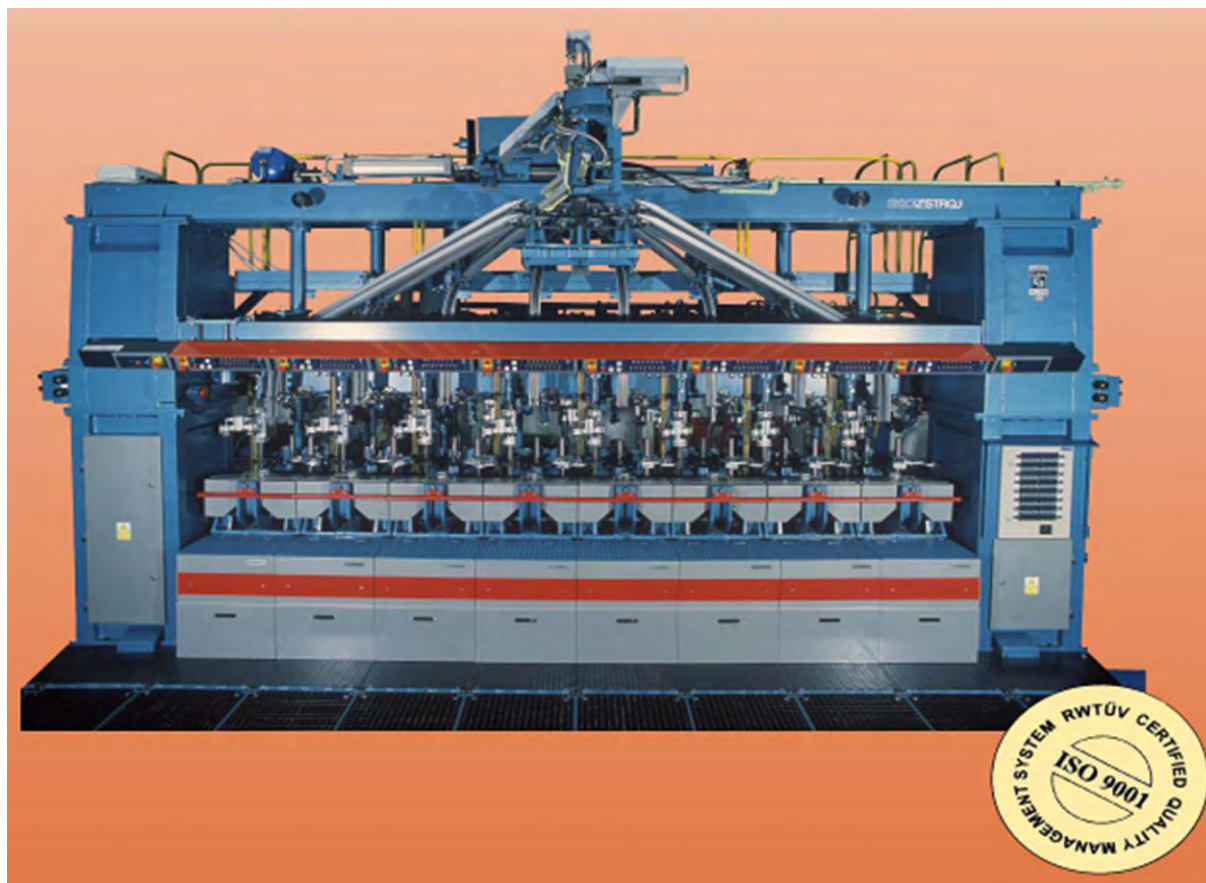
2 Popis sklářského IS stroje.

IS stroj je zařízení, které se skládá z jedné centrální společné části. Jejímž úkolem je připravit kapky skla a přivést je do jedné ze sekcí stroje, které jsou určeny k vlastnímu vytvarování lahve. Podle toho kolik lahví se během jednoho taktu v sekci vyrábí, se stroj nazývá 2-5 kapkový stroj. Nejčastější jsou dnes 2 kapkové, nebo 3 kapkové stroje.

2.1 Centrální část stroje

Nad vlastním strojem se nachází kanál přivádějící rozžhavené sklo ze sklářské pece. Na konci kanálu se nachází takzvaný dávkovač, který má ve spodní části 2-5 děr, dle toho o kolika kapkový stroj se jedná. Aby vytékalo těmito děrami správné množství skla, je sklo do děr vtlačováno takzvaným plunžrem. Pod dávkovačem se nachází nůžky, které tekoucí sklo stíhají na kapky. Jen pro představu u 12-ti sekčního stroje při rychlosti jedné sekce blížící se

dnes cca 20 taktům za minutu. Musí tyto nůžky řízené servomotorem provádět až 240 stříhů za minutu! Pod nůžkami se nachází takzvaný distributor, který volně padající kapky směruje postupně do svodů (žlabů) vedoucích do jednotlivých sekcí!



Obrázek 1: Celkový pohled na sklářský IS stroj firmy Sklostroj typu AL

2.2 Tvarovací sekce stroje

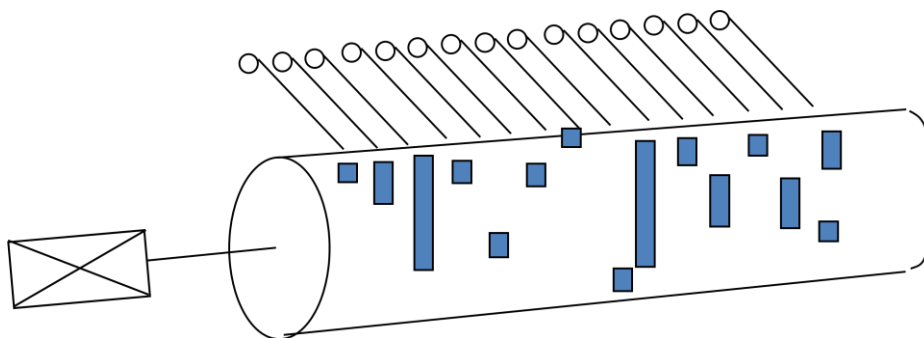
Každá sekce se skládá ze dvou skupin forem a některých mechanismů kolem nich. Ve chvíli kdy kapky skla dopadnou do rozevřené přední formy, tato se sevře a vytvaruje ústí lahve. Zároveň se do ní zespoda vsune ústník, který udělá do hrdla díru a vytvaruje tak zárodek budoucí lahve. Zárodek lahve je tedy „vzhůru nohama“. Přední forma se rozevře ale nikoliv malá forma hrdla. Tato forma je umístěna v tak zvaném obrazení tedy mechanismu, který se otáčí o 180° a přenáší lahev z přední do konečné formy. Po tomto přenesení a otočení je láhev sevřena v konečné formě a pomocí tlakového vzduchu vyfouknuta do konečného tvaru. Poté je láhev přenesena takzvaným přesouvačem do chladicího místa, kde se krátce dochladí. Z tohoto místa je takzvanou odstávkou přesunuta na dopravník, který láhev odveze do dochlazovací pece. Která zajistí pomalé ochlazení láhve. Výše zmiňovaná odstávka je opět jeden z velmi rychlých a extrémně přesně synchronizovaných mechanismů, protože se lahvemi musí trefit na rychle jedoucí dopravník do mezery mezi lahvemi z jiných stanic!

2.3 Shrnutí popisu sklářského stroje

Ve stroje je tedy velké množství mechanismů, které mají za úkol obsluhovat hlavní pracovní mechanismy, což jsou formy. Čas strávený ve formě je užitečný čas a většinou nejde zkrátit, v každé formě musí vždy láhev chvíli čekat a mírně se zchladit, tak aby se po postavení před odstávkou nezhroutila, ale zachovala si svůj tvar. Naopak všechny obslužné

pohyby například obraceče či odnímače jsou časy neužitečné a tudíž jejich zkrácením jsme schopni zvýšit výkon stroje.

Dříve se tyto stroje řídily pomocí válce s mnoha výčnělky, které spínali ventily různých mechanismů. Princip „flašinetu“.



Obrázek 2: Historický způsob řízení sklářského stroje

U současných velmi rychlých strojů jednotlivé mechanismy nečekají jeden na druhý, ale ve stejném okamžiku se pohybují ve společném prostoru, kde se míjí jen o několik málo milimetrů. Proto jakákoliv chyba v synchronizaci a časování těchto mechanismů může způsobit kolizi a tedy poškození stroje. Nároky na řídicí systémy tohoto stroje jsou tudíž extrémně vysoké, nepostačí jen „obyčejné“ i když velmi rychlé PLC, ale musí to být Motion Control systém. Tedy systém který umí v jednom ř.s. spojit v reálném čase jak PLC funkce tak různé druhy synchronizací, elektronických vaček a podobně. Takovým to systémem je bezesporu moderní řídicí Motion Control systém Simotion od firmy Siemens.

3 Požadavky na řídicí systém stanovené na začátku vývoje stroje ISS

Na základě zkušeností z jiných rozsáhlých strojů i zkušeností firmy Sklostroj Turnov CZ s.r.o. jsme si stanovili několik kritérií, které nové řízení musí splňovat, jednalo se o tyto kritéria:

Synchronizace všech prvků stroje

Modulární řešení

Moderní komunikační sběrnice

Bezpečnostní koncept

Jaké jsou důvody pro tyto kritéria a jak jsme je naplnili, si popíšeme v následujících kapitolách.

4 Popis řešení řídicího systému

4.1 Synchronizace všech prvků stroje

Z výše uvedeného opisu stroje je zřejmé, že stroj je složen z velkého množství různých mechanismů, které se často pohybují ve společném prostoru, nebo které na sebe technologicky i časově navazují. Pokud se mechanismy pohybují uvnitř společného prostoru, musí být velmi přesně řízeny. Za předpokladu, že stroj jede vždy stejnou rychlostí, dá se to řešit tak, že budu pouze velmi přesně zapínat pohony v běžné polohové vazbě, které mají pro daný pohyb definovány pevné parametry, jako jsou cílová poloha, rychlost, zrychlení, zpomalení, případně „jerk“ (derivace zrychlení). Pohon většinou tyto parametry udržuje

opakovatelně bez obtíží. Dá se tedy i dobře nastavit parametry pohybů a časy spínání, aby nedošlo ke kolizi různých mechanismů. U takto řízeného stroje nastává problém, pokud chci například zrychlit. V tom okamžiku potřebuji, aby pohon vykonal pohyb rychleji, jinak může dojít ke kolizi. Proto je u takto řízeného stroje při změně rychlosti třeba provést kompletní nové nastavení časování, což je velmi zdlouhavé a neefektivní.

Abychom minimalizovali tyto problémy při změnách rychlosti, zvolili jsme princip vačkové synchronizace. Uvnitř stroje máme definovanou virtuální osu, která určuje rychlost stroje, podobně jako válec s výčnělky v obrázku 2. Dle této virtuální osy se spínají například různé ventily. Většina reálných mechanismů má pak definovanou elektronickou vačku, která popisuje průběžně požadovanou polohu reálné osy v závislosti na poloze virtuální osy.



Obrázek 3: Příklad elektronické vačky

Řídicí systém Simotion umožňuje definovat vačku buď bodově, nebo také polynomičky. Většinou využíváme polynomičky, kdy jednotlivé úseky jsou popsány polynomem až 6 řádu + goniometrická funkce. Bodové vačky využíváme zřídka, jelikož vyžadují definici velkého množství bodů.

4.2 Modulární řešení

Z konstrukce stroje vyplývá, že se dá logicky rozdělit na část centrálních mechanismů a několika naprosto shodných částí konkrétních sekcí. Proto jsme se rozhodli, že zde bude jeden rozvaděč s řídicím systémem Simotion D445 a pohony pro společné části stroje. A dále několik identických sekčních rozvaděčů, které budou mít každý svůj vlastní řídicí systém Simotion D445. Posléze jsme se z prostorových důvodů rozhodli, že v jednom rozvaděči bude ř.s. a pohony pro dvě sekce. Tento modulární koncept přináší řadu výhod.

Jednou z výhod je to, že sekční rozvaděče se dají vyrábět sériově a nemusí se moc řešit, o kolika sekčních strojích se jedná. Což přináší minimální konstrukční náklady u konkrétních strojů.

Další výhodou je jednodušší programování. Byl vytvořen jeden program pro společné části stroje, který obsahuje odkazy na rozhraní každé sekce. V programu jsou vloženy rozhraní pro nejvyšší možný počet sekcí a dle konkrétního stroje se můžou některé sekce zakázat. Výhodou je i to, že v případě závažné poruchy, či údržby na jedné sekci se tato sekce

zablokuje a její případné chyby neovlivňují zbytek stroje. Dále byl vytvořen program pro jednu sekci stroje. Tento program plně řeší řízení celé sekce. Tento jeden program se nahrává do všech sekcí stroje. Rozhraní mezi společnou částí stroje a konkrétní sekcí tvoří sada několika logických informací, ale především master systém vysílá všem sekcím svoji vlastní virtuální osu, se kterou se synchronizují interní virtuální osy každé sekce. Využívá se funkce takzvaného „Distributed Gearing“. Což zajistí velmi vysokou přesnost všech částí celého stroje. Chyba mezi časovou značkou všech sekcí a centrální částí je menší jak 1µs.

4.3 Moderní komunikační sběrnice

Vzhledem k rozsahu aplikace, množství komunikovaných dat, požadavkům na modularitu a přesnost synchronizace padla volba jednoznačně na moderní komunikační sběrnici PROFINET, založenou na formátu Industrial Ethernet. V těchto aplikacích se používá IRT(Isochronous Real Time) PROFINET, který nabízí skutečně velký objem přenášených dat při velmi krátkých cyklech a extrémní přesnosti synchronizace.

Mezi zajímavé funkce související s PROFINETem, které byly v poslední době implementovány do programovacího prostředí Simotion SCOUT je nahrávání přes alternativní IP adresu. Proč je tato funkce zajímavá? Jak jsem zmínil výše máme pro sekční stroj jen jeden program, který má v konfiguraci jen jednu výchozí IP adresu, která ale na žádné sekci reálně není. Tudíž musíme mít možnost nahrát program do zařízení s jinou IP adresou.

Jak se do ř.s. Simotion dostane jiná IP adresa než nahraná? Řídicí systém Simotion má systémové funkce umožňující prozkoumat okolní připojená zařízení (strukturu sítě). Na základě této analýzy lze napsat program, který například u řídicího systému sekce sklářského IS stroje určí svoji polohu, tedy o kterou sekci se jedná. Na základě tohoto vyhodnocení zase další systémovou funkcí nastaví své PROFINET jméno a IP adresu.

4.4 Bezpečnostní koncept

Jedním z požadavků firmy Sklostroj Turnov CZ s.r.o. bylo integrovat bezpečnostní koncept do řídicího systému. Toto jsme realizovali nadřazením bezpečnostního PLC řady Simatic S7-300, respektive jednou z variant tohoto PLC ET200S-PN/DP CPU. Toto PLC má jednotky bezpečnostních vstupů a výstupů, které přijímají signály z tlačítek, závor atd. a svými výstupy ovládají bezpečnostní funkce integrované v měničích Sinamics S120. Na budoucích strojích plánujeme ovládání bezpečnostních funkcí přes protokol PROFISAFE na sběrnici PROFINET.

Dříve se v případě nouzového stopu například vypínaly stykače na vstupu rozvaděče, což nebylo příznivé, jelikož se vypnuly všechny měniče. Nyní je často požadavek na možnost zablokovat jen jeden měnič a ostatní nechat běžet. Například při potřebě vstupu do sledované zóny pohonů, kde se nachází jen určité pohony a je požadavek jen na jejich zablokování, aniž by byly odpojeny od sítě. Což měniče Sinamics S120 plně podporují svými integrovanými bezpečnostními funkcemi.



Obrázek 4: Celkový pohled na sklářský stroj ISS

5 Závěr

V současné době je stroj realizován a studené testy splnili očekávání firmy Sklostroj Turnov CZ s.r.o. Stroj by měl být v brzké době nasazen do reálného provozu. Již předvedení tohoto stroje na veletrhu Glastec 2011 v Dusseldorfu, vzbudilo velký zájem u odborné veřejnosti. Ačkoliv je stroj funkční a firma Sklostroj je spokojená, již se plánují další vylepšení a nové funkce, které by měly dále posílit pozici firmy Sklostroj Turnov CZ s.r.o. jako jedné z nejvýznamnějších firem udávajících trendy v IS strojích na výrobu obalového skla. Použití systému Simotion v podobně komplexních projektech dále potvrzuje mnohostrannost a výkonnost tohoto Motion Control systému.

Reference/References

- [1] Maximum Precision and Consistent Accuracy, GlassFocus 2010.

www.siemens.com/glass

CAPACITIVE SENSORS WITH PRECALCULABLE CAPACITANCE

Stanislav ĎAĎO

Katedra měření, Fakulta elektrotechnická, České vysoké učení technické v Praze

Technická 2, 166 27 Praha 6

E-mail: dado@fel.cvut.cz

Abstract:

The advantages of capacitance techniques for measurement (contactless sensing, relatively easy and flexible construction) are well known and can be further enhanced by application of the methods of design based on possibility to calculate their capacitance solely from electrodes dimensions. Capacitive sensors with precalculable capacitance (CSWPC) represent absolute sensors with high stability and immunity against the spurious effects of objects located close to sensor. The analytical calculation of capacitance is limited only to capacitors having cylindrical geometry of electrodes. Nevertheless by conformal mapping the majority of capacitive sensors can be designed by CSWPC methods. In this article besides the basic theory an example of CSWPC for measurement angle of rotation is presented.

Keywords: Kelvin guard ring, capacitive sensors, analytical methods of capacitance calculation, Thompson -Lampard theorem

1 Basic theory of pre-calculable capacitor

The accuracy of capacitance calculations by analytical methods is much higher than one reached by the numerical methods based on the finite element approach (see example of fringe field of sensor with circular electrodes and guard ring calculated by FEM on Figure 1 d). The principle of Kelvin's guard ring capacitor (Figure 1 a) stimulated many works on generalization of this principle.

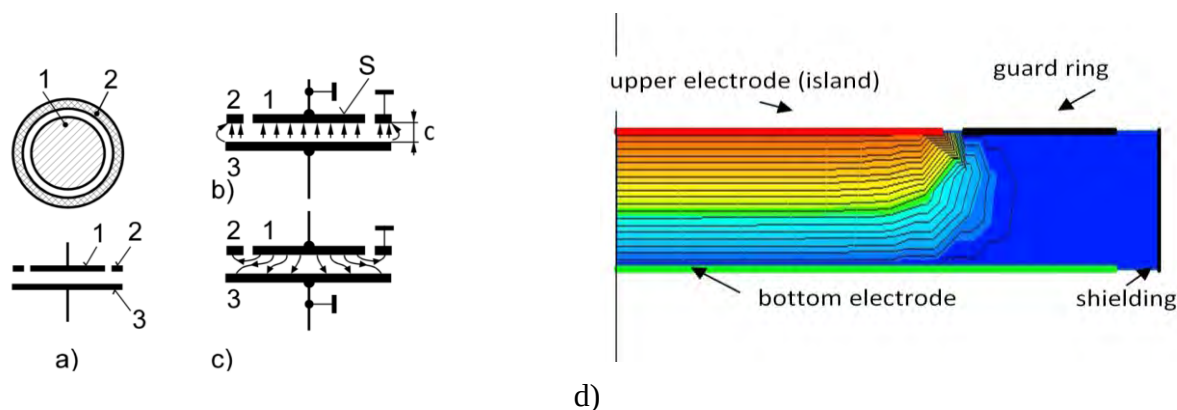


Figure 1. a) The principle of Kelvin guard ring capacitor and demonstration of independence of capacitance on field homogeneity. Capacitance in cases b) and c) is equal. d) Fringe field numerical model (FEM).

The capacitance between two electrodes i, j according Maxwell definition is given by equation

$$C_{ji} = \frac{Q_{ji}}{V_j - V_i} \quad (1)$$

where Q_{ji} is charge on electrode j induced by potential difference $V_j - V_i$ between electrodes. This means that for all other electrodes only their *presence and not their potential* contributes to the capacitance C_{ji} . Analytical calculation of capacitance consists in solving of potential V distribution in volume (space) in question (capacitor) following the surface charge density ρ calculations.

Potential V is related to charge density by *Poisson's equation* (1).

$$\nabla^2 V = -\frac{\rho}{\varepsilon_0} \quad (2)$$

In the charge-free region of space and for cylindrical coordinates (r, φ, z) Poisson equation transforms to Laplace equation, (3) (4).

$$\nabla^2 V = 0 \quad (3)$$

$$\frac{\delta^2 V}{\delta r^2} + \frac{1}{r} \frac{\delta V}{\delta r} + \frac{1}{r^2} \frac{\delta^2 V}{\delta \theta^2} + \frac{\delta^2 V}{\delta z^2} = 0 \quad (4)$$

For the case, when observed space and boundary conditions are rotation symmetrical, the potential V does not depend on angle θ , leading to reduction of *Laplace equation* into form (5).

$$\frac{\delta^2 V}{\delta r^2} + \frac{1}{r} \frac{\delta V}{\delta r} + \frac{\delta^2 V}{\delta z^2} = 0 \quad (5)$$

By means of the method of variables separation the potential V can be written as a sum of several products of functions of r and z .

In general the solution has a form of sum of Bessels and trigonometric functions. But complete solution is possible only in a few cases – when section shape of observed volume is simple and *volume is surrounded by electrode boundaries with prescribed potential overlay*.

For instance, where one of three dimensions can be considered as infinite compared with two others, the Laplace equation becomes linear and two-dimensional. Several different electrode geometries can be solved analytically by using the complex function theory and conformal mapping.

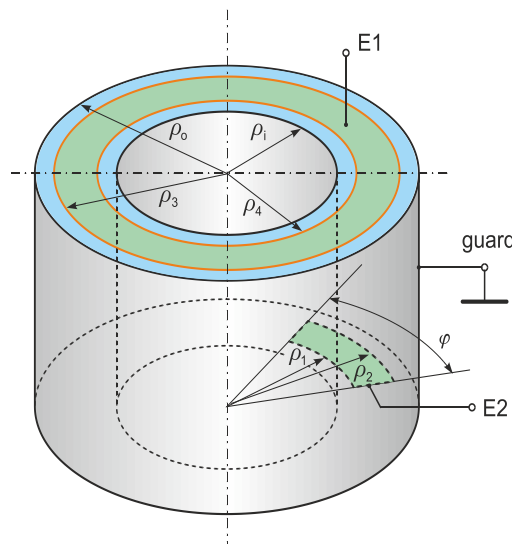


Figure 2. The archetype of torus shaped capacitive sensor

This approach is used for e.g. for analytical calculations of Thompson-Lampard primary standard of capacity [8].

One of the most important geometry of volume is the cylinder with rectangular section profile and rotational symmetric potential distribution on the boundaries (Figure 2.). In this case Laplace equation is not linear but analytical solution is possible. Segment or ring

electrodes overlapping or non-overlapping can be located on top or bottom surfaces (3 combinations) or as strips on inner and outer cylinder, or only on inner, or one electrode on inner and second on bottom surfaces (3 combinations). The analytical calculation of capacitance is feasible only if conducting surfaces lying beyond the electrodes are on ground potential, i.e. they form equivalent of guard rings. Five basic sensor archetypes can be derived from archetype of toroidal capacitor on Figure 2.

For example for configuration, when ring electrode E1 is located on top surface and the sector electrode E2 lies on bottom surface with all other part of the torus acting as guard material following formula for capacitance is valid ($\rho_j = \pi r_j / d$):

$$C_1 = \frac{\varepsilon_0 \varepsilon_r 2\pi d}{\pi^2} \sum_{n=1}^{\infty} (-1)^n \{ \rho_3 W'_n(\rho_3, \rho_i) \times [\rho_1 W'_n(\rho_1, \rho_0) - \rho_2 W'_n(\rho_2, \rho_0)] - \rho_4 W'_n(\rho_4, \rho_0) \times [\rho_1 W'_n(\rho_1, \rho_i) - \rho_2 W'_n(\rho_2, \rho_i)] \} [W'_n(\rho_0, \rho_i)]^{-1} + \frac{1}{4} (\rho_2^2 - \rho_1^2) \quad (6)$$

The following short notations were used here:

$$W_n(x, y) = I_0(nx)K_0(ny) - K_0(nx)I_0(ny) \quad (7)$$

and

$$W'_n(x, y) = \frac{1}{n} \frac{\partial}{\partial x} [W_n(x, y)] = I_1(nx)K_0(ny) + K_1(nx)I_0(ny) \quad (8)$$

The modified Bessel functions of the first and second kind and of zero and first order are denoted as $I_0(u), K_0(u), I_1(u), K_1(u)$. The torus geometry of archetype converts to cylindrical when internal diameter $\rho_i \rightarrow 0$.

Using conformal mapping the cylindrical shape archetypes of sensors can be converted to planar equivalents.

For conformal mapping of cylinder in w-plane (Figure 3.) to planar shape of capacitor (Figure 4.) in z-plane the transformation is as follows

$$z = \frac{1}{j} \ln \frac{1-w}{1+w} \quad (9)$$

where

$$w = \rho e^{j\varphi} \text{ and } z = x + jy \quad (10)$$

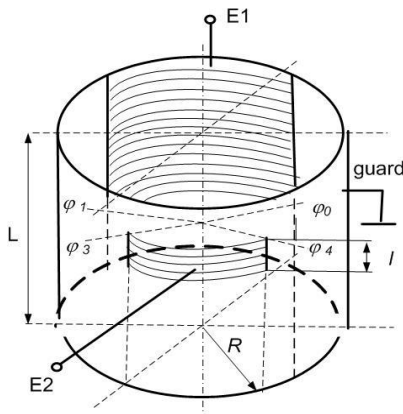


Figure 3. Cylindrical sensor

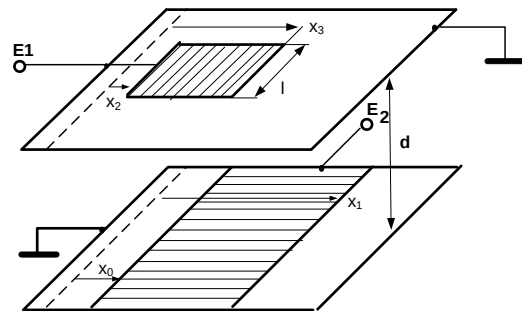


Figure 4. Conformal mapping of sensor from Figure 3.

The capacitance of cylindrical sensor can be derived using formulae for archetype torus shaped capacitor. For parameters shown in Figure 3. capacitance is given by the Equation (11)

$$C = \frac{\varepsilon_0 \varepsilon_r}{\pi} l \ln \left\{ \frac{\sin((\phi_3 - \phi_0)/2) \sin((\phi_2 - \phi_1)/2)}{\sin((\phi_2 - \phi_0)/2) \sin((\phi_3 - \phi_1)/2)} \right\} \quad (11)$$

It is remarkable that of all dimensions of the sensor exclusively the length l of the shortest electrode, measured along the cylinder determines the capacitance and neither the length L of the longest electrode nor the radius R of the tube.

The capacitance of strip sensor (Figure 4.) derived from cylinder by mapping is then given by formally similar formula

$$C = \frac{\varepsilon_0 \varepsilon_r}{\pi} l \ln \left\{ \frac{\cosh(\pi(x_3 - x_0)/2d) \cosh(\pi(x_2 - x_1)/2d)}{\cosh(\pi(x_2 - x_0)/2d) \cosh(\pi(x_3 - x_1)/2d)} \right\} \quad (12)$$

2 Calculation of cross capacitances

Typical example is calculation of circular cross-capacitances for the capacitor derived from the basic shape of Kelvin guard ring capacitor (Figure 5).

The formulas for two possible cross-capacitances C_{C1} and C_{C2} (Figure 5 and Figure 6) were found analytically [4], and expressed by means of the first order modified Bessel functions I_1 and K_1 , (13) (14).

$$C_{C1} = \frac{\varepsilon_r \varepsilon_0 (4\pi R_1 R_2)}{d} \sum_{n=1}^{\infty} (-1)^{n+1} I_1 \left(\frac{n\pi R_1}{d} \right) K_1 \left(\frac{n\pi R_2}{d} \right) \quad (13)$$

$$C_{C2} = C_{C1} + \frac{\varepsilon_r \varepsilon_0 \pi (R_2^2 - R_1^2)}{d} \quad (14)$$

It is important to notice that for $R_1 = R_2 = R$ equation (13) determines the capacitance of *circular version of Thompson-Lampard* capacitor with pre-calculable capacitance.

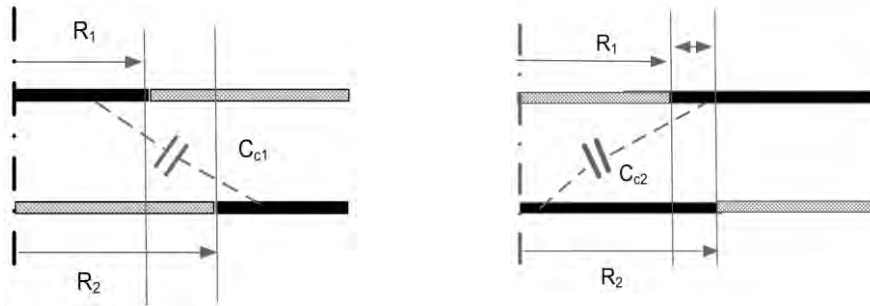


Figure 5. Cross-capacitances of circular electrodes

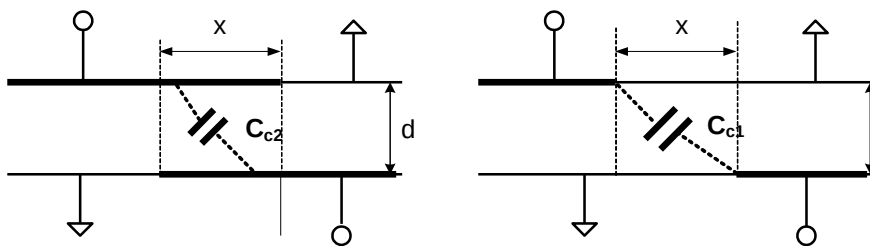


Figure 6. Cross-capacitances of rectangular electrodes

Interesting and for practice very useful fact is validity of equations for C_{C1} and C_{C2} also for the cases when one electrode has the shape of ring sector having angle Φ . Then factor π in equations (13) and (14) should be replaced by $\Phi/2$.

The basic equation valid for circular capacitor can be modified for calculation of capacitance of capacitor having rectangular electrodes. Expanding the average radius R to

infinity, leaving the length of arc $l = \Phi.R$ finite and replacing overlapping portion of radii R_2-R_1 by x , the capacitor with circular plates is transformed to rectangular type $l.x$ (Figure 5 a). Reflecting these modifications mathematically yields following relations for cross-capacitances (15) and (16).

$$C_{C1} = \varepsilon_r \varepsilon_0 \frac{l}{\pi} \left\{ \ln \left[2 \cosh \left(\frac{\pi x}{2d} \right) \right] - \frac{\pi x}{2d} \right\} \quad (15)$$

$$C_{C2} = \varepsilon_r \varepsilon_0 \frac{l}{\pi} \left\{ \ln \left[2 \cosh \left(\frac{\pi x}{2d} \right) \right] + \frac{\pi x}{2d} \right\} \quad (16)$$

And again for $x = 0$ it changes into the relation for the linear Thompson-Lampard primary standard capacitor whose capacitance is determined solely by the dimension l [8].

$$C_{C1} = C_{C2} = C_C = \varepsilon_r \varepsilon_0 \frac{l}{\pi} \ln 2 \quad (17)$$

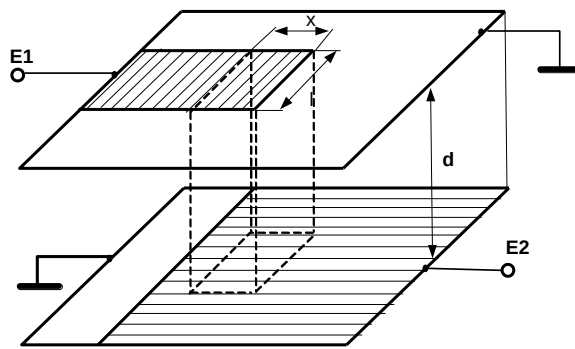


Figure 7. Asymptotic expansion of circular sector and ring capacitor

3 Heerens's rules for design of capacitor with pre-calculable capacitance

The founder of the theory of the capacitive sensors with pre-calculable capacitance, W.Ch. Heerens, worked systematically on the development of rules and principles for design and construction of these kinds of sensors, [3]. The most important rules come out from relations for cross-capacitances, (15) and (16).

By simple analysis of relation for C_{C2} can be deduced, that this capacitance depends nearly linearly on the area of electrodes overlapping $l.x$ when argument in $\cosh(\cdot)$ reaches high value, i.e. when

$$\frac{\pi x}{2d} \gg 1 \quad (18)$$

If this condition is fulfilled then

$$2 \cosh \frac{\pi x}{2d} \cong e^{\frac{\pi x}{2d}} \quad (19)$$

and

$$\ln e^{\frac{\pi x}{2d}} = \frac{\pi x}{2d} \quad (20)$$

Capacitance C_{C2} then equals to

$$C_{C2} = \frac{\varepsilon_0 \varepsilon_r}{\pi} l \cdot 2 \frac{\pi x}{2d} = \frac{\varepsilon_0 \varepsilon_r}{d} l x = C_{C2lin} \quad (21)$$

The capacitance, in accordance with relation for capacitance of planar electrodes placed in a homogenous field, is linearly proportional to the area $l.x$ of electrodes overlapping.

The relative deviation from linearity, δ , can be defined as

$$\delta = \frac{C_{C2} - C_{C2lin}}{C_{C2lin}} = \frac{d}{x\pi} \ln \left[2 \cosh \left(\frac{\pi x}{2d} \right) \right] - \frac{1}{2} \quad (22)$$

The graph describing dependence of δ [%] on x/d is at the figure a). It documents that linearity error is less than 1 ppm for $x/d > 5$ as it is given by Heerens rules. Fulfilling the condition (18) guarantees the linear dependence of capacitance on the change of the overlapping area caused by movement of electrodes, x , in horizontal plane.

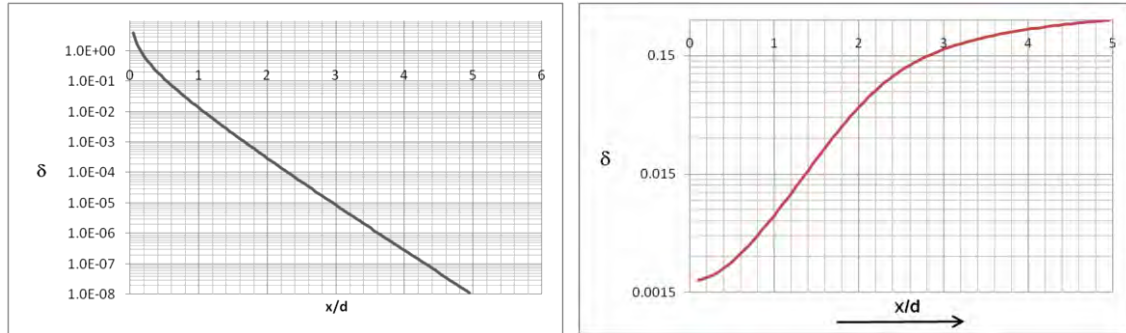


Figure 8. Linearity error, a) single sensor, b) differential sensor, $x_0 = 2d$

Similarly the capacity of differential sensor in which stator electrodes overlap the moving rotor electrode by x_0 can be found.

As expected differential configuration doubles sensor sensitivity and the ideal linear dependence of capacity $C(x)$ describes formula (23).

$$C_{lin} = \epsilon_r \epsilon_0 l \frac{2x}{d} \quad (23)$$

In case when condition (18) is not fulfilled, but the concept of capacitors with pre-calculable capacitance is preserved, it is still possible to calculate capacitance from active electrode dimensions.

From the air gap distance d , given usually by mechanical or electrical restrictions, minimum value of electrodes overlapping area $l \cdot x_{min}$ and maximum allowable linearity error for the width e of planar electrode should be chosen according criterion (18).

The next essential design parameter is the geometry of guard electrodes – usually grounded. Their main role is to eliminate the effects of fringe field on the capacitance and to avoid any influence of electrically active bodies in the vicinity of active electrodes. The relation (16) determining cross-capacitance can provide this function.

In order to minimize cross-capacitance C_{C1} the first and the second term of (15) should be equal. In practice it means that the distance x between the edge of active electrode and border of neighbouring conductive body should be chosen again according the condition (18). The acceptable influence of conducting bodies in the vicinity of capacitor can be found from (22) following the concept of capacitors with pre-calculable capacitance the volume, separating (guarding) capacitor from conductive surrounding plays role of guard ring. So condition (18) and equation (22) can be used for the design of the width of guard electrodes.

On the basis of the previous analysis W.C. Heerens defined following *rules of thumb* for design capacitors with pre-calculable capacitance:

- For the deviation from linearity (i.e from theoretical formula) $\delta \leq 1$ ppm and precision of calculation of capacitance from dimensions using theoretical formula the minimum dimension causing the change of overlapping area of planar electrodes must be $> 5d$, where d is the air gap – distance between electrodes. Based on this rule the dimension of active electrode can be chosen.

- b) The minimum guard area dimension g (e.g. width of guard ring) satisfying this inequality guarantees, that the relative deviation of accurate and calculated capacitance δ will be less than 1 ppm ($g \geq 5$)
- c) The difference between true and calculated capacitance will be less than 1 ppm, provided that width of insulating gap s between active electrodes and guarding electrodes fulfils the condition $s < d/5$.
- d) The boundaries of active electrodes can be found in the center of these gaps
- e) The gap influence decreases according relation $\exp(-\pi d/s)$
- f) Fringe field influence decreases proportionally to $\exp(-\pi y/d)$.

If two electrodes - which are situated on both opposite electrode carriers and which are forming together the properly guarded capacitor - have fringe field effect, accuracy of calculation better than 1 ppm is achieved if the sideways distance y between fringes is equal to $5d$.

4 Capacitive sensor of angle of rotation

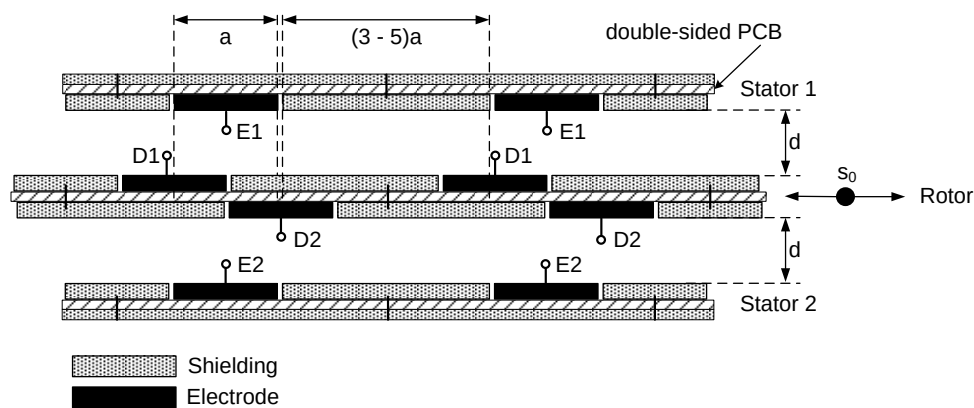


Figure 9. Front view (along the arrow) on vertically placed electrodes

As an example sensor for angle of rotation in horizontal plane is introduced. Sensor is based on change of the area of electrodes overlapping and could fulfill requirement of linearity of transfer characteristics for angle of rotation measurement. To avoid the basic drawback of this configuration, i.e. influence of air gap variation, and enhance sensitivity, the differential type of sensor with multiple electrodes was used (Figure 9.).



Figure 10. Capacitive sensor of angle of rotation

Sensor is composed from vertical three-electrodes capacitor. Two of them, the bottom and the upper one, are fixed. The third, the middle one, is mounted on a seismic mass and can move within a small horizontal range between the others, Figure 10. This type used as a special kind of differential arrangement of capacitive sensor which is independent on the vertical displacement of the middle electrode, [9].

The outline of all electrodes is the same – sector of a circle. This shape is suitable for manufacturing technology on PCB. The electrodes are designed as angular rays going from the

centre of the circle made on the each desk. The electrodes are connected by a bus at the inner periphery.

As illustrated at Figure 9 electrodes on stator are marked E1 and E2 and on rotor D1 and D2. Desks are manufactured using printed circuit board (PCB) technology. On the rotor are both sides used for electrodes with angle shift the same as the electrode width is. Electrodes on stator are the same and face each other. All the other conductive parts are used as shielding and can be grounded.

In the ideal zero position electrodes D1 and D2 of rotor overlap in the middle of stator electrodes E1 and E2. The move of rotor causes the change ΔS of overlapping area and consequently increases the capacitance of one couple and decreases the second one.

For the measurement of angle of rotation the configuration, in which electrodes E1 and E2 are joined, shielding parts of stator are grounded and shielding parts of rotor are left not connected.

Stator electrodes overlap half width of rotor electrode at the initial position, s_0 . Using the differential sensor the horizontal movement shifts the rotor electrode by $\pm \Delta s$ from the s_0 . As it can be expected the differential sensor shows substantially better sensitivity (Figure 8).

For the purpose of angle of rotational sensing the linearity of order 1 % appears to be satisfactory allowing the decrease of electrodes dimensions. From mechanical reason the air gap distance d cannot be substantially less than 1 mm. The full scale of angle was chosen to be ± 2.5 mrad. Based on the graph at Figure 8 for linearity error 1 % and full scale angle, corresponding to $\Delta s = 0.6$ mm (radius of rotor segment $r = 500$ mm), ratio $\Delta s/d = 0.6$. The angle width of electrode $a = 5$ mrad and guarding segment angle width $4a = 20$ mrad was chosen. Sensitivity is given by ENOB of measuring circuit and reached the value of 170 fF.

4.1 Ratiometric method of measurement

The critical issue of the proper function of rotational seismometer is as small as possible sensitivity to vertical movement (i.e. variation of d). The suppression of vertical movement influence can be achieved by placing rotor between two stator electrodes and using ratio-metric method of measurement. Stator electrodes E1 and E2 are joined and guarding electrodes of stators are grounded. Capacitance between D1 and combined E1 and E2 is sum of particular capacitances between D1–E1 and D1–E2 (i.e. C_{D1E1} and C_{D1E2}). At the initial position overlapped area is equal to S_0 and air gap is equal to d_0 . The influence of capacitance between electrodes E1, E2 and guarding area of stator is negligible due to the three terminal configuration of measuring circuit – it acts in parallel to the voltage source having very small output impedance.

The ratio-metric method of measurement eliminates an influence of air gap d on the output capacitance. This is proportional to the relative change of electrode overlapping area caused by rotation in the horizontal plane.

$$\frac{C_{D1E1E2} - C_{D2E1E2}}{C_{D1E1E2} + C_{D2E1E2}} = \frac{2 \cdot \epsilon_r \epsilon_0 \cdot V^{-1} \cdot (S_0 + \Delta S - S_0 + \Delta S) \cdot d_0^{-1}}{2 \cdot \epsilon_0 \epsilon_r \cdot V^{-1} \cdot (S_0 + \Delta S + S_0 - \Delta S) \cdot d_0^{-1}} = \frac{\Delta S}{S_0} \quad (24)$$

Changes of vertical vibrations, x , are efficiently reduced by the mechanical construction of sensor. The value x/d_0 is then close to zero and factor V approaches value 1. Consequently the vertical component of vibration has minimal effect on difference of capacities $C_{D1E1E2} - C_{D2E1E2}$.

4.2 Parallel connection of n electrodes

The main advantage of connection of n parallel electrodes is n -times enhancement of useful signal (proportional to ΔS) and increase of SNR at measuring circuit. It leads to

increase of resolution. Signal proportional to the difference of the capacities can be realized by the differential configuration of the measuring circuit that offers twice higher resolution without requirement to measure C_{D1E1E2} and C_{D2E1E2} separately.

$$\frac{C_{D1E1E2} - C_{D2E1E2}}{C_{D1E1E2} + C_{D2E1E2}} = \frac{2 \cdot \varepsilon_r \varepsilon_0 \cdot V^{-1} \cdot n \cdot (S_0 + \Delta S - S_0 + \Delta S) \cdot d_0^{-1}}{2 \cdot \varepsilon_0 \varepsilon_r \cdot V^{-1} \cdot n \cdot (S_0 + \Delta S + S_0 - \Delta S) \cdot d_0^{-1}} = \frac{n \Delta S}{n S_0} \quad (25)$$

Formally equation (24) and (25) give the same results. But the practical advantage on multiple electrode structure is the enhancement of SNR which might play decisive role for measurement of extremely small changes of capacitance. Using n groups of parallel connected electrodes n -times increases useful signal.

5 Measuring Circuit- C/D Convertor

For the proper function of SPC the elimination of parasitic capacitances of leads to electrodes. The simplest way is to use circuit based on Figure 11.

Influence of capacitances of leads is eliminated as they are connected in parallel to ideal voltage source or to virtual ground of operational amplifier.

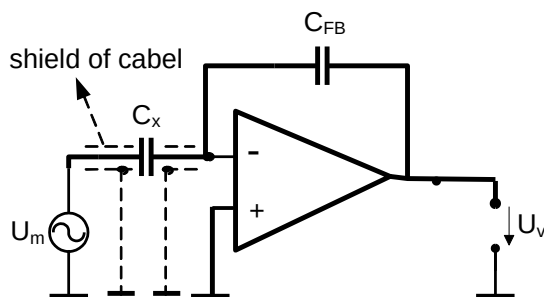


Figure 11. Elimination of parasitic capacitances of leads.

Analog Devices offers sophisticated Capacitance-to-Digital Converter with up to 24-bit resolution, with input circuit connected as shown in Figure 12. [1].

Usage of ratio-metric method dictates to measure both capacitances ideally synchronously. In order to measure two capacitances concurrently using just one CDC one-chip micro-computer controls directly values in registers of AD7746 and provide nearly *simultaneous* measurement of two single-ended capacitances connected alternatively to CIN1(+) and CIN2(+).

According to the AD7746 technical specification it permits independent single-ended measurement of the first or the second channel and also differential measurement. In configuration according Figure 12 b) switching between *channels* is performed only via change of register content. This property is also important for reaching the maximal possible sampling rate in case of time varying capacitance measurement.

6 Conclusions

The typical configuration of SWPC is definition of potential overlay of all electrodes. Designing capacitive sensors according the rules of pre-calculable capacitance enables construction of absolute sensor with capacity determined solely dimensions of electrodes. In practice it brings welcome and important property of avoiding the influence of all objects in the vicinity of sensor. Yet another advantageous property is possibility to calculate capacity from the formulae for “ideal” capacitor with homogenous field with accuracy predicted from dimensions of “guarding elements”(Heerens rules). This fact allows determination of transfer curve between measured quantity and capacitance and eventually to calculate e.g. the deviation from ideal function.

The certain disadvantage of SWPC are larger dimensions (guarding area!) in case when function of sensor requires preservation of ideal theoretical formula for capacitance calculation.

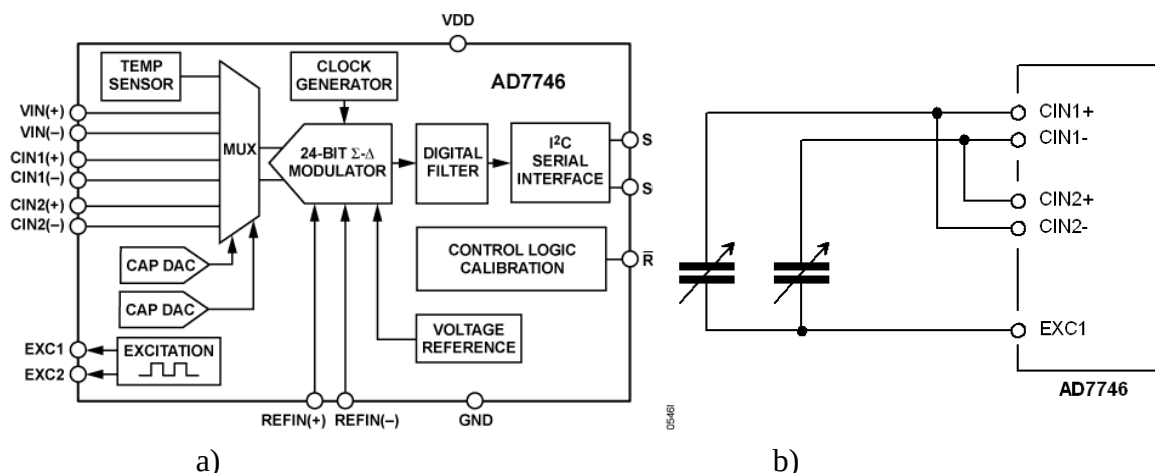


Figure 12 a) Capacitance –to Digital Converter AD 7746 by Analog Devices b) Connection of capacitors to CDC

References

- [1] AD7745/AD7746 (2005), 24-Bit Capacitance-to-Digital Converter with Temperature Sensor, Datasheet, Analog Devices, Rev.0,
Online: http://www.analog.com/static/imported-files/data_sheets/AD7745_7746.pdf (as of May, 2010)
- [2] Heerens W.C. (1986), Application of capacitance techniques in sensor design, J.Phys.E:Sci.Instrum., Vol. 19, pp. 897-906
- [3] Heerens W.C. (1982-1983), Basic principles in designing highly reliable multi-terminal capacitor sensors and performance of some laboratory test models, Sensors and Actuators A, Vol. 3, pp. 137-48
- [4] Heerens W.C. (1982), Multi-terminal capacitor sensors, J.Phys.E:Sci.Instrum., Vol. 15, No. 1, pp. 137-41
- [5] Heerens W.C. (1976), The solution of Laplace's equation in cylindrical and toroidal configurations with rectangular sectional shapes and rotation – symmetrical boundary conditions, Journal of Applied Physics, Vol.47, No. 8, pp. 3740-44
- [6] Strunc J., Buben J., Dado S. (2005): Rotation Component of Seismic Signal; Poster, EGU 2nd General Assembly, Vienna, Austria, SM1-1TU2P-0037
- [7] Dado S., Kreidl M.: Senzory a měřicí obvody, Vydavatelství ČVUT, Praha 1996
- [8] Bohacek J. (1994), Metrologie elektrických veličin, Monography, CVUT, Praha. ISBN 80-01-01152-6 (in Czech)
- [9] Strunc J.: Rotational Seismometer on the Capacitive Principle. Doctoral Thesis, Faculty of Electrical engineering, CTU, Prague. July 2010. Tutor of Thesis: Dado S.

PRODEJ A OCHRANA DUŠEVNÍHO VLASTNICTVÍ

Tomáš Halva

Jihomoravské inovační centrum
U Vodárny 2, 616 00 Brno
E-mail: halva@jic.cz

Abstrakt:

Příspěvek popisuje možnosti a procesy používané při ochraně duševního vlastnictví vzniklého při výzkumné práci vědeckých univerzitních pracovníků. Současný trend jednoznačně ukazuje na nutnost rozšíření spolupráce mezi firemní sférou a výzkumnými organizacemi. Jedna z možných forem spolupráce je tzv. transfer technologií, tedy prodej výsledků vědecké práce do firem.

Klíčová slova/Keywords: transfer technologií, komercializace, prodej, licence, duševní vlastnictví

1 Ochrana duševního vlastnictví

Aby bylo vůbec možné duševní vlastnictví vzniklé na univerzitě efektivně nabízet do firmy a úspěšně jej zkomercializovat, je bezpodmínečně nutné je předem ochránit. V České republice se tematikou ochrany duševního vlastnictví zabývá Úřad průmyslového vlastnictví (ÚPV), který zabezpečuje veškerou agendu spojenou s administrací dokumentů potřebných k ochraně.

Vhodným partnerem při procesu ochrany duševního vlastnictví a pro komunikaci s ÚPV je zkušený patentový zástupce, který zpravidla provádí tzv. patentovou rešerši, jejíž cílem je odhalit, zda technologie je opravdu unikátní. Další důležitou rolí patentového zástupce je připravit veškerou dokumentaci potřebnou pro podání přihlášky na Úřad průmyslového vlastnictví.

2 Formy ochrany duševního vlastnictví

2.1 Vynálezy / patenty

Předmětem patentové ochrany jsou vynálezy. Jedná se o ochranu myšlenek, které v praxi umožňují vyřešení konkrétního problému v oblasti techniky, včetně zemědělství, služeb, obchodu. Po přijetí tzv. patentové přihlášky zkoumá ÚPV, zda jsou splněny tzv. podmínky patentovatelnosti. Patří k nim tyto:

- novost
- vynálezecká činnost
- průmyslová využitelnost

Za vynález nelze považovat:

- objevy
- vědecké teorie a matematické metody, estetické výtvořky
- plány, pravidla a způsoby vykonávání duševní činnosti
- hraní her
- počítačové programy
- způsob chirurgického léčení, terapeutického ošetřování lidského a zvířecího těla

- vynálezy, jejichž využití by odporovalo veřejnému pořádku nebo dobrým mravům
- odrůdy rostlin a plemena zvířat a biologické způsoby pěstování rostlin či chovu zvířat

Pro vynálezy, které byly objeveny v rámci činnosti zaměstnance v rámci jeho pracovního poměru, platí, že vlastnické právo na patent přechází na zaměstnavatele. Zaměstnanec je povinen zaměstnavateli nový vynález ohlásit.

2.2 Užité vzory

Užitný vzor je udělen pro technická řešení, která jsou nová, a jejich průmyslová využitelnost přesahuje rámec odborné dovednosti.

Právní rámec ochrany užitných vzorů je srovnatelný s vynálezy, užitné vzory jsou však administrativně jednodušší. Na druhou stranu mají užitné vzory kratší dobu ochrany.

2.3 Průmyslové vzory

Chrání to, co dělá výrobek atraktivní, tedy vzhled výrobku nebo jeho části. Jedná se o ochranu designu výrobku či jeho prostorové/grafické uspořádání.

2.4 Ochrané známky

Jedná se o ochranu označení, jež je schopné grafického znázornění. Důležité je zde jasné spojení s konkrétním výrobkem nebo službou.

Možné druhy ochranných známek:

- Slovní
- Obrazové
- Kombinované
- Trojrozměrné
- Zvukové
- Blokační
- Certifikační
- Kolektivní
- Výrobní
- Obchodní
- Zásobní

3 Prodej duševního vlastnictví

V okamžiku, kdy je duševní vlastnictví řádně ochráněno, je možné spustit proces komercializace, tedy realizovat obchodní aktivity s cílem nabídnout technologii potenciálním zákazníkům, připravit licenční smlouvy a uzavřít obchod.

Transfer technologií může být chápán dvěma směry. Prvním směrem je komercializace již existujícího duševního vlastnictví, tedy směr popsany výše. Druhým směrem však může být i opačný postup, tedy zprostředkování tzv. smluvního výzkumu. V tomto případě firma zadá výzkumníkům požadavek na to, jaké vlastnosti by měla výsledná technologie mít. Výzkumník je tedy postaven před úkol vymyslet a dodat tzv. technologii na míru, dle požadavků klienta, a to v dohodnutém termínu, kvalitě a formě.

4 Závěr

Transfer technologií může být – pokud je vhodně uchopen – obrovským přínosem pro obě strany, tedy jak pro výzkumníky, tak pro firemní sektor. Firmám může ušetřit mnoho času a práce s vývojem, výsledky bývají často lepší než vývoj uskutečněný vlastními silami a často firmám chybí i dostatečné odborné znalosti v daném oboru. Univerzitám naopak tato spolupráce může zajistit částečný zdroj jejich financování a případně nasměrovat jejich výzkum směrem využitelným v praxi.

MODELOVÁNÍ A PREDIKTIVNÍ ŘÍZENÍ SPALOVACÍHO MOTORU

Daniel PACHNER, Jaroslav PEKAŘ

Honeywell Laboratory, Automation & Control Solutions

V Parku 2326/18, 148 00 Praha 4

E-mail: Daniel.Pachner@honeywell.com, Jaroslav.Pekar@honeywell.com

Abstrakt:

Společnost Honeywell vyvíjí a úspěšně testuje systém prediktivního řízení spalovacích motorů založeném na systematickém návrhu na základě modelu: H-ACT (Honeywell Advanced Control Technology). Jedná se o aplikaci prediktivního řízení na relativně rychlý nelineární systém s krátkou vzorkovací periodou, která využívá poslední výsledky rychlých numerických optimalizačních metod. Hlavním cílem je zrychlení vývoje algoritmů řídicích jednotek motorů a tím urychlení testování a zavádění nových technologií konstrukce spalovacích motorů, které si obecně kladou za cíl snížení emisí a snížení spotřeby paliva.

Klíčová slova: přeplňované motory, emise, prediktivní řízení

1 Úvod

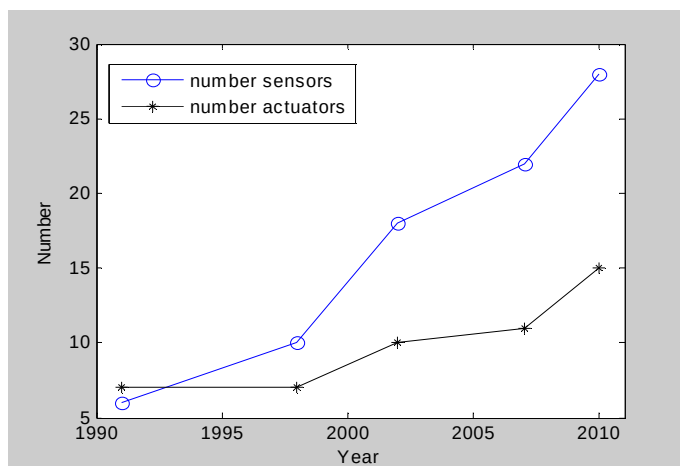
Prediktivní řízení na základě modelu MPC (Model Based Predictive Control) se úspěšně aplikuje v pokročilých průmyslových řídicích systémech především díky systematickosti návrhu a z toho vyplývající snazší udržitelnosti systému a jeho ladění, snadné aplikovatelnosti na mnohazměrové systémy s křížovými vazbami mezi veličinami, a především díky možnosti respektovat omezení na vstupní, výstupní i vnitřní veličiny soustavy. Díky výpočetní náročnosti se MPC nejprve aplikovalo především na řízení pomalých chemických procesů (vzorkovací perioda kolem 60s), později např. na řízení výroby energie v tepelných elektrárnách (systém Honeywell Advanced Energy Solutions, vzorkovací perioda pod 5s). Nyní vyvíjí a testuje společnost Honeywell prediktivní řízení spalovacího motoru s vzorkovací periodou pod 5ms, o kterém pojednává tento příspěvek.

2 Křížové interakce v řízení emisí motorů

Díky tomu, že prediktivní řízení přirozeně používá mnohazměrový popis řízené soustavy, je jeho aplikace atraktivní především na soustavy s mnoha vstupy a výstupy, kde se může návrh řízení založený na jednotlivých smyčkách jedné řízené a jedné akční veličiny ukázat jako velmi obtížný v důsledku křížového ovlivňování smyček. Typické příklady aplikací mnohazměrového prediktivního řízení jsou seřazeny podle počtu signálů regulátoru v tabulce:

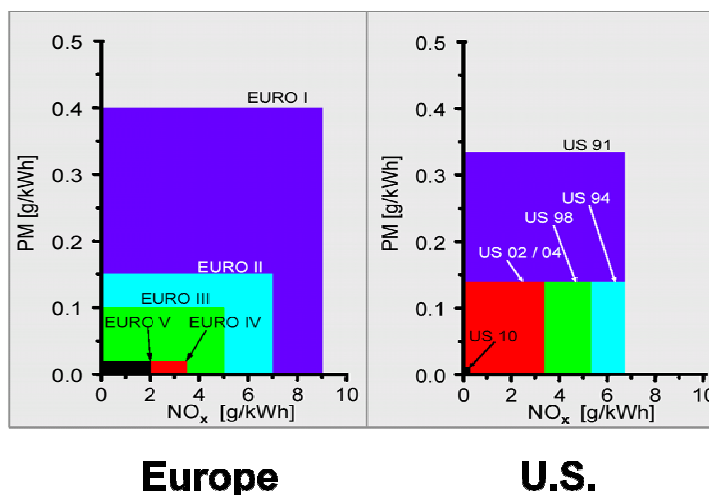
Tabulka 1: Přibližný počet řízených a akčních veličin v realizovaných řídicích systémech Honeywell

Aplikace	Akčních veličin	Měření
Řízení vzduchové cesty motoru	5	5
Výroba polyethylenu	15	40
Rafrinerie, destilační kolona	30	80
Papírenský stroj	600	8000



Obrázek 1 Typické počty akčních veličin a senzorů přeplňovaného motoru v posledních letech.

Díky zpřísnujícím se emisním limitům především na emise oxidů dusíku a pevných částic (sazí), viz Obrázek 2, byli výrobci motorů nuceni zavést v posledních letech na motorech další řízení a akční veličiny, jak ukazuje Obrázek 1. Bylo také nutno přistoupit k zavedení přeplňování turbodmychadlem ke snížení hmotnosti a zvýšení účinnosti motoru. Tím se úloha řízení zkomplikovala.



Obrázek 2: Emisní limity pro nákladní vozidla

Například pro snížení emisí vysoce toxického oxidu dusíku bylo nutno zavést recirkulaci ochlazených výfukových plynů (EGR, Exhaust Gas Recirculation) do sání tak, aby se špičková teplota ve válci snížila pod teplotu, kdy dochází k oxidaci vzdušného dusíku. Recirkulované spaliny představují vlastně inertní plyn, jehož tepelná kapacita absorbuje část tepla uvolněného hořením. Recirkulace je ovládána zpětným ventilem mezi výfukem a sáním. Hmotnostní tok recirkulovaných spalin je ovlivňován hlavní akční veličinou úhlem otevření EGR ventilu. Je ale zjevné, že je neméně ovlivněn tlakovou diferencí na ventilu, tj. rozdílem plnicího tlaku a tlaku za kompresorem a před turbínou turbodmychadla. Tyto tlaky se velmi dynamicky mění podle zatížení motoru a samotný plnicí tlak je mimoto řízenou veličinou podle požadavků na moment motoru. Například během zvýšení výkonu motoru (náhlém stlačení pedálu akcelérátoru) se nejprve podle konstrukce turbodmychadla uzavře klapka

obtoku turbíny (wastegate), nebo dojde k natočení lopatek turbíny tak, aby se zvýšil výkon turbíny a došlo k růstu plnicího tlaku. Tím dojde k brždění toku spalin a zvýšení tlaku na vstupu EGR ventilu. Následně dojde, bez patřičné korekce polohy EGR ventilu, k prudkému zvýšení hmotnostního toku přes EGR ventil, který bude pozvolna klesat tak, jak se bude zvyšovat plnicí tlak. Důsledkem je krátkodobé snížení koncentrace kyslíku ve válci až pod stochiometrické množství a nedokonalé spalování paliva, které se projeví viditelným kouřením motoru (ohlašujícím někdy předjížděcí máněvr kamiónu). K tomu nutně dochází, pokud je řídicí strategie recirkulace založena na jediné akční veličině úhlu otevření EGR ventilu, bez uvažování intrakce řízení turbodmychadla. Tento přístup je přitom dosud nejpoužívanější. EGR ventil se pak při popisovaném skokovém zvýšení požadavku výkonu uzavře až se zpožděním až několika sekund, až když se vliv změny tlakové difference projeví na měřených veličinách.

Z příkladu je vidět, že i tak jednoduchá věc, jako je recirkulace spalin, nemusí být z pohledu návrhu řízení triviální a nelze ji dobře realizovat PID regulátorem. Další zpřísnění limitů emisí oxidů dusíku nicméně situaci zesložilo, neboť vznikla nutnost zařadit do výfukového potrubí zařízení katalytické redukce oxidů dusíku SCR (Selective Catalytic Reduction). Toto zařízení redukuje oxidy dusíku aminiakem na vodu a dusík. Aminiak vzniká v SCR reaktoru tepelným rozkladem močoviny, jejíž 32,5% roztok je znám jako AdBlue. Tím vznikají další interakce. Rychlost chemických reakcí v SCR reaktoru například závisí na jeho teplotě, kterou lze ovlivnit řízením motoru.

3 Omezení řízených a řídicích veličin

Další výhodou pokročilých metod řízení MPC je schopnost respektovat omezení na řídicí i řízené veličiny. Mezi hlavní omezení v úloze řízení motoru patří: omezení na akční veličiny, případně rychlost jejich změny. Dále je vhodné respektovat omezení poměru paliva a spalovacího vzduchu na množství mírně vyšší než stochiometrické. Dále existují materiálová omezení, především teploty turbíny a rychlosti otáčení turbodmychadla. Je zbytečné zdůrazňovat, že nerespektování materiálových omezení má přímý vliv na životnost a spolehlivost motoru. Technologie MPC definuje výpočet optimálního akčního zásahu jako optimalizační problém za respektování těchto omezení (constrained optimization).

4 Modelování motorů

Modely pro řízení lze obecně rozdělit do několika kategorií. První jsou modely zařízení, v tomto případě motoru, které ještě neexistuje, tzv. proaktivní modely. V případě vývoje nového motoru se někdy vyvíjí takový model, který je typicky popsán soustavou nelineárních parciálních diferenciálních rovnic. Jejich řešení je velice náročné a simulace jedné minuty běhu motoru může trvat několik hodin. Vzhledem k tomu je takový model pro návrh řídicího systému nevhodný. Návrh řídicího systému motoru probíhá typicky pro již existující motor, používáme proto spíše tzv. reaktivní modely, které obsahují neurčité parametry odhadované na základě dat. Tím lze dosáhnout požadované přesnosti s jednoduchým modelem. Tyto modely se dále dělí na tzv. black box a grey box modely. Struktura black box modelu je definována náhodně, většinou tak, aby byla numerická identifikace parametrů co nejsnazší. Struktura grey box modelů je naopak definována na základě alespoň přibližných fyzikálních vztahů (termodynamika, zákony zachování) a pouze koeficienty rovnic jako např. pneumatické odpory, koeficienty přestupu tepla a tepelné kapacity jsou částečně neurčité a je třeba je identifikovat.

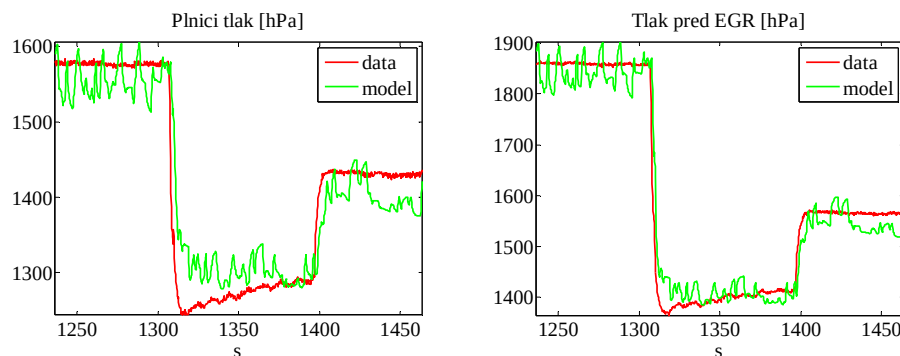
Model vhodný pro návrh MPC motoru patří do kategorie grey box „mean value“ modelů, což znamená, že modeluje střední hodnoty veličin za otáčku klikového hřídele. Detailnější modely motorů modelují např. průběh tlaku ve výfukovém potrubí během otáčky

motoru, ne jen jeho střední hodnotu při určitých otáčkách a zatížení. To je dáno tím, že během otáčky hřídele bude mít většina akčních členů v podstatě konstantní polohu a proto detailnější popis nemá pro řízení smysl.

5 Systematický přístup návrhu řídicího systému

Systém H-ACT definuje systematický přístup návrhu řídicího systému motoru krok za krokem a implementuje příslušné algoritmy. Uživatel komunikuje se systémem přes přehledné grafické rozhraní „H-ACT design suite“. Popíšeme nyní jednotlivé kroky návrhu:

- 1) Konfigurace modelu: Uživatel nakreslí konfiguraci motoru a definuje potřebné konstanty, například objem motoru, objem potrubí, kompresorové a turbínové mapy. Tato konfigurace se přeloží do soustavy obyčejných diferenciálních rovnic s neurčitými parametry. Uživatel definuje polohu akčních členů a senzorů.
- 2) Proběhne návrh experimentů, který je nutno na motoru provést v otevřené smyčce. Vzhledem k nelinearitě soustavy je potřeba proměřit reakce na polohy aktuátorů v určité mřížce v mnoha pracovních bodech motoru.
- 3) Sebraná data jsou automaticky použita na identifikaci neurčitých parametrů modelu metodami identifikace nelineárních systémů. Typická přesnost modelu je $\pm 10\%$ v ustáleném stavu. Ve změnách (trendech) je model přesnější. Věrnost predikcí modelu šestiválcového přeplňovaného dieselového motoru ukazuje Obrázek 3. Zde jsme pro ilustraci zvolili dva tlaky napříč EGR ventilem, o kterých byla řeč výše. Ve skutečnosti model modeluje kolem padesáti měřených veličin. Zašumění predikcí modelu je dáno šuměním měření vstupních veličin modelu.

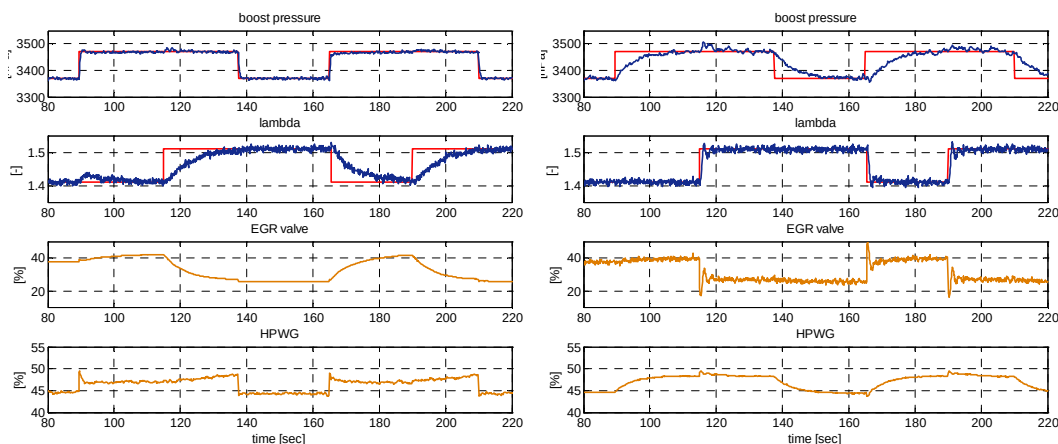


Obrázek 3 Ukázka věrnosti predikce reakce tlaků před a za EGR ventilem na skokové změny polohy EGR ventilu a ventilu obtoku turbíny (wastegate).

- 4) Uživatel definuje vstupy a výstupy regulátoru, dovolené rozsahy veličin v závislosti na režimu motoru a dovolené rychlosti změny akčních veličin. Dále definuje tabulky žádaných hodnot řízených veličin a požadavky na rychlost regulace a robustnost.
- 5) Je automaticky vygenerován odpovídající optimální mnohazarozměrový regulátor, který je testován proti nelineárnímu modelu motoru. Je možné vrátit se k předchozímu bodu a měnit specifikace, testovat různé řídicí strategie. Testují se odezvy na skokové změny referenčních a poruchových signálů.
- 6) V další fázi uživatel testuje standardní emisní jízdní cykly definované normami Euro a podobně.
- 7) Regulátor je možné okamžitě nahrát do RP (Rapid Prototyping) systému nebo řídicí jednotky motoru a testovat na zkušební stolici s motorem.

Vzhledem k tomu, že akční zásah MPC regulátoru je výsledkem minimalizace aditivního kritéria, kde jednotlivé sčítance odpovídají jednotlivým cílům řízení, většinou

jednotlivým veličinám, spočívá ladění regulátoru jednoduše v iterativním ladění váhových koeficientů, jak to ukazuje Obrázek 4. Ukazuje, jak lze jednoduše balancovat přesnost řízení dvou veličin dvěma akčními členy: EGR ventilem a wastegate vysokotlaké turbíny (dvojitě přepínání).



Obrázek 4 Ukázka změn vah složek kritéria. Vlevo je akcentováno přesné sledování plnicího tlaku (boost), vpravo přesné sledování poměru paliva a vzduchu ve válci (lambda).

6 Závěr

Společnost Honeywell se snaží o systematické řešení řízení spalovacích motorů využitím nástrojů fyzikálního modelování a rychlé numerické optimalizace. V praktických testech se ukazuje, že v současné době existují algoritmy, které umožňují řídit tzv. vzduchovou cestu spalovacího motoru v podstatě automaticky vygenerovaným MPC algoritmem a dosahovat výsledků minimálně srovnatelných s klasickými algoritmy, které jsou založeny na vyhledávacích tabulkách a PID smyčkách, které vyžadují velmi pracné a časově náročné iterativní ladění. Redukce času potřebného pro vývoj řídicího systému novým přístupem je minimálně 80%. V praktických testech se navíc často ukazuje možnost dosáhnout snížení emisí NOx a pevných částic.

Poděkování

Technologická agentura České republiky podpořila rozšíření výzkumu prediktivního řízení vzduchové cesty motoru o zahrnutí systému katalytické redukce NOx a filtru pevných částic jako výzkumný projekt programu Alfa TA01030170.

Reference/References

- [1] Stewart, G.E., Borrelli, F., Pekar, J., Germann, D., Pachner, D., and Kihns, D.: Toward a Systematic Design for Turbocharged Engine Control, in Automotive Predictive Control, L. del Re, Ed. New York: Springer Verlag, 2009.
- [2] Bemporad, A., Morrari, M.: The explicit linear quadratic regulator for constrained systems, in Automatica 38, 3-20, 2002.
- [3] Kolmanovsky, I.V., Stefanopoulou, A.G., Moraal, P.E., and Van Nieuwstadt, M.: Issues in Modeling and Control of Intake Flow in Variable Geometry Turbocharged Engines”, 18th IFIP Conference on System Modelling and Optimization, 1997.

PŘÍŘAZENÍ JORDANOVY FORMY ZPĚTNOU VAZBOU: METODA, APLIKACE, PODPŮRNÝ SOFTWARE

Miloš Schlegel

Katedra kybernetiky, Západočeská univerzita v Plzni

Univerzitní 8, 306 14 Plzeň

E-mail: schlegel@kky.zcu.cz

Abstrakt:

Cílem článku je poskytnout jednoduchý úvod do důležitého a obtížného problému modálního řízení lineárních systémů. Pomocí jednotného přístupu založeného na Sylvestrově maticové rovnici je předloženo řešení problému přiřazení Jordanovy formy matici dynamiky uzavřeného systému. Metoda obsahuje úplné i částečné přiřazení stavovou i výstupní zpětnou vazbou a to v parametrickém tvaru s minimálním počtem parametrů. Oproti obdobným známým metodám (např. metodě přiřazení pólů v Matlabu) nabízí všechna řešení. Tento fakt lze s výhodou využít při návrhu robustních regulátorů a regulátorů s omezenou strukturou, kdy volné parametry zpětné vazby mohou být využity ke splnění dalších návrhových požadavků.

Klíčová slova: Modální řízení, přiřazení pólů, Jordanova forma, robustní návrh, regulátory s omezenou strukturou.

1 Úvod

Úloha modálního řízení (přiřazení pólů) je považována za základní problém lineární teorie řízení [11]. V té či oné podobě se skutečně vyskytuje téměř ve všech algoritmech syntézy lineárních řídicích systémů. Pro její těsný vztah s fyzikálně názornými pojmy jako vlastní frekvence, časová konstanta, tlumení, rychlost odezvy, šířka frekvenčního pásma a podobně je velmi populární v praxi. Nejjednodušší variantou modálního řízení je oblíbená metoda geometrického místa kořenů (Evansova metoda). Standardní metodu reprezentuje přiřazení pólů stavovou zpětnou vazbou, složitější variantou je přiřazení pólů výstupní statickou zpětnou vazbou, která v sobě zahrnuje i případ dynamické zpětné vazby [9]. Obecně přísnější návrhový požadavek než přiřazení pólů je přiřazení Jordanovy formy. Nyní uvedeme formulaci této úlohy.

Uvažujeme lineární t-invariantní systém

$$\dot{x}(t) = Ax(t) + Bu(t) \quad (1)$$

$$y(t) = Cx(t) \quad (2)$$

kde $\dot{x}(t)$ značí $\dot{x}(t)$ pro spojitý systém a $x(t+1)$ pro diskrétní systém, $x(t) \in \mathbb{R}^n$ je stavový vektor, $u(t) \in \mathbb{R}^m$ je vstup, $y(t) \in \mathbb{R}^p$ je výstup a A, B, C jsou konstantní reálné matice příslušných rozměrů. Systém popsáný rovnicí (1) budeme označovat dvojicí (A, B) , podobně trojice (A, B, C) bude označovat systém (1), (2).

V dalším budeme rozlišovat dvě základní úlohy přiřazení Jordanovy formy matici dynamiky uzavřeného systému. V prvním případě jde o přiřazení stavovou zpětnou vazbou

$$u(t) = Fx(t), \quad F \in \mathbb{R}^{m \times n}, \quad (3)$$

v druhém, komplikovanějším případě půjde, o přiřazení výstupní statickou zpětnou vazbou

$$u(t) = Ky(t), \quad K \in \mathbb{R}^{m \times p}. \quad (4)$$

Dále rozlišíme úplné a částečné přiřazení. Pod úplným přiřazením budeme rozumět přiřazení Jordanovy formy $L \in \mathbb{R}^{n \times n}$ matici dynamiky uzavřeného systému A_c (tj. matici $A+BF$ resp. $A+BKC$). Pod částečným přiřazením budeme rozumět též přiřazení zvolené Jordanovy formy $L \in \mathbb{R}^{s \times s}$, $1 \leq s < n$, avšak pouze pro některý invariantní podprostor matice dynamiky uzavřeného systému. Připomeňme, že $\mathcal{X}$ je invariantní podprostor matice M , jestliže $M\mathcal{X} \subset \mathcal{X}$.

Důležitost přiřazení Jordanovy formy stavovou zpětnou vazbou vyplývá ze skutečnosti, že mezi trajektoriemi autonomních systémů

$$\dot{x}(t) = A_c x(t) \quad (5)$$

a

$$\dot{z}(t) = Lz(t), \quad (6)$$

kde A_c je matice dynamiky uzavřeného systému a L je přiřazovaná Jordanova forma, existuje vzájemně jednoznačné zobrazení

$$x(t) = Tz(t) \quad \text{resp.} \quad z(t) = T^{-1}x(t), \quad \forall t,$$

právě tehdy, jestliže $A_c = TLT^{-1}$, jak plyne z následující rovnosti

$$\dot{x}(t) = T\dot{z}(t) = T L z(t) = T L T^{-1} x(t) = A_c x(t)$$

platící pro všechny možné trajektorie $x(t)$. Tedy systémy (5) a (6) jsou z topologického hlediska ekvivalentní právě tehdy, jestliže jsou matice A_c a L podobné, neboli mají-li stejnou Jordanovu formu. Poněvadž podobné matice mají stejná vlastní čísla, přiřazení Jordanovy formy zajišťuje i přiřazení požadovaných vlastních čísel (pólů). Obráceně to však neplatí. Například matice

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad a \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

mají stejná vlastní čísla, ale různé Jordanovy formy. Podobně lze ukázat, že v případě násobných vlastních čísel je přiřazení Jordanovy formy přísnější požadavek, než pouhé přiřazení vlastních čísel. Poznamenejme, že tento přísnější požadavek má uplatnění v teorii řízení.

Jako ilustrační příklad, kdy nestačí pouhé přiřazení pólů uveďme diskrétní stavový regulátor konečného počtu kroků. Za tímto účelem uvažujeme diskrétní systém

$$x(t+1) = Ax(t) + Bu(t)$$

se stavovou zpětnou vazbou $u(t) = Fx(t)$. Uzavřený systém je tedy popsán diferenční rovnicí

$$x(t+1) = (A + BF)x(t).$$

Aby jeho odezva na nenulové počáteční podmínky $x(0)$ vymizela za konečný počet kroků, musí být matice $A + BF$ nilpotentní, neboli všechna její vlastní čísla musí být nulová. Až potud bychom vystačili s přiřazením vlastních čísel (pólů). Požadujeme-li však navíc, aby

odezva pro libovolný počáteční stav měla minimální možný počet kroků, je nutné požadovat, aby matice $A + BF$ byla podobná nilpotentní matici L , jejíž Jordanovy bloky mají řády rovné indexům řiditelnosti dvojice (A, B) . Poznamenejme, že o tom, zda matice L je přiřaditelná i v obecném případě rozhodují indexy řiditelnosti a tzv. stupně nekonstantních invariantních polynomů matice L . O obou těchto číselných charakteristikách bude pojednáno v oddílu 2.

Cílem tohoto článku je poskytnout jednoduchý, avšak pokud možno nezkreslený úvod do problému přiřazení Jordanovy formy stavovou a výstupní zpětnou vazbou. Výklad vychází z původních prací [5],[7], [8], [6]. Použitý přístup je založen na Sylvestrově maticové rovnici a metodě řešení soustav polynomiálních rovnic technikou Gröbnerovy báze. V tomto článku jsou však uvedeny pouze některé jednodušší výsledky z výše uvedených prací a to bez důkazu. Hlavní důraz je kladen na srozumitelnost a na nové možnosti, které příslušné metody a výpočtové procedury [8] poskytují v aplikacích. Z tohoto důvodu je zmíněna též volně přístupná knihovna (Jordan Form Assignment Package - JFAP) pro MAPLE 13 [3]. Přednosti této knihovny spočívají především v tom, že výpočtové procedury poskytují všechna řešení a to v parametrickém tvaru s minimálním možným počtem parametrů. V důsledku toho lze volné parametry využívat pro splnění dodatečných návrhových požadavků, například na robustnost ve stabilitě. Vzhledem k tomu, že všechny výpočty jsou prováděny symbolicky, lze přiřazovat i Jordanovy formy, které jsou specifikovány jedním nebo více volnými parametry. Takový způsob návrhu lze například využít pro snadno ručně přeladitelné regulátory zachovávající konstantní tlumení.

2 Základní pojmy

Pro přesnou formulaci řešení problémů modálního řízení budeme potřebovat některé speciálnější pojmy z teorie matic a lineárního řízení

2.1 Komplexní a reálná Jordanova forma

Nejdůležitější v následujícím textu budou pojmy Jordanova forma (tvar) matice a reálná Jordanova forma [2].

Je-li λ komplexní číslo a l přirozené číslo, nazývá se Jordanovým blokem $J_l(\lambda)$ čtvercová matice l -tého řádu

$$J_l(\lambda) \triangleq \begin{bmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}.$$

Říkáme, že matice J je v Jordanově tvaru, je-li blokově diagonální a její každý diagonální blok je Jordanův blok:

$$J \triangleq \text{diag}\{J_{l_1}(\lambda_1), J_{l_2}(\lambda_2), \dots, J_{l_r}(\lambda_r)\} \quad (7)$$

Čísla λ_i , která nemusí být navzájem různá, jsou vlastní čísla matice J . V případě, že některé vlastní číslo matice J je komplexní, je i Jordanova forma (7) komplexní matice. Tato skutečnost je na závadu při výpočtech omezených na reálný obor. Z těchto důvodů se zavádí reálná Jordanova forma matice, která se snadno získá z komplexního Jordanova tvaru, jak je patrné z následujícího příkladu.

Příklad 1. Uvažujeme komplexní Jordanův tvar

$$J = \text{diag}\{J_2(a + jb), J_2(a - jb), J_1(c)\}, \quad \text{kde } a, b, c \in \mathbb{R},$$

potom reálný Jordanův tvar matice J má následující podobu

$$L = \begin{bmatrix} a & b & 1 & 0 \\ -b & a & 0 & 1 \\ & & a & b \\ & & -b & a \\ & & & & c \end{bmatrix},$$

kde Jordanovy bloky stejného řádu příslušné komplexně sdruženým vlastním číslům vytvářejí jeden reálný Jordanův blok dvojnásobného řádu a reálné Jordanovy bloky zůstávají beze změn.

2.2 Stupně nekonstantních invariantních faktorů

Dále z komplexní Jordanovy formy snadno určíme stupně nekonstantních invariantních polynomů této matice. Nechť J je matice ve tvaru (7) řádu s , která obsahuje

r_1 Jordanových bloků příslušných k vlastnímu číslu λ_1 s rozměry $n_1^1 \geq n_2^1 \geq \dots \geq n_{r_1}^1$,

r_2 Jordanových bloků příslušných k vlastnímu číslu λ_2 s rozměry $n_1^2 \geq n_2^2 \geq \dots \geq n_{r_2}^2$,

...

r_l Jordanových bloků příslušných k vlastnímu číslu λ_l s rozměry $n_1^l \geq n_2^l \geq \dots \geq n_{r_l}^l$.

Definujeme

$$\nu_i = \sum_{j=1}^l n_i^j, \quad i = 1, 2, \dots, k,$$

kde $k = \max\{r_1, r_2, \dots, r_l\}$ a $n_i^j = 0$ pro $i > r_j$. Snadno lze ověřit, že platí

$$\nu_1 \geq \nu_2 \geq \dots \geq \nu_k$$

a

$$\nu_1 + \nu_2 + \dots + \nu_k = s$$

O libovolné matici L , která je podobná matici J budeme říkat, že její stupně nekonstantních invariantních polynomů jsou $\nu_1, \nu_2, \dots, \nu_k$ a tuto skutečnost budeme zapisovat: $\nu(L) = \{\nu_i\}_{i=1}^k$.

Příklad 2. Pro matici J z příkladu 1 platí $k = 2$, $\nu_1 = 3$, $\nu_2 = 2$ tj. $\nu(J) = \{3, 2\}$. Poznamenejme ještě, že je-li $\nu(L) = \{s\}$ (tj. je-li $k=1$), potom matici L nazýváme cyklickou. Matice s vesměs různými vlastními čísly je zřejmě vždy cyklická.

2.3 Indexy řiditelnosti

Pro rozhodnutí, zda je daná Jordanova forma přiřaditelná stavovou zpětnou vazbou či nikoliv, budeme potřebovat znát tzv. indexy řiditelnosti dvojice (A, B) . Předpokládejme, že dvojice (A, B) je řiditelná, tj.

$$\text{rank} [B, AB, \dots, A^{n-1}B] = n,$$

a dále

$$B = [b_1, b_2, \dots, b_m] \text{ a } \text{rank } B = m.$$

Utvořme posloupnost vektorů

$$b_1, \dots, b_m, Ab_1, \dots, Ab_m, \dots, A^{n-1}b_1, \dots, A^{n-1}b_m.$$

V této posloupnosti postupujeme zleva doprava a vyškrtneme každý vektor, který je lineárně závislý na předchozích vektorech této posloupnosti. Po ukončení tohoto procesu a po příslušném přerovnání a případném přechíslování sloupců matice B obdržíme posloupnost

$$b_1, Ab_1, \dots, A^{\mu_1-1}b_1, b_2, Ab_2, \dots, A^{\mu_2-1}b_2, \dots, b_m, Ab_m, \dots, A^{\mu_m-1}b_m,$$

kde

$$\mu_1 \geq \mu_2 \geq \dots \geq \mu_m > 0$$

a

$$\sum_{i=1}^m \mu_i = n$$

Čísla $\mu_1, \mu_2, \dots, \mu_m$ budeme nazývat indexy řiditelnosti dvojice (A, B) a tuto skutečnost budeme zapisovat následovně: $\mu(A, B) = \{\mu_i\}_{i=1}^m$.

3 Přiřazení Jordanovy formy stavovou zpětnou vazbou

Uvažujeme řiditelný systém (A, B) , $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, kde matice B má plnou sloupcovou hodnotu. Hledáme všechny stavové zpětnovazební matice $F \in \mathbb{R}^{m \times n}$ patřící do množiny

$$\mathcal{F}_s(A, B, L) \triangleq \left\{ F \in \mathbb{R}^{m \times n} : (A + BF) \sim \begin{bmatrix} L & * \\ 0 & * \end{bmatrix} \right\},$$

kde $L \in \mathbb{R}^{s \times s}$, $1 \leq s \leq n$ je daná matice v reálném Jordanově tvaru, $*$ označuje libovolnou reálnou matici vhodných rozměrů a symbol $\sim$ označuje relaci podobnosti matic. Je-li $s = n$, potom množina $\mathcal{F}_s(A, B, L)$ obsahuje zpětnovazební matice F , které splňují podmínku $A + BF \sim L$. V tomto případě budeme hovořit o úplném přiřazení Jordanovy formy, neboť matice dynamiky uzavřeného systému $A + BF$ je úplně změněna podle požadavku vyjádřeného zadanou maticí L . V opačném případě $s < n$ budeme hovořit o neúplném přiřazení, neboť požadovaná změna dynamiky nastala pouze na s -rozměrném invariantním podprostoru stavového prostoru. Zbylé módy (vlastní čísla) se ovšem v obecném případě změní též, nikoliv však podle našeho požadavku.

Na otázku, kdy je množina řízení $\mathcal{F}_s(A, B, L)$ neprázdná odpovídá následující věta.

Věta 1. [7], [4] Nechť $\mu(A, B) = \{\mu_i\}_{i=1}^m$ a $\nu(L) = \{\nu_i\}_{i=1}^k$, potom $\mathcal{F}_s(A, B, L)$ je neprázdná právě tehdy, jestliže $k \leq m$ a pro všechna $j = 1, 2, \dots, k$ platí nerovnost

$$n - s + \sum_{i=1}^j (\nu_i - \mu_i) \geq 0.$$

Věta 1 specifikuje v jakém smyslu musí být vnitřní struktura systému (A, B) , charakterizována indexy řiditelnosti $\mu(A, B) = \{\mu_i\}_{i=1}^m$, kompatibilní s vnitřní strukturou matice L , chrakterizované posloupností $\nu(L) = \{\nu_i\}_{i=1}^k$, aby L byla přiřaditelná stavovou zpětnou vazbou. Důležitý důsledek věty 1 je fakt, že libovolná cyklická matice L je vždy přiřaditelná řiditelnému systému (A, B) .

Nyní naznačíme podstatu metody přiřazení Jordanovy formy stavovou zpětnou vazbou založenou na Sylvestrově maticové rovnici. Nechť $F \in \mathcal{F}_s(A, B, L)$, potom z definice podobných matic plyne, že existuje regulární matice $T \in \mathbb{R}^{n \times n}$ taková, že

$$A + BF = TMT^{-1},$$

neboli

$$AT - TM + BFT = 0, \quad (8)$$

kde

$$M = \begin{bmatrix} L & R \\ 0 & S \end{bmatrix}$$

a R, S jsou vhodné matice. Položme $T = [X, V]$, $X \in \mathbb{R}^{n \times s}$, $V \in \mathbb{R}^{n \times (n-s)}$, potom z (8) plyne

$$AX - XL + BFX = 0,$$

neboli

$$AX - XL + BH = 0 \quad (9)$$

kde $H \triangleq FX \in \mathbb{R}^{m \times s}$ je nějaká (neznámá) matice. Zatím jsme tedy ukázali, že je-li $F \in \mathcal{F}_s(A, B, L)$, potom existují matice X a H takové, že platí rovnice (9). Předpokládejme, že matice H je známá, potom (9) představuje maticovou rovnici pro neznámou matici X . V takovém případě je rovnice (9) nazývána Sylvestrovou maticovou rovnicí. Všimněme si, že jde o soustavu lineárních rovnic pro neznámé prvky matice X . Za předpokladu, že matice A a L nemají společná vlastní čísla (tuto skutečnost budeme dále zapisovat následovně: $\sigma(A) \cap \sigma(L) = \emptyset$), existuje pouze jediné řešení $X = X(H)$ maticové rovnice (9). Hledanou zpětnovazební matici F můžeme potom určit z rovnice

$$H = FX(H).$$

To je opět soustava lineárních rovnic pro prvky matice F . Její obecné řešení je ve tvaru

$$F = H [X^T(H)X(H)]^{-1} X^T(H) + F_0, \quad (10)$$

kde F_0 je libovolná matice splňující podmínku $F_0X(H) = 0$. K tomu, aby pravá strana rovnosti (10) byla definovaná, musí mít řešení $X(H)$ rovnice (9) plnou sloupcovou hodnotu. Z tohoto důvodu definujeme množinu

$$\mathcal{H}_s(A, B, L) \triangleq \{H \in \mathbb{R}^{m \times s} : \text{rank } X(H) = s\}.$$

Zjemněním předchozích úvah dospějeme k následující větě, která poskytuje explicitní parametrizaci množiny $\mathcal{F}_s(A, B, L)$ pomocí neminimálního počtu ms parametrů.

Věta 2. [8] Nechť $\sigma(A) \cap \sigma(L) = \emptyset$, $\mathcal{F}_s(A, B, L) \neq \emptyset$, potom libovolná matice $F \in \mathcal{F}_s(A, B, L)$ může být vyjádřena ve tvaru

$$F = \varphi(H, F_0) \triangleq H [X^T(H)X(H)]^{-1} X^T(H) + F_0, \quad (11)$$

kde $H \in \mathcal{H}_s$, $X(H)$ je řešení maticové rovnice (9) a $F_0 \in \mathbb{R}^{m \times n}$ splňuje podmínku $F_0X(H) = 0$. Je-li $s = n$, potom $F_0 = 0$. Obráceně, za uvedených předpokladů je matice $F = \varphi(H, F_0)$ daná vztahem (11) prvkem množiny $\mathcal{F}_s(A, B, L)$.

Poněvadž lze snadno dokázat, že doplněk $\mathcal{V}$ množiny $\mathcal{H}_s$ v prostoru $\mathbb{R}^{m \times s}$, za předpokladu $\mathcal{F}_s(A, B, L) \neq \emptyset$, je sjednocení konečného počtu nadploch nižší dimenze než ms , skoro každá matice H patří do množiny $\mathcal{H}_s$. V důsledku toho můžeme vztah (11) využít pro výpočet libovolné $F \in \mathcal{F}_s(A, B, L)$. Zvolíme-li $H \in \mathbb{R}^{m \times s}$ například náhodně, s pravděpodobností 1 platí $H \in \mathcal{H}_s$ a vztah (11) můžeme použít pro výpočet příslušné stavové zpětné vazby. Jinými slovy, probíhá-li matice H všechny matice z prostoru $\mathbb{R}^{m \times s}$, probíhá F , daná vztahem (11), všechny matice množiny $\mathcal{F}_s(A, B, L)$. Problém s inverzí v (11) může nastat pouze pro „patologické“ matice H . Má-li řízený systém A, B více vstupů ($m > 1$), potom množina $\mathcal{F}_s(A, B, L)$ obsahuje nekonečně mnoho prvků a volnost ve volbě H může být využita ke splnění dalšího návrhového požadavku. Například robustnosti vzhledem ke stabilitě uzavřené smyčky. Bohužel vztah (11) nenabízí parametrizaci pomocí minimálního počtu parametrů. Taková „minimální“ parametrizace je uvedena v práci [8]. Z důvodu její technické složitosti se jí zde nebudeme podrobně zabývat. Naznačíme však v čem spočívá.

Ukazuje se, že zafixujeme-li některé prvky matice H ve vztahu (11), „neztrácíme“ žádné matice F z množiny $\mathcal{F}_s(A, B, L)$. Zaměníme tedy matici H za tzv. parametrickou matici $Q(\alpha)$, která má některé prvky konstantní a zbylé tvoří volné složky nového návrhového vektorového parametru α . Poznamenejme, že strukturu fixních a volných prvků matice $Q(\alpha)$ lze snadno určit z přiřazované matice L v reálném Jordanově tvaru. Jde však o relativně složitou závislost a skutečnou konstrukci matice $Q(\alpha)$ je vhodné svěřit počítači (viz procedura *ParametricMatrix* v knihovně JFAP, [3]). Pro zajímavost uveďme ještě jaká je dimenze nového (redukovaného) návrhového parametru α . Omezíme se na jednodušší případ úplného přiřazení $s = n$. Zde máme v tomto případě v původní parametrizaci (11) mn volných parametrů. Nahradíme-li H parametrickou maticí $Q(\alpha)$ bude dimenze nového návrhového parametru $\alpha \in \mathbb{R}^\gamma$ rovna výrazu

$$\gamma = mn - \sum_{i=1}^k (2i - 1)\nu_i, \quad (12)$$

kde k a ν_i jsou určeny maticí L podle vztahu $\nu(L) = \{\nu_i\}_{i=1}^k$. Je-li L cyklická ($k = 1$, $\nu_1 = n$), obdržíme

$$\gamma = n(m - 1).$$

V některých speciálních případech však může být redukce volných návrhových parametrů mnohem dramatičtější. Například má-li L pouze jedno násobné vlastní číslo a platí-li $n = 9$, $m = 3$, $k = 3$, $\nu_1 = 3$, $\nu_2 = 3$, $\nu_3 = 3$, potom z (12) plyne

$$\gamma = 3 \cdot 9 - (3 + 3 \cdot 3 + 5 \cdot 3) = 0.$$

Tedy neexistuje žádná volnost ve výběru matice F z množiny $\mathcal{F}_s(A, B, L)$. Aby však v tomto případě existovalo řešení ($\mathcal{F}_s(A, B, L) \neq \emptyset$), musí podle věty 1 pro indexy řiditelnosti dvojice (A, B) platit: $\mu_1 = 3$, $\mu_2 = 3$, $\mu_3 = 3$.

Věta 2 a její zesílená varianta [8] má významné aplikace v návrhu regulátoru s omezenou strukturou. Nyní velmi stručně naznačíme proč. Ze vztahu (11) plyne, že každý prvek matice F je racionální funkcí v návrhových parametrech (prvky H nebo α a prvky F_0). Požadujeme-li, aby určité vybrané prvky matice F byly nulové, obdržíme soustavu polynomiálních rovnic v návrhových parametrech. Má-li tato soustava řešení, existuje i regulátor s příslušně zvolenou omezenou strukturou přiřazující danou Jordanovu formu L . Tato univerzální technika může být použita pro návrh výstupní nebo decentralizované zpětné vazby, pro návrh dynamického regulátoru daného řádu a dokonce i pro návrh klasického PID regulátoru.

Příklad 3. Uvažujeme následující problém úplného přiřazení Jordanovy formy:

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad L = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Platí $\mu(A, B) = \{3, 1\}$, $\nu(L) = \{3, 1\}$. Dále

$$\gamma = 2 \cdot 4 - 3 - 3 \cdot 1 = 2$$

a

$$Q(\alpha) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \alpha_1 & \alpha_2 & 0 & 1 \end{bmatrix}.$$

Řešení Sylvestrovské rovnice (9) je dáno vztahem

$$X(\alpha) = \begin{bmatrix} -1 & -\alpha_1 & 2\alpha_1 - \alpha_2 - 1 & 0 \\ -\alpha_1 & -\alpha_1 - \alpha_2 - 1 & \alpha_2 - 3\alpha_1 + 1 & -1 \\ \alpha_1 - 1 & \alpha_2 - 3\alpha_1 + 1 & 5\alpha_1 - 3\alpha_2 - 3 & 1 \\ 0 & \alpha_1 & \alpha_2 - \alpha_1 - 1 & 0 \end{bmatrix}$$

a

$$\det X(\alpha) = \alpha_1^2$$

Z parametrizace (11) plyne,

$$F(\alpha) = \frac{1}{\alpha_1^2} \begin{bmatrix} 0 & -\alpha_1^2 & -\alpha_1^2 & -2\alpha_1^2 \\ -2\alpha_1^2 + \alpha_2 - \alpha_1 + 1 & \alpha_1^2 - \alpha_2 + \alpha_1 - 1 & 2\alpha_1^2 + \alpha_1 - \alpha_2 - 1 & 3\alpha_1^2 - \alpha_2 - 1 \end{bmatrix}. \quad (13)$$

V [6] je ukázáno, že má-li L pouze jedno vlastní číslo a je-li $\mu(A, B) = \nu(L)$, potom je množina $\mathcal{F}_n(A, B, L)$ nadrovinou v prostoru matic $\mathbb{R}^{m \times n}$. Tedy parametrizace množiny $\mathcal{F}_s(A, B, L)$ může být afinní. Speciálním procesem implicitizace (13) obdržíme novou afinní parametrizaci

$$F(\beta) = \begin{bmatrix} 0 & -1 & -1 & -2 \\ -2 + \beta_1 & 1 - \beta_1 & 2 - \beta_1 & 3 - \beta_2 \end{bmatrix}. \quad (14)$$

4 Přiřazení Jordanovy formy výstupní zpětnou vazbou

Přiřazení Jordanovy formy (podobně jako přiřazení pólů) výstupní zpětnou vazbou je mnohem obtížnější problém než v případě stavové zpětné vazby. Přes pozoruhodné pokroky v této oblasti [10] neexistuje jeho uspokojivé řešení. Důvod spočívá v jeho silné nelinearitě. K efektivnímu řešení tohoto problému bychom potřebovali umět řešit soustavy polynomiálních rovnic v reálném oboru. To znamená umět nalézt všechna reálná řešení a nikoliv pouze některá numerickou metodou. V současné době existuje několik málo kandidátů na takové řešení. Všechny však trpí závažnými omezeními a nedostatky. Nejelegantnější z nich se zdá technika Gröbnerovy báze [1], kterou lze stručně popsat jako dalekosáhlé zobecnění Gaussovy eliminační metody. Tato metoda dovoluje v principu řešit libovolnou soustavu polynomiálních rovnic s racionálními koeficienty v konečném počtu kroků. Bohužel zaručený konečný počet kroků této metody nazabezpečuje rozumné nároky na čas výpočtu a potřebnou paměť počítače. Technika Gröbnerovy báze je nicméně podporována v mnoha nástrojích pro symbolické výpočty (Maple, Mathematica, ...) a proto je použita i zde.

Problém přiřazení Jordanovy formy výstupní zpětnou vazbou spočívá v nalezení všech výstupních zpětnovazebních matic $K \in \mathbb{R}^{m \times p}$ patřících do množiny

$$\mathcal{K}_s(A, B, C, L) \triangleq \left\{ K \in \mathbb{R}^{m \times p} : (A + BKC) \sim \begin{bmatrix} L & * \\ 0 & * \end{bmatrix} \right\},$$

kde $L \in \mathbb{R}^{s \times s}$, $1 \leq s \leq n$ je daná matice v reálném Jordanově tvaru, $*$ označuje libovolnou reálnou matici a symbol $\sim$ relaci podobnosti matic. Je-li $s = n$, resp. $s < n$, potom podobně jako v případě stavové zpětné vazby jde o úplné resp. částečné přiřazení.

Otázka, za jakých podmínek je množina $\mathcal{K}_s(A, B, C, L)$ neprázdná, není dosud uspokojivě vyřešena. Nejslibnějším výsledkem v tomto směru pro úplné přiřazení je tvrzení, že pro generickou trojici (C, A, B) a libovolnou cyklickou matici $L \in \mathbb{R}^{n \times n}$ je $\mathcal{K}_n(A, B, C, L) \neq \emptyset$ jestliže $n < mp$ [10]. Poněvadž nutná podmínka je pro tento případ zřejmě $n \leq mp$, zbývají tedy dořešit pouze případy $n = mp$. Ovšem předpoklad generičnosti systému (C, A, B) je v praxi většinou nepoužitelný, neboť některé prvky matic A, B, C jsou nulové nebo přesné konstanty.

Nyní stručně naznačíme, jak lze větu 2 použít pro výpočet hledané výstupní zpětné vazby K . Nechť $K \in \mathcal{K}_s(A, B, C, L)$, potom podle definice existuje $F \in \mathcal{F}_s(A, B, L)$ taková, že $F = KC$. Tedy podle věty 2 existuje $H \in \mathbb{R}^{m \times s}$ a F_0 takové, že

$$H [X^T(H)X(H)]^{-1} X^T(H) + F_0 = KC, \quad (15)$$

kde $X(H)$ je řešení maticové rovnice (9) a $F_0 X(H) = 0$. Vynásobením (15) zprava maticí $X(H)$ obdržíme

$$H = KCX(H).$$

Zaměníme-li matici H za parametrickou matici $Q(\alpha)$ příslušnou k matici L a $X(H)$ za příslušné $X(\alpha)$, podobně jako v případě stavové zpětné vazby, obdržíme lineární rovnici

$$Q(\alpha) = KCX(\alpha) \quad (16)$$

pro neznámé prvky $\alpha \in \mathbb{R}^s$ a matice K . Řešením (16) technikou Gröbnerovy báze získáme hledanou matici K .

5 Podpůrný software

Pro účely vývoje nových metod přiřazení Jordanovy formy výstupní zpětnou vazbou byla vyvinuta knihovna procedur JFAP [3] pro Maple 13. Autor tohoto článku se domnívá, že již tato předběžná verze knihovny může být užitečná i pro řešení praktických úloh (alespoň pro nízké dimenze $mp \leq 10$). Následující příklad naznačuje způsob jak může být knihovna využita.

Příklad 4. (Návrh PID regulátoru) Uvažujme řízený systém popsany přenosem

$$P(s) = \frac{1}{(s+4)^4(s+1)},$$

zapojený v regulační smyčce s PID regulátorem

$$C(s) = k_p + k_I \frac{1}{s} + k_D s.$$

naším cílem je určit parametry regulátoru k_p , k_I a k_D tak, aby navržený systém měl tři zadané dominantní póly $-p$, $-p \pm jp$, $p > 0$. Snadno se lze přesvědčit, že tato úloha lze převést na následující neúplné přiřazení Jordanovy formy

$$L = \begin{bmatrix} -p & & \\ & -p & -p \\ & p & -p \end{bmatrix},$$

výstupní zpětnou vazbou pro systém (C, A, B) :

$$A = \begin{bmatrix} -4 & 0 & 0 & 0 & 0 & 0 \\ 1 & -4 & 0 & 0 & 0 & 0 \\ 0 & 1 & -4 & 0 & 0 & 0 \\ 0 & 0 & 1 & -4 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, C = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 & 0 \end{bmatrix}.$$

Zpětnovazební matice K obsahuje jako své prvky parametry regulátoru k_p , k_I a k_D . Konkrétně $K = [-k_p, -k_I, -k_D]$. Výpočtem metodami z oddílu 4 obdržíme

$$\begin{aligned} k_p &= -10p^5 + 238p^4 - 1120p^3 + 1408p^2 - 256, \\ k_I &= -p^3(10p^3 - 170p^2 + 672p - 704), \\ k_D &= -p^4 + 85p^3 - 560p^2 + 1056p - 512. \end{aligned}$$

Ukázka řešení této úlohy pomocí knihovny JFAP v prostředí Maple je uvedena v dodatku 1. Poznamenejme, že podrobný návod k využití každé procedury z knihovny JFAP lze získat najetím kurzoru na jméno procedury a stisknutím klávesy „F2“.

6 Závěr

Článek zpřístupňuje některé výsledky práce [8] zabývající se přiřazením Jordanovy formy stavovou a výstupní zpětnou vazbou. Jeho cílem je podat dostatečné objasnění základního problému a metody řešení, které postačí pro plnohodnotné využití nově vyvinuté knihovny JFAP pro Maple 13. Pro nezjednodušená tvrzení s úplnými důkazy je zájemce odkazován na originální článek [8].

Reference

- [1] Cox, D.; Little, J., O'Shea, D.: Ideals, varieties and algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra. 2nd ed. New York, Springer, 1997.
- [2] Fiedler, M.: Speciální matice a jejich použití v numerické matematice. SNTL, Praha, 1981.
- [3] Königsmarková, J.: Přiřazení pólů lineárnímu dynamickému systému výstupní zpětnou vazbou, Bakalářská práce. Plzeň: Západočeská univerzita v Plzni, 2010
- [4] Rosenbrock, H. H.: State-space and multivariable theory., John Willey and Sons, Inc., 1970.
- [5] Schlegel, M.: Parametrization of the class of deadbeat controllers. IEEE Trans. Automat. Contr., vol. AC-27, 727-729, 1982.

- [6] Schlegel, M.: Modální řízení konečněrozměrných lineárních systémů, Dizertační práce, VŠSE, Plzeň, 1984, (na www.schlegel.zcu.cz/downloads).
- [7] Schlegel, M.: A generalization of Rosenbrock's theorem. IEEE Trans. Automat. Contr., 12: 660-665, 1986.
- [8] Schlegel, M., Königsmarková, J.: Parametric Jordan Form Assignment Revisited., přijato na IFAC WC, Milan 2011.
- [9] Syrmos, V.L., Abdallah, C.; Dorato, P., "Static output feedback: a survey", Proceedings of the 33rd IEEE Conference on Decision and Control, vol.1, pp.837-842, 1994.
- [10] Wang, X.: Grassmanian, central projection and output feedback pole assignment of linear systems. IEEE Trans. Autom. Control, vol. AC-41, 786-794, 1996.
- [11] Wothman, W. M.: On pole assignment in multi-input controllable linear systems. IEEE Trans. Automat. Contr., vol. AC-27, 727-729, 1967.

Dodatek 1

Návrh PID regulátoru metodou neúplného přiřazení Jordanovy formy pro

$$\text{řízený systém : } P(s) = \frac{1}{(s+4)^4(s+1)}$$

$$\text{regulátor : } C(s) = kp + \frac{ki}{s} + kd \cdot s$$

```

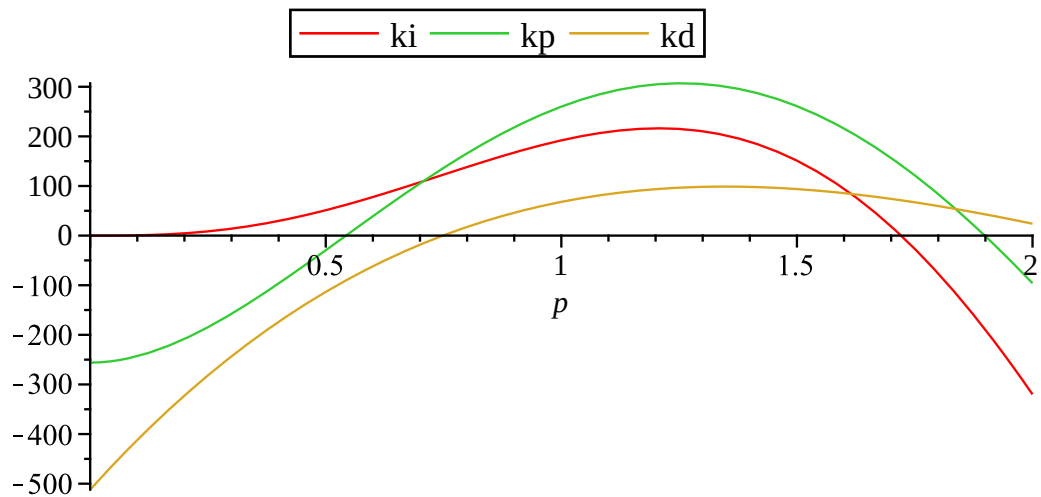
> with(LinearAlgebra) :
> with(JordanFormAssignment);
[ControllabilityIndexes, DeadbeatFeedback, OutputFeedback, OutputFeedback2,
  OutputFeedback3, ParametricMatrix, SolveElements, StateFeedback, StateFeedback2,
  Sylvesters]
> A := Matrix( [[ -4, 0, 0, 0, 0, 0], [1, -4, 0, 0, 0, 0], [0, 1, -4, 0, 0, 0], [0, 0, 1, -4, 0, 0], [0, 0,
  0, 1, -1, 0], [0, 0, 0, 0, 1, 0]] ) :
B := Matrix( [[1], [0], [0], [0], [0], [0]] ) :
C := Matrix( [[0, 0, 0, 0, 1, 0], [0, 0, 0, 0, 0, 1], [0, 0, 0, 1, -1, 0]] ) :
> Transpose(Eigenvalues(A));
      [ 0  -1  -4  -4  -4  -4 ]
(1)

> L1 := Matrix( [[ -p]] ) :
L2 := Matrix( [[ -p, -p], [p, -p]] ) :
L := [L1, L2];
Q := maticeQ(A, B, L);
      L := [ [ -p ], [ [ -p  -p ] ]
      Q := [ 1  1  0 ]
(2)

> K := OutputFeedback(A, B, C, L) :
K := K[1];
kp := -K[1, 1];
ki := -K[1, 2];
kd := -K[1, 3];
      "1 solution "
K := [ 1120 p^3 - 238 p^4 + 10 p^5 - 1408 p^2 + 256, 672 p^4 - 170 p^5 + 10 p^6 - 704 p^3, p^4 - 85 p^3
      + 560 p^2 - 1056 p + 512 ]
      kp := -1120 p^3 + 238 p^4 - 10 p^5 + 1408 p^2 - 256
      ki := -672 p^4 + 170 p^5 - 10 p^6 + 704 p^3
      kd := -p^4 + 85 p^3 - 560 p^2 + 1056 p - 512
(3)

> plot( {kp, ki, kd}, p=0..2);
(4)

```

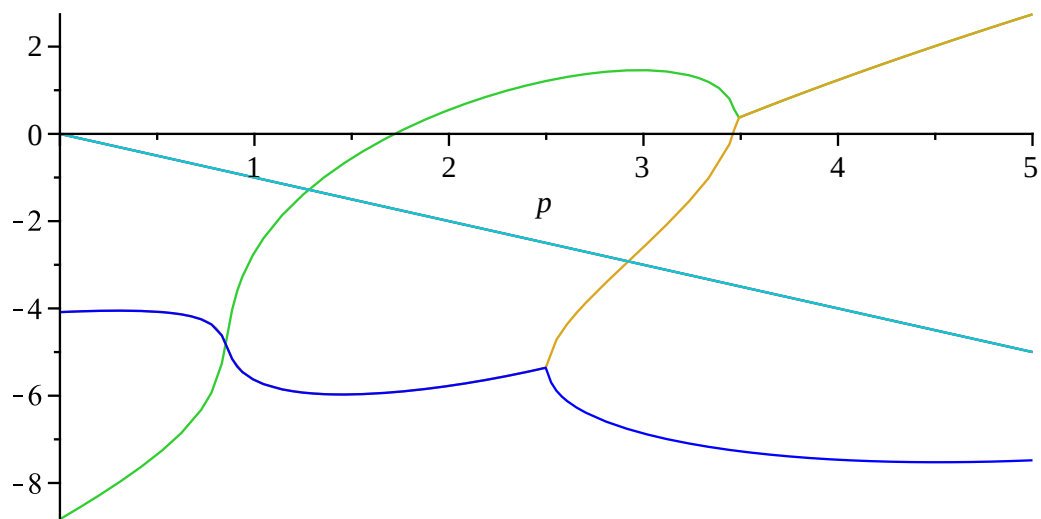


```
> Ac := A + B.K.C :
  spAc := Eigenvalues(Ac) :
  λ1 := spAc[1] : λ2 := spAc[2] : λ3 := spAc[3] : λ4 := spAc[4] : λ5 := spAc[5] : λ6
    := spAc[6] :
  [[λ1, λ2, λ3]]
```

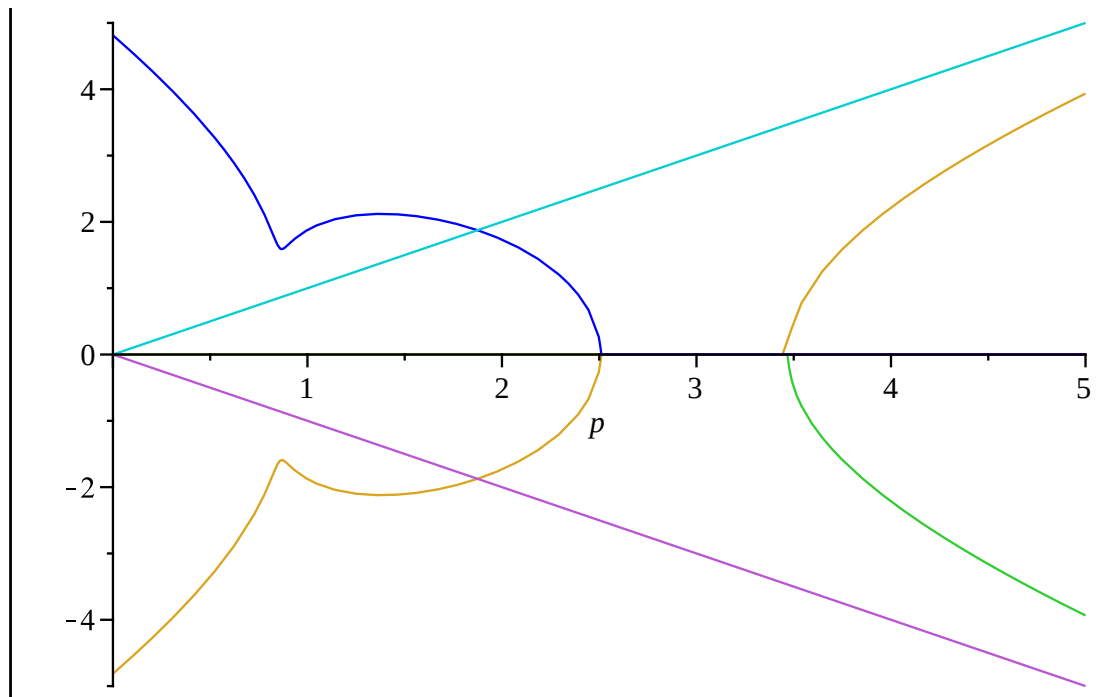
```
[[ -p, (-1 + I) p, (-1 - I) p]]
```

(5)

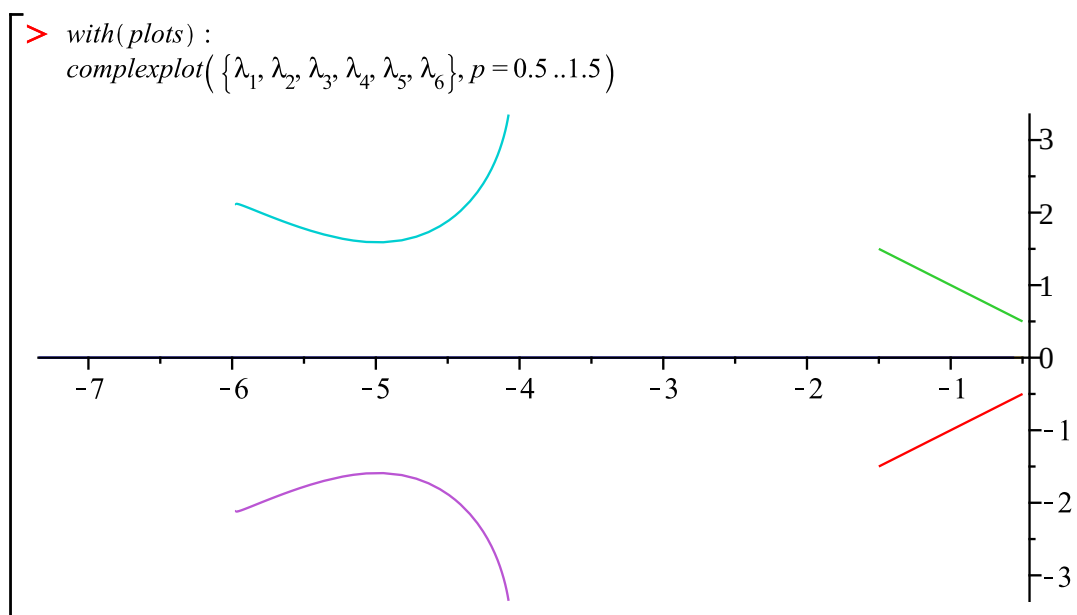
```
> plot({Re(λ1), Re(λ2), Re(λ3), Re(λ4), Re(λ5), Re(λ6)}, p = 0 .. 5);
```



```
> plot({Im(λ1), Im(λ2), Im(λ3), Im(λ4), Im(λ5), Im(λ6)}, p = 0 .. 5);
```



Interval parametru p pro rucni ladeni $p=0.5..1.5$



ŘÁDOVÁ ANALÝZA SIGNÁLŮ Z TOČIVÝCH STROJŮ S PROMĚNLIVÝMI NEBO NEUSTÁLENÝMI OTÁČKAMI

Jiří TŮMA

VŠB – Technická univerzita v Ostravě
17. listopadu 15, 708 33 Ostrava
E-mail: jiri.tuma@vsb.cz

Abstrakt/Abstract:

Rotační stroje se velmi často neotáčejí naprosto konstantními otáčkami se stabilitou srovnatelnou s frekvencí vzorkování při záznamu signálu jako je například hluk nebo vibrace. Naproti tomu vybuzené složky spektra těchto signálů jsou přesnými násobky frekvence otáčení. Pro účely diagnostiky nebo kontroly kvality je třeba tyto složky sledovat metodami, které se nazývají řádová analýza. Jiné označení je souběhová filtrace nebo analýza. Referát popisuje některé nejvíce používané metody. Je to řádová analýza na FFT signálových analyzátoch, dále kvadrurní směšování a postup s velkým potenciálem praktického uplatnění, který má název Vold-Kalmanova řádová filtrace.

Klíčová slova/Keywords: order analysis, quadrature mixing, Vold-Kalman order filter

1 Úvod

Při analýze chování rotačních strojů nebo periodicky pracujících mechanismů je třeba použít řádovou analýzu. Řád (order) v tomto smyslu je vybuzená frekvenční složka signálu, jejíž frekvence je určitým fixním násobkem základní frekvence stroje, kterou je například frekvence otáček nebo opakování vratných pohybů. Rezonanční frekvence nejsou závislé na otáčkách. Výstupem řádové analýzy je časový průběh amplitudy nebo RMS složky signálu se zvoleným řádem. Tato přednáška se bude zabývat třemi metodami sledování řádů:

- Výpočtem řádových spekter
- Kvadrurním směšováním
- Vold-Kalmanovou filtrací

Výsledkem výpočtu řádového spektra je běžné frekvenční spektrum s pozmeněnou frekvenční osou. Další dvě metody poskytují časový průběh amplitudy složky s frekvencí, která se mění v závislosti například na otáčkách stroje. Řádová analýza může být použita pro ustálené otáčky stroje nebo jeho rozběh a doběh. Ustálené otáčky nelze prakticky dosáhnout, proto sledování řádů reaguje i na drobné změny otáček. Všechny popisované metody vyžadují znalost okamžité rychlosti otáčení. Metodou měření frekvence otáček se referát nezabývá. Je předpokládáno, že okamžitou rychlost otáčení lze interpolovat z vyhodnocení otáček z tachosignálu.

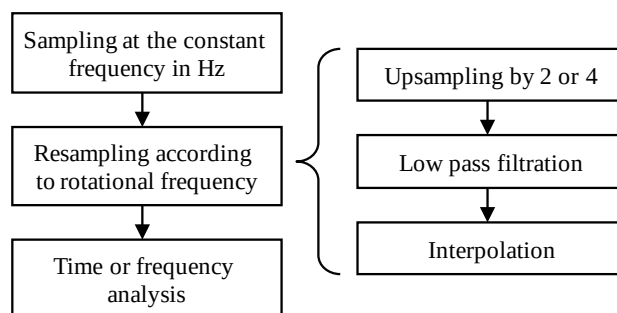
2 Řádová spektra

Řádová spektra se neopírají o frekvenci v hertzech, ale v řádech [1]. Frekvenční osa řádových spekter je bezrozměrná na rozdíl od frekvenčních spekter, kde je v Hz. Rozbor signálů hluku a vibrací rotačních strojů je nejčastější aplikací řádové analýzy. Nemusí jít o výpočet řádového spektra, ale i o záznamy, ve kterých je měřítko času nahrazeno úhlem otočení, přičemž základní periodě odpovídá stálý počet vzorků. Vzorkovací frekvence v počtu vzorků za sekundu je nahrazena počtem vzorků za jednu perioduotočení, například otáčku hřídele.

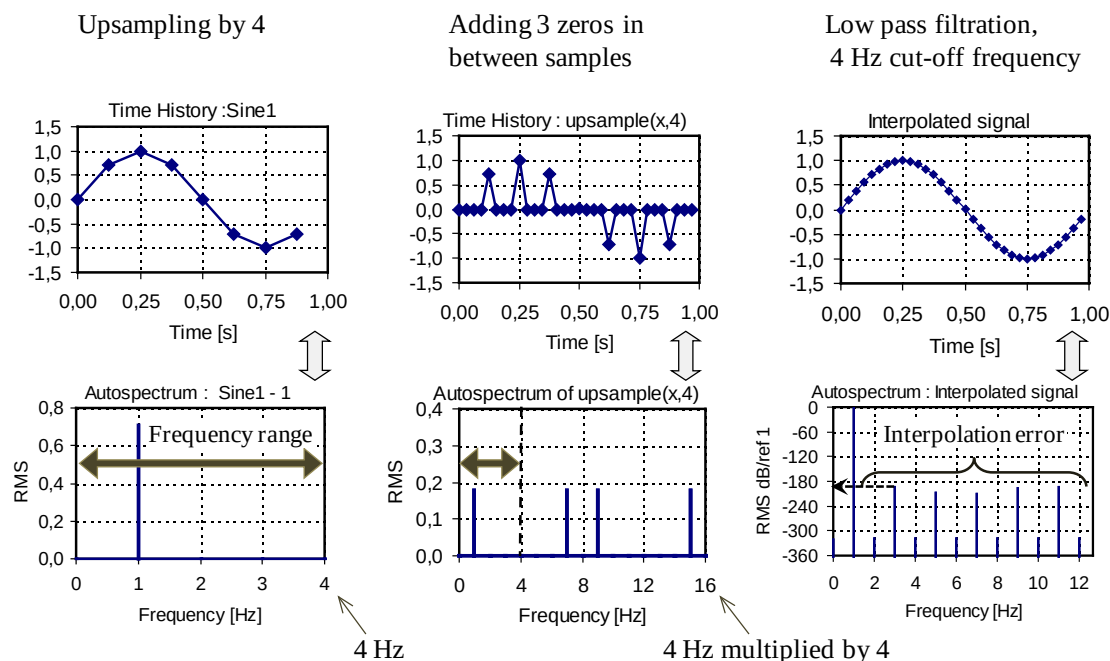
Hned v úvodu poznamenejme, že otáčky strojů nelze udržet na stálé hodnotě souměřitelné svou přesností se stabilitou elektrického oscilátoru, který řídí vzorkování signálu při měření. V tomto ohledu je analýza vlastností strojů mnohem komplikovanější než analýza elektronického obvodu. Konstantní počet vzorků na otáčku stroje byl u signálových analyzátorů v minulosti zajišťován externě řízeným vzorkováním například od IRC snímače produkujícího počet impulsů v počtu stovek, nejčastěji mocninu dvou, aby záznam vyhovoval výpočtu FFT. Je však obtížné připojit na konec hřídele encoder a zajistit jeho naprosto synchronní otáčky s analyzovaným strojem. Často bylo možné získat jen jeden impuls za otáčku. K signálovým analyzátorům se proto dodávaly frekvenční násobičky jako například BK 5050. Regulace frekvence impulsů na výstupu násobičky nemohla být z podstaty funkce fázového závěsu přesná. Frekvenční násobičky se proto už nepoužívají, jejich místo zastoupilo software analyzátoru, kterému se říká digital order analysis.

2.1 Digital order analysis

Tento způsob řádové analýzy se opírá o vzorkování stálou frekvencí a následný přepočet vzorků signálu na rozložení nových vzorků tak, aby jejich počet pro jednu periodu, například otáčku, byl fixní a pro případ výpočtu FFT roven mocnině dvou. Převzorkování znamená interpolaci nových vzorků mezi původními vzorky. Postup je znázorněn na obrázku 1. Pro převzorkování je uměle zvýšena vzorkovací frekvence obvykle dvakrát nebo čtyřikrát a až pak se provádí interpolace.



Obrázek 1: Postup převzorkování

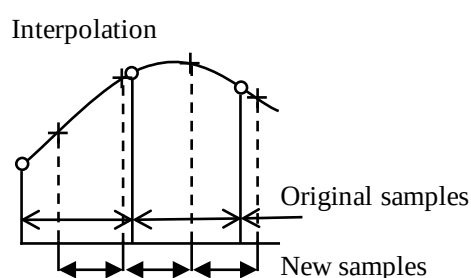


Obrázek 2: Zředění signálu se čtyřnásobným zvětšením vzorkovací frekvence

Shannonův vzorkovací teorém vyžaduje aspoň dva vzorky za periodu harmonického signálu. Teoreticky je třeba pro jednu periodu tohoto signálu 2,56 vzorků při respektování přechodového pásma antialiasingového filtru. Pro tento hraniční počet by byla interpolace nepřesná. Další zpřesnění lze dosáhnout například čtyřnásobným zvětšením vzorkovací frekvence operací, která se nazývá zředění signálu, což v daném případě znamená čtyřnásobné zvětšení vzorkovací frekvence. Mezi sousední vzorky vloží tři nuly. Zředění signálu však změní frekvenční spektrum. Původní signál lze obnovit dolnoproputnou filtrací. Algoritmus zvýšení vzorkovací frekvence zředěním nulami je demonstrován na obrázku 2.

Frekvenční spektrum s nulovými vzorky se po výpočtu FFT změní. Zobrazovaná polovina spektra obsahuje dvě úplná spektra ve směru frekvenční osy vedle sebe. První úplné spektrum má nenulové složky o frekvencích 1 a 7 Hz a druhé spektrum má nenulové složky o frekvencích 9 a 15 Hz. Toto spektrum je zobrazováno jako ostatní spektra reálných signálů jen polovičním počtem složek. Úplné spektrum obsahuje celkem 4 úplná spektra výchozího nezředěného signálu, protože se vzorkovací frekvence tímto způsobem zvětšila čtyřikrát.

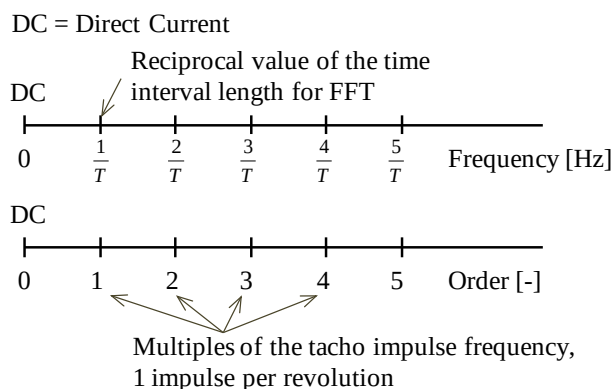
Dolnoproputná filtrace v uvedeném příkladu byla provedena pro jednoduchost prostřednictvím filtrace ve frekvenční oblasti. Vložení vzorků znamenalo zvětšení vzorkovací frekvence na celočíselný násobek původní vzorkovací frekvence. Další krok výpočtu vyžaduje již změnu vzorkování danou neceločíselným násobkem vzorkovací frekvence, tj. teoreticky umístění nového vzorku mezi původní vzorky na libovolné místo, jak ukazuje obrázek 3. K interpolaci mezi vzorky signálu se zvětšenou vzorkovací frekvencí může být využit například Newtonův interpolační polynom.



Obrázek 3: Interpolace mezi vzorky

2.2 Řádové spektrum

Jak již bylo uvedeno, frekvenční spektrum má frekvenční stupnici v hertzech. Rozdíl frekvencí sousedních složek spektra je roven převrácené hodnotě doby trvání záznamu, ze kterého se počítá FFT. Necht' je FFT převzorkovaného signálu počítána ze záznamu o délce odpovídající přesně jedné otáčce. V tomto případě bude frekvenční vzdálenost složek rovna frekvenci otáček. Složky spektra lze pak označovat pořadím jako harmonické frekvence otáčení, tj. bezrozměrné řády – Order nebo ve zkratce Ord, viz. obrázek 4.



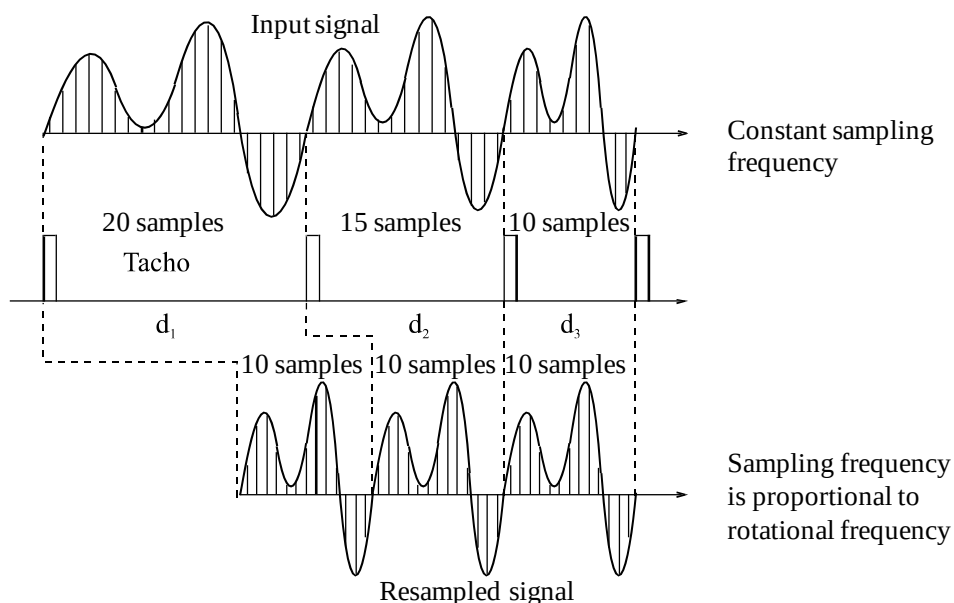
Obrázek 4: Interpolace mezi vzorky

Jestliže bude pro výpočet FFT vzat záznam o délce n -násobku celého otočení, pak spektrum bude obsahovat zlomkové řády o frekvenční vzdálenosti $1/n$.

Výpočet řádového spektra s použitím FFT vede ke stanovení amplitudy a fáze zvoleného řádu.

2.3 Příklad řádové analýzy

Příklad převzorkování je znázorněn na obrázku 5. Perioda signálu se zkracuje, což odpovídá vzrůstu rychlosti otáčení. V daném případě bude po převzorkování obsahovat každá otáčka stejný počet vzorků. Řádové spektrum bude obsahovat jen harmonické složky základní frekvence otáček.

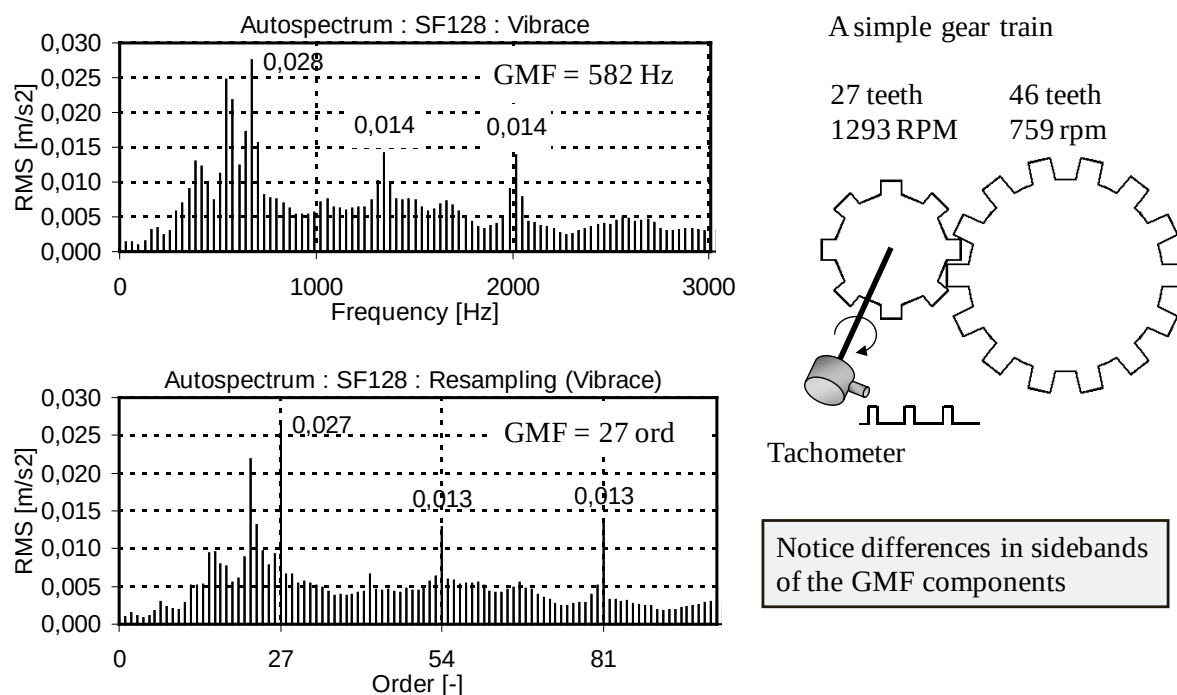


Obrázek 5: Příklad převzorkování signálu při rozběhu točivého stroje

Další příklad řádové analýzy je uveden na obrázku 6. V tomto obrázku je pro srovnání znázorněno frekvenční spektrum s frekvenční osou v hertzech. Pro určení odezvy záběru ozubených kol je třeba znát otáčky vstupní nebo výstupní hřídele a samozřejmě počet zubů soukolí. Záběrová frekvence ozubeného kola (GMF – gear meshing frequency) je dána součinem frekvence otáčení ozubeného kola a jeho počtu zubů. Řádové spektrum je možné vypočítat buď pro frekvenci otáčení vstupního nebo výstupního hřídele. Tato frekvence se určí z frekvence tachosignálu. Převzorkování záznamu se provádí pro časové intervaly ohraničené sousedními impulsy tachosignálu, přesněji mezi okamžiky překročení náběžné nebo sestupné hrany tachopulsů trimovací úrovní. Počet vzorků záznamu po převzorkování je volen tak, aby bylo možné vypočítat spektrum užitím algoritmu FFT.

Řádové spektrum má frekvenční osu v Order, která je bezrozměrná a udává frekvenci jako násobek frekvence otáčení vstupního hřídele s pastorkem o počtu zubů 27. Základní frekvence odezvy záběru zubů má tedy bezrozměrnou frekvenci 27 ord. Amplituda této komponenty spektra je vypočtena přesně bez zkreslení vlivem známého jevu, proti kterému se čelí používáním časových oken. V řádové analýze signálů, které obsahují jen komponenty s celočíselným násobkem základní frekvence, není třeba použít speciální časové okno. Vlivem kolísání otáček jsou složky o záběrové frekvenci neostré a navíc i zkresleny.

Frequency and order spectrum of a gearbox vibration



Obrázek 6: Frekvenční a řádové spektrum vibrací převodovky s ozubenými koly

3 Kvadrurní směšování

Kvadrurní směšování (anglicky quadrature mixing) je metoda demodulace modulovaného harmonického signálu. Nosná frekvence je rovna frekvenci složky, jejíž amplitudu a fázi je žádáno sledovat. Kvadrurní směšování posune sledovanou frekvenční složku signálu o zmíněné frekvenci do nuly tak, že po dolnoproustné filtraci lze určit jak její amplitudu, tak fázi. Necht' je harmonický signál popsán vzorcem ve tvaru

$$x(t) = A(t) \sin(\omega_0 t + \Phi(t)) \quad (1)$$

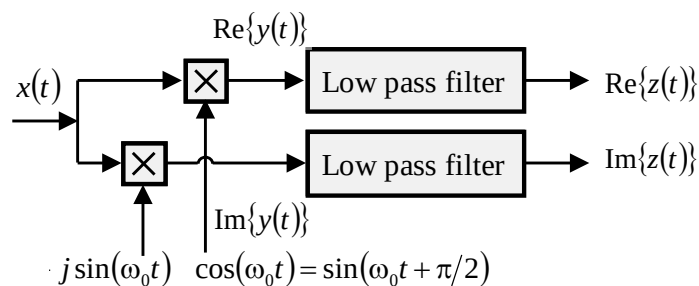
kde $A(t)$ je v čase proměnná amplituda, $\Phi(t)$ je v čase proměnná fáze a $\omega_0 = 2\pi f_0$ je úhlová frekvence. Modulovaný signál lze rozložit na dva modulované signály s fázovým posunem 90°

$$x(t) = I(t) \sin(\omega_0 t) + Q(t) \cos(\omega_0 t) \quad (2)$$

kde pro jednotlivé časově proměnné amplitudy platí

$$I(t) = A(t) \sin(\Phi(t)), \quad Q(t) = A(t) \cos(\Phi(t)) \quad (3)$$

Složka s funkcí kosinus se nazývá kvadrurní. Oddělit od sebe obě složky lze kvadrurním směšováním, jehož princip je znázorněn na obrázku 7. Výchozí signál $x(t)$ je vynásoben funkcí $\cos(\omega_0 t)$ a $j \sin(\omega_0 t)$ ($j = \sqrt{-1}$ je imaginární jednotka).



Obrázek 7: Kvadrurní směšování

Výsledek násobení neboli kvadrurního směšování je komplexní signál, který je označen $y(t)$

$$y(t) = A(t) \sin(\omega_0 t + \Phi(t)) (\cos(\omega_0 t) + j \sin(\omega_0 t)) \quad (4)$$

Reálná a imaginární složka komplexního signálu je následující

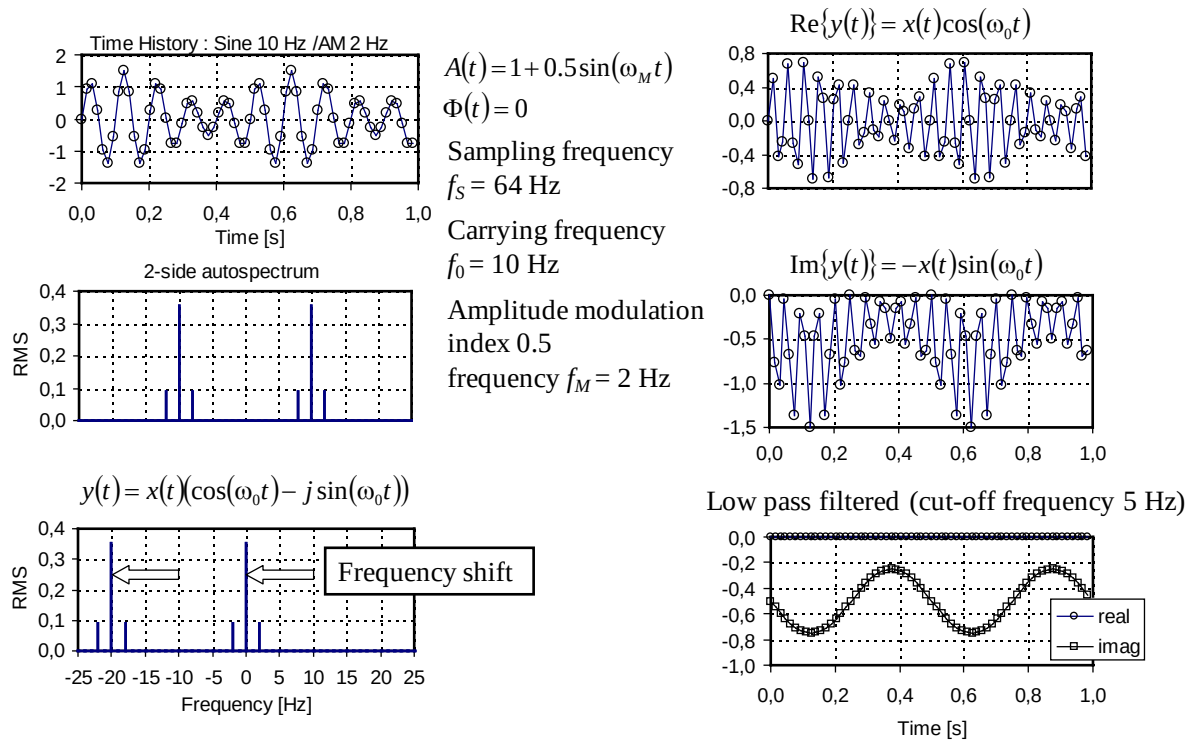
$$\text{Re}\{y(t)\} = A(t) \sin(\omega_0 t + \Phi(t)) \cos(\omega_0 t) = A(t) (\sin(\Phi(t)) + \sin(2\omega_0 t)) / 2 \quad (5)$$

$$\text{Im}\{y(t)\} = A(t) \sin(\omega_0 t + \Phi(t)) \sin(\omega_0 t) = A(t) (\cos(\Phi(t)) - \cos(2\omega_0 t)) / 2 \quad (6)$$

Původní nosný signál zdvojnásobil v obou složkách frekvenci. Po dolnoproputné filtraci lze dostat

$$\text{Re}\{z(t)\} = A(t) \sin(\Phi(t)) / 2 \quad (7)$$

$$\text{Im}\{z(t)\} = A(t) \cos(\Phi(t)) / 2 \quad (8)$$



Obrázek 8: Příklad na kvadrurní směšování

Absolutní hodnota signálu po kvadrurního směšování je obálka nebo amplituda složky signálu o úhlové frekvenci $\omega_0 = 2\pi f_0$ a fáze tohoto signálu je fáze měřená oproti modulačnímu signálu

$$\text{Abs}\{z(t)\} = |z(t)| = A(t)/2 \Rightarrow A(t) = 2|z(t)| \quad (9)$$

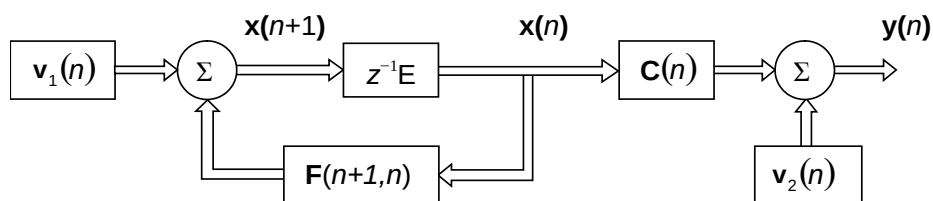
$$\text{Re}\{z(t)\}/\text{Im}\{z(t)\} = \tan(\Phi(t)) \Rightarrow [\Phi(t)]_{\text{WRAPPED}} = \arctan(\text{Re}\{z(t)\}/\text{Im}\{z(t)\}) \quad (10)$$

Příklad na kvadrurní směšování je uveden v obrázku 8. V tomto příkladu je místo modulační složky $\sin(\cdot)$ použita složka $-\sin(\cdot)$. Výsledek výpočtu fázového posunu se liší jenom ve znaménku oproti teoretickým vzorcům. Kvadrurní směšování vyžaduje znalost okamžité frekvence složky, kterou je žádáno sledovat.

4 Vold-Kalmanova filtrace

Výstupem Kalmanova filtru je časový průběh vývoje stavu lineárního systému, který je dostupný jen prostřednictvím nepřímého pozorování časového vývoje jistého vektoru dat. Algoritmus odhadování je rekurzivní, tedy zvláště vhodný pro číslicové zpracování dat v reálném čase, a chyba odhadu je minimální ve smyslu statistickém, což znamená, že střední hodnota druhé mocniny chyby je minimální. Bez bližšího popisu podstaty rekurzivních výpočtů bude v úvodu popsáno jen zadání úlohy filtrace jako referenční metody k Vold-Kalmanově filtraci. Označení veličin bude převzato z knihy [2].

Základem filtru, jehož blokové schéma je na obrázku 9, jsou dvě rovnice. První rovnice (process equation) popisuje vývoj stavu procesu popsaného v časovém okamžiku n vektorem stavových proměnných $\mathbf{x}(n)$ tak, že je definována souvislost mezi stavovým vektorem v okamžiku n a v následujícím okamžiku $n+1$. Stav procesu v okamžiku $n+1$ je ovlivněn náhodným vektorem $\mathbf{v}_1(n)$. Složky tohoto náhodného vektoru jsou členy bílé posloupnosti, tj. lze je označit za nekorelovaný šum. Druhá rovnice (measurement equation) modeluje měření stavu. Zatímco první rovnice byla v principu dynamická – diferenční, představuje druhá rovnice prostou transformaci hodnot vektoru stavu prostřednictvím transformační matice na hodnoty přístupné měření $\mathbf{y}(n)$. Transformovaný stav ovlivněn aditivním šumem $\mathbf{v}_2(n)$. Vektor šumu obsahuje také složky představující bílé posloupnosti. Blokové schéma vzájemných vazeb lze znázornit na následujícím diagramu. Symbol z^{-1} znázorňuje posunutí o jeden vzorek ($\mathbf{E}$ je jednotková matice)



Obrázek 9: Kalmanův filtrání

Výstupem Kalmanova filtru jsou odhady vektoru stavu procesu $\hat{\mathbf{x}}(n|\mathbf{y}(1), \dots, \mathbf{y}(n))$ v časovém okamžiku n na základě měřených hodnot $\mathbf{y}(1), \dots, \mathbf{y}(n)$ a odhadu počátečního stavu $\hat{\mathbf{x}}(1|\mathbf{y}(1))$ a korelačního vektoru chyby tohoto počátečního odhadu.

Zadání úlohy Kalmanovy filtrace lze aplikovat na problém extrakce harmonické složky z měřeného signálu. Stav procesu lze v této úloze považovat za hledanou harmonickou složku, která v součtu se zbývajícími složkami signálu představuje měřený signál. Popis vývoje stavu procesu představující tuto harmonickou složku s případně měnící se frekvencí bude předvedeno v další části popisu. Modelování měření představuje sumaci harmonické složky se

zbytkem složek obsažených v signálu. Jistý problém však představují budící signály $\mathbf{v}_1(n)$ a $\mathbf{v}_2(n)$ uvedeného signálového diagramu na obrázku 9. Jejich rozptyly nejsou předem známy a navíc signál $\mathbf{v}_2(n)$ nemusí být dominující bílý šum, ale další harmonická složka o jiné frekvenci.

První práce, které publikoval Vold a Leuridan [3], se datují rokem 1993. Detaily postupu výpočtu však nejsou popisovány. V úvahách Volda se respektuje neznalost rozptylů obou zmíněných budících posloupností a ve výpočtu se ovládá jen jejich vzájemný poměr, který ovlivňuje frekvenční vlastností filtru. Nový způsob filtrace autorů Vold a Leuridan je inspirující a otvírá nové možnosti řádové analýzy signálů se známou frekvencí, která je odvozena z měření otáček. Popularizace této metody filtrace byla zajištěna její implementací do signálového analyzátoru PULSE firmy Brüel & Kjær [4].

4.1 Vold-Kalmanův filtr první generace

Účelem první generace filtru tohoto typu bylo extrahování modulované harmonické složky o známé frekvenci pro analýzu časového průběhu změn její amplitudy například za rozběhu stroje. Zvýšená amplituda může indikovat například rezonanci. Vold-Kalmanův filtr může být použit také k odfiltrování zvolené části spektra pro systémy analyzující kvalitu zvuku.

4.1.1 Datová rovnice

Jestliže z měřeného signálu se vzorky $y(n)$ je žádoucí extrahovat jednu harmonickou složku, pak datová rovnice vyjadřuje prostý fakt, že měřený signál je složen z extrahované harmonické složky se vzorky označenými $x(n)$ a zbytku signálu se vzorky označenými $\eta(n)$, kde $n = 1, 2, \dots, N$ znamená pořadí vzorku. Její tvar pro vzorky s pořadím n je následující

$$y(n) = x(n) + \eta(n) \quad (11)$$

Harmonická složka může být amplitudově nebo fázově modulována. Zbytkové složky signálu mohou představovat nejen širokopásmový šum, ale také další harmonické složky.

Obecný tvar datové rovnice pro vzorky s pořadím n a s počtem P sledovaných harmonických složek $x_i(n)$ je

$$y(n) = \sum_{i=1}^P x_i(n) + \eta(n) \quad (12)$$

Tento vztah platí pro každý naměřený a extrahovaný vzorek signálu.

4.1.2 Strukturální rovnice pro harmonický signál

V čase spojitý harmonický signál

$$x(t) = A \cos(\omega t) \quad (13)$$

je řešením diferenciální rovnice druhého řádu s nulovým součinitelem tlumení. Pro případ vzorkování signálu v diskrétních časových okamžicích $t_n = n \Delta t$, $n = 0, 1, 2, \dots$ je harmonický signál řešením diferenční rovnice rovněž druhého řádu, jejíž charakteristická rovnice má dva komplexně sdružené kořeny

$$z_1 = \exp(j\omega \Delta t) \quad z_2 = \exp(-j\omega \Delta t) \quad (14)$$

kde ω je úhlová frekvence. Obecné řešení diferenční rovnice druhého řádu rovnice se zapisuje ve tvaru

$$x(n) = C(z_1^n + z_2^n) \quad (15)$$

Charakteristický polynom tvoří součin kořenových činitelů $(z - z_1)(z - z_2)$, ze kterých lze rekonstruovat výchozí homogenní diferenční rovnici

$$x(n) - 2 \cos(\omega \Delta t) x(n-1) + x(n-2) = 0 \quad (16)$$

příčemž její koeficient u $x(n-1)$ je vhodné označit

$$c(n) = 2 \cos(\omega \Delta t). \quad (17)$$

K výpočtu posloupností vzorků postačují hodnoty prvních dvou vzorků a samozřejmě velikost úhlové frekvence. Výsledkem řešení homogenní diferenční rovnice s konstantní velikostí ω je harmonický signál o konstantní amplitudě.

Harmonické složky měřených signálů z rotačních strojů mění ve skutečnosti svou frekvenci a amplitudu, tj. stávají se modulovanými harmonickými signály. Jednu harmonickou složku může popisovat rovnice

$$x(n) - 2 \cos(\omega \Delta t) x(n-1) + x(n-2) = \varepsilon(n) \quad (18)$$

kde $\varepsilon(n)$ je budící funkce, která spolu s velikostí okamžité úhlové frekvence ω určuje řešení diferenční rovnice s pravou stranou.

Přenos budící funkce $\varepsilon(n)$ na posloupnost vzorků $x(n)$ není stabilní, protože kořeny charakteristické rovnice leží na jednotkové kružnici. Zesílení je nekonečné pro složku o úhlové frekvenci ω .

4.1.3 Soustava datových a strukturálních rovnic

Datové rovnice pro všechny naměřené vzorky signálu a neznámé harmonické složky o celkovém počtu N s individuálními vzorky lze zapsat takto

$$\begin{bmatrix} y(1) \\ y(2) \\ \dots \\ y(N) \end{bmatrix} - \begin{bmatrix} x(1) \\ x(2) \\ \dots \\ x(N) \end{bmatrix} = \begin{bmatrix} \eta(1) \\ \eta(2) \\ \dots \\ \eta(N) \end{bmatrix} \quad (19)$$

Pohodlnější je používání zkráceného zápisu pomocí vektorů

$$\mathbf{y} - \mathbf{x} = \boldsymbol{\eta} \quad (20)$$

Velikost zbytkového signálu jako vektoru lze ohodnotit jeho euklidovskou normou (zobecněnou délkou), která v případě reálných signálů představuje součet čtverců jednotlivých vzorků $\eta(n) = y(n) - x(n)$, $n = 1, \dots, N$. Vektorový zápis této euklidovské normy zjednodušuje transpozice vektoru

$$\boldsymbol{\eta}^T \boldsymbol{\eta} = (\mathbf{y}^T - \mathbf{x}^T)(\mathbf{y} - \mathbf{x}) \quad (21)$$

Podobně jako datové rovnice lze také strukturální rovnice v počtu $N-2$ pro jednotlivé vzorky uspořádat následujícím způsobem

$$\begin{bmatrix} 1 & -c(1) & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & -c(2) & 1 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 1 & -c(N-2) & 1 \end{bmatrix} \begin{bmatrix} x(1) \\ x(2) \\ \dots \\ x(N) \end{bmatrix} = \begin{bmatrix} \varepsilon(3) \\ \varepsilon(4) \\ \dots \\ \varepsilon(N) \end{bmatrix} \quad (22)$$

Menší počet rovnic než N je dán řádem diferenční rovnice. Jako v případě datových rovnic je pohodlnější pracovat s maticovým zápisem soustavy strukturálních rovnic.

$$\mathbf{A} \mathbf{x} = \boldsymbol{\varepsilon}, \quad (23)$$

kde $\mathbf{A}$ je pásová třídiagonální obdélníková matice s N sloupci a $N-2$ řádky.

Podobně jako zbytkový signál $\eta(n)$ lze také velikost budící funkce se vzorky $\varepsilon(n)$, $n=1, \dots, N$ ohodnotit její euklidovskou normou. Transpozice vektoru zjednodušuje zápis součtu čtverců

$$\boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} = \mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} \quad (24)$$

Násobením matice $\mathbf{A}^T$ (horní index T označuje transpozici) maticí $\mathbf{A}$ vznikne čtvercová symetrická matice o rozměru $N \times N$. Euklidovská norma vektoru obsahujícího budící funkce má tvar kvadratická formy

$$\mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} \quad (25)$$

pro libovolný nenulový vektor budících funkcí $\mathbf{x} \neq \mathbf{0}$. Tato vlastnost se nazývá pozitivní semidefinitnost čtvercové symetrické matice $\mathbf{A}^T \mathbf{A}$. Pásovost se zvýší na 5, přičemž nenulové prvky jsou na dvou diagonálách pod a nad hlavní diagonálou této čtvercové pásové matice.

4.1.4 Globální řešení datových a strukturálních rovnic

Výsledkem filtrace má být nalezení neznámého vektoru $\mathbf{x}$, který musí vyhovovat soustavě datových a strukturálních rovnic. Počet prvků vektoru $\mathbf{x}$ je N a počet rovnic $N + (N - 2) = 2N - 2$. Mezi neznámé patří také vektory $\boldsymbol{\varepsilon}$ s $N - 2$ prvky a $\boldsymbol{\eta}$ s N prvky, pak počet neznámých $3N - 2$ je vyšší než počet rovnic, tj. soustava rovnic je vzhledem ke zmíněným třem neznámým vektorům nedourčena. K soustavě rovnic mohou být připojeny podmínky, podle kterých je požadováno, aby vektory $\boldsymbol{\varepsilon}, \boldsymbol{\eta}$ měly minimální euklidovskou normu, a aby tyto normy byly v žádoucím vztahu, tj. aby bylo minimalizováno kritérium ve tvaru

$$J = r^2 \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} + \boldsymbol{\eta}^T \boldsymbol{\eta} = \mathbf{x}^T (r^2 \mathbf{A}^T \mathbf{A} + \mathbf{E}) \mathbf{x} - \mathbf{y}^T \mathbf{x} - \mathbf{x}^T \mathbf{y} \quad (26)$$

kde r je váhový koeficient, který určuje předepsaný vzájemný vztah mezi euklidovskými normami zbytkového signálu a budící funkce. Jak bude dále ukázáno, podmínka minimální velikosti váženého součtu a zvolený váhový koeficient dávají vypočtenému vektoru výhodné vlastnosti, protože realizují filtr s výhodným průběhem frekvenční charakteristiky představující pásmovou propust. Hodnoty váhového koeficientu blízké nule zmenšují vlivnost strukturální rovnice a filtr ztrácí účinnost.

Podle monografie o speciálních maticích [5] lze dokázat, že jestliže je k hlavní diagonále pozitivně definitní matice přičteno libovolně malé číslo, pak je výsledná matice pozitivně definitní, tj. pro libovolný nenulový vektor budících funkcí $\mathbf{x} \neq \mathbf{0}$ platí ostrá nerovnost $\mathbf{x}^T (r^2 \mathbf{A}^T \mathbf{A} + \mathbf{E}) \mathbf{x} > 0$. Protože symetrická čtvercová matice $(r^2 \mathbf{A}^T \mathbf{A} + \mathbf{E})$ je pozitivně definitní, je také regulární a tedy existuje její inverzní matice.

První derivace skalárního kritéria J podle vektoru $\mathbf{x}$ se pro minimalizaci tohoto kritéria položí rovna nulovému vektoru

$$\frac{\partial J}{\partial \mathbf{x}} = 2r^2 \mathbf{A}^T \mathbf{A} \mathbf{x} + 2(\mathbf{x} - \mathbf{y}) = \mathbf{0} \quad (27)$$

což lze přepsat do tvaru

$$\mathbf{B} \mathbf{x} = \mathbf{y} \quad (28)$$

kde $\mathbf{B} = \mathbf{r}^2 \mathbf{A}^T \mathbf{A} + \mathbf{E}$.

Řešení soustavy rovnic je dáno následujícím vzorcem

$$\mathbf{x} = \mathbf{B}^{-1} \mathbf{y} \quad (29)$$

4.1.5 Adaptace váhového koeficientu na frekvenci a šířku propustného pásma

Analýza frekvenčních vlastností pásmového filtru byly na základě výpočtu harmonického signálu s konstantním váhovým součinitelem, který se uplatnil v kritériu. Ve vektoru budící funkce

$$\mathbf{A} \mathbf{x} = \boldsymbol{\varepsilon} \quad (30)$$

je vhodné individuálně vážit každou její složku tak, aby pro okamžitou frekvenci naladění filtru byla zajištěna vhodná hodnota váhového součinitele pro zvolenou šířku propustného pásma filtru, tj.

$$\mathbf{K} \mathbf{A} \mathbf{x} = \mathbf{K} \boldsymbol{\varepsilon}_K \quad (31)$$

kde $\mathbf{K}$ je diagonální čtvercová matice s $N-2$ sloupce a $N-2$ řádky s kladnými diagonálními prvky závislými na okamžité frekvenci naladění filtru a požadované šířce propustného pásma filtru $k_{i,i} = k(f, \Delta f)$. Euklidovská norma vektoru budící funkce bude obsahovat součin čtvercových matic s váhovými koeficienty

$$\boldsymbol{\varepsilon}_K^T \mathbf{K}^T \mathbf{K} \boldsymbol{\varepsilon}_K = \mathbf{x}^T \mathbf{A}^T \mathbf{K}^T \mathbf{K} \mathbf{A} \mathbf{x} \quad (32)$$

Kritérium pro globální řešení datových a strukturálních rovnic nabude tvar

$$J = \boldsymbol{\varepsilon}_K^T \mathbf{K}^T \mathbf{K} \boldsymbol{\varepsilon}_K + \boldsymbol{\eta}^T \boldsymbol{\eta} \quad (33)$$

Vlastnosti matic se nezmění a řešení lze nalézt jako pro předchozí formulaci úlohy.

4.2 Vold-Kalmanův filtr druhé generace

Výsledek filtrace Vold-Kalmanovým filtrem první generace je časový průběh sledované harmonické složky signálu, která je frekvenčně a amplitudově modulována. Okamžitá frekvence je součástí zadání filtrace. Zájem se proto soustřeďuje jen na obálku (amplitudu) sledované složky. Takto definovaná úloha filtrace patří Vold-Kalmanovu filtru druhé generace.

4.2.1 Datová rovnice

Obálku signálu lze vypočítat z amplitudy sinové složky a amplitudy kosinové složky. Tento postup užívá zbytečně mezivýpočet dílčích údajů a komplikuje řešení. Výhodnější je modelovat sledovaný signál jako komplexní harmonický signál s komplexní amplitudou. Datová rovnice pro jednu sledovanou složku má tedy tvar

$$y(n) = x(n) \exp(j\Theta(n)) + \eta(n) \quad (34)$$

kde $x(n)$ je komplexní amplituda nebo také obálka signálu, j je imaginární jednotka, $\Theta(n)$ je průběžná fáze signálu od počátku záznamu a n je pořadí vzorků, přičemž $n = 1, \dots, N$. Shodně jako v datové rovnici filtru první generace je $y(n)$ měřený signál a $\eta(n)$ zbytkový signál.

Průběžná fáze signálu je obecně rovna časovému integrálu úhlové frekvence. V případě vzorkovaných signálů se vypočte podle vzorce

$$\Theta(n) = \sum_{i=0}^n \omega(i) \Delta t \quad (35)$$

kde $\omega(n)$ je okamžitá úhlová frekvence a $\Delta t = 1/f_s$ je interval vzorkování.

Zápis soustavy datových rovnic pro všechny naměřené vzorky lze uspořádat do vektorového tvaru

$$\mathbf{y} - \mathbf{C}\mathbf{x} = \boldsymbol{\eta} \quad (36)$$

kde $\mathbf{C} = \mathbf{C}^T$ je diagonální komplexní matice s diagonálou obsahující vzorky fáze

$$\mathbf{C} = \text{diag}\{\exp(j\Theta(1)), \exp(j\Theta(2)), \dots, \exp(j\Theta(N))\} \quad (37)$$

Tato matice stejně jako vektory $\mathbf{x}$, $\boldsymbol{\eta}$ je komplexní. Pro hodnocení průběhu zbytkového signálu $\boldsymbol{\eta}$ se použije opět euklidovská norma vektoru. Protože jsou vektory komplexní, je třeba nejen vektor transponovat, ale také změnit znaménko imaginární složky. Tento transponovaný komplexně sdružený vektor se označí $\boldsymbol{\eta}^H$. Součin tohoto řádkového vektoru a původního sloupcového vektoru je reálné číslo. Operace transponování matice a obrácení znaménka imaginární části jejích prvků se označuje horním indexem H . V případě matic se výsledná matice označuje jako hermitovská. Pro euklidovskou normu vektoru zbytkového signálu platí

$$\boldsymbol{\eta}^H \boldsymbol{\eta} = (\mathbf{y}^T - \mathbf{x}^H \mathbf{C}^H)(\mathbf{y} - \mathbf{C}\mathbf{x}) \quad (38)$$

4.2.2 Strukturální rovnice pro obálku harmonického signálu

Strukturální rovnice generuje účinkem budící posloupnosti komplexní amplitudu sledované modulované harmonické složky. Nejjednodušší požadavek na posloupnost amplitud je, aby se po sobě jdoucí hodnoty odlišovaly o určitou chybu představující budící signál, tj.

$$x(n) = x(n-1) + \varepsilon(n) \quad (39)$$

Poslední rovnice obsahuje jeden zpožděný vzorek amplitudy a nahrazuje průběh obálky po úsecích konstantou. Tuto rovnici lze také zapsat užitím operátoru zpětné difference

$$\nabla x(n) = \varepsilon(n). \quad (40)$$

Může být také požadováno, aby rozdíl dvou sousedních diferencí mezi vzorky obálky se odlišoval o určitou chybu. Časový průběh amplitudy signálu se nahrazuje (aproximuje) lineárními úseky. Pro aktuální vzorek platí

$$x(n) = x(n-1) + x(n-1) - x(n-2) + \varepsilon(n). \quad (41)$$

Užitím operátoru zpětné difference lze tento vzorec zapsat ve tvaru

$$\nabla^2 x(n) = \nabla x(n) - \nabla x(n-1) = \varepsilon(n). \quad (42)$$

Odvozená podmínka obsahuje dva zpožděné vzorky amplitudy. Tři zpožděné vzorky amplitudy obsahuje rovnice, která nahrazuje časový průběh obálky po úsecích parabolou

$$x(n) = 3x(n-1) - 3x(n-2) + x(n-3) + \varepsilon(n). \quad (43)$$

Operátorový tvar je následující

$$\nabla^3 x(n) = \nabla^2 x(n) - \nabla^2 x(n-1) = \varepsilon(n). \quad (44)$$

Navržené strukturální rovnice obsahují na pravé straně budící funkci. Řešení těchto diferenčních rovnic bez budícího signálu, tj. jako homogenních s nulovou pravou stranou, je v prvním případě konstanta. Ve druhém případě je řešením lineární funkce pořadí vzorků nebo

jinak řečeno polynom prvního stupně v proměnné, kterou je pořadí vzorku. Ve třetím případě jde o kvadratickou funkci pořadí vzorků n . Jestliže budící funkce bude považována za vstupní posloupnost lineární soustavy, pak její přenos na posloupnost výstupní, kterou jsou vzorky obálky, je nestabilní. Charakteristická rovnice má obecně násobný kořen v bodě 1 komplexní roviny.

Zobecněním předchozích rovnic lze formulovat podmínku pro aproximační polynom obecného stupně k , podle níž je difference aproximačního polynomu řádu $k + 1$ rovna nule, tj.

$$\nabla^{k+1}x(n) = 0. \quad (45)$$

Protože se obálka signálu mění v důsledku amplitudové modulace sledované složky signálu, je třeba strukturální rovnici doplnit budící funkcí. Obecný tvar strukturální rovnice je

$$\nabla^{k+1}x(n) = \varepsilon(n). \quad (46)$$

Strukturální rovnice pro tyto filtry mají tvar

Počet pólů Strukturální rovnice

1	$x(n) - x(n-1) = \varepsilon(n)$
2	$x(n) - 2x(n-1) + x(n-2) = \varepsilon(n)$
3	$x(n) - 3x(n-1) + 3x(n-2) - x(n-3) = \varepsilon(n)$
4	$x(n) - 4x(n-1) + 6x(n-2) - 4x(n-3) + x(n-4) = \varepsilon(n)$

Soustavu strukturálních rovnic pro všechny výstupní vzorky lze obecně popsat maticovou rovnicí, která byla použita u analýzy Vold-Kalmanova filtru první generace, tj. s maticí soustavy $\mathbf{A}$, vektorem neznámých hodnot komplexní obálky $\mathbf{x}$ a vektorem pravých stran $\mathbf{\varepsilon}$.

Matice soustavy $\mathbf{A}$ obsahuje jen konstantní prvky v závislosti na řádu filtru. Její rozměr pro jednopólový filtr je $(N-1) \times N$, pro dvoupólový filtr je rozměr matice soustavy $(N-2) \times N$ a pro třípólový filtr je rozměr $(N-3) \times N$. Jednopólovému filtru odpovídá dvoudiagonální obdélníková matice následujícího tvaru

$$\mathbf{A} = \begin{pmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & \dots & \dots & \\ & & & 1 & -1 \end{pmatrix}. \quad (47)$$

4.2.3 Globální řešení datových a strukturálních rovnic pro obálku harmonického signálu

Oproti postupnému zobecňování v případě filtru první generace bude u filtru druhé generace předpokládáno, že z důvodu adaptace na měnící se frekvenci sledované složky signálu jsou vzorky budící funkce, které tvoří vektor, násobeny individuálně váhovými koeficienty. Tyto reálné váhové koeficienty tvoří hlavní diagonálu diagonální matice $\mathbf{K}$. Účelem je adaptovat váhový koeficient na průběžně se měnící frekvenci sledované složky. Euklidovská norma charakterizující velikost vektoru složeného ze vzorků budící funkce, které jsou násobeny váhovými koeficienty, se změní s ohledem na skutečnost, že budící funkce je komplexní stejně jako vektor obálky $\mathbf{x}$. K transpozici vektoru pro výpočet normy je třeba přidat také změnu znaménka imaginární části jeho prvků, aby výsledek násobení vektorů bylo reálné číslo, tj.

$$\boldsymbol{\varepsilon}^H \mathbf{K}^T \mathbf{K} \boldsymbol{\varepsilon} = \mathbf{x}^H \mathbf{A}^T \mathbf{K}^T \mathbf{K} \mathbf{A} \mathbf{x} = \mathbf{x}^H \mathbf{D} \mathbf{x}. \quad (48)$$

S ohledem na počet neznámých je soustava rovnic, které jsou k dispozici pro výpočet obálky, nedourčena. Stejně jako v případě filtru první generace, také pro řešení těchto rovnic se doplní pomocným kritériem z váženého součtu euklidovských norem budící funkce a zbytkového signálu ve tvaru

$$J = \mathbf{x}^H \mathbf{D} \mathbf{x} + (\mathbf{y}^T - \mathbf{x}^H \mathbf{C}^H)(\mathbf{y} - \mathbf{C} \mathbf{x}). \quad (49)$$

kde $\mathbf{D} = \mathbf{A}^T \mathbf{K}^T \mathbf{K} \mathbf{A}$ je pomocná čtvercová matice zjednodušující zápis vzorců. Minimum kritéria se vypočte stejným postupem jako u filtru první generace

$$\mathbf{x} = \mathbf{B}^{-1} \text{diag}\{\exp(-j\Theta(1)), \exp(-j\Theta(2)), \dots, \exp(-j\Theta(N))\} \mathbf{y}. \quad (50)$$

resp.

$$\mathbf{x} = \mathbf{B}^{-1} \mathbf{C}^H \mathbf{y}, \quad (51)$$

kde $\mathbf{B} = \mathbf{A}^T \mathbf{K}^T \mathbf{K} \mathbf{A} + \mathbf{E}$.

4.3 Křížování řádů

V předchozím textu byl popsán výpočet výstupu Vold-Kalmanova filtru pro sledování jednoho řádu, tj. komponenty spektra s frekvencí měnící se v závislosti na čase. Některé stroje obsahují dva rotory s velmi volně vázanou frekvencí otáčení. Zvolené násobky frekvence otáčení mohou splývat, což se označuje jako křížování řádů. Obecně lze řešit úlohu s rozkladem na P harmonických komponent. Datová rovnice pro případ, kdy je hledána obálka několika komponent současně, bude obecnější než v předchozích případech

$$y(n) = \sum_{k=1}^P x_k(n) \exp(j\Theta_k(n)) + \eta(n) \quad (52)$$

Kritérium pro řešení soustavy nedourčených rovnic bude také obecnější než v předcházejících případech

$$J = \sum_{k=1}^P r^2 \boldsymbol{\varepsilon}_k^T \boldsymbol{\varepsilon}_k + \boldsymbol{\eta}^T \boldsymbol{\eta} = \sum_{k=1}^P r^2 \mathbf{x}_k^H \mathbf{A}^T \mathbf{A} \mathbf{x}_k + \left(\mathbf{y}^T - \sum_{k=1}^P \mathbf{x}_k^H \mathbf{C}_k^H \right) \left(\mathbf{y} - \sum_{k=1}^P \mathbf{C}_k \mathbf{x}_k \right). \quad (53)$$

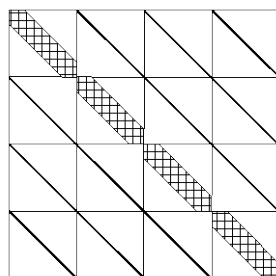
Neznámé vektory $\mathbf{x}_i^H, i=1, \dots, P$ v počtu P mají minimalizovat předcházející kritérium. Postup výpočtu P neznámých vektorů, určujících obálku, je shodný jako v předcházejících případech. Výsledkem postupného derivování podle těchto neznámých vektorů je P výrazů, které se musí v extrému kritéria J rovnat nule

$$\frac{\partial J}{\partial \mathbf{x}_i^H} = \mathbf{B}_i \mathbf{x}_i + \mathbf{C}_i^H \sum_{\substack{k=1 \\ k \neq i}}^P \mathbf{C}_k \mathbf{x}_k - \mathbf{C}_i^H \mathbf{y} = \mathbf{0}, \quad i=1, \dots, P \quad (54)$$

V rovnicích byla použita substituce $\mathbf{B}_k = r^2 \mathbf{A}^T \mathbf{A} + \mathbf{E}$ a $\mathbf{C}_i^H \mathbf{C}_i = \mathbf{E}$. Řešení soustavy přímo je prakticky nemožné. Matice soustavy už není pásová, ale je složena z pásových bloků na hlavní diagonále a diagonálních bloků mimo tuto diagonálu. Matice soustavy rovnic $\mathbf{B} \mathbf{x} = \mathbf{b}$ včetně vektoru neznámých veličin a vektoru pravé strany má následující strukturu

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_1 & \mathbf{C}_1^H \mathbf{C}_2 & \dots & \mathbf{C}_1^H \mathbf{C}_P \\ \mathbf{C}_2^H \mathbf{C}_1 & \mathbf{B}_2 & \dots & \mathbf{C}_2^H \mathbf{C}_P \\ \dots & \dots & \dots & \dots \\ \mathbf{C}_P^H \mathbf{C}_1 & \mathbf{C}_P^H \mathbf{C}_2 & \dots & \mathbf{B}_P \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \dots \\ \mathbf{x}_P \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} \mathbf{C}_1^H \mathbf{y} \\ \mathbf{C}_2^H \mathbf{y} \\ \dots \\ \mathbf{C}_P^H \mathbf{y} \end{pmatrix} \quad (55)$$

Pásové matice $\mathbf{B}_i$ s reálnými prvky jsou symetrické a pozitivně definitní a protože pro komplexní diagonální matice $\mathbf{C}_i^H \mathbf{C}_j$ platí $\mathbf{C}_i^H \mathbf{C}_j = (\mathbf{C}_j^H \mathbf{C}_i)^H$, je matice $\mathbf{B}$ soustavy rovnic hermitovská. Struktura nenulových prvků matice například pro $P=4$ je znázorněna na obrázku 10. Jednotlivé bloky jsou čtvercové matice o počtu řádků dosahujících desítek tisíc. Blokované matice na hlavní diagonále jsou pásové matice s několika nenulovými diagonálami (mezní počet několik desítek). Mimo hlavní diagonálu jsou blokované matice diagonální.



Obrázek 10: Struktura matice soustavy rovnic

Vhodným postupem se jeví metoda iterační, který bude těžit z vlastností matice soustavy. Tento postup pro Vold-Kalmanův filtr dříve navrhli také Feldbauer, Ch. & Holdrich, R. [6]. Iteračních metod pro řešení soustavy lineárních rovnic je velmi mnoho. Monografie popisující iterační řešení soustav lineárních rovnic (Saad: Iterative Methods for Sparse Linear Systems) nebo dokumentace k Matlabu označují jako zvláště vhodnou a navíc oblíbenou metodu pro řídké pozitivně definitní symetrické matice (symmetric positive definite – SPD), která se nazývá Preconditioned Conjugate Gradient (PCG) Algorithm. PCG metoda kombinuje iterační řešení

$$\mathbf{B} \mathbf{M}^{-1} \mathbf{u} = \mathbf{b} \quad (56)$$

s přímým výpočtem

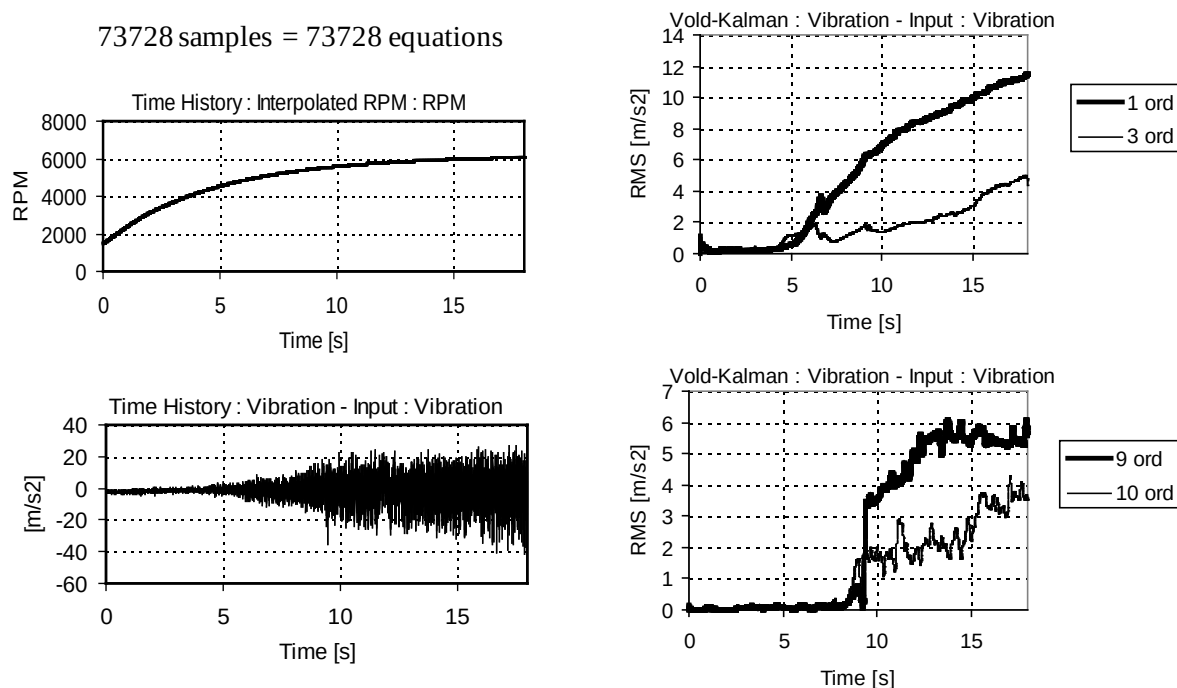
$$\mathbf{x} = \mathbf{M}^{-1} \mathbf{u} \quad (57)$$

kde $\mathbf{M}$ je předpokládaná snadno invertovatelná matice. Iterační výpočet vyžaduje první odhad řešení.

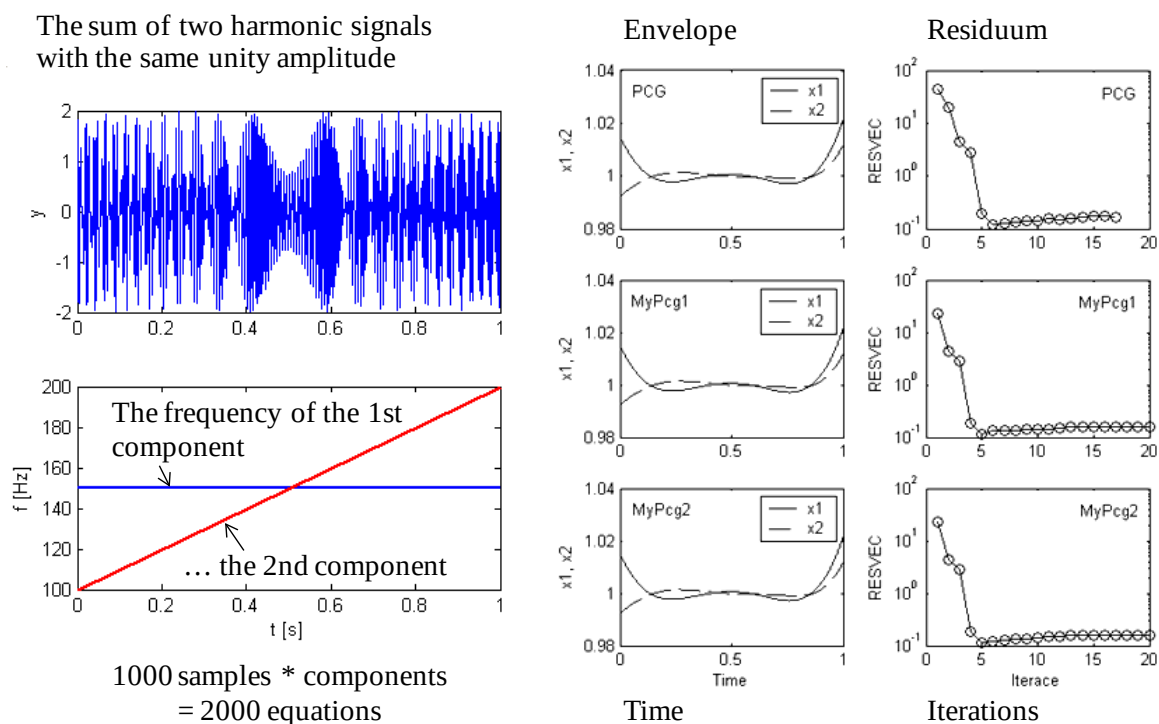
4.4 Příklady

První příklad na obrázku 11 vypočítává obálku několika složek signálu zrychlení, který byl naměřen na elektrickém motoru během rozjezdu. Jde o 1., 3., 9. a 10. Harmonickou frekvence otáček.

Druhý příklad z obrázku 12 odděluje dva harmonické signály se shodnou amplitudou rovnou jedné, přičemž první signál má konstantní frekvenci 150 Hz a druhý mění svou frekvenci lineárně s časem z frekvence 100 Hz na frekvenci 200 Hz. Odchyly od jednotkové amplitudy jsou na začátku a konci v rozsahu 1 až 2%. Po 5 iteracích se chyba řešení zmenší skoro tisíckrát.



Obrázek 11: Rozběh elektrického motoru



Obrázek 12: Oddělení křížujících řádů

4.5 Aplikace Vold-Kalmanovy filtrace v praxi

Vold-Kalmanova filtrace je metoda nevhodná pro reálný čas. Výpočet časového průběhu vyfiltrované složky signálu nebo její obálky je možný jen po ukončení měření, tj. v režimu postprocesingu. Pro filtr byly nalezeny aplikace z oblasti testování kvality výrobků, kdy je zapotřebí rychlý rozjezd a dojezd s měřením vibrací a hluku. Při tomto měření se sledují vybrané řády (násobky základní frekvence stroje) a diagnostikuje se překročení

mezních hodnot. Příkladem je aplikace filtru pro rozbor vnějšího hluku vozidel, jehož frekvenční složení závisí na otáčkách motoru [11], [12].

Ke smysluplné aplikaci Vold-Kalmanova filtru je třeba mít pod kontrolou šířku propustného pásma filtru [13], [14]. Zajímavý je rovněž náběh filtru [7] až [10]. Na zdokonalování řádové analýzy se ve světě intenzivně pracuje [15] až [19].

5 Závěr

Přínos práce spočívá hlavně v popisu teoretických základů a způsobu použití metod řádové analýzy signálů z rotačních strojů. Pozornost je zaměřena na výpočet řádových spekter, kvadrurní směšování a Vold-Kalmanovu řádovou filtraci. Uživatel se dozví, jak vzpomenuté metody fungují a co je jejich výsledkem.

Poděkování/Acknowledgement

The research has been supported by Czech Science Foundation under the project GA 102/09/0894.

Reference/References

- [1] Tůma, J. Zpracování signálů získaných z mechanických systémů. 1. vyd. Praha : Sdělovací technika, 1997. 174 s. ISBN 80-901936-1-7
- [2] Haykin, S. Adaptive filter theory. Prentice Hall PTR, Upper Saddle River, New Jersey 07458, 1996.
- [3] Vold, H. & Leuridan, J. Order Tracking at Extreme Slew Rates, Using Kalman Tracking Filters. SAE Paper Number 931288.
- [4] Gade, S., Herlufsen, H., Konstantin-Hansen, H., Vold, H., "Characteristics of the Vold-Kalman Order Tracking Filter", B&K Technical Review No 1 - 1999.
- [5] Fiedler, M. Speciální matice a jejich použití v numerické matematice. SNTL Praha, 1981.
- [6] Feldbauer, Ch. & Holdrich, R. Realisation of a Vold-Kalman Tracking Filter – A Least Square Problem, Proceedings of the COST G-6 Conference on Digital Audio Effects (DAFX-000, Verona Italy, December 7-9, 2000.
- [7] Tůma, J. Vold-Kalman filtration in MATLABu (in Czech). In Proceedings of Eleventh MATLAB Conference. Praha : Humusoft Praha, 25. 11. 2003, s. 575-586
- [8] Tůma, J. Frequency Response of the Vold Order Tracking Filter. In XXVI. ASR '2001 Seminar, Instruments and Control, Ostrava, April 26 - 27, 2001. 1-8 PP.
- [9] Tůma, J. Vold-Kalman Order Tracking Filtration as a Tool for machine diagnostics. In Engineering Mechanics 2001, Svratka, Czech Republic. 1st ed. Praha : Institute of Thermomechanics – Academy of Science, May 14 - 17, 2001. P136. ISBN 80-85918-64-1.
- [10] Tůma, J. Sound Quality Assessment Using Vold-Kalman Tracking Filtering. In Proceedings of XXIX. Seminary ASR '04 "Instruments and Control". Ostrava : Katedra ATR, VŠB-TU Ostrava, 2004, pp. 305-308. ISBN 80-248-0590-1.
- [11] Tuma, J. Dedoplerisation in Vehicle External Noise Measurements. In Proceedings of Eleventh International Congress on Sound and Vibration. St. Petersburg : IIAV, 5.-8.7. 2004, P 151-158.

- [12] Pelant, P. - Tůma, J. - Beneš, T. Vold-Kalman Order Tracking Filtration in Car Noise and Vibration Measurements. In Proceedings of the 33rd International Congress and Exposition on Noise Control Engineering – Internoise 2004. Prague : Czech Acoustical Society, 22.-25..8. 2004, P 1/8-8/8. ISBN 5-7325-0816-3.
- [13] Tůma, J. Setting the passband width in the Vold-Kalman order tracking filter. In: Twelfth International Congress on Sound and Vibration, (ICSV12). Lisabon, July 11-14, 2005, Paper 719.
- [14] Tůma, J. The passband Width of the Vold-Kalman Order Tracking Filter. Sborník vědeckých prací VŠB-TU Ostrava, řada strojní r. LI, 2005. č. 2, příspěvek č. 1485, s. 149-154. ISSN 1210-0471. ISBN 80-248-0880-X.
- [15] KeSheng Wang, Approaches to the improvement of order cracking techniques for vibration based diagnostics in rotating machine, Doctoral thesis, University of Pretoria, 2010
- [16] KeSheng Wang, P.S. Heyns, The combined use of order tracking techniques for enhanced Fourier analysis of order components. Mechanical systems and signal processing (2010), doi: 10.1016/j.ymssp.2010.10.005.
- [17] Yu Guo, Kok Kiong Tan, High efficient crossing-order decoupling in Vold–Kalman filtering order cracking based on independent komponent analysis, Mechanical Systems and Signal Processing, 24(2010) 1756–1766.
- [18] Yu Guo , Yilin Chi, New decoupling approach for Vold-Kalman filtering order tracking, 郭瑜... - 振动与冲击, 2009 - jvs.sjtu.edu.cn
- [19] Yu Guo; Kok Kiong Tan; Sunan Huang; Yi Zhang; Noise removal in Vold-Kalman order tracking based on independent component analysis. International Conference on Automation and Logistics, Qingdao, China September 2008. <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4636216&tag=1 >

NEJNOVĚJŠÍ TRENDY VÝZKUMU INTELIGENTNÍCH ŘÍDICÍCH SYSTÉMŮ

Pavel Vrba

Rockwell Automation s.r.o., Výzkumné a vývojové středisko Praha
Pekařská 695/10a, 150 00 Praha 5
E-mail: pvrba@ra.rockwell.com

Abstrakt:

Cílem příspěvku je seznámit čtenáře s dlouhodobými výzkumnými a vývojovými aktivitami společnosti Rockwell Automation v oblasti distribuovaných inteligentních řídicích systémů. Ty nabízejí mnohem flexibilnější a robustnější řešení pro systémy s vysokou mírou složitosti a dynamicky se měnícími podmínkami, pro něž klasické centralizované přístupy často selhávají. Řídicí systém je v tomto pojetí tvořen soustavou distribuovaných autonomních jednotek, které obsahují jak nízkoúrovňovou vrstvu řízení pro přímou interakci s ovládaným procesem v reálném čase, tak vrstvu inteligentního rozhodování, zodpovědnou za optimalizaci, diagnostiku a koordinaci spolupráce s ostatními jednotkami. V posledních letech dochází k významnému posunu ve využití technologií původně navržených pro sémantický web, které nabízejí pokročilé způsoby reprezentace a výměny znalostí v řídicím systému. Spolu s masivním nástupem technologií pro tvorbu uživatelských rozhraní a aplikací běžících v internetovém prohlížeči se otevírají nové perspektivy pro efektivní řešení průmyslové automatizace budoucnosti.

Klíčová slova: automatizace, distribuované inteligentní řízení, multi-agentní systémy, sémantika, ontologie, webové aplikace, HTML5

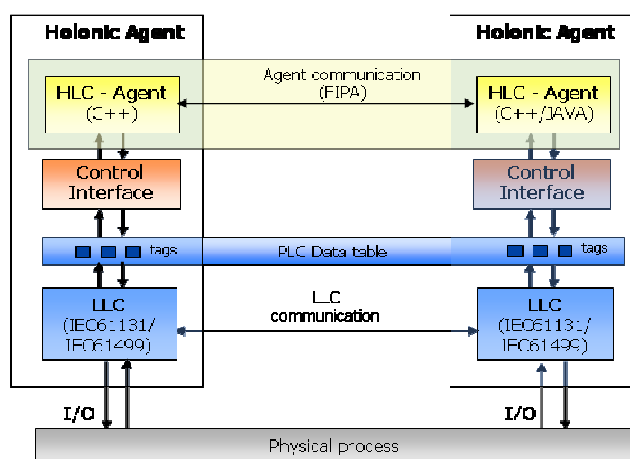
1 Úvod

Již po více než dvě desetiletí se veřejné i soukromé výzkumné instituce snaží prosadit nové metody v průmyslové automatizaci s cílem poskytnout mnohem efektivnější, robustnější a flexibilnější řešení. Obecně řečeno, hlavními charakteristikami těchto moderních přístupů jsou *distribuovanost*, *kooperace*, a *intelligence*. Řídicí systém již není koncipován jak monolitická aplikace s centralizovanými rozhodovacími procesy, ale jako síť distribuovaných samostatných jednotek s vlastní inteligencí, které dokáží koordinovat své aktivity pomocí komunikace tak, aby vzájemně směřovaly ke splnění globálně definovaných cílů [1].

Tento trend započal v rámci aktivit kolem konsorcia Holonic Manufacturing Systems (HMS) reinkarnací pojmu *holon*, který reprezentuje inteligentní, autonomní a kooperativní modul jakožto základní stavební jednotku výrobního systému [4]. Inspirací jsou biologické nebo sociální systémy, ve kterých se každá jednotka organizace skládá z dalších jednotek (např. buňka z jádra a plazmy) a přitom současně představuje část nějakého většího celku (buňky tvoří svalovou tkáň). Další vlastností je autonomnost, která dává jednotce jistou míru nezávislosti tak, aby dokázala zvládat běžné i nepředvídané situace samostatně bez žádosti o pomoc od vyšších autorit. Současně je ale jednotka ovládána z vyšších vrstev tak, aby přispívala ke správné funkci celku. Ukazuje se, že takovéto hybridní soustavy, nazývané *holarchie*, představují velmi stabilní formy, které jsou odolné vůči poruchám a dovedou se snáze adaptovat na nové podmínky.

2 Holonické a multi-agentní řídicí systémy

Uvažujeme-li o aplikaci holonických systémů v průmyslovém řízení, je stále nutné klást důraz na zajištění odezvy řídicího systému v reálném čase. Proto byla navržena architektura tzv. holonického agenta, znázorněná na obrázku 1, kde je kromě bloku HLC (High Level Control), který zajišťuje inteligentní a kooperativní chování, též subsystém nízkourovňového řízení (LLC – Low Level Control) zodpovědný za odezvu v reálném čase [6]. V rámci již zmiňovaného konsorcia HMS bylo věnováno značné úsilí návrhu nového standardu pro implementaci nízkourovňové vrstvy řízení, který by reflektoval distribuovaný a modulární charakter nové generace řídicích aplikací. Výsledkem je standard IEC 61499 známý též jako *funkční bloky* (angl. function blocks), který je mnoha experty považován za slibného nástupce standardu IEC 61131 jazyků pro programovatelné řídicí automaty PLC (Programmable Logic Controllers). Přes mnohé výhody nenašel tento nový standard prozatím uplatnění v běžné průmyslové praxi a jeho aplikace se stále omezuje pouze na laboratorní podmínky [2]. Proto je pro modul LLC stále využíván některý z jazyků standardu IEC61131, jako např. žebříčková logika. Modul LLC je tak vlastně standardním řídicím podprogramem běžícím v PLC v klasické smyčce, kdy PLC nejprve načte vstupní hodnoty ze senzorů v řízeném procesu prostřednictvím analogových nebo digitálních vstupně-výstupních karet, pak řídicí kód provede příslušnou akci, což zahrnuje výpočet výstupních hodnot, které jsou následně propagovány k akčním členům v řízeném systému.



Obrázek 1. Architektura holonického agenta obsahující blok nízkourovňového řízení (LLC), běžící na PLC, agentový modul (HLC) a komunikační rozhraní (Control Interface)

Návrh modulů vysokoúrovňového řízení HLC je vesměs založen na aplikaci teorie multi-agentních systémů. V nich jsou rozhodovací a řídicí procesy distribuovány v komunitě autonomních, vzájemně kooperujících jednotek – agentů – v ideálním případě bez jakéhokoli centrálního rozhodovacího prvku. Každý agent je specializovaný na určitý typ úloh, které dokáže řešit samostatně s využitím znalostí získávaných ze svého bezprostředního okolí. V případě potřeby však agent může požádat jiného agenta o poskytnutí určité informace či provedení dané akce, která nespadá do jeho kompetence. Tyto interakce probíhají pomocí zasílání krátkých zpráv, nejčastěji v textové podobě. Vzhledem k uplatnění technik distribuované umělé inteligence, jako je inteligentní zpracování znalostí, vyjednávání a koordinace, organizace, strojové učení, apod., je nutné pro implementaci modulů HLC použít některý z vyšších programovacích jazyků jako je Java nebo C++. To s sebou přináší určité potíže v souvislosti s výběrem vhodné cílové platformy pro běh agentů. V laboratorních podmínkách jde téměř vždy o řešení využívající osobní počítač. V průmyslovém nasazení je tato varianta nerealizovatelná, proto se hledají možné alternativy. Jednou z nich je prostředí ACS (Autonomous Cooperative System), vyvinuté společností Rockwell Automation, které

umožňuje díky modifikovanému firmwaru provozovat agenty naprogramované v jazyce C++ přímo na PLC řady ControlLogix, a to na stejném procesoru, kde jsou vykonávány i podprogramy nízkourovňového řízení. Další variantou, která je v současné době testována, je využití specializovaného modulu, který se zasunuje do šasi PLC, který je vybaven operačním systémem Linux s možností běhu programů (včetně agentů) implementovaných jak v jazyce C++ tak v jazyce Java, a to nezávisle na hlavním procesoru PLC.

Důležitou komponentou navržené architektury je rozhraní (blok Control Interface na obr. 1), které umožňuje obousměrnou komunikaci mezi bloky HLC a LLC [6]. Jedním směrem dostává agent notifikace od nižší vrstvy o důležitých událostech, které vyžadují jeho intervenci, jako je např. diagnostická informace o poruše, zpráva o dokončení výroby určitého výrobku, apod. Opačným směrem jsou rozhodnutí agenta propagována do bloku LLC, typicky jde o novou konfiguraci parametrů výrobního procesu, určenou součinností více agentů jako reakce na vzniklou situaci. Multi-agentní řídicí systém je tak vlastně koncipován jako nadřazená optimalizační vrstva nad klasickým řízením, která poskytuje pokročilé metody zpracování informací, které by bylo obtížné, ne-li nemožné realizovat s programovým vybavením klasického PLC. Celý systém je navržen tak, aby za normálních podmínek pracovalo nízkourovňové řízení v PLC samostatně, bez časté intervence agentní vrstvy, což by mohlo ohrozit zaručení odezvy řídicího systému v reálném čase. K takové intervenci dojde pouze v případě potřeby, kdy je nutné rekonfigurovat výrobní systém (např. naplánovat novou trasu pro výrobky dopravním systémem v případě poruchy dopravního pásu), přeplánovat výrobu (při poruše stroje přesunout naplánované úlohy na alternativní stroje), apod.

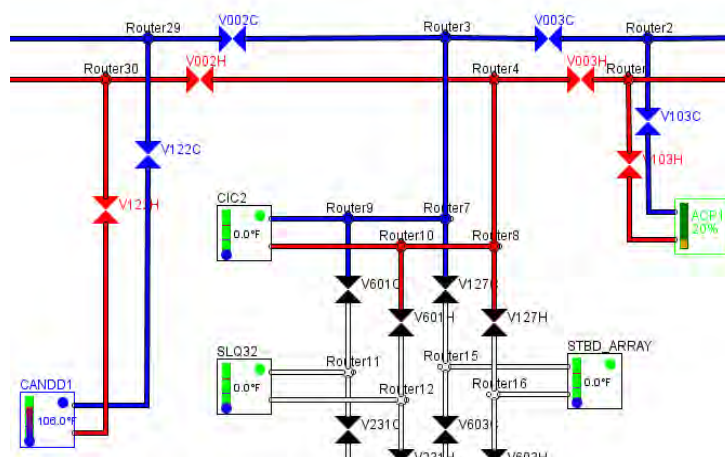
Vzhledem k tomu, že cílovou platformou pro popisovaný řídicí systém je stále klasické PLC, rozhodli jsme se implementovat komunikační rozhraní mezi bloky HLC a LLC tím nejpřirozenějším způsobem. Tím je sdílení dat přímo v paměti řídicího automatu, v tzv. datové tabulce, kde jsou uloženy všechny proměnné, tzv. *Tagy*, se kterými řídicí programy pracují. Pomocí komunikačního rozhraní může agent přímo číst a zapisovat v datové tabulce; kromě toho může dostávat automatická upozornění o změně určité hodnoty.

3 Vybrané aplikace multi-agentních řídicích systémů

Pro ilustraci uvedeme dvě koncepčně podobné aplikace holonických a multi-agentních přístupů k průmyslové automatizaci. Obě aplikace byly vyvinuty v rámci aktivit výzkumného střediska společnosti Rockwell Automation v Praze.

3.1 Chladicí systém lodi

První aplikací vytvořenou ve spolupráci s americkým Úřadem pro námořní výzkum (ONR) je řídicí systém vodního chlazení systémů válečné lodi [6]. Cílem bylo vytvoření robustního řídicího systému, který by byl schopen pracovat i za ztížených podmínek, kdy může být jeho část poškozena nebo zcela zničena. Jako referenční testovací systém posloužil zmenšený model skutečné lodi, který se skládá z dvou chladičů, redundantní soustavy potrubí s třicetidvěma ventily a jedenácti chlazených zařízení, jako je radar, sonar, zbraňový systém, apod. (viz obr. 2). Každá z těchto komponent je reprezentována jedním agentem; celý systém se tak skládá z téměř šedesáti agentů, které jsou distribuované ve dvaadvaceti řídicích automatech typu ControlLogix a FlexLogix, vzájemně propojených pomocí sítí typu ControlNet a Ethernet.



Obrázek 2. Výřez chladicího systému lodi znázorňující výměnu studené (modrá barva) a teplé vody (červená barva) mezi chladičem ACP1 (vpravo nahoře) a chlazeným zařízením CANDD1 (vlevo dole) sítí potrubí a ventilů.

Systém funguje tak, že v okamžiku, kdy agent reprezentující chlazené zařízení zaznamená překročení teploty nad určitou mez (o tom je notifikován z vrstvy nízkourovňového řízení, implementované v žebříčkové logice), začne vyjednávat s dostupnými agenty chladičů o možnosti dodávky studené vody. Agenty chladičů se ve spolupráci s agenty ventilů, které mají informaci o topologii sítě potrubí, snaží nalézt nejkratší cestu od chladiče k chlazenému zařízení (jde o dynamické sekvenční prohledávání grafu). Délka této cesty skombinovaná s vlastní volnou kapacitou pak reprezentuje cenu, za níž chladič nabídne svoji službu chlazenému zařízení. Agent chlazeného zařízení pak vybere nejlepší nabídku a následně požádá chladič o zahájení dodávky studené vody. Ten vyjedná s agenty ventilů jejich otevření či uzavření tak, aby voda proudila po předem dohodnuté cestě (ve skutečnosti je nutné uvažovat dvě cesty potrubím, zvlášť pro studenou a zvlášť pro teplou vodu, vracející se zpět do chladiče).

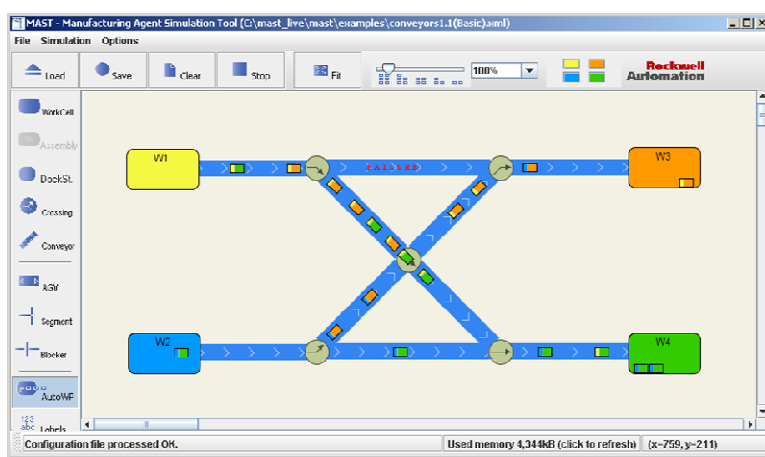
Zmíněné rozdělení na úroveň agentní a řídicí s využitím žebříčkové logiky zde hraje klíčovou roli pro běh v reálném čase. Například při detekci poškození potrubí je nezbytné okamžitě uzavřít přívod vody, aby se zamezilo jejímu dalšímu úniku, což se provádí za pomoci žebříčkové logiky na úrovni řídicí. Následně je pak spuštěna diagnostika na úrovni agentní, jejímž cílem je lokalizovat místo úniku, postupným otevíráním kohoutů (opět se uplatňuje komunikace mezi agenty) a detekcí poklesu tlaku. Po lokalizaci a uzavření kohoutů kolem místa úniku se pak agenty snaží nalézt alternativní cestu chladicím systémem tak, aby pokračovalo chlazení všech předešle chlazených zařízení.

3.2 Systém přepravy materiálu

Další z multi-agentních řídicích aplikací je systém MAST (zkratka angl. Manufacturing Agent Simulation Tool). Původním záměrem bylo vytvořit simulační nástroj, který by přehlednou formou ukazoval hlavní výhody použití principů multi-agentních systémů v automatizaci, tj. odolnost vůči poruchám a schopnosti jednoduché rekonfigurace. Postupně systém MAST dospěl do plnohodnotného nástroje pro návrh, simulaci a reálné řízení výrobních systémů, s hlavním zaměřením na přepravu výrobků v rámci výrobního závodu [6].

Hlavní myšlenkou je umožnit dynamické plánování trasy výrobku sítí dopravních pásů mezi jednotlivými výrobními jednotami, jako jsou stroje, skladovací jednotky, překládací stanice, apod. Systém je opět navržen tak, že každá komponenta je reprezentována jedním agentem, jenž je vytvořen z obecné předlohy (v objektově orientovaném pojetí jako instance třídy) reprezentující daný typ řízené komponenty. Základní množina typů agentů zahrnuje

výrobní jednotku, dopravní pás, a výhybku. Výhybka rozděljuje tok výrobků z jednoho ze svých vstupních pásů na dva či více svých výstupních pásů, popř. spojuje tok výrobků ze dvou pásů na jeden (viz obr. 3). Naplánování optimálních tras systémem proběhne při inicializaci agentů, které se vzájemně informují o tom, kdo je na koho napojen. Opět se uplatní sekvenční dynamické šíření informace o dostupnosti cílových výrobních jednotek celou dopravní sítí. Ve výsledku je tak každá výhybka vybavena znalostí, které všechny cílové jednotky jsou dostupné po každém ze svých výstupních pásů a také za jakou cenu (součet délek pásů do cíle). Tuto znalost, tzv. směrovací tabulku, předá agent výhybky své komponentně nízkourovňového řízení, která podle ní pak řídí v reálném čase směrování výrobků. V okamžiku, kdy dojde k detekci poruchy (např. zastavení nebo zablokování dopravního pásu), iniciují agenti přepočítání optimálních tras systémem stejným algoritmem jako ve fázi inicializace.

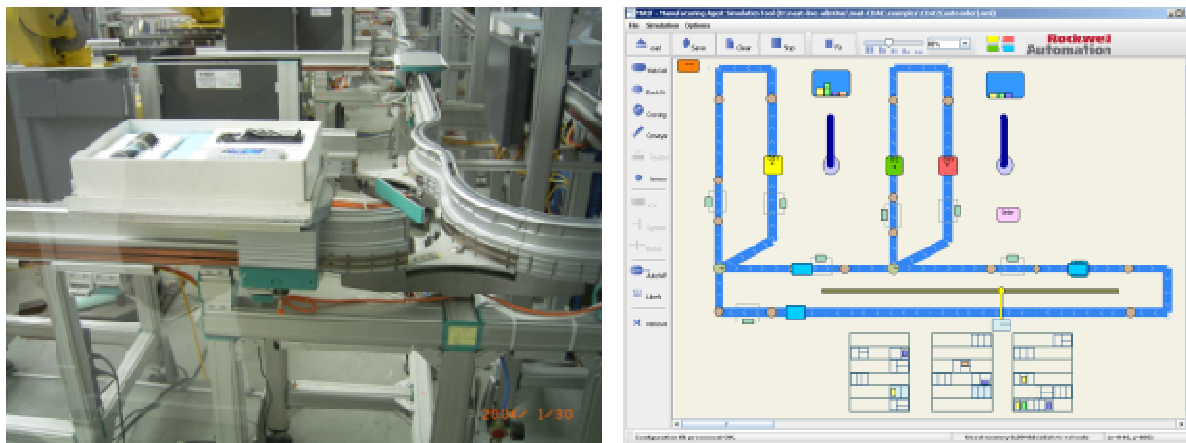


Obrázek 3. Multi-agentní řídicí systém MAST pro přepravu materiálu ve výrobním závodě mezi výrobními komponentami redundantní sítě dopravních pásů a výhybek

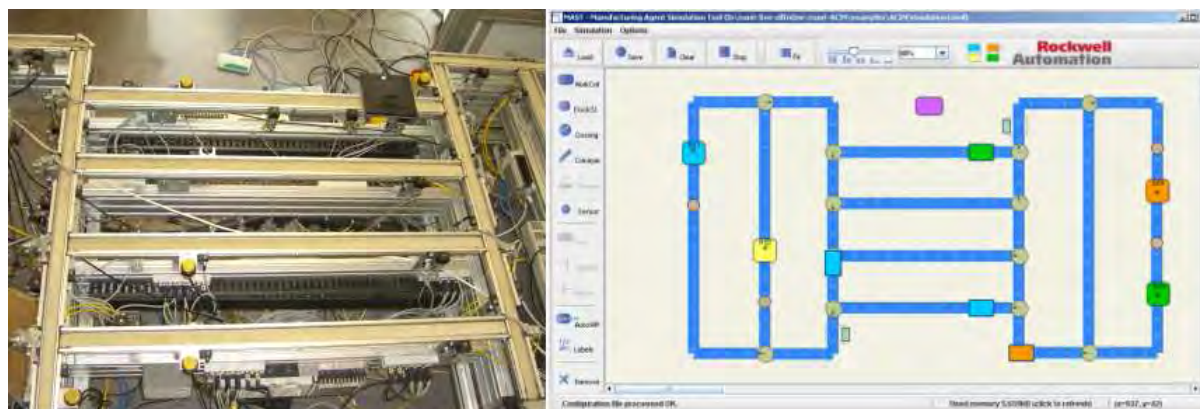
Schopnosti systému MAST byly otestovány na dvou reálných aplikacích. První z nich je prototyp balicí linky, zkonstruovaný v laboratoři DIAL (Distributed Information and Automation Laboratory) na Univerzitě v Cambridge [3]. Linka realizuje balení dárkových balíčků, kdy každý balíček může obsahovat libovolnou kombinaci tří ze čtyř dostupných holicích propriet (gel, deodorant, pěnu a žiletku). Vlastní vkládání do balíčku je zajišťováno dvěma nezávislými pracovními jednotkami, kdy každá z nich obsahuje robota a meziklady holicích potřeb. Doprava mezi pracovními jednotkami a hlavním policovým skladem je zajišťována dopravníkovým systémem Montech (viz obr. 4). Systém MAST byl použit pro simulaci operací balicí linky, jako distribuované řešení, alternativní k předchozímu centralizovanému řídicímu systému. MAST byl rozšířen o řadu nových typů agentů, reprezentujících komponenty balicí linky. Navíc byl navržen typ agenta, reprezentující jeden konkrétní kus výrobku (v tomto případě dárkového balíčku), který sám řídí svoji vlastní výrobu. Tento agent, vytvořený pro každý výrobek zvlášť, proaktivně vyjednává vložení požadovaných holicích potřeb s agenty pracovních stanic.

Další reálnou aplikací, kde byl systém MAST otestován, je prototypová výrobní linka v institutu ACIN (Automation and Control Institute) na Technické univerzitě ve Vídni [6]. Zde byl systém MAST aplikován na reálné řízení paletového dopravního systému skládajícího se ze čtrnácti dopravních pásů, dvanácti výhybek, a čtyřech dokovacích stanic (viz obr. 5). Dopravní systém přepravuje palety s výrobky mezi dokovacími stanicemi, kde je vždy paleta pozastavena, aby mohla být provedena daná operace s výrobkem. Všechna vstupně-výstupní data jsou přenášena sítí DeviceNet do PLC řady ControlLogix, kde jsou též sdílena pomocí dříve popsaného rozhraní s agentní řídicí vrstvou. V podmínkách reálného systému byl

úspěšně otestován scénář poruchy dopravního pásu, kdy agenti našli alternativní trasu, která se vyhýbala místu s poruchou.



Obrázek 4. Fotografie balicí linky v laboratoři DIAL na Univerzitě v Cambridge (vlevo); simulace stejné linky v systému MAST



Obrázek 5. Fotografie paletového dopravního systému v laboratoři ACIN na Technické univerzitě ve Vídni (vlevo); řízení stejné linky v systému MAST

4 Využití sémantiky a nových technologií pro vizualizaci

Během několika posledních let pozorujeme pronikání nových trendů z oblasti internetových technologií, konkrétně sémantického webu (Semantic web), servisně orientovaných architektur (Service Oriented Architecture) a webových uživatelských rozhraní. Sémantické technologie, konkrétně tzv. *ontologie*, pomáhají vytvořit obecný explicitní popis struktury informací vztahujících se k výrobnímu systému a to v podobě kategorií objektů (stroj, operace, výrobek, ...) a jejich vazeb (stroj provádí operaci na výrobku). Inteligentní jednotky ve vrstvě HLC pak drží informace o aktuálním stavu výrobního systému (stroj S_{32} vykonává operaci O_{15} na výrobku V_6) v sémantické podobě v tzv. *znalostních bázích*. Pomocí ontologie jsou také popsány výrobní postupy pro jednotlivé typy výrobků v podobě sekvencí výrobních operací s příslušnými parametry. Hlavní výhodou tohoto přístupu je, že chování agentů není fixně naprogramováno na konkrétní typy výrobků nebo na konkrétní uspořádání výrobního systému. Agenty jsou schopny realizovat jakýkoliv, předem neznámý výrobní postup tím, že automaticky procházejí jeho sémantický popis, vyhledávají operace, které je nutné provést a dynamicky plánují jejich přiřazení k dostupným strojům [6]. Souvisejícím trendem je uplatnění servisně orientovaných architektur a sémantických webových služeb, které poskytují

moderní nástroj pro implementaci systémů založených na distribuovaných, vzájemně spolupracujících jednotkách [5].

Další technologií, která začíná hrát v průmyslové automatizaci velmi podstatnou roli, je návrh uživatelských aplikací a zejména grafických uživatelských rozhraní jako aplikací webových, tj. provozovaných ve webovém prohlížeči. Hlavním benefitem tohoto přístupu je přenositelnost mezi různými hardwarovými zařízeními (osobní počítač, tablet, chytrý mobilní telefon, operátorský panel, ...) a operačními systémy (MS Windows, Linux, Mac OS, Android, iOS, ...) a také to, že již není nutné distribuovat, instalovat a spravovat software na cílových uživatelských zařízeních. Místo toho poslouží standardní webový prohlížeč, jako tzv. *tenký klient*, dostupný dnes na všech cílových platformách.

Mnoho vývojářských týmů nyní zaměřuje pozornost na nové standardy HTML5¹ a CSS3², které přinášejí zajímavé grafické možnosti jako je canvas (plátno pro kreslení ve webovém prohlížeči), SVG (škálovatelná vektorová grafika), přehrávání videa (video tag), nebo animace. Přesto, že standard HTML5 stále ještě nezískal statut finální verze, je již podporován, alespoň do určité míry, ve všech hlavních webových prohlížečích. I když nebudeme uvažovat nový HTML 5 standard, již nyní jsou dostupné nástroje pro tvorbu interaktivních uživatelských rozhraní s moderním vzhledem. Je to především AJAX³ (Asynchronous JavaScript and XML), společně s kaskádovými styly CSS (Cascade Style Sheet), které umožňují asynchronní výměnu dat mezi klientem (webový prohlížeč) a serverem a dynamické překreslení pouze korespondující části webové stránky bez nutnosti jejího kompletního znovunačtení. Vývoj takových aplikací byl významně usnadněn, když společnost Google uvolnila zdarma vývojové prostředí GWT (Google Web Toolkit)⁴. Jeho velkou výhodou je, že vývojář nemusí mít hluboké znalosti o HTML, JavaScriptu nebo CSS pro to, aby mohl navrhovat a ladit AJAX aplikace. Vývojář používá pouze jazyk Java pro vývoj serverové i klientské části aplikace. Kompilátor GWT zajistí převod klientské části aplikace z jazyka Java do optimalizované podoby Javascriptu, která je vnořena do HTML stránky, vygenerované automaticky společně s kaskádovým stylem. Vývojář též dostane do ruky snadno použitelný mechanismus komunikace mezi klientem a serverem, který umožňuje z klienta asynchronní volání metod na serveru. Součástí GWT je též bohatá knihovna grafických prvků (jako tlačítek, textových políček, tabulek, apod.), ze kterých je možné vytvářet interaktivní uživatelská rozhraní, kde je vždy překreslena pouze část webové stránky typicky jako reakce na příjem dat ze serveru.

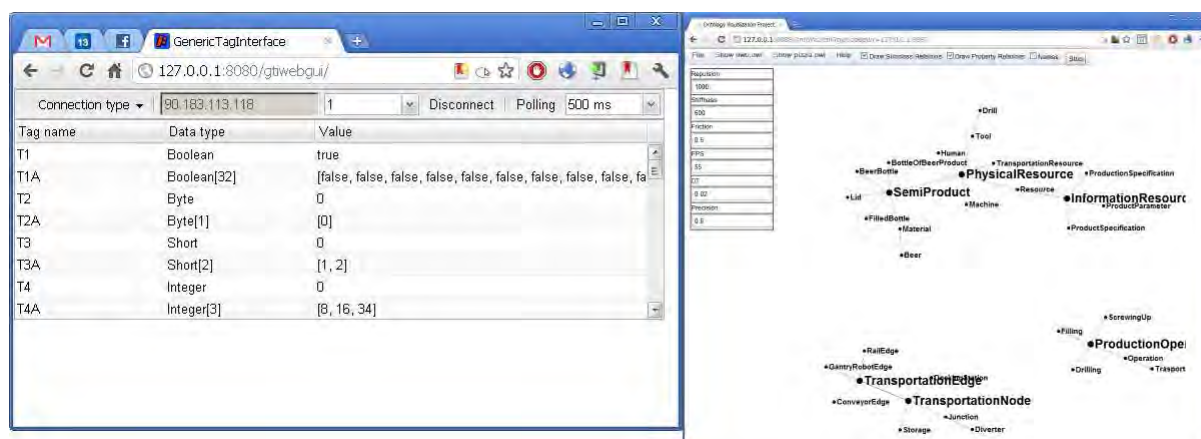
Prostřednictvím výše zmíněných technologií jsme implementovali dvě pilotní aplikace. První z nich slouží k zobrazování hodnot tagů PLC ve webovém prohlížeči (viz levá část obr. 6). Druhá z aplikací, vyvinutá v rámci našich prvních experimentů s HTML5, zobrazuje ontologie v pobobě samoorganizujících se grafů (viz pravá část obr. 6). Aplikace je integrována s multi-agentním řídicím systémem MAST, popsáným v předchozí kapitole, a umožňuje tak zobrazovat v reálném čase ontologie a znalostní báze jednotlivých agentů. Obě výše zmíněné aplikace byly úspěšně otestovány na mobilních zařízeních, konkrétně na tabletu iPad a telefonech HTC Desire a Nokia E52 (v tomto případě pouze první z nich).

¹ <http://www.w3.org/TR/html5/>

² <http://www.w3.org/TR/CSS/>

³ [http://en.wikipedia.org/wiki/Ajax_\(programming\)](http://en.wikipedia.org/wiki/Ajax_(programming))

⁴ code.google.com/webtoolkit/



Obrázek 6. Aplikace pro zobrazení hodnot tagů řídicího automatu PLC ve webovém prohlížeči (vlevo); aplikace využívající HTML5 pro zobrazení ontologií ve webovém prohlížeči

5 Závěr

Záměrem příspěvku bylo ukázat některé z moderních přístupů v průmyslové automatizaci. Přesto, že technologie holonických a agentových systémů stále ještě nenašly širší uplatnění v reálných řídicích aplikacích, filozofie distribuovaných, inteligentních a kooperujících entit je v poslední době reinkarnována v podobě službově orientovaných architektur. Ty společně s novými internetovými technologiemi, jako je HTML5 a cloudové služby, představují do budoucna perspektivní směr vývoje automatizačních řešení. Vizí je plné nahrazení lokálně instalovaných uživatelských aplikací řešením založeným na přístupu ke cloudovým službám prostřednictvím webového prohlížeče. Ten se stává obecným klientem pro renderování uživatelských rozhraní s perspektivou využití chytrých mobilních zařízení jako jsou tablety či telefony pro novou generaci operátorských panelů.

Reference

- [1] S. Bussmann, N. R. Jennings, and M. Wooldridge, Multiagent systems for manufacturing control: A design methodology, Springer Verlag, Berlin-Heidelberg, 2004.
- [2] Fletcher, M., Brennan, R.W.: Designing a Holonic Control System with IEC 61499 Function Blocks. Int. Conf. Intelligent Modeling and Control, Las Vegas (2001)
- [3] Fletcher, M., McFarlane, D., Lucas, A., Brusey, J., Jarvis, J., 2003. The Cambridge Packing Cell - A Holonic Enterprise Demonstrator. In: *Multi-Agent Systems and Applications III* (Mařík, V., Müller, J., Pěchouček, M. (Ed.)), LNAI 2691, Springer Verlag, Berlin, Heidelberg 533-543.
- [4] Höpf, M.: Holonic Manufacturing Systems (HMS) – The Basic Concept and a Report of IMS Test Case 5. In: Knudsen, J.K.H. et al. (Eds): *Sharing CIM Solutions*, IOS Press (1994)
- [5] Martinez Lastra, J.L., Delamer, I.M.: Semantic Web Services in Factory Automation: Fundamental Insights and Research Roadmap. *IEEE Trans. Industrial Informatics*, vol. 2, no. 1, pp. 1--11 (2006)
- [6] Vrba, P., Tichý, P., Mařík, V., Hall, K., Staron, R., Maturana, F., Kadera, P. Rockwell Automation's Holonic and Multi-agent Control Systems Compendium. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, Vol. 41, Issue 1, 2011, pp. 14-30.

Řídicí technika

NĚKTERÉ APLIKACE ŘÍDICÍHO SYSTÉMU REX

Pavel BALDA, Miloš SCHLEGEL

Katedra kybernetiky, Fakulta aplikovaných věd, Západočeská univerzita v Plzni
Univerzitní 8, 301 00 Plzeň

E-mail: pbalda@kky.zcu.cz, schlegel@kky.zcu.cz

Abstrakt:

Řídicí systém REX je systematicky vyvíjen již více než 10 let. Tento příspěvek nepopisuje architekturu řídicího systému REX, ale demonstruje jeho možnosti na některých aplikacích ze dvou oblastí. První oblastí je řízení laboratorních mechatronických modelů a robotů realizovaných na katedře kybernetiky Fakulty aplikovaných věd Západočeské univerzity v Plzni. Jako příklad je uveden lanový robot pro řízení jevištní techniky, tzv. dvojitý SCARA robot s paralelní architekturou a model „robotického lachtana“ balancující míč na jehle. Druhou oblastí je řízení speciálních technologií (s velkým poměrem výzkumu, vývoje a potřeby simulací), prototypů a testovacích systémů (stendů). Jako příklad je uveden řídicí systém výkonu jaderného reaktoru LR0 v ÚJV Řež a řídicí systém teplotního šokového stendu pro testování automobilových komponent vyvinutý pro koncern ZF.

Klíčová slova: řídicí systém, vestavěný řídicí systém, Simulink, mechatronický model, paralelní robot, operační systém reálného času, průmyslový Ethernet.

1 Úvod

Řídicí systém REX (viz např. [3]) je softwarový systém vhodný pro řízení jak laboratorních úloh, prototypů a testovacích systémů (stendů), tak i strojů, mechatronických systémů, speciálních robotů a technologických procesů. Přestože lze REX považovat za tzv. Soft PLC, nemůže konkurovat a ani nekonkuruje programovatelným automatům velkých firem jako jsou např. Siemens, Rockwell, ABB, B&R, Beckhoff, apod. Pro některé typy aplikací, zejména inovačních aplikací s velkým podílem výzkumu, vývoje a simulací pokročilých řídicích algoritmů je však vhodnější než standardní řešení dodávaná velkými firmami.

Důvodem je jeho kompatibilita se simulačním systémem Matlab-Simulink pomocí rozsáhlé knihovny funkčních bloků (např. [10], [6]), která umožňuje v průběhu vývoje řídicího systému provádět průběžné simulace, aniž by bylo třeba přecházet mezi různými vývojovými platformami. Jinými slovy, pokud uživatel vlastní licenci systému Matlab-Simulink, může vyvíjet vlastní algoritmus řídicího systému přímo v grafickém editoru Simulinku. Poznamenejme, že systém REX obsahuje vlastní konfigurační program RexDraw, který umožňuje plnohodnotnou konfiguraci systému i bez Simulinku.

Za více než 10 let vývoje řídicího systému REX vzniklo více než 50 prací týkajících se architektury tohoto řídicího systému, pokročilých algoritmů zabudovaných do systému REX, podpory standardu OPC [8], ovladačů, převodů do různých operačních systémů a na různé hardwarové platformy, implementace pokročilých algoritmů řídicího systému REX do systémů třetích stran, apod. Přehled (neúplný) těchto prací a jednotlivých platform systémů REX je uveden v rozsáhlém článku [2].

Cílem tohoto příspěvku je na příkladech konkrétních realizovaných aplikací naznačit

možnosti řídicího systému REX. V sekci 2 jsou uvedeny příklady řídicích systémů několika laboratorních modelů z pracoviště autorů, v sekci 3 jsou pak zmíněny průmyslové aplikace, pracující od okamžiku nasazení v trvalém provozu.

2 Ukázky řízení laboratorních modelů

Řídicí systém REX je v laboratoři automatického řízení na katedře kybernetiky používán pro řízení všech modelů a jen v této laboratoři je nainstalován více než patnáctkrát. Studenti a doktorandi, kteří se podílejí na vývoji řídicích algoritmů, se naučí pracovat se systémem REX během několika hodin, protože všichni procházejí kurzem systému Simulink a přechod na REX je především záležitostí změny knihovny funkčních bloků (blockset).

Následující podkapitoly informují o některých modelech vyvinutých a oživených za poslední dva roky.

2.1 Lanový robot se šesti stupni volnosti

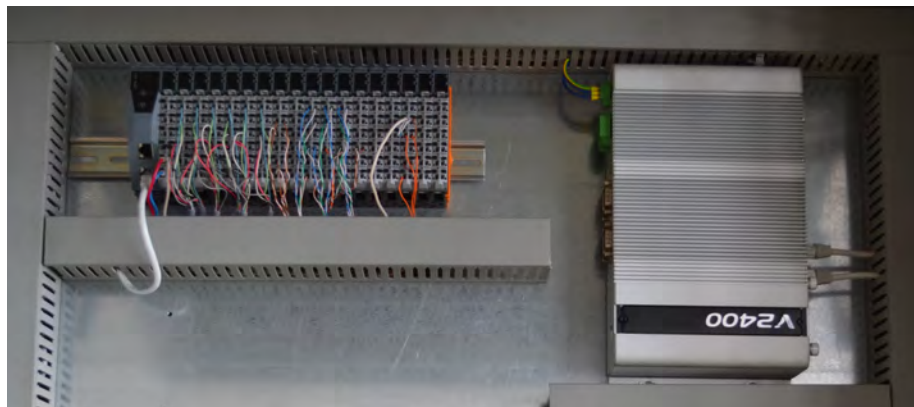
Konstrukce paralelního robotu se šesti stupni volnosti [1] je patrná z obr. 1. Robot se skládá ze čtvercového rámu o rozměrech cca 3x3 m připevněnému ke stropu laboratoře (na obr. 1 je patrná jen část tohoto rámu). Rám je rozdělen na dvě poloviny, v každé z nich jsou umístěny tři posuvné příčky a na každé z těchto příček je na jednom konci příčky vodící kladka a ve zvoleném bodě přestavitelná koncová kladka. Koncové kladky s výstupními vodícími očky určují body, ze kterých vycházejí lanka zakončená na kovové platformě (na obr. 1 je vidět šestiúhelníková platforma v popředí). Druhý konec lanek je veden přes soustavu kladek na motoricky řízené navijáky, umístěné na zadní stěně laboratoře.



Obrázek 1: Pohled na platformu lanového robotu

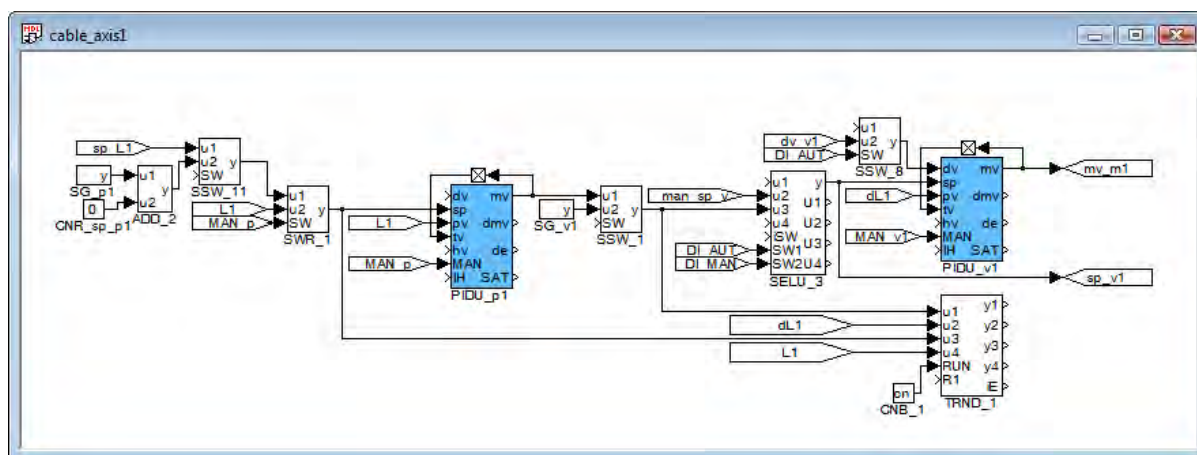
Hardware řídicího systému je umístěn v rozvaděči pod navijáky, viz detailní snímek na obr. 2. V levé části jsou vstupně výstupní moduly systému X20 firmy B&R, v pravé pak řídicí počítač MOXA V2400 (v provedení bez jakýchkoliv pohyblivých částí) s procesorem

Intel Atom. Řídicí systém REX je provozován na tomto průmyslovém PC v operačním systému Linux a jeho rozšíření reálného času Xenomai a komunikuje se vstupně-výstupními moduly prostřednictvím průmyslové komunikace Ethernet POWERLINK [5], která je do systému REX začleněna jako speciální ovladač [11].



Obrázek 2: Řídicí systém lanového robotu

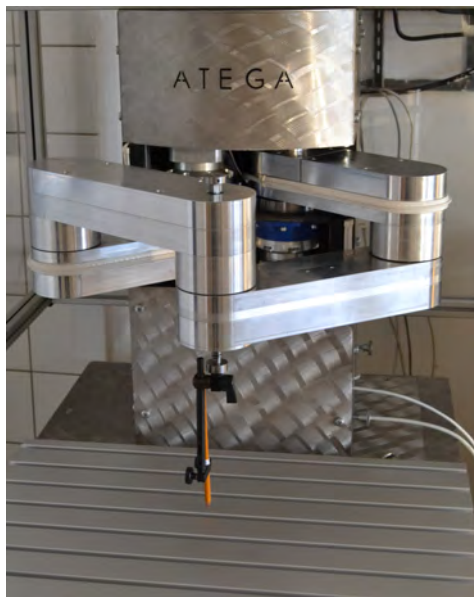
Konfigurace jedné konkrétní úlohy řídicího systému byla vytvořena v programu RexDraw a její část řídicí jeden pohon (naviják) je znázorněna na obr. 3. Jedná se o kaskádní regulaci rychlosti (regulátor PIDU_v1) a polohy (regulátor PIDU_p1). Regulátor proudu (momentu) je uvnitř řídicí jednotky použitého pohonu firmy Maxon. Poznamenejme, že generování trajektorie pro jednotlivé pohony tak, aby koncový efektor (platforma) se pohybovala po požadované trajektorii a s definovanými úhly natočení vzhledem k jednotlivým souřadným osám je složitá nelineární úloha, která přesahuje rámec tohoto příspěvku.



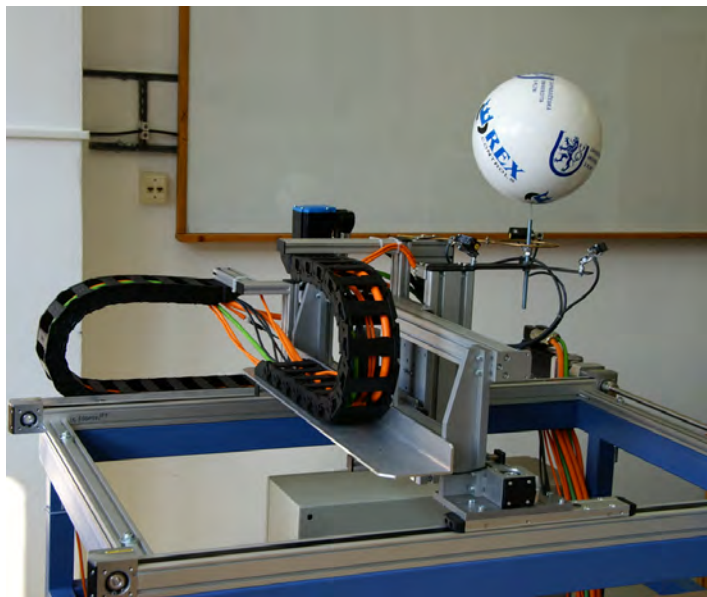
Obrázek 3: Kaskádní regulace jedné osy lanového robotu

2.2 Paralelní robot se čtyřmi stupni volnosti

Robot [1] se skládá ze čtyř vzájemně propojených otočných ramen, z nichž dvě (zadní ramena na obr. 4a) jsou poháněna servomotory s převodovkami, takže mají dva stupně volnosti. Koncový efektor je umístěn na platformě, kterou prochází osa spojení předních dvou ramen (na obr. 4a je tato platforma v popředí a je přizpůsobena pro přidržování tužky). Platforma může navíc vykonávat posuvný vertikální pohyb a rotační pohyb, což



(a)



(b)

Obrázek 4: Paralelní robot se čtyřmi stupni volnosti (a) a „robotický lachtan“ (b)

jsou další dva stupně volnosti. Motory s převodovkami pro tyto dva stupně volnosti jsou umístěny uvnitř zakryté skříně robotu za motory pohybuujícími rameny. Celkem má tedy tento robot čtyři stupně volnosti.

Frekvenční měniče k servopohonům (obojí od firmy Lenze) jsou umístěny v rozvaděčové skříně, tvořící stůl (podstavec) robotu. Pro řízení robotu je použito průmyslové PC MOXA (stejně jako v případě lanového robotu), které však tentokrát komunikuje s frekvenčními měniči pomocí komunikace EtherCAT [4], opět začleněné do systému REX jako speciální ovladač. Řídicí systém REX je i v tomto případě provozován v operačním systému Linux a řídicí algoritmus je opět zkonfigurován v programu RexDraw.

2.3 „Robotický lachtan“

Poslední robotický model znázorněný na obr. 4b je řízen tak, aby balancoval míč na špičce jehly, a je nazýván robotickým lachtanem (lachtan umí totéž na špičce nosu opravdu skvěle). Model vychází z tříosého portálového robotu umožňující pohyb vlevo/vpravo, dopředu/dozadu a nahoru/dolů. Na svislé ose je připevněna ocelová jehla, která během balancování podepírá míč v nestabilní rovnovážné poloze. Řízení tohoto robotu je opravdu náročné [7], balancování v obou vodorovných osách od sebe nejde oddělit, zjednodušeně řečeno, pokud by probíhalo balancování pouze v jedné horizontální ose, v druhé by míč mezitím spadl a naopak. Vzhledem k relativně malému poloměru je pro řízení míče použit průmyslový počítač Advantech vybavený zásuvnými kartami. Na tomto počítači je provozován řídicí systém REX v operačním systému reálného času ETS (dříve Phar Lap ETS) s periodou 1 ms.

3 Ukázky řízení průmyslových aplikací

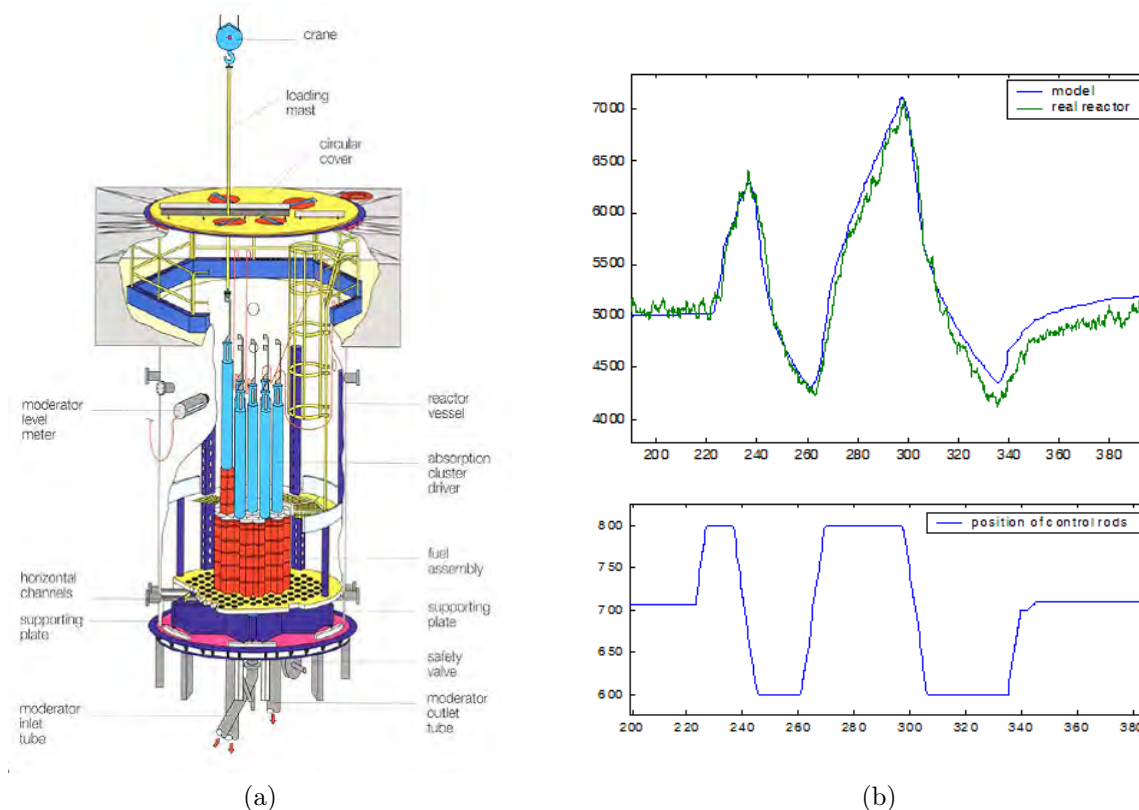
Jak již bylo řečeno v úvodu, řídicí systém REX lze použít nejen k řízení laboratorních modelů, ale i pro výzkum a vývoj řídicích systémů speciálních zařízení nebo prototypů, kde je třeba iteračně kombinovat simulaci s vlastním řízením. V této sekci jsou uvedeny

dva takové příklady.

3.1 Regulátor výkonu jaderného reaktoru LR0

V roce 2008 byl v Ústavu jaderného výzkumu Řež a.s. modernizován řídicí systém jaderného reaktoru LR0 (viz schéma na obr. 5a), kde byl pro regulaci výkonu použit řídicí systém REX [9]. Důvodem pro použití systému REX byla potřeba simulace jaderného reaktoru včetně řídicího systému v prostředí Matlab-Simulink i vlastní řízení skutečného reaktoru. Oba cíle byly splněny konfigurací algoritmů v řídicím systému REX, ve kterém lze snadno přecházet mezi simulací v Simulinku a vlastním řízením, aniž by musel být řídicí algoritmus komplikovaně ručně převáděn do odlišného prostředí cílového řídicího systému.

Výsledný řídicí systém je provozován na průmyslovém počítači firmy NexCom v operačním systému ETS a komunikuje s nadřazeným ochranným systémem reaktoru pomocí komunikace Modbus TCP/IP jako slave. Komunikační protokol Modbus je opět zařazen do řídicího systému REX ve formě speciálního ovladače. Aplikace obsahuje téměř 200 funkčních bloků a je provozována s periodou 50 ms.



Obrázek 5: Schéma reaktoru LR0 (a) a dosažená shoda experimentu s modelem (b)

Obr. 5b nahoře ukazuje velmi dobrou shodu mezi chováním simulačního modelu a skutečného reaktoru při změnách požadovaného výkonu, obr. 5b dole pak pohyb regulačních tyčí reaktoru během regulace.

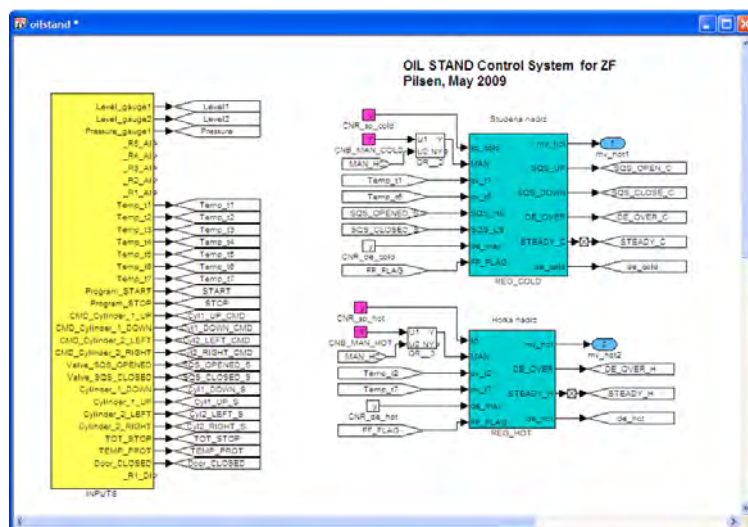
3.2 Regulace teplotního šokového stendu

Poslední z aplikací uvedených v tomto příspěvku je řídicí systém teplotního šokového stendu pro testování komponent v automobilovém průmyslu vyvinutý pro firmu ZF Engi-

neering Plzeň s.r.o. Testovací systém znázorněný na obr. 6a obsahuje dva koše naplněné automobilovými komponentami, které se střídavě ponořují do dvou nádob naplněných olejem, jedna je chlazená a druhá ohřívána. Systém umožňuje provádět zadané množství testovacích cyklů a do nadřazeného systému LabVIEW, sloužícího pro vizualizaci, exportovat experimentální data.



(a)



(b)

Obrázek 6: Teplotní šokový stend (a) a část řídicího schématu s hlavními subsystémy (b)

Řídicí systém je provozován v operačním systému Windows CE v. 5.0 na stanici WinPAC-8000 s periodou 15 ms a obsahuje 470 funkčních bloků. Část hlavní řídicí úlohy s několika subsystémy je znázorněna na obr. 6b.

4 Závěr

Príspevek se snažil stručně představit několik laboratorních i průmyslových aplikací řídicího systému REX, a tím demonstrovat jeho praktické možnosti. Ve všech zmíněných případech se jedná o netriviální řídicí systémy, jejichž realizace vyžaduje velkým podíl výzkumné a vývojové práce. Právě pro takové typy aplikací je řídicí systém REX vhodným a velmi efektivním nástrojem.

Demonstrační verze řídicího systému REX je volně dostupná na internetové adrese www.rexcontrols.cz, kde lze nalézt i další užitečné informace.

Poděkování

Príspevek vznikl díky částečné podpoře projektů Ministerstva průmyslu a obchodu FR-TI1/077 a FR-TI1/394 a projektu Fondu rozvoje vysokých škol 1051/2010. Děkuje.

Reference

- [1] Balda, Pavel. Robotická a mechatronická laboratoř. Závěrečná zpráva projektu FRVŠ 1051/2010. Plzeň : Západočeská univerzita v Plzni, 2010. s. 1-8.
- [2] Balda, Pavel; Schlegel, Miloš. REX - Nástroj pro budování laboratorních i průmyslových řídicích systémů jednotným způsobem. In Sborník přednášek z technické kon-

- ference: Automatizace, regulace a procesy. Praha : DIMART s.r.o., 2009. s. 35-52. ISBN: 978-80-903844-3-9.
- [3] Balda, Pavel; Schlegel, Miloš; Štetina, Milan. Advanced control algorithms + Simulink compatibility + Real-time OS = REX. Proceedings of the 16th IFAC world congress, Oxford : Elsevier, 2006. ISBN: 0-08-045108-X.
- [4] EtherCAT Technology Group: www.ethercat.org
- [5] Ethernet POWERLINK V2.0 Communication Profile Specification. EPSG (Ethernet Powerlink Standardisation Group), 2006.
- [6] Funkční bloky systému REX – Referenční příručka. Vydání 8. REX Controls s.r.o., Plzeň, 2010.
- [7] Gouběj, Martin; Schlegel, Miloš; Škarda, Radek. Mechatronic models for education – Robotic sea lion. Submitted to the 18th IFAC World Congress, 2011.
- [8] OPC Foundation. Data Access Custom Interface Standard Version 3.00. March 4, 2003.
- [9] Schlegel, Miloš; Balda, Pavel. Power Controller of Nuclear Reactor. In Process Control 2008. Pardubice : University of Pardubice, 2008, s. 1-7. ISBN: 978-80-7395-077-4.
- [10] Schlegel, Miloš; Balda, Pavel; Štetina, Milan. Knihovna C MEX bloků pro průmyslovou regulaci s aplikačními příklady. Konference Matlab 2001, Humusoft s.r.o., Praha.
- [11] Štetina, Milan; Balda, Pavel; Pišl, Roman; Gouběj, Martin; Schlegel, Miloš. Motion control algorithms in the new version of the REX control system. In Process Control 2010. Brno : Tribun EU, 2010. s. 1-9. ISBN: 978-80-7399-951-3.

REKUPERACE ENERGIE MOBILNÍCH ROBOTŮ

František BURIAN, Luděk ŽALUD

Centrum aplikované kybernetiky, Vysoké Učení Technické v Brně

Kolejní 2906/4, 612 00 Brno

E-mail: burianf@feec.vutbr.cz, zalud@feec.vutbr.cz

Abstrakt:

V současné době mnoho pracovišť na světě provádí výzkum problematiky úspor při provozu mobilních strojů. Jedna z možností, jak snížit energetické nároky stroje je účinně a efektivně řídit cyklus energetických toků mezi baterií a pohonnými jednotkami, pokud možno s minimálními ztrátami. V článku je rozebrána jedna z variant, kterou je možné potlačit hlavní nevýhodu baterií, která brání efektivní rekuperaci energie.

Klíčová slova: baterie, rekuperace, tok energie

1 Úvod

Rekuperace energie je dávno známý problém, který je řešen již od počátku, kdy začalo hojně používání elektrických motorů v průmyslu. Tento fenomén je umožněn vlastnostmi motoru, který se v okamžiku brzdění může chovat jako zdroj elektrické energie v režimu brzdění. Bez rekuperace je plná energie transformována na neúčinné teplo v brzdových destičkách.

Současná řešení rekuperace, známá zejména u hybridních automobilů mají velmi nízkou účinnost, kdy z rekuperace je možné získat zpět kolem 20 % elektrické energie, kterou vozidlo v pohybu má. Tato účinnost ještě závisí na rychlosti jízdy vozidla, nejúčinnější je při brzdění ve vyšších rychlostech.

V tomto článku se budeme zabývat zejména použitím u mobilních robotů, kde je kladen velký důraz na velmi nízkou energetickou náročnost a zároveň nehluký provoz ve všech režimech jízdy.

2 Robotický systém Orpheus

Robotický systém Orpheus slouží pro průzkum a prohledávání oblastí, které jsou lidskému životu nepřístupné nebo nebezpečné. Oblastí po útoku teroristy, s chemickým, biologickým či radiačním zamořením, nebo po živelní události s nestabilním podložím. Rozměry a maximální hmotnost robotu byly voleny s ohledem na prostupnost terénem, a na možnost osazení speciálními senzory sloužícími k průzkumu.

V současnosti je energetický subsystém mobilního robotu sestaven z jedné baterie LiFePo₄, která dodává energii pro pohon podvozku, pro bezdrátovou komunikaci a pro senzorické vybavení. Do budoucna se počítá s rozšířením na dvě baterie, což umožní lépe reagovat na nutnost rekuperace, a sníží se tím ztráty a zvýší dojezd.

Robot [6] [7] má hmotnost 60 kg, a dosahuje maximální rychlosti 10 km/h. Pro další výpočty budeme předpokládat jízdu průměrnou rychlostí $v = 5$ km/h. Baterii obsahuje o napětí 24V a kapacitě 12 Ah

3 Energie uchovaná v baterii

Pro další výpočty je potřebné znát energii uchovanou v jedné baterii. Výpočet je zřejmý přes nominální napětí na baterii a nominální kapacitu baterie. Vše dokládá rovnice 1

$$E_{bat} = UIt = 24 \times 12 \times 3600 = 1,04 MJ \quad (1)$$

4 Energie potřebná pro chod systémů

Z baterie robot odebírá trvalý klidový proud 1,5A. Pokud převedeme tento proud na úbytek energie za pomoci vztahu 2, získáme kontinuální úbytek energie z baterie na sekundu provozu robotu.

$$\frac{dE_0}{dt} = UI = 24 \times 1,5 = 36 J/s \quad (2)$$

Pokud vztah 2 dáme do poměru se vztahem 1, lze vyvodit, že stojící robot na místě může být provozován po dobu, uvedenou ve vztahu 3, tedy po dobu o málo větší než 8 hodin.

$$t = \frac{E_{bat}}{\frac{dE_0}{dt}} = \frac{1,04 \cdot 10^6}{36} = 28888s = 8h \quad (3)$$

5 Energie uchovaná v rychlosti pohybu robotu

Vypočítanou energii, která je k dispozici pro přeměnu z mechanické setrvačnosti na elektrochemickou uvnitř baterie jsem použil vzorec 4 Pro robot o hmotnosti $m = 60$ kg, jedoucí průměrnou rychlostí $v = 5$ km/h (1,4 m/s) lze zjistit, že robot má využitelnou kinetickou energii 59 Js.

$$E_{mech} = \frac{1}{2}mv^2 = \frac{1}{2}60 \times 1,4^2 = 58,8 J \quad (4)$$

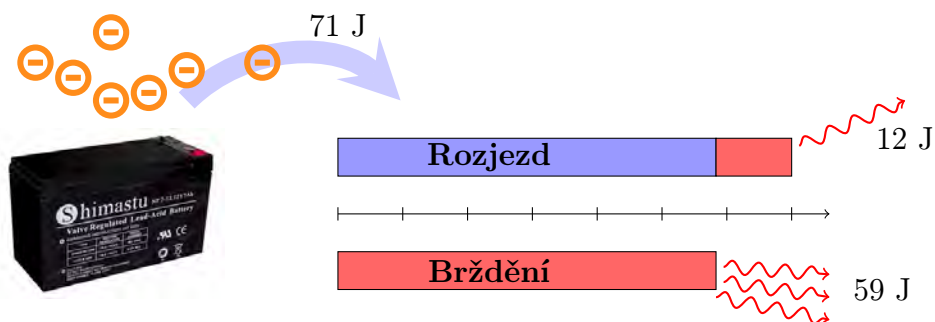
Tato energie je však pouze energie, kterou má jedoucí robot. Pro výpočet kolik energie spotřeboval robot tím, že vytvořil tuto rychlost, nebo naopak kolik energie můžeme očekávat že se vrátí zpět do baterie, musíme použít účinnost vlastního motoru s převodem. Pro standardní elektricky poháněná soukolí se uvádí účinnost 80 %. Pokud bude robot na tuto rychlost zrychlovat, vydá energii E_A , naopak při zpomalování až do zastavení robot musí absorbovat energii E_B . Obě veličiny je možné matematicky vyjádřit vztahem:

$$\begin{aligned} E_A &= (1 + (1 - 0,8)) \times E_{mech} = 1,2 \times 58,8 = 70,5 J \\ E_B &= (0,8) \times E_{mech} = 0,8 \times 58,8 = 47 J \end{aligned} \quad (5)$$

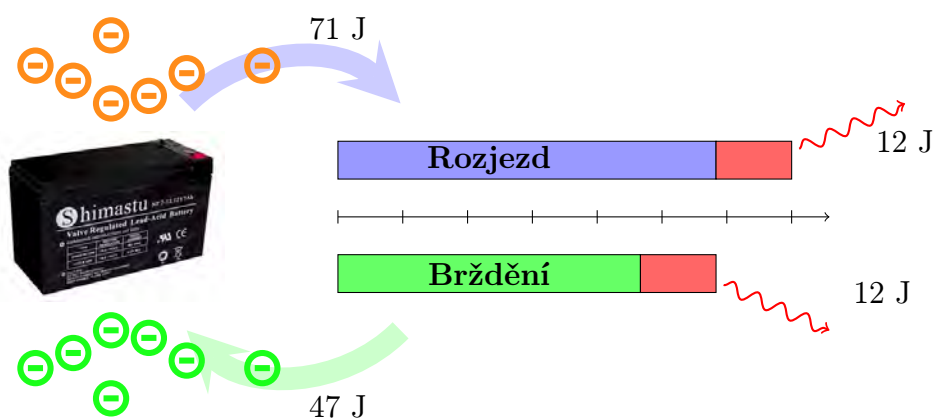
Zbývajících 20 procent energie je v obou případech nutno vyzářit ve formě tepelné energie v převodovce a motoru.

Z rovnice 5 plyne, že při úplném zabrzdění a maximální účinnosti rekuperace dokážeme získat zpět energii rovnou 47 J, což odpovídá 67 % úspora energie. Tento princip je názorně srovnán na obrázcích 5 a 5. Na obrázku 5 je vidět v grafu energetická bilance v

okamžiku, pokud brzdíme klasickou kotoučovou brzdou, na obrázku 5 je vidět možnost brzdění motorem a navrácení energie zpět do baterie. Červeně zvýrazněné plochy v grafu odpovídají tepelným ztrátám.



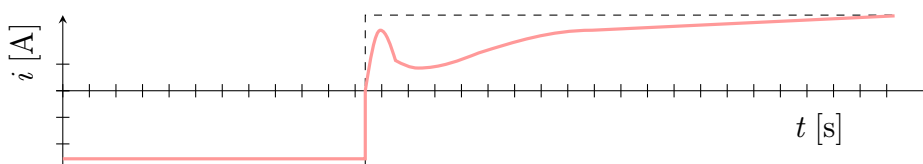
Obrázek 1: Energetická bilance bez rekuperace



Obrázek 2: Energetická bilance s rekuperací

6 Problém dynamiky baterie

Mezi hlavní problémy rychlé rekuperace kinetické energie na elektrickou nepatří jen účinnost vlastního motoru a měniče, ale je zde též problém s rychlostí elektrochemické reakce. Baterii, která dodává proud do motoru není možné jednoduchým způsobem přimět k rychlé reverzaci proudu. Chování baterie popisuje obrázek 6. [1] [2] [3]



Obrázek 3: Reverzace proudu baterie

Z grafu je patrné, že při pokusu o reverzaci proudu z baterie (čárkovaná čára) je skutečný proud (plná červená čára) reverzován jen částečně, a objeví se mezera, jejímž

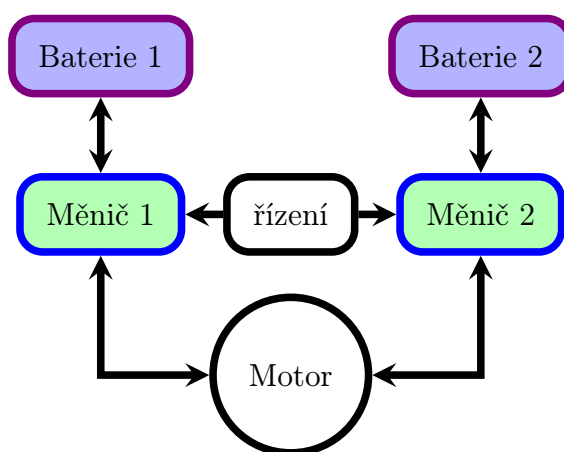
výsledkem je okamžité zahřívání desek baterií, což je nežádoucí. První překmit je chemického rázu, kdy desky, pokryté oxidy z původní reakce která vytvářela elektrodu se v mžiku rozpusí a dokážou tak velmi rychle absorbovat první proudový impuls. Čím nižší je hodnota vnitřního odporu baterie (oxid má tenčí vrstvu), tím menší plochu tento impuls zabírá.

Díky této vlastnosti je třeba navrhnout takové řízení, aby se reverzace proudu baterií děla pokud možno minimálně. Poté budou tepelné ztráty na baterii minimální.

7 Řešení rozdělením baterií

Jednoduché a intuitivní řešení, které se naskýtá, v sobě obsahuje rozdělení hlavní baterie do dvou sekcí. To umožňuje udržovat jednu baterii stále v nabíjecím cyklu a druhou stále ve vybíjecím cyklu. Při potřebě impulsního proudu je možné obě baterie přimět k dodávce výkonu ven z baterie s přihlédnutím k nižší účinnosti v tomto režimu.

Nástin tohoto řešení je možné vidět na obrázku 7



Obrázek 4: Řešení dělením na dvě baterie

Systém řízení pracuje tak, že měnič 1 je nastaven na trvalé dodávání výkonu do motoru, měnič 2 bere energii z motoru pouze v okamžicích, kdy je nutné brzdit, případně zvyšuje výkon měniče 1 v okamžiku, kdy je potřeba vyššího výkonu.

Toto řešení má řadu výhod, zejména je možné jmenovat výhodu modulárnosti a fault-tolerant návrhu, který umožní činnost zařízení i v okamžiku kritického selhání jednoho z řídicích modulů (druhý měnič může plně převzít kontrolu nad motorem).

8 Závěr

V článku popisovaná metoda řízení motoru umožňuje lépe využít kapacitu baterií umítných na mobilním robotu. Díky rozdělení baterií je umožněna efektivnější práce při rekuperaci a zároveň zajištěna činnost při chybě. Díky důrazu na rekuperaci energie je možné snížit energetickou stopu výroby a používání mobilních robotů.

Poděkování

The research was performed in Center for Research and Utilization of Renewable Energy Sources. Authors gratefully acknowledge financial support from European Regional

Development Fund under project No. CZ.1.05/2.1.00/01.0014.

Reference

- [1] Crompton T.R.: Battery Reference Handbook. 752s. 1996. ISBN 978-0750625678
- [2] Bogenschütz A. F.: Technical dictionary for Batteries and Advanced Energy Conversion. 200s. ISBN 978-3870970024
- [3] Wenzl H.: Batterietechnik. 253s. ISBN 978-3816916918
- [4] Wise E.: Applied Robotics, 312s. 1998. ISBN 0-7906-1184-8
- [5] Everett HR.: Sensors for Mobile Robots, Theory and Applications, AK Peters, Ltd., USA, 1995. ISBN: 1-56881-048-2.
- [6] Zalud L.: Universal Autonomous and Telepresence Mobile Robot Navigation. In: 32nd International Symposium on Robotics - ISR 2001I, s 1010-1015, Seoul, Korea. 2001.
- [7] Zalud L.: ORPHEUS - Reconnaissance Teleoperated Robotic System. In: 16th IFAC World Congress. s. 1-6. Praha, Czech Republic. 2005.

Jednotná platforma pro sdílení dat a dokumentů při řešení projektů výzkumnými konsorcií

Filip BURŠÍK, Martin SÝKORA, Martin HENEŠ

Vysoké učení technické v Brně

Fakulta elektrotechniky a komunikačních technologií

Ústav telekomunikací

Purkyňova 118, 612 00 Brno

E-mail: xbursi01@stud.feec.vutbr.cz, xsykor01@stud.feec.vutbr.cz, xhenes@stud.feec.vutbr.cz

Abstrakt:

V článku je popsána realizace systému podporujícího komunikaci mezi výzkumnými pracovišti. Součástí systému je informační systém pro zprávu řešených projektů, dále videokonferenční propojení mezi jednotlivými subjekty konsorcia a systém centralizovaného úložiště jednotlivých součástí projektu s podporou workflow.

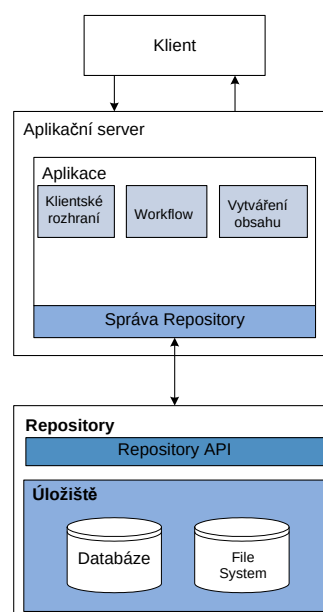
Klíčová slova/Keywords: sdílení dat, workflow, videokonferenční systém, repository

1 Úvod

V případě vzniku výzkumného konsorcia mezi výzkumnými pracovišti, vzniká problém efektivní komunikace. Často jsou tyto subjekty od sebe výrazně vzdáleny a osobní komunikace se stává časově značně náročnou. Komunikace pomocí běžných prostředků (telefon, fax, email apod.) je sice možná, ale nemůže dosahovat úrovně osobní komunikace. Také sdílení částí řešeného projektu je velmi problematické, předávání si části řešení projektu např. elektronickou poštou je sice možné, ale velmi nepřehledné. Neexistují efektivní komunikační prostředky, které by byly optimalizovány pro druh komunikace mezi výzkumnými subjekty sdruženými ve výzkumné konsorcium. Tento problém je z našich dosavadních zkušeností také velmi výrazný a značně komplikuje vědecko výzkumnou činnost vícero subjektů na jednom tématu.

2 Návrh systému

Součástí systému musí být informační systém pro zprávu řešených projektů, dále videokonferenční propojení mezi jednotlivými subjekty konsorcia a systém centralizovaného úložiště jednotlivých součástí projektu s podporou workflow. Další součástí systému by mělo být centralizované úložiště dokumentací k jednotlivým projektům řešeným v rámci konsorcia. Elektronickým dokumentem v tomto pojetí rozumíme nejen texty, ale fotodokumentaci, soubory dat z měřicích přístrojů, vstupní a výstupní soubory technologických souborů, programové produkty, video a audio soubory, účelové databáze a jistě i celá řada dalších, zde nezmíněných dokumentů. Elektronická média ve spojení s komunikačními sítěmi tak poskytují pružnější a rychlejší distribuci dokumentů, kdy se obrací tradiční model. Žadatel o dokument



Obrázek 1: Blokové schéma systému

prohledává datové úložiště a potřebná data si stahuje na lokální počítač. Takový systém podporuje vzdálený přístup k datům a je zajištěno sdílení zdrojů mezi více subjekty a je schopný dělit žadatele o dokument do různých skupin podle přístupových práv. Navíc dokumenty v elektronické formě mohou obsahovat obrazová nebo zvuková data a další formy multimediálních dat, která nelze klasickým způsobem přenášet.

Při stanovování koncepce systému je nejvhodnější zvolit modulární strukturu, která je složena z několika základních částí, jak je znázorněno na obrázku 1. Jedná se zejména o moduly pro vytváření obsahu, workflow, grafické rozhraní pro přístup uživatele k vlastnímu multimediálnímu obsahu a centrální datové úložiště (repository). Funkce hlavních modulů je následující:

2.1 Vytváření obsahu

Modul vytváření obsahu slouží ke vkládání dokumentů k jednotlivým projektům. Příkládat lze soubory libovolného typu. Součástí tohoto modulu je také zpřístupnění videokonferenčních nahrávek. Jak je uvedeno v úvodu, důležitým prvkem komunikace je také videokonferenční propojení jednotlivých pracovišť. V některých případech je výhodné konferenci nahrát a záznam zpřístupnit. Jedná se zejména o různé prezentace nebo přednášky a také nahrávky interního charakteru, kdy bude záznam přístupný pouze omezené skupině uživatelů.

2.2 Workflow

Modul workflow je velice důležitý pro řízení projektů [5]. Zjednodušeně se dá říci, že workflow specifikuje, jak probíhá určitý proces od jeho vzniku do konce. Workflow tvoří logika procesů a řídicí pravidla [1], [2]. Logika procesů definuje sekvenci úkolů, které mají být provedeny a řídicí pravidla pak definují v jakých závislostech a limitech mohou být úkoly vykonávány [3]. Podle standardu CSCW (Computer Supported Cooperative Work) je možno klasifikovat workflow do následujících skupin [4].

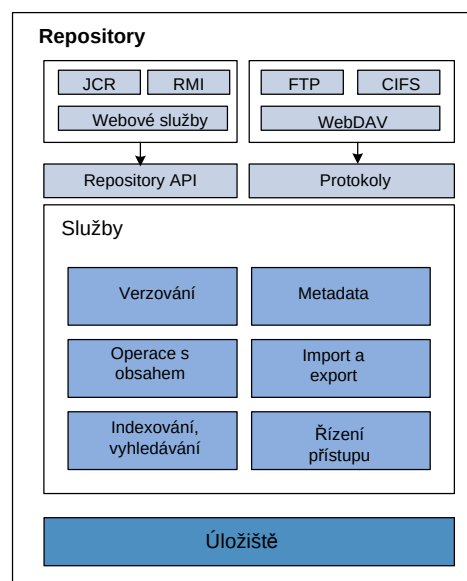
- Administrativní workflow zajišťuje rutinní administrativní činnosti.
- Kolaborativní workflow podporuje týmovou spolupráci, která je klíčová při řešení projektů. Typická je přítomnost nějakého dokumentu, pomocí něhož si účastníci vyměňují své poznatky a který je v podstatě výsledkem jejich společné práce. Většinou tyto procesy obsahují opakující se cyklus několika iterací téhož kroku, dokud se všichni neshodnou. Jedná se například o tvorbu dokumentace, změnu funkčnosti výrobku, či návrh postupu testování. Vždy je jako výstup očekáván dokument, na kterém spolupracuje několik uživatelů, a který prochází různými schvalovacími cykly.
- Produkční workflow
- Ad hoc workflow je založeno na náhodnosti vzniku workflow procesu. Jedná se o procesy, jejichž průběh není předem znám. Nemohou být proto standardizované a proto se definují až při jejich vzniku. Tímto typem workflow se řeší odchylky, výjimky a různé typy unikátních situací.

3 Centrální datové úložiště

Jádrum celého systému by mělo být kvalitně navržené datové úložiště - repository. Repository je datové úložiště, které obsahuje databázi, multimediální soubory (obsah), jejich popisy (metadata), různé druhy přístupu a širokou nabídku služeb.

Repository zajišťuje tyto stěžejní služby:

- verzování,
- správu metadat,
- import a export,
- správa indexů a vyhledávání,
- řízení přístupu.



Obrázek 2 : Blokové schéma repository

3.1 Správa metadat

V metadatach jsou uloženy informace umožňující efektivně spravovat, vyhledávat a publikovat obsah. Metadata lze rozdělit do dvou kategorií.

a) *Metadata určená ke správě* – každá položka obsahu musí mít uložena metadata, která slouží ke správě dat. Mezi tato metadata patří autor obsahu, recenzent, datum modifikace, publikace, expirace, a titul obsahu.

b) *Metadata určená k vyhledávání* – obsah, který je uložen v repository. Pro zajištění snadného hledání musí repository obsahovat vyhledávací mechanismus, který umožní najít relevantní data. Vyhledávání se neprovádí přímo v multimediálních datech, ale v metadatach. U větších objemů dat je zapotřebí mít k dispozici hierarchickou strukturu kategorií (taxonomii), do které lze jednotlivé položky obsahu vkládat. Informace o kategorii, ve které je obsah zařazen je uložena v metadatach, podle kterých interní vyhledávač omezí množinu výsledků.

3.2 Verzování

Verzování představuje jednu ze stěžejních služeb. Pokud uživatel pracuje nad objemným souborem a provádí v něm řadu změn, musí mít k dispozici přístup k původnímu obsahu. Důvodem je přehled nad provedenými změnami a také možnost vracet se o několik kroků zpět v případech, kdy jsou provedeny nevhodné změny. Princip je obdobný jako při editaci klasických textových dokumentů. Pro multimediální záznamy je navíc požadavek, aby multimediální záznam zůstal zachován také v původní podobě, kvůli prokázání autentičnosti. Změny, které se budou provádět, se zapisují pouze do metadat, která budou verzována.

3.3 Import a export

Systém musí umožňovat export a import obsahu. Tyto funkce slouží především k zálohování, či přechodu mezi různými instancemi úložišť stejného typu.

3.4 Přístup k repository

K centrálnímu datovému úložišti je definováno několik typů rozhraní. Všechny rozhraní reprezentují obsah ve formě adresářů a souborů. V některých případech je výhodnější přistupovat do repozitáře přímo do adresářové struktury. Jedná se například o výstupy z měřicích přístrojů, kdy se přístroji zpřístupní adresář z repozitáře a on data ukládá přímo do něj. Data jsou ihned přístupná všem uživatelům zainteresovaných do projektu a také jsou ihned zálohovaná a nehrozí jejich ztráta.

WebDAV – (Web-based Distributed Authoring) je jeden z W3C standardů, který vznikl jako rozšíření http. WebDAV umožňuje vytvářet, upravovat, přesouvat obsah na vzdáleném serveru.

FTP – (File Transfer Protocol) je jednoduchý protokol sloužící k výměně souborů v rámci TCP/IP sítě, kde je komunikace založena na systému klient-server.

CIFS – (Common Internet File System) je protokol, který definuje vzdálený přístup k souborům mezi platformami [6].

4 Standardizovaný přístup k repository – JCR

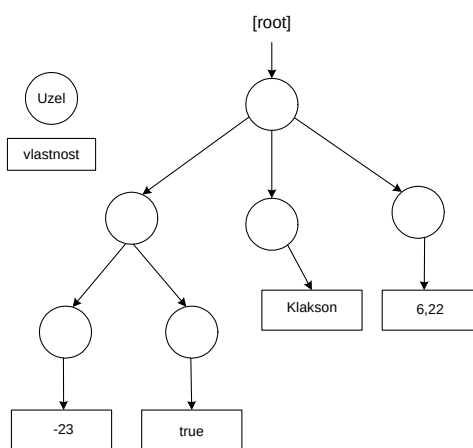
Content Repository for Java Technology API (JCR) je specifikace standardního aplikačního rozhraní pro přístup k systémům pro správu dokumentů. Tento standard byl přijat v roce 2005 jako JSR 170 [7] s cílem sjednotit přístup k různým implementacím systému pro správu dokumentů.

API definuje základní požadavky na vlastnosti datového úložiště a metody pro přístup k datům uloženým v informačním systému pro správu dokumentů. Data jsou uložena v tzv. uzlech, které jsou uspořádány do hierarchické struktury a každý uzel má sadu vlastností jejichž struktura je uvedena na obrázku 3.

Úložiště může být rozděleno do více nezávislých struktur - workspace a podpora jmenných prostorů zabráňuje kolizím jmen a umožňuje odkazovat na uzly v jiných workspace. Standard definuje podporu pro vyhledávání dat, verzování dokumentů, zajištění přístupových práv, dotazy pomocí XPath syntaxe nebo transakce.

XPath je vyhledávací jazyk, který byl původně navržen k vybírání elementů z XML dokumentu. XPath je vhodný k hledání v JCR obsahu, protože repository je tvořeno stromem uzlů a vlastností, který je analogický stromu XML dokumentu. XPath dotazovací výrazy jsou prováděny nad XML dokumentovým pohledem. Provedený dotaz vrátí tabulku obsahující jména a obsah vlastností, které splnily hledaný výraz.

Standard JCR obsahuje dvě úrovně funkcionality, podle kterých se určuje kompatibilita systému s JCR. Systém bude splňovat obě úrovně standardu JCR.



Obrázek 3 : Hierarchická struktura uzlů a vlastností

Na I. úrovni je prováděno:

- Procházení a získávání uzlů a vlastností.
- Čtení hodnot.
- Jmenný prostor - jejich účelem je minimalizace kolize názvů.
- Export do XML.
- Dotazy pomocí XPath syntaxe.
- Zjištění definovaných typů uzlů.

Na II. úrovni je prováděno:

- Přidávání a odstraňování uzlů vlastností.
- Zápis hodnot vlastností.
- Import z XML.
- Přiřazování typů uzlům.

5 Vyhledávání v systému

Systém ukládá metadata v databázi a vlastní obsah v souborovém systému. V souborovém systému jsou také ukládány indexy s klíčovými slovy fulltextového vyhledávání realizovaného knihovnou Lucene (<http://lucene.apache.org/>). Přístup k databázi je realizován pomocí objektově-relačního mapování. Tato konfigurace může v dané kombinaci využívat výhod obou řešení (viz. Tabulka 1), stejně tak jsou oba nástroje plně konfigurovatelné. Díky použití specializovaných nástrojů není také nutné řešit nízkoúrovňové přístupy až již k souborovému systému, nebo databázi.

Tabulka 1 : Výhody jednotlivých řešení

<i>Výhody využití knihovny Lucene</i>	<i>Výhody databázového řešení</i>
– Výkonnější než databázové řešení	– Vyhledávání bývá součástí DB systému
– Plně konfigurovatelné	– Jednoduše použitelné
– Možnost použití pro více doménových modelů	– Podpora více národních jazyků

5.1 Hodnocení efektivity vyhledávání

Mezi zásadní požadavky patří rychlost zpracování požadavků, jistý uživatelský komfort pro zadávání požadavků (interakce se systémem), způsob zadávání dotazů a způsob výstupu výsledků tohoto vyhledávání. Pro posouzení efektivity vyhledávacího systému se používají dva koeficienty udávající, jak výsledky vyhledávání odpovídají požadavkům tazatele (poskytnutí relevantních dokumentů), koeficient přesnosti a koeficient úplnosti.

Koeficient přesnosti představuje pravděpodobnost, že vybrané dokumenty jsou relativní, koeficient úplnosti představuje pravděpodobnost, že byl vybrán relevantní dokument.

$$P = \frac{|A \cap B|}{|B|}$$

$$R = \frac{|A \cap B|}{|A|} \quad (1)$$

Kde P vyjadřuje koeficient přesnosti (Precision), R odpovídá koeficientu úplnosti (Recall), B vyjadřuje celkový počet relevantních dokumentů a A celkový počet všech vybraných dokumentů. Ideální případ by nastal, kdyby oba koeficienty byly rovny 1, ovšem tohoto případu není možné v praxi dosáhnout. Uživatel, který po prvním dotazu dostane příliš mnoho nerelevantních dokumentů, začne dotaz zpřesňovat (zmenšuje se míra úplnosti a zvětšuje míra přesnosti), díky čemuž dosáhne vyšší přesnosti na výstupu, ovšem za cenu zmenšení absolutního počtu relevantních dokumentů.

Pro zjednodušení informací o efektivitě byly zavedeny metody zobrazující přesnost a úplnost do jednodimenzionálního prostoru, jednou z nich je Van Risjbergenova F -míra.

$$F_{\beta} = 1 - \frac{1 + \beta^2}{\frac{\beta^2}{R} + \frac{1}{P}} = \frac{(1 + \beta^2)|A \cap B|}{\beta^2|A| + |B|} \quad (2)$$

Kde β vyjadřuje poměrovou důležitost mezi přesností a úplností.

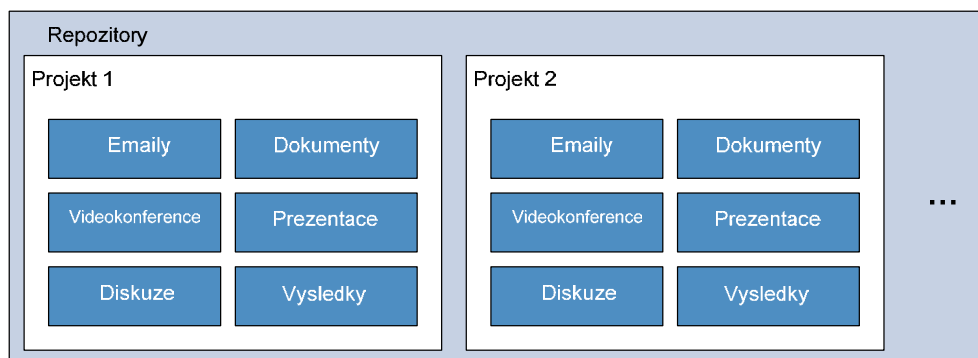
6 Realizovaný systém

Na základě výše uvedené specifikace byl navržen systém, který vyhovuje potřebám výzkumných konsorcií. Především podporuje komunikaci jednotlivých subjektů na více úrovních, aby sdílení, publikování a získávání znalostí a výsledků bylo co nejjednodušší a zároveň bezpečné.

Jelikož jsou základní potřeby každého projektu stejné, byla navržena obecná struktura projektu. Mezi základní části, jak je zobrazeno na obrázku 4, patří :

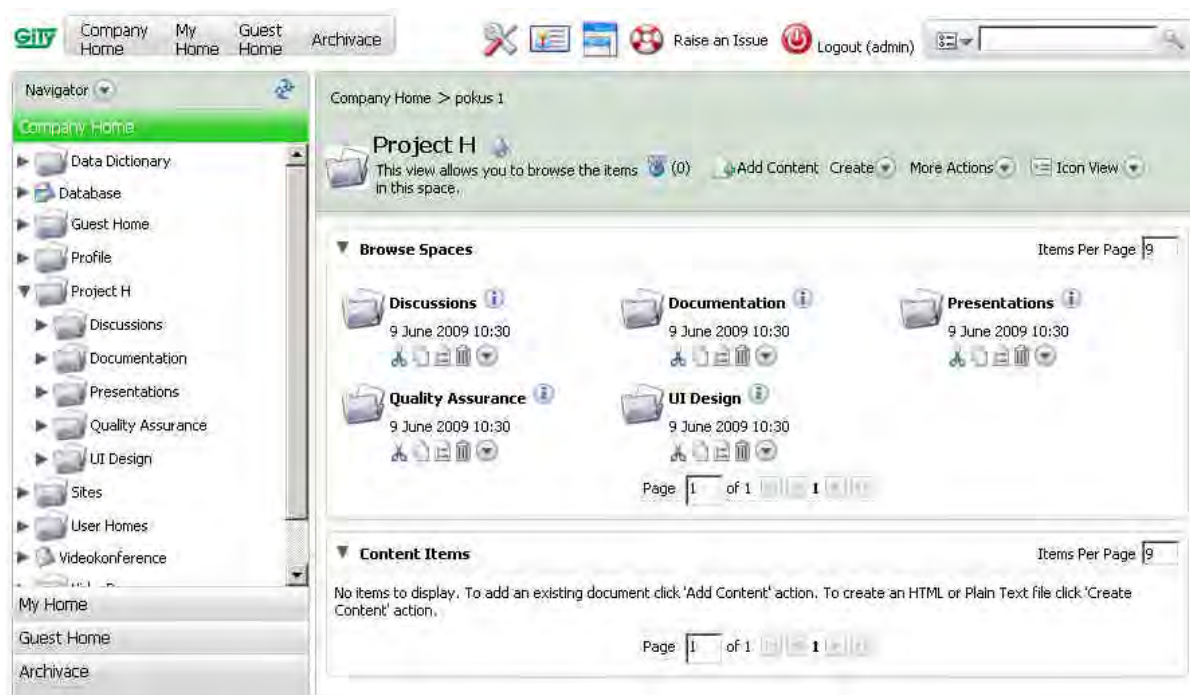
- Emaily
- Dokumenty
- Videokonference
- Prezentace
- Diskuze
- Výsledky

Dokumenty, prezentace a výsledky jsou založené na kooperativním workflow, aby uživatelé měli možnost nad těmito dokumenty diskutovat a upravovat je po vzájemné shodě. Oproti tomu emaily a diskuze jsou prosté prostory, kde se archivuje komunikace mezi členy týmu, aby již vyřešené problémy nebyly zapomenuty. Poslední částí projektu jsou archivovaná videokonferenční sezení. Jejich archivací je získána obsáhlá databáze problémů, které jsou vyřešené a které teprve čekají na vyřešení. V této sekci se nearchivují pouze samotné videozáznamy, ale i související data, jako jsou prezentace či jiné dokumenty související se záznamem. Záznamy a příložené soubory by měli být přístupny skrze kalendář, kde budou uloženy podle data vzniku.



Obrázek 4 : Struktura projektů v systému

Jelikož každý projekt má svá specifika, musí systém také umožňovat v rámci projektu definovat nové prostory s novými funkcionalitami. Na následujícím obrázku je náhled navrženého systému.



Obrázek 5 : Náhled vytvořeného systému

7 Závěr

V případě vzniku výzkumného konsorcia mezi výzkumnými pracovišti, vzniká problém efektivní komunikace. Často jsou tyto subjekty od sebe výrazně vzdáleny a osobní komunikace se stává časově značně náročnou. Z tohoto důvodu bylo vhodné navrhnout a realizovat systému pro podporu výzkumných konsorcií. Součástí systému je informační systém pro zprávu řešených projektů, dále videokonferenční propojení mezi jednotlivými subjekty konsorcia a systém centralizovaného úložiště jednotlivých součástí projektu s podporou workflow. Systém podporuje vzdálený přístup k datům a je zajištěno sdílení zdrojů mezi více subjekty a je schopný dělit žadatele o dokument do různých skupin podle

přístupových práv. Navíc dokumenty v elektronické formě mohou obsahovat obrazová nebo zvuková data a další formy multimediálních dat, která nelze klasickým způsobem přenášet.

Reference

- [1] A. Cichocki et. al., Workflow and process automation: Concepts and Technology (Kluwer academic Publisher, Dordrecht, 1998)
- [2] D. Hollingswort, Terminology & Glossary (Workflow Management Coalition,1999, WPMC-TC-1011)
- [3] L. Fischer, 2007 BPM & Workflow Handbook (Future Strategies Inc., 2007)
- [4] Reference Model for an Open Archival Information System (OAIS) (Consultative Committee for Space Data Systems, 2002) [online]
<http://public.ccsds.org/publications/archive/650x0b1.pdf>
- [5] S. W. Sadiq, M. E. Orlowska, W. Sadiq, Specification and validation of process constraints for flexible workflows, Information Systems, vol.30, 2005, 349 - 378
- [6] P.J. Leach, D.C. Naik, A Common Internet File System (CIFS/1.0) Protocol (Microsoft Documentation, 1998), [online]
<ftp://ftp.microsoft.com/developr/drg/CIFS/>
- [7] D. Nuescheler, JSR 170: Content Repository for Java™ technology API (Day Software AG, 2005), [online] <http://www.day.com/specs/jcr/1.0/>
- [8] E. Hatcher, O. Gospodnetić, M. McCandless, Lucene in Action, Second Edition (Manning Publications Co., 2008)
- [9] Apache Lucene – Overview (The Apache Software Foundation, 2009), [online]
<http://lucene.apache.org/java/docs/>
- [10] C. J. Rijsbergen, Information Retrieval – Second Edition (Butterworth, 1979).

REALIZACE GRAVÍROVACÍHO ZAŘÍZENÍ

Pavel ČÍP, Tomáš BABINEC

Centrum aplikované kybernetiky, Vysoké Učení Technické v Brně
Ústav automatizace a měřicí techniky, Vysoké Učení Technické v Brně
Kolejní 2906/4, 612 00 Brno

E-mail: xcippa00@stud.feec.vutbr.cz, xbabin01@stud.feec.vutbr.cz

Abstrakt:

Gravírování je odstraňování vrstvy z povrchu materiálu za účelem realizace grafiky (symbolu, nápisu apod.). Slouží k nesmazatelnému značení výrobků či reklamním účelům. Gravírování lze provést mechanicky do kovu či kamene pomocí fréz nebo vodního paprsku, nebo vypálením pomocí laseru do dřeva. V našem případě jsme z bezpečnostních důvodů prozatím realizovali gravírovací zařízení na bázi plotteru – tedy kreslení grafiky na X-Y stolku. V budoucnu se předpokládá výměna fixy za laser a realizace laserového gravírování do dřevěné překližky či balzy.

Klíčová slova: gravírování, SVG, zpracování obrazu

1 Úvod

Podle technického slovníku znamená gravírování totéž jako rytí či odstraňování vrstvy z povrchů tvrdého materiálu (kov, dřevo, kámen, sklo apod.) za účelem zobrazit na něm grafiku, symbol, informační hodnotu. V průmyslu se strojové gravírování velmi rozšířilo v posledních letech zejména s rozvojem nových technologií. Dříve bylo možné gravírovat pouze pomocí tvrdšího materiálu (rydla, hrotu) do měkčího (dřeva, kamene). Vznikaly tak náhrobní kameny, dřevěné formy, štítky na dveře apod.

Dnes je však známo mnohem více nových moderních přístupů. Grafiku je možné realizovat vodním paprskem do kamene či oceli, laserem do kovu, dřeva apod. Vzhledem k současné technice je možné vytvářet gravírované předměty mnohem rychleji a přesněji. Své uplatnění si najdou nejen v průmyslu, ale i při výrobě reklamních předmětů.

1.1 Automatické gravírovací stroje

Automatické gravírovací stroje, obvykle realizovány pomocí CNC strojů, jsou určeny pro mechanické gravírování, řízené počítačem a dosahují vysokého stupně automatizace. Gravírování je možné realizovat pomocí frézky či diamantového hrotu velmi přesně a rychle, za velmi příznivé ceny. Posun frézovací hlavy lze realizovat ve třech osách, přičemž hloubka záběru určuje i vzhled výsledné grafiky. Například při použití skládané překližky o více vrstvách, lze takto docílit zajímavého efektu.

1.2 Návrh zařízení

V našem případě se však oddělíme od mechanického rytí a přejdeme k návrhu gravírování grafiky pomocí laserového paprsku. Návrh našeho zařízení vychází zejména z již realizovaného X-Y stolku, často nazývaného také jako plotter. Ten je ovládán dvěma krokovými motory POWERMAX II s velmi jemným krokem. Převod točivého momentu na posuv je realizován pomocí šroubového pohybového mechanismu a pomocí kluzného nebo valivého uložení. Tím je dosaženo jemného krocování. Předpokládáme, že laserový paprsek, lze zaostřit na desítky mikrometrů a tímto jemným posuvem, lze posléze realizovat i velmi detailní grafiku.

Plotter je vytvořen ze čtvercových ocelových profilů a nese na sobě rovnoběžné vodící lišty, které zajišťují pohyb plotterové desky v ose X, která je kolmá na osu Y. Tato deska má ve svém povrchu řadu vodících drážek, které slouží k upnutí gravírovaného materiálu (obrázek 1).

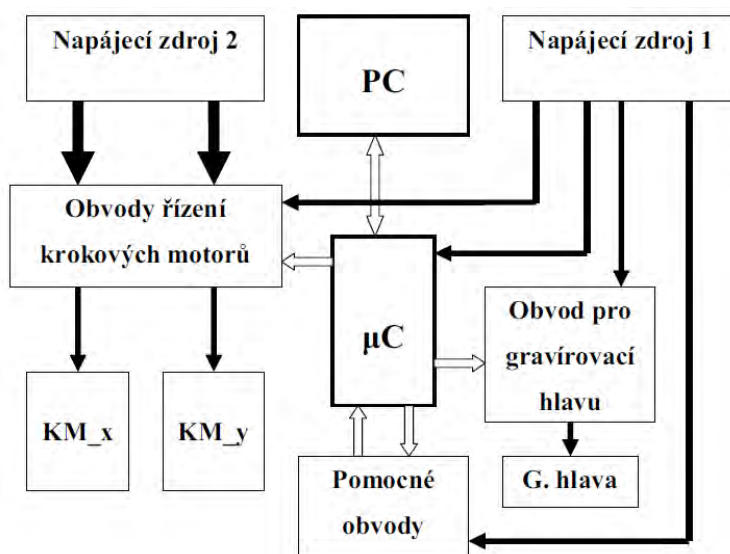


Obrázek 1: Navržené řešení X-Y stolku s dvěma krokovými motory [2]

Grafika je v současné chvíli realizována fixem na papír. Příčiny této dočasné náhrady budou ještě dále vysvětleny. V počítači je jeden obrazový bod nazýván pixelem – ten je bezrozměrný. Při jeho realizaci na papír (či jiný materiál) dochází k jeho transformaci na pixel fyzický, který již má rozměr daný minimální rozlišovací schopností interpreta (tedy hrotem fixu, šířkou paprsku laseru, hrotem frézky apod.).

1.3 Ovládací jednotka

Celé zařízení ovládající gravírování si lze názorně ukázat na obrázku 2. Je složen zejména z mikroprocesoru AT89C51ED2, řídicí stanice PC, napájecích zdrojů, budících obvodů pro motory a dalších pomocných prvků.



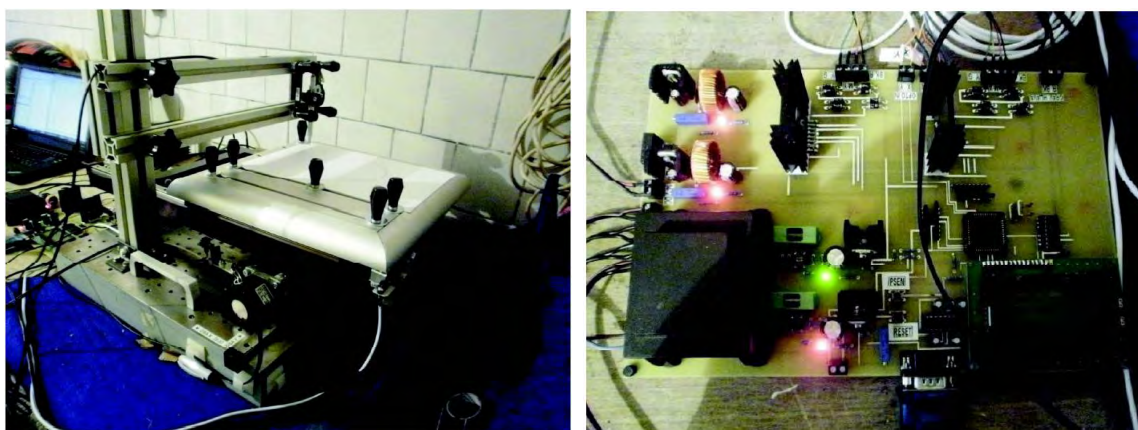
Obrázek 2: Schématické znázornění řízení gravírovacího zařízení

Mikrokontrolér komunikuje po sériové lince s počítačem. Díky vestavěné funkci ISP jej lze snadno programovat přímo v zařízení. Napájecí obvody byly záměrně zvoleny dva. Důvodem je oddělení napájení signálové části od výkonových obvodů budičů motorů, které by při zatížení mohly způsobit pokles napětí na napájení mikroprocesoru a tím i jeho reset.

Pro napájení mikrokontroléru bylo vstupní střídavé napětí transformátorem sníženo na 15V a stabilizátorem stabilizováno na 5V.

Transformátor obsahuje dvě oddělené větve vinutí – jedna pro napájení mikrokontroléru a druhá je určena pro napájení gravírovací hlavičky. K napájení budících obvodů krokových motoru POWERMAX II bylo stanoveno napětí 5V s odběrem 1.4 A přivedené z externího zdroje.

Jaké bude nejvhodnější zapojení krokových motorů a následně jejich řízení, bylo z části dáno již výrobcem. Jde o motory s bipolárním sériovým zapojením. Dosáhne se tak vyššího točivého momentu v oblastech nižších krokovacích frekvencí. K řízení těchto motorů se často používají H-můstky. Jde o čtveřici tranzistorů s ochrannými diodami spínaných v určitém pořadí tak, aby v danou chvíli protékal proud vinutím krokového motoru v požadovaném směru [1,2].



Obrázek 3: vlevo) realizované gravírovací zařízení; vpravo) deska plošného spoje řídicí krokové motory, komunikaci s PC a řízení gravírovací hlavičky [2]

1.4 Pomocné obvody

Z obrázku 2 je patrné, že řízení celého zařízení potřebuje nutně zpětnou vazbu. Ta se v našem případě využívá k určení počáteční polohy – bod [0,0]. Jedná se o dvojici optických závor z fototranzistorů a diody. Na pohyblivém stolku jsou umístěny mechanické zábrany. Pokud se zábrana při pohybu stolku dostane do optické závory, přeruší světelný tok a v mikrokontroléru je toto vyhodnoceno jako přerušení a tedy dojezd na nulovou výchozí polohu. Závor jsou umístěny dvě, pro každou osu jedna. V rozšířené verzi by se mohly doplnit ještě další dvě optické závor, které by hlídaly i maximální polohu plotteru. Prozatím je tato hranice hlídána softwarově tak, že nelze gravírovat grafiku větší než povolené limity.

Avšak i kdyby selhala softwarová ochrana, není gravírovací stůl ponechán bez záložních čidel. Od optické závory jsou ve vzdálenosti 1cm umístěny spínací mikrotlačítka (v klidu trvale sepnuta), která při dojezdu stolku na jejich stlačecí plochu, se rozpojí a přeruší tak napájení krokových motorů. Tlačítka jsou na stolku čtyři, dva pro každou osu. Touto ochranou je zaručeno, že motory nebudou hřídeli otáčet až za hranice možností pohybu a nedojde tak k jejich poškození.

1.5 Gravírovací hlavička

Gravírovací hlavička je prozatím z bezpečnostních důvodů realizována pouze pomocí slabého laseru o výkonu 5W (550nm). Ten plní spíše funkci signalizační a při ladění vývojové aplikace alespoň nehrozilo nebezpečí oslepení. Další užitek byl bezesporu ten, že nebyla velká spotřeba materiálu. Grafika je prozatím realizována rozšířením této gravírovací hlavičky

o posuv v ose Z, kterou tvoří stejnosměrný motor a jemná mechanika. V té je upnut fix, jenž je hlavicí unášen nad papírem a při realizaci pixelu je automaticky sesunut blíže k papíru. Dochází tak ke kreslení. Výhoda tohoto rozšíření hlavice o Z pohyb je bezesporu ten, že v budoucnu, lze do něj upnout, stejně jako v komerčních výrobcích, frézovací hlavicí a celé zařízení tak transformovat velmi rychle na výrobní sestavu pro plošné spoje.

1.6 Gravírování bitmapy

Zařízení dokáže gravírovat obrázky (bitmapy) a vektory. Vstupní snímek může být libovolného formátu, který dokáže být otevřen pomocí program MATLAB. Jako nejlepší se jeví BMP, PNG. V krajním případě lze použít i velmi rozšířený JPEG, ale je nutné si uvědomit, že jde o formát se ztrátovou kompresí, která nám bohužel vytváří nežádoucí efekty na ostrých hranách. Jelikož zařízení prozatím není vybaveno laserem a u použité fixy není možné určit, jakou sytostí má být fyzický pixel realizován, využíváme techniky nazývané dithering. To je technika užívaná v počítačové grafice k vytvoření iluze měnících se odstínů šedi. Výstupem algoritmu je obvykle pouze binární obraz, kde iluzi odstínu tvoří právě shluky těchto binárních bodů. Existuje celá řada metod na tuto transformaci, v programu MATLAB je implementována Floyd-Steinbergova metoda [3]. Výsledek transformace je patrný z obrázku 4.



Obrázek 4: vlevo: vstupní šedotónový obraz; vpravo: po převodu na binární obraz s využitím ditheringu metodou Floyd-Stejnberg



Obrázek 5: Realizace bitmapové grafiky pomocí ditheringu a gravírovacího zařízení

Takto převedený vstupní obraz je následně řádek po řádku posílán do zařízení s mikroprocesorem a ten pomocí posuvu hýbe X-Y stolkem. Jakmile je zaslána log.1, zařízení gravírovací hlavičky realizuje fyzický pixel. Výstup takto gravírovaného obrazu je na obrázku 5. Pokud bude k zařízení připojen laser, předpokládáme, že dobou svitu na plochu dřevěné překližky dokážeme realizovat fyzický pixel ve více odstínech šedé (alespoň 16) a metoda ditheringu bude již nepotřebná.

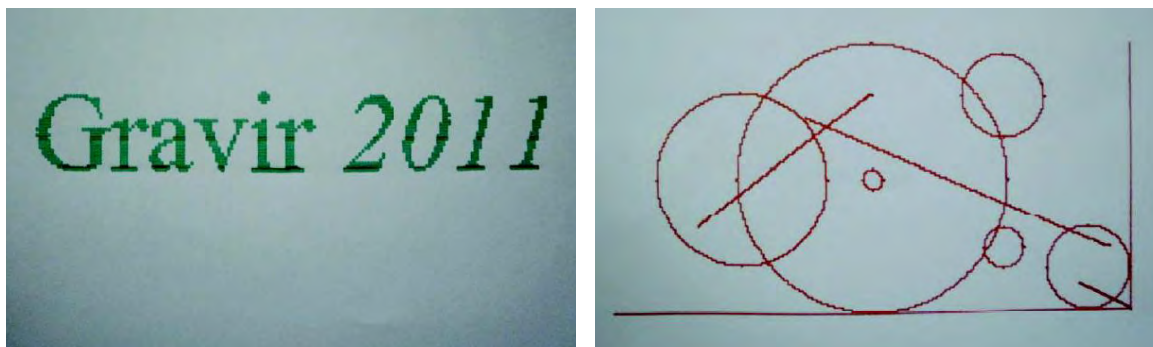
Pohyb X-Y stolku je realizován v CIK-CAK pořadí tzn., že nejprve se načte první řádek bitmapy v daném pořadí a v druhém řádku (stolek je v danou chvíli na konci prvního řádku) se pixely realizují v obráceném pořadí. Přechod z koncové pozice do počátku bez gravírování by přinesl značnou časovou prodlevu mezi zbytečnými přejezdy.

1.7 Gravírování vektoru

Ne vždy je výhodné gravírovat grafiku pomocí bitmapy. Vzhledem k pomalé rychlosti posuvů X-Y stolku a také tomu, že kdybychom gravírovali bitmapu obsahující např. dvě čáry, gravírovali by se vždy řádek po řádku, což by bylo velmi pomalé. Přiložená aplikace dokáže vyčíst vektorová data uložená ve formátu SVG – scalable vector graphics [4]. Tento formát pro ukládání vektorů je velmi rozšířený a podporován hojnou řadou editorů. Pakliže by mělo být zařízení v budoucnu použito i jinak, např. pro výrobu plošných spojů, je vektorové vykreslování velmi žádoucí.

Navržený systém prozatím počítá s vektory typu – úsečka, obdélník, kruh (obrázek 6, vpravo). Další vektorové objekty typu elipsa či volné tvary tvořené Bézierovými křivkami jsou určeny pro další rozšíření aplikace v blízkém časovém horizontu.

Řídicí aplikace, která načte formát SVG jej rozloží na jednotlivé elementy a vybere si pouze výše uvedené známé objekty. V současné chvíli si z parametrů uloží pouze pozici, počáteční body a úhel natočení. Šířka čáry či výplň jsou prozatím ignorovány. Takto získaná data jsou pomocí komunikačního paketu zaslána mikroprocesoru a ten z tvaru paketového rámce rozpozná, o jaký tvar se jedná. Veškerý výpočet souřadnic tak probíhá již na úrovni mikroprocesoru. Toto řešení bylo navrženo zejména pro redukci přenášené informace mezi řídicí stanicí a mikroprocesorem. U výpočtů souřadnic kruhu je ovšem nutné znát přibližné hodnoty funkcí sinus a kosinus. Jelikož mikroprocesor nemá tyto funkce zabudované, byly jejich přibližné hodnoty předpočítány a nahrány do paměti programu. Toto řešení samozřejmě i zrychluje běh řídicí aplikace. Další možnou modifikací algoritmu by bylo nahrazení kruhu N-úhelníkem, takže by se pozice N-1 bodů počítaly pomocí funkcí sinus, cosinus a zbytek proložil úsečkami. Je zřejmé, že v následujícím řešení pro Bézierovy křivky jejich výpočet je časově náročnější, budeme muset realizovat výpočty přímo v řídicí stanici PC. Samotné přenášení paketů s jednoduchými vektory však zůstane nadále podporováno.



Obrázek 6: vlevo: tisk bitmapy pixel po pixelu; vpravo: načtené vektorové objekty ze souboru

2 Technické parametry zařízení

Stoupání hřídele je 36mm na otočku, tzn. po jednom kroku krokového motoru se horní plottrová deska posune o 0,1mm. Bezpečná pracovní plocha horní plošiny činí výměr o rozměrech 240 x 160 mm. Ovládací program umožňuje nastavení rozlišení snímku od 1 do 10 bodů na mm. Je tedy možné při nejnižších rozlišení (1 bod – 1mm) gravírovat bitmapu o rozměrech 240x160 bodů. Obrázky 4 a 5 zobrazují ukázkou gravírování obličeje ženy s rozměrem 150x150 bodů se vzdáleností bodů 0,7mm. Realizace bitmapy je bohužel prozatím časově náročná. Můžou za to větší časové prodlevy mezi sestupem a výstupem fixy při realizaci papíru. Předpokládá se, jak již bylo naznačeno, výměna hrotu fixy za laser. Výsledná rychlost gravírované grafiky tak bude dána výkonem laseru a optikou, kterou lze zaostřit laser na co nejmenší možnou plochu. Omezující bude i použitý materiál, u něhož je důležitým koeficientem zápalná teplota. V průmyslových gravírovacích strojích se často používají koherentní laserové diody o výkonu 60W a jejich paprsek je rozveden pomocí optického vlákna a soustavou zrcátek až na čočku umístěnou na pohyblivém rameni. Maximální rychlost posuvů X-Y stolku je dána několika omezujícími parametry. Jednak dynamikou celé sestavy a také tím, že při vyšších rychlostech posuvu dochází k rezonanci desky, která vydává při posuvu hluk. Rychlost posuvu je tedy 6mm za vteřinu.

3 Závěr

Gravírovací zařízení bylo realizováno pomocí pohyblivého X-Y stolku ovládaného krokovými motory s převodem točivého momentu na pohyb pomocí hřídele a valivých kuličkových ložisek. V našem navrženém řešení se pohybuje gravírovaný materiál a gravírovací hlavice je umístěna stacionárně. K tomuto řešení nás vedl hlavně fakt, že stůl byl již součástí laboratoře a jeho modifikace na gravírovací zařízení nikterak neovlivnila jeho stávající funkčnost. Zařízení je schopné gravírovat šedotónové obrázky (bitmapy) s využitím metody ditheringu, což je optická iluze vyplývající z převodu odstínu šedé na shluk binární bodů. Dalším rozšířením je možnost gravírovat vektorovou grafiku (úsečky, obdélníky a kružnice) načtenou ze souboru ve formátu SVG. V budoucnu předpokládáme výměnu hrotu fixy za laser a gravírovat do dřevěné překližky či balzy.

Poděkování

Tato publikace vznikla za podpory grantu „Podpora výzkumu moderních metod a prostředků v automatizaci“ financované z Interní grantové agentury Vysokého učení technického v Brně (číslo grantu FEKT-S-11-6) a MŠMT v rámci projektu číslo 1683/G1 – Laserové gravírovací zařízení.

Reference

- [1] ERIKSSON, F.: *Stepper motor basics* [online], 1998 [cit 1.6.2011]. Dostupné z <<http://www.solarbotics.net/library/pdf/lib/pdf/motorbas.pdf>>
- [2] HLOŽEK, V.: *Realizace gravírovacího přístroje pro vykreslování grafiky*, VUT Brno, 2011, Diplomová práce 78s., vedoucí diplomové práce: Karel Horák, konzultant práce: Pavel Číp
- [3] FLOYD, R. W., - L. STEINBERG, *An Adaptive Algorithm for Spatial Gray Scale* International Symposium Digest of Technical Papers, Society for Information Displays, 1975
- [4] *SVG – scalable vector graphics* [online], [cit 2.6.2011]. Dostupné z <<http://www.w3.org/Graphics/SVG/>>

ADAPTIVNÍ UČENÍ MODELU ARMAX REKURZIVNÍ KLOUZAVOU METODOU LEVENBERG-MARQUARDT

Jakub DOKOUPIL, Vladimír BURLAK

Ústav automatizace a měřicí techniky, Vysoké Učení Technické v Brně

Kolejní 2906/4, 612 00 Brno

E-mail: xdokou05@stud.feec.vutbr.cz, xburla00@stud.feec.vutbr.cz

Abstrakt:

V článku je rozšířen přístup adaptivní nelineární identifikace s metodou učení Levenberg-Marquardt (LM) využívající techniku numerické optimalizace. Byla nově odvozena rekurzivní varianta původní dávkové metody LM s exponenciálním zapomínáním pracující na klouzavém horizontu pozorování. V příkladu simulace pro pseudolineární model ARMAX dosáhl navržený přístup vyšší přesnosti estimace parametrů oproti lineární metodě nejmenších čtverců s technikou klouzavého datového horizontu.

Klíčová slova: rekurzivní identifikace, metoda Levenberg-Marquardt, exponenciální zapomínání, konečná délka datového okna

1 Úvod

Pseudolineární modely mohou být trénovány lineárními i nelineárními dávkovými algoritmy, např.: [5], rovněž mohou být trénovány po vzorech s aktualizací na základě nově příšlé informace z procesu a s využitím předchozí znalosti modelu. Učení po vzorech umožňuje rekurzifikaci, čímž podstatně snižuje výpočetní nároky. Ve spojitosti s rekurzivní identifikací dochází obecně k poklesu schopnosti adaptace snižováním vlivu korekčního členu způsobeného klesající hodnotou směrové matice, která je v ustáleném stavu mnohem menší než při inicializaci. Motivací je proto začlenění techniky klouzavého datového horizontu za účelem sledování časového vývoje parametrů modelu procesu [1, 3].

Některé z modelů, zejména pak třídy modelů s chybou výstupu jako jsou OE (Output Error) a BJ (Box-Jenkins) modely, obvykle vyžadují nasazení nelineární optimalizační metody. Jedna z vhodných robustních nelineárních optimalizačních metod je právě metoda Levenberg-Marquardt umožňující řešení problému, který není injektivní vzhledem k povaze parametrů modelu.

Hlavním těžištěm práce je proto propojením techniky konečného exponenciálně váženého datového horizontu a rekurzivního učení Levenberg-Marquardt (EF-SWRLM).

2 Metoda EF-SWRLM

Prediktor modelu obecného dynamického procesu lze zapsat ve tvaru funkční závislosti

$$\hat{y}_{k|k-1,\theta} = g(\varphi_{k,\theta}, \theta) \quad (1)$$

kde $\hat{y}_{k|k-1,\theta}$ je predikovaná hodnota výstupu, $g(\cdot)$ je nějaká vektorová funkce mapující vstupy modelu na jeho výstup, $\varphi_{k,\theta}$ je regresní vektor dat (v případě pseudolineární regrese vstupy vektoru představují měřené i odvozené signály z procesu) a θ je vyšetřovaný vektor parametrů modelu.

Úlohu nalezení vektoru θ lze interpretovat jako požadavek dosažení shody predikovaných výstupních hodnot modelu $\hat{Y}_{k,N,\theta}$ s odpovídajícími měřenými hodnotami z procesu $Y_{k,N}$

$$Y_{k,N} = [y_k \quad \cdots \quad y_{k-N+2} \quad y_{k-N+1}]^T \quad (2)$$

$$\hat{Y}_{k,N,\theta} = [\hat{y}_{k|k-1,\theta} \quad \cdots \quad \hat{y}_{k-N+2|k-N+1,\theta} \quad \hat{y}_{k-N+1|k-N,\theta}]^T \quad (3)$$

Za účelem odvození rekurzivní klouzavé varianty metody vyjdeme z konceptu dávkového učení minimalizujícího pro dávku vstupně výstupních dat $\{u(i) \ y(i)\}_{i=1}^N$ hodnotu kriteriální funkce zvolené ve tvaru

$$V_N(\theta) = \|\hat{E}_{k,N,\theta}\|_{Q_N}^2 = \|Y_{k,N} - \hat{Y}_{k,N,\theta}\|_{Q_N}^2 \quad (4)$$

kde vektor chyb predikcí nese značení

$$\hat{E}_{k,N,\theta} = [\varepsilon_{k,\theta} \quad \cdots \quad \varepsilon_{k-N+2,\theta} \quad \varepsilon_{k-N+1,\theta}]^T \quad (5)$$

a pozitivně definitní váhová čtvercová matice Q_N obsahuje na hlavní diagonále faktory exponenciálního zapomínání $\lambda_e \in (0; 1)$ v příslušných mocninách

$$Q_N = \begin{bmatrix} \lambda_e^0 & & \\ & \ddots & \\ & & \lambda_e^{N-1} \end{bmatrix} \quad (6)$$

čímž zajišťuje snižování vlivu starších dat v okně na odhady parametrů θ . Vyjdeme-li z počátečního odhadu $\hat{\theta}_k^0$, obecný gradientní algoritmus s technikou lokální linearizace je pro $i = 1, 2, \dots, i_{max}$ dán iterativním výpočtem

$$\hat{\theta}_k^i = \hat{\theta}_k^{i-1} - \mu_k^{i-1} (R_{k,N}^{i-1})^{-1} V_N^{i-1}(\hat{\theta}_k^{i-1}) \quad (7)$$

kde $\hat{\theta}_k^i$ představuje odhad parametrů θ_k v i -té iteraci výpočtu, $R_{k,N}^i \in \mathbb{R}^{n_\theta \times n_\theta}$ ($n_\theta = \dim \theta$) je směrová matice modifikující směr prohledávání, který je přímo úměrný krokem učení μ_k^i gradientu $V_N^{i-1}(\hat{\theta}_k^{i-1}) \in \mathbb{R}^{n_\theta}$.

U metody Levenberg-Marquardt je směrová matice $R_{k,N}^i$ ve tvaru

$$R_{k,N}^{i-1} = \Psi_{k,N,\hat{\theta}_k^{i-1}}^T Q_N \Psi_{k,N,\hat{\theta}_k^{i-1}} + \lambda^{i-1} I \quad (8)$$

kde

$$\Psi_{k,N,\hat{\theta}_k^{i-1}} = \left(\frac{\partial \hat{Y}_{k,N,\theta}^{i-1}}{\partial \theta_k^{i-1}} \right)^T = \begin{bmatrix} \psi_{k,\hat{\theta}_k^{i-1}}^T \\ \vdots \\ \psi_{k-N+2,\hat{\theta}_k^{i-1}}^T \\ \psi_{k-N+1,\hat{\theta}_k^{i-1}}^T \end{bmatrix} \quad (9)$$

s výrazem $\lambda^i \geq 0$ zajišťujícím regulárnost matice a tím konvergenci metody i v numericky špatně podmíněných případech. Krok učení μ_k^i proto pokládáme roven $\mu_k^i = 1$. Jedna z interpretací výrazu $\lambda^i I$ je nahrazení členu druhých derivací Hessovy matice, v práci [2] je však podán důkaz, že konstrukce směrové matice u dané metody je exaktní matematickou formulací vícekritériálního problému. Pro obnovu parametrů θ podle metody Levenberg-Marquardt ve výsledku píšeme

$$\hat{\theta}_k^i = \hat{\theta}_k^{i-1} + \left(\Psi_{k,N,\hat{\theta}_k^{i-1}}^T Q_N \Psi_{k,N,\hat{\theta}_k^{i-1}} + \lambda I \right)^{-1} \Psi_{k,N,\hat{\theta}_k^{i-1}}^T Q_N \hat{E}_{k,N,\theta_k^{i-1}} \quad (10)$$

Protože přímé analytické řešení $V_N(\theta)$ a tím nalezení $\hat{\theta}_k$ jednorázovým výpočtem není proveditelné, vyjma lineárního případu, tak i formulace rekurzivního algoritmu hledání $\hat{\theta}_k$ vyžaduje zavedení určité aproximace. Jednotlivé aproximace, které budou uvedeny pro

odvození rekurzivní varianty metody Levenberg-Marquardt pracující na klouzavém exponenciálně váženém datovém okně, budou exaktním rekurzivním řešením problému lineárního. Předpokládejme jednokrokovou obnovu podle (10)

$$\hat{\theta}_k = \hat{\theta}_{k-1} + P_{k,N} \Psi_{k,N,\hat{\theta}}^T Q_N \hat{E}_{k,N,\hat{\theta}} \quad (11)$$

kde z důvodu realizovatelnosti rekurzifikace byly použity následující aproximace [6]

$$\Psi_{k,N,\hat{\theta}} \approx \Psi_{k,N,\hat{\theta}_k} \quad (12)$$

$$\hat{E}_{k,N,\hat{\theta}} \approx \hat{E}_{k,N,\hat{\theta}_k} \quad (13)$$

přestože pro většinu modelů výpočet vektoru $\hat{E}_{k,N,\hat{\theta}_k}$ a jeho derivace vyžaduje aktualizaci všech elementů v každém kroku výpočtu.

Rekurentním vyjádřením směrové matice $P_{k,N}$ akumulující v sobě N pozorování obdržíme

$$\begin{aligned} P_{k,N} &= (\Psi_{k,N,\hat{\theta}}^T Q_N \Psi_{k,N,\hat{\theta}})^{-1} \\ &= \left([\psi_{k,\hat{\theta}} \quad \Psi_{k-1,N-1,\hat{\theta}}^T] \begin{bmatrix} \lambda_e^0 & 0 \\ 0 & \lambda_e Q_{N-1} \end{bmatrix} \begin{bmatrix} \psi_{k,\hat{\theta}}^T \\ \Psi_{k-1,N-1,\hat{\theta}} \end{bmatrix} + \lambda I \right)^{-1} \\ &= \left([\psi_{k,\hat{\theta}} \quad \Psi_{k-1,N,\hat{\theta}}^T] \begin{bmatrix} \lambda_e^0 & 0 \\ 0 & \lambda_e Q_N \end{bmatrix} \begin{bmatrix} \psi_{k,\hat{\theta}}^T \\ \Psi_{k-1,N,\hat{\theta}} \end{bmatrix} - \lambda_e^N \psi_{k-N,\hat{\theta}} \psi_{k-N,\hat{\theta}}^T + \lambda I \right)^{-1} \\ &= (\lambda_e P_{k-1,N}^{-1} + \psi_{k,\hat{\theta}} \psi_{k,\hat{\theta}}^T - \lambda_e^N \psi_{k-N,\hat{\theta}} \psi_{k-N,\hat{\theta}}^T + \lambda I)^{-1} \end{aligned} \quad (14)$$

Výraz $\Psi_{k,N,\hat{\theta}}^T Q_N \hat{E}_{k,N,\hat{\theta}}$ vystupující v (11) vyjádříme jako

$$\Psi_{k,N,\hat{\theta}}^T Q_N Y_{k,N} = \psi_{k,\hat{\theta}} \varepsilon_{k,\hat{\theta}} - \lambda_e^N \psi_{k-N,\hat{\theta}} \varepsilon_{k-N,\hat{\theta}} \quad (15)$$

s předpokládanou hodnotou gradientu na intervalu pozorování od $(k - N)$ až po $(k - 1)$ rovnou

$$\lambda_e \Psi_{k-1,N,\hat{\theta}}^T Q_N \hat{E}_{k-1,N,\hat{\theta}} = 0 \quad (16)$$

Přímé řešení matice $P_{k,N}$ podle (14) není z důvodu výpočetní náročnosti příliš praktické, navíc, častěji vyžadujeme znalost matice $P_{k,N}$ a ne její inverze. Ovšem užitím lemma o inverzi matice přímo na $P_{k,N}$ podle (14) není přínosem, protože aditivní regularizační člen λI má hodnotu n_θ . Uspokojivé zjednodušení problému je umožněno přičítáním $n_\theta \lambda$ k $\psi_{k,\hat{\theta}} \psi_{k,\hat{\theta}}^T$ vždy s jedním nenulovým prvkem na hlavní diagonále, jak je navrženo v [4], tedy

$$P_{k,N} = (\lambda_e P_{k-1,N}^{-1} + \psi_{k,\hat{\theta}} \psi_{k,\hat{\theta}}^T - \lambda_e^N \psi_{k-N,\hat{\theta}} \psi_{k-N,\hat{\theta}}^T + \lambda n_\theta Z_{k,n_\theta})^{-1} \quad (17)$$

zde $Z_{k,n_\theta} \in \mathbb{R}^{n_\theta \times n_\theta}$ je nulová matice vyjma jediného diagonálního prvku na pozici $(k \bmod n_\theta) + 1$, který je jednotkový. Výraz (17) je po vykonání periody n_θ výpočtů prakticky totožný s (14).

Výhodou varianty (17) je možnost přepisu na tvar

$$P_{k,N} = (\lambda_e P_{k-1,N}^{-1} + \tilde{\psi}_{k,\hat{\theta}} (\Delta_k)^{-1} \tilde{\psi}_{k,\hat{\theta}}^T - \lambda_e^N \psi_{k-N,\hat{\theta}} \psi_{k-N,\hat{\theta}}^T)^{-1} \quad (18)$$

kde byl zaveden rozšířený vektor $\tilde{\psi}_{k,\hat{\theta}} \in \mathbb{R}^{n_\theta \times 2}$

$$\tilde{\psi}_{k,\hat{\theta}} = \begin{bmatrix} 0 & \dots & 0 & \psi_{k,\hat{\theta}}^T & \dots & 0 \end{bmatrix}^T \quad (19)$$

↑
pozice = $(k \bmod n_\theta) + 1$

a odpovídající váhová matice

$$(\Delta_k)^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & \lambda n_\theta \end{bmatrix} \quad (20)$$

Rovnici pro výpočet matice $P_{k,N}$ (18) ekvivalentně přepíšeme zavedením pomocné matice $P_{k,N+1}$ akumulující v sobě informace z posledních $N + 1$ pozorování

$$P_{k,N+1} = (\lambda_e P_{k-1,N}^{-1} + \tilde{\psi}_{k,\hat{\theta}} (\Delta_k)^{-1} \tilde{\psi}_{k,\hat{\theta}}^T)^{-1} \quad (21)$$

$$P_{k,N} = (P_{k,N+1}^{-1} - \lambda_e^N \psi_{k-N,\hat{\theta}} \psi_{k-N,\hat{\theta}}^T)^{-1} \quad (22)$$

Aplikací lemmy o inverzi matice na výše položené rovnice (21), (22) a dosazením dílčího výsledku výpočtu gradientu (15) do vztahu (11), ve výsledku obdržíme rekurzivní zápis obnovy parametrů na klouzavém exponenciálně váženém datovém horizontu ve shodě s učením Levenberg-Marquardt

$$P_{k,N+1} \frac{1}{\lambda_e} \left[P_{k,N} - P_{k,N} \tilde{\psi}_{k,\hat{\theta}} (\lambda_e \Delta_k + \tilde{\psi}_{k,\hat{\theta}}^T P_{k,N} \tilde{\psi}_{k,\hat{\theta}})^{-1} \tilde{\psi}_{k,\hat{\theta}}^T P_{k,N} \right] \quad (23)$$

$$P_{k,N} = P_{k,N+1} - \lambda_e^N P_{k,N+1} \psi_{k-N,\hat{\theta}} (-1 + \lambda_e^N \psi_{k-N,\hat{\theta}}^T P_{k,N+1} \psi_{k-N,\hat{\theta}})^{-1} \psi_{k-N,\hat{\theta}}^T P_{k,N+1} \quad (24)$$

$$\hat{\theta}_k^{N+1} = \hat{\theta}_{k-1} + P_{k,N} \psi_{k,\hat{\theta}} \varepsilon_{k,\hat{\theta}} \quad (25)$$

$$\hat{\theta}_k = \hat{\theta}_k^{N+1} - \lambda_e^N P_{k,N} \psi_{k-N,\hat{\theta}} \varepsilon_{k-N,\hat{\theta}} \quad (26)$$

Dimenze matice ve vztahu (23) nad níž je požadována operace inverze je nyní nanejvýš dvě pro libovolný počet vstupů ve vektoru $\hat{\theta}_k$.

3 Model ARMAX a jeho gradient

Model ARMAX (Autoregressive Moving Average with Exogenous Input) se řadí mezi modely s chybou rovnice. Optimální ARMAX prediktor je [5]

$$\hat{y}_{k|k-1,\theta} = \frac{B(q)}{C(q)} u_k + \left(1 - \frac{A(q)}{C(q)}\right) y_k \quad (27)$$

kde

$$A(q) = 1 + a_1 q^{-1} + \dots + a_{n_a} q^{-n_a} \quad (28)$$

$$B(q) = b_1 q^{-1} + \dots + b_{n_b} q^{-n_b} \quad (29)$$

$$C(q) = 1 + c_1 q^{-1} + \dots + c_{n_c} q^{-n_c} \quad (30)$$

jsou diskrétní polynomy příslušného stupně s operátorem posuvu q ve významu $q^{-1} y_k = y_{k-1}$. Vytváří stabilní prediktor i v případech, kdy samotný model ARMAX je nestabilní. S porovnáním s modelem ARX (Autoregressive with Exogenous Input), je model ARMAX více flexibilní neboť rozšiřuje model stochastické části procesu o filtr $C(q)$. Na druhou stranu zavedené rozšíření vyžaduje začlenění techniky pseudolineární regrese, čímž se predikce modelu stává nelineární funkcí vzhledem k působení vektoru θ .

Technika numerické optimalizace gradientními algoritmy využívá znalosti gradientu kritériální funkce. Za tímto účelem proto sestavíme gradient predikce modelu. Násobením rovnice (27) polynomem $C(q)$ získáme

$$C(q)\hat{y}_{k|k-1,\theta} = B(q)u_k + (C(q) - A(q))y_k \quad (31)$$

Derivací rovnice (31) podle a_i obdržíme

$$\frac{\partial \hat{y}_{k|k-1,\theta}}{\partial a_i} = -\frac{1}{c(q)}y_{k-i} \quad (32)$$

Derivací rovnice (31) podle b_i obdržíme

$$\frac{\partial \hat{y}_{k|k-1,\theta}}{\partial b_i} = -\frac{1}{c(q)}u_{k-i} \quad (33)$$

Derivací rovnice (31) podle c_i obdržíme

$$\frac{\partial \hat{y}_{k|k-1,\theta}}{\partial c_i} = -\frac{1}{c(q)}(y_{k-i} - \hat{y}_{k-i|k-i-1,\theta}) \quad (34)$$

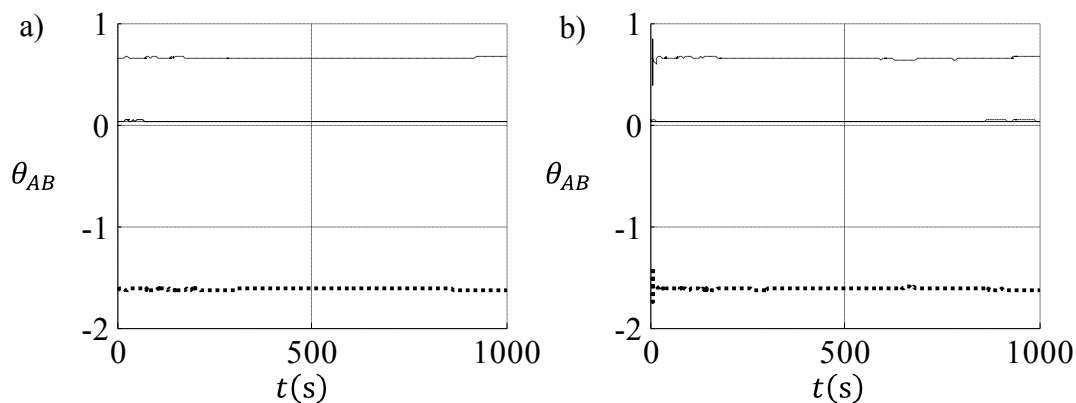
4 Výsledky dosažené simulací

V předchozí sekci byly uvedeny teoretické aspekty navržené rekurzivní metody Levenberg-Marquardt pracující na exponenciálně váženém klouzavém datovém horizontu (EF-SWRLM). V této části budou vyšetřeny vlastnosti metody na příkladech simulací na matematickém modelu. Výsledky dosažené metodou EF-SWRLM budou porovnány s výsledky dosaženými lineární rekurzivní metodou nejmenších čtverců pracující na klouzavém exponenciálně váženém datovém horizontu (EF-SWRLS) [1, 3]. Metoda EF-SWRLS se od EF-SWRLM odlišuje zjednodušeným předpokladem lineární závislosti predikce modelu na odhadech parametrů a položením $\lambda = 0$. Předpokládáme proces ARMAX ve tvaru

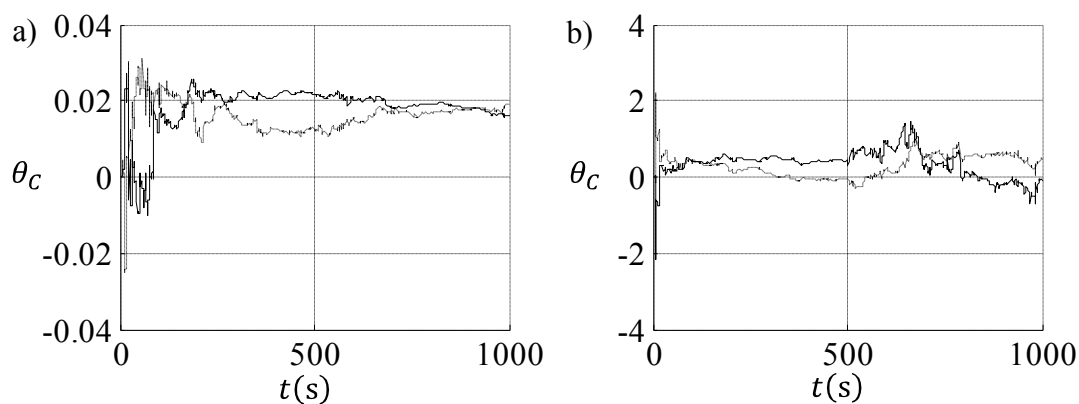
$$\begin{aligned} y_k - 1.6253y_{k-1} + 0.6592y_{k-2} \\ = 0.0363u_{k-1} + 0.0316u_{k-2} + v_k + 0.02v_{k-1} + 0.02v_{k-2} \end{aligned} \quad (35)$$

kde $\{v_k\}$ představuje neměřitelnou veličinu s vlastnostmi bílého šumu s rozptylem $\sigma_v^2 = 0.01^2$, $\{u_k\}$ tvoří sekvenci persistentně budícího signálu s nulovou střední hodnotou a jednotkovým rozptylem. Z důvodu korektního porovnání bylo u obou metod voleno $\lambda_e = 0.999$, počáteční inicializace $P_0 = 10^9 I$ a délka časového okna $N = 200$. A pro EF-SWRLM bylo zvoleno $\lambda = 1$.

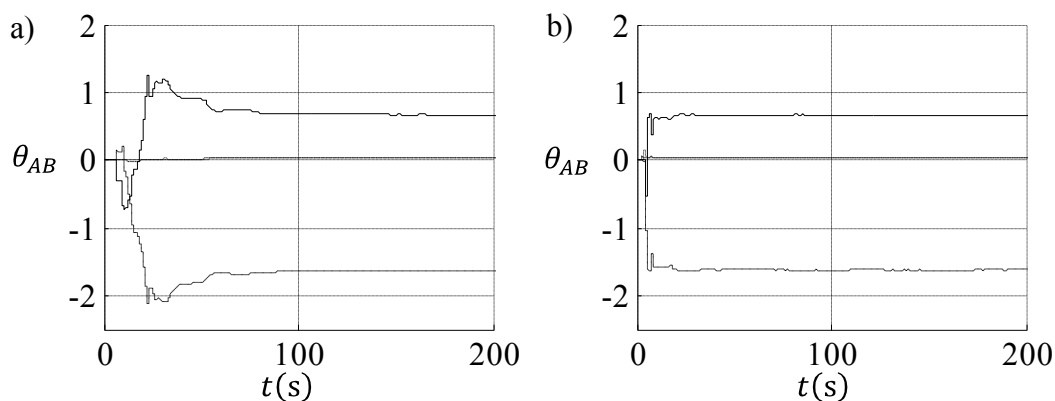
Výsledky estimace s užitím obou metod jsou prezentovány na následujících obrázcích. Obrázek (1) zachycuje vývoj parametrů části $\theta_{AB} = [b_1 \ b_2 \ a_1 \ a_2]$ a Obrázek (2) pak části $\theta_C = [c_1 \ c_2]$ při počáteční inicializaci obou metod modelem ARX. Číselné vyhodnocení kvality estimace s ohledem na akumulovanou chybu $\delta = \sum_{i=0}^{T_t} \|\hat{\theta}(i) - \theta\|_2^2$ pro celý časový horizont $T_t = 1000s$ je uvedeno v Tabulce (1). Z vyšetřovaných průběhů je patrné, že vyšší kvality estimace dosáhl při počáteční inicializaci nelineární přístup s metodou EF-SWRLM. Na druhou stranu, bez kvalitní počáteční inicializace jak zachycuje Obrázek (3), nelineární přístup významně pomaleji konverguje. Urychlení konvergence lze podmínit vhodnou délkou horizontu pozorování N a teoreticky snížením hodnoty parametru λ , jak je dokázáno v [2]. Poznamenejme, že pro volbu $\lambda = 0$ nebyla metoda v daném případě z počátečního odhadu schopna konvergovat.



Obrázek 1: Průběh estimace inicializovaných parametrů části procesu $\theta_{AB} = [b_1 \ b_2 \ a_1 \ a_2]$. a) Metoda EF-SWRLM. b) Metoda EF-SWRLS



Obrázek 2: Průběh estimace parametrů části procesu $\theta_C = [c_1 \ c_2]$. a) Metoda EF-SWRLM. b) Metoda EF-SWRLS



Obrázek 3: Průběh estimace parametrů části procesu $\theta_{AB} = [b_1 \ b_2 \ a_1 \ a_2]$. a) Metoda EF-SWRLM. b) Metoda EF-SWRLS.

Tabulka 1: Akumulovaná kvadratická chyba estimace parametrů modelu procesu

	EF-SWRLS	EF-SWRLM
$\delta_{A,B} = \sum_{i=0}^{T_t} \ \hat{\theta}_{AB}(i) - \theta_{AB}\ _2^2$	14.466	0.0025
$\delta_C = \sum_{i=0}^{T_t} \ \hat{\theta}_C(i) - \theta_C\ _2^2$	282.7667	0.9171

5 Závěr

Tento článek prezentuje a odvozuje rekurzivní klouzavou metodu Levenberg-Marquardt s exponenciálním zapomínáním umožňující adaptivní trénování obecně nelineární struktury modelu. Výsledky experimentu naznačují, že navržená metoda je schopna dosáhnout přesnějšího odhadu estimace než lineární metoda nejmenších čtverců za stejných podmínek simulace. Rovněž vykazuje vyšší schopnost konvergence s porovnáním s metodou v literatuře označované jako „rekurzivní metoda maximální věrohodnosti (rekurzivní Gauss-Newtonova metoda aplikovaná na model ARMAX [6])“. Oblast dalšího výzkumu bude vedena na vývoj výpočetně efektivních rekurzivních algoritmů s vyšším stupněm linearizace schopných aproximovat nelineární proces s vyhovující přesností.

Poděkování

Tato práce vznikla za podpory Grantové agentury České republiky (102/09/H081 SYNERGY – Senzory a inteligentní senzorové systémy) a za podpory grantu “Podpora výzkumu moderních metod a prostředků v automatizaci“ financované z Interní grantové agentury Vysokého učení technického v Brně (číslo grantu FEKT-S-11-6).

Reference

- [1] Ding, F., Xiao, Y.: A Finite-Data-Window Least Squares Algorithm with a Forgetting Factor for Dynamical Modeling. Applied Mathematics and Computation, sv. 186, č. 1, s. 184-192, 2007. ISSN 0096-3003
- [2] Dokoupil, J.: Adaptivní Parametrická Identifikace v Klouzavém Režimu. Pojednání o disertační práci. Brno: Vysoké Učení Technické v Brně, 2011
- [3] Liu, H., Zhenya, H.: A Sliding-Exponential Window RLS Adaptive Filtering Algorithm: Properties and Applications. Signal Processing, sv. 45, č. 3, s. 357-368, 2007. ISSN 0165-1684
- [4] Ljung, L.: Recursive Identification Algorithms. Circuits, Systems, and Signals Processing, sv. 21, č. 1, s. 57-68, 2002. ISSN 1531-5878
- [5] Nelles, O.: Nonlinear System Identification. Berlin: Springer, 2001. 785 stran. ISBN 3-540-67369-5
- [6] Söderstrom, T., Stoica, P.: System Identification. Cambridge: Prentice Hall, Int., 1989. 612 stran. ISBN 0-13-881236-5

VYLEPŠENÍ VLASTNOSTÍ SAMOČINNĚ SE NASTAVUJÍCÍCH REGULÁTORŮ S IDENTIFIKACÍ ZALOŽENOU NA NEURONOVÝCH SÍTÍCH

Vlastimil Lorenc

Ústav automatizace a měřicí techniky, Vysoké Učení Technické v Brně

Kolejní 2906/4, 612 00 Brno

E-mail: xloren05@stud.feec.vutbr.cz

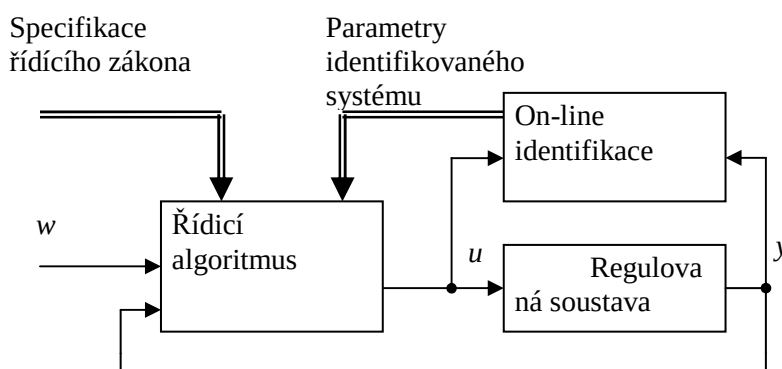
Abstrakt/Abstract:

Článek pojednává o problematice samočinně se nastavujících regulátorů (dále STC) z pohledu vylepšení robustnosti a spolehlivosti. V první části článku je popsána metoda zajišťující výběr vhodných tréninkových vzorů pro neuronovou síť v případě on-line identifikace regulované soustavy. Druhá část článku se zabývá možností zvýšení robustnosti návrhu STC za použití paralelního robustního regulátoru. Je navržen algoritmus, který je schopen připojit do uzavřené smyčky řídicí zákon, který je v daném okamžiku vyhodnocen jako více vhodný dle předem daného kritéria.

Klíčová slova/Keywords: samočinně se nastavující regulátory, neuronové sítě, řízení se supervizorem

1 Úvod

Po řízení časově invariantních systémů, tedy systémů s časově proměnnými parametry, jsou známy dva základní přístupy. Prvním je použití robustního algoritmu, který je stabilní pro celou množinu přenosů, které může daný proces nabývat. Toto řešení však vzhledem k nutnosti vytvoření rezervy vede na nižší výslednou kvalitu regulačního děje. Další možností je použití některé ze známých forem adaptivního řízení. V tomto článku se budeme zabývat STC regulátorem, jehož blokové schéma je na obrázku 1 [4].



Obrázek 1: Blokové schéma STC

Je zřejmé, že algoritmus STC regulátoru se logicky rozpadá na dvě základní části: identifikační algoritmus a samotný řídicí zákon včetně jeho návrhu na základě identifikovaných dat. V této práci je jako identifikační algoritmus použit identifikace založená na neuronové síti. V literatuře se často setkáváme s tímto přístupem k identifikaci a byly vyzkoušeny mnohé známé modely sítí, neuronů i metod učení. Málo kdy je však řešen

samotný problém on-line identifikace, která je problémová zejména z pohledu tvorby tréninkové množiny. Tento článek řeší právě výběr vhodných dat, která se následně použijí pro učení neuronové sítě.

Funkční identifikační algoritmus je samozřejmě klíčový pro správnou a bezpečnou funkci STC regulátoru. Pokud identifikace selže a poskytne chybný odhad parametrů, dojde následně k syntéze řídicího zákona, který neodpovídá realitě. Tato práce kombinuje přístup adaptivní a robustní regulace za použití nadřazeného algoritmu (supervizoru) založeného na přístupu známém jako „unfalsified control theory“. Pomocí tohoto přístupu lze dosáhnout výrazného zlepšení spolehlivosti regulace v případech, kdy adaptivní systém není schopen generovat dostatečně kvalitní akční zásahy.

2 On-line identifikace pomocí neuronové sítě a tvorba vhodné trénovací množiny

Nejrozšířenější metodou pro odhad parametrů regulovaného systému je metoda nejmenších čtverců (RLS). Tato metoda bývá rozšířena o princip exponenciálního zapomínání, který zajišťuje časovou penalizaci informací vstupujících do identifikačního algoritmu a tím vylepšuje jeho schopnost sledovat změny dynamiky identifikovaného systému. S metodou exponenciálního zapomínání ovšem souvisí jev v literatuře označován jako estimator windup [4]. Pokud je soustava regulátorem nedostatečně vybudena, tedy nově přichodící data nenesou dostatečnou informaci o dynamice systému, může dojít k selhání celého identifikačního procesu. V literatuře lze nalézt mnohé způsoby řešení toho to problému. Například bylo navrženo takzvané směrové zapomínání, pomocí kterého je dosaženo zapomínání staré informace pouze v případě, že nová data v dostatečné míře reprezentují dynamiku identifikovaného systému [4]. Pokud nezamýšlíme STC regulátor provozovat pouze v laboratorních podmínkách je řešení ustálených stavů systému v identifikačním algoritmu bezpodmínečně nutno řešit.

Výhodou identifikačních algoritmů založených na neuronových sítích je to, že u nich lze určitým způsobem upravovat data sloužící pro identifikaci. U klasických algoritmů jako je RLS musíme na základě každé nově přichodící dvojice $u(k)$, $y(k)$ provést iteraci identifikačního algoritmu. U neuronových sítí, vzhledem k tomu že zde máme k dispozici trénovací množinu, můžeme nejprve sestavit nový trénovací vzor a na základě definované metody usoudit, zda nový vzor je pro zařazení do učícího procesu vhodný, či nikoliv. Na tomto principu je založen algoritmus v práci [3], kde jsou v množině trénovacích vzorů hledány vždy dva nejbližší vzory a posléze je ten starší z nich odstraněn. Tímto způsobem bylo dosaženo toho, aby množina trénovacích vzorů obsahovala více pestré informace o dynamice identifikovaného systému.

Pro určení vzdálenosti jednotlivých vzorů je třeba zavést vhodnou metriku. Mějme transformaci R , pomocí které transformujeme regresní vektor pro systém n -tého řádu

$$\varphi^T(k-1) = [-y(k-1), -y(k-2), \dots, -y(k-na), u(k-1), u(k-2), \dots, u(k-nb)]$$

do nového prostoru:

$$\varphi' = R\varphi. \quad (1)$$

Vzdálenost dvou vektorů v tomto novém prostoru je pak:

$$\|\varphi'_i - \varphi'_j\| = \|R\varphi_i - R\varphi_j\| = \|R\Delta\varphi\| = \sqrt{\Delta\varphi^T R^T R \Delta\varphi}. \quad (2)$$

Zavedením metriky

$$Q = R^T R, \quad (3)$$

se vzdálenost dvou vektorů vypočítá podle:

$$\sqrt{\Delta\varphi^T R^T R \Delta\varphi} = \sqrt{\Delta\varphi^T Q \Delta\varphi} = \|\Delta\varphi\|_Q = \|\varphi_i - \varphi_j\|_Q. \quad (4)$$

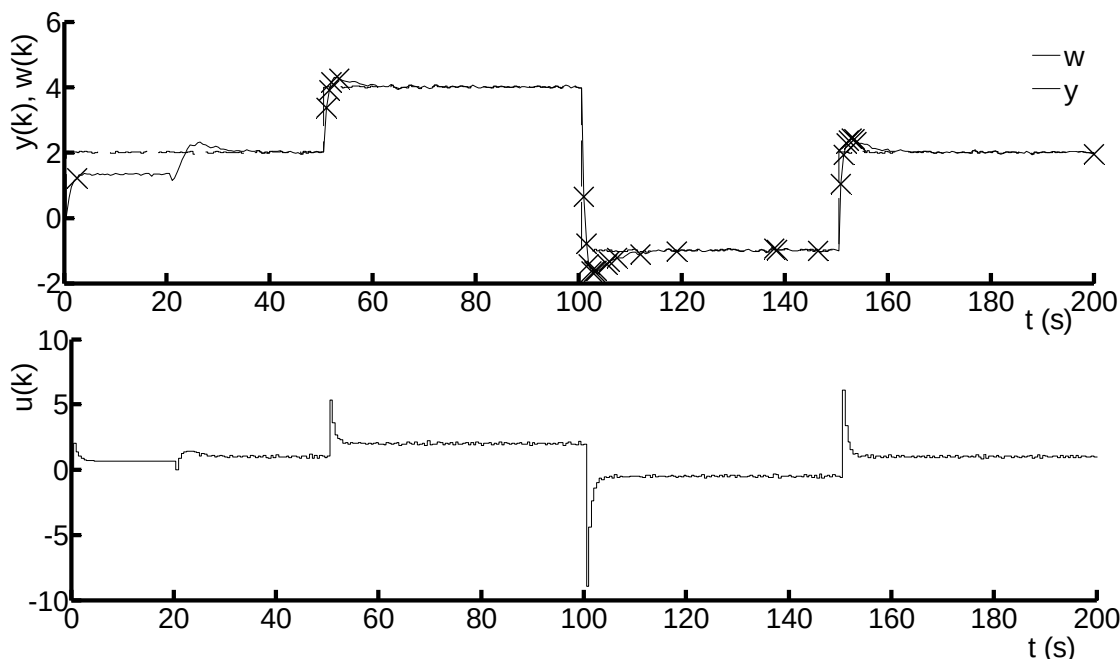
Nevýhodou této metody je nezávislost rozhodování o výběru vzorů do trénovací množiny z hlediska času. Může se totiž stát například to, že na proces zapůsobí porucha a tím vzniknou trénovací vzory, jejichž vzdálenost bude od ostatních velmi výrazná. Tyto vzory pak v trénovací množině, při použití uvedeného algoritmu, zůstanou po dlouhou dobu. Tento fakt, může výrazně zhoršit adaptační schopnosti regulátoru. Z výše popsaného důvodu byl algoritmus pro výběr vzorů nově rozšířen o exponenciální časové vážení vstupních vzorů. Důsledek tohoto rozšíření je schopnost algoritmu penalizovat a vyřazovat s větší pravděpodobností starší trénovací vzory. V tomto novém algoritmu se všechny vzdálenosti násobí koeficientem γ , který je určen následujícím vztahem

$$\gamma = e^{(1-\lambda)\max(t_1, t_2)}, \quad (5)$$

kde: λ koeficient určující míru penalizace starých dat,

t_1, t_2 proměnné určující stáří dvou vzájemně porovnávaných vzorů.

Následující graf ukazuje rozložení trénovacích vzorů při použití doplnění metody výběru vzorů o exponenciální penalizaci starších dat.



Obrázek 2: Výběr vzorů do trénovací množiny (křížkem jsou označena data, která byla po 200 sekundách experimentu obsažena v trénovacím setu).

3 Zvýšení robustnosti STC regulátoru pomocí nadřazeného členu

Většina časově invariantních procesů, na které by bylo vhodné aplikovat adaptivní regulátor je možno regulovat pomocí robustně navrženého řídicího zákona (v praxi se s touto variantou setkáváme velmi často). Snahou této části je dosáhnout spojení adaptivního regulátoru, který je teoreticky schopen generovat optimální akční zásahy v každém pracovním bodě procesu, a robustního regulátoru garantující jakousi minimální kvalitu regulačního pochodu.

Spojení těchto dvou přístupů může být dosaženo paralelním výpočtem robustního a adaptivního řídicího zákona a supervizoru, který je schopen vyhodnotit který z algoritmů poskytuje v daném čase lepší výsledky a zařadí jej do zpětné vazby.

3.1 Výběr vhodného řídicího zákona

Metoda výběru vhodného řídicího zákona je založena na metodě v anglickém jazyce zvané „unfalsified control“ [1, 2]. V této podkapitole bude stručně popsán její princip.

Mějme nejvhodnější řídicí zákon C ze skupiny časově invariantních řídicích zákonů C_i , $i = 1, \dots, N$. Kritérium kvality pro každý zákon označme J_i . Problém je jak vyhodnotit kritérium kvality pro řídicí zákony, které nejsou aktuálně zařazeny do zpětné vazby. Pro tento účel je nezbytné definovat takzvanou fiktivní žádanou hodnotu:

$$\tilde{w}_i(k) = C_i^{-1}u(k) + y(k) \quad (6)$$

V rovnici (6), se vyskytuje inverze matice řídicího zákona, nicméně v [2] bylo dokázáno, že tento výpočet není vždy nezbytný. Nyní můžeme definovat fiktivní odchylku:

$$\tilde{e}_i(k) = \tilde{w}_i(k) - y(k) \quad (7)$$

Posledním krokem je určit tvar ztrátové funkce $J(k)$:

$$J_i(k) = \frac{\|\tilde{e}_i^2\|_k + \alpha \|u^2\|_k}{\|\tilde{w}_i^2\|_k} \quad (8)$$

kde α je nezáporná váhová konstanta a dále

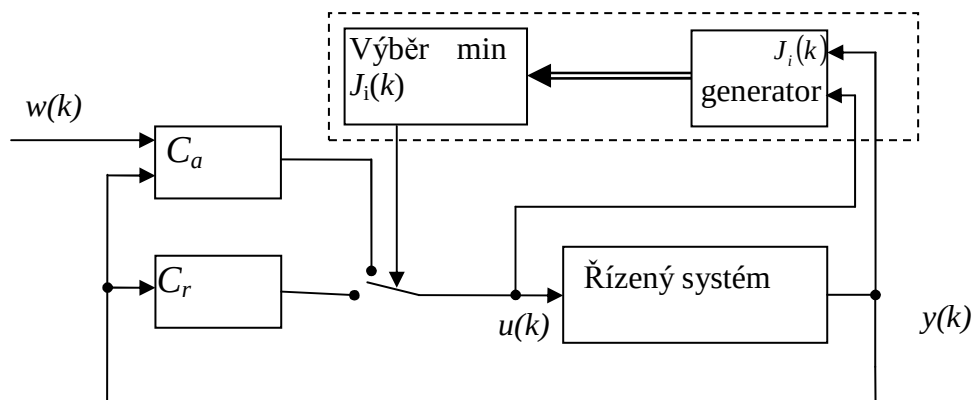
$$\|x\|_k = \sum_{j=k-N}^k e^{-\lambda(k-j)} x(j) \quad (9)$$

je norma signálu s exponenciálním zapomínáním.

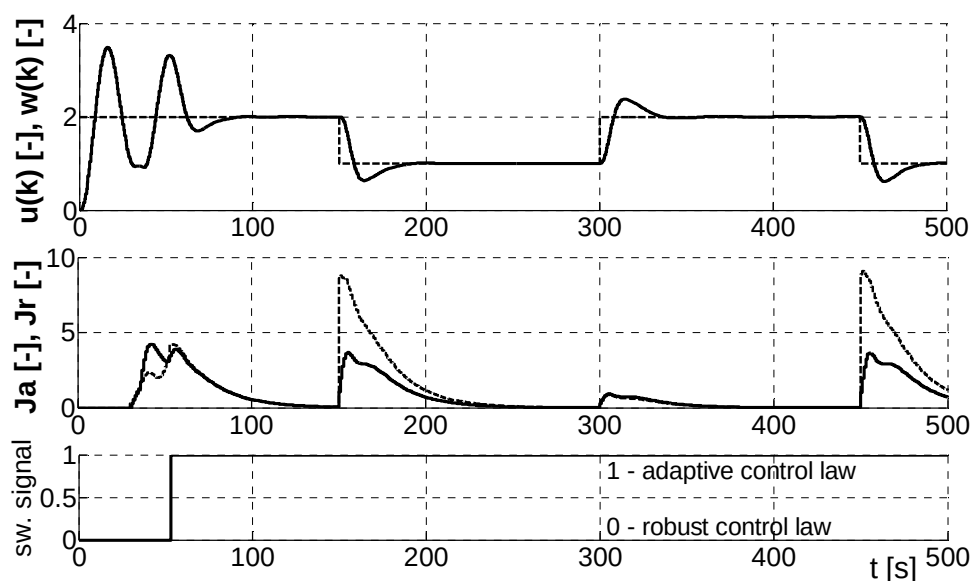
Výpočtem funkce $J(k)$: pro každý C_i lze určit, který řídicí zákon je v daném okamžiku vzhledem ke zvolenému kritériu nejoptimálnější.

3.2 Spojení robustního a adaptivního řídicího zákona

Na základě obecného principu popsaného v předchozí kapitole byl sestaven algoritmus kombinující vlastnosti adaptivní a robustní regulace. Množina řídicích zákonů ze které je vybíráno má v tomto případě pouze dva členy: C_a a C_r (tedy adaptivní a robustní zákon).



Obrázek 3: Princip supervizoru.



Obrázek 4: Experiment s přepínáním řídicích zákonů.

4 Závěr

V první části článku byla ukázána metoda výběru vzorů do trénovací množiny. Jak lze vidět z obrázku 2, algoritmus byl schopen zařadit do učícího procesu pouze ty vzory, které byly informačně bohaté. Vzory pořízené při ustáleném pochodu systému byly vyřazeny. Ve druhé části článku byl nastíněn algoritmus nadřazeného členu pro adaptivní regulátor, který je schopen vyhodnotit nevhodné chování a v takovém případě připojit k řízené soustavě předem navržený robustní regulátor. Na obrázku 4 je znázorněn experiment, při kterém byl popsán princip vyzkoušení. Nejprve byl do obvodu zařazen pevný regulátor. Je zřejmé, že parametry nebyly zvoleny nejlépe, regulační pochod je však stabilní. V čase kolem 50 sekund došlo k automatickému přepnutí na adaptivní řídicí zákon. V tomto čase již byl systém identifikován s takovou přesností aby výsledkem syntézy řídicího zákona byly akční zásahy, které se vykazovaly menší hodnotou kontrolovaného kritéria.

Reference/References

- [1] Ayanendu, P., Stefanovic, M.: Safonov, M., Akar, M.: *Multi-Controller Adaptive Control (MCAC) for a Tracking Problem using an Unfalsification approach*, Proceedings IEEE Conf. on Decision and Control and European Control Conference, 2005
- [2] Dehghani, A., Anderson, B., Lanzon A.: 2007, *Unfalsified Adaptive Control: A New Controller Implementation and Some Remarks*, Proceedings of the European Control Conference 2007
- [3] Vychodil, H.: *Geometric properties of the on-line identification (Geometrické vlastnosti on-line identifikace)*. Doctoral thesis in Czech, Brno, ÚAMT FEKT VUT, 2005.
- [4] Bobál, V., Bohm J., Prokop R. and Fessl J.: *Practical Aspects of Self-Tuning Controllers, Algorithms and Implementation (Praktické aspekty samocinne se nastavujících regulátorů, algoritmy a implementace)*, VUTIUM, 1999, ISBN 80-214-1299-2.

AKCELERACE HASHOVACÍHO ALGORITMU MD5 ZA POMOCI FPGA

Daniel PIŠI

Ústav automatizace a měřicí techniky, Vysoké Učení Technické v Brně

Kolejní 2906/4, 612 00 Brno

E-mail: xpsid00@stud.feec.vutbr.cz

Abstrakt:

Tento dokument diskutuje možnosti hardwarové akcelerace hashovacího algoritmu MD5 na hradlovém poli Xilinx řady Spartan 3A ve spojení se softwarovým mikrokontrolérem MicroBlaze. Datovou propustnost dosaženou za pomoci PLB periférie akcelerující výpočet srovnává s výkonem softwarové implementace běžící na dvou různých konfiguracích jádra MicroBlaze a mikrokontroléru s jádrem ARM7TDMI.

Klíčová slova: MD5, FPGA, MicroBlaze, ARM, hash, akcelerace.

1 Úvod

S hashovacími funkcemi se lze setkat především v aplikacích řešících autorizaci, autentizaci, zaručení integrity předávaných zpráv a databázových systémech, kde usnadňují vyhledávání záznamů a odhalování duplicit. Přestože pro většinu běžných použití postačuje softwarová implementace, nalezneme řadu aplikací, kde je datová propustnost, zpoždění a deterministické chování hashovací funkce kritické, případně příkon, složitost zapojení a cena výkonných procesorů neakceptovatelná. Jedná se o síťový hardware, systémy reálného času, zabezpečovací systémy a platební systémy, viz [5].

Hashovací funkce jsou zpravidla tvořeny řadou matematicko-logických operací, což vybízí k paralelizaci výpočtů za pomoci FPGA. Řada úspěšných implementací ukazuje, že se jedná o efektivní řešení. Článek [4] ukazuje aplikovatelnost tohoto přístupu na prakticky všechny běžně používané hashovací algoritmy a [9], že s využitím FPGA lze mnohonásobně překonat výkon běžných desktopových procesorů.

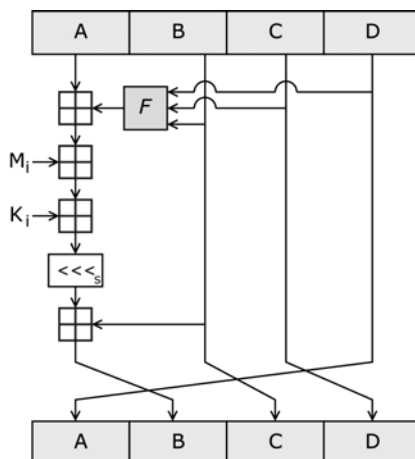
Tento článek staví na práci provedené v rámci semestrální práce [7], zaměřuje se na možnosti akcelerace algoritmu MD5 na FPGA určených pro cenově citlivé aplikace postavené na softwarovém mikrokontroléru MicroBlaze.

2 Algoritmus MD5

Přestože byl hashovací algoritmus MD5 uveden již roku 1992 a dnes ho lze považovat za prakticky prolomený [1], stále se s ním v praxi často setkáváme. Je využíván především pro kontrolu integrity souborů nebo v kombinaci s jinými hashovacími algoritmy a technikami potlačujícími odhalené slabiny pro jednosměrné ukládání a ověřování hesel.

Činnost MD5 je popsána v internetovém standardu RFC 1321 [8], zde budou shrnuty pouze klíčové body. Algoritmus vytváří otisk velikosti 128 bitů ze zprávy libovolné délky. Vstupní data jsou rozdělena na bloky 512 bitů, respektive šestnáct 32-bitových čísel. Doplnění zprávy na násobek délky bloku, tzv. padding, je realizován přidáním bitu hodnoty 1 následovaného nulovými bity do velikosti o 64 bitů nižší násobku 512. Posledních 64 bitů je rezervováno pro celé číslo reprezentující délku původní zprávy v bitech.

Algoritmus pracuje nad 128-bitovým stavem, resp. čtyřmi 32-bitovými slovy označenými A, B, C a D. Tyto jsou před zahájením výpočtu inicializovány na definované hodnoty. Následně MD5 na základě příchozích 512 bitových bloků zprávy stav modifikuje. Zpracování každého bloku sestává z 4 navzájem podobných stupňů. Každý stupeň představuje 16 opakujících se operací tvořených nelineární funkcí F, modulo sčítáním a bitovou rotací vlevo. Na místě funkce F jsou postupně dle zpracovávaného stupně dosazovány funkce



Obrázek 1: Jedna operace algoritmu MD5, zdroj obrázku: Wikipedia

$$F(X, Y, Z) = (X \wedge Y) \vee (\neg X \wedge Z)$$

$$G(X, Y, Z) = (X \wedge Z) \vee (Y \wedge \neg Z)$$

$$H(X, Y, Z) = X \oplus Y \oplus Z$$

$$I(X, Y, Z) = Y \oplus (X \vee \neg Z)$$

3 Softwarový mikrokontrolér MicroBlaze

MicroBlaze je 32-bitový RISC softwarový mikrokontrolér navržený a optimalizovaný společností Xilinx speciálně pro FPGA tohoto výrobce. Dle požadavků konkrétní aplikace lze volit tři nebo pět stupňové zřetězení instrukcí, velikost vyrovnávacích pamětí, jednotku pro správu paměti, hardwarové násobičky a další. K jádru lze za pomoci sběrnice PLB, případně dvoubodové linky FSL připojovat řadu předpřipravených periférií stejně tak jako periférie definované uživatelem. Podrobné informace o architektuře a vlastnostech mikrokontroléru lze nalézt v referenční příručce [10].

4 Softwarová implementace MD5

Pro možnost srovnání výpočetního výkonu různých konfigurací mikrokontroléru byl vytvořen jednoduchý program využívající open source C knihovnu RFC1321-based (RSA-free) MD5 library [3]. Tato byla zvolena pro svou přenositelnost a možnost výpočtu MD5 za běhu po libovolně velkých datových blocích. Knihovnu lze přeložit pro MicroBlaze i ARM7TDMI bez zásahu do zdrojových kódů.

5 Akcelerace výpočtu za pomoci hardware

Pro akceleraci výpočtu MD5 byla v jazyce VHDL navržena periférie s rozhraním PLB (Processor Local Bus). Jedná se o sběrnici s vysokou propustností pro všeobecné použití. Umožňuje mapování periférie do paměťového prostoru mikrokontroléru.

Jádro implementující algoritmus MD5 vychází z projektu FPGA MD5 Cracker [6]. Mimo refaktoringu relevantní části kódu došlo k rozšíření jádra o rozhraní ke sběrnici PLB, transparentní konverzi kódování little a big endian a doplnění logiky zjednodušující a akcelerující zápis dat do periferie za pomoci automatického adresování datového bufferu, čítače zapsaných bytů a automatické nulování bufferu. Tyto funkce snižují režii spojenou s přenosem dat do periferie a čas potřebný pro vykonání závěrečné fáze algoritmu MD5.

Periferie pracuje synchronně s jádrem mikrokontroléru. Jedná se o iterativní návrh, v jednom hodinovém cyklu zpracovávající jeden krok algoritmu MD5. Zpracování jednoho bloku zprávy, tj. 512 bitů, vyžaduje 65 strojových cyklů – 64 pro vlastní algoritmus MD5 a jeden pro přípravu na zpracování dalšího bloku dat.

Zvolená architektura je vhodná pro menší hradlová pole. V případě dostupnosti většího množství hardwarových prostředků lze algoritmus MD5 rozbalit a vykonávat v jediném hodinovém cyklu. Autorům článku [2] se tímto způsobem podařilo dosáhnout 2,1 krát vyšší propustnosti za cenu 6,3 krát vyšších hardwarových nároků.

6 Datová propustnost

Předmětem měření je počet instrukčních cyklů strávených výpočtem hashe MD5. Měření je provedeno za pomoci časovačů jednotlivých mikrokontrolérů pro soubory velikostí 1 kB, 10 kB, 100 kB a 1 MB získané za pomoci pseudonáhodného generátoru dat. Data jsou přenášena po asynchronní sériové lince a výpočet prováděn po 64 bytových blocích. Doba strávená přenosem dat mezi PC a vývojovou deskou osazenou FPGA započtena není. Pro překlad programu pro ARM i MicroBlaze byl použit překladač GNU GCC s optimalizací -O2.

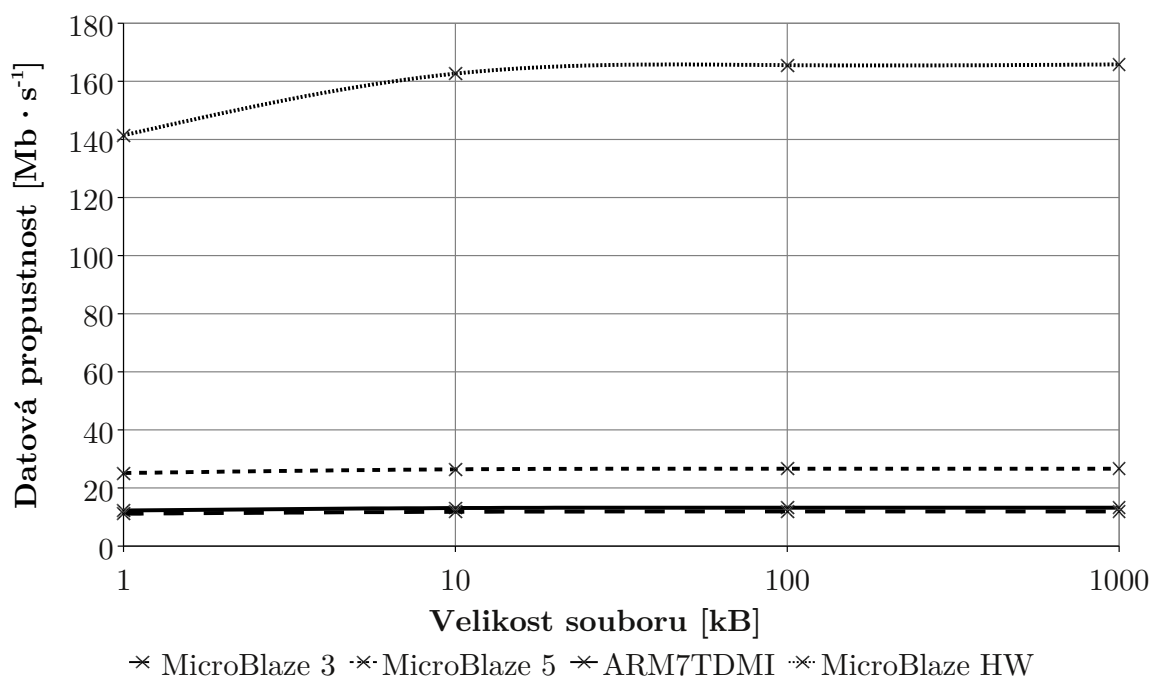
Výpočetní výkon softwarového mikrokontroléru je porovnán s mikrokontrolérem Atmel AT91SAM7S64 postaveným na jádře ARM7TDMI taktovaném 48 MHz. Testovány jsou 3 konfigurace mikrokontroléru MicroBlaze. Odlehčená varianta “MicroBlaze 3” je postavena na jádře s 3-stupňovým zřetězením instrukcí. Druhá varianta, označována jako “MicroBlaze 5”, disponuje 5-stupňovým zřetězením instrukcí, hardwarovou násobičkou MUL64, děličkou a jednotkou barrel shifter. Hodinový kmitočet této dvojice mikrokontrolérů činí 66 MHz. Poslední varianta “MicroBlaze HW” obohacuje “MicroBlaze 5” o periferii akcelerující výpočet MD5. Periferie omezuje taktovací frekvenci mikrokontroléru na 46 MHz.

Velikost souboru [kB]	Datová propustnost [$\text{Mb} \cdot \text{s}^{-1}$]			
	ARM7TDMI	MicroBlaze 3	MicroBlaze 5	MicroBlaze HW
1	12,2	11,1	25,2	141,4
10	13,1	11,8	26,4	162,7
100	13,2	11,9	26,6	165,6
1000	13,2	11,9	26,6	165,8

Tabulka 1: Datová propustnost MD5 jednotlivých mikrokontrolérů

I přes omezený hodinový kmitočet $f = 46 \text{ MHz}$ zůstává MicroBlaze s hardwarovou akcelerací výpočtu 5,4 až 6,2 krát výkonnější než “Microblaze 5”. Za povšimnutí stojí i výkonnostní rozdíl mezi konfiguracemi “MicroBlaze 3” a “MicroBlaze 5”. Je způsoben především absencí/přítomností jednotky “barrel shifter”, která akceleruje bitový posun.

U souborů malé velikosti je patrný vliv inicializace a především pak závěrečné části algoritmu MD5 spojené s paddingem.



Obrázek 2: Závislost datové propustnosti MD5 na velikosti souboru

Za předpokladu dokonalého zřetězení všech činností by bylo možno s periferií zvolené architektury dosáhnout propustnosti

$$th = \frac{b}{n} \cdot f = \frac{512}{65} \cdot 46 \cdot 10^6 = 362 \text{ Mb} \cdot \text{s}^{-1} \quad (1)$$

kde b je velikost datového bloku MD5, n počet instrukčních cyklů potřebných pro zpracování tohoto bloku a f pracovní frekvence mikrokontroléru.

Bohužel, reálná propustnost dosahuje pouze $th = 165,8 \text{ Mb} \cdot \text{s}^{-1}$. Tento výkonnostní propad je způsoben režii potřebnou pro přenos dat po sběrnici. Zde zůstává prostor pro optimalizaci. Nabízí se např. využití techniky double buffering.

7 Obsazenost hradlového pole

Periferie akcelerující výpočet MD5 realizuje množství matematicko-logických operací, obsahuje řadu registrů, konstant a pomocných obvodů. To se odráží v požadavcích kladených na zdroje FPGA. Samotná periferie obsazuje 47 % dostupných LUT tabulek použitého FPGA Xilinx XC3S400A. Systém jako celek pak 94 % LUT tabulek napříč 98 % slice jednotek. Možnosti dalšího rozšíření jsou tedy velmi omezené. Pro implementaci mikrokontroléru MicroBlaze spolu s periferií realizující jednokrokový výpočet MD5 je třeba zvolit FPGA větší kapacity.

8 Závěr

Článek ukazuje možnosti akcelerace hashovacího algoritmu MD5 a využití softwarového mikrokontroléru MicroBlaze na FPGA Xilinx XC3S400A. Systém využívající hardwarové akcelerace dosáhl datové propustnosti $165,8 \text{ Mb} \cdot \text{s}^{-1}$, což je hodnota 12,5 krát vyšší než u mikrokontroléru s jádrem ARM7TDMI taktovaném 48 MHz. Přínosem dedikované periferie pro výpočet MD5 je vedle datové propustnosti a minimalizace zátěže CPU rovněž její deterministické chování. Toto je klíčové pro systémy reálného času.

9 Poděkování

Tato publikace vznikla za podpory grantu “Podpora výzkumu moderních metod a prostředků v automatizaci” financované z Interní grantové agentury Vysokého učení technického v Brně (číslo grantu FEKT-S-11-6).

Reference

- [1] Black, J.; Cochran, M.; Highland, T.: A study of the MD5 attacks: Insights and improvements. In *Fast Software Encryption*, Springer, 2006, s. 262–277.
- [2] Deepakumara, J.; Heys, H.; Venkatesan, R.: FPGA implementation of MD5 hash algorithm. In *Electrical and Computer Engineering, 2001. Canadian Conference on*, ročník 2, IEEE, 2001, s. 919–924.
- [3] Deutsch, L.: RFC1321-based (RSA-free) MD5 library. 2009, [Online; poslední aktualizace 17. 7. 2009], <<http://libmd5-rfc.sourceforge.net/>>.
- [4] Gonzalez, I.; Gomez-Arribas, F.: Ciphering algorithms in MicroBlaze-based embedded systems. In *Computers and Digital Techniques, IEE Proceedings-*, ročník 153, IET, 2006, s. 87–92.
- [5] Jarvinen, K.; Tommiska, M.; Skytta, J.: Hardware implementation analysis of the MD5 hash algorithm. In *System Sciences, 2005. HICSS'05. Proceedings of the 38th Annual Hawaii International Conference on*, IEEE, 2005, s. 298a–298a.
- [6] Meyer, G.; Iverson, E.: FPGA MD5 Cracker. 2009, [Online; poslední aktualizace 16. 6. 2009], <<http://infinityexists.com/2009/06/16/fpga-md5-cracker/>>.
- [7] Piši, D.: Mikrokontrolér v hradlovém poli XC3S400. Technická zpráva, Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2009, vedoucí semestrální práce Ing. Pavel Kučera, Ph.D.
- [8] Rivest, R.: RFC1321: The MD5 message-digest algorithm. *RFC Editor United States*, 1992, <<http://www.ietf.org/rfc/rfc1321.txt>>.
- [9] Wang, Y.; Zhao, Q.; Jiang, L.; aj.: Ultra High Throughput Implementations for MD5 Hash Algorithm on FPGA. *High Performance Computing and Applications*, 2010: s. 433–441.
- [10] Xilinx: *MicroBlaze Processor Reference Guide*. 2011, [Online; poslední aktualizace 1. 3. 2011], <http://www.xilinx.com/support/documentation/sw_manuals/xilinx13_1/mb_ref_guide.pdf>.

SYNTÉZA REGULÁTORU ASYNCHRONNÍHO MOTORU METODOU SMÍŠENÝCH CITLIVOSTNÍCH FUNKCÍ

Lukáš POHL

Ústav Automatizace a Měřicí Techniky,
Vysoké Učení Technické v Brně
Kolejní 2906/4, 612 00 Brno
E-mail: xpohl01@stud.feec.vutbr.cz

Abstrakt:

Tento článek popisuje metodu návrhu regulátoru asynchronního motoru v d-q souřadnicovém systému. Cílem návrhu je lineární regulátor aplikovaný na nelineární model asynchronního motoru bez použití běžných linearizačních metod. Stabilita regulátoru a kvalita přechodného děje je dosažena pomocí $\mathcal{H}_\infty$ metody, která minimalizuje normu systému ze vstupu na výstup [5]. Na nelinearity v rovnicích asynchronního motoru je nahlíženo jako na neurčitosti, jejichž vliv na regulační smyčku navržený regulátor potlačuje. Připojením váhových funkcí jsou do určité míry ovlivněny vlastnosti regulačního obvodu. Tyto váhové funkce definují frekvenční charakteristiky citlivostní a komplementární citlivostní funkce.

Klíčová slova: $\mathcal{H}_\infty$, Smíšené citlivostní funkce, Robustní řízení.

1 Úvod

Metoda popsaná v tomto článku má oproti ostatním metodám řízení motorů jednu nespornou výhodu - již při návrhu regulátoru lze specifikovat jaké vnější vlivy či poruchy na regulovanou soustavu působí a výsledný regulátor tak obstojí i při změnách těchto parametrů v předem definovaném rozsahu. Obdobně lze i na ty parametry soustavy, které způsobují nelinearitu, nahlížet jako na neurčitosti za podmínky alespoň částečné znalosti jejich velikosti.

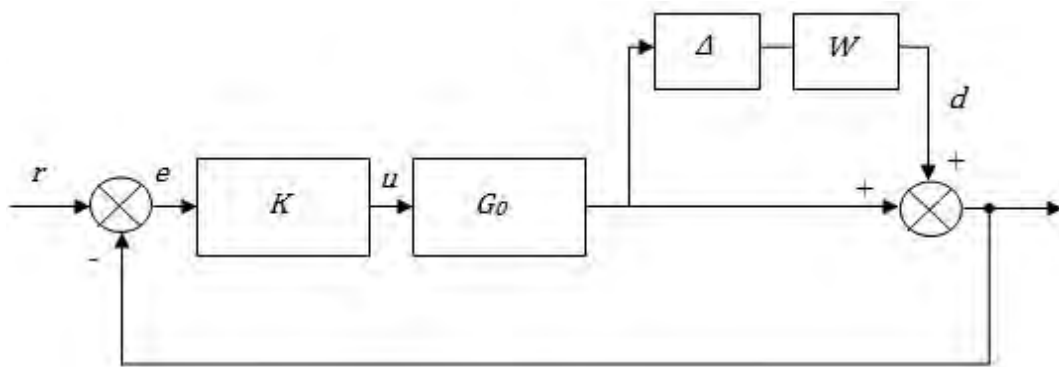
2 Návrhová metoda $\mathcal{H}_\infty$

Pro potřeby syntézy robustního regulátoru definujeme obecnou strukturu regulačního obvodu, obsahující blok Δ , který v sobě zahrnuje veškeré neurčitosti regulovaného systému. Tento blok se skládá z diagonální matice Δ , jejíž prvky Δ_i reprezentují jednotlivé zdroje nejistot. Pro všechny změny neurčitosti $\Delta_i = W\Delta$ splňující normu $\|\Delta_i\|_\infty \leq 1$ je možné obecný systém modelovat jako nestrukturovanou multiplikativní nejistotu $G = [I + W\Delta]G_0$ [2]. Kde W je stabilní přenosová funkce charakterizující frekvenční závislost neurčitosti Δ_i (Obrázek 1). Vyřazením bloku Δ lze vyjádřit přenos mezi výstupem a vstupem těchto neurčitostí. Pro výstupní multiplikativní nejistotu zahrnutou do regulačního obvodu platí následující přenos z výstupu nejistot na vstup:

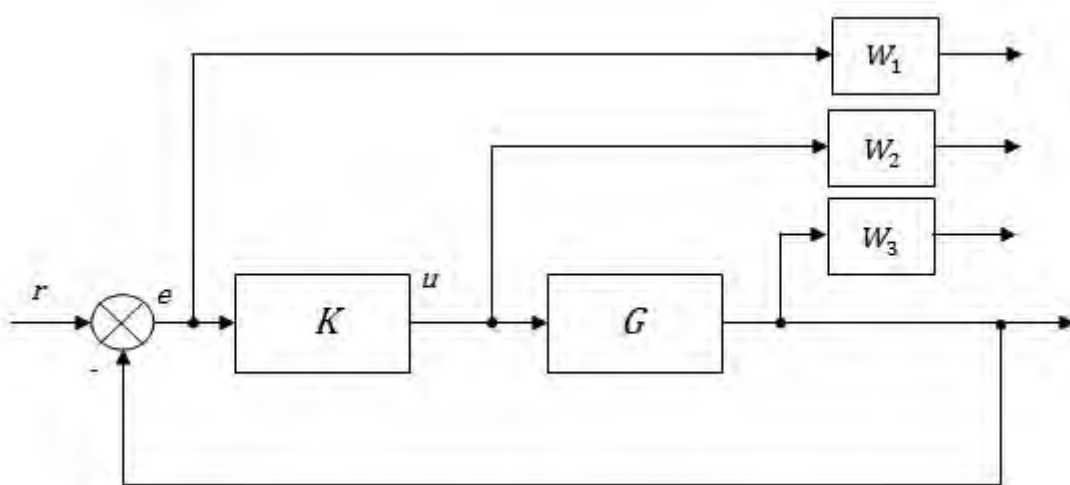
$$M = GK[I + GK]^{-1}W = T \quad (1)$$

Vhodným umístěním váhových funkcí do regulačního obvodu lze při iterativním nebo výpočetním návrhu regulátoru dosáhnout požadovaných vlastností vyjádřených těmito váhami. Pokud je dodržena následující nerovnost, bude zpětnovazební systém stabilní pro všechny Δ :

$$\|WT\|_\infty \leq 1 \quad (2)$$



Obrázek 1: Zpětnovazební systém s multiplikativní nejistotou



Obrázek 2: Struktura zpětnovazebního systému s váhovými funkcemi

Pokud je systém (Obrázek 1) bez vlivu nejistot ($\Delta = 0$) stabilní, pak lze určit nejmenší hodnotu Δ_i , která systém přivede na mez stability.

$$\bar{\sigma}(\Delta_i(j\omega)) = \frac{1}{\bar{\sigma}(T(j\omega))} \quad (3)$$

Z rovnice (3) plyne, že volba menších hodnot $\bar{\sigma}(T(j\omega))$ dovoluje větší rozptyl nejistot Δ_i a tím i větší zásobu stability. Aby byla dosažena dostatečně nízká hodnota $\bar{\sigma}(T(j\omega))$, tak se stanoví horní limit $\|T(j\omega)\|$ následovně:

$$\bar{\sigma}(T(j\omega)) \leq |W_3^{-1}(j\omega)| \quad (4)$$

Váhová funkce W_3 je šablonou specifikující návrhové podmínky na chování uzavřeného systému. Obrázek 2 zobrazuje typické propojení systému a váhových funkcí. Při použití všech váhových funkcí lze při návrhu ovlivnit kromě zásoby stability také rychlost přechodného děje a velikost akčního zásahu. Pokud některý z výše uvedených požadavků není stanoven, lze jej z návrhu vypustit.

$$\left\| \begin{array}{c} W_1 S \\ W_2 K S \\ W_3 T \end{array} \right\|_{\infty} \quad (5)$$

Rovnice (5) vyjadřuje, jakým způsobem jsou stanoveny hranice pro jednotlivé přenosy citlivostní funkce, komplementární citlivostní funkce a akčního zásahu. S, T a KS jsou na sobě vzájemně závislé, proto volba jednotlivých vah je kompromisem mezi robustní stabilitou, rychlostí a velikostí akčního zásahu. Obvykle se návrh rozdělí na jednotlivé frekvenční oblasti.

3 Návrh proudového regulátoru na asynchronní motor

Model asynchronního motoru se souřadnicovým systémem spojeným s rotorovým tokem motoru je vyjádřen těmito rovnicemi:

$$\begin{aligned}
 \frac{d\phi_{qs}}{dt} &= -\frac{R_s}{\sigma L_s} \phi_{qs} - \omega \phi_{ds} + \frac{L_m R_s}{L_s L_r - L_m^2} \phi_{qr} + u_{qs} \\
 \frac{d\phi_{ds}}{dt} &= \omega \phi_{qs} - \frac{R_s}{\sigma L_s} \phi_{ds} + \frac{L_m R_s}{L_s L_r - L_m^2} \phi_{dr} + u_{ds} \\
 \frac{d\phi_{qr}}{dt} &= \frac{L_m R_s}{L_s L_r - L_m^2} \phi_{qs} - \frac{R_r}{\sigma L_r} \phi_{qr} - \omega_s \phi_{dr} \\
 \frac{d\phi_{dr}}{dt} &= \frac{L_m R_s}{L_s L_r - L_m^2} \phi_{ds} + \omega_s \phi_{qr} - \frac{R_r}{\sigma L_r} \phi_{dr} - \frac{n_p}{J_m} T_L + \frac{3n_p^2 |\phi_r^2|}{2R_r J_m} \omega \\
 \frac{d\omega_r}{dt} &= -\frac{n_p}{J_m} \left(\frac{3n_p}{2R_r} |\phi_r| + C_1 \right) \omega_r
 \end{aligned} \tag{6}$$

S těmito parametry:

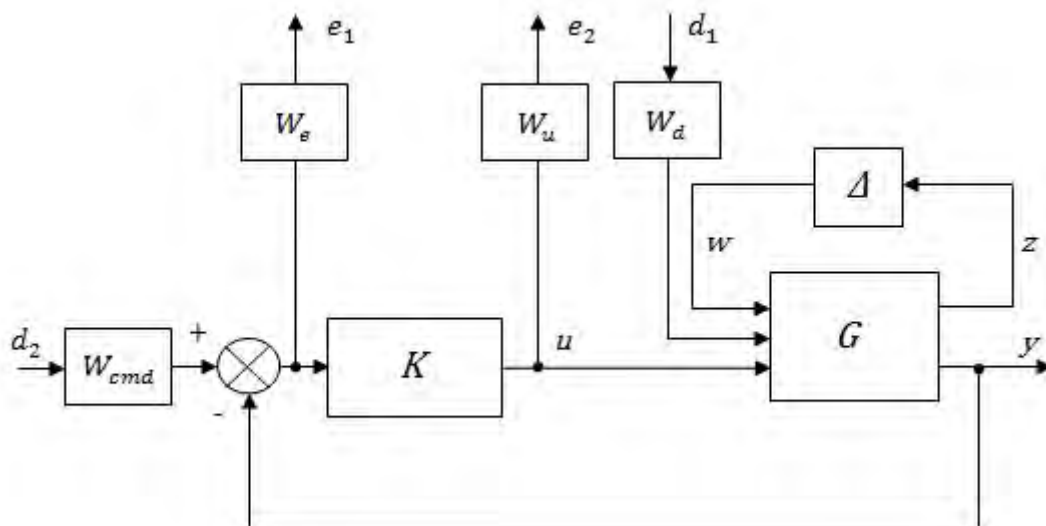
Tabulka 1: Parametry asynchronního motoru

Parametr	Velikost	Jednotky
$\sigma = 1 - \frac{L_m^2}{L_s L_r}$	0,3016	[-]
R_s	7,1	Ω
R_r	5,8	Ω
L_s	0,3105	H
L_r	0,3105	H
L_m	0,28456	mH
J_m	0,0038	kgm^2
C_1	0,0015	Nms
n_p	2	[-]

Rovnice modelu motoru (6) jsou závislé na časově proměnných parametrech ω a ω_s . Tyto parametry jsou ošetřeny tak, že se v daném pracovním rozsahu motoru vyjádří jako strukturované neurčitosti, které na model motoru působí jako vnější vliv [1]. Rozsah parametrů ω a ω_s byl zvolen $\omega = 0 - 380$ a $\omega_s = 0 - 20$. Tento parametr může být pomocí následujícího zápisu vyjádřen nominální hodnotou variací $\delta \in [-1, 1]$:

$$\begin{aligned}
 \omega &= \bar{\omega}(1 + \delta_1) \\
 \omega_s &= \bar{\omega}_s(1 + \delta_2)
 \end{aligned} \tag{7}$$

Systémová matice A je rozložena na afinní parametrický zápis:



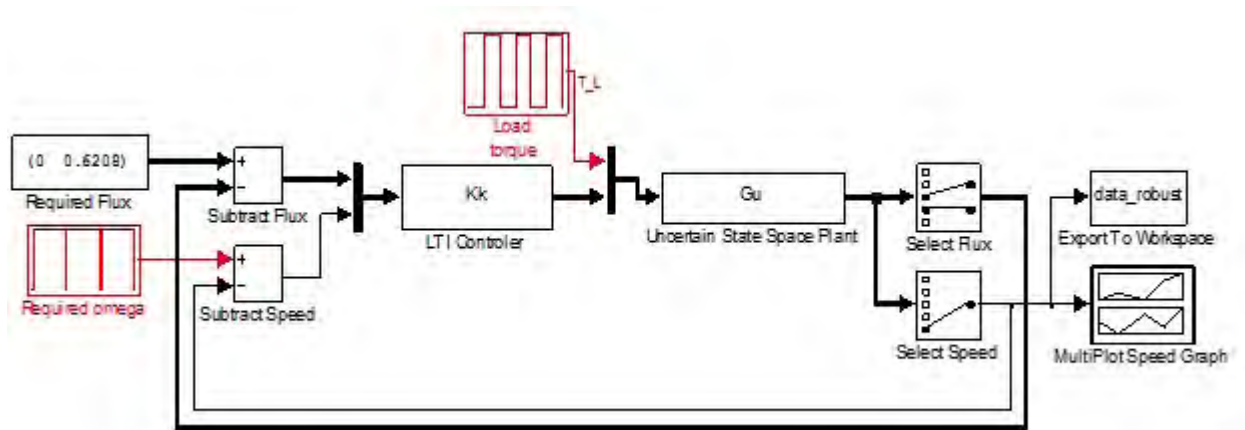
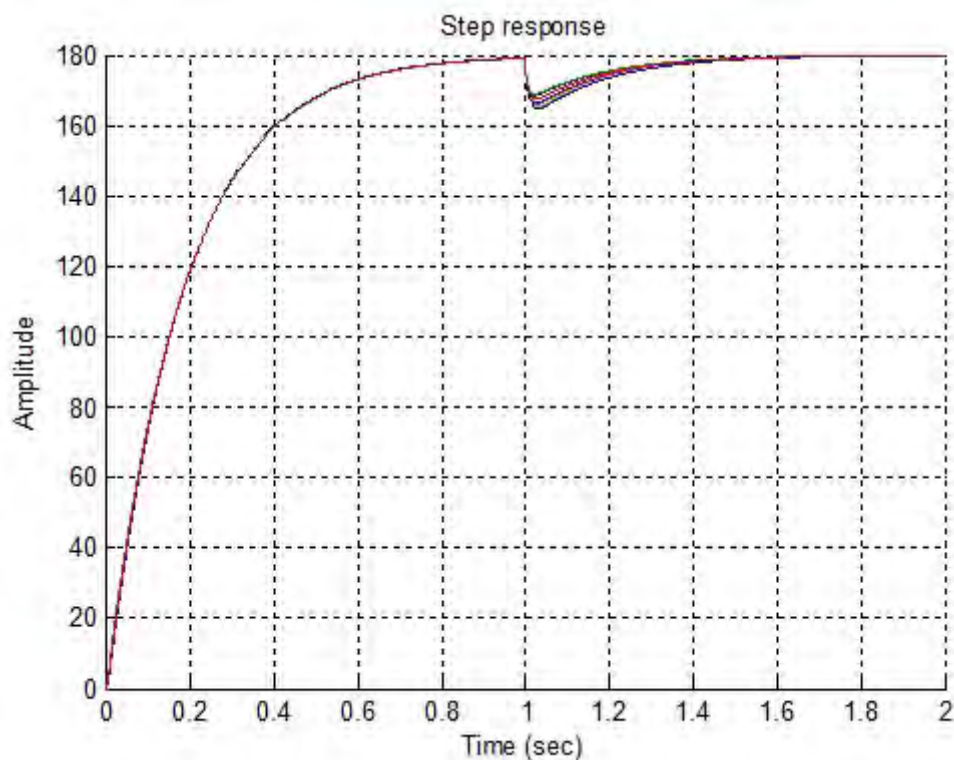
Obrázek 3: Struktura řízení asynchronního motoru s váhovými funkcemi

$$\begin{aligned}
 A &= A_0 + \delta_1 A_1 + \delta_2 A_2 \\
 A_0 &= \begin{bmatrix} -\frac{R_s}{\sigma L_s} & -\bar{\omega} & \frac{L_m R_s}{L_s L_r - L_m^2} & 0 & 0 \\ \bar{\omega} & -\frac{R_s}{\sigma L_s} & 0 & \frac{L_m R_s}{L_s L_r - L_m^2} & 0 \\ \frac{L_m R_s}{L_s L_r - L_m^2} & 0 & -\frac{R_r}{\sigma L_r} & -\bar{\omega}_s & 0 \\ 0 & \frac{L_m R_s}{L_s L_r - L_m^2} & \bar{\omega}_s & -\frac{R_r}{\sigma L_r} & 0 \\ 0 & 0 & 0 & 0 & -\frac{n_p}{J_m} \left(\frac{3n_p}{2R_r} |\phi_r| + C_1 \right) \omega_r \end{bmatrix} \\
 A_1 &= \begin{bmatrix} 0 & -\bar{\omega} & 0 & 0 & 0 \\ \bar{\omega} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \bar{\omega}_s & 0 \\ 0 & 0 & \bar{\omega}_s & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}
 \end{aligned} \tag{8}$$

Obdobným způsobem lze ze stavového popisu asynchronního motoru získat matice B, C, D závislé na jednotlivých parametrech. Rozložením matic A_1 a A_2 na součiny $E_1 G_1$ a $E_2 G_2$ získáme matematický popis závislosti řízených a měřených výstupů na vnějších (poruchových) a řídicích veličinách:

$$\begin{bmatrix} \dot{x} \\ z_1 \\ z_2 \\ y \end{bmatrix} = \begin{bmatrix} A_0 & E_1 & E_2 & B \\ G_1 & 0 & 0 & 0 \\ G_2 & 0 & 0 & 0 \\ C & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ w_1 \\ w_2 \\ u \end{bmatrix} \tag{9}$$

Systém rozšířený o váhové funkce definující požadované vlastnosti uzavřené smyčky je zobrazen na obrázku 3. Pro výpočet regulátoru byla použita $\mathcal{H}_\infty$ optimalizační metoda na rozšířené zpětnovazební zapojení (vytvořeno funkcí `sconnect` v Matlabu). Pro samotný výpočet regulátoru podle šablony definované váhovými funkcemi byla použita funkce `hinflmi` z Matlab Robust Control Toolboxu. Simulace celého kaskádního zapojení byla provedena s robustním regulátorem na řízení statorových proudů a jednoduchým statickým zesílením na řízení požadované hodnoty otáček. Žádané hodnoty byly nastaveny následovně: $\phi_{qr}^* = 0Wb$, $\phi_{dr}^* = 0.6208Wb$, $\omega_r^* = 180rad/s$.

Obrázek 4: Simulační schéma zpětnovazebního regulátoru systému G_u s neurčitostmi

Obrázek 5: Odezva asynchronního motoru na skokovou změnu požadovaných otáček

4 Závěr

Tento článek představil metodu návrhu proudového regulátoru asynchronního motoru pomocí smíšených citlivostních funkcí. Na otáčky motoru zde bylo nahlíženo jako na neznámé parametry nabývající určitého rozsahu. Tento způsob popisu matematického modelu asynchronního motoru umožnil použití jednoduchého statického zesílení jako otáčkového regulátoru motoru. Předdefinované šablony (váhové funkce) přesně vymezily oblast, v které se má pohybovat frekvenční odezva citlivostní a komplementární citlivostní funkce výsledného zpětnovazebního zapojení. Z obrázku odezvy otáček asynchronního motoru na

skokovou změnu žádané veličiny (Obrázek 5) je patrné, že všechny křivky simulace pro 20 náhodných hodnot otáček v rozsahu 0 - 380 rad/s se překrývají. V jediném momentu dojde k postřehnutelnému rozdílu pro různé otáčky a to při změně zatěžovacího momentu a jen na omezenou dobu, než dojde k opětovnému vyregulování této poruchy. Nevýhodou této metody je příliš velká konzervativnost návrhu regulátoru mající za následek pomalé sledování požadovaných otáček motoru.

Poděkování

Tento projekt byl podpořen českou grantovou agenturou GA P103/10/0647 "Intelligent Electrical Drives Predictive and Robust Control Algorithms".

Reference

- [1] BOTURA, C. P., NETO, M. F. S., FILHO, S. A. A.: Robust Speed Control of an Induction Motor: An $\mathcal{H}_\infty$ Control Theory Approach with Field Orientation and μ -Analysis, IEEE Transactions on Power Electronics, Vol. 15, No. 5, September 2000.
- [2] OLIVEIRA, V. A., TOGNETTI, E. S, SIQUEIRA, D.: Robust Controllers Enhanced With Design and Implementation Processes, IEEE Transactions On Education, vol. 49, no. 3, August 2006
- [3] GU, DA-WEI , PETKOV, PETKO HR, KONSTANTINOV, MIHAIL M.: Robust control design with MATLAB, (2005). ISBN-10: 1852339837.
- [4] POHL L.: Robustní řízení elektromechanických systémů. Diplomová práce. FEKT VUT v Brně, 2010, 63 s..
- [5] SKOGESTAD, S., POSTLETHWAITE, I.: Multivariable Feedback Control Analysis and Design, (2005). ISBN 0-470-01167-X.