ANNALS

OF THE

UNIVERSITY OF CRAIOVA

**Series: AUTOMATION, COMPUTERS,
ELECTRONICS and MECHATRONICS**

Vol. 10 (37), No. 1, 2013

ISSN 1841-0626

EDITURA UNIVERSITARIA
Craiova, 2013

ANNALS OF THE UNIVERSITY OF CRAIOVA

Series: AUTOMATION, COMPUTERS, ELECTRONICS AND MECHATRONICS

Vol. 10 (37), No. 1, 2013 **ISSN 1841-0626**

***Note*:** The **"Automation, Computers, Electronics and Mechatronics Series"** emerged from **"Electrical Engineering Series" (ISSN 1223-530X)** in 2004.

Address for correspondence:

Vladimir RĂSVAN
University of Craiova
Faculty of Automation, Computers and Electronics
Al.I. Cuza Street, No. 13
RO-200585, Craiova, Romania
Phone: +40-251-438198, Fax: +40-251-438198
Email: vrasvan@automation.ucv.ro

This issue has been published under the responsibility of Emil PETRE.

We exchange publications with similar institutions from country and from abroad

CONTENTS

New Approaches in Orthopedic Medical Education by E-Learning

E. T. Avramescu*, O. Lupescu, Gh. I. Popescu**, C. Patru**, M. Nagea****

** University of Craiova, Romania, Craiova, 13 A.I.Cuza Str. (e-mail: taina_mistico@yahoo.com)*
*** Clinical Emergency Hospital, Bucharest*

Abstract: The goal of the training programme for residency in orthopaedics in Romania is to produce orthopaedic surgeons who are technically competent and knowledgeable of the literature in the field of orthopaedic surgery. In addition, we hope to stimulate interest in solving clinical and basic science problems in the field of orthopaedic surgery. To accomplish this, the present paper offers solutions for improving resident training by using new approaches, such as an innovative e-training method that is able to provide the trainees with a range of case studies and an advanced training curriculum. These solutions are developed within the framework of a Leonardo da Vinci project Transfer of innovation, named "A Web-based E-Training Platform for Extended Human Motion Investigation in Orthopedics". The project addresses to medical professionals, proposing formation of specialists that will systematically apply the principles of medical and bioengineering sciences in finding solutions that lead to improve health condition.

The main outcome of the project is a Virtual Training & Communication Center ORTHO-eMAN for innovative education - on-line education and training material accessed via a standard web browser, which provides an integrated on-line learning environment. The e-learning platform includes a repository of training material with real clinical case studies using digital imaging and accompanying notes, an interactive multimedia database system containing full reports on patients receiving orthopedic treatment. It will be used as a method of dynamic distribution of course information, but with innovative and more interactive uses. By following this programme we hope that the residents will develop patterns of life-long learning about the field of orthopaedic surgery, as well as an interest in making contributions to this field of knowledge.

Keywords: e-learning, orthopedics, medical education, case studies, human motion investigation.

1. BACKGROUND

E-learning, or the use of ICTs in support of teaching and learning, is often mentioned in the same breath as educational renewal and the construction of the knowledge society. Nowadays the value of e-learning in medical education, especially in LLL, is well recognised and there are a lot of approaches (http://ec.europa.eu/ education/leonardo-da-vinci/ transfer_en).

To meet better the health needs of the population, significant changes take place in medical education all over the world promoting the quality of higher education as essential priority. Integration of basic and clinical sciences, integration of theory and practice, implementation of problem-based learning, e-Learning and continuing all-life learning are important part of these changes, in unison of the recommendations of the World Health Organization, the World Federation for Medical Education, the European Commission (Directorate-General for Education and Culture), the national strategy for higher education and the strategy of the Medical Universities. Brown *et al.* (2007) identifies a transformative role for e-learning, with changes to views of learning and to the nature and operation of the tertiary institutions and the tertiary system.

A growing number of reports draw attention to the need of adjustment of the offer in medical education to labour market needs and the knowledge-based society. Increasing relevance and compatibility of bachelor university programs in relation to labour market needs and changes induced by the knowledge society by introducing e-learning in higher education institutions of medicine at multiregional level are some of priorities (Casebeer *et al.* (2010)). The main problem raised is represented by the reserved attitude of medical professionals regarding the use of e-learning and mainly using it for practical issues. Most of professionals believe that medical education is more suitable for theoretical concepts (guideline, procedures, protocols, legal frameworks, patient approach). Steps in infirming this concept were made in the last year by creation of virtual patients and simulations (Valcke and De Wever (2006)).

Recent reviews of e-learning literature in diverse medical education contexts reveal that e-learning is at least as good as, if not better than, traditional instructor-led methods (lectures) in contributing to demonstrated learning. Gibbons and Fairweather (1998) cite several studies, including two meta-analyses, which compared the utility of computer-based instruction to traditional teaching methods. Chumley-Jones *et al.* (2002) reviewed

76 studies on the utility of web-based learning from the medical, nursing, and dental literature, demonstrating evidence for more efficient learning via web-based instruction. Similar findings were reported by Schopf and Flytkjaer (2011) that described the benefits of online learning in the health sector and by Huckstadt and Hayes (2005) that mentioned online learning as a successful method in the education of advanced practice nurses.

The present paper presents a LDV/TOI project named "A Web-based E-Training Platform for Extended Human Motion Investigation in Orthopedics", coordinated by the University of Craiova, Romania. The project consortium is formed by higher education institutions (2 universities from Romania and Greece), 2 research centres (from Greece and Spain), an emergency hospital (from Romania).

The goal of our project is to encourage advanced orthopaedic care through achieving a greater understanding of musculo-skeletal conditions and treatments. Our aim is to develop new innovations in orthopaedic treatment, explore applications for new technologies, and enhance the results of surgical procedures. We also impact surgical outcomes through the design, assessment, and evaluation of knowledge and therapies.

2. METHODOLOGY

The main result of our project is a Virtual Training & Communication Centre ORTHO-eMAN for innovative education - on-line education and training material, accessed via a standard web browser, which provides an integrated on-line learning environment. It will be used as a method of dynamic distribution of course information, but with innovative and more interactive uses, by including a lifelong learning platform consisting of three main components: (a) E-Learning, (b) E-Communicating and (c) E-Mentoring. The platform provides a repository of training material with real clinical case studies using digital imaging and accompanying notes, an interactive multimedia database system containing full reports on patients receiving orthopaedic treatment. The learners have control over content, learning sequence, pace of learning, time, and often media, allowing them to tailor their experiences and to meet their personal learning objectives. Interactivity allows trainees to test their knowledge and provides immediate feedback using images and cases that they could encounter in clinical practice. We developed an innovative e-training method that is able to provide the trainers with a range of case studies.

The target groups that will be reached during the life of the project, are:

- In terms of persons: medical, physiotherapy and bioengineering trainees in different types of vocational training, specialists in orthopaedic practice (CME by postgraduate courses), physiotherapists and biomedical engineers involved in implants development and manufacturing and movement analysis, employees and managers in medical institutions, research entities, orthopaedic device manufacturers and marketing, educational staff in heath compartment;
- In terms of organizations (target sectors): universities, public and private medical institutions, manufacturing, import and distribution companies and associations of healthcare technologies, stakeholders, EU structures in the field (education and health).

In this way the project benefits three different communities of users: Universities and University Associations, Enterprises and Enterprise Associations and trainees and their associations.

The target groups will be reached by intensive dissemination by various modalities and by using the partner's professional networking.

Foreseen benefits (impact) for the target group are:

- *increased level of knowledge and skills* of the trainees by new training content and curriculum elaborated by professional educational institutions with regard to the existing needs;
- *increased flexibility, attractiveness and accessibility of training* by providing easy usable teaching and a learning tool suitable for lifelong learning;
- *offering equal chances to knowledge* by developing an e-learning platform accessible for professionals in different locations and from different partner countries;
- a higher degree of *efficiency* in terms of easy access and the ability to put into practice by physicians of newly available information. Due to increased level of interactivity owned by "e learning" (online courses, etc.) professional development via the Internet proved, compared to traditional methods of education;
- *economic impact***,** lowering the costs of training (travel/accommodation in university centres);
- *increasing virtual and real mobility and improvement of language skills* by the use of a multilingual e-learning platform;
- *improvement of VET systems* in bioengineering and orthopaedic sector in partner countries; especially in Romania a few degree programs use e-learning programs and emphasize applications in medicine and health care. There is a paucity in online learning opportunities, it is quite novel and not widely spread in Romania, it would make education and information more accessible without the need to travel to certain institutes, and it could also be used anytime, not only on the rare training courses once or twice a year, not to mention it would be free which is also a considerable fact;
- *harmonisation in higher education* for training curricula, learning methods, communication systems and specific language, *as first steps in the development of an Educational Network in Orthopaedic Medicine*;

Main methodological / didactic approaches:

- defining the target groups in terms of structure and number, identifying and analysing user demands and e-learning needs for LLL and EBM;
- defining on adaptation of e-learning platform (feasibility of transfer and methodology in transfer of results);
- defining a new platform development (technical and didactic teams, teaching level, specific content of each module, multimedia approaches as text, graphics, animation, audio, video, etc, evaluation system of the trainees);
- defining the alternative applications of on line learning (new approaches/adaptations, virtual classrooms; measuring added benefits);
- elaboration of the new e-learning platform and training modules; integration of modules within the platform (training packages);
- on line training; assessment; certification;
- integration in national and/or sectoral training systems as part of the resident training or/and continuous training.

The partnership's medical practitioners revised existing materials and created new training modules (best practice) delivered via the internet. The technical teams took care of adaptation of the platform and integration of training materials. Tutoring for trainers and trainees is foreseen.

The platform and training materials were first developed as working versions/prototypes. Pilot testing of the prototype by the users took place by selecting a test group from the target group that we address to. The test group evaluated if the prototype respond to the user's needs and correspond to quality requirements (level of knowledge, user interface, language level, graphical approach, level of interaction) by questionnaires. Feedback from the users was gathered by a trialling team with the use of a Trial Checklist in order to assess the materials against quality criteria and to identify potential changes that will be incorporated into subsequent forms. Evaluation of products effectiveness and impact (expectations and customer satisfaction questionnaire for the target group) and evaluation of project usability and transferability by questionnaires addresses to end users, stakeholders and open public are foreseen, as well as creation of a Focus Group.

Technologies from another project that has been running separately in recent years were used in an integrated way. The previous project entitled "e-MedI – Virtual Medical School* had as main result a web-based training environment, utilizing a multimodal breast imaging paradigm. The e-MedI architecture was based on a client-server 3-tier architecture that consists of the following core elements:

A. Learning Management System (LMS);

B. Visual Authoring Tool;

C. Trainee's interactive e-training environment;

D. Keyword-guided Clinical Case search tool.

After analysing the feasibility of transfer we decided that the e-MedI platform can be modified and adapted to our objective, an extended human motion investigation for orthopaedics (ORTHO-eMAN), as shown in Figure 1. The LMS is by definition a complex administrative system used to deliver electronic content in the form of lessons and to organize people who attend these lessons and the LMS core can be used in the same form for ORTHO-eMAN. These requirements led the ORTHO-eMAN project to the decision to implement the Authoring Tool as a plug-in of the LMS. The chosen LMS was Moodle and the development effort was led to the ORTHO-eMAN plug-in.

The plugin consists of 3 parts:

- plugin-core: handles the integration with the Moodle (authentication, data. grading), written in PHP;
- Authoring Tool: handles creation and editing of the content. It is a GWT application meaning that it compiles to JavaScript and runs the client side in the user's browser. It calls out the plug-in-core for services regarding Moodle integration;
- Display Tool: handles the presentation of the content to the trainees. It calls out to plugin-core for services regarding Moodle integration.

The presentation tier is adapted to include additional data provided by specific modern investigation methods of ORTHO-eMAN, including medical imaging, video files of motion analysis, force graphs, muscle and joint reactions, numerical data, contact pressure diagrams, etc.

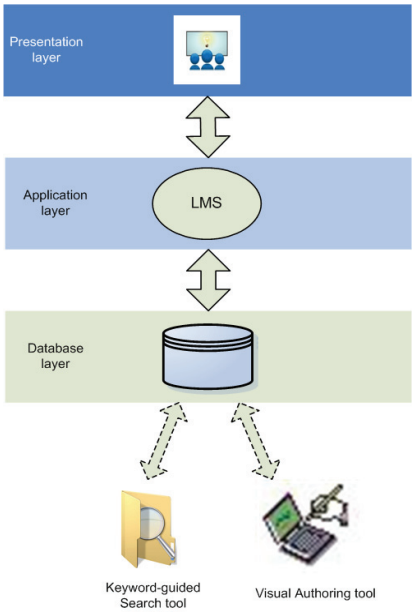


Fig. 1. Client-server 3-tier architecture used in the development of ORTHO-eMAN platform.

Some added values of this project are:

- Development of interactive training modules by using out-line courses presentations and on-line courses on theoretical issues, case studies (patient medical background, clinical signs, paraclinical evaluations, X-Rays and MRIs, clinical tests and valid biomechanical tests- clinical gait analysis). Each case study indicates the correct diagnosis for trainees (interactive feedback).
- Using of an Authoring Tool that creates case studies consisting of patient's general information and a set of consecutive stages Each Stage is a data point in patient's diagnostic time line. And is associated with one or more visual content objects (i.e. 2D, 3D images and video sequence), which are the acquisition outputs of one of the available modalities (MRI, Tomography, video processing, pressure pattern). The theoretical studies include the most important information regarding the anatomical and biomechanical aspects of human motion; the case studies were chosen due to the fact that there is a complete gap within orthopaedic training nowadays, concerning imagistic diagnosis. Although the residents are theoretically taught the specific elements of skeletal pathology, very little is done for testing their ability to evaluate the outcome of orthopaedic treatment. That is why the cases are presented with their Xrays/ CTs/MRIs and motion analysis so as the trainee should identify the most important imagistic findings, correlate them with motion analysis and establish a proper treatment

The Authoring Tool creates a XML description of the whole course and it is used by the Display Tool to reconstruct the course and present it to the trainees. The XML document is stored along with the other resources in the Moodle database. Regarding its main features, the Authoring Tool itself is capable of creating multimedia content by letting the users upload images and video. The teacher can then create quizzes of the following type:

- Find Region Of Interest (ROI). Multiple ROI may be requested by the teacher. This is important for the trainees because it improves their ability to evaluate an image as a whole and find the pathologic findings. For example, the trainee is asked to indicate the mechanical problems within a certain type of osteosynthesis, which requires a proper understanding of the biomechanics of the initial injury and knowledge regarding the correct treatment , otherwise the trainee will not be able to identify the pitfalls. Having that performed will definitely generate two effects: first of all , a " don't do like that" alarm sign which will be activated when the resident will have to treat a similar injury by himself; secondly, the trainee must establish a correct treatment in a very real situation, in a defined amount of time , as in daily medical practice. (Fig. 2)
- Quiz: Multiple answer quiz. (None, one, or more answers can be correct). The questions will be similar to those that the practitioner has to answer to when treating the patient, so they have a very clear practical character; although theoretical knowledge is strongly needed, the trainee will not be able to answer the quiz unless they integrate theory and apply it to each particular case, as it happens in reality in orthopaedics. This very strong character of reality-reflecting gives value to our system since little has been done until now for "simulating" orthopaedic patients and apply the principle of interactivity in orthopaedic training.

Fig. 2. Identifying the regions with pathologic elements.

- Range Quiz: A question that has a number for an answer. The teacher specifies the minimum and maximum allowed value for the answer to be considered as correct.

The Authoring Tool also has some standard image processing tools such as brightness, contrast, inverse that are heavily used in X-Ray setups. Angle calculation and cross-hair tool were added specifically for the ORTHO-eMAN project.

Angle calculation was considered of particular interest since treatment of deformities in skeletal pathology starts with their angular evaluation which also rules their treatment; different angles mean different treatment, so this e-learning tool gives the trainee the possibility to evaluate his ability of finding and calculating the angles unanimously accepted as significant (Fig. 3).

In the following picture (Fig. 4) the Authoring Tool is depicted as part of a Moodle installation.

The Display Tool supports several tools that were not available in the previous Flash implementation

- back and forward functionality
- tracking (trainees' progress)

The display tool is implemented in HTML5 language and analytically, it adheres to the user interface skeuomorphism design guidelines and principles. Therefore, the display tool emulates for each lesson the function of the book. Moreover, each page corresponds to each containing lesson case (Fig. 5).

Fig. 3. Calculating the angles.

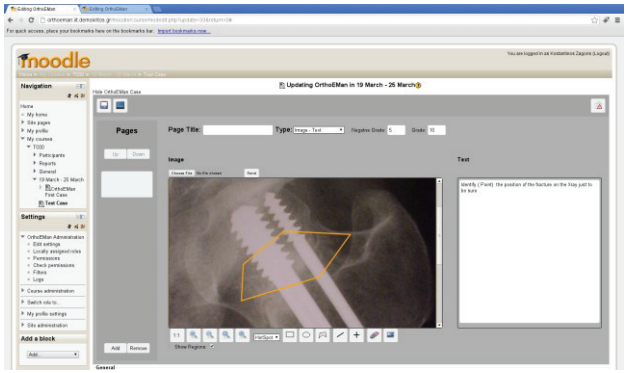


Fig. 4. Authoring Tool fully integrated to Moodle LMS: ROI - Image – Text.

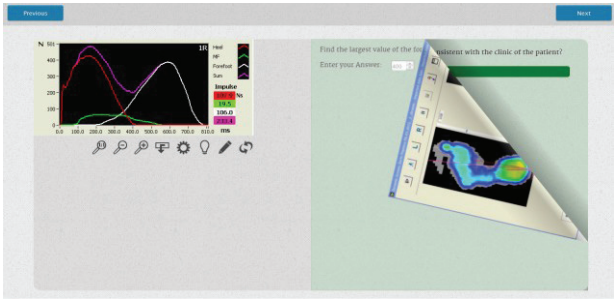


Fig. 5. Display Tool.

The original scientific issue of this approach consits mainly in the fact that the broad-based educational tool was developed using evidence-based medicine and new biomechanical technologies related to real-life scenarios that are relevant for the user groups - residents and specialists in orthopaedic practice. In this way the e-learning platform developed by us does not deal only with a technological problem for trainers in defining a medical case study but aimes to develop strong conceptual skills, helping the trainee to think creatively and solve problems.

3. RESULTS

Medical doctors will follow pressure plate case studies in relation to orthopaedic case studies. By using the integrated learning environment and making use of a fair number of real case studies, the medical trainees will be able to identify, classify, diagnose and propose the appropriate action or treatment, identify the risk zones, appreciate the efficiency of the treatment or rehabilitation programme.

The e-training environment is multilingual (English, Greek, Romanian and Spanish languages), as shown in Fig. 6.

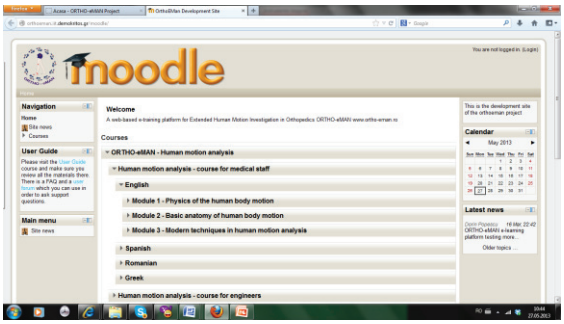


Fig. 6. The structure of the Human Motion Analysis Course.

The case studies include measurements recorded for persons without known orthopaedic pathology and measurements recorded for persons with gait abnormalities. Measurements refer to clinical tests and clinical gait analysis and are analysed before and after surgery and/or rehabilitation. The trainees must use the system specifications that allow comparison of measurements. In Fig. 7 is shown the start of a case study and in Fig. 8 an example of case study development.

The Authoring Tool provides simple tools to draw region of interest like rectangles ellipses and polygons to mark the region of interest. Our implemented features within the interactive e-learning platform for pressure plate (Footscan) measurements allow the trainee to:

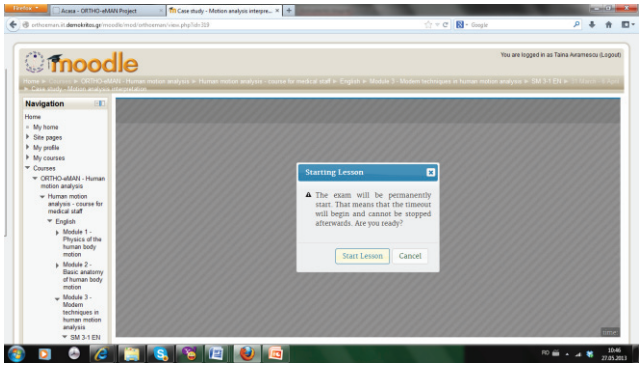


Fig. 7. Starting lesson - case study.

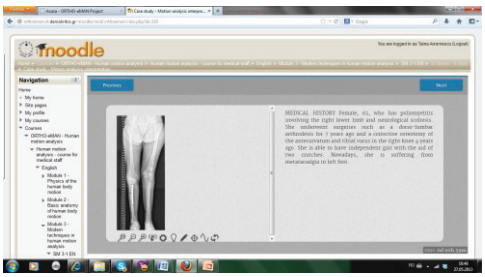


Fig. 8. An example of case study development – clinical analysis.

- select a point in the image by mouse click or designate an area also by employing the mouse. This will be used for example to indicate zones of abnormal high or low pressures, identify the centre of pressure, the highest impulse area, the contact percentage, foot axis or foot angles;
- measure angles;
- draw a horizontal and a vertical line on the graph image; both lines form a crosshair-like object that will allow to target any desired point on the graph and estimate it's values for x and y axes;
- in some cases, text boxes provide the trainee a way of inputting measured or estimated graph values.

Some cases include images in motion, which are actual recordings of pressures exerted on patient's feet during roll-off. The e-training platform offers basic movie player controls such as Play/Pause/Stop functionalities and Rewind. It also provides a way to play the video material regarding the movement for both feet. The trainee will have to watch the video (the evolution in time of the stepping process) and answer a quiz regarding abnormal features of the movement.

4. CONCLUSIONS

Our course in Human Motion Analysis is developed in order to help the trainee to understand and solve case studies regarding data acquisition and analysis of human motion using specific devices and software as pressure plate and video motion analysis by Simi Motion. The broad-based educational tool was developed using evidence-based medicine and new biomechanical technologies related to real-life scenarios that are relevant for the user groups - residents and specialists in orthopedic practice.

The developed e-training platform was designed to support e-learning, to manage access to e-learning materials, consensus on technical standardization, methods for peer review of these resources. Our solution presents numerous research opportunities for the target group, along with continuing challenges for professional development and it combines the design and problem solving skills of engineering with medical sciences, improving healthcare diagnosis, monitoring and therapy.

The proposed e-learning methodology for achieving these tasks is a viable alternative, as it motivates the development of a number of on-line learning/training environments.

Achievement of these aspects allows a close cooperation, by working in a team, between doctors and biomedical engineers for diagnostic and treatment targeting.

The integration of e-learning into graduate and continuing medical education will promote a shift toward adult learning in medical education. Integrating the results into national systems - mainstreaming of project results into education systems is previewed as the best solution to guarantee that the project has a long-term impact. Decision-makers will be regularly addressed by a coherent advocacy strategy and the project outcomes will be developed in order to be easily adapted for use in other contexts. The project would be supported and continued by the recognition of training courses and certification on national level for all participant countries as mentioned in WP6. We foresee the following ways:

- Medical education: residents - this e-learning platform could be part of the practical, where we could introduce surgical videos with explanations during class, the description of surgical techniques/ rehabilitation protocols could be considered extracurricular activities for extra credits.
- Postgraduate: Trainees in orthopaedics have certain mandatory courses they must attend, and there is one called "Surgical Techniques", and one called "Rehabilitation", both are 1 week long courses, the e-learning platform could be added to the teaching material, or alternatively, it could serve as an extra help for the trainees to be able to complete these courses.
- Orthopaedic surgeons: they must revalidate their license in every five years, and in order to do that, they must attend courses/congresses/give lectures etc. to get the required amount of credit. The e-learning platform could be accredited by university, and serve as an online teaching tool, and allow surgeons to learn/get extra credits whenever it suits them.

ACKNOWLEDGMENT

This work is supported by LLP-LdV-ToI-2011-RO-008 grant. ORTHO-eMAN ("A Web-based E-Training Platform for Extended Human Motion Investigation in Orthopedics") is a two year European funded project 2011-1-RO1-LEO05-15321 (Contract LLP-LdV/ToI/ 2011/RO/008). The project has been funded with support from the European Commission. This publication reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

REFERENCES

Brown, M., Anderson, B., and Murray, F. (2007). E-learning policy issues: Global trends, themes and tensions, *Proceedings ascilite Singapore*, pp. 75-81.

Casebeer, L., Brown, J., Roepke, N., Grimes, C., Henson, B., Palmore, R., Granstaff, U., and Salinas, G.D. (2010). Evidence-based choices of physicians: a comparative analysis of physicians participating in Internet CME and non-participants, *BMC Medical Education* (www.biomedcentral. com).

Chumley-Jones, H.S., Dobbie, A., and Alford, C.L. (2002). Web-based learning: sound educational method or hype? A review of the evaluation literature, *Acad Med.*, vol. 77, Suppl. 10, pp. 86-93.

Gibbons, A.S., and Fairweather, P.G. (1998). *Computer-Based Instruction: Design and Development*. Englewood Cliffs, NJ: Educational Technology Publications.

http://ec.europa.eu/education/leonardo-da-vinci/transfer_en.

Huckstadt, A. and Hayes, K. (2005). Evaluation of interactive online courses for advanced practice nurses, *J. Am. Acad. Nurse Pract.*, vol. 17, no. 3, pp. 85-89.

Schopf, T., and Flytkjær, V. (2011). Doctors and nurses benefit from interprofessional online education in dermatology, *BMC Med. Educ.*, no. 11, p. 84, Oct. 14.

Valcke, M., and De Wever, B. (2006). Information and communication technologies in higher education: evidence-based practices in medical education", *Med. Teach.*, 28, 1, pp. 40-48.

Modelling of the Upper Limb Wearable Exercisers

Beniamin Chetran*, Simona Noveanu,**
Gelu Răducanu*, Olimpiu Tătar****,**
Dan Mândru***

Department of Mechatronics and Machine Dynamics,
Faculty of Mechanical Engineering, Tehnical University of Cluj-Napoca, Romania
** e-mail: Beniamin.Chetran@mdm.utcluj.ro*
*** e-mail: Simona.Noveanu@mdm.utcluj.ro*
***** e-mail: Olimpiu.Tatar@mdm.utcluj.ro*
****** e-mail: Dan.Mandru@mdm.utcluj.ro*

Abstract: Robot-assisted rehabilitation procedures have clear advantages, already recognized. The researches in this field often involve the development of models of the upper limb and robotic system. In this paper, some models of portable exoskeleton-type rehabilitation robotic systems are emphasized and an original MATLAB model of upper limb – wearable exerciser ensemble, is given. The results of simulations for different combination of anatomical movements are presented and used for actuators selection.

Keywords: Physical therapy, rehabilitation, exerciser, modeling, exoskeleton, upper limb

1. INTRODUCTION

The wearable rehabilitation robotic exoskeletons become useful tools for more efficient rehabilitation processes, being able to repeat with the patients large range movements which would otherwise be done by the therapists. Moreover, a useful feedback to the therapists is provided, based on precise recording of forces and movements. It has shown that this type of robot-assisted rehabilitation therapy allows objective assessment for recovery process, improves morale and motivation, promotes home exercises and training, Rui et al. (2011).

Usually, the wearable rehabilitation robotic exoskeletons have a serial architecture in accordance with upper limb kinematics. Much research in this field refers most often to the modelling of the robotic systems, useful in the design of their mechanical structure, actuation and control systems. The kinematics of the 5 DOF exoskeleton rehabilitation robot for upper limb described in Qinling et al. (2009), is analyzed with a model based on Denavit-Hartenberg (D-H) method, in order to conceive a two-stage (passive and active-resistant) rehabilitation procedure. The same D-H method was applied in Wu et al. (2011) for the kinematic model of upper limb – elbow exoskeleton ensemble. In addition, an analytical model of joint torques was developed for design of the elastic elements which provide necessary resistance for shoulder and elbow movements. A 2 DOF wrist exerciser has been kinematically modelled using modified D-H notations, Rahman et al. (2010). A dynamic simulation is also proposed, with a nonlinear sliding mode control technique. D-H method is used for the positional analysis of a 4 DOF shoulder and elbow exoskeleton, while the Lagrange-Euler method is used for its dynamic analysis; the human arm and rehabilitation robot were modelled in MATLAB SimMechanics, Parasuraman et al. (2009). Specific model is proposed in Mustafa et al. (2005) for the 3 DOF shoulder module of biologically-inspired anthropocentric exerciser which consists of cable-driven parallel mechanisms.

Browsing the references it results that the wearable robotic exoskeletons for functional compensation provide the required mechanical power being controlled through the signals provided by the users. For active-resistant exercises, the exoskeletons have more complex functions and different control methods, as well as those for interactive exercises, which react to user's command to provide the desired assistance. A précis control of velocity and torque is given by the actuation systems based on DC motors. Besides these, other actuators were implemented in wearable robotic exoskeletons: pneumatic actuators, McKibben artificial muscles, electroactive polymers, etc., considering a balance between power and weight.

2. THE PROPOSED MATLAB MODEL

In our previous works, a 7 DOF kinematic model of upper limb was developed, Mandru et al. (2012) for which have been obtained the equations for both forward and inverse kinematics. Also, a model for a wearable exoskeleton robotic exerciser with 3 DOF (shoulder, elbow and wrist flexion-extension) was developed. It gives the opportunity to deduce the forward and inverse kinematics equations for above-mentioned exerciser.

The aim of the proposed MATLAB model is to obtain useful information about kinematic and dynamic parameters specific to a robotic exoskeleton exerciser attached to the upper limb, during different movements, in order to establish the final design of the exerciser, including its actuators, links, brakes, clutches and control system. At the beginning the upper limb model was developed by using an imported 3D CAD mannequin (Fig. 1), The MATLAB reference coordinate system was attached to the trunk. For each segment of the upper limb, the characteristics like length, centre of gravity and so on

were set in accordance with the anthropomorphic data, Drillis et al. (1964).

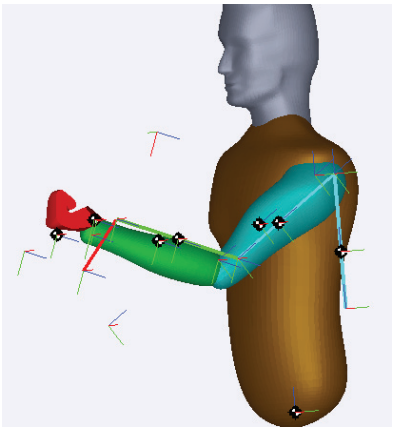


Fig. 1.The exoskeleton attached on the upper limb.

The upper limb model is connected with the exoskeleton device at the level of each segment of the limb: arm, forearm and hand ensuring the coincidence between the axes of the robotic system and biomechanical ones. Both upper limb and exoskeleton models consist of bodies corresponding to each segment connected by revolute joints with imposed constrains. The actuators corresponding to upper limb were deactivated; instead, actuators for each joint of the exoskeleton were active. Fig. 2 shows the block diagram of the developed model.

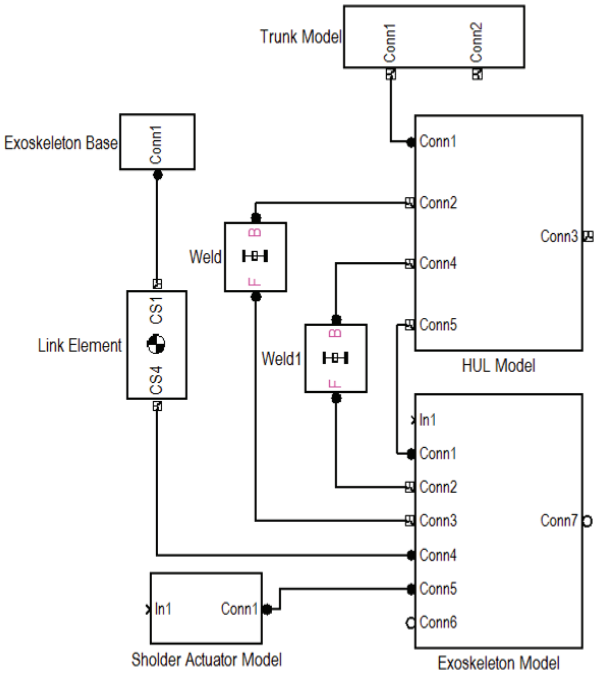


Fig. 2. The developed MATLAB model of the exoskeleton attached to the upper limb.

For forward kinematics of a 3DOF exoskeleton presented in Fig. 3, the θ_1, θ_2, θ_3, variables are assumed as known. The result is the position O_{30} and orientation $R_{3,0}(\theta_1,\theta_2,\theta_3)$ expressions of the end element attach by the hand in respect with the main coordinate system attached on the shoulder joint $O_0x_0y_0z_0$ which is considerate stationary.

The homogeneous transformation matrix $T_{3,0}$ is determinate with the following expression:

$$T_{3,0}=A_{10}\ A_{21}\ A_{32}= \\ =\begin{bmatrix} \theta(R_{3,0}\ (\theta_1,\theta_2,\theta_3)) & O_{3,0}\ (\theta_1,\theta_2,\theta_3) \\ 0 & 1 \end{bmatrix}= \quad (1)$$

$$=\begin{bmatrix} c_{123} & -s_{123} & 0 & l_{1c_1}+l_{2c_{12}}+l_{3c_{123}} \\ s_{123} & c_{123} & 0 & l_{1c_1}+l_{2c_{12}}+l_{3c_{123}} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

where $s_1=\sin\theta_1$, $c_1=\cos\theta_1$, $s_{12}=\sin(\theta_1+\theta_2)$, $c_{12}=\cos(\theta_1+\theta_2)$, $s_{123}=\sin(\theta_1+\theta_2+\theta_3)$ and $c_{123}=c\,\mathrm{os}(\theta_1+\theta_2+\theta_3)$

Nine parameters are describing the end element orientation but the exoskeleton mechanism is planary the proposed mechanism orientation will have the expression in respect with only one parameter noted with Θ:

$$R_{3,0}(\theta_1,\theta_2,\theta_3)=R_{z,(\theta_1,\theta_2,\theta_3)}=R_{z,\Theta} \\ \Theta=\theta_1+\theta_2+\theta_3 \quad (2)$$

Position O_{30} of the exoskeletons end element is determined from (1) being the terms from the fourth column.

$$O_{30}=\begin{bmatrix} x_3 \\ y_3 \\ z_3 \end{bmatrix}=\begin{bmatrix} l_1c_1+l_2c_{12}+l_3c_{123} \\ l_1s_1+l_2s_{12}+l_3s_{123} \\ 0 \end{bmatrix} \quad (3)$$

The position and orientation of the end element have the following expression:

$$\begin{cases} x_3=l_1c_1+l_2c_{12}+l_3c_{123} \\ y_3=l_1s_1+l_2s_{12}+l_3s_{123} \\ \Theta=\theta_1+\theta_2+\theta_3 \end{cases} \quad (4)$$

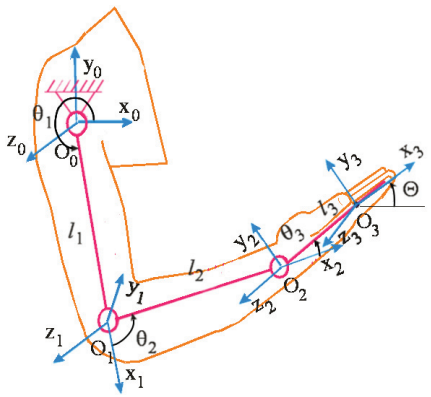
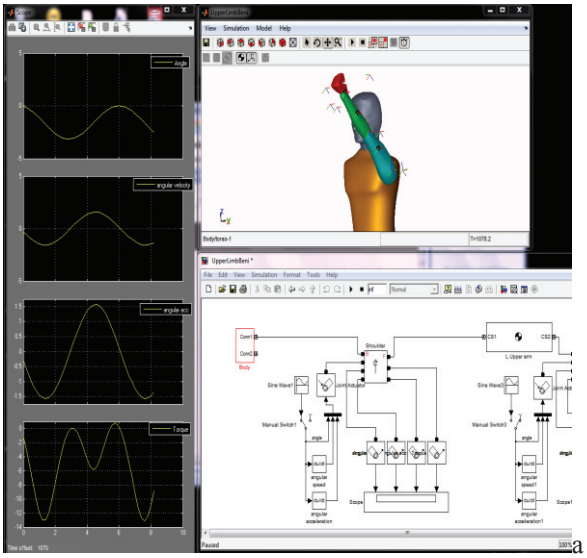
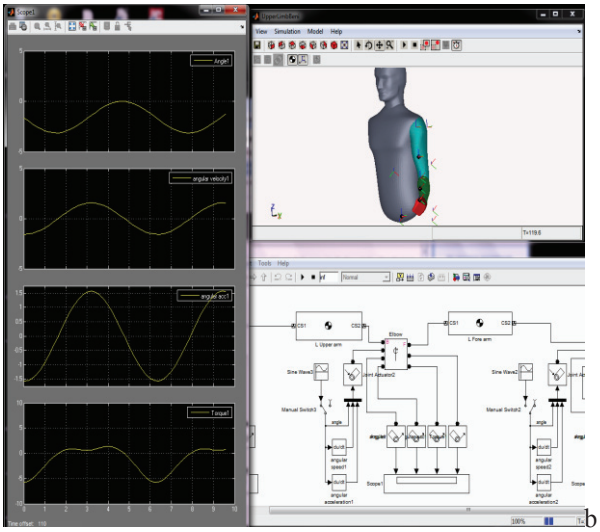
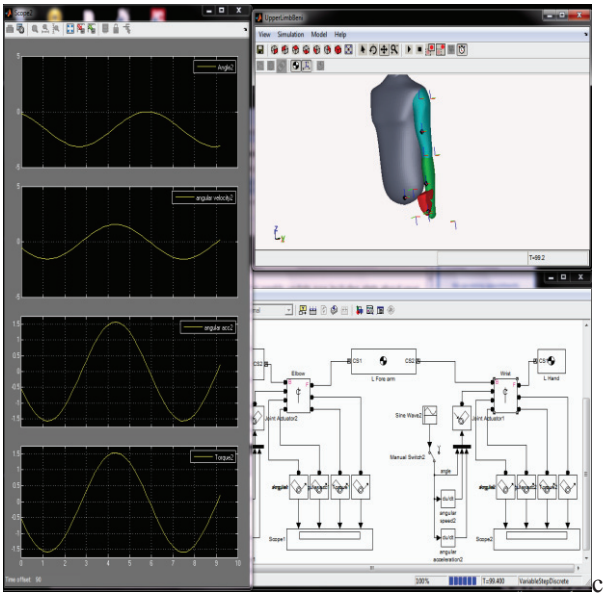


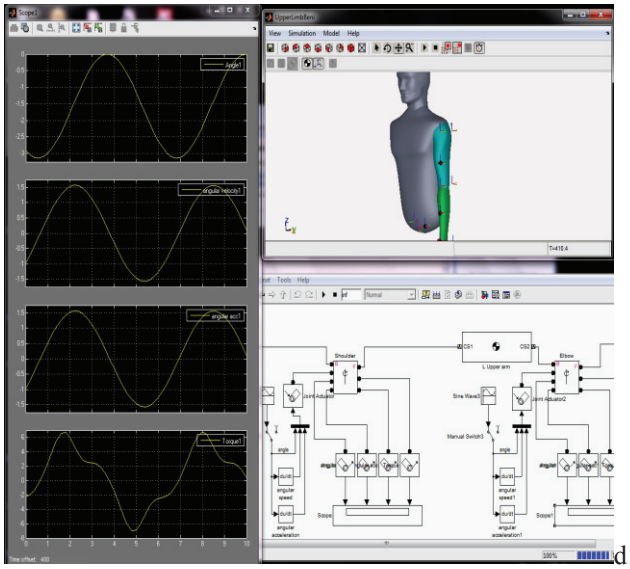
Fig. 3. Kinematic analysis diagram for a 3 DOF exoskeleton

Various simulations for different combination of anatomical movements were performed (Fig. 4). Angular velocity, angular acceleration and torque are computed while the amplitude of movement is imposed.


a


b


c


d

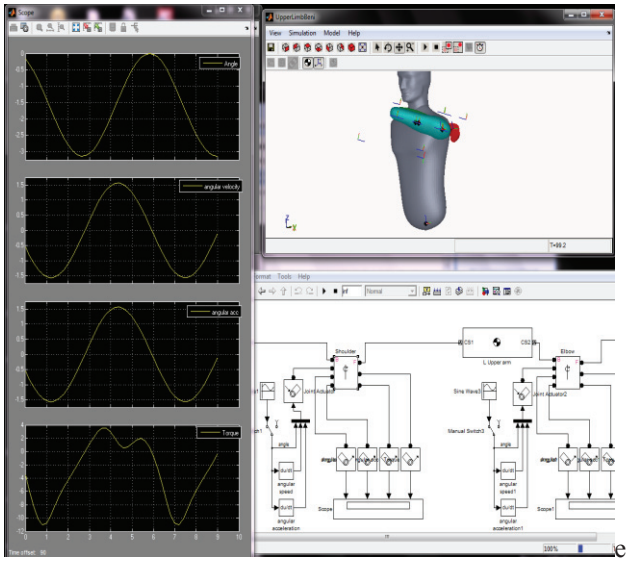

e

Fig. 4. Simulations for different combination of limb movements: a -shoulder joint, forward/backward projection of the arm; b - elbow joint, flexion/extension of the forearm; c - wrist joint, flexion/extension of the hand; d - shoulder angular velocity, acceleration and torque when the shoulder forward/backward projection and elbow flexion/extension movements are combined; e - shoulder forward/backward projection; elbow flexion/ extension and wrist flexion/extension movements are combined.

The biggest values of the required actuator torque for passive movements of the arm in forward/backward projection and adduction/abduction are around 35 Nm, except the starting spike (Fig. 5). These values are determined by the weights of the limb segments and weight of the exoskeleton (the considered weights if its modules are: 2 kg for upper arm; 2 kg for forearm and 0.46 kg for the hand). In a first approach, the maximum values of the required torques are used for actuator selection for prototype development. A few criteria for the selection of the actuators are following: must provide the required power for passive movements; high power to

weight ratio; must be able to develop large angular displacements; the energy storage devices must ensure certain autonomy. The electrical actuators (DC, stepper, brushless, synchronous and asynchronous) seem to be the most appropriate but the pneumatic and hydraulic actuators represent a viable alternative

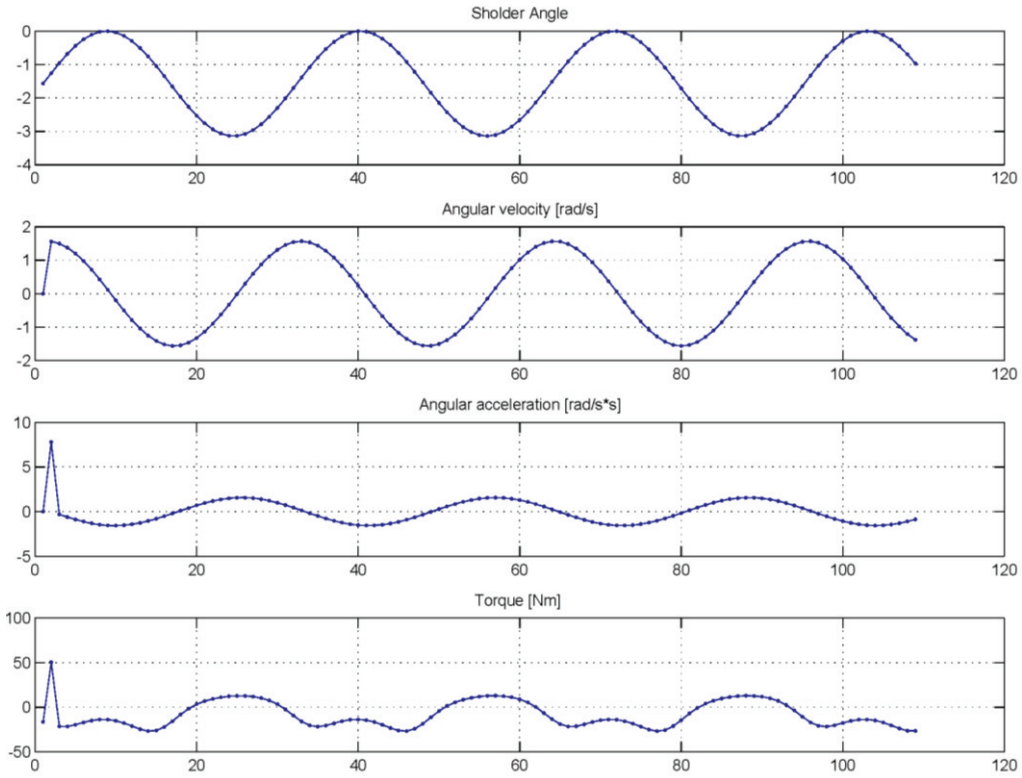


Fig. 5. Results of MATLAB simulations: shoulder angle vs. time plot represents arm angular displacement at the shoulder level; arm angular velocity vs. time at the shoulder level; arm angular acceleration vs. time at shoulder level; the torque that must be applied for arm forward/backward projection.

Electric actuators are known for their high precise motion, accuracy and flexibility. Also pneumatic and hydraulic actuators can produce precise motion but their limits cannot rich the electrical actuator limits.

Pneumatic or hydraulic actuators use efficiently their input energy. Their speed and force or torque is grater then electrical actuators.

The compressed fluid actuator comes in two categories: linear and rotary. In both cases the input energy is not lost though mechanisms that convert linear motion into circular motion or otherwise around. The speed and power of hydraulic and pneumatic actuators can be easily adjusted and are independent one of each other in comparison with electrical actuators which don't exhibit this property.

Pneumatic and hydraulic actuators are lighter and smaller then electrical actuators with the same output parameters: force or torque, displacement, velocity and acceleration.

First and most important disadvantage of using this type of actuators is represented by the high cost of producing a compress fluid. The compressed fluid is obtained through a mechanical system droved by another actuator. The most common actuator used is electrical or combustion motors. This multiple energy transformations resumes to a high loss of power at the compressor input. Even so a good design of the compressed fluid system can reduce to loss given by the multiple energy transformations. A good design tends to use as much actuators as possible into the highest limit of the compressor capacity.

Another disadvantage in case of pneumatic actuators is represented by the noise emitted by the compressed air at the actuator output. The compressor is also a source of noise and vibrations because of its moving mechanical components.

The electrical actuators for linear displacement are more expensive and they have less significant output force then compressed fluid actuators. In order to achieve the compressed fluid forces and torques, electrical actuators needs complementary mechanisms for rising the torque or force. www.azorobotics.com (2013)

Table 1 gives a selection of DC motors and gear reducers that can be selected for this application www.leroy-somer.com (2013), www.maxonmotor, (2013), www.valeoservice, (2013). Some servomotors need additional transmissions in order to achieve the required torque. According to Table I, the electrical motors or servomotors that satisfy the high output torque are heavy and also require high current, power amplifiers for drivers, what influence negatively the portability of such robotic system. The lightly motors have an insufficient output torque to drive passive exercises, especially at the shoulder level.

Table 1. Suitable actuators for shoulder joint of the wearable rehabilitation exerciser

Producer	*Cod*	*Torque [Nm]*	*Speed [rpm]*	*Red. ratio "i"*	*Power [W]*	*Voltage [V]*	*Current [A]*	*Weight [kg]*	*Obs.*
Leroy Somer	MBT 1141L	1.7	3000	99.36	0.55	24/48	-	9	brake
Leroy Somer	MBT 82 IL MBT 1141S	0.98	1500	49.68	0.15	12/48	-	4.5 6	brake
Maxon	370354	0.405	5950	100	200	24	10	1.1	-
Maxon	370355	0.418	5680	100	200	36	7.07	1.1	-
Maxon	370356	0.420	4900	71	200	48	4.58	1.1	-
Maxon	370357	0.452	2760	71	200	70	1.89	1.1	-
Valeo	404722/0231B	12	200	39.7	-	24	20	0.65	wormgear
Valeo	404465/0231A	10	150	39.7	-	24	24	0.65	wormgear

3. CONCLUSION

The developed MATLAB model covers the specific cases of a family of wearable exercisers for the upper limb, with up to six degrees of freedom (in our model, forearm pronation-supination was not considered), starting with shoulder exercisers, elbow or wrist exercisers, and different combinations thereof.

Design with joints and links that correspond to those of the upper limb, compact and lightweight design, are important characteristics for an adequate portability of the exoskeleton-type rehabilitation robotic systems. If our target is passive mobilisation of the different segments of the upper limb, in order to assist /compensate some functions or in order to carry out exercises, the actuators issue becomes relevant in terms of their torque to weight ratio. Different simulations for a large range of anatomical movement combinations give us useful data concerning the required torques of the actuators. These values were partially experimentally validated by our measurements performed at the shoulder and wrist level. Results obtained from simulation lead us to two directions in the development of a functional prototype:

- reducing the weight of all components of the wrist and elbow modules, including their actuators, as they constitute a significant load for the actuators of the shoulder module;

- inclusion in the structure of the wearable rehabilitation robotic system of gravity-balancing mechanisms.

ACKNOWLEDGMENT

This work is supported by PCCA Project, no. 180/2012, A Hybrid Fes-Exoskeleton System to Rehabilitate the Upper Limb in Disabled People (EXOSLIM).

REFERENCES

Rui C.V., Harwin W.S., Nagai K. and Johnson M. (2011) *Advances in upper limb stroke rehabilitation: a technology push*, in Med. Biol .Eng. Comput. 49, pp. 1103-1118.

Mandru D. et al. (2012), *A Hybrid Fes-Exoskeleton System to Rehabilitate the Upper Limb in Disabled People (EXOSLIM)*, Research Report, Technical University of Cluj-Napoca.

Qinling L., Minxiu K, Zhijiang D. and Lining S. (2009), *A novel 5-DOF exoskeletal rehabilitation robot system for upper limb* in High Technology Letters, vol. 15, no.31, pp.245-290.

Rahman M. H., Saad M., Kenne J. P. and Archambault P. S. (2010), Modeling and development of an Exoskeleton Robot for Rehabilitation of Wrist Movements, in IEEE/ASME Int. Conf. in Advanced Intelligent Mechatronics, Montreal, pp. 25-30.

Parasuraman A. W., Oyong and Ganapathy V. (2009), Development of Robot Assistend Stroke Rehabilitation System of Human Upper Limb, in 5th Annual IEEE Conf. on Automation Science and Engineering, Banlagore, pp.256-261.

R. Drillis, R. Contini, and M. Bluestein (1964), *Body Segment Parameters a Survey of Measurement Techniques*, in Artificial limbs Vol. 8.1, pp. 44-66.

Mustafa S. K., Yeo S. H., Pham C. B., Yang G. and Lin W. (2005), *A Biologically-Inspired Anthropocentric Shoulder Joint Rehabilitator: Workspace, Analysis & Optimization*, in Proc. of the IEEE Int. Conf. on Mechatronics & Automation, Niagara Falls, pp. 1045-1050.

Wu T-M., Wang S-Y. and Chen D-Z. (2011) *Design of an exoskeleton for* strengthening *the upper limb muscle for over extension injury prevention* in Mecanism and Machine Theory, 46, pp. 1825-1839

www.leroy-somer.com/catalogue-industrie/fichiers/MBT-fr.pdf, (2013)

www.maxonmotor.com/maxon/view/content/products, (2013)

www.valeoservice.com/html/export/en/produits.catalogueproduits.php, (2013)

www.azorobotics.com/Article.aspx?ArticleID=44, (2013)

Intelligent Monitoring and Control System of Environmental Parameters for Spaces Designed for Radio Equipment

Dicu Gheorghe Doru

E-mail: dicu_doru@yahoo.com

Abstract: This paper develops an integrated heating, cooling and ventilation solution for spaces designed for telecommunications equipments that are monitored and controlled by a controller. This ensures a significant reduction in energy consumption, a decrease in administration costs, increased safety and extended life operation of telecommunications equipments.

Keywords: Parameters, ambient temperature, monitoring, ventilation, equipment, telecommunications.

1. INTRODUCTION

Special containers in which are installed telecommunications equipments are referred to in the literature as shelters, and require control of ambient temperature and humidity parameters. Keeping these parameters should take into account both day / night, seasons, and the climate, the location on the sea coast or at altitudes above 1500 m where environmental conditions, temperature and humidity are extreme.

Although there are concerns about maintaining environmental parameters this is very difficult to accomplish when the shelters are located outside cities, usually on the outskirts of towns or edge of the forests where thermal insulation is undersized, or inexistent, also in winter condensation can cause serious problems to equipments, in summer very often air-conditioning fail due to uninterrupted operation for months and last but not least the low costs that are allocated to those equipments considered related.

Shelters are found in the endowment of Cosmote, Vodafone, Orange mobile networks, the endowment of the Ministry of Administration and Interior, Ministry of National Defence, Institute of Earth Physics, National Administration of Meteorology, and other agencies who can not perform activities such as measuring water levels, weather stations, radiation, measuring environmental parameters, telecommunications, etc.

Environmental requirements imposed by telecommunications equipments are to maintain a inside temperature between 15 ºC - 30 ºC and humidity less than 90%, at the variations of the outside temperature between -25 ^{o}C to 60 ^{o}C and the humidity more than 90%.

To maintain temperature and humidity parameters we followed the development of an intelligent monitoring and control system which consists of: -microcontroller; - inside / outside temperature sensors; - ventilators; - cooling and dehumidification system for the summer season; - heating system for the winter season;

2. FUNCTIONAL STRUCTURE OF THE SHELTER

To ensure these requirements Shelter is equipped with the following equipment:
- Temperature sensors for measuring temperature inside and outside of shelter and inside Rack equipment;
- Heating equipment (electric heater) used when the indoor temperature is below 18 ° C, to avoid condensation;
- Two air fans that provide ventilation air to the outside, when the temperature is between 15 ^{o}C and 30 ^{o}C;
- A cooling system (air conditioning) used if the temperature is above 25 ° C, providing and regulating air humidity;
- A Rack Equipment with the Communications and Control system can be equipped with functioning all their fans.

For special geographical conditions Shelter can be provided with suitable equipment such as passive heating systems desert and mountain areas exploiting large temperature differences between day and night or solar electric panels in areas where there is no possibility of power supply. (www.intertec.info, 2012). Communications are a constant concern and permanent methods are developed for data acquisition and monitoring equipment and maintaining a climate imposed by the manufacturer of telecommunications equipment. On this line, a major concern of researchers in the field is the algorithms and implementation of on-line detection and localization of faults occurring or telecommunications equipment or to maintain climatic parameters inside a shelter (Lubrittoa, 2011).

The electronic module control cabinet SCS (System Control Shelter) is a digital system with the main task to optimize the functioning of equipment inside the enclosure provides weather conditions, monitoring their operation, the memory of events that result in exceeding the optimum temperature regimes and launch alarms for each event.

Figure 1 shows the block diagram and connections mounted equipment shelter and in Figure 2 is presented

physical layout of equipment and components for ambient monitoring and control.

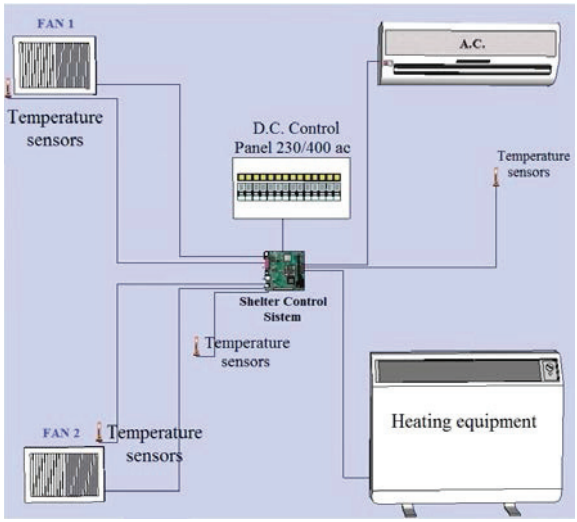


Fig. 1. Block diagram and connections mounted equipment shelter

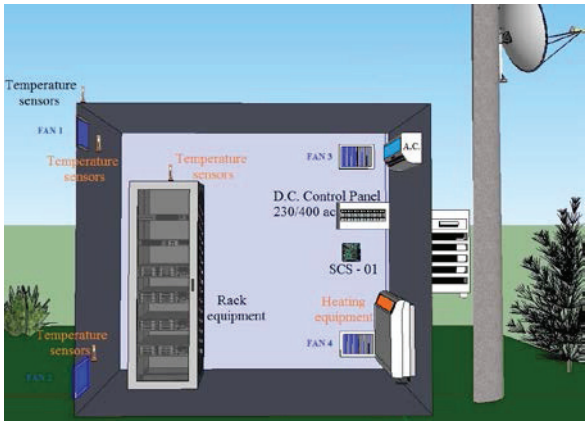


Fig. 2. Physical layout of equipment and components for ambient monitoring and control.

Shelter Control System module application is based on a microcontroller and programming module aims at optimal operating conditions telecommunications equipment installed inside the cabinet. Shelter System Control Module Wiring within the cabinet and interconnection with temperature sensors and equipment whose operation he manages is shown in Figure 1.

For this purpose, Shelter Control System monitors multipoint control indoor temperature and then cooling elements fitted to the cabinet: Fan, coil (electric heater), air conditioning equipment.

Depending on the power dissipated inside the enclosure and outside temperature, Shelter Control System activates more than one means of conditioning mentioned above.

This results in the following modes:

a. cooling when air conditioning is enabled;

b. ventilation where fans are activated and temperature exchange between the shelter and the external environment;

c. natural air cooling in the case in which none of the components of air-conditioning is not available.

Damage to all temperature sensors or a faulty power Shelter System Control System activates both alarm temperature.

3. THE CONTROL AND MONITORING SYSTEM

The Controller that performs monitoring environmental parameters in communication shelters is a process system conducted with ColdFire V1 MCF51JM64 microcontroller, manufactured by Freescale (USA). Intelligent Monitoring and Control System can achieve air conditioning of communication containers by continuous running of a pre-recorded software in the flash memory of the microcontroller. This application is launched in execution automatically upon powering, and does not require any initial settings or operator intervention and is scheduled to provide, within the shelter is installed, an internal temperature between 15 º C and 30 º C.

External environmental conditions that can be installed the shelters provided with SIMC-01 are: ambient temperature between -33 º C and +50 º C; maximum relative humidity 100%; maximum air speed 50 m / s; maximum intensity of rain in 6 mm / min. In parallel with the air conditioning operation, SIMC-01 also undertakes monitoring environmental parameters in shelter. For this the system performs the following operations:

• collects data on: the status of alternative power supply (absence or presence); the status of continuous electrical power supply system (battery voltage); external and internal shelter temperature at predetermined intervals;

• if the internal temperature permissible limits exceed, generates temperature alarms;

• stores the information collected and the alarms in files on Micro SD CARD support FAT32 format.

• every day SCMS-01 makes a text file in which the abovementioned data are stored.

3.1. Functional description

From the constructive and functional points of view, SIMC-01 consists of two modules:

• DATA LOGGER Module;

• FAN CONTROLLER Module.

The two modules are connected to each other through two connectors, one of 36 pins for low voltage logical and analogical signals and one of 5 pins for -48V supply.

3.2. DATA LOGGER Module

DATA LOGGER module is embedded system, is a microcomputer running a dedicated application. The basic component of an embedded system is the microcontroller.

The microcontroller for DATA LOGGER module is MCF51JM64 with 32-bit operating system. The MCF51JM64 microcontroller features enabled implementation of the following functional units in the module DATA LOGGER: Micro SD CARD memory interface; LCD-Display with 2 lines of 16 characters each; -Memory for constants EEPROM type with I2C interface; Real-time clock and calendar; - Port USB; Serial Port with RS232 interface, (COM Port); -Four analogue inputs; - Four digital inputs; -Command buttons.

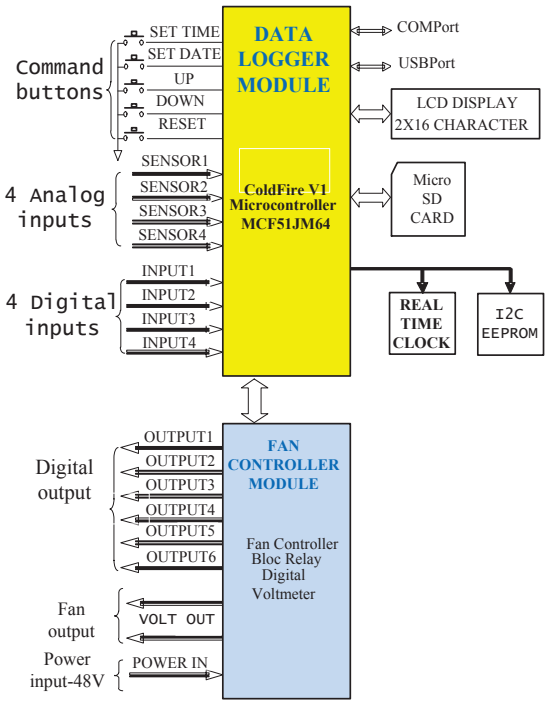


Fig. 3. Intelligent Monitoring and Control System -01 Block Diagram.

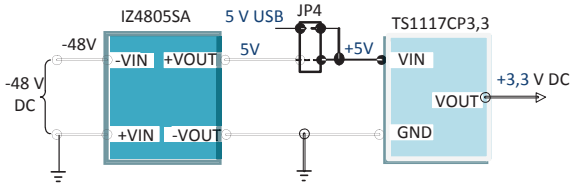


Fig. 4. Power block.

Except the functional units mentioned, DATA LOGGER module also contains a power unit, figure 4, consisting of two dc-dc voltage converters 5V / 3,3 V.

● LCD display 2x16 characters. The LCD display, model DEM16216SYH-PY, is a liquid crystal display with two lines of 16 characters each.

The Intelligent Monitoring and Control System -01 can pass through the following states during the air conditioning process: heating - HM (heating mode) is shown on the display; natural air conditioning - VO is shown on the display; six-step ventilation - V1 , V2, V3, V4, V5 or V6 is shown on the display; cooling - CM (cooling mode) is shown on the display;

● Interface of micro SD memory CARD. DATA LOGGER module is provided with an interface for the micro SD memory card.

The interface of micro SD memory CARD may operate in two ways: SD CARD mode; SPI mode.

● Real time clock and calendar. Provides for SIMC-01 system the information on the local time and date as follows: date (day of the week, date, month, year); time (seconds, minutes, hours).

This information is displayed permanently on the LCD display of the system, but it is also used by FAT32 file system in order to create the temporary attributes of files and folders. Real time clock and calendar are performed with the specialized DS1337 integrated circuit, manufactured by DALLAS-MAXIM. The last two features enabled DS1337 real time clock to be powered directly from a lithium battery.

3.3. FAN CONTROLLER Module

● Digital outputs

● DC fan control outputs

To this end, SIMC-01 monitors in many areas, the inside / outside temperature and then orders the conditioning items with which it is equipped the shelter: ventilators; air conditioning equipment; electric heater. Depending on the temperature dissipated inside the shelter by the equipments and the outdoor temperature, SIMC-01 activates at most one element of conditioning.

Resulting in the following modes of operation: cooling when air conditioning is enabled; ventilation, where the ventilators are activated; heating, when the electric heater is activated; and natural air cooling, if any of the conditioning elements are not available (Ciobanu, 2006).

Where the temperature inside the shelter exceeds permitted limits, SIMC-01, can generate the following two alarms: - the alarm for upper limit temperature exceed; - alarm for lower limit temperature exceed.

To control the temperature inside the cabinet SIMC-01 uses three temperature sensors. If one of the sensors is defective or not connected, the temperature control is done using the other sensors remained in the system.

The redundancy of sensors ensures increased system reliability and also optimizes the internal temperature of the cabinet (Vinatoru and Iancu, 1999).

4. PROCEDURE FOR TESTING

Testing in terms of functionality the Intelligent Control System -01, can be done in the laboratory and there are necessary the following equipments:

● Controller for monitoring and conditioning of the shelter, which will be tested,

● Direct current laboratory source, output voltage of 48V and minimum output current of 2A;

● One or two DC fans, model EBM PAPST 6248N, connected to VOLT OUT 1 and VOLT OUT 2;

● Temperature sensors type KTY81-110, equipped with bond wires connected to SENSOR 4;

● 2200 ohm multi-turn potentiometer to simulate temperature changes connected to SENSOR 1.

From a functional perspective Intelligent Monitoring and Control System -01 is an automated whose entries are:

● Temperature sensors: SENSOR1, SENSOR2 and SENSOR3 and

● Digital input INPUT3, and whose outputs are:

● Controlled voltage outputs for fan control: VOLT OUT1 and VOLT OUT2

● OUTPUT5 for cooling equipment command (air conditioning),

● OUTPUT6 Output for heating equipment order (electric heater),

● OUTPUT3 for temperature alarm,

● OUTPUT4 for under-temperature alarm.

Temperature sensor of the SENSOR4 input is used to measure the temperature outside the shelter. Sensors of input: SENSOR1, SENSOR2 and SENSOR3, measure the internal temperature of the shelter (Iancu and Vinatoru, 2003).

Air conditioning algorithm makes decisions based on the maximum temperature measured by these three sensors. If one or two sensors are disconnected air conditioning algorithm ignores it or them. Therefore you can use a single sensor to test the operation of SIMC-01. It has three internal sensors for safety reasons during functioning and for the levelling of internal temperature.

If all three temperature sensors: SENSOR1, SENSOR2 and SENSOR3 are disconnected, air conditioning algorithm can not work any more and in this situation both fans are commanded at full speed, and temperature alarms are activated to identify the situation (Vinatoru, 2001). KTY81-110 is a silicon temperature sensor with positive temperature coefficient. The range of temperatures in which it can be used is comprised between -55ºC and +150 ºC.

Within this range the resistance of the sensor modifies from 490 ohms to 2211 ohms.

INPUT3 jumper simulates the presence of voltage inside the network (electric-power-supply network of 220Vc.a.). By disconnecting the jumper the absence of network voltage is signalled. In this situation air conditioning controller can only ventilate, situation in which the equipments are powered from DC power supply, (-48V). Heating or cooling commands are inhibited even if the internal temperature would impose the activation of either.

After AC power supply is restored, the activation of cooling or heating commands (if need be) is made with a one minute delay.

5. DESIGNING OF FUNCTIONING FLOW CHARTS

The conditioning algorithm can distinguish seven internal temperature ranges (see figure 5).

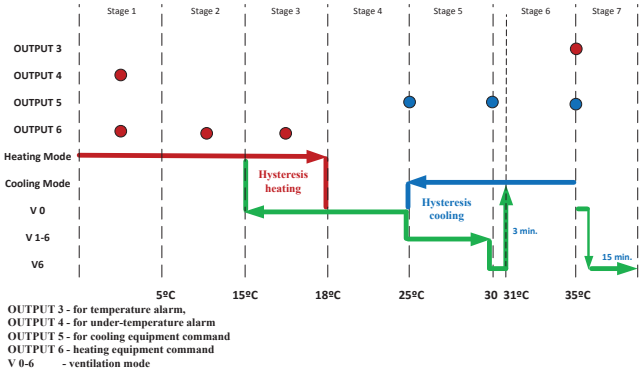


Fig. 5. Functioning chart of the Intelligent Monitoring and Control System - 01

5.1. Temperature less than or equal to 5° C (t ≤ 5° C).

In this stage the low temperature alarm is activated, the led corresponding to OUTPUT4 output turns off. The heating mode is activated (if the mains voltage is present) and the led corresponding to OUTPUT6 output is on.

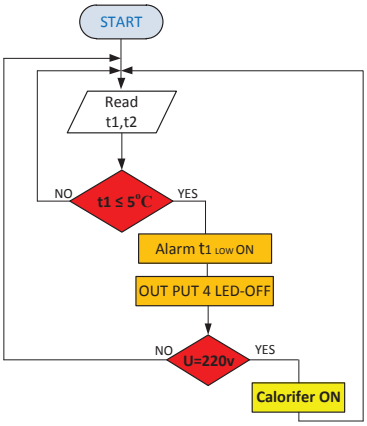


Fig. 6. Temperature less than or equal to 5 º C

Otherwise there is no command of any climate element.

5.2. Temperatura strictly higher than 5 ° C and strictly lower than 15 ° C.

In this stage low temperature alarm is disabled, the led corresponding to OUTPUT4 output is on. Heating mode is activated (if mains voltage is present) and the led corresponding to OUTPUT6 output is lit. Otherwise there is no command of any climate element.

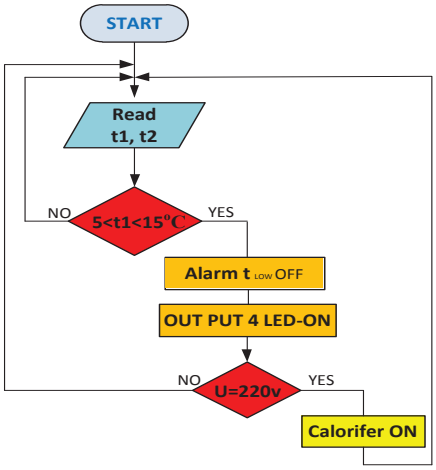


Fig. 7. Temperature strictly higher than 5 ° C and strictly lower than 15 ° C

5.3. Temperatura higher than or equal to 15 ° C and strictly lower than 18 ° C.

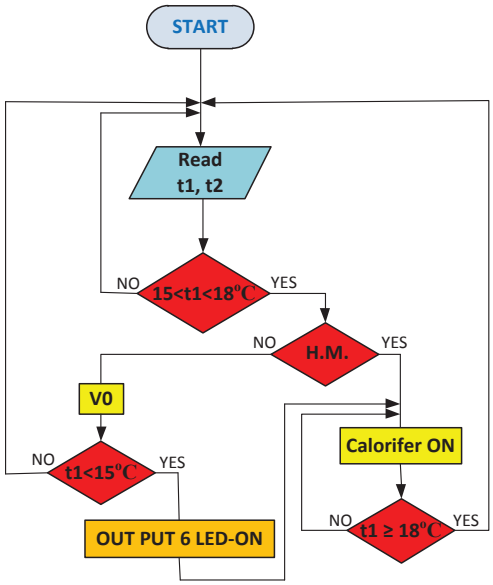


Fig. 8. Temperature higher than or equal to 15 ° C and strictly lower than 18 ° C.

At this stage of temperature is the heating mode hysteresis area. If the conditioning controller is on heating mode, HM (heating mode) is displayed in the bottom right corner of the display when this mode is maintained until the internal temperature of the container is higher than or equal to 18 ° C. If the conditioning controller is on natural cooling mode, V0 is displayed in the bottom right corner of the display then this mode is maintained until the temperature inside the container falls below 15 ° C. At this point switches on heating mode, HM, and the led corresponding to OUTPUT6 output is on.

5.4. Temperatura higher than or equal to 18 ° C and strictly lower than 25 ° C.

At this stage of temperature the conditioning controller is on natural cooling mode, V0 is displayed in the lower right corner of the display.

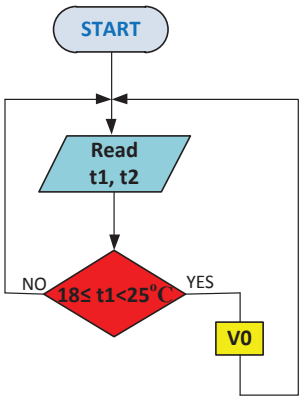


Fig. 9. Temperature higher than or equal to 18 ° C and strictly lower than 25 ° C.

5.5. Temperature higher than or equal to 25 ° C and lower than or equal to 30 ° C.

At this stage of temperature is the cooling mode hysteresis area. If the conditioning controller is on cooling mode, CM (cooling mode) is displayed in the bottom right corner of the display, then this mode is maintained until the internal temperature of the container is lower than or equal to 24 ° C.

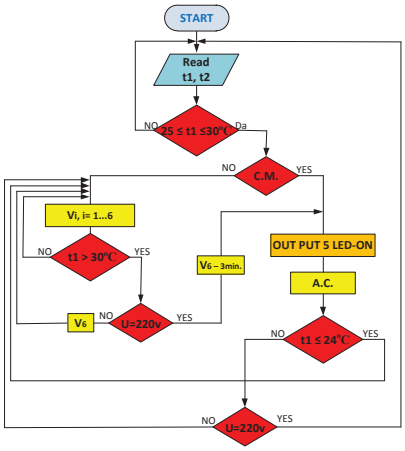


Fig. 10. Temperature higher than or equal to 25 ° C and lower than or equal to 30 ° C.

When SIMC-01 is on ventilation mode, Vi, i = 1-6 is shown in the bottom right corner of the display, and this mode is maintained until the temperature inside the container rises above 30 ° C. On ventilation mode SIMC-01 has 6 operation steps (25 ° C level V1, at 26 ° C level V2, at 27 ° C level V3, at 28 ° C level V4, at 29 ºC level V5 and at 30 ° C level V6). If the temperature inside shelter reaches 31ºC is initiated transition on cooling mode and the led corresponding to OUTPUT5 output is on.

At this moment the indoor unit of the air conditioning equipment is power-supplied, but the outdoor unit starts with a delay of no more than three minutes. In this 3-minute interval, Air Conditioning Controller remains in ventilation mode level V6. After these 3 minutes, the ventilation stops and Air Conditioning Controller switches actually to cooling mode and CM (cooling

mode) is displayed in the lower right corner of the display. The switch to cooling mode is of course subject to the presence of power-supply voltage, otherwise it remains in ventilation mode level V6.

5.6. Temperature strictly higher than 25 ° C and lower than or equal to 35 ° C.

At this stage of temperature the conditioning controller remains on cooling mode, CM (cooling mode) is displayed in the bottom right corner of the display if the mains voltage is present otherwise is passed on ventilation mode level V6.

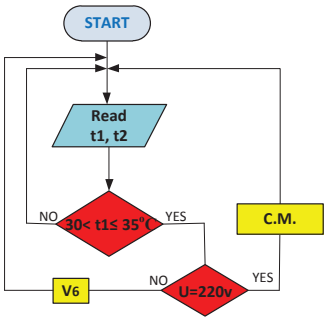


Fig. 11. Temperature strictly higher than 25 ° C and lower than or equal to 35 ° C.

4.7. Temperature strictly higher than 35 ° C.

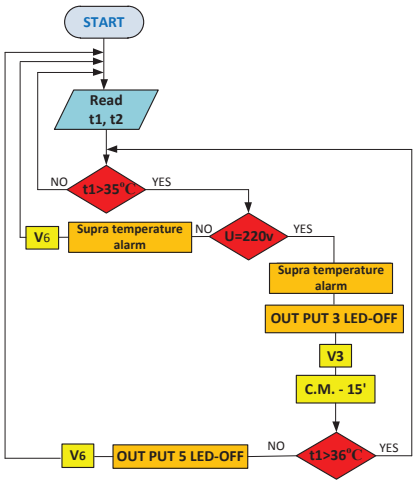


Fig. 12. Temperature strictly higher than 35 ° C.

At this stage of temperature the conditioning controller is initially on cooling mode CM, over-temperature alarm is activated and the led corresponding to OUTPUT3 turns off, and also the ventilation at level V3 is activated. If over-temperature alarm does not disappear in the next 15 minutes, this means if the temperature does not decrease below 36º C, while cooling and ventilation at V3 are operating, cooling is stopped, the led corresponding to OUTPUT5 turns off, ventilation increases at V6 and Air Conditioning Controller is switched to ventilation mode, V6 is displayed in the bottom right corner of the display.

In the absence of power-supply voltage and internal temperature is strictly higher than 35º C, over-temperature alarm is activated and the air conditioning controller is switched to ventilation mode, maximum level. V6 is displayed in the bottom right corner of the display.

For testing the conditioning controller is *slowly* adjusted the potentiometer connected to SENSOR 1 input such that T1 temperature on the display to go through all temperature ranges mentioned above. It is verified for each temperature range shown above, if the operation of the conditioning controller conforms to the description given.

5. CONCLUSIONS.

Throughout the paper we aimed as main idea to ensure maximum safety in the operation of telecommunications equipments; this can be ensured only by an intelligent monitoring and control system of environmental parameters inside the shelters, leading to extension of life of both the telecommunications and air conditioning equipments. The safety of equipment operation reduces the number of alarms and accidental stops, leading to increased efficiency of voice and data communications and consequently of gains for cellular telephony operators. At the same time also was aimed the intelligent management of electrical power, resulting in real savings and therefore in lower maintenance costs and not least environmental impact. On telecommunication shelters market, that are in operation at this time at the main telephone operators, I can say that there is no product to manage heating, air conditioning and ventilation equipments. Although these heating and air conditioning equipments are provided, they are connected periodically manually by maintenance personnel who involve movement to these objectives and management costs increase.

REFERENCES

Ciobanu, L. (2006). *Sensors and transducers*, Ed. MatrixRom, Bucuresti.

Iancu, E. and Vinatoru, M. (2003). *Analitical Methods for Fault Detection and Isolation* – case studies, Ed. Universitaria Craiova.

Lubritto, C. et al. (2011). Energy and environmental aspects of mobile communication systems, *Energy,* Vol. 36, Issue 2, pp. 1109-1114.

Purnima, S.R.N. Reddy (2012). Design of Remote Monitoring and Control System with Automatic Irrigation System using GSM-Bluetooth, *International Journal of Computer Applications* (0975–888), Vol. 47, No. 12.

Vinatoru, M. (2001). *Industial Plant Control*, Ed. Univesitaria. Craiova

Vinatoru, M. and Iancu, E. (1999). *Fault detection and location in dynamic systems,* Ed. SITECH, Craiova.

*** (2012). Passive Cooling Systems, *INTERTEC-Hess GmbH · Raffineriestraße* 8 ·D-93333 Neustadt/Donau, www.intertec.info.

A FG-MOS Continuous-Time Current-Mode Fully-Differential Integrator for Low Supply Voltage Applications

Elena Doicaru

Faculty of Automation, Computer Science and Electronics, University of Craiova, Craiova, Romania (e-mail: dmilena@electronics.ucv.ro)

Abstract: This paper presents a CMOS current-mode fully-differential integrator realized by using floating-gate (FG)-MOSFETs. The possibility of controlling the apparent threshold voltage of FG-MOSFETs whit multiple control gates offer an extra degree of freedom which can be exploited with success in design of low voltage circuits. Also, the voltage signal summation at the floating gate was effectively used to simplify the circuit configuration. The proposed current-mode integrator was analyzed for 0.5µm standard CMOS process. The simulations show that the circuit implemented in 0.5µm standard CMOS process operate as integrator, for the 1V supply voltage, in the frequency range 8MHz-800Mz. In the paper firstly, the FG-MOS device is presented. Then the current-mode integrator is presented and analyzed. Finally, the simulation results for the sub-micron technologies are presented.

Keywords: Analogue signal processing circuit, low voltage circuit, CMOS analogue integrator, FG-MOSFET.

1. INTRODUCTION

Signal processing in its various forms has an explosive growth, both technically and economically. A rapid development of digital MOS VLSI technology has enabled the integration of powerful digital processing systems into a single chip. The next step consists in integration of systems with mixed signals on same chip. Unfortunately, because of optimization of the technology for the implementation of digital circuits, the design of analogue circuits becomes even more difficult. In addition, along with the development of the portable electronics products and the mobile communications, the design of low-voltage and low-power systems with mixed signals has gained importance. For systems like hearing aids, implantable cardiac pacemakers, cell-phones, hand-held multimedia terminals and so on, the battery is the main or the single source of power. Using the DC-DC conversion in order to obtaining higher voltages is not adequate for analogue circuits. So, these circuits must be able to operate at supply voltage values contained in the range [0.5V; 3V].

As the feature size of the CMOS process reduces, the supply voltage has to be reduced in order to reduce the power dissipation and the intensity of electrical field in thinner oxide layers. The reduction in supply voltage leads to degradation of circuit performances, specially the available bandwidth and the voltage swing. By scaling down the size, the performances of the digital circuits are improved but the benefit for the analogue circuits is reduced. The size of devices in analogue circuits cannot become very low because the requirements of noise and offset.

For the low-voltage high performance analog circuit design current mode approach is a good alternative. In this technique, for CMOS technologies, four approaches are more used: (1) translinear approach with those four variants: with MOS transistors in saturation regime, with MOS transistors in ohmic regime, with MOS transistors in subthreshold regime and with back-gate control, (2) subthreshold operation of MOS transistors approach, (3) FG-MOS approach and (4) so called "log-domain" filtering.

Generaly, in current-mode technique the addition and subtraction of signals are power consuming operations and the design flexibility is limited by the nonlinear I-V and V-I conversions. These problems can be overcome by using FG-MOSFET approach. The floating-gate MOS transistor can perform the linear addition of signals in voltage mode and more this addition does not consume DC power. But these devices offer another property very useful in realization of low-voltage analogue building blocks, the possibility to linear controlling of the threshold voltage. For this property I choose the FG-MOSFET approach for realisation of basic building bloc of filters - the integrator.

The integrator presented in this paper is realized by using the floating-gate-MOS devices that can be easily implemented in standard double-poly CMOS technologies. By adopting a fully-differential structure, the influence of the common-mode signals and noise are minimised. The proposed current-mode integrator was analyzed and its main characteristics are derived. Also, for a 0.5µm CMOS process technologies the SPICE simulation results are also provided.

2. FLOATING-GATE MOSFET

2.1. The Characteristics I-V of FG-MOSFET

The FG-MOS transistor is an important element in the design of micro-power and low-voltage systems. It can be use in neural networks e.g. Borgstrom et al. (1990), Sin et al. (1992), Vittoz et al. (1991) and in analogue processing signals as a summing node to perform mathematical operations in the charge domain e.g. Yang et al. (1993). In Fig. 1 are presented: (a) a cross section through simplified structure of an FG-MOS device with a single control gate, (b) the simplified model of structure and (c) the symbol used for FG-MOS devices with multiple control gates. In a standard double-poly-silicon process, the floating gate is the first poly-silicon layer and the control gate is the second poly-silicon layer (see Fig. 1.a). The floating gate is capacitive coupled with the control gate, source, drain and substrate (see Fig 1.b).

The I-V characteristics for the FG-MOST can be obtained by modifying the equations for the conventional MOS transistor e.g. Wang (1979). The main differences to the conventional MOS transistor are a consequence of the fact that (1) the control gate is capacitive coupled with floating gate and (2) the floating gate has no DC path to the ground so a charge can be stored here, charge that directly affects the drain current and threshold voltage. Primarily the FG-MOSFETs have been used in EPROM and EEPROM circuits, but this second feature led to the use of these devices as analog memories (Field Programmable Analogue Areas – FPAA). A generalization of equations developed by Wang (1979) for a device with multiple control gates and if no leakage current through insulator oxide is presented below.

For an *n* channel FG-MOSFET, in the ohmic region, the drain current can be expressed as

$$I_{DS}^{o} = \mu_n C_{ox} \frac{W}{L} \cdot \left[\left(\sum_{i=1}^{N} \frac{C_{gi}}{C_T} V_{CGSi} + \frac{C_{gd}}{C_T} V_{DS} + \frac{C_{gb}}{C_T} V_{BS} + \right. \right. \\ \left. \left. + \frac{Q_{FG}}{C_T} - V_T \right) \cdot V_{DS} - \frac{1}{2} \cdot V_{DS}^2 \right] \tag{1}$$

and in saturation region as

$$I_{DS}^{s} = \frac{1}{2} \mu_n C_{ox} \frac{W}{L} \cdot \left(\sum_{i=1}^{N} \frac{C_{gi}}{C_T} V_{CGSi} + \frac{C_{gd}}{C_T} V_{DS} + \frac{C_{gb}}{C_T} V_{BS} + \right. \\ \left. + \frac{Q_{FG}}{C_T} - V_T \right)^2 \tag{2}$$

where V_{CGSi} is the control gate *i*-to-source voltage, Q_{FG} is the amount of charge stored on the floating gate, V_T is the threshold voltage of standard MOS, W and L are the MOS channel width and length, μ_n is the electron mobility, C_{ox} is the oxide capacitance per unit area, C_{gi} is the coupling capacitance from the control gate i to the floating gate, C_{gd}, C_{gs} and C_{gb} are the capacitances from floating gate to drain and to bulk, N is the number of control gates and C_T is the total capacitance on the floating gate,

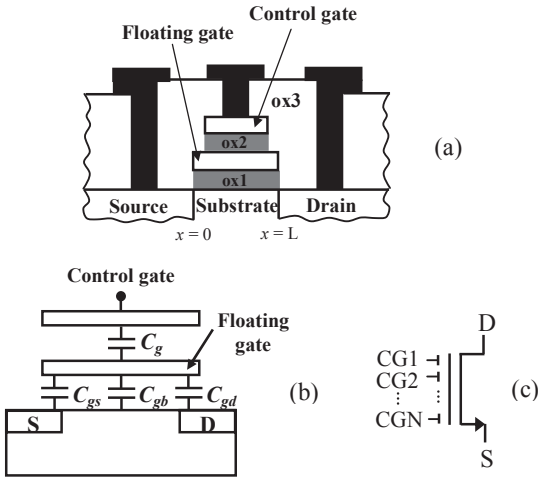


Fig. 1. FG-MOSFET: a) a cross section through simplified structure of an FG-MOS device with a single control gate, b) the simplified model of structure, c) the symbol used for FG-MOS devices with multiple control gates.

$$C_T = \sum_{i=1}^{N} C_{gi} + C_{gs} + C_{gd} + C_{gb} \ . \tag{3}$$

Equation (2) relives that in the saturation region, if the drain is strongly capacitive coupling with the floating gate, the drain current is not independent of the drain voltage. In order to minimized this effect, in standard double poly CMOS process, the capacitances between floating gate and the control gates, C_{gi}, can be readily made much greater than the parasitic capacitances by using layout techniques for the floating gate e.g. Yang et al. (1992), Thomsen and Brooke (1991). So, assuming $C_{gi} >> C_{gd}$, C_{gb}, (1) and (2) can be approximated as

$$I_{DS}^{o} = \mu_n C_{ox} \frac{W}{L} \cdot \left[\left(\sum_{i=1}^{N} \frac{C_{gi}}{C_T} V_{CGSi} + \frac{Q_{FG}}{C_T} - V_T \right) \cdot V_{DS} - \right. \\ \left. - \frac{1}{2} \cdot V_{DS}^2 \right] \tag{4}$$

$$I_{DS}^{s} = \frac{1}{2} \mu_n C_{ox} \frac{W}{L} \cdot \left(\sum_{i=1}^{N} \frac{C_{gi}}{C_T} V_{CGSi} + \frac{Q_{FG}}{C_T} - V_T \right)^2 \tag{5}$$

To illustrate how one can achieve the threshold voltage control via the control gate in FG-MOSFET, let us consider a device with two control gates, one on which apply the useful signal and one on which apply a constant voltage. According to (4) and (5) get the following expressions for drain current

$$I_{DS}^{o} = \mu_n C_{ox} \frac{W}{L} \cdot \frac{C_{g1}}{C_T} \cdot \left\{ \left[V_{CGS1} - \left(\frac{C_T}{C_{g1}} V_T - \frac{Q_{FG}}{C_{g1}} - \right. \right. \right. \\ \left. \left. \left. - \frac{C_{g2}}{C_{g1}} V_{CGS2} \right) \right] \cdot V_{DS} - \frac{1}{2} \cdot V_{DS}^2 \right\} = \tag{6}$$

$$= \beta_{eff}^{o} \cdot \left[\left(V_{CGS1} - V_{T,eff} \right) \cdot V_{DS} - \frac{1}{2} \cdot V_{DS}^{2} \right] \tag{6}$$

$$I_{DS}^{s} = \frac{1}{2} \mu_n C_{ox} \frac{W}{L} \cdot \frac{C_{g1}}{C_T} \cdot \left[V_{CGS1} - \left(\frac{C_T}{C_{g1}} V_T - \frac{Q_{FG}}{C_{g1}} - \right. \right. \\ \left. \left. - \frac{C_{g2}}{C_{g1}} V_{CGS2} \right) \right]^2 = \beta_{eff}^{s} \cdot \left(V_{CGS1} - V_{T,eff} \right)^2 \tag{7}$$

where

$$\beta_{eff}^{o} = \mu_n C_{ox} \frac{W}{L} \cdot \frac{C_{g1}}{C_T} \quad \text{and} \quad \beta_{eff}^{s} = \frac{1}{2} \mu_n C_{ox} \frac{W}{L} \cdot \frac{C_{g1}}{C_T} \tag{8}$$

are the effective structural parameters of the device, relative to signal gate, in ohmic and saturation regions, and

$$V_{T,eff} = \frac{C_T}{C_{g1}} V_T - \frac{Q_{FG}}{C_{g1}} - \frac{C_{g2}}{C_{g1}} V_{CGS2} \tag{9}$$

is the effective (or apparent) threshold voltage relative to signal gate.

On can see that the effective threshold voltage $V_{T,eff}$ can be adjusted or by controlling the charge on the floating gate or by voltage applied to a control gate. So, one can consider that the FG-MOS transistor is a device with programmable threshold voltage. Generalizing (9) for the case in that are used more control gates to adjust the threshold voltage

$$V_{T,eff\,i} = \frac{C_T}{C_{gi}} V_T - \frac{Q_{FG}}{C_{gi}} - \sum_{\substack{j=1 \\ j \neq i}}^{N} \frac{C_{gj}}{C_{gi}} V_{CGSj} \tag{10}$$

In the saturation operation of FG-MOSFET, the trans-conductance gain and the output conductance relative to a gate are

$$g_{mi} \cong \sqrt{2 \beta_{eff}^{s} I_{DS}} = \frac{C_{gi}}{C_T} \cdot g_m^c \tag{11}$$

$$g_{di} \cong \lambda I_{DS} + \frac{C_{gd}}{C_{gi}} g_{mi} = g_d^c + \frac{C_{gd}}{C_{gi}} g_{mi} \tag{12}$$

where λ is the channel length modulation parameter, g_m^c and g_d^c are the trans-conductance and output conductance of conventional MOS transistor.

2.2. The simulation models for FG-MOSFET

Floating gate MOS transistors have proven to be extremely useful devices in the development of analogue systems. However, the inability to properly simulate these devices has delayed their adoption. There have been several attempts at modelling FG-MOS devices in the past, but these models have been simplistic and very limited. However, there are some models that have become very popular within analogue circuit designers. The most popular simulation model for the device was: (1) the Ramirez-Angulo's model, Ramirez-Angulo et al. (1997), (2) the "coupling" model developed by Mondagon-Torres et al. (2002) and presented in detail by Sanchez-Sinencio (2005) and (3) the synapse model used by Rahimi et al. (2002) and Gray et al. (2008). Generally, the SPICE simulator needs at least one path to ground or an initial condition in each point in the circuit in order to obtain the convergence to a solution. Due to the nature of FGMOS (no DC path from the coupling capacitors to ground), a DC model is needed to simulate it.

The Ramirez-Angulo's model provides the DC part by a huge resistor and controlled voltage sources (see Fig. 2). This model is very simple to implement and has been used widely for simulating FG circuits for analogue applications. The major limitation is the inability to represent the charge movement from the gate leakage.

The "coupling" model resolves the DC convergence by adding large resistors in parallel with each capacitor. The resistors are selected such that they are sufficiently large to be neglected in the AC case. More, the sizes of the resistors are also selected such that the RC product of each resistor-capacitor pair to be equal. In Fig. 3 is presented a bettered variant of this model, proposed by the Sanchez-Sinencio (2005). A controlled voltage source and an independent voltage source used to apply the initial charge, Q_{FG}, are added to the initial model. The value of the controlled voltage source is the same with V_{FG} in Ramirez-Angulo's model.

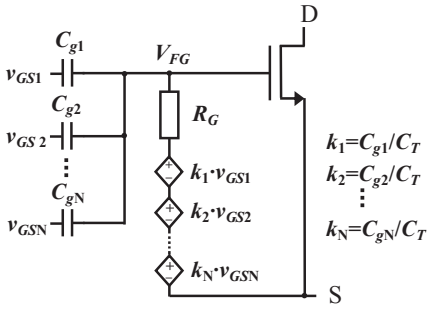


Fig. 2. The Ramirez-Angulo's simulation model.

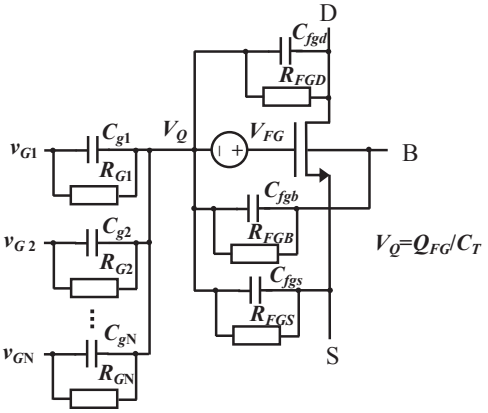


Fig. 3. The Sanchez-Sinencio's "coupling" model.

It can see that this model could be used to test the DC and AC responses of an analogue FG-MOSFET system that does not use any programming functionality during normal operation (the Hot-electron injection and Fowler-Nordheim tunnelling are not effectively implemented in the model, only the initial operating point consider the amount of charge stored on the floating gate). The major drawback in this model is incapability to take in account the gate leakage or charge movement. Practically, in non programming systems, implemented in over-500nm CMOS process, the results of simulations are acceptable. In the sub-500nm CMOS process, due to the very thin silicon oxide at the gate, the direct tunnelling current become a growing issue, so it is need to take in account by adding new elements at the simulation model.

The model proposed by Rahimi et al., often referred as "synaptic" model because his behaviour is similar to that of the biological synapses of neuron, focus on the aspect of charge's alteration in the operation of FG-MOSFET, in a manner suitable for use in adaptive systems. Three currents are modelled in this structure by three voltage-controlled current sources (see fig. 4): the current flowing from the tunnelling junction to the floating node, the injection current flowing from the channel to the floating node and the well current flowing from the channel to the well of the device. However, this model is unable to perform any DC analyses and does not account for the reduction in output resistance.

The model that I have used to simulate the operation of an FG-MOSFET with multiple control gates is shown in Fig. 5. Since the circuit in which it uses the FG-MOS devices does not require programming features, the elements that implement the Hot-electron injection and Fowler-Nordheim tunnelling were removed. The influence of applied voltage on the device terminals over the floating gate voltage is taken into account by introducing a controlled voltage source V_{FG1}

$$V_{FG1} = \sum_{i=1}^{N} \frac{C_{gi}}{C_T} V_{CGi} + \frac{C_{gd}}{C_T} V_D + \frac{C_{gb}}{C_T} V_B + \frac{C_{gs}}{C_T} V_S \,. \tag{13}$$

The influence of the initial charge, Q_{FG}, is simulated by an independent voltage source V_{FG2}

$$V_{FG2} = Q_{FG} / C_T \tag{14}$$

For to simulate the influence of the direct tunnelling current over the voltage of floating gate, three ABM (analogue behaviour modelling) elements have been added: an element that provides the leakage current according to the voltage V_{FG}, an integrator and a subtractor. The implemented mathematical model is

$$\begin{aligned} I_{\text{leakage}} &= A \cdot \exp\left(B \cdot V_{FG}\right) \\ V_{\text{leakage}} &= \frac{1}{C_T} \int_0^t I_{\text{leakage}} \mathrm{d}t \\ V_{FG}^{/} &= V_{FG} - V_{\text{leakage}} \quad V_{FG} = V_{FG1} + V_{FG2} \end{aligned} \tag{15}$$

where A and B are the technology and biasing conditions dependent parameters.

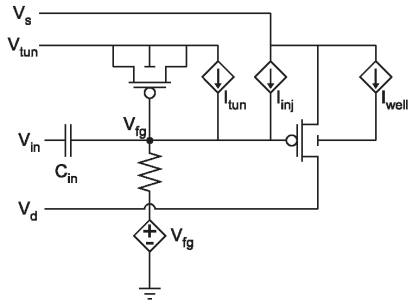


Fig. 4. The synaptic model.

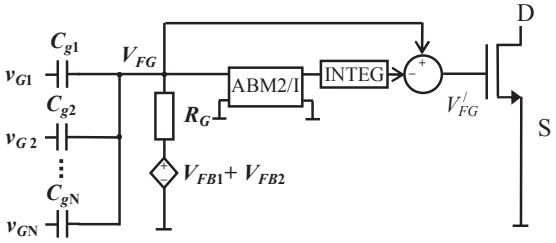


Fig. 5. The structure of simulation model used in this paper for an n-type FG-MOSFET.

The parameters A and B will increase as the technology scaled more to 100nm.

3. THE FG-MOS CURRENT-MODE FULLY-DIFFERENTIAL INTEGRATOR

The current-mode integrator at the bloc diagram level is presented in Fig. 6. The fundamental building blocks for the current-mode integrator are the current amplifiers (see Fig. 7) and the transistor driver trans-conductors. The ideal current gain for are the current amplifier is

$$\begin{aligned} A_I = \frac{i_{out}}{i_{in}} &= \frac{(W/L)_{M4}}{(W/L)_{M3}} \cdot \frac{(W/L)_{M10}}{(W/L)_{M11}} = \\ &= \frac{K \cdot (W/L)}{(W/L)} \cdot \frac{K \cdot (W/L)}{K \cdot (W/L)} = K \end{aligned} \,. \tag{16}$$

The ideal differential-mode current gain for the current-mode integrator is

$$A_{dd} = \frac{i_o^+ - i_o^-}{i_{in}^+ - i_{in}^-} \,. \tag{17}$$

In order to computation the A_{dd} expression firstly we determine the voltages across the capacitors:

$$\begin{aligned} v_C^+ &= i_C^+ \cdot \frac{1}{\mathrm{s}C} = \left[K \cdot (i_{in}^+ - g_{m2} v_C^-) - K \cdot g_{m1} v_C^+ \right] \cdot \frac{1}{\mathrm{s}C} \\ v_C^- &= i_C^- \cdot \frac{1}{\mathrm{s}C} = \left[K \cdot (i_{in}^- - g_{m2} v_C^+) - K \cdot g_{m1} v_C^- \right] \cdot \frac{1}{\mathrm{s}C} \end{aligned} \tag{18}$$

From (18) yields the differential-mode input current

$$i_{in}^+ - i_{in}^- = \left(\frac{\mathrm{s}C}{K} + g_{m1} - g_{m2} \right) (v_C^+ - v_C^-) \tag{19}$$

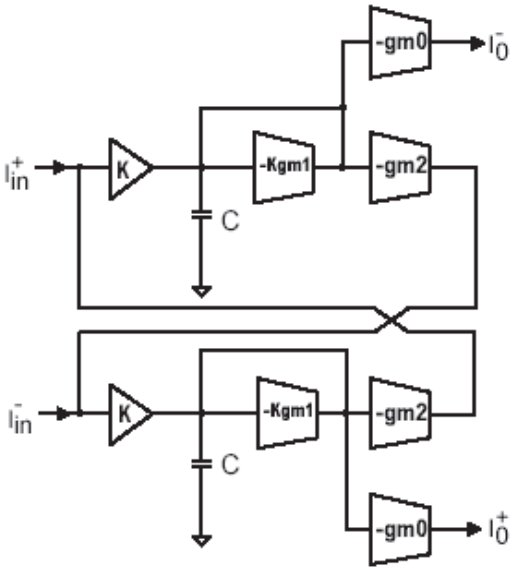


Fig. 6. The bloc diagram of current-mode integrator

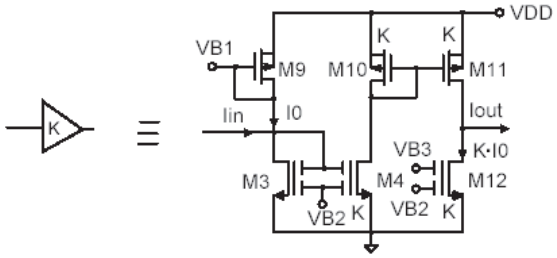


Fig. 7. The current amplifier.

The outputs currents are

$$i_o^- = -g_{m0} v_C^+ \quad i_o^+ = -g_{m0} v_C^- \tag{20}$$

and the output differential-mode current is

$$i_o^+ - i_o^- = g_{m0} \cdot (v_C^+ - v_C^-) . \tag{21}$$

So the ideal differential-mode current gain of the first-order model of integrator is

$$A_{dd} = \frac{i_o^+ - i_o^-}{i_{in}^+ - i_{in}^-} = \frac{K \cdot g_{m0}}{\mathrm{s}C + K \cdot (g_{m1} - g_{m2})} = \\ = \frac{K \cdot g_{m0}/C}{\mathrm{s} + K \cdot (g_{m1} - g_{m2})/C} \tag{22}$$

The ideal common-mode current gain is

$$A_{cc} = \frac{i_o^+ + i_o^-}{i_{in}^+ + i_{in}^-} = \frac{-g_{m0}}{\frac{\mathrm{s}C}{K} + g_{m1} + g_{m2}} = \\ = -\frac{K \cdot g_{m0}/C}{\mathrm{s} + K \cdot (g_{m1} + g_{m2})/C} \tag{23}$$

The common-mode rejection ratio is

$$CMRR = \left| \frac{A_{dd}}{A_{cc}} \right| = \left| \frac{\mathrm{s}C + K \cdot (g_{m1} + g_{m2})}{\mathrm{s}C + K \cdot (g_{m1} - g_{m2})} \right| \tag{24}$$

From (24) we obtain

$$CMRR\big|_{\mathrm{s=j0}} = \left| \frac{g_{m1} + g_{m2}}{g_{m1} - g_{m2}} \right| \tag{25}$$

and

$$CMRR\big|_{\mathrm{s=j\infty}} = 1 \tag{26}$$

It can see that by setting g_{m1} equal with g_{m2} both a high CMRR at low frequencies and an ideal lossless integrator can be obtained. So

$$A_{dd}\big|_{g_{m1}=g_{m2}} = K \cdot \frac{g_{m0}/C}{\mathrm{s}} . \tag{27}$$

However, at high frequencies the CMRR is poor. Grace to fully differential structure of integrator, the signal path is immune to the common-mode noise even at high frequencies. But, when the common-mode signal level is too large, the low value of CMMR at high frequencies causes distortion in differential output signal.

This integrator can have multiple outputs by simply adding additional output trans-conductors (similar with – gm0, see Fig. 6).

The implementation of architecture from Fig. 6 is shown in Fig. 8. In this implementation some groups of transistors have similar dimensions. These groups are presented in table 1.

Table 1. Groups of transistors with similar dimensions

Group no.	Transistors	Aspect ratio
1	$M_1 \equiv M_2$	$B\cdot(W/L)_n$
2	$M_3 \equiv M_5$	$(W/L)_n$
3	$M_4 \equiv M_6 \equiv M_{12} \equiv M_{16}$	$K\cdot(W/L)_n$
4	$M_7 \equiv M_8$	$A\cdot(W/L)_n$
5	$M_9 \equiv M_{13}$	$(W/L)_p$
6	$M_{10} \equiv M_{11} \equiv M_{14} \equiv M_{15}$	$K\cdot(W/L)_p$
7	$M_{17} \equiv M_{19}$	$A\cdot(W/L)_p$
8	$M_{18} \equiv M_{20}$	$B\cdot(W/L)_p$

The transistors involved in implementation of trans-conductor blocs in Fig. 6 are:

- M_1 and its pair M_2 implement the pairs $-K\cdot g_{m1}$ at the control gates 1 and $-g_{m2}$ at the control gates 2

$$K \cdot g_{m1} = \frac{C_{g11}}{C_{T1}} \cdot g_{m1}^c = \frac{C_{g12}}{C_{T2}} \cdot g_{m2}^c = \\ = \frac{K\, C_{g21}}{C_{T1}} \cdot g_{m1}^c = \frac{K\, C_{g22}}{C_{T2}} \cdot g_{m2}^c = \frac{K\, C_{g21}}{C_{T1}} \cdot B\, g_m^c \tag{28}$$

$$g_{m2} = \frac{C_{g21}}{C_{T1}} \cdot g_{m1}^c = \frac{C_{g22}}{C_{T2}} \cdot g_{m2}^c = \frac{C_{g22}}{C_{T2}} \cdot B\, g_m^c \tag{29}$$

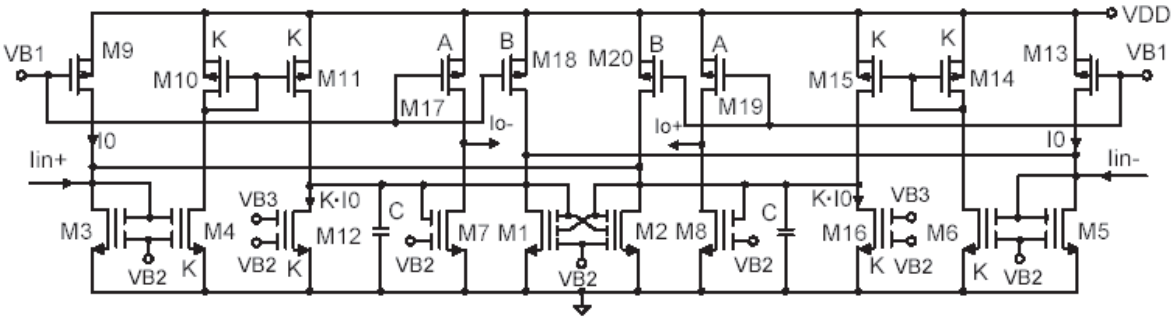


Fig. 8. The FG-MOS current-mode fully-differential integrator

$$\begin{aligned} &C_{g11} = C_{g12} = KC_{g21}, \quad C_{g21} = C_{g22} \\ &g_{m1}^c = g_{m2}^c = B\,g_m^c \\ &C_{T1} = C_{T2} \end{aligned} \tag{30}$$

where C_{gij}, $i = \overline{1,3}$ and $j = \overline{1,2}$, is the coupling capacitance from the control gate i to the floating gate of the transistor M_j, C_{gdj}, C_{gsj} and C_{gbj} are the capacitances from floating gate to drain and to bulk of the transistor M_j, g_{mj}^c are the trans-conductance of the conventional MOS transistor M_j and

$$\begin{aligned} &C_{T1} = \sum_{i=1}^{3} C_{gi1} + C_{gs1} + C_{gd1} + C_{gb1} \\ &C_{T2} = \sum_{i=1}^{3} C_{gi2} + C_{gs2} + C_{gd2} + C_{gb2} \end{aligned} \quad ; \tag{31}$$

- M_7 and its pair M_8 implement the $-g_{m0}$ trans-conductor blocs

$$g_{m0} = \frac{C_{g17}}{C_{T7}} \cdot g_{m7}^c = \frac{C_{g8}}{C_{T8}} \cdot g_{m8}^c = \frac{C_{g8}}{C_{T8}} \cdot A\,g_m^c \tag{32}$$

$$\begin{aligned} &C_{g17} = C_{g18}, \quad g_{m7}^c = g_{m8}^c, \quad C_{T7} = C_{T8} \\ &C_{T7} = \sum_{i=1}^{2} C_{gi7} + C_{gs7} + C_{gd7} + C_{gb7} \\ &C_{T8} = \sum_{i=1}^{2} C_{gi8} + C_{gs8} + C_{gd8} + C_{gb8} \end{aligned} \tag{33}$$

- $M_3 \div M_6$ and $M_9 \div M_{16}$ implements the two current amplifiers of the differential integrator.

All FG-MOS transistors with two control gates have equal ratio C_{gij}/C_{Tj}, i = 1,2 and j = 3÷8, 12, 16. The FG-MOS transistor with three control gates M_1 and M_2 have

$$\begin{aligned} &C_{g21}/C_{T1} = C_{g31}/C_{T1} = C_{g22}/C_{T2} = C_{g32}/C_{T2} \\ &C_{g11}/C_{T1} = C_{g12}/C_{T2} = K \cdot C_{g21}/C_{T1} \end{aligned} \tag{34}$$

In order to ensure closed-loop common-mode stability the inequality

$$B \cdot 2C_{g21}/C_{T1} > A \cdot C_{g17}/C_{T7} \tag{35}$$

must be satisfied.

If the finite output trans-conductor resistances and current loss through coupling capacitance are considered, the differential transfer and common-mode functions become:

$$A_{dd} = \frac{K\,g_{m0}/C_{ech-dm}}{\mathrm{s} + B\,g_0/C_{ech-dm}} = \frac{K\,\dfrac{A\,g_m^c\,C_{g17}/C_{T7}}{C_{ech-dm}}}{\mathrm{s} + B\,g_0/C_{ech-dm}} \tag{36}$$

$$\begin{aligned} A_{cc} &= \frac{-K\,g_{m0}/C_{ech\text{-}cm}}{\mathrm{s} + \left(2K\dfrac{C_{g21}}{C_{T1}}B\,g_m^c + 2Kg_0 + B\,g_0\right)\Big/C_{ech\text{-}cm}} = \\ &= \frac{K\,\dfrac{A\,g_m^c\,C_{g17}/C_{T7}}{C_{ech-cm}}}{\mathrm{s} + \left(2K\dfrac{C_{g21}}{C_{T1}}B\,g_m^c + 2Kg_0 + B\,g_0\right)\Big/C_{ech\text{-}cm}} \end{aligned} \tag{37}$$

where

$$\begin{aligned} C_{ech-dm} = C - 2K\,C_{g13}(1-a) + (1-b)C_{g17} + KC_{g21} - \\ - C_{g21}(K-1)\frac{KC_{g21} - C_{g21} + C_{gd1}}{C_{T1}} \end{aligned}, \tag{38}$$

$$\begin{aligned} C_{ech\text{-}cm} = C + 2KC_{g13}(1-a) + C_{g17}(1-b) + KC_{g21} - \\ - C_{g21}(K+1)\frac{KC_{g21} + C_{g21} + C_{gd1}}{C_{T1}} \end{aligned}, \tag{39}$$

$B\,g_m^c$ is the trans-conductance of conventional MOS transistors M_1 and M_2, $B \cdot g_0$ is the sum between the output conductances of transistors M_1 and M_{18} (or M_2 and M_{20}), $K \cdot g_0$ is the sum between the output conductances of transistors M_{12} and M_{11} (or M_{16} and M_{15}), a is the ratio v_{fgj}/v_C^+ and v_{fgi}/v_C^-, j = 5,6 and i = 3,4, b is the ratio v_{fg7}/v_C^+ or v_{fg8}/v_C^-. The values of ratios b and a are slightly subunit.

The common-mode rejection ratio becomes

$$\begin{aligned} &CMRR = \frac{\mathrm{s}\,C_{ech\text{-}cm} + 2K\dfrac{C_{g21}}{C_{T1}}B\,g_m^c + 2Kg_0 + B\,g_0}{\mathrm{s}\,C_{ech-dm} + B\,g_0} \\ &CMRR\big|_{\mathrm{s}=\mathrm{j}0} = 2K\frac{C_{g21}}{C_{T1}}\frac{g_m^c}{g_0} + 2\frac{K}{B} + 1 \end{aligned} \tag{40}$$

$$CMRR\big|_{s=j\infty} = C_{ech\text{-}cm} / C_{ech-dm} \ . \quad (40)$$

The minimum supply voltage required for this circuit is given by the relation:

$$V_{DD} > V_{T,eff} + V_{sat,I_0} + \sqrt{2 \cdot I_0 / \left[\mu_n C_{ox} \left(W/L \right)_n \right]} \ . \quad (41)$$

The output noise of the integrator is direct proportional with the bias current A·I$_0$, $\overline{i_{nout}^2} = P_n \cdot A I_0$ where P_n depend on technology and temperature. The dynamic range can then be expressed as:

$$DR = \frac{I_{o\max}}{\sqrt{i_{n,out}^2}} = \frac{2 \cdot \eta \cdot A \cdot I_0}{\sqrt{P_n \cdot A \cdot I_0}} = 2 \cdot \eta \cdot \sqrt{\frac{A \cdot I_0}{P_n}}, \eta \in (0,1) \quad (42)$$

In conformity with Ismail and Fiez (1994) and Sanchez-Sinencio and Andreou (1998) the distortion of the integrator can be given by:

- the third intermodulation distortion

$$IM_3 \cong (3/32) \left[i_{0d} / (2AI_0) \right]^2 \quad (43)$$

- the THD

$$TDH \cong \left[i_{0d} / (2I_B) \right]^2 / 32 \quad (44)$$

- the third order harmonic distortion due to the threshold voltage deviation of transistors

$$HD_3 \cong \left[i_{0d} / I_B \right]^2 \cdot \Delta V_T / \left[32 (V_{GS} - V_T) \right] \quad (45)$$

4. SIMULATION RESULTS FOR A DESIGN EXAMPLE

The SPICE simulations were performed in order to point out the integrator specific performances like module-frequency and phase-frequency characteristics for differential gain, noise, distortions and dynamic range.

For simulation, level 6 PSPICE parameters furnished by American Microsystems Inc. for CMOS 0.5 μm technology were used.

In table 2 are presented the values of elements that allow the computation of transistor sizes. In this table A_{fj} and A_{gij} are the areas of floating-gate and control gate i of FG-MOS transistor j.

Table 2. The values of elements that allow the computation of transistor sizes

Element	Value	Element	Value
$(W/L)_n$	12μm/2μm	A_{fj}, $j = 3 \div 8, 12, 16$	1722μm
$(W/L)_p$	26μm/2μm	$A_{g1j}=A_{g2j}$, $j = 3 \div 8, 12, 16$	736μm
K	5	$A_{g2j}=A_{g3j}$, $j = 1,2$	736μm
A	1,5	A_{g1j}, $j = 1,2$	3680μm
B	3	A_{fj}, $j = 1,2$	5512μm
C	2pF	V_{B2}	V_{DD} =1V

For a supply voltage of 1V and a current source I$_0$ of 5μA the simulation results for module-frequency and phase-frequency characteristics for differential transfer function are presented in Fig. 9 and Fig.10. One can see that the frequency range of the current-mode fully-differential integrator is 8MHz - 800MHz for -90° ± 2.5° in phase.

The simulations also point out, for an 8 μA pick to pick sine-wave input current, a dynamic range grater than 45db and a THD less then 2% for frequencies less then 200MHz. For the THD computation was considerate a 5% mismatch in the dimensional ratio W/L of the paired FG-MOS transistors. The power consumption is around 155μW. The power supply rejection ratio is slightly grater than 50dB.

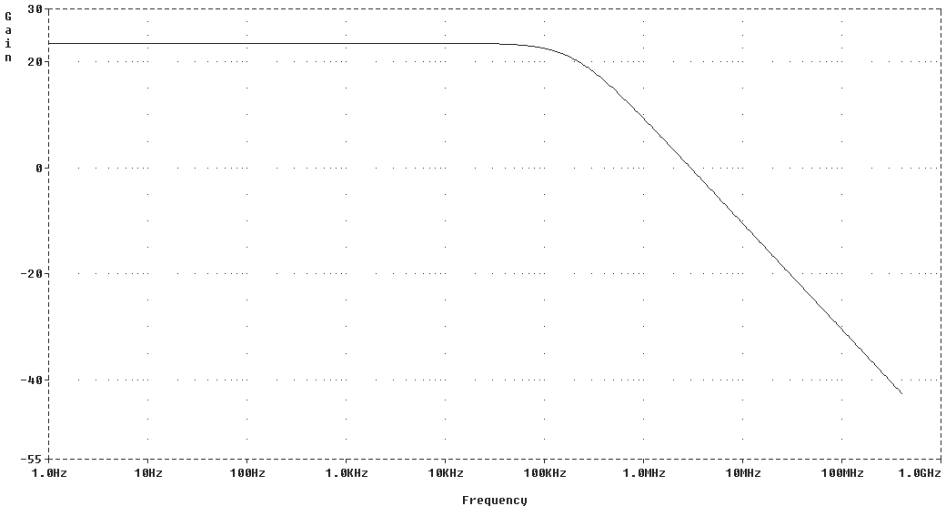


Fig. 9. The differential transfer function: module-frequency characteristic.

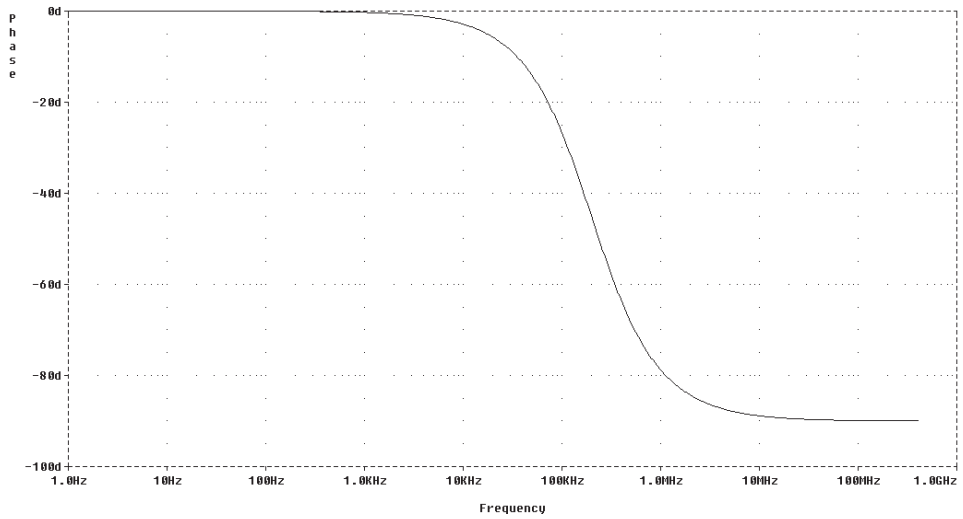


Fig. 10. The differential transfer function: phase-frequency characteristic

5. CONCLUSIONS

The suitable topologies for signal processing at low values of supply voltages are the circuit that operate in current domain because in this way the very small voltage swings are avoided. This paper presented a current mode integrator suitable for static and dynamic analogue signal processing and low supply voltage operation. By adopting a fully-differential structure, the influence of the common-mode signals and noise are minimised. By using FG-MOS transistors the circuit configuration is simplified and the voltage supply is lowered. The minimum value of required supply voltage is given by the sum of the FG-MOSFET effective threshold voltage and approximate two drain-source saturation voltages.

The SPICE simulations point out that the circuit implemented in 0.5μm standard CMOS process operate as integrator, for the 1V supply voltage, in the frequency range 8MHz-800Mz and for an 8 μA pick to pick sine-wave input current has a dynamic range grater than 45db and a THD less then 2% for frequencies less then 200MHz. For the THD computation was considerate a 5% mismatch in the dimensional ratio *W/L* of the paired FG-MOS transistors. The power consumption is around 155μW. The power supply rejection ratio is slightly grater than 50dB.

REFERENCES

Borgstrom, T.H., Ismail, M., and Bibyk, S.B. (1990). Programmable current-mode neural network for implementation in analog MOS VLSI, *IEE Proc.,* volume 137, pt. G, no.2, pp.175-184.

Gray, J., Robucci, R., and Hasler, P. (2008). The design and simulation model of an analog floating-gate computational element for use in large-scale analog reconfigurable systems. *Proceedings of the IEEE Midwest Symposium on Circuits and Systems,* volume 1, pp. 253-256, Knoxville.

Ismail, M. and Fiez, T. (1994). Analog VLSI signal and information processing, McGraw-Hill, New York.

Mondagon-Torres, A., Schneider, M., and Sanchez-Sinencio, E. (2002). Extraction of electrical parameters of floating gate devices for circuit analysis, simulation, and design. *Proceedings of the IEEE Midwest Symposium on Circuits and Systems*, volume 1, pp. I-311-I-314, Tulsa.

Rahimi, K., Diorio C., and Brockhausen, M. (2002). A simulation model for MOS synapse transistors, *Proceedings of the International Symposium on Circuit and Systems*, volume 2, pp. II-532-II-535, Phoenix

Ramirez-Angulo, J., Gonzalez-Altamirano, G., Choi, S.C. (1997). Modeling multiple-input floating-gate transistors for analog signal processing. *Proceedings of the IEEE Midwest Symposium on Circuits and Systems*, volume 3, pp. 2020–2023, Hong Kong.

Sanchez-Sinencio and Andreou (1998). *Low-Voltage/Low-Power Integrated Circuits and Systems. Low-Voltage Mixed-Signal Circuits*, IEEE Press Series on Microelectronics Systems, New York.

Sanchez-Sinencio, E. (2005). Floating gate techniques and applications, available at http://amesp02.tamu.edu/sanchez/607-2005-Floating Gate Circuits.pdf.

Sin, C.K., Kramer, A., Hu, V., Chu, R.R., and Ko, P.K. (1992). EEPROM as an analog storage device, with particular application in neural networks, *IEEE Trans. Electron Devices*, volume ED-39, no.6, pp.1410-1419.

Thomsen, A., and Brooke, M.A. (1991). A Floating-Gate MOSFET with Tunneling Injector Fabricated Using a Stansard Double-Polysilicon CMOS Process. *IEEE Electron Device Letters,* volume EDL-12, pp. 111-113.

Vittoz, E., Oquey, H., Maher, M.A., Nys, O., Dijkstra, E., and Chevroulet, M. (1991). Analog Storage of adjustable Synaptic Weights, VLSI Design of Neural Networks, Kluwer Academic, Boston.

Wang, S. T. (1979). On the I-V Characteristic of floating-gate MOS transistors, *IEEE Trans. Electron Devices*, volume ED-26, pp. 346-348.

Yang, H., Sheu, B.J., and Lee, J.C. (1992). A Nonvolatile Analog Neural Memory Using Floating-Gate MOS Transistors. *Analog Integrated Circuits Signal Process*, volume 2, pp. 19-25.

Yang, K., and Andreou, A.G. (1993). Multiple Intput Floating-Gate MOS Differential Amplifiers and Applications for Analog Computation, Proceedings of the Thirty-Sixth IEEE Midwest Symposium on Circuits and Systems, pp.1212-1216, Detroit.

On the robust stability in the presence of input disturbance of a pilot-vehicle system using a controller synthesized by the H_∞ method

Ionel Iorga *

* *Doctoral School of Control Engineering and Computers, University of Craiova, A. I. Cuza str., no. 13, RO - 200585 Craiova, Romania (e-mail: ioneliorga@yahoo.com).*

Abstract: In this paper the robust stability of a pilot-aircraft system with input disturbances, using a robust controller synthesized by the H_∞ method, is studied. The aircraft is modelled by the ADMIRE short-period dynamics together with an unsaturated first-order actuator and the human operator influence is represented through the use of a proportional-integrator type element. Through a trial of numerical cases considered, the robustness of the proposed control law is investigated in the context of step, sinusoidal and white noise input disturbances.

Keywords: stability robustness, H_∞ method, input disturbances, short-period dynamics, algebraic Riccati equation, white noise

1. INTRODUCTION

Considering the linear time invariant system

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + Dd(t) \\ y(t) &= Cx(t) \end{aligned} \tag{1}$$

the robust stability is put into discussion in the context of disturbance input.

Remark 1. In the above system the used notations have the following meaning:

- x represents the state vector;
- u is the control input;
- d denotes the disturbance input;
- y the output (measured) vector;
- $A \in R^{n\times n}$, $B \in R^{n\times m}$;
- $C \in R^{p\times n}$, $D \in R^{n\times m}$.

In the context of system 1, when using the next performance output

$$\|z\|^2 = x^T Qx + u^T Ru \tag{2}$$

then the Figure 1 describes the general problem.

2. THEORETICAL CONSIDERATIONS

In what follows, a series of definitions is given for reviewing the theoretical frame.

Definition 2. The pair (A, B) is stabilizable if $\exists\ K$, $K \in R^{m\times p}$ so that $A - BK$ is (asimptotically) stable.

Definition 3. The pair (A, C) is detectable if $\exists\ L$, $L \in R^{m\times n}$ so that $A - LC$ is stable.

Definition 4. The system (1) is closed-loop stabilizable if $\exists K \in R^{m\times p}$ so that $A - BKC$ is stable.

Definition 5. The gain L_2 of the (1) is bounded by γ if

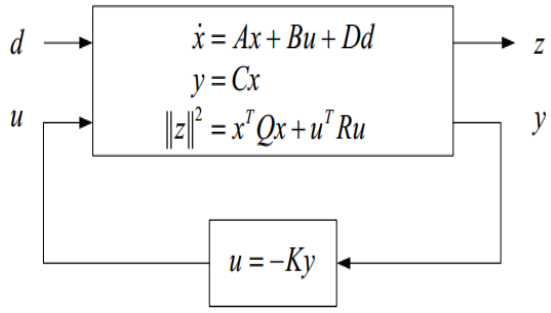


Fig. 1. System with perturbations (after Khalaf et al. (2006))

$$\frac{\int_0^\infty \|z\|^2}{\int_0^\infty \|d\|^2} = \frac{\int_0^\infty (x^T Qx + u^T Ru)dt}{\int_0^\infty (d^T d)dt} \le \gamma^2 \tag{3}$$

The L_2-gain problem is to find a time-invariant gain K so that

$$u = -Ky = -KCx \tag{4}$$

and it is desirable that the closed-loop system to be stable and the L_2-gain to be bounded by a prescribed value γ.
In what follows, the parsimony principle will be applied, skipping too detailed explanations (for thoroughgoing study the reader is sent to the mentioned reference).
Let the following theorem:

Theorem 6. Khalaf et al. (2006) - *necessary and sufficient conditions for closed loop static control obtained by H_∞ method.* Let $Q \ge 0$ and $(A, \sqrt{Q})$ detectable then the system (1) is closed-loop stabilizable, with L_2 bounded by γ, $\Leftrightarrow$

- (A, B) is stabilizable and (A, C) is detectable

- $\exists K^*$ and L so that

$$K^*C = R^{-1}(B^TP + L) \tag{5}$$

where $P > 0$, $P^T = P$ is a solution of the Riccati algebraic matrix equation

$$PA + A^TP + Q + \frac{1}{\gamma^2}PDD^TP - PBR^{-1}B^TP + \\ + L^TR^{-1}L = 0 \tag{6}$$

From the above mentioned considerations the demonstration is skipped and an algorithm for determination of $K = K^*$ is given (Khalaf et al. (2006))

- Input γ, Q, R; Output: K
- Initialise $L_0 = 0$, $L \in R^{m\times n}$
- At step $(n > 0)$ solve $P_n = P$ from (6) and evaluate

$$\begin{aligned} K_{n+1} &= R^{-1}(B^TP_n + L_n)C^T(CC^T)^{-1} \\ L_{n+1} &= RK_{n+1}C - B^TP_n \end{aligned} \tag{7}$$

if K_{n+1} and K_n are sufficiently close go to final step, otherwise increment current step: $n = n + 1$
- Finish:

$$K = K^* = K_{n+1} \tag{8}$$

3. LOW-ORDER ADMIRE DYNAMICS: THE SIMPLIFIED GAM-ADMIRE

In this section, basing on the simplified ADMIRE aircraft model - Balint and Balint (2011) - together and the "low-order GAM-ADMIRE" theoretical aircraft [Ioniţă et al. (2008)] - the quotes are putted because this is a nomenclature which has not been used in the original cited paper - the simplified GAM-ADMIRE model is described [Iorga (2013)] as follows:

$$\begin{aligned} \dot\alpha &= \zeta_\alpha\alpha + q + \frac{g}{V_0}\cos\theta + \zeta_{\delta_e}\delta_e \\ \dot q &= \mu_\alpha\alpha + \mu_q q + \mu_f\alpha q + \frac{g}{V_0}(\mu_c\cos\theta + \mu_s\sin\theta) + \mu_{\delta_e}\delta_e \\ \dot\theta &= q \end{aligned} \tag{9}$$

Remark 7. In this paper α is angle of attack, θ - pitch angle, $q = \frac{d\theta}{dt}$ (see for example Etkin and Reid (1996)) and, for simplified GAM-ADMIRE, δ_e is the turning angle of the elevon.

The model (9) is a synthetic one, his coefficients will take either one of the two values of each cell of Table 1.

$\zeta_\alpha \in \left\{ \begin{matrix} -.7986, \\ -1.6 \end{matrix} \right\}$	$g = 9.81\ \frac{m}{s^2}$	$V_0 \in \left\{ \begin{matrix} 84.5, \\ 84.569 \end{matrix} \right\} \frac{m}{s}$
$\zeta_{\delta_e} \in \left\{ \begin{matrix} -.2603, \\ -0.5209 \end{matrix} \right\}$	$\mu_\alpha \in \left\{ \begin{matrix} -6.5315, \\ 1.7251 \end{matrix} \right\}$	$\mu_{\delta_e} \in \left\{ \begin{matrix} -8.2668, \\ -9.9729 \end{matrix} \right\}$
$\mu_q \in \left\{ \begin{matrix} -.6957, \\ -22.61 \end{matrix} \right\}$	$\mu_c \in \left\{ \begin{matrix} -.162, \\ -5.2642 \end{matrix} \right\}$	$a \in \left\{ \begin{matrix} -.2424, \\ -.485 \end{matrix} \right\}$
$\mu_f \in \left\{ \begin{matrix} 4.1254, \\ 0 \end{matrix} \right\}$	$\mu_s \in \left\{ \begin{matrix} -1.424, \\ -.7154 \end{matrix} \right\}$	

Table 1. Coefficients for the simplified GAM-ADMIRE model

Remark 8. As a mnemonic the displayed values of each cell of the Table 1 are structured as follows

$$\left\{ \begin{matrix} \text{low-order GAM-ADMIRE,} \\ \text{simplified ADMIRE model} \end{matrix} \right\} \tag{10}$$

3.1 Symbolic determination of the equilibrium points

With an arbitrary, but fixed $\bar\delta_e$, bounded by

$$|\bar\delta_e| \le \frac{\sqrt{\varrho_a^2 + \varrho_b^2}}{|\varrho_c|} \tag{11}$$

- where the next notations (12) were employed -

$$\begin{cases} \varrho_a = \frac{g}{V_0}(\mu_c - \frac{\mu_\alpha}{\zeta_\alpha}) \\ \varrho_b = \mu_s\frac{g}{V_0} \\ \varrho_c = \mu_{\delta_e} - \frac{\zeta_{\delta_e}}{\zeta_\alpha}\mu_\alpha \end{cases} \tag{12}$$

the following results were obtained

$$\bar\theta_{\{1,\ 2\}} = 2\tan^{-1}(\gamma_{a_{\{1,2\}}}) \tag{13}$$

where

$$\bar\alpha = -\frac{1}{\zeta_\alpha}\left(\frac{g}{V_0}\cos\bar\theta + \zeta_{\delta_e}\bar\delta_e\right) \tag{14}$$

$$\gamma_{a_{\{1,\ 2\}}} = \frac{-\varrho_b \pm \sqrt{\varrho_a^2 + \varrho_b^2 - \varrho_c^2\bar\delta_e^2}}{\bar\delta_e\varrho_c - \varrho_a} \tag{15}$$

3.2 The linearized system of the simplified GAM-ADMIRE model

The system (9) is linearized around the equilibrium point $\bar x = (\bar\alpha, 0, \bar\theta)^T$. The resulted linearized system has the form

$$\begin{cases} \Delta\dot x = A_\Lambda\Delta x + b_\Lambda\Delta\delta_e \\ \Delta y = c_\Lambda^T\Delta x \end{cases} \tag{16}$$

where:

$$A_\Lambda = \begin{pmatrix} \zeta_\alpha & 1 & -\frac{g}{V_0}\sin\bar\theta \\ \mu_\alpha & (\mu_q + \mu_f\bar\alpha) & \frac{g}{V_0}(\mu_s\cos\bar\theta - \mu_c\sin\bar\theta) \\ 0 & 1 & 0 \end{pmatrix} \\ \Delta x = \begin{pmatrix} \Delta\alpha \\ \Delta q \\ \Delta\theta \end{pmatrix};\ b_\Lambda = \begin{pmatrix} \zeta_{\delta_e} \\ \mu_{\delta_e} \\ 0 \end{pmatrix};\ c_\Lambda^T = (0\ 0\ 1) \tag{17}$$

Remark 9. In (16) the following notations were been used:

$$\Delta x = x - \bar x;\ \Delta\delta_e = \delta_e - \bar\delta_e \tag{18}$$

3.3 Short-period approximation with first-order unsaturated actuator model

For system (16) the short-period approximation is used - see, for example, Ioniţă (2009) - resulting

$$\begin{aligned} \Delta\dot\alpha &= \zeta_\alpha\Delta\alpha + \Delta q + \zeta_{\delta_e}\Delta\delta_e \\ \Delta\dot q &= \mu_\alpha\Delta\alpha + \mu_{qf}\Delta q + \mu_{\delta_e}\Delta\delta_e \end{aligned} \tag{19}$$

where

$$\mu_{qf} \overset{def}{=} \mu_q + \mu_f\bar\alpha \tag{20}$$

Remark 10. A natural question is if always someone can use the short-period dynamics of an airplane in order to draw conclusions about the stability of the pilot-aircraft system on the longitudinal axis. The answer is not positive because the modes of the longitudinal motion (the short-period and the phugoid) can not be always decoupled. A precise answer is given by the theory singular perturbations - Vidyasagar (1978) - or by some necessary and sufficient conditions for decoupling - Falb and Wolovich

(1967). The following assertion is considered: when $\theta = \alpha$, i.e. when the airplane is over the runway in the process of landing or when just flies in a straight line, the longitudinal dynamics can be decoupled.

The first order actuator model with SAS (**S**tability **A**ugmentation **S**ystem) is

$$\dot{\delta}_e = -\omega_0(\kappa_\alpha \alpha + \kappa_q q + \delta_e) \tag{21}$$

is considered, where $\{\kappa_\alpha,\ \kappa_q > 0\}$ are the SAS gains and $\omega_0 > 0$ represents the actuator constant.
From (19) and (21) the following model is derived

$$\begin{aligned} \Delta\dot{\alpha}(t) &= \zeta_\alpha \Delta\alpha(t) + \Delta q(t) + \zeta_{\delta_e}\Delta\delta_e(t) \\ \Delta\dot{q}(t) &= \mu_\alpha \Delta\alpha(t) + \mu_{qf}\Delta q(t) + \mu_{\delta_e}\Delta\delta_e(t) \\ \Delta\dot{\delta}_e(t) &= -\omega_0 \Delta\delta_e(t) + u(t) \end{aligned} \tag{22}$$

where $u(t)$ has components from feedforward and feedback (SAS) paths and will be explicited in the next section.

4. PROBLEM MODEL

In what follows, the robust stability in the presence of input disturbances, is putted into discussion, for a system expressed by: unsaturated actuator model - including SAS - together with short period approximation dynamics of the simplified GAM-ADMIRE theoretical aircraft (22), proportional-integrator pilot model and low-pass filter for attenuating the frequencies higher than the cutoff frequency.

Considering the next figure

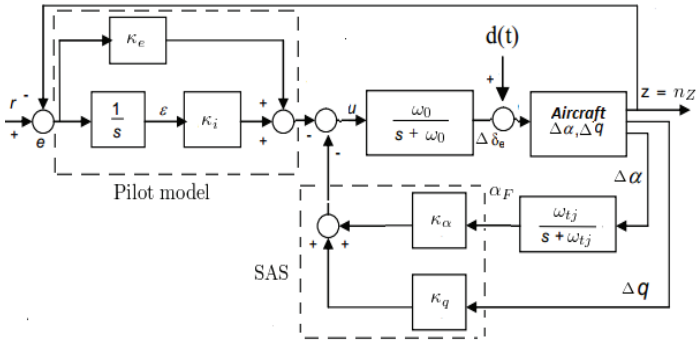


Fig. 2. System with perturbations of 5^{th} order (adapted from Khalaf et al. (2006))

then the associated open loop system is in the form of (1)

$$\begin{aligned} \frac{dx_1}{dt} &= Ax_1 + Bu + Dd \\ y &= Cx_1 \end{aligned} \tag{23}$$

$$A = \begin{pmatrix} \zeta_\alpha & 1 & \zeta_{\delta_e} & 0 & 0 \\ \mu_\alpha & \bar{\mu}_{qf} & \mu_{\delta_e} & 0 & 0 \\ 0 & 0 & -\omega_0 & 0 & 0 \\ \omega_{tj} & 0 & 0 & -\omega_{tj} & 0 \\ \xi_\alpha & \xi_q & \xi_{\delta_e} & 0 & 0 \end{pmatrix},\ B = \begin{pmatrix} 0 \\ 0 \\ \omega_0 \\ 0 \\ 0 \end{pmatrix}$$

$$D = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix},\ C = \begin{pmatrix} 0 & 0 & 0 & 57.3 & 0 \\ 0 & 57.3 & 0 & 0 & 0 \\ \xi_\alpha & \xi_q & \xi_{\delta_e} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{24}$$

with:

- the state-vector - $x_1^T = (\Delta\alpha\ \Delta q\ \Delta\delta_e\ \alpha_F\ \varepsilon)$;
- α_F represents the angle of attack measured through a low-pass filter, for reducing the noise-errors (perturbations d), with the cut-off frequency ω_{tj};
- ε - the output from the forward path for reducing the stationary errors;
- the output vector is $y^T = (\alpha_F\ \Delta q\ e\ \varepsilon)$.

Remark 11. $\bar{\mu}_{qf}$ from matrix A is the arithmetic average of the expressions (20) - for every value of $\bar{\alpha}$ (14).

Remark 12. The value 57.3, from the matrix C, is added for converting α_F and q from radians to degrees.

Remark 13. The pilot model embedded in the previous figure is proportional-integrator type - Anderson and Page (1995), McRuer and Krendel (1974) - and consist of

- fixed gain κ_e,
- integrator for reducing the stationary error (together with the constant gain κ_i)

in the feedforward path.

Remark 14. The closed loop system associated to system (23) is

$$\begin{aligned} \frac{dx_1}{dt} &= Ax_1 + Bu + Dd \\ y &= Cx_1 \\ u &= -k^T y \end{aligned} \tag{25}$$

where

$$k^T = (\kappa_\alpha\ \kappa_q\ \kappa_e\ \kappa_i) \tag{26}$$

Remark 15. Although the perturbations affect the system as described in the previous figure (on the turning angle of the elevon) the procedure employed from Khalaf et al. (2006) uses the system (1) which corresponds to the assumption that the disturbance acts on input, like in the next figure

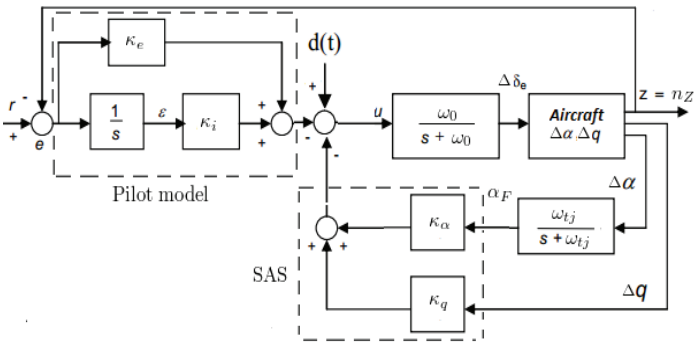


Fig. 3. System with perturbations of 5^{th} order

Remark 16. $z = n_Z$ from the previously mentioned figures is the normal acceleration - obtained through change of the turning angle of the elevon (or depth rudder - altitude control) at constant speed - Ioniţă (2009).

Remark 17. The robust stability in the presence of input disturbances of the considered pilot-vehicle system using a controller synthesized by the H_∞ method - by employing a system of the form (1) - is verified by simulations, where the input disturbances acts on the turning angle of the elevon (which is the input to the short period simplified GAM-ADMIRE model).

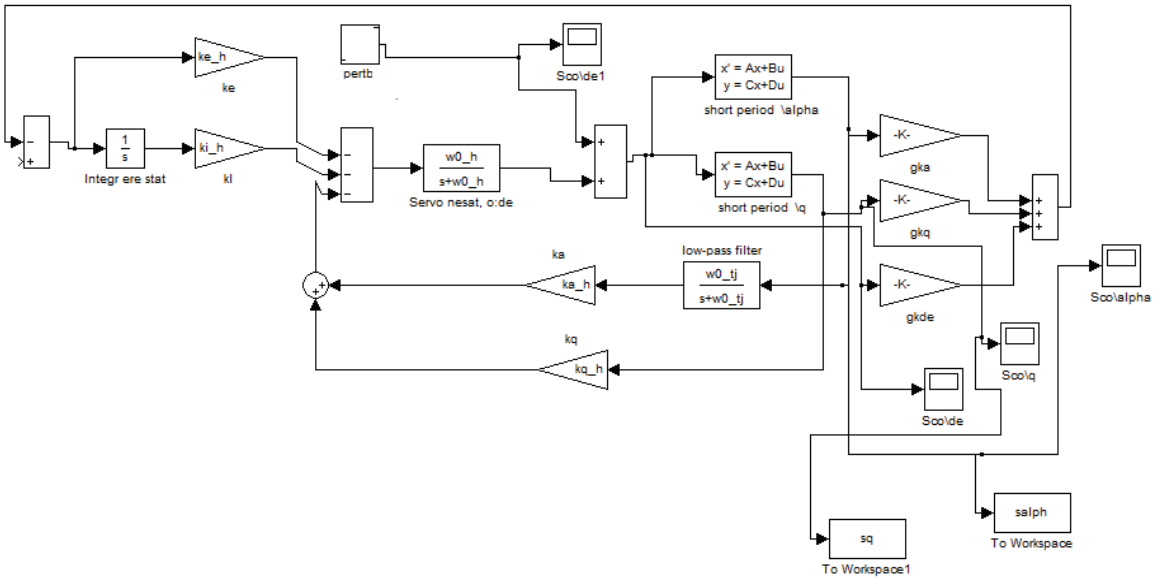


Fig. 4. The SIMULINK sheme used for simulations

Remark 18. The simulink scheme used in this paper is the one shown in the previous figure.

5. NUMERICAL CONSIDERATIONS

From Khalaf et al. (2006), the values used for computing the normal acceleration are

$$(\xi_\alpha \ \xi_q \ \xi_{\delta_e}) = (-13.26 \ \ -0.9788 \ 0.4852) \tag{27}$$

and the cut-off frequency is selected to be $\omega_{tj} = 10$ rad/s. In the case of low-order GAM-ADMIRE model, the system (23) is evaluated to

$$A = \begin{pmatrix} -0.7986 & 1 & -0.2603 & 0 & 0 \\ -6.5315 & -0.0632 & -8.2668 & 0 & 0 \\ 0 & 0 & -\omega_0 & 0 & 0 \\ 10 & 0 & 0 & -10 & 0 \\ -13.26 & -0.9788 & 0.4852 & 0 & 0 \end{pmatrix}, \ B = \begin{pmatrix} 0 \\ 0 \\ \omega_0 \\ 0 \\ 0 \end{pmatrix} \tag{28}$$

and, in the case of simplified ADMIRE theoretical aircraft, (23) is represented by

$$A = \begin{pmatrix} -1.6 & 1 & -0.5209 & 0 & 0 \\ 1.7251 & -22.61 & -9.9729 & 0 & 0 \\ 0 & 0 & -\omega_0 & 0 & 0 \\ 10 & 0 & 0 & -10 & 0 \\ -13.26 & -0.9788 & 0.4852 & 0 & 0 \end{pmatrix}, \ B = \begin{pmatrix} 0 \\ 0 \\ \omega_0 \\ 0 \\ 0 \end{pmatrix} \tag{29}$$

Remark 19. In systems (29) and (28) the value of the actuator is constant: $\omega_0 = 20 \ \frac{rad}{s}$.

Remark 20. In the case of low-order GAM-ADMIRE model, $\bar{\mu}_{qf} = -0.0632$ and, in the case of simplified ADMIRE model, $\bar{\mu}_{qf} = -22.61$ (see Remark 11).

Remark 21. In the both mentioned models, the D matrix from (23) is expressed by

$$D = \begin{pmatrix} 0 & 0 & 0 & 57.3 & 0 \\ 0 & 57.3 & 0 & 0 & 0 \\ -13.2600 & -0.9788 & 0.485 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{30}$$

The matrices Q and R, from (2), are choosed in the following manner

$$Q = \begin{pmatrix} 264 & 16 & 1 & 0 & 0 \\ 16 & 60 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 100 \end{pmatrix}, \quad R = (0.1) \tag{31}$$

and the scalar from inequality (3) is $\gamma = 0.2$.

In the case of low-order GAM-ADMIRE model, after the previous defined algorithm is applied, K^T is evaluated to:

$$K^T = (0 \ \ -0.260329 \ 8.13985 \ 31.722063)$$

with the following closed-loop poles

$$\{-46.6908 \pm \imath 44.07654887, \ -3.2346 \pm \imath 2.769, \ -10\}$$

For the simplified ADMIRE theoretical airplane, using the mentioned algorithm is it obtained

$$K^T = (0 \ \ -0.1562 \ 5.5214 \ 31.722)$$

with the next closed-loop poles

$$\{-43.2346 \pm \imath 53.461775, \ -5.66 \pm \imath 3.33515, \ -10\}$$

5.1 Step disturbance input

In this subsection, the step disturbance input is choosed to be 0.1222 *rad* (7 *deg*) with step time 0 and the initial value considered 0.

In the case of low-order GAM-ADMIRE the graphical answer is shown in Figure 5

and, in the case of simplified ADMIRE theoretical aircraft, the graphic is given in Figure 6

5.2 Sinusoidal disturbance input

The sinusoidal disturbance input was chosed with amplitude 0.1222 *rad*, frequency 1 and zero phase.

In the case of low-order GAM-ADMIRE the answer is plotted in Figure 7.

In the case of simplified ADMIRE theoretical aircraft, the graphical answer is depicted in Figure 8.

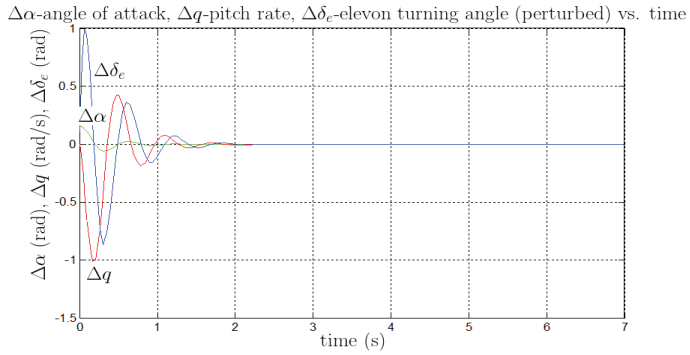


Fig. 5. Short-period low-order GAM-ADMIRE disturbance step response

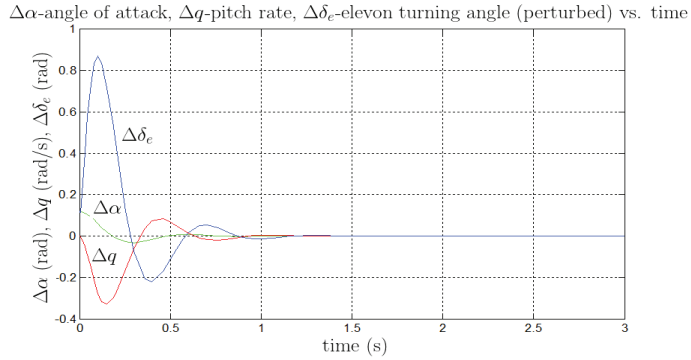


Fig. 6. Short-period simplified ADMIRE theoretical aircraft disturbance step response

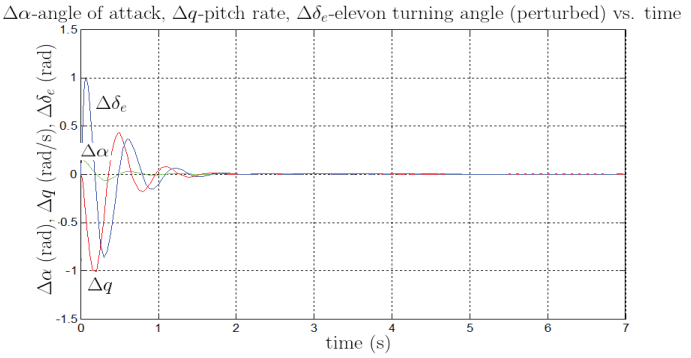


Fig. 7. Short-period low-order GAM-ADMIRE sinusoidal disturbance response

5.3 White-noise disturbance input

The white-noise disturbance input was chosed with noise-power 0.003, and corresponding maximum amplitude of 23.7 deg.

In the case of low-order GAM-ADMIRE the input was the one from Figure 9, with the output from the Figure 10.

In the case of simplified ADMIRE theoretical aircraft, the input was the one from Figure 11, with the output given in Figure 12.

6. CONCLUSIONS

Through a trial of numerical cases considered, the robustness of the proposed control law was proved in the context

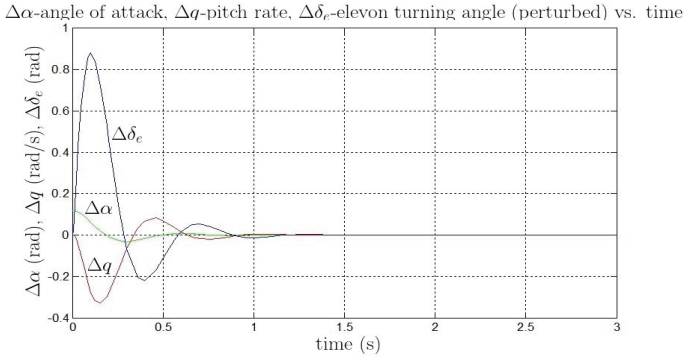


Fig. 8. Short-period simplified ADMIRE theoretical aircraft sinusoidal disturbance response

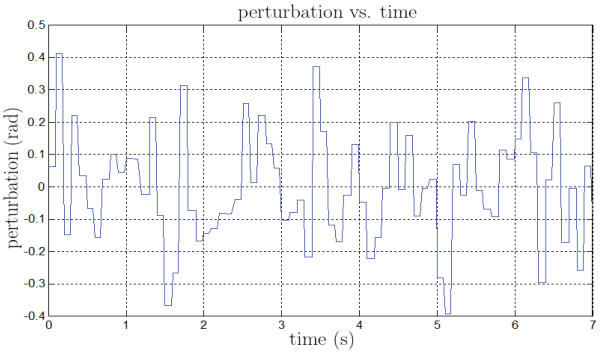


Fig. 9. Short-period low-order GAM-ADMIRE white noise disturbance input

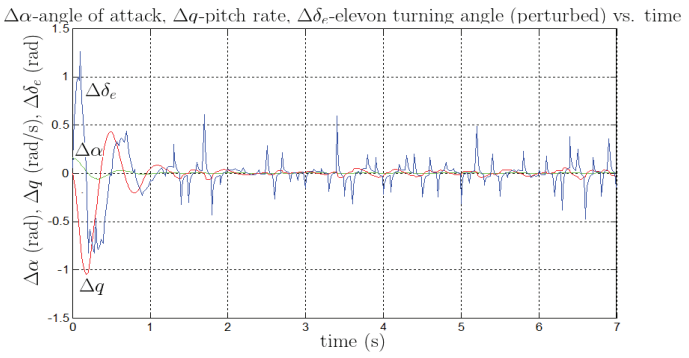


Fig. 10. Short-period low-order GAM-ADMIRE white noise disturbance response

of step, sinusoidal and white noise input disturbances.

For future work, an actual problem can be represented by the relation between the human pilot delay and the stability boundaries, in the conditions of the given system, with simulations for both nonlinear and linearized systems.

Another line of research may be done with the consideration of a bigger order dynamics (e.g coupled phugoid and short-period).

ACKNOWLEDGEMENTS

Thanks to professor Vl. Răsvan for his guidance in all my research activity.

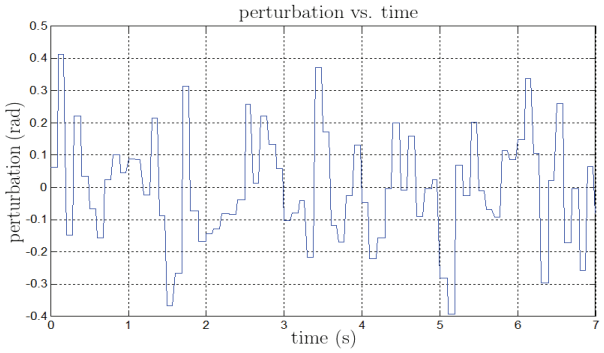


Fig. 11. Short-period simplified ADMIRE theoretical aircraft white noise disturbance input

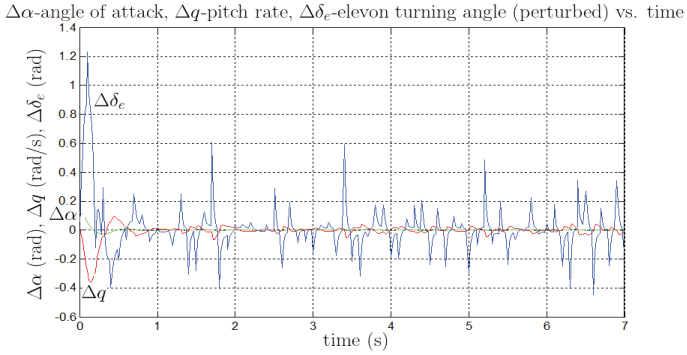


Fig. 12. Short-period simplified ADMIRE theoretical aircraft white noise disturbance response

REFERENCES

Anderson, M. and Page, A. (1995). Unified Pilot-Induced Oscillation Theory, Vol. III: PIO analysis using multivariable methods. *Wright Laboratory, Wright-Patterson Air Force Base.*

Balint, A. and Balint, S. (2011). Oscillations susceptibility of an unmanned aircraft whose automatic flight control system fails. In *Advances in Flight Control Systems*, 275–296. InTech, 1st edition.

Etkin, B. and Reid, L. (1996). *Dynamics of flight. Stability and control (3rd Edition)*. John Wiley & Sons, ISBN 0-471-03418-5, New York.

Falb, P. and Wolovich, W. (1967). Decoupling in the design and synthesis of multivariable control systems. *IEEE Transact. on Autom. Control, Vol. AC-12, No. 6*, 651–659.

Ioniţă, A. (2009). *Dinamica avionului cu întârzieri în comandă (in romanian)*. Editura Academiei Tehnice Militare, ISBN 978-973-640-177-0, Bucureşti.

Ioniţă, A., Balint, A., and Balint, S. (2008). Limit cycle behaviour in aircraft longitudinal terminal phase. 7^{th} *International Conference on Mathematical Problems in Engineering, Aerospace and Sciences, Genova*, 276–287.

Iorga, I. (2013). On the stability of a pilot-aircraft system with input delay using a controller obtained by Artstein transform (accepted). *Journal of Control Engineering and Applied Informatics.*

Khalaf, A., Huang, J., and Lewis, F. (2006). *Nonlinear H_2/H_∞ Constrained Feedback Control.* ISBN 1-84628-349-3, Springer-Verlag, London.

McRuer, D. and Krendel, E. (1974). Mathematical models of human pilot behaviour. Technical report, NATO AGARD Report No. 188.

Vidyasagar, M. (1978). *Nonlinear System Analysis.* Prentice-Hall, ISBN 0-13-623280-9, New Jersey.

Design, Modelling and Simulation of a Hexapod Robot for Basic Locomotion Strategies over Obstacles Using Matlab

Sorin Mănoiu-Olaru*, Mircea Nițulescu**

**Department of Automation, Electronics and Mechatronics, University of Craiova, Romania (e-mail:manoiusorin2006@yahoo.com)*
*** Department of Automation, Electronics and Mechatronics, University of Craiova, Romania (e-mail:nitulescu@robotics.ucv.ro)*

Abstract: In this paper the authors present a new software platform that has been made using Matlab, for studying hexapod robot stability in gravitational field and locomotion simulations over common types of obstacles and two particular forms of obstacles. For the proposed design of the leg the kinematical model was calculated and a workspace analysis of the leg was made. The trajectory generator for the leg tip was implemented using piecewise cubic spline interpolation method. Next the most common types of obstacles are presented and the influence upon robot locomotion. The analysis of the robot static stability is made for two different situations. The paper includes some simulation results related to the static gravitational stability depending on the support polygon, single leg control and locomotion over common types of obstacles.

Keywords: hexapod robot, kinematic model, Matlab, simulation, robot locomotion.

1. INTRODUCTION

In nature, most arthropods have six legs to easily maintain static stability, and it has been observed that a larger number of legs do not necessarily increase walking speed (Bensalem et al. 2009).

The large diversity of the existing walking animals offers innumerable examples of locomotion possibilities. There are two main types of legs configurations that are biological inspirited. The main difference is in the way the leg swing relative to body. The first type is similar to cats, humans and birds where the leg swings around a horizontal axis. The second type is more similar to insects where the legs swing around a vertical axis providing better stability.

Since most of the earth's surface is inaccessible to regular vehicles there is a need for mobile robots that can handle difficult terrain. Legged locomotion minimizes the terrain modification because they don't need continuous contact with the ground and legs provide the capability of manoeuvring within confined space. As can be seen from nature, legs are not used only for walking (e.g.: manipulators, sensors).

Legged walking machines show robustness in case of leg faults. An important drawback of legged machines is the complexity of the control required to achieve walking even on completely flat and horizontal surface in which much simpler wheeled machines work perfectly well.

Many scientists (Conrad et al. 2009, Jakimovsk 2011) use animal's similarities in their mechanical design for leg configuration and leg design leading often to simplifications of the actual biological system.

The number of the legs of a robotic structure affects stability (more legs on the ground mean better stability) and the workspace.

In legged animal locomotion the periodical excitation of the flexor and extensor muscles is needed in order to produce effective walking movements. For a walking structure to be able to traverse rough terrain, at least 2 DoF on each leg are necessary to move forward or backward or lift over objects.

Building a mechanical structure that mimics an animal leg is still in progress (Tao et al. 2011, Wang ZhouYi et al. 2011). The most common leg structure among current studies has 3 dof of which 2 dof are considered located in the first joint or the segment between the first and the second joint is very small.

The authors proposed a leg with 3 dof in which the segment between the first and the second joint is much longer. With this in mind the robot body was considered narrowed and longer to compensate for the enlargement of the segment.

The legged locomotion on natural terrain presents a set of complex problems (foot placement, obstacle avoidance, load distribution, general stability) (Krzysztof et al. 2008) that must be taken into account both in mechanical construction of vehicles and in development of control strategies.

One way to handle these issues is using models that mathematically describe the different situations. Therefore modelling has become a useful tool in understanding systems complexity and for testing and simulating different control approaches (Lewinger et al. 2011, Jung-Min Yang 2009, Silva et al. 2011).

2. HEXAPOD ROBOT

The successful design of a legged robot depends mostly on the mechanical characteristics of the chosen leg. Since all aspects of locomotion are ultimately governed by the physical limitations of the leg, it is important to select a mechanical design that will maximize motion and impose fewer constraints during walking.

2.1 Robot Leg Kinematics

A three-revolute kinematical chain (RRR) has been chosen for each leg mechanism in order to mimic the leg structure (Fig. 1). A direct geometrical model for each leg is formulated between the moving frame $O_i(x_i,y_i,z_i)$, i=0…3, attached to the leg base, and the fixed frame, $O_G(x_G,y_G,z_G)$, attached to the world.

The assignment of link frames follows the Denavit - Hartenberg (D-H) (Schilling 1990) direct geometrical modeling algorithm and is presented in Fig. 1.

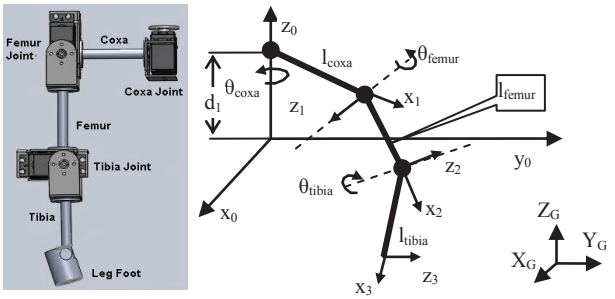


Fig. 1. CAD design of the robot's leg (left) and associate model used for kinematics (right).

The robot leg frame starts with link 0 which is the point on the robot where the leg is attached; link 1 is the *coxa*, link 2 is the *femur* and link 3 is the *tibia*. The overall transformation from coxa frame to leg tip frame is obtained as a product between three transformation matrixes:

$$T_{coxa}^{tip} = T_{coxa}^{femur} \cdot T_{femur}^{tibia} \cdot T_{tibia}^{tip} \tag{1}$$

Considering Fig. 1 and using (1) the coordinates of each leg tip are:

$$\begin{aligned} x &= [l_1 + l_2 \cdot \cos(\theta_2) + l_3 \cdot \cos(\theta_2 - \theta_3)] \cdot \cos(\theta_1) \\ y &= [l_1 + l_2 \cdot \cos(\theta_2) + l_3 \cdot \cos(\theta_2 - \theta_3)] \cdot \sin(\theta_1) \\ z &= d_1 + l_2 \cdot \sin(\theta_2) + l_3 \cdot \sin(\theta_2 - \theta_3) \end{aligned} \tag{2}$$

where:

d_1 – the distance from global frame to coxa joint frame along Z axis,
l_i – length of the coxa, femur and tibia segment,
θ_i – coxa, femur and tibia joint variables.

The centre of gravity of each link is positioned relative to the link frame by a position vector $p_i=[x_i,y_i,z_i]^T$. To find the position of the centre of gravity of each link relative to leg frame, the coordinates p_i are multiplied with the D - H transformation:

$$p_{CoG_i} = T_0^i \cdot p_i \text{ , i=1…3} \tag{3}$$

The geometrical model described at (2) establishes a connection between the joint variables and the position and orientation of the end frame. In order to transform the motion assigned to the end frame into the corresponding joint angle motions we need to calculate the joint variables, θ_1, θ_2, and θ_3, as a function of leg tip coordinates. Using (2) and considering the following constraints: all joints allow rotation only about one axis, femur and tibia always rotate on parallel axes, and the physical limitation of each joint we can determine the joint angle as follows:

$$\begin{aligned} \theta_1 &= \begin{cases} \arctan(\frac{y}{x}) + \pi, & x < 0 \\ \arctan(\frac{y}{x}), & otherwise \end{cases} \\ \theta_2 &= \arccos\left(\frac{l_2^2 + x'^2 + y'^2 - l_3^2}{2 \cdot l_2 \cdot \sqrt{(x'^2 + y'^2)}}\right) + \arctan(\frac{y'}{x'}) \\ \theta_3 &= \pi - \arccos\left(\frac{l_3^2 + l_2^2 - (x'^2 + y'^2)}{2 \cdot l_2 \cdot l_3}\right) \end{aligned} \tag{4}$$

where:

x, y – the coordinates of the leg tip in the leg frame,
x', y' – the coordinates of the leg tip in the coxa frame.

2.2 Workspace Analysis of the Leg

Workspace must be defined in order to maximize motion but in the same time to minimize singularities or any other pitfalls. Workspace analysis is useful for walking algorithm to prevent collision of adjacent legs by imposing restrictions for direct or inverse kinematics.

In order to minimize leg collisions θ_1 was limited between $-\pi/4$ and $\pi/4$. To get the most out of motion θ_2 was limited from $-\pi/2$ to $\pi/4$ and θ_3 from 0 to $3\pi/4$. In our case there is a singularity when the x and y coordinates of the leg tip are zero and θ_1 is arbitrary. To avoid this scenario we can define the x coordinate greater than zero.

In Fig. 2 are presented the extreme interest points from the workspace envelope.

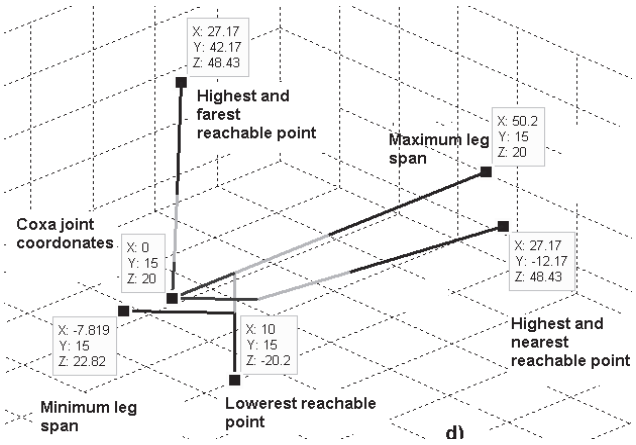


Fig. 2. Leg workspace. Extreme interest points.

2.3 Hexapod Robot Leg control

The legged locomotion on natural terrain presents a set of complex problems (foot placement, obstacle avoidance, load distribution, general stability) that must be taken into account both in mechanical construction of vehicles and in development of control strategies.

To facilitate analysis and control during movement the motion of leg tip is split in two:

- support phase (stance) in which the robot uses the leg as support and propulsion,
- transfer phase (swing) which represents the movement of the leg from one point to another.

One step consists of one stance cycle and one swing cycle but to maintain a smooth motion of the leg it's important that the swing cycle continues where the stance cycle ended, and that the swing cycle ends where the new stance cycle starts. This does not only hold for the position, but also for the speed and the direction of the motion.

If we consider a joint control level the leg tip coordinates are:

$$\begin{aligned} x &= (step\_length)\cdot\sin(\theta_1) \\ y &= (step\_length)\cdot\cos(\theta_1) \\ z &= \begin{cases} step\_height, & \text{for swing} \\ 0, & \text{for stance} \end{cases} \end{aligned} \tag{5}$$

where:

step_length – the traveled distance for the leg,
step_heigth – the distance from the ground to leg tip.

3. HEXAPOD ROBOT MODEL

It is well known that to maintain a structure stable in a 3-D space requires three point of support. Machines with three or more legs continuously in contact with the ground are said to be statically balanced if they maintain their projection (G') of the centre of gravity (G) within the polygon determined by the legs on the support plane. The polygon is known as "the support polygon" (Fig. 3). The motion of legged robots can be divided into statically and dynamically stable. Static stability means that the robot is stable at all times during its gait cycle. Dynamic stability means that the robot is stable only when it is moving. For legged robots, static stability demands that the robot has to have at least three legs on the ground at all times and the projection of the robot's centre of gravity stays inside the support polygon, i.e. the convex polygon formed by the legs supporting the robot (Fig. 3).

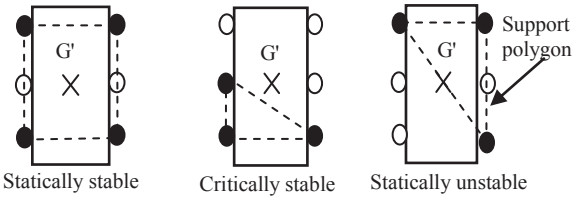


Fig. 3. Stability cases for a hexapod robot: stable, unstable and critically stable.

The robot structure considered (Fig. 4) has 6 identical legs (Fig. 1). All the relevant points have been put on the model: coordinates of the centre of gravity of each leg G_i, i=1…6, coordinates of the centre of gravity of the robot G, projection of the centre of gravity into the support polygon G', robot's centre of symmetry with the attached frame $O_R(X_R,Y_R,Z_R)$, the global frame $O_G(X_G,Y_G,Z_G)$, and direction of motion.

The overall transformation from global frame to each leg tip is obtained as follows:

$$T_G^{tip_i} = T_G^R \cdot T_R^{coxa_i} \cdot T_{coxa_i}^{tip_i}, \text{ i=1…6} \tag{6}$$

where:

T_R^G – transformation matrix between the global frame and the robot frame and defines the rotation about *Y, X, Z* (roll, pitch, yaw matrix),
$T_{coxa_i}^R$ – transformation matrix from the robot centre to coxa_i

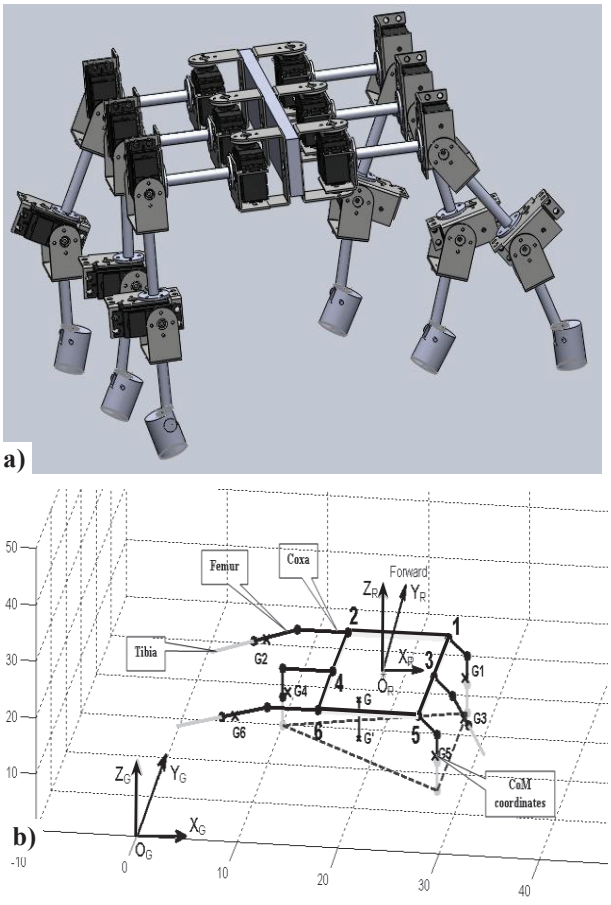


Fig. 4. CAD model of the hexapod robot (a) and the corresponding model using Matlab (b).

The position of robot's centre of gravity is calculated using the following relations:

$$X_G = \frac{\sum_{i=1}^{6} x_{gi}\cdot m_{Li}}{\sum_{i=1}^{6} m_{Li}}, \quad Y_G = \frac{\sum_{i=1}^{6} y_{gi}\cdot m_{Li}}{\sum_{i=1}^{6} m_{Li}}, \quad Z_G = \frac{\sum_{i=1}^{6} z_{gi}\cdot m_{Li}}{\sum_{i=1}^{6} m_{Li}} \tag{7}$$

where:

$m_{Li} = \sum_{j=1}^{3} m_{lj}$, m_{li} – the mass of each leg segment,

x_{gi}, y_{gi}, z_{gi} – the coordinates of the center of gravity for each leg.

4. TYPES OF OBSTACLES

Earth provides an extremely large variety of terrains, but the obstacles encountered can be classified in 4 main groups as can be seen from figure below (Fig. 5). The definition of the obstacles types will allow the development of better locomotion strategies for each case.

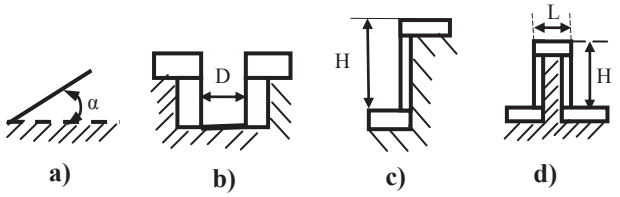


Fig. 5. Common types of obstacles. a) Inclined plane; b) Ditch; c) Step; d) Crest.

Next we will present a short description of each type of obstacle highlighting the most important parameters that influence the robot motion and what constrains may introduce during locomotion.

4.1 Locomotion on Inclined Plane

The main difference between locomotion on level ground and inclined plane is that the projection of the centre of gravity on to the support polygon shifts. Doing that the static stability, which is the minimum distance from the projection of the centre of gravity to the polygon side, becomes smaller. Locomotion on inclined plane can be achieved in two ways: along the longitudinal axis of the plan (Fig. 6a) and across the plane (Fig. 6b). In Fig. 6 we noted longitudinal stability at $\alpha=0$ with S_0, longitudinal stability on inclined plane with S, the dotted bar, the height of the robot with H and the angle of the robot across the plane with β.

The total stability on inclined plane (S) for Fig. 6.a is:

$$S = (S_0 - H \cdot \tan\alpha) \cdot \cos\alpha \tag{8}$$

From (8), considering $S = 0$, we can determine the maximum height of the robot on the inclined planes (H_{max}) or the maximum angle of the plane that the robot traverses (α_{max}):

$$H_{\max} = S_0 \cdot \tan\alpha \qquad \alpha_{\max} = \arctan\frac{S_0}{H} \tag{9}$$

For case b mentioned in Fig. 6 the projection of the centre of gravity is shifted across with $H \cdot \tan\alpha$.

This distance can't be greater than the half of the width (l) of the robot:

$$H \cdot \mathrm{tg}\,\alpha = \frac{l}{2} \tag{10}$$

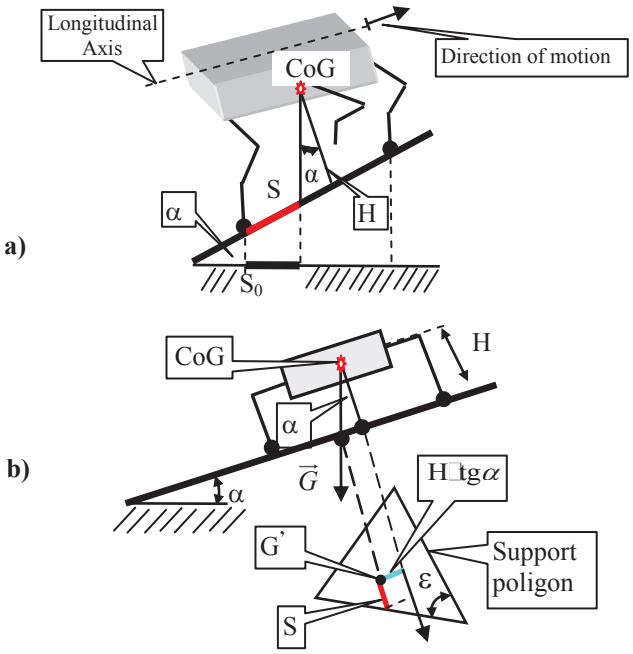


Fig. 6. Locomotion over inclined plane. a) Along the plane; b) Across the plane.

From (10) we can obtain the maximum angle of the plane without the robot becoming unstable and the static stability:

$$\alpha_{\max} = \arctan\frac{l}{2 \cdot H} \tag{11}$$

The total stability on inclined plane (S) is:

$$S = S_0 - H \cdot \frac{\tan\alpha}{\tan\varepsilon} \tag{12}$$

For certain angles of the plane we can calculate the maximum height of the robot:

$$H_{\max} = S_0 \cdot \frac{\tan\varepsilon}{\tan\alpha} \tag{13}$$

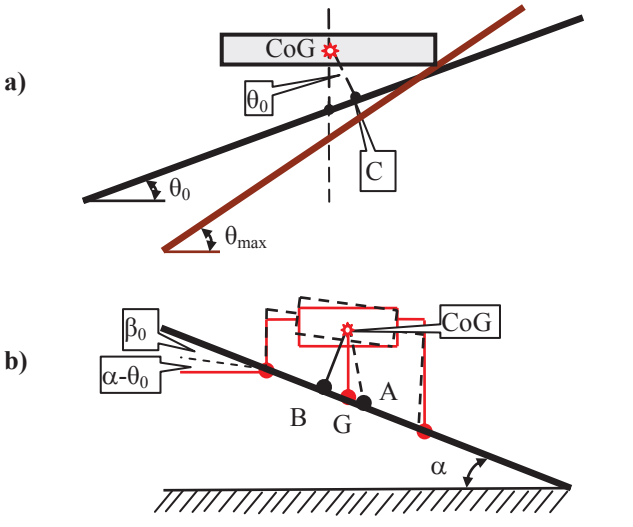


Fig. 7. Locomotion over inclined plane maintaining the robot structure parallel with the ground. a) Along the plane; b) Across the plane.

Another way to deal with this type of obstacle is for the robot to maintain its body parallel with the ground as in Fig. 7.

For the first case presented in Fig. 7a the static stability is the same with the one on the horizontal plane. The robot can be position parallel with the ground on the inclined plane only if:

$$\theta_{\max} = \operatorname{arctg}\frac{2 \cdot R_z}{l} \tag{14}$$

where:

R_z – represents the step length on the Z axis;
l– width of the robot body.

If the angle of the inclined plane, α, is greater than the maximum tilt of the robot, α_0 the static stability in this case is:

$$S = \left(\frac{S_0}{\cos\alpha} - dist\left(CoG, C\right) \cdot \left(\tan\alpha - \tan\theta_0\right) \right) \cdot \cos\alpha \tag{15}$$

where:

$dist\left(CoG, C\right) = \left(H - \frac{R_z}{2} \right) \cdot \cos\left(\theta_0\right)$ - distance from the centre of gravity to the inclined plane.

For the second case presented in Fig. 7b the maximum tilt of the robot is:

$$\beta_0 = \arctan\frac{R_z}{l} \tag{16}$$

If the angle of the plane is greater than β_0 the body of the robot can not be made parallel with the ground. In this case the static stability and the maximum angle of the inclined plane are:

$$S = S_0 - \frac{AG}{\tan\left(\varepsilon\right)}, \text{ for } BG \le \frac{l'}{2} \tag{17}$$

$$\alpha_{\max} = \operatorname{arctg}\left(\frac{S_0 \cdot \tan\varepsilon}{OB} + \tan\beta_0 \right) \tag{18}$$

where:

$$l' = \frac{l}{\cos\beta_0},$$

$$AG = OB \cdot (\operatorname{tg}\alpha - \operatorname{tg}\beta_0) - R_y \cdot \frac{\cos\beta_0}{2},$$

$$OB = \left(H_0 - \frac{R_z}{2} \right) \cdot \cos\beta_0 - R_y \cdot \frac{\sin\beta_0}{2}.$$

4.2 Locomotion Over a Step

A step is only defined by its height as can be seen from the Fig. 5c. There can be 3 phases of locomotion during overcoming this type of obstacle. The first one is presented next.

The maximum tilt of the robot is achieved when the rear part of the robot touches the ground:

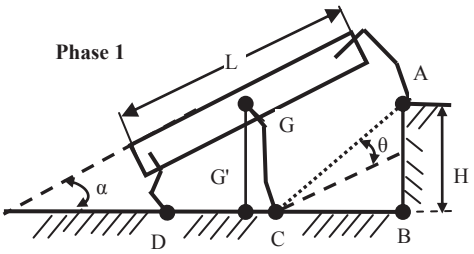


Fig. 8. Locomotion over a step. Maximum step length for front legs.

$$\theta_{\max} = \operatorname{arctg}\left(\frac{2 \cdot R_{z0}}{L - R_{x0}} \right) \tag{19}$$

where:

L – length of the robot,
R_{z0} – step heigth,
R_{x0} – step length.

The middle legs are considered extended to the maximum and placed at a distance BC from the step. The rear legs can be position anywhere after point G' and by doing the calculations, the height step and BC distance are:

$$\begin{aligned} H &= AC \cdot \sin(\alpha + \theta) \\ BC &= AC \cdot \cos(\alpha + \theta) \end{aligned} \tag{20}$$

where:

$\theta = \operatorname{arctg}\frac{R_{z0}}{P_0 + R_{x0}}$,

$\mathrm{AC} = \frac{R_{z0}}{\sin\theta}$,

P_0 – the longitudinal distance of stepping.

In phase two of climbing (Fig. 9) the four legs are near the top of the step and the middle legs are on the top of the step. The projection of the centre of gravity is on BC and like in the previous case the height of the step and the length of BC are the same. In order to maintain the robot stable the rear legs must position after G'.

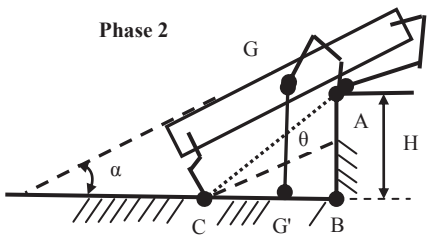


Fig. 9. Locomotion over a step. Maximum step length for middle legs.

$$\begin{aligned} H &= AC \cdot \sin(\alpha + \theta) \\ BC &= AC \cdot \cos(\alpha + \theta) \end{aligned} \tag{21}$$

where:

$$\theta = \operatorname{arctg}\frac{R_{zo}}{P_0 + \frac{R_{xo}}{2}},$$

$\mathrm{AC} = \frac{R_z}{\sin\theta}$

From Fig. 10 it can be seen that if we raise the height of the step, the length of *BC* decreases. The maximum step length of rear legs corresponds when the centre of gravity is over the step edge. The expressions for step height and *BC* length are the same. Also the maximum tilt of the robot is when *BC* equals zero.

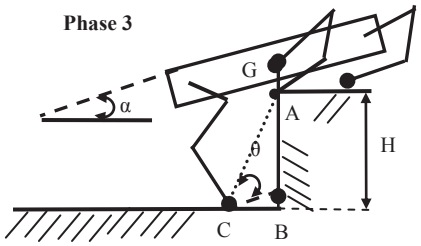


Fig. 10. Locomotion over a step. Maximum step length for rear legs.

$$H = AC \cdot \sin(\alpha + \theta)$$
$$BC = AC \cdot \cos(\alpha + \theta) \tag{22}$$

where:

$$\theta = \operatorname{arctg} \frac{R_{z0}}{P_0 + \frac{R_{x0}}{2} - (H_0 - R_{z0}) \cdot \operatorname{tg}\alpha},$$

$$\mathrm{AC} = \frac{R_{z0}}{\sin\theta},$$

H_0 – maximum height of the robot.

If we compare the 3 cases presented results that for a certain tilt of the robot, the height of the step is the smallest in the third case and also the maximum height of the step.

4.3 Locomotion Over a Ditch

A ditch obstacle (Fig. 5b) is characterized only by one parameter, its width (*D*). The positioning of 2 legs from one side can not be simultaneously made in the same point but in a neighbourhood (Fig. 11). Next the strategy that determines the maximum width of the ditch was considerate:

$$D = 0.5 \cdot \lambda + \delta \tag{23}$$

where:

λ – the total travelled distance of each leg,

δ – the positioning error relative to the corresponding front leg.

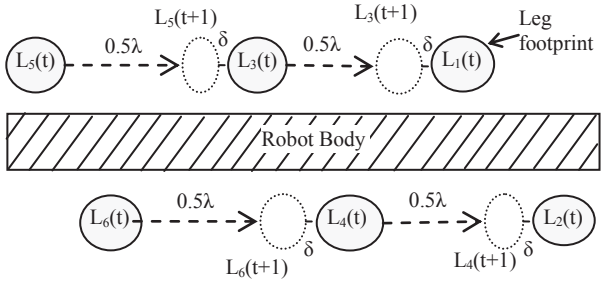


Fig. 11. Evolution of leg footprints.

Using this strategy the motion can be divided into the following locomotion phases (Fig. 12):

- the positioning of the front legs from the initial plane to the final plane (Fig. 12b),
- positioning of the middle legs to the final plane (Fig. 12c),
- the transfer of the rear legs from the initial plane to the final plane (Fig. 12d).

Between these phases, mentioned above, additional movements take place to ease the overcoming of this type of obstacle.

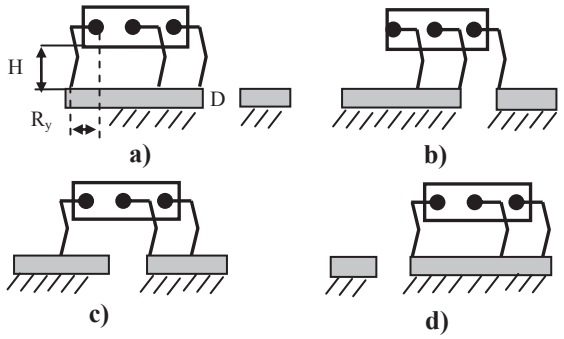


Fig. 12. Locomotion over a ditch. a) Initial position; b) Transfer of front legs over the ditch; c) Transfer of rear legs over the ditch; d) Final position of the robot.

In order for the robot to approach this type of motion strategy the maximum width of the ditch must be:

$$D = P_0 + \frac{R_y}{2} \tag{24}$$

where:

P_0 – the longitudinal distance of stepping,

R_y – the length of the step on Y axis.

If we consider the horizontality of the structure and maximum height of the robot, the maximum width of the ditch is:

$$D_{\max} \le 2 \cdot l_1 \cdot \sin\left(\frac{\theta_1}{2}\right), \quad H = l_2 + l_3 \tag{25}$$

If we consider a variable height of the robot the width of the ditch is:

$$D_{\max} < 2 \cdot [l_1 + l_2 \cdot \cos(\theta_2) + l_3 \cdot \cos(\theta_2 - \theta_3)] \cdot \sin\left(\frac{\theta_1}{2}\right) \tag{26}$$

4.4 Locomotion Over a Crest

The crest type obstacle is presented in Fig. 5d and is characterized by 2 parameters: width (*L*) and height (*H*). If the width of the crest is large enough to support the simultaneous positioning of 2 pair of legs from one side the locomotion strategy is the same as ascending and descending a step (presented earlier).

We considered that this condition is not fulfilled and the strategy proposed for overcoming this type of obstacles is presented in Fig. 13. Another restriction introduced was that the robot body has to remain parallel with the ground during locomotion.

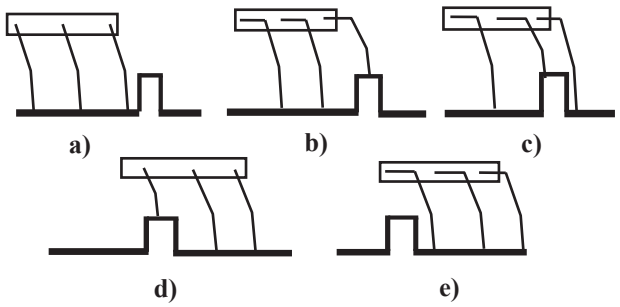


Fig. 13. Locomotion over a crest. a) Initial position; b) Position front legs on the crest; c)Transfer of front legs over the crest and positioning of the middles legs on the crest; d)Transfer of middle legs over the crest and positioning of the rear legs on the crest. e)Transfer of rear legs on the ditch.

The strategy for leg movement presented in Fig. 13 is similar with the locomotion over a ditch. By imposing the horizontality of the body during motion we can differentiate 3 heights for the crest: $H_{crest} \sim H_{robot}$, $H_{crest} << H_{robot}$ and $0 < H_{crest} < 0.5 \cdot H_{robot}$. The last condition is more practical and it is explained below (Fig. 14).

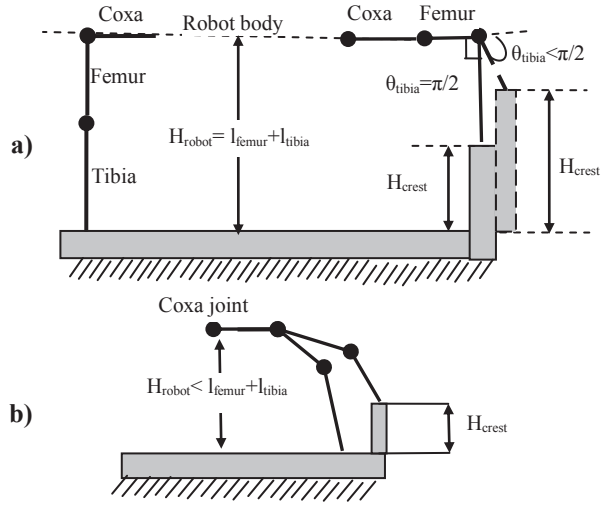


Fig. 14. Calculus of the crest height.

According to Fig. 12a the height of the crest is:

$$H_{crest} = H_{robot} - l_3 \cdot \sin(\theta_3)$$
$$H_{robot} = l_2 + l_3 \quad (27)$$

For case b the height of the crest is:

$$H_{crest} = H_{robot} - \left[l_2 \cdot \sin(\theta_2) + l_3 \cdot \sin(\theta_2 - \theta_3)\right]$$
$$H_{robot} = l_2 \cdot \cos(\theta_2) + l_3 \cdot \cos(\theta_2 - \theta_3) \quad (28)$$

4.5 Locomotion Over Stairs

Stair represents one of the most common constructions (obstacle) in human life. There are two types of stairs: straight or spiral. Locomotion over stairs is more difficult in relation with obstacles presented above.

A stair can be defined as number of steps disposed horizontal and vertical. An important issue regarding stairs is that its height increases with each step.

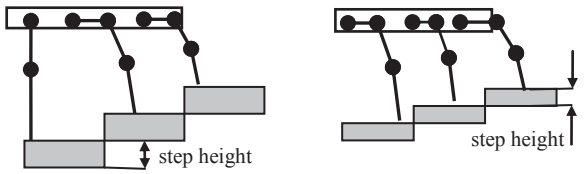
It has been imposed the condition of horizontality for the robot body during locomotion. For this condition we had to calculate the maximum height of a single step. We know from previous cases that if the robot height decreases so the height of the obstacle should decrease.

In this case we'll do two analyses, one for the robot height equal with the sum of femur and tibia links (Fig. 15a) and one where the height is smaller (Fig. 15b). For the first case when $H_{robot}=l_{femur}+l_{tibia}$ the height of the step is:

$$H_{step} = \frac{H_{robot} - \left[l_2 \cdot \sin(\theta_2) + l_3 \cdot \sin(\theta_2 - \theta_3)\right]}{3} \quad (29)$$

where:

θ_{femur} , θ_{tibia} – the angles for frontal legs;
H_{robot} – the height calculated according to the length of rear legs.



a) $H_{robot}=l_{femur}+l_{tibia}$ **b)** $H_{robot}<l_{femur}+l_{tibia}$

Fig. 15. Calculus of step height.

For case b from Fig.15 the height of the step can be calculate using:

$$H_{step} = \frac{H_{robot} - \left[l_2 \cdot \sin(\theta_2) + l_3 \cdot \sin(\theta_2 - \theta_3)\right]}{3},$$
$$H_{robot} = l_2 \cdot \cos(\theta_2) + l_3 \cdot \cos(\theta_3) \quad (30)$$

where:

θ_{femur}, θ_{tibia} – the angles for frontal legs;
H_{robot} – the height calculated according to the length of rear legs.

5. SIMULATION PLATFORM

The main purpose of this simulation platform (Fig.13) is to control all the aspects of the hexapod robot. The simulation program was made using MATLAB G.U.I.D.E. (Brian et al. 2001, Knight 2000, Marchand et al. 2001).

The developed software platform can be used for static stability analysis in gravitational field and locomotion analysis. Stability analysis in gravitational field allows simulating how the robot behaves in case of malfunction. The locomotion analysis allows walking simulations of the robot over common types of obstacles: ditch, crest, step, inclined plane, straight stairs and spiral stairs.

The graphical representation of the robot also allows seeing the shape of the support polygon which is updated according to the legs on the ground. The leg is considered on the ground if the distance from the tip to the ground is smaller than a threshold. This threshold is modifiable and was introduced as a way to compensate certain position errors that may occur due to real servomotors.

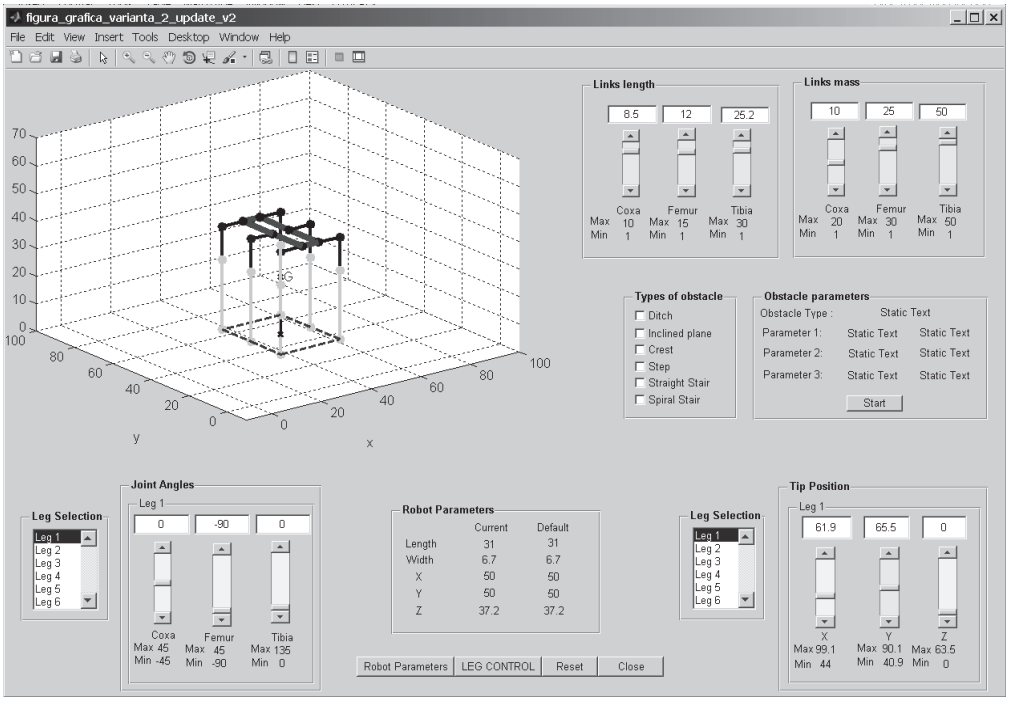


Fig. 16. Matlab graphical user interface (GUI) developed for simulation.

The simulation program shows all the stages the robot goes through for a better understanding. We have elaborated an algorithm in order to achieve the goal of analyzing the static stability of a hexapod robot in gravitational field.

The algorithm is structured in 5 steps as following:

- Setting the joints values.
- Determine the mechanical configuration.
- Determine which legs are on the ground.
- Evaluation of the static stability condition.
- While (condition of static stability == false)
  - Determine the rotation line using the minimum distance from *G'* (Fig. 4) to support polygon's sides,
  - Rotate the robot about the line found,
  - Determine which legs are on the ground,
  - Evaluation of the static stability condition.

The determination of static stability condition is resumed at finding if the projection of robot's centre of gravity is inside the support polygon. For this we used the following algorithm:

- Determine the convex polygon.
- Determine the area of the convex polygon (*A*).
- Form the *n* triangle using 2 consecutive points of the support polygon and the projection of centre of gravity, *G'*.
- Determine the areas of the n triangles formed (Ai).
  - If (ΣAi == A) then condition=true
  - else condition=false

6. EXPERIMENTAL RESULTS

The parameters, in cm, of the hexapod robot model considered for stability analysis are: length 20, width 10, coxa 5, femur 5 and tibia 5. The parameters, in cm, of the hexapod robot model considered for locomotion analysis are: length 45, width 6.7, coxa 8.5, femur 12.5 and tibia 25.

6.1 Stability analisys during malfunction

On this type of analysis some joints of the robot are considerate malfunctioning from whatever causes. The only force that acts upon the robot is the gravitational force. For a given set of joint values the robot passes through many transitory stages until it becomes statically stable (Fig. 17a, b, and c). Legs that have contact with the ground determine the shape of the support polygon (triangle, quadrilateral, pentagon or hexagon). In order to know if the robot achieves static stability the projection of *G* (*G'*) must be inside the support polygon. To solve this problem the above algorithm is applied. For this analysis the joints of the leg are considerate malfunctioning.

As it can be seen in Fig. 17a, the 6th leg loses contact with the ground and the robot is unstable. The projection of the centre of gravity, *G'*, is not inside the support polygon and the robot rotates about the line determined by the leg 2 and leg 3. The rotation line is determined by calculating the minimum distance from *G'* to polygon's sides. In this case the first leg closest to ground that will provide support is leg 5.

In Fig. 17b it can be seen that even in this configuration *G'* is not inside the support polygon and the robot continues to rotate about the line determined by leg 2 and leg 5 until the first leg touches the ground, which in this case, is leg 4.

In Fig. 17c a new configuration is formed and if the algorithm described above it is applied, point *G'* is inside the support polygon and the robot becomes statically stable.

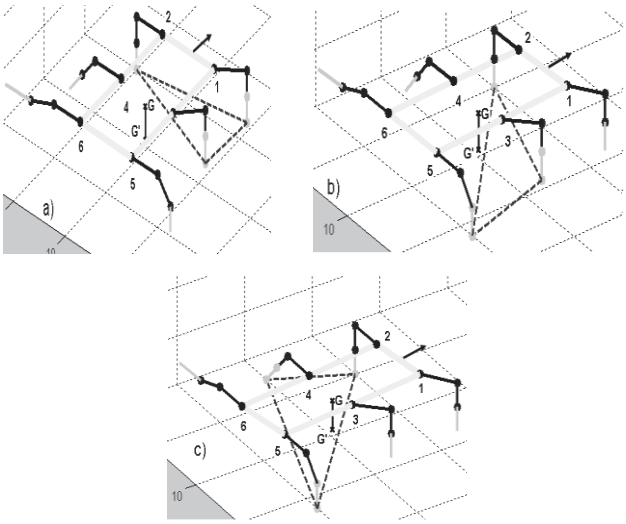


Fig. 17. Stability analysis. a) Phase one – unstable; b) Phase two – unstable; c) Phase three – statically stable.

6.2 Locomotion on Inclined Plane

The most important configurations of the robot during locomotion on inclined plane are presented in Fig. 18.

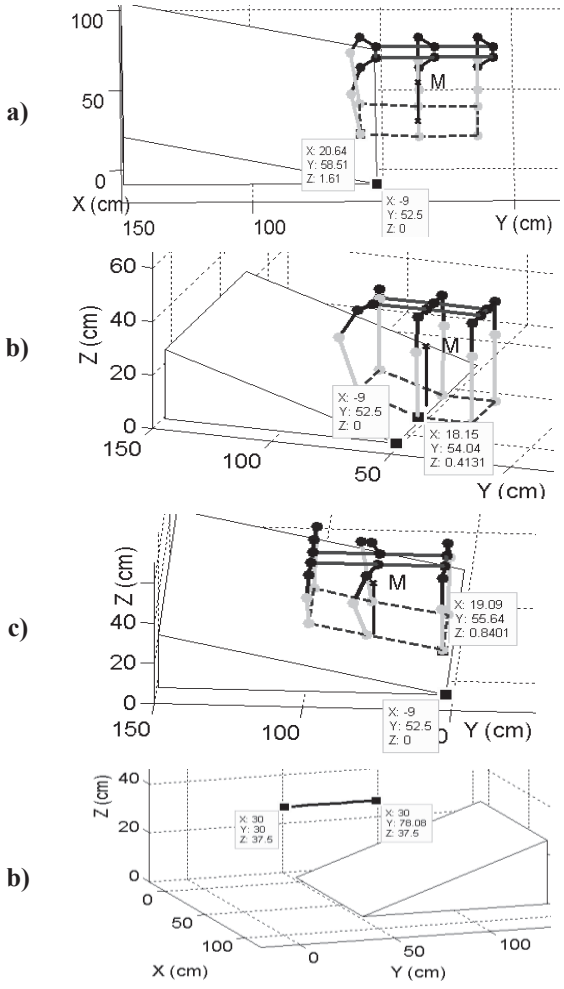


Fig. 18. Locomotion on inclined plane.

Locomotion on inclined plane is achieved by moving all legs from one side one by one, body drag and then moving all the legs from the other side. Restrictions imposed for locomotion were: keeping the robot structure horizontal with respect to the ground during locomotion, maximum step length, keeping the maximum height of the robot body with respect to the ground, angle of the plane is 15^{o}.

In Fig. 18.a-c the robot positions the front legs, the middle legs and the rear legs on the inclined plane. In Fig. 18.d is presented the trajectory of the centre of symmetry of the robot body during locomotion.

6.3 Locomotion Over a Step

The most important configurations of the robot during locomotion over a step are presented in Fig. 19.

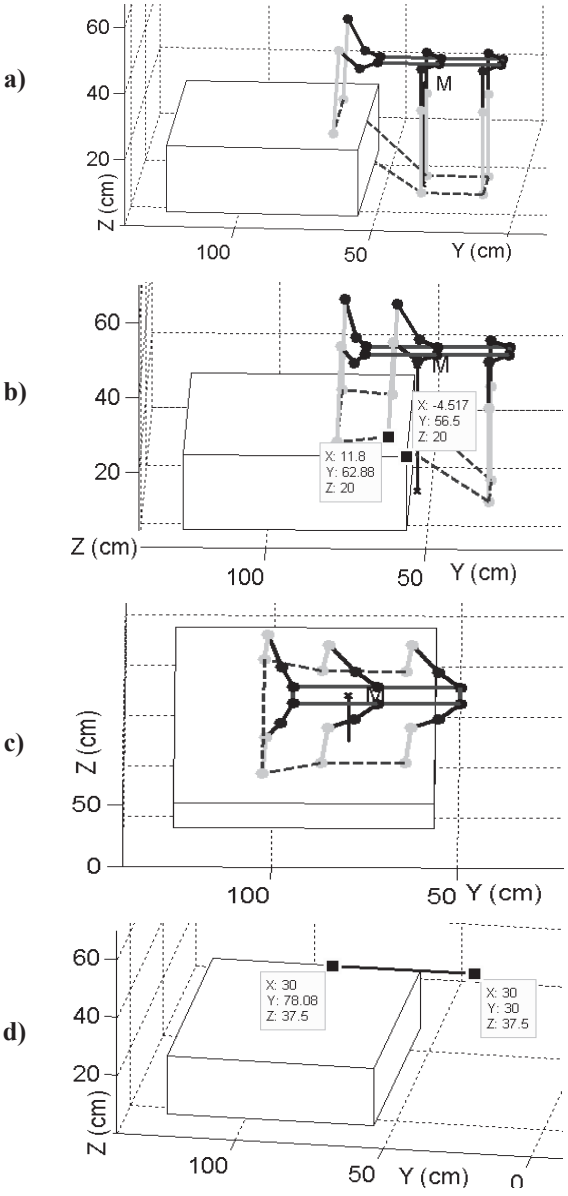


Fig. 19. Locomotion over a step.

As in the previous case the legs follow the same algorithm presented above. Restrictions imposed to the robot are: keeping the robot structure horizontal with respect to the ground during locomotion, maximum step length, keeping the maximum height of the robot body with respect to the ground, the height of the step is 20 cm.

In Fig. 19.a-c the robot positions the front legs, the middle legs and the rear legs on the step. The robot position from Fig. 19.c is the final position of the robot. In Fig. 19.d is presented the trajectory of the centre of symmetry of the robot body during locomotion.

6.4 Locomotion Over a Ditch

After running the simulation several times we discovered a better starting leg configuration than the one presented in Fig. 12a. The new configuration allows locomotion over a bigger ditch then the one presented in (24). The restrictions imposed during locomotion are the same as in previous cases. The width of the ditch is 9 cm. The results of the simulations are presented in Fig. 20.

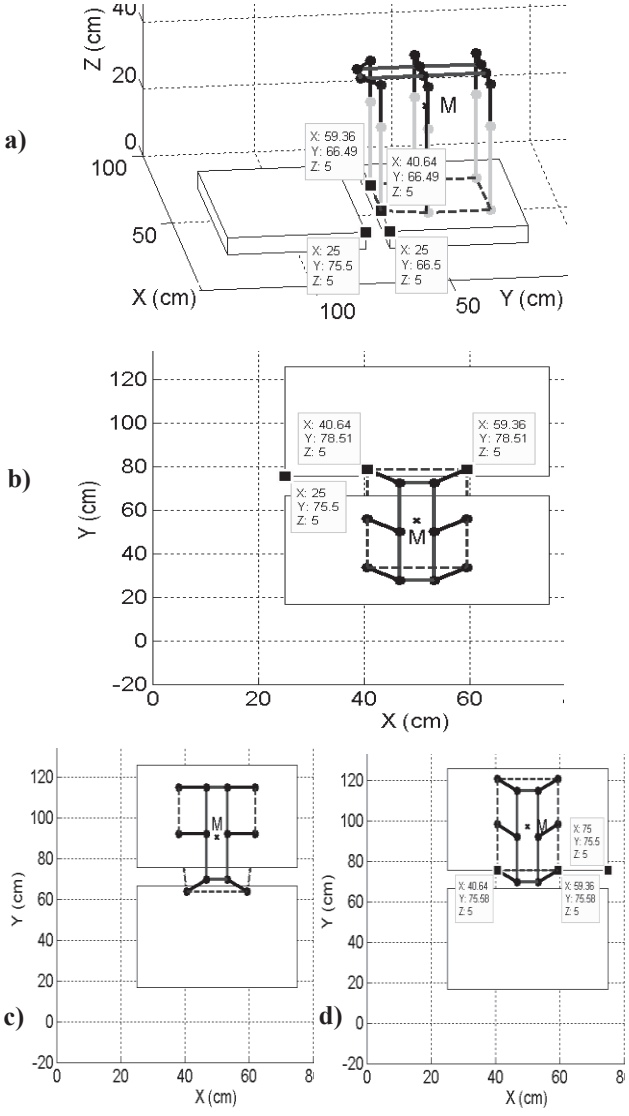


Fig. 20. Locomotion over a ditch. a) Initial position of the robot; b) Intermediary position of the robot during locomotion over a ditch; c) Intermediary position of the robot for transferring the rear legs; d) Rear legs on the final plane.

In Fig. 20.c are created the conditions for the rear leg to be transferred over the ditch. The conditions are similar with those in Fig. 20.a. In Fig. 20.d the rear legs are transferred over the ditch.

6.5 Locomotion Over a Crest

As in the previous cases the legs follow the same algorithm presented above. The restrictions imposed during locomotion are the same as in previous cases. The height of the crest is 15 cm and the width of the crest is 10 cm. The results of the simulations are presented in Fig. 21.

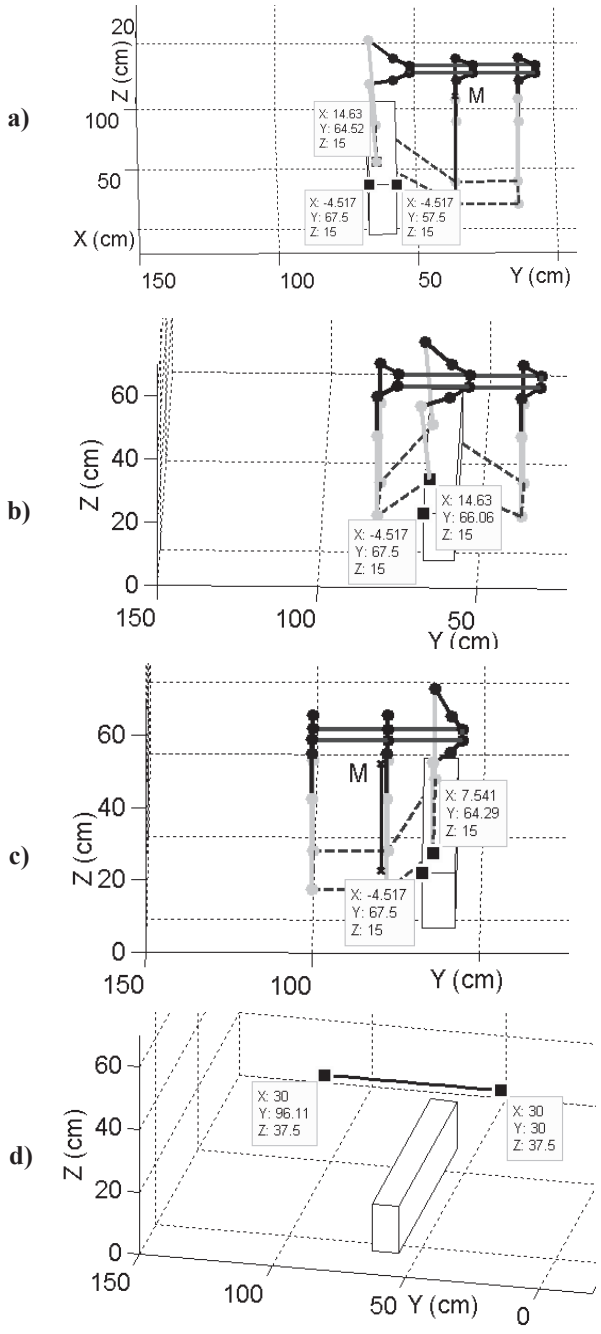


Fig. 21. Locomotion over a crest. a) Front legs on crest; b) Middle legs on crest; c) Rear legs on crest; d) Trajectory of the centre of symmetry of the robot body during locomotion.

In Fig. 21.a-c the robot positions the front legs, the middle legs and the rear legs on the step. In Fig. 21.d is presented the trajectory of the centre of symmetry of the robot body during locomotion.

6.6 Locomotion Over Straight Stairs

The most important configurations of the robot during locomotion on inclined plane are presented in Fig. 22.

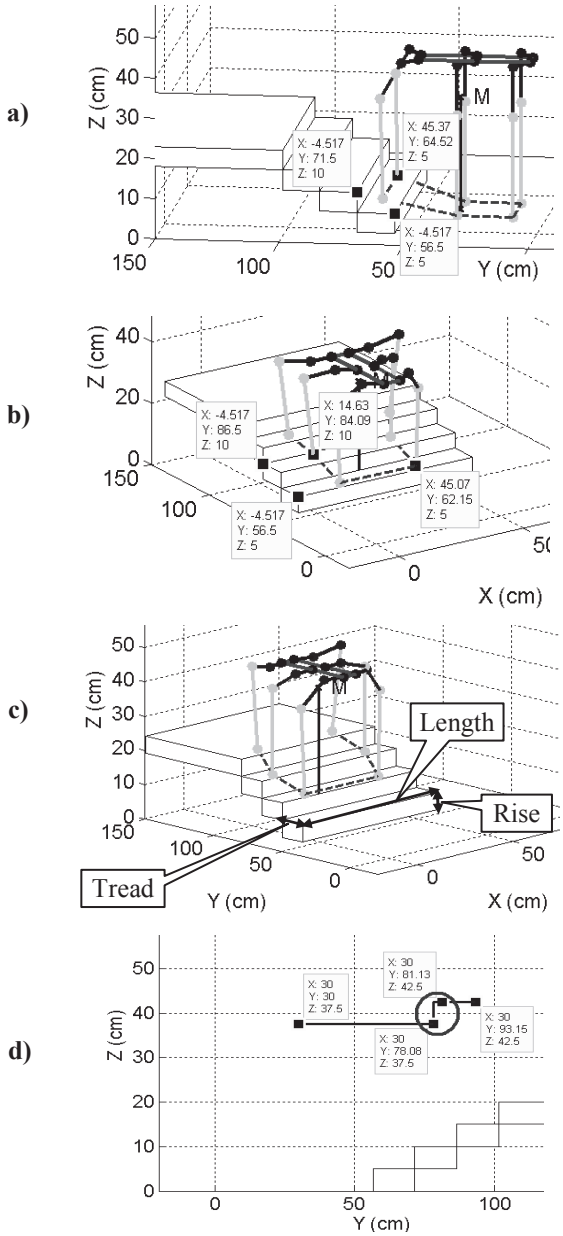


Fig. 22. Locomotion over straight stair. a) Positioning of the front legs on the first step; b) Intermediary position of the robot on stair; c) Position of the robot after rising the body; d) Trajectory of the centre of symmetry of the robot body during locomotion.

Locomotion over straight stairs is similar with climbing a step but there are some differences. The restrictions imposed during locomotion are the same as in previous cases. The parameters considered for stairs were: rise 5 and is ~1/8 of the height of the robot, tread 15 and is ~1/3 of the length of the robot. The stair is form of 4 steps. The length of the stair must be larger than twice the sum of width of the robot and twice the length of coxa.

The position of the robot from Fig.22.b represents the starting point for the robot to change its height because during locomotion the height of the robot decreases and the robot couldn't move.

The algorithm for changing the height of the robot is:

- position every pair of legs on individual steps,
- position every leg tip below each femur joint,
- raise the robot using the tibia and femur joints.

The position of the robot presented in Fig.22.c is the results of the raising process mentioned above. In Fig.22.d is presented the trajectory of the centre of gravity (G) of the robot during locomotion and is highlighted the moment when the raising process starts.

6.7 Locomotion Over Spiral Stairs

The parameters considered for stair steps were: height 5 cm, width 37 cm. For spiral stair the width of the step is larger because of the way is built. The spiral stair is form of 3 steps. The spiral stairs were simulated using the implemented Matlab helix equations, but with some adaptations:

$$\begin{aligned} &z = 0 : (step\_heigth) : (steps\_number \cdot step\_heigth \cdot \pi); \\ &angle = z / 2 \cdot \arcsin(step\_width / step\_length); \\ &x = x_0 + (step\_length) \cdot \cos(angle); \\ &y = y_0 + (step\_length) \cdot \sin(angle); \end{aligned} \tag{31}$$

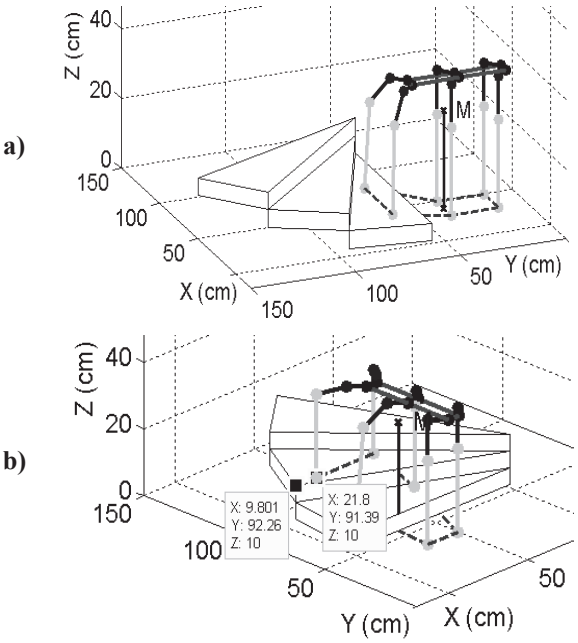


Fig. 23. Locomotion over a spiral stair. a) Position of the frontal legs on the first step. b) Intermediary position of the robot on the spiral stair.

For the robot to be able to overcome this type of obstacle it needs to rotate during locomotion. The algorithm for rotating the robot is:

- calculate the trajectory for the legs that don't provide suport,
- raise the legs 2, 3 and 6,
- rotate the robot using the coxa joint of the legs 1, 4 and 5,
- rotate and lower the legs 2, 3, and 6.

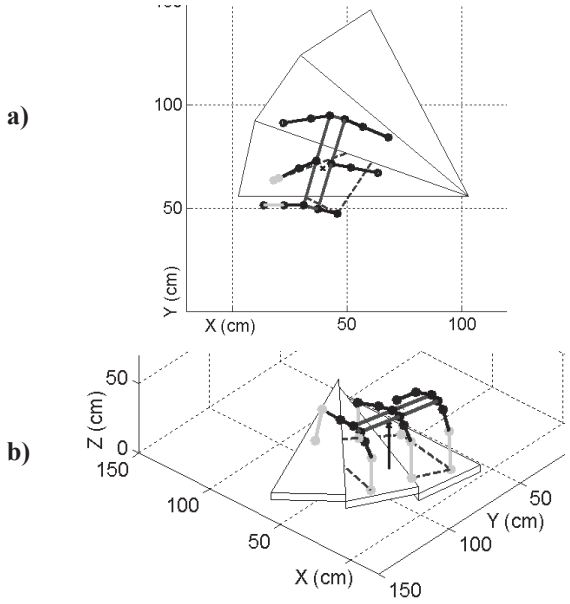


Fig. 24. Locomotion over a spiral stair. a) The position of the robot after the rotation process is over; b) Intermediary position of the robot during locomotion.

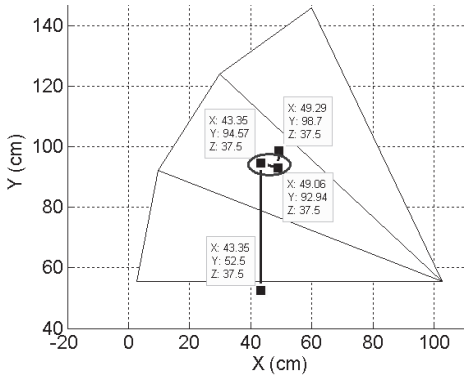


Fig. 25. Trajectory of the centre of symmetry of the robot body during locomotion.

In Fig. 25 is presented the trajectory of the centre of symmetry of the robot and its highlighted the moment when the robot starts to rotate.

Locomotion on spiral stairs is one of the most challenging due to the fact that the robot must rotate about its axis. Examining the simulations, there are some tweaks needed for the robot to perform better.

7. CONCLUSIONS

In this paper a simulation platform for legged mobile robots was presented, that allows stability analysis and locomotion analysis for a hexapod robot. Stability analysis for the robot in case of malfunction is useful for investigating what happens with the robot due to loses of contact, servomotor failure, power supply failure.

Also the paper presented the main classes of obstacles met in everyday life and the influence upon the hexapod robot structure along with a series of locomotion simulations for a hexapod robot platform over these types of obstacles.

Regarding the robot leg, the direct kinematical model and the inverse kinematical model for the mechanical structure for the leg were presented. In the paper are calculated the positions of the centre of gravity for each leg and for the whole robot. These points are important for stability analysis in different walking algorithms. For the same leg structure a workspace analysis has been made in order to highlight the restrictions when developing the walking algorithms of the robot. The movement of the leg tip along the trajectory was made using piecewise cubic spline interpolation method.

The Matlab graphical user interface was designed to be simple and intuitive and to offer the user a simple and efficient way to manage every aspects of the robot (angles, masses, lengths).

The software platform can also be used for legged mobile robots with 4 or 8 legs using minor code modifications and in the future will be upgraded permitting additional controls and functions for stability analysis (including dynamic stability) on uneven ground and implementing collision detection algorithms.

Also the results of these studies represent the bases for different strategies of locomotion on different terrains. A real hexapod has been constructed and we are trying to connect it to the platform in order to be able to control it.

ACKNOWLEDGMENT

This work was partially supported by strategic grant POSDRU/88/1.5/S/50783, Project ID 50783 (2009), co-financed by the European Social Fund – Investing in People, within the Sectoral Operational Programme Human Resource Development 2007 – 2013.

REFERENCES

Bensalem, S., Gallien, M., Ingrand, F., Kahloul, I., Thanh -Hung N. (2009). Designing autonomous robots. *IEEE Robotics & Automation Magazine*, Vol. 16.

Brian, R., Hunt, R., Lipsman, L., Rosenberg, J.M. (2001). *A guide to MATLAB for beginners and experienced users*. ISBN:978-0-521-00859-4, Cambridge University Press.

Conrad, J.M., Mills, J.W. (2009). The History and Future of Stiquito: A hexapod insectoid robot. *Artificial life models in hardware*, DOI 10.1007/978-1-84882-530-7_1, ISBN 978-1-84882-529-1.

Jakimovsk, B. (2011). Biologically inspired approaches for locomotion of a hexapod robot OSCAR. *Biologically inspired approaches for locomotion, anomaly detection and reconfiguration for walking robots*, pp 35-66, DOI: 10.1007/978-3-642-22505-5_5, ISBN 978-3-642-22504-8.

Jung-Min Yang (2009). Fault-tolerant gait planning for a hexapod robot walking over rough terrain. *Journal of intelligent and robotic systems*, Vol 54, No 4, pp. 613-627, DOI 10.1007/s10846-008-9282-x, ISSN: 0921-0296.

Knight, A., (2000). *Basic of Matlab and Beyond*, ISBN: 8493-2039-9, CRC Press LLC.

Krzysztof, W., Dominik, B., Andrzej K. (2008). Control and environment sensing system for a six-legged robot, *Journal of automation, mobile robotics & intelligent systems*, Vol. 2 No. 3.

Lewinger, W. A., Quinn, R.D. (2011). Neurobiological-based control system for an adaptively walking hexapod, *Industrial robot, an international journal*, Vol. 38, No. 3.

Marchand, P., Thomas, H.O.(2001). *Graphics and GUIs with MATLAB third edition*, ISBN 1-58488-320-0, Chapman & Hall/CRC.

Schilling, R.J. (1990). Fundamentals of robotics, analysis and control, ISBN: 0-13-344433-3, Prentice Hall, New Jersey, USA.

Silva, M.F., Tenreiro Machado, J.A. (2011). Fractional control of legged robots, *Springer Proceedings in mathematics,* Vol. 2, pp. 647-650.

Tao, L., Ceccarelli, M.(2011), Additional actuations for obstacle overcoming by a leg mechanism, *Proceedings of the 18th IFAC World Congress*, Vol. 18, Part. 1, pp. 6898-6903, DOI:10.3182/20110828-6-IT-1002. 00351, Italy.

Wang ZhouYi, Wang JinTong, Ji AiHong, Li HongKai, Dai ZhenDong (2011). Movement behavior of a spider on a horizontal surface, *Chinese science bulletin*, Vol. 56, No. 25, DOI: 10.1007/s11434-011-4584-y, pp. 2748− 2757.

Implementing a Complex Fast Hybrid Sudoku Solver

Ileana-Diana V.D. Nicolae*, Anca-Iuliana P.M. Nicolae**

**Department of Computers and Information Technology*
University of Craiova, Decebal Blv. No. 107
(e-mail:nicolae_ileana@software.ucv.ro)
***Department of Cybernetics and Economic Statistics*
Doctoral School of Bucharest, University of Economic Studies
(e-mail:ancaiuliana.nicolae@gmail.com)

Abstract: The efforts made by authors to implement an efficient Sudoku solver based on 3-d arrays, successfully tested for classic grids of small and medium difficulty were continued. An improved version, able to solve all types of grids, was conceived and tested. New functions were added : hidden/naked quads are now detected and exploited, the functions dealing with the techniques known in gamers world as „x-wing" and "y-wing" were improved such as to deal with more general indirect reduction techniques, known as "Extreme strategies" : "Perfect and finned (non) Sashami x-wing" and "Reduction based on Almost Locked Sets" respectively. Tests made on 35 grids with different degrees of difficulty (17 of them being classified as "hardest") revealed significant mean, maximum and minimum runtime savings (around 50%, in some cases reaching hundreds of seconds) as compared to a classic backtracking algorithm using a 2d representation of the grid and to free online solvers respectively. The Matlab code used by solver does not make use of any specialized libraries. It is portable to C++ or Java with minimum modifications. The hybrid solver can be easily adapted to solve grids with additional restrictions (e.g. restrictions corresponding to diagonals, symmetric grids etc.).

Keywords: Constraint satisfaction problems, Backtracking, Algorithms, Computer programming

1. INTRODUCTION

Sudoku puzzle with 9 rows (lines) and columns containing digits from 1 to 9, whose popularity is especially remarkable among gamers all over the world since 2005, originated in Japan. The rows/columns are divided into 9 3x3 zones (marked with thicker frames on the grid – Fig. 3). In its original (classic) form, restrictions are imposed such as no duplicates are allowed along a row, a column and respectively inside a zone.

In one of their studies on Sudoky puzzles (Ercsey-Ravasz and Toroczkai , 2012) the authors began by reminding that the mathematical structure of these puzzles is akin to hard constraint satisfaction problems lying at the basis of many applications (including ground-state problem of glassy spin systems, protein folding etc.) and proved that the difficulty of Sudoku translates into a transient chaotic behaviour exhibited by continuous-time dynamical systems.

Sudoku puzzles have also been used as test cases for many varied computer problem-solving techniques including constraint programming, reduction to satisfiability, harmony search , metaheuristics, genetic algorithms , particle swarms , belief propagation , projection onto the Birkhoff polytope, Gröbner bases, mixed integer programming , and compressed sensing (Eppstein , 2012).

Also, deductive puzzle solving algorithms may cast light on human psychology, by helping us understand what kinds of thought processes are easier or more difficult for humans (Eppstein, 2012).

Every properly-designed Sudoku puzzle has a unique solution, and in most cases this solution can be found by humans via a sequence of deductions that does not involve trial and error or backtracking. Sudoku puzzlists typically learn a repertoire of deductive rules that allow them to make steps towards solving a puzzle whenever they can find certain patterns in it (Eppstein, 2012). Rules such as this, and exercises to help learn them, may be found in guidebooks for Sudoku puzzle solving (e.g. Gordon and Longo, (2006) , Stephens, (2007) and Stuart, (2010)).

It is not difficult for computers to solve 9x9 Sudoku puzzles using backtracking together with simple deductive rules to determine the consequences of each choice and prune the search tree whenever an inconsistency is discovered (Brouwer, A. E., (2006), Eppstein , (2005)). Despite the ease of solving Sudoku using these methods, there is still a need for computer puzzle solvers that eschew both trial and error and sophisticated computer search techniques, instead using mimic human deduction (Eppstein , 2012). A grid's hardness depends on many factors, including the positioning pattern of the clues, as shown in Rosenhouse and Taalman, 2011.

2. ALGORITHMS AND SOLVERS – STATE OF THE ART

Backtracking techniques represent a common approach in solving Sudoku grids, often used as study case for

computing techniques in universities (Zelenski, 2008) when intending to solve classic grids, where no additional restrictions (e.g. symmetries) are imposed. In their standard form these techniques guarantee the detection of all possible solutions, but tests performed on large sets of grids revealed significant maximum and mean runtimes.

The huge interest in solving Sudoku puzzles all over the world with pure or hybrid backtracking techniques is revealed by a great number of programs and routines (many available online), developed in various programming environments: in Java (Daniel Liang, 2007), in PHP (Eferanto, 2008), in JavaScript (Detar, 2010), in Python (Python Fiddle, Eppstein, 2005) in ECLiPSe (Simonis, 2005) etc. There are also backtracking-free software tools (for example Stuart, (2013a), which helped enormous in our work), conceived mainly for gamers, which rely exclusively on some direct and indirect restriction techniques and therefore cannot find easily solutions for certain grids.

Consulting the online available content of the above mentioned solvers we could make a synthetic picture, presented below.

Emanuel Eferonato (Eferanto, 2008) implements a variant of backtracking algorithm, with no indirect restriction techniques. Frequent calls to dedicated PHP functions for handling arrays are involved, so specialized libraries must be available. When the algorithm tries to find a new solution for a certain place, randomization is used considering all the possible solutions, and therefore results' reproducibility with respect to runtimes can not be provided. Memory consumptions are significant as the current (unsolved) configuration of the grid is entirely saved in a stack structure at every step. All these remove this solver from the list of candidates for time critical applications.

Detar's online solver (Detar, 2010) relies on a simple recursive function and a stack data structure respectively. Therefore timeouts may appear in some browsers for some puzzles. Using it on the test grids presented in Section 6 we obtained systematically runtimes at least double than that provided by our hybrid algorithm. Its latency definitely does not recommend it for time critical applications.

Not implementing indirect restriction techniques, the Python solver from (Python Fiddle) is also not competitive relative to fastness for hard grids.

The work of David Eppstein (Eppstein , 2012) is impressive. Striving to avoid backtracking, instead mimicking human rule-based reasoning, he conceived a hybrid solver, written in Python, which implements many (in)direct restriction techniques. Consulting it at (Eppstein, 2005) we could identify most of the useful techniques presented in (Stuart, 2013a). Still we could not detect any function responsible for dealing with finned (non) Sashami x-wings or with Almost Locked Sets, which proved to be very powerful techniques, generating many indirect restrictions and runtime savings in the examples analyzed by us. Moreover, the author did not provide any information on runtimes, instead mentioning the difficulty of getting meaningful timings in a slow interpreted language such as Python.

The most spectacular results found by us were those authored by Simonis (2005). Considering a constraint formulation of the Sudoku puzzle, he demonstrates that one can use different modeling techniques to find puzzle solutions without search. His programs are written in ECLiPSe 5.8 and use the IC library and its extension ic global. A propagator for a hyper arc-consistent all different was added for the mentioned study. With an impressive mathematical and software support, using a 1.5 GHz/1Gb laptop, he obtained mean runtimes under 1 second and maximum runtimes of tens of seconds, this being definitely much smaller then any results obtained with the tools analyzed by us. Unfortunately the memory consumptions involved by his techniques are significantly high.

In this context we conceived and tested a hybrid algorithm to implement a Sudoku solver in Matlab 2008, using a 2.53 GHz/1.38 GB laptop. In (Nicolae and Nicolae, 2012), we presented a previous, more simplified version of it. The use of auxiliary (9x9) matrices containing flags related to restrictions associated with rows, columns and zones, along with execution branches that implement strategies from gamers' world (known as „x-wing", „y-wing", „hidden/naked pairs/triples", „box-to-line/column restrictions" etc.) and with a minimal (ideally zero) segment of backtracking code, exhibited very good runtimes (not exceeding 0.4 s for classic grids of medium difficulty and respectively 0.15 s for easy puzzles). We continued our effort and released an improved version, as described below.

3. 3D BACKTRACKING ALGORITHM

When implementing our vanilla 3d Backtracking algorithm, a Java implementation of the classic 2d backtracking algorithm (Daniel Liang, 2007) was used as draft. Modifications were made such as to take advantage on the flags stored by the 3d array $c(9,9,10)$ used as main data structure for the grid content and restrictions.

In the 3d version, unlike at the 2d algorithm, only the values whose corresponding positions in the third index of $c(i,j,1:9)$ are 1 are tested as possible solution for the place identified by (i,j). For example if only $c(i,j,3)$ and $c(i,j,5)$ are 1 and the rest of $c(i,j,:)$ are 0, then only 2 possibilities are tested instead of 9 from the classic 2d backtracking algorithm. Only the values placed within the range [*min_value_in_place*(i,j), *max_value_in_place*(i,j)] and flagged as possible (as revealed by 1-s in $c(i,j,1:9)$) are considered when searching the solution for the currently tested place. When the solution for a place (i,j) is known as being x, then $c(i,j,1:9)=0$ and $c(i,j,10)=x$.

To account for the last value tested during the analysis of the previous tested place (identified by the indices of row and column stored in *freePlaceList*$(k\text{-}1,1)$ and *freePlaceList*$(k\text{-}1,2)$), a vector dinamically updated at each step (*last_tested* (*freePlaceList*$(k\text{-}1,1)$, *freePlaceList*$(k\text{-}1,2)$)) is used.

The price paid for the runtime drastic reduction is the necessity to store 3 (9x9) matrices as local variables for the use of the „search" function employed by the 3d backtracking algorithm. They provide the minimum and maximum possible values and respectively the last tested value for each place.

4. IMPLEMENTATION OF INDIRECT RESTRICTION TECHNIQUES BASED ON ALMOST LOCKED SETS

4.1. Fundaments

The principle of indirect restriction techniques relying on the so called „Almost Locked Sets" – ALS is briefly explained in (Stuart, 2013a). An ALS consisting of n places can host combinations of only n+1 distinct digits, the combinations' lengths being variable within the range [2 , n+1].

Fig. 1 depicts a configuration (with zoom over the last 5 columns) in which 2 ALS sets are marked with special symbols. The first one, established within the 8-th zone and marked with ‚xx', consists of 4 places, containing combinations of only 5 digits (namely 4,5,6,7,8).

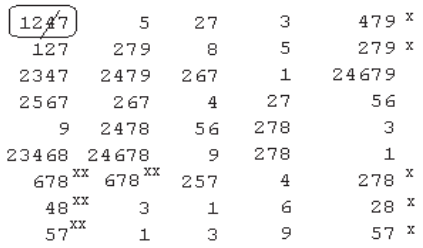


Fig. 1. First example of a grid with ALS sets.

Similarly another set (denoted by ‚x') is formed along the 9- th column, relying on 5 places which can host combinations of only 6 distinct digits: 2,4,5,7,8 and 9.

Four major conditions must be accomplished by a pair P of ALS sets (S1 and S2) in order to use them for indirect restrictions over certain digits – digits known as „Candidates for Restrictions" - *CR*.

The first condition involves the perfect disjunction among ALS sets (no intersection between their places is allowed).

The 2nd condition requires the presence of some (sets of) place(s) in both ALS sets, „fully viewing" each other with respect to one or more digits which are present in both sets, digits known as „Restricted Commons"- *RC*. A digit found as *RC* with respect to P cannot be restricted during the processing of P.

In our example, the digit 8 appears in both ALS sets. All places hosting 8 in S1 can „see" all places hosting 8 in S2 and vice versa. So, 8 can be considered as *RC*. Similarly 5 will be considered too as *RC*. On the other hand, even if the places from both ALS sets containing 7 in the rows 7...9 see each other, 7 cannot be considered a *RC* because the places (1,9) and (2,9) from the second ALS set cannot see some places hosting 7 in the bottom rows of the second ALS set.

The third condition imposed to the currently processed P is a non-null intersection between the sets of digits stored by S1 and respectively S2. The digits common to S1 and S2, other than *RC*s, represent the *CR*s for P.

The 4-th condition requires the presence of "observers" - O. A place from grid is O for P relative to a certain *CR* if:
- it represents an unsolved place and can host *CR* (for our example only 4 can be *CR*);
- it does not belong neither to S1 nor to S2;
- it can „see" all places from S1 and S2 hosting *CR*.

The currently discussed P obeys the conditions required to restrict the digit 4 from the place (1,5) (which shares the same line with the place (1,9) and the same column with the place (8,5)) and therefore 4 will be restricted from this place.

A more complex example is provided by Fig. 2.

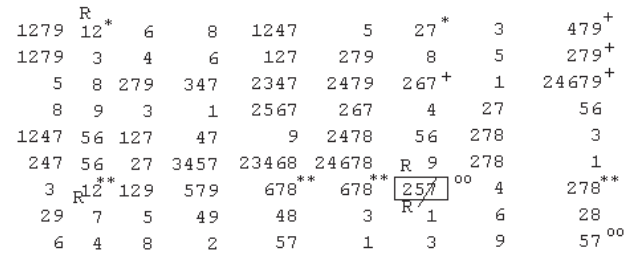


Fig. 2. Second example of a grid with ALS sets.

Let us analyze the 4 ALS sets identified through symbols as follows: S1 -> '*', S2-> '**', S3 -> '+' and S4 -> 'oo'. The sets S1 and S2 see each other by means of the places (1,2) and (7,2) containing 1 and 2. As only the digit 1 appears in the mutually visible places, it represents the only possible *RC* for the pair (S1, S2). The digit 2 cannot play this role, as the second place from S1 is not visible from the first place containing (1,2) in S2. The rest of digits present in both ALS are 2 and 7 and they represent *CR*s.

There is no O for P on the grid corresponding to the digit 2. On the other hand, the place (7,7) is the only O for P relative to '7', because it can see the only place able to host 7 from S1 and both places able to host 7 from S2. Therefore 7 will be restricted from this place.

An analysis of all possible pairs which may be formed with the sets emphasized in Fig. 2 is presented below:
- *(S1,S3)* . Visibility between (1,7) and all places from S3. *RC*=7; *CR*=2. There is no O.
- *(S1,S4)*. Visibility between (1,7) and (7,7). There is no *RC*.
- *(S2,S3)* . Visibility between (1,7) and places from lines 1...3, column 9. There is no *RC*.
- *(S2,S4)*. Visibility along column 9 and line 7. *RC*=2. *CR*=7. There is no O.
- *(S3,S4)*: Visibility between (3,7) and (7,7), respectively between (9,9) and places from lines 1...3, column 9. There is no *RC*.

An interesting property is related to a *CR* belonging to other ALS sets, not involved in the currently analyzed pair. As seen from Fig. 2, the restriction imposed by the

pair (*S*1, *S*2) corresponding to *CR*=7 over the place belonging to *S*4 does not influence the definition of *S*4 – as it still has both its places and still contains combinations of all 3 digits used initially for its building.

4.2. Implementation

For the beginning, indices of places able to „see" a certain place (*i,j*) are calculated for each major restricted unit (row, column and zone) and stored in a 3d array – *V*. The indices are stored in a compact form – as a number *xy*, where *x* represents the row and *y* represents the column.

Fig. 3 represents with 'o' the places able to see the place (4,7). As the order in which these indices is „line first, zone last", for the mentioned place, the following components are deduced: *V*(4,7)=[41,42,..46,48,49,17,27, 37,57,67,77,87,97,48,49,57,58, 59,67,68,69].

						o		
						o		
						o		
o	o	o	o	o	o	**x**	o	o
						o	o	o
						o	o	o
						o		
						o		
						o		

Fig. 3. Visual representation of places "able to view" the place (4,7).

The second stage assumes the building of all possible ALS sets over all major restricted units considering the combination of digits whose lengths vary from 2 to 6.

The modality we used to construct all the ALS sets of length 3 for the column *j* is presented in Fig. 4, where *l* is used to denote the total number of unsolved places in the column, *ic1*, *ic2* and *ic3* are used to generate all possible combination of 3 distinct unsolved places within the column *j* which might form an ALS set, whilst the vectors *local_i* and *local_j* are used to identify the places' indices within the grid.

```
for ic1=1:l-2
  for ic2=ic1+1:l-1
    for ic3=ic2+1:l
     s=0;
      for k=1:9
       c1=c(local_i(ic1), local_j(ic1),k);
       c2=c(local_i(ic2), local_j(ic2),k);
        c3=c(local_i(ic3), local_j(ic3),k); s=s+c1 || c2 || c3;
      end;
      if s==4
           i_ v_ALS_indices = i_ v_ALS_indices +1;
             obtain v by merging (ic3,j), (ic2,j) and (ic1,j) ;
                v_ALS_indices( i_ v_ALS_indices)=v;
                end; end; end; end;
```

Fig. 4. Generating ALS sets of lenght 4 over the column *j*.

When *n* currently analyzed places are found as forming an ALS set of length *n*, their indices are merged and saved in a special array - *v_ALS_indices*.

Let us consider the 4 places marked with ‚xx' in Fig. 1 and their pair of (*i,j*) coordinates: (7,5), (7,6), (8,5) and (9,5) respectively. Assuming that the currently discussed ALS set was discovered immediately after four other ALS sets, then *v_ALS_indices*(5)= 75768595.

To detect the presence of at least one „visibility" between places coming from both currently analyzed ALS sets (*S*1 and *S*2), the sequence of code from Fig. 5 is used.

```
viz_ALS=false;
for ic1=1:length_set1
    for ic2=1:length_set2
      c1=i_ALS_set1(ic1)= =i_ALS_set2(ic2);
      c2=j_ALS_set1(ic1)= =j_ALS_set2(ic2); c3=iz1= = iz2;
                    if c1 || c2 || c3
                       viz_ALS=true;
                        break;
               end; end; end;end;
```

Fig. 5. Detecting visibility between currently analyzed ALS sets.

The indices *icm* (*m*=1 or 2) were used to cycle within the places belonging to ALS sets whilst *izm* was used to denote the index of zone including the place identified by *i_ALS_setm*(*icm*) and *j_ALS_setm*(*icm*) respectively, these last vectors storing the indices *i* and *j* of places belonging to *S*1 and *S*2. As the mappings (line, column)->(number of zone) are made only once in the main program, they are transmitted as global variables to all functions using them.

The digits hosted by both sets are obtained after an AND operation between two vectors of flags - *mask1* and *mask2*, which contain 1-s in the position corresponding to a hosted digit. For example, for the ALS sets from Fig. 1, *mask1*=[000111110] and *mask2*= [010110111], so the candidates for *RC* and *CR* can be found within the set {4,5,7 and 8} and are stored in the vector „*mask*".

The *RC*-s are determined using two boolean values – *r1* and *r2*. For a certain digit *k*, *r1* is true if all places from *S1* containing *k* are visible from all places containing *k* from *S2*. *r2* is computed similarly, but considering the visibility of places of *S2* from the places belonging to *S1*. If both *r1* and *r2* are true for *k*, then *k* is removed from „mask" and a flag is set to signalize the presence of at least one *RC*.

After this step, the mask containing the *CR* for the example from Fig. 1 will be [000100100], so only the digits 4 and 7 are qualified as *CR*.

To determine *O*s for a *CR* equal to *k* and the ALS sets *S1* and *S2*, the next major steps are performed:
- gathering in a single vector, all indices *i* corresponding to the places hosting *k* and belonging to *S1* or *S2* (we will denote them by "*places_k*") and in another vector all corresponding *j* indices of *places_k* . The vectors *v_i_k* and *v_j_k* will be obtained;
- to identify "the cloud" of unsolved places outside *S1* and *S2* flagged as able to host *k* and view all the places from

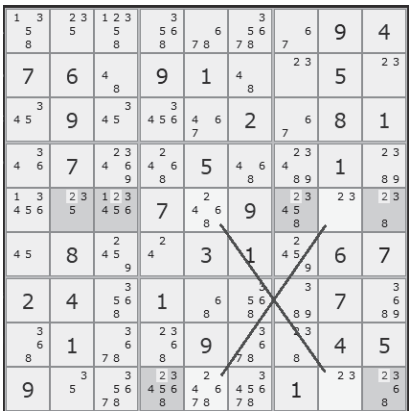


Fig. 6. Example for x-wing forming and restrictions.

S1 and *S2* containing *k*, we had to perform a reunion of all *V(ik, jk)*, (where *ik*, *jk* are provided by *v_i_k* and *v_j_k* candidates for *O* over *k* (a matrix *R* will be obtained);
- parsing each component from *R* and if the place identified by *R*(*i,j*) "sees" the places identified by all corresponding pairs of indices (*v_i_k*, *v_j_k*), then *O* is observer and *k* will be restricted from *R*(*i,j*).

Because finding the restrictions over some digits may result in the detection of other ALS sets, the code dealing with this type of restrictions is reiterated until no more restrictions are detected.

5. INDIRECT RESTRICTION TECHNIQUES RELYING ON GENERALIZED X-WING STRUCTURES

5.1. Main configurations

A. Perfect X-Wings

A special combination of coordinates of places and their content is known as "x-wing", an example of a perfect x-wing being given in Fig. 6 (Stuart, 2013a).

A perfect x-wing is delimited by places forming a rectangle, each of them hosting a certain digit *x* (in this example *x*=2). On two of the diagram's "units" of the same type (columns or lines) which overlap a pair of similar units forming the rectangle's sides, *x* appears only twice (in this example 2 appears only twice along the columns 5 and 8 respectively, which overlap the x-wing's vertical sides) and both apparitions are in the x-wing corners. The

Non-Sashami finned x-wing oriented along columns, fins on the right column														
Fins in NE								***Fins in SE***						
..	..	..	..	..	..	..		..	..	..	..	..	..	..
..	..	..	..	F(i1-2,j2)	..	..		..	C(i1,j1)	..	..	C(i1,j2)	..	..
..	..	..	..	F(i1-1,j2)	..	..		..	..	..	..	..	..	..
..	C(i1,j1)	..... R	R	XD(i1,j2)	R R...	..		..	..	..	..	F(i2-2,j2)	..	..
..		..		F(i1+1,j2)	..	..		..	..	..	..	F(i2-1,j1)	..	..
..	..	..		F(i1+2,j2)	..	..		..	C(i2,j1)	...R	R	XD(i2,j2)	R R...	..
..		..		..	..	..		..	..	..	..	F(i2+1,j2)	..	..
..	C(i2,j1)	..		C(i2,j2)	..	..		..	..	..	..	F(i2+2,j2)	..	..
..	..	..	..	..	..	..		..	..	..	..	..	..	..
Non-Sashami finned x-wing oriented along columns, fins on the left column														
Fins in NW								***Fins in SW***						
..	..	..	..	..	..	..		..	..	..	..	..	..	..
..	..	F(i1-2,j1)	..	..	..	..		..	..	C(i1,j1)	..	..	C(i1,j2)	..
..	..	F(i1-1,j1)	..	..	..	..		..	..	...	..	..	..	..
..	.. R R	XD(i1,j1)	R	R.....	C(i1,j2)	..		..	..	F(i2-2,j1)	..	..	..	..
..	..	F(i1+1,j1)	..	..		..		..	..	F(i2-1,j1)	..	..	..	..
..	..	F(i1+2,j1)	..	..	..	..		..	.. R R	XD(i2,j1)	R	R.....	C(i2,j2)	..
..	..	..	..	..		..		..	..	F(i2+1,j1)	..	..	..	..
..	..	C(i2,j2)	..	..	C(i2,j2)	..		..	..	F(i2+2,j1)	..	..	..	..
..	..	..	..	..	..	..		..	..	..	..	..	..	..

Non-Sashami finned x-wing oriented along lines, fins on the upper line														
Fins in NW								***Fins in NE***						
..	..	..	..	..	..	..		..	..	..	..	..	..	..
..	..	R	..	..	..	..		..	..	..	..	R	..	..
..	..	R	..	..	..	..		..	..	..	..	R	..	..
...F(i1,j1-2)	F(i1,j1-1)	XD(i1,j1)	F(i1,j1+1)	F(i1,j1+2)	..	C(i1,j2)...		...	C(i1,j1)	...F(i1,j2-2)	F(i1,j2-1)	XD(i1,j2)	F(i1,j2+1)	F(i1,j2+2) ..
..	..	R	..	..	..	..			..	..	..	R	..	..
..	..	R	..	..	..	..		..	..	..	..	R	..	..
..	..	..	..	..	..	..		..	..	..	..	..	..	..
..	..	C(i2,j1)	..		...	C(i2,j2)...		...	C(i2,j1)	..	..	C(i2,j2)	..	..
..	..	..	..	..	..	..			..	..	..	..	..	..
Non-Sashami finned x-wing oriented along lines, fins on the lower line														
Fins in SW								***Fins in SE***						
..	..	..	..	..	..	..		...	..	..	..	..	..	..
..		C(i1,j1)			...	C(i1,j2)...		...	C(i1,j1)	..	..	C(i1,j2)		
..	..	R	..	..	..	..				..	..	R	..	..
..	..	R	..	..	..	..				..	..	R	..	..
..	..	..	..	..	..	..		..	..	..	..	..	..	..
...F(i2,j1-2)	F(i2,j1-1)	XD(i2,j1)	F(i2,j1+1)	F(i2,j1+2)	..	C(i2,j2)...		...	C(i2,j1)	...F(i2,j2-2)	F(i2,j2-1)	XD(i2,j2)	F(i2,j2+1)	F(i2,j2+2)..
..	..	R	..	..	..	..		...		..	..	R	..	..
..	..	R	..	..	..	..		...		..	..	R	..	..
...		..	..	..	..	..		...		..	..	..	..	..

Fig. 7. Non-Sashami finned x-wing configurations.

restrictions imposed by x-wing follow the rule (Stuart, 2013a) "when there are only two candidates for a value, in each of 2 different units of the same kind, and these candidates lie also on 2 other units of the same kind, then all other candidates for that value can be eliminated from the latter two units". In this example, 2-s are forbidden in the shaded places, which follow the upper and lower sides of the x-wing.

B. Non-Sashami and Sashami Finned X-Wing

The first type of special generalized x-wing configurations which allow for the indirect restriction application is known as „non-Sashami finned x-wing" in which a corner is not „perfect". The digit (addressed from this point forward as *XD*) used to build a finned x-wing oriented along columns:

- must appear in the x-wing corners denoted by *C*;
- must appear in the single x-wing corner denoted by *XD* (in the case of non-Sashami finned x-wings, otherwise a solved place should be there);
- must be present within at most 2 places from the area *F* (at least one apparition is compulsory, otherwise the structure is a perfect x-wing), the area *F* consisting in those places that share the same zone and column with *XD* (zone addressed as *ZF*). There will not be considered for building a x-wing structure those cases when the digit on which the x-wing is constructed, lying in *F*, and the other corner from the same column, share the same zone;
- will be restricted only in the places denoted by *R* which belong to *ZF*, over the line of *XD*, in places denoted by *R*.

All possible „non-Sashami finned x-wing" configurations are depicted by Fig. 7. The non-Sashami finned x-wing oriented along columns requires the presence of exactly 2 places containing *XD* along one of the columns. In the other column one corner must follow the perfect x-wing schematic rules, whilst the other one can be surrounded by at least one and at most 2 fins, placed in at most 2 places compulsory nearby the imperfect corner and sharing the same zone with it. The schematics corresponding to non-Sashami finned x-wings oriented along lines can be deduced from those oriented along columns, after a rotation by angles of 90^0 or 270^0 respectively.

The 3rd variant of x-wing schematics is represented by „Sashami finned x-wings". In this case, the place denoted by XD must be a solved place, as represented with rhomb in Fig. 8.

1389	6	1238	13789	179	4	5	27	79
5	4789	1248	1789	1679	26789	3	267	4679
349	3479	234	379	5	23679	8	1	4679
489	2	468	5789	4679	1	467	3	5678
7	489	1468	589	3	5689	146	568	2
1348	5	13468	2	467	678	1467	9	1678
3468	1	5	3479	2	379	679	678	3678
2346	34	9	13457	8	357	1267	567	13567
238	38	7	6	19	359	129	4	1358

Fig. 8. Sashami finned x-wing SE with one fin, oriented along the lines 4 and 9, detected and exploited at test no. 20.

5.2. Algorithmic considerations.

Owing to the great number of possible x-wing schematics, the function used to deal with x wings has a significant number of execution branches. The code segment used for x-wings oriented across columns is briefly described by Fig. 9.

```
for j1=1:8 //first column
 for j2=9:-1:j1+1 //second column
  for XD=1:9  //x digit
   c_exact_2=(sum(c(:,j1,XD))==2) && (sum(c(:,j2,XD))==2)
   c_exact_2_on_col1=(sum(c(:,j1,XD))==2);
   c_exact_2_on_col2=(sum(c(:,j2,XD))==2);
   c_min_3_on_col1=(sum(c(:,j1,XD))>=3);
   c_min_3_on_col2=(sum(c(:,j2,XD))>=3);
   if c_exact_2
      build i_col1, i_col2 /*vectors containing the first index of
                        places hosting XD over j1 and j2 */
       if i_col1(1)==i_col2(1) && i_col (2)==i_col2(2)
        //perfect x-wing over columns
          c(i_column1(1:2),1:9,XD)=0; // apply restrictions
          c(i_column1(1:2),(j1,XD)=1; //recovering corners
          c(i_column1(1:2),(j1,XD)=1;
               else
          if  (3 corners well positioned and one wrong positioned)
                        //possible x-wing Sashami with 1 fin
             if (conditions for Sashami with 1 fin accomplished)
               apply restrictions over the clue's row in clue's zone
               recover corners
             end;    end;
       else
            if c_exact_2_on_col1 &&  c_min_3_on_col2 //**
              build i_col1 and i_col2 to store the first indices
                 of  places containing XD on j1 and j2 ;
                   if (XD is properly positioned in j2)
                  //possible finned x-wing
                       if (all finns are in the zone Z with a single
                      corner - CO placed on j2)
                      apply restrictions over the line of CO in Z ;
                      recover corners from j2;
                        else
                       //possible Sashami x-wing with 2 fins
                        if (all fins are in zone Z with the clue Sashami
                        and the other corners are correctly aligned)
                        apply restrictions over the line of clue
                         Sashami for places from Z;
                         recover corners;
                       end;end;end; //**
                    else
                         if c_exact_2_on_col2 &&  c_min_3_on_col2
                             .... //a similar sequence of instructions
                               //as that delimited by ** is executed
                    end;end;end;
```

Fig. 9. General structure of the function dealing with x-wing structures oriented along columns

The modality of dealing with a Sashami finned with 1 fin NW oriented along columns is depicted by Fig. 10.

6. TESTS

Tests were performed considering 35 grids with different degrees of difficulty, 17 of them being evaluated as "hardest".

The results with respect to runtimes, number of unsolved places after the applying of direct restrictions and number of indirect restrictions generated by our original hybrid 3d algorithm are gathered by Tables 1 and 2.

```
c1 =c(i_col2(1),j1,10);
c2=(i_col1(2)==i_col2(2));
nr_c=number_of_zone(i_col2(1),j1);/ *identifier of restricted
                                        zone*/
c3= number_of_zone (i_col1(1),j1)==nr_c;  /*testing zones'
                                            matching */
 if c1 && c2 && c3
  for k1=j_left(nr_c):j_right(nr_c) /*only j-s from zone are
                                      considered */
        c(i_col2(1),k1,XD)=0;// applying restrictions
  end;
  c(i_col2(1),j2,XD)=1; //recovering corners
 end;
```

Fig. 10. Dealing with a Sashami finned NW with 1 fin along columns.

The symbols used in Table 1 to specify the sources of information were assigned as follows: '*' - (Nicolae, and Nicolae, 2012); '+'- (The new Sudoku players forum, 2009a); 'o'- (Stuart, 2013a), '$'- (Stuart, 2013b), and '++'- (Eppstein , 2012).

Three methods were used on the same test data, denoted in Table 2 as M1, .., M3: the classic 2d backtrack algorithm, our "vanilla" 3d backtracking algorithm (when the functions dealing with indirect restrictions are not used) and our original hybrid 3d algorithm respectively. T1...T4 were used in Table 2 to symbolize the number of indirect restrictions yielded by the following techniques: "Hidden/naked locked sets" (Nicolae, and Nicolae, 2012), "Line/column to box" (Nicolae, and Nicolae, 2012),, "ALS" and "Generalized X-wing".

Table 3 gathers the average and extreme runtimes along with ratios between runtimes generated through all mentioned methods.

An interesting aspect must be approached firstly relative to the grids with numbers 17, 18 and 20. They were classified as "extreme unsolvable", because the algorithm presented at (Stuart, 2013b) could not find a backtracking-free solution for them. Our hybrid 3d algorithm could detect indirect restrictions, as depicted by Figs. 8, 11...14 but the backtracking branch had to be activated to get the solution.

Table 1. Test Grids. Source, Difficulty and Initial Configuration .

Test no.	Source	Nickname (if available)	Degree of difficulty	Initial configuration of unsolved grid
1	*	-	EA	706800000000004810900000000004000000005900200080000502000600000000000030000476
2	*	-	T	309000400200709000087000000750060230600904008028050040000005900001060000060001104
3	+	-	H	120300000340000100005000000602400500000060070000008006004200300000070009000009080
4	o	Almost Locked Sets	E	060000509002400600905000070009010086000206905056080700090000162008007493604000857
5	o	Hard	H	002090300805000000100000000090060040000000058000000001070000200300500000000100000
6	o	Arto Inkala	H	800000000003600000070090200050007000000045700000100030001000068008500010090000400
7	o	Intersection removal	T	000921003009000060000000500080403006007000800500700040003000000020000700800195000
8	o	Sue de Coq	E	150432069904186005003070004002000040000090002030020500000000404034675000881024005 6
9	o	Quad Forcing Chains	E	008907050070040010060501007630000900009000800007400035700603020040070060006104700
10	o	-	E	006805030034600850580000010893100400000090003000000901300000040075003160648201390
11	o	Sashimi Finned	D	900008503080020097006000004300406002060080359800203001700860205650070018008500006
12	+	Discrepancy	HA	120400300300010050006000100700090000040603000003002000500080700007000005000000098
13	+	Cheese	HA	020050700400100006800003000200008003040020500000600010002090000090000005704000900
14	+	Fata Morgana	HA	000000003001005600090040070000009050700000008050402000800200900035001006000000000
15	+	Platinum Blonde	HA	000000012000000003002300400001800005060070800000009000085000009000405004700060000
16	o	Hiden triples	T	300000000970010000600583000200000900500621003008000005000435002000090056000000001
17	$	-	EU	600008940900006100070040000200610000000000200089002000000600050000000308000016000
18	$	-	EU	080009000000000206900050030002400170300020005007003000030090001204000000006000500
19	$	-	EU	004005806000009000802000100000630700037082000704100000400070200090000030085007000
20	$	-	EU	060004500000000300000050810020001030700030002050200090015020000009080000007600040
21	+	tarx0011	HA	003000000400080036008000100040060073000900000000020050040700686000000007006005000
22	+	-	HA	100050000007009030009007540004003070060000000090800000000790020000002403002000000
23	+	-	HA	003000000400080036008000100040060073000900000000002005004070068600004000700000500
24	+	-	HA	003000000400080036008300100040060073000900000000002005004070068600000000700000500
25	+	-	HA	003090000400080036008000100040060073000900000000002000004070068600000000700000504
26	+	-	HA	003000000400080036008300100040060073000900000000002005004070068600000000700000500
27	+	-	HA	003000000400080036008000100040060073000900000000002005004070068600004000700000500
28	+	-	HA	100400009056009000000010060060000805000040909000050100700002006000010500003000000
29	+	Golden Nugget	HA	000000039000001005003050800008090006070002000100400000009080050020000600400700000
30	+	Kolk	HA	120300000400000300003050000004200500000800090600050700015002000009006000000 7008
31	+	Imam_bayildi	HA	003006080000100200000070004009008060030040001070200000300005000005000600980000050
32	+	Coly013	HA	000090050010000030002300700004500070800000200000006400090010000080060000005400007
33	+	Red Dwarf	HA	120300004350000100004000000005400200600070000000008090003100500000009070000060008
34	++	-	E	309000074005000902000049000600003020000205000030100007000780000803000700910000508
35	++	-	D	026005001090300000130700000058006000000080000000400970000009043000004080900500210

[a] Legend: H = hard; E=extreme; T=tough; D=Diabolical; EA=Easy; HA=hardest

Table 2. Test results – number of unsolved places after direct restrictions, runtimes and number of digits indirectly restricted.

Test no.	Unsolved places after direct restrictions	Runtimes [s.]			Number of digits indirectly restricted from sets of possible solutions of unsolved places			
		M1	*M2*	*M3*	*T1*	*T2*	*T3*	*T4*
1	0	14.72	0.03	0.03	0	0	0	0
2	39	1.35	0.11	1.32	0	10	14	0
3	59	493.18	220.91	222.19	5	0	0	0
4	46	1.28	0.53	1.36	5	2	1	1
5	61	176.28	64.53	0.31	9	10	25	0
6	60	16.75	7.02	7.36	0	0	0	0
7	50	13.95	1.72	0.61	2	10	29	134
8	45	2.32	1.09	1.51	0	2	5	7
9	50	5.77	1.7	2.56	6	4	2	1
10	10	5.31	2.23	1.26	14	1	13	4
11	45	0.25	0.12	1.64	11	49	17	5
12	58	96.75	44.64	45.37	0	0	0	0
13	58	0.08	0.05	0.62	0	0	0	1
14	59	641.6	230.8	231.21	0	0	0	0
15	59	361.31	114.27	116.49	0	0	0	0
16	51	2.57	0.43	0.12	7	172	0	0
17	59	981.15	552.89	459.45	0	4	0	0
18	50	3.04	1.05	1.90	0	2	1	0
19	51	1.10	0.20	0.58	0	0	0	0
20	56	4.72	1.95	1.78	0	11	0	2
21	21	0.69	0.34	0.95	0	4	0	0
22	22	0.41	0.20	0.87	0	9	0	0
23	23	0.58	0.27	0.84	0	4	0	0
24	24	0.42	0.20	0.80	0	4	0	0
25	25	0.09	0.05	0.64	0	4	0	0
26	27	0.39	0.20	0.78	0	4	0	0
27	28	0.58	0.28	0.86	0	4	0	0
28	29	0.23	0.09	0.43	0	0	0	0
29	30	98.2	42.07	43.74	0	0	0	0
30	31	254.61	103.77	105.19	0	0	0	0
31	32	0.17	0.11	0.34	0	0	0	0
32	60	695.94	348.92	161.93	0	9	0	0
33	34	49	18.42	19	0	0	0	0
34	51	15.31	1.38	1.7	30	5	20	1
35	52	0.91	0.094	0.48	0	9	17	0

Table 3. Runtimes

	Runtimes [s]			Runtimes ratios		
	Backtra-cking classic	Vanilla Bactra-cking 3d (VB3d)	Hybrid Backtra-cking 3d (HB3d)	VB3d over Backtracking classic	HB3d over Backtracking classic	HB3d over VB3d
Mean	95.44	41.59	41.03	0.44	0.43	0.99
Max.	981.15	552.89	459.45	0.56	0.47	0.83
Min.	0.08	0.03	0.03	0.39	0.39	1

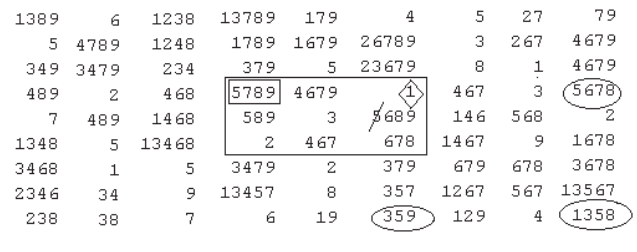


Fig. 11. Sashami finned x-wing NW with one fin - test no. 20.

The results revealed that searching for cases when indirect restrictions may be applied can be extremely fruitful as compared to both vanilla backtracking algorithms (e.g. tests no. 1, 5, 7, 17 and 32 at which tens or hundreds of seconds are saved or tests no. 10, 16 and 20, where

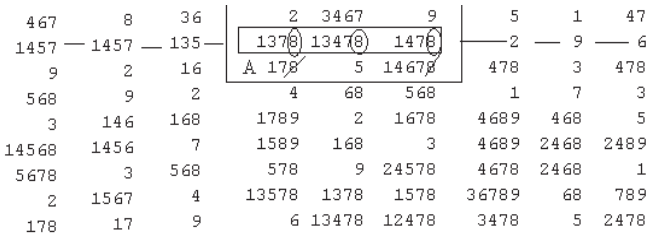


Fig. 12. Line-to-box restriction - test no. 18.

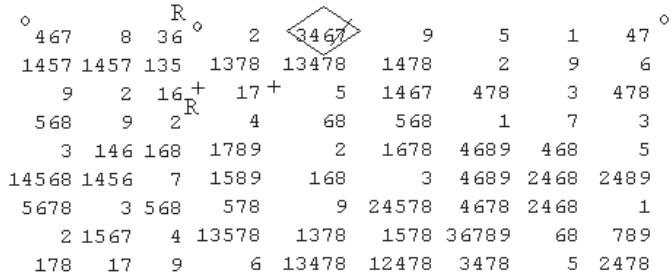


Fig. 13. Restriction using ALS - test no. 18.

seconds are saved). On the other side, for some cases, owing to the runtimes needed to run the indirect restrictions-related special functions (weather yielding indirect restrictions or not) slightly higher runtimes were revealed comparative to the 2d backtracking method, e.g. tests no. 4,11,13,21...28,31, but usually only tenths of seconds make the difference (at most 1.4 sec), which worth being sacrificed in some cases if hundreds of seconds are saved in others and the mean, maximum and minimum runtimes are reduced by over 50%.

For the analyzed cases, mainly at hardest grids, most of the runtime saving is actually provided by the implementation of the 3d variant of the backtracking code. As the runtime consumed by it is still high, our future work will focus on the implementation of strategies meant to limit even more the number of places to be solved by backtracking at very hard grids (but not limited to them) with advanced techniques like the detection of automorphism (The new Sudoku players forum, 2007) or of symmetric groups are (The new Sudoku players forum, 2009b).

Because the performances of the hybrid 3d algorithm related to the maximum, mean and minimum runtimes are clearly superior to those exhibited by the other analyzed algorithms, this algorithm should be the best option for time critical applications.

7. CONCLUSIONS

The logic, implementation, testing and evaluation of an efficient Sudoku solver, successfully tested for classic grids of all degrees of difficulty, were presented. An efficient 3-d array containing information on grid's content and on possible candidates for its unsolved places at every computation step was used. Auxiliary matrices containing flags related to restrictions associated with the restriction units were used, along with original functions that implement special strategies for indirect applying of restrictions („Hidden/naked pairs, triples or quads", „Box-

to-line/column restrictions" , "Almost Locked Sets" and "Perfect and Finned (non) Sashami x-wings"). These techniques result into the beneficial reducing of the number of possible solutions for places to be solved by backtracking, usually saving significant runtimes.

The tests made on grids with all degrees of difficulty revealed significant mean, maximum and minimum runtime savings (around 50%, in some cases reaching hundreds of seconds) as compared to a classic 2d backtracking algorithm and to free online solvers respectively.

The "built-in" features of the program (no libraries are required, only basic instructions being used) and the moderate additional memory relative to classic algorithms let us conclude that this algorithm represents a reliable option for applications where the time is critical.

REFERENCES

Brouwer, A. E. (2006). Sudoku puzzles and how to solve them. *Nieuw Archief Wisk.* (Ser. 5) 7:258–263, available at http://www.win.tue.nl/~aeb/preprints/ sudoku.pdf.

Daniel Liang, Y. (2007). Online free Java code to solve Sudoku with Backtracking, *Armstrong Atlantic State University*, available at , www.cs.armstrong.edu/liang /intro7e/book/Sudoku.java.

Detar, C. (2010). Automatic Sudoku Solver, available at tirl.org.software.sudoku/.

Eferanto, E. (2008). Sudoku creator/solver with PHP, available at www.emanueleferonato.com/ 2008/12/09/sudoku-creatorsolver-with-php

Eppstein, D. (2005). Sudoku Py, available at www.ics.uci.edu/~eppstein /PADS/Sudoku.py

Eppstein, D. (2005). Nonrepetitive paths and cycles in graphs with application to Sudoku, *ACM Computing Research Repository*, arXiv:cs.DS/0507053.

Eppstein, D. (2012). Solving Single-digit Sudoku Subproblems, *Lecture Notes in Computer Science*, Vol. 7288, pp 142-153, http://arxiv.org/pdf/ 1202.5074.pdf

Ercsey-Ravasz, M. and Toroczkai, Z. (2012). The Chaos Within Sudoku, *Nature*, article no. 725, http://arxiv.org/pdf /1208.0370v1.pdf.

Gordon, P. and Longo, F. (2006). *MENSA Guide to Solving Sudoku*. Sterling Pub., New York, USA.

Nicolae, I.D. and Nicolae, A.I. (2012). Limiting Backtracking in Fast Sudoku Solvers, *Annals of the University of Craiova, Automation, Computers, Electronics and Mechatronics*, Volume 9(36) no. 2, pp 25-32.

Python Fiddle, Shortest Sudoku Solver in Python, available at http://pythonfiddle.com/shortest-sudoku-solver-in-python/.

Rosenhouse, J. and Taalman, L. (2011). Taking Sudoku Seriously: The Math Behind the World's most Popular Pencil Puzzle, Oxford University Press, New York.

Simonis, H. (2005). Sudoku as a Constraint Problem, *CP Workshop on Modeling and Reformulating Constraint Satisfaction Problems* , pp. 13-27, http://citeseerx.ist. psu.edu/ viewdoc/download?doi= 10.1.1.88.2964 &rep=rep&type=pdf.

Stephens, P. (2007). *Mastering Sudoku Week By Week*, Sterling Pub, New York, USA.

Stuart, A. (2010). *The logic of Sudoku*, Michael Mepham Editorial Services, Somerset, UK.

Stuart, A, (2013a). Sudoku Solver, available at Sudokuwiki.org.

Stuart, A. (2013b). The Weekly Extreme 'Unsolveable' Sudoku Puzzle , available at http://www.sudokuwiki. org/weekly_sudoku.asp.

The new Sudoku players forum (2007). Automorphic Sudokues, available at http://forum.enjoysudoku.com/ automorphic-sudokues-t5588.html.

The new Sudoku players forum (2009a). The hardest Sudoku (New Thread), available at http://forum.enjoysudoku. com/the-hardest-sudokus-new-thread-t6539.html.

The new Sudoku players forum (2009b). About Red Ed's Sudoku symmetry group, available at http://forum.enjoysudoku.com/about-red-ed-s-sudoku-symmetry-group-t6526-75.html.

Zelenski, J. (2008). Lecture 11 - Programming Abstractions, *Stanford University*, available at http://www.youtube. com/watch?v=p-gpaIGRCQI.

Case Studies of Human Motion Analysis for e-learning in Biomedical Engineering

D. Popescu*, J.L. Pascual, M. Marian*, M. Ilie*, I.B. Boch****

* Faculty of Automation, Computers and Electronics, University of Craiova, Romania (e-mail: dorinp@robotics.ucv.ro)
** Biomechanics Institute of Valencia, Spain

Abstract: E-learning proposes a new method of learning in many educational fields. The project named "A Web-based E-Training Platform for Extended Human Motion Investigation in Orthopedics" proposes formation of engineers that will apply the principles of medical and biomedical engineering sciences for the use, adaptation, evaluation, projecting and/or distribution of technological and medical solutions that lead to improvement of the patients' health. The paper presents this project, its aims and the e-learning platform that offers to engineers interested in the medical field, a learning tool, with interdisciplinary approaches, trying to close the gap between engineering and medicine, by creating a new e-learning environment for human motion analysis.

Keywords: biomedical engineering, e-learning, motion analysis, case studies, pressure platform.

1. INTRODUCTION

The progress of technology, generally, and of electronics, informational technology, particularly, requires engineers connected to the latest findings that may learn lifelong and carry out properly work. This means "continuing professional development" or "lifelong learning". From this reason many countries introduced national lifelong learning programmes and, too, the European Commission's introduced a Lifelong Learning Programme that enables people at all stages of their lives to take part in continuing professional development, as well as helping to develop the education and training sector across Europe (http://ec.europa.eu/education/lifelong-learning-programme). This programme funds a range of actions and sub-programmes. One of them is Leonardo da Vinci Programme that funds several types of projects in the field of vocational education and training: Mobility actions; Partnerships; Multilateral projects (http://ec.europa.eu/education/lifelong-learning-programme).

This paper presents partially the Leonardo da Vinci project (transfer of innovation), named "A Web-based E-Training Platform for Extended Human Motion Investigation in Orthopedics" (ORTHO-eMAN), developed by institutions from Romania, Greece and Spain. The paper presents developed case studies of human motion analysis for e-learning in biomedical engineering.

2. E-LEARNING IN BIOMEDICAL ENGINEERING

Probably the fastest changes appear in engineering, which is facing a challenge by development of the new multidisciplinary fields. One of them is biomedical engineering (application of engineering principles and techniques to the medical field).

The evolution of biomedical engineering (BME) has only recently emerged as its own discipline. The work in this interdisciplinary field consists of research and development, spanning a broad array of sub-areas. Significant biomedical engineering applications include the development of various diagnostic and therapeutic medical devices ranging from clinical equipment to micro-implants, common imaging equipment such as MRIs and EEGs, regenerative tissue growth, pharmaceutical drugs and therapeutic biological.

Clinical engineering is the branch of biomedical engineering dealing with the actual implementation of medical equipment and technologies in hospitals or other clinical settings. One of the roles of clinical engineers is to serve as technological consultants for physicians.

Biomedical engineers and bioengineers require significant knowledge of both engineering and biology. As interest in BME increases, many engineering colleges now have a Biomedical Engineering Department or Programme, with offerings ranging from the undergraduate to doctoral levels. The number of biomedical engineers is expected to rise as both a cause and effect of improvements in medical technology. A comprehensive study about the BME education in Europe is presented in several reports like (http://www.worldwidelearn.com/online-education-guide/engineering/bioengineering-major.htm; http://www.biomedea.org/Status%20Reports%20on%20BME%20in%20Europe.pdf; Salerud *et al.* (2006)).

According to Choules (2007), biomedical engineering can be defined as "a profession exerted in healthcare institutions: the biomedical (or clinical) engineer is responsible of applying and implementing medical technology to optimise healthcare delivery".

In this context there is also an increased need for e-learning methods, tools and platforms to be used in

biomedical engineering education. This is also reflected in the number of European projects launched on this subject in the past few years.

Teaching and learning have moved towards active, student-centred, problem-based, challenge-based, inquiry-based, cooperative and self-directed learning (Kybartaite (2010)). The practice of using technology to deliver coursework has also created new opportunities for teaching and learning. For example, audio and video records, CD-ROMs and DVDs, personal computers (PCs), iPods, Internet and Web applications, i.e., wikis, blogs, podcasts (Boulos *et al.* (2006)) have all been adapted for educational purposes (Kybartaite (2010)). Nowadays, the terms "e-learning" or "virtual learning" are commonly considered as umbrella terms describing any type of learning that depends on or is enhanced by the latest information communication technology (Kybartaite (2010)).

It is possible to develop, implement, test and demonstrate the significance of educational theories, technologies and models through an open, free of charge, modern technology-based, high quality teaching and learning environment. An example of such an environment is the European Virtual Campus for Biomedical Engineering (EVICAB) (Kybartaite (2010)).

Learning can be seen as an investment into the future. Moreover, learning in times of dynamic change calls for methods which teach large audiences up-to-date topics in a cost efficient way. E-learning solutions are perceived as a possibility to increase the quality of education, while simultaneously lowering costs (Frank (2003)).

E-learning is the unifying term to describe the fields of online learning, web-based training, and technology-delivered instruction. E-learning is a very dynamic domain, in continuous growth, which refers to educational content or learning experiences delivered or mediated by means of digital technologies. The development of this domain will lead to a growth in the quality of instruction, cost reductions and a more efficient implementation of distance and life-long learning. E-learning can incorporate many elements that make learning new material, a new process or a new program more fun. Making learning more fun or interesting is what makes it more effective.

Today's e-learning is dominated by the Learning Management Systems (LMS), such as Moodle, Atutor, Sakai, ILIAS, Dokeos, Blackboard, Desire2Learn; these represent integrated systems which offer support for a wide area of activities in the e-learning process.

The use of e-learning in Romanian higher education institutions is continuously expanding. The universities are adapting their curriculum to allow the creation of new technology-enhanced learning settings in various areas (Istrate (2007)). There are both small-scale research initiatives (e.g. Popescu (2010, 2012)), as well as large-scale uses of dedicated platforms, such as learning management systems. The most widely used LMS in Romania is Moodle.

As far as biomedical engineering area is concerned, the e-learning approach is not well-developed in Romania. Dedicated undergraduate programs are offered at the Faculty of Medical Bioengineering, G.T Popa University of Medicine and Pharmacy Iasi, but they don't include technology-enhanced distance learning settings.

The demand for engineers in the conception, design, installation and training in the use of medical equipment and instruments has grown in line with the advances in medical technology. These essential aspects of health technologies are currently covered by the EU directives and the laws of all developed countries.

The objective of this LLP-LdV project is to train qualified engineers for health systems and medical instrumentation technology support positions in hospitals, health industry and research centers.

The ORTHO-eMAN project seeks to close the gap between engineering and medicine, by creating a new e-learning environment for human motion analysis. For this reason, the project aims to offer to orthopedic doctors and to engineers interested in medical field, a common learning tool, with interdisciplinary approaches using learning methodologies and experience from previous EU project ("e-MedI – Virtual Medical School").

The ORTHO-eMAN consortium is formed by University of Craiova (Romania), Democritus University of Thrace (Greece), National Center for Scientific Research "DEMOKRITOS" (Greece), Biomechanics Institute of Valencia (Spain) and Clinical Emergency Hospital Bucharest (Romania). All partners have experience in orthopedics or biomedical engineering, with demonstrated skills, recognized expertise and competence required to carry out all aspects of the project.

The main objective of the ORTHO-eMAN project consists in developing of a web-based e-training platform for human motion investigation in biomedical engineering. The project's e-training platform provides a repository of training material with real clinical case studies using digital imaging (Fig. 1).

The chapters of ORTHO-eMAN curriculum are:

1. Physics of the human body motion. Basic knowledge on materials and biomaterials properties.
2. Basic anatomy of human body motion.
3. Modern techniques in human motion analysis.

Only through an interdisciplinary approach and achievement of border knowledge by both involved (medical and engineering) will be possible to improve quality of care.

E-learning is a viable alternative, motivating the development of a number of learning/training environments. Recent reviews of e-learning literature in medical or biomedical engineering education contexts reveal that e-learning is at least as good as, if not better than, traditional instructor-led methods/lectures (Choules (2007); Konstantinidis *et al.*(2007); Salerud *et al.* (2006);

Johnson *et al.* (2004); European Virtual Campus for Biomedical Engineering, *EVICAB*, http://www.evicab.eu; Kybartaite (2010)).

Fig. 1. ORTHO-eMAN website (www.ortho-eman.ro)

3. NEEDS ASSESSMENT FOR E-TRAINING IN MEDICAL AND BIOMEDICAL ENGINEERING EDUCATION

The main objective of this survey was to set the characteristics of the training course that will be developed during the ORTHO-eMAN project, and to evaluate the learning necessities in human motion investigation in orthopedics. The survey constituted the point of departure for the integration and adaptation of the teaching contents that will be carried out during the project.

The specific objectives of the study have been:

- to know the socio-demographic and academic profile of the potential users of the learning platform;
- to identify the knowledge level at the job incorporation time and after some years of experience;
- to know if the formation received by these professionals was adequate;
- to value the possibility of carrying out a complementary formation;
- to identify the main learning areas for each profile;
- to provide information useful for the development of learning actions adapted to the necessities of the future trainees.

Data collection was done by auto-fulfilled questionnaire. Independent surveys were conducted for the four profiles defined of interest for the project development:

- managers of medical departments;
- academic medical staff;
- medical residents and medical doctors;
- engineers, bioengineers.

The project partners achieved the survey in Romania, Greece and Spain concerning the learning needs in human motion analysis of all groups of potential users of the training course.

The first part of all questionnaires was focused on characteristics (regional distribution, work place, experience, specialization etc.) of the participants. From the point of view of regional distribution the results can be considered representative for these three countries.

The second part of the questionnaires was focused on the participants' familiarity with online tools and e-learning. All groups answered similarly concerning the information sources, they are using. The majority in all groups uses online resources, then they use books and finally they use magazines to update their knowledge. The least favorable source is newspapers. The majority in all groups use these resources daily, then weekly and monthly.

The type of information they demand depends on their profession, medical issues required by academic medical staff and medical residents and doctors, and engineering issues required by engineers.

Another question was concerned with their familiarity with e-tools. The participants generally showed knowledge of the available tools. The most known and used e-tools are chat, audio conferencing and e-mail groups. The least known e-tool is mobile learning. The level of familiarity seems to be less for academic medical staff, increasing for medical doctors and giving top values for engineers.

All groups access the Internet daily and all of them use it to update their professional knowledge. The majority uses the Internet daily or weekly to improve their professional career.

The majority of the participants from Romania was no familiar with online courses and has no participated in one before. But the majority is interested on participating into an online course, which shows a very positive attitude towards e-learning and our online course on human motion analysis.

The third part of the questionnaire was about the participants' opinion on gait analysis. One of conclusions is that gait analysis is not widespread in Romania (comparative with Spain) and if it is achieved is usually performed via clinical evaluation, performed by physiotherapists. All groups of participants believe that gait analysis has medium or high degree of usefulness. In Romania if gait analysis is achieved is usually performed by doctors, not by engineers.

The fourth part of the questionnaire contains information about the contents and methodology of a potential online course on human motion analysis. The training interests for Spanish, Greek and Romanian residents and medical doctors are presented in Fig. 2 and for engineers in Fig. 3.

In conclusion, we found differences in the interests and competences between engineers and clinicians. These results reinforce the initial idea in the project of developing different contents for these two categories. It seems that there is interest in all the training topics

proposed to the potential users. All the profiles in the survey showed interest in theoretical and practical training and case studies.

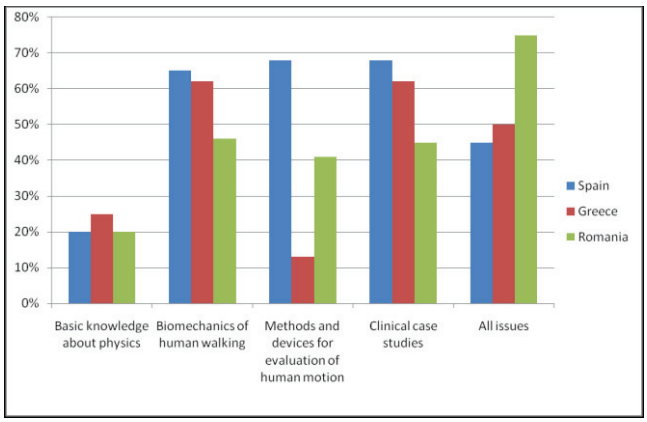


Fig. 2

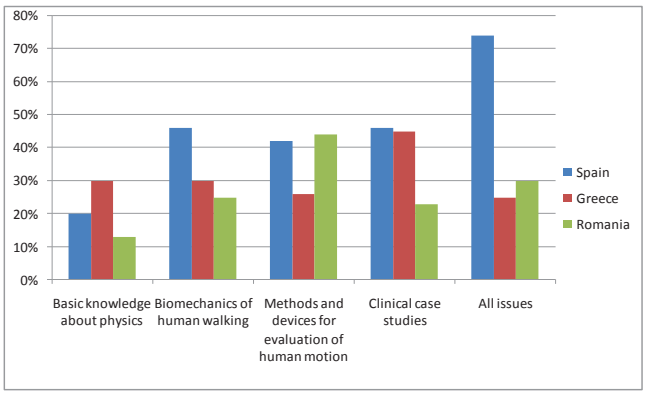


Fig. 3.

4. ORTHO-EMAN E-LEARNING PLATFORM

The target groups of ORTHO-eMAN project are formed by two main categories of trainees: orthopedic trainees (residents, doctors, physiotherapists) and bioengineers or engineers.

Human motion is a complex process achieved through a highly coordinated mechanical interaction between bones, muscles, ligaments and joints within the musculoskeletal system. To be able to perform and interpret the results of a motion analysis, a global training including knowledge of the musculoskeletal system, physics and human biomechanics and the use of measurement techniques is required.

For this purpose, there have been developed the following educational materials in the Ortho-eMan e-learning platform:

1. A theoretical part including all issues regarding motion analysis (anatomy, biomechanics, principles and main technical approaches);
2. Case studies from clinic orthopedic practice;
3. Case studies using up-to-date investigation methods (data acquisition and analysis using various techniques).

The teaching units for the new developed course are structured on multiple modules that contain: off-line courses (ppt), medical imaging, video files of motion analysis, graphs of forces, muscle and joint reaction forces, numerical data, contact pressure diagrams, an evaluation and self-assessment software. The system provides feedback by showing trainees the correct diagnosis. The training modules are available in English and partners' languages.

5. CASE STUDIES

In a case study, a description of a real or imagined situation is presented, and the trainee must analyze the data given in order to answer a particular question. To solve the case studies developed in the ORTHO-eMAN course, the trainees will need to apply the knowledge acquired in the three learning modules, at the same time that they will get some of the technical skills that they will need to perform motion analysis studies in the future.

The e-training platform includes case studies for: imaging for orthopedics pathology provided by classic imaging methods; motion analysis methods and applications; pressure analysis using Footscan; video motion analysis using SimiMotion.

The teacher creates the case studies using the Authoring Tool, developed in the ORTHO-eMAN project. Multiples possibilities have been implemented to make the learning easy and attractive for the student. It provides tools that enable the manipulation of multimedia content (image and video), the construction of quiz questions as well as a set of drawing abilities which identify critical regions or points that require identification by the trainee.

By means of this learning approach, the following goals are foreseen:

- Establishment of the theoretical concepts regarding motion analysis;
- Development of new knowledge and technical skills to perform motion analysis studies;
- Development of the ability of solving problems and taking decisions in real situations;
- Learning how to use and interpret the results of specific motion analysis systems.

A. Motion analysis methods and applications

Very often human movement alterations appear as consequence of a physical or a neurological impairment. Motion analysis has become a useful tool for clinicians, enabling them to make an assessment of the functional impact of a disease. Due to its utility, there has been a rapid evolution of the technology and its clinical applications in the recent years.

The *Module 3 - Modern techniques in human motion analysis* intends to give to the trainee a global view on the most important aspects of motion analysis:

- Description of motion analysis techniques and methods;

- Development of motion analysis studies;
- Interpretation of the results.

The case studies are the best way to transmit these practical concepts to the trainees. They offer to the trainee the possibility of work with real data from several measurement systems and they may receive feedback on the correct manner to perform the analysis.

When planning and performing a motion analysis, many questions may arise: e.g. kinematic or kinetic analysis? Which is the adequate measurement technique? How to analyze the results? In Module 3, the main kinematic and kinetic measurement techniques are described and a number of case studies deal with the methodological aspects regarding its use. An example can be seen in Fig. 4.

Other case studies are focused on photogrammetry and dynamometric platforms, as gold standard measurement systems. In these case studies, a deeper attention was put on the performance of the motion analysis, the measurement protocols and the biomechanical interpretation of the results.

Fig. 4. Case study on methodological concepts.

The inclusion of multimedia content from motion analysis studies performed with NedAMH/IBV (Lafuente *et al.* (2000)), Kinescan/IBV (Sánchez Zuriaga *et al*. (2011)), and Biofoot/IBV (Martínez Nova *et al.* (2007)) is an added value for the trainee, making easier the assimilation of knowledge and skills.

B. Footscan pressure plate

A set of case studies has been developed to train the trainee on the use of a specific system for pressure analysis, the 0.5 meter Footscan® pressure plate. The pressure plate is able to carry out both static (Fig. 6) and dynamic measurements (walking or running). The system outputs ground reaction force and the developed pressure during stance or walking (*** Footscan plate documentation – Footscan User Manual).

Gait analysis can be done by studying synthetic numerical data included in tables or image analysis included in dynamic screen, impulse screen, balance screen, comparing screen or gait screen. The trainee can analyse: information about the time distribution of plantar pressure level, force-time distribution and load on each area, the contact surface which is active, subtalar axis and angle foot, foot balance, gravity center position.

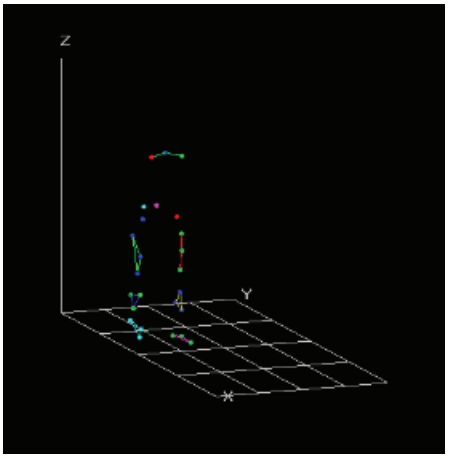


Fig. 5. Case study on methodological concepts.

For force plate measurements the authoring tool and the trainee's interactive e-training environment is able to:

- indicate zones of abnormal high or low pressures, identify the center of pressure, the highest impulse area, the contact percentage, foot axis or foot angles, center of mass (the red area, Fig. 6). With help of biomedical engineer who observed left-right differences, the medical doctor can identificate the possible causes for these differences like: proprioceptical problems, length difference of the lower limbs or alignment problems with orthotics-prostethics, and should investigate them further on. Also other differences in dividing of the percentage of the body weight can be observed: front-backwards difference that can mean possible static problems, and diagonal differences (possible pelvic rotation).

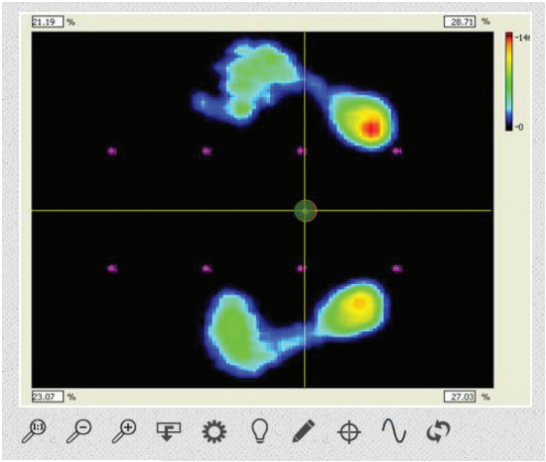


Fig. 6. Case study.

- measure angles.
- watch videos (recordings of pressures exerted on patient's feet during roll-off). The trainee will have to answer a quiz regarding abnormal features of the movement.

- target any desired point on the graph and estimate it's values for x and y axes (by drawing a horizontal and a vertical line on the graph image), or to find the largest or smallest value on a specified curve (Fig. 7). The value will have to be estimated. For example, in the image below, the trainee will be asked to find the largest value of the force exerted on the fourth metatarsal.

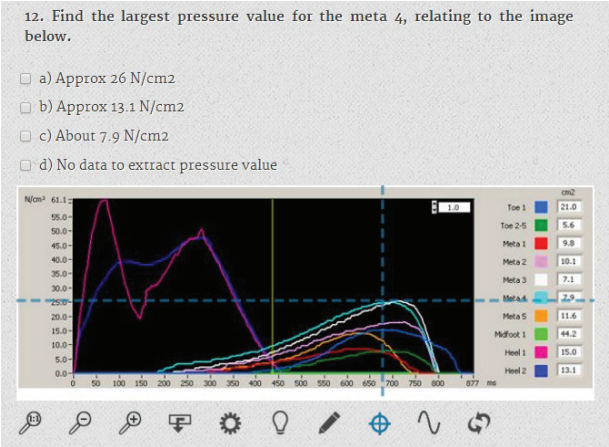


Fig. 7. Force/pressure graphs.

C. Video motion analysis using SimiMotion

Compared with most other methods of measurement, human motion image analysis has the advantage that it has no direct repercussions (Moeslund *et al.* (2001); Sezan *et al.* (1993); Aggarwal *et al.* (1997); Jun *et al.* (2005)). Motion is understood physically as the change in coordinates in a certain time span.

The trainees will be trained on the performance of 2D kinematic analysis of human body motion using SIMI Motion sofware (*** SimiMotion manual). Fist phase in motion analysis is description of movements with analytical stand out of motion system. Next phase is recording of motion features and analysis of motion. The obtained data are processed, graphics presented and analysed.

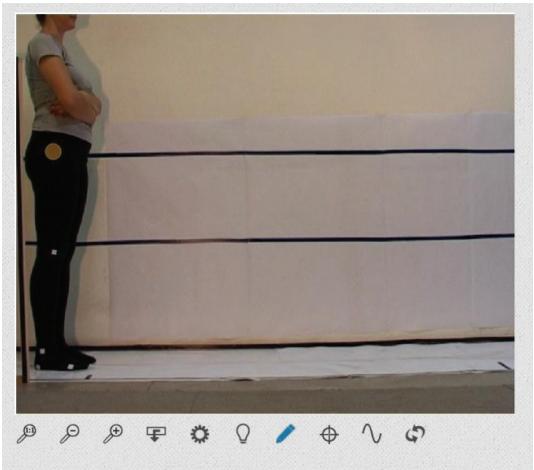


Fig. 8. 2-D recorded images for motion analysis.

If automatic tracking is used, light or dark markers (e.g. stickers) must be placed on the points which are to be captured (e.g. human joints) before the recording is made. The user then has to assign each point once and thereafter tracking of the video sequence is carried out by the computer.

The trainee has to solve some case studies. For example he has to identify the list of points/markers are to be captured (Fig. 8) or to approximate how much time does the patient needs to complete one step (Fig. 9).

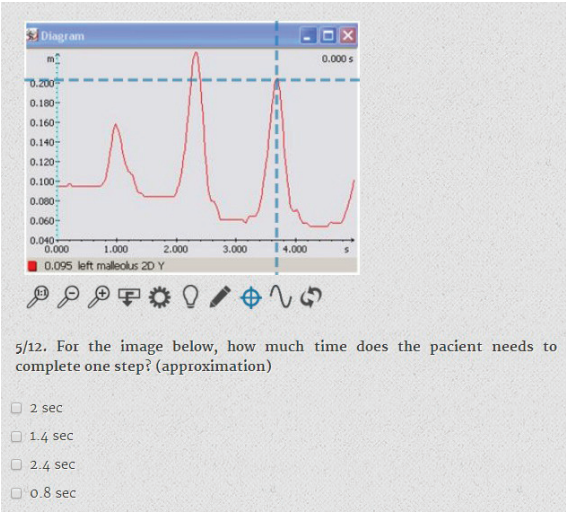


Fig. 9.

6. PILOT TESTING

Prior to the course launch, a pilot testing was conducted during the consortium meeting at Biomechanics Institute of Valencia (IBV) facilities in Valencia. Eight potential users of the ORTHO-eMAN e-learning platform followed an evaluation checklist and freely navigate through the course with the main aim of identifying usability and technical issues to be corrected.

The pilot testing was directed by IBV personnel, who solved the trainees' doubts and took note of their comments. The overall result of the learning platform evaluation was good, with positive opinions from the trainees and getting a mean grade of 7.75 in terms of usability.

The course enrolment, the User's Guide consultation, limited space in the screen for reading documents and the use of the case studies tools, were the main usability problems identified in the session. An error when working with videos with old versions of Mozilla Firefox browser was the most relevant technical problem detected.

The possible solutions were debated during the consortium meeting and were finally implemented in the following weeks. Thus, the results of the pilot testing permitted the improvement of the quality of the learning platform, in order to be adequate to the users' needs.

7. CONCLUSION

Nowadays, cross-disciplinary activities that integrate the engineering sciences with the biomedical sciences and clinical practice appear more and more. The ORTHO-

eMAN project seeks to close the gap between engineering and medicine, by creating a new e-learning environment for human motion analysis. Achievement of this e-learning platform will allow a close cooperation, by working in a team, between doctors and bioengineers for diagnostic and treatment targeting.

Our project adresses bringing together knowledge from the field of biomedical engineering and medicine in order to develop new procedures in human motion analysis.

One of the project target groups is represented by engineers. The project aims at the formation of specialists that will apply the principles of bioengineering science for the use, adaptation, evaluation, designing and/or distribution of technological and biomedical solutions that lead to improvement of the patient health.

The developed e-training platform was designed to support e-learning, to manage access to e-learning materials, consensus on technical standardization, methods for peer review of these resources.

The integration of e-training into biomedical engineering education can catalyze the shift toward applying on-line learning, where professors will not be mainly the distributors of content, but will become more involved as facilitators of learning and assessors of competency.

ACKNOWLEDGMENT

This work is supported by LLP-LdV-ToI-2011-RO-008 grant. This project has been funded with support from the European Commission. This publication reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

REFERENCES

Aggarwal, J.K., and Cai, Q. (1997). Human motion analysis: a review, Proc. of Nonrigid and Articulated Motion Workshop, pp. 90-102.

Boulos, M.N.K., Maramba, I., Wheeler, S. (2006) Wikis, blogs and podcasts: a new generation of Web-based tools for virtual collaborative clinical practice and education, BMC Medical Education, vol. 6, no. 41.

Choules, A.P. (2007). The use of elearning in medical education: a review of the current situation, in Postgraduate Medical Journal, 83, pp. 212-216.

EAMBES - What is Biomedical Engineering? Available at: http://www.eambes.org/about (Accesed April 2013).

European Virtual Campus for Biomedical Engineering, EVICAB, http://www.evicab.eu/ (Accesed April 2013).

Frank, C. (2003). Conceptual Design of the Web-Based Case Method - A Pedagogical Perspective, PhD thesis, University of Paderborn, Germany.

http://www.biomedea.org/Status%20Reports%20on%20BME%20in%20Europe.pdf (Accessed June, 2013).

http://ec.europa.eu/education/lifelong-learning-programme/ (Accesed April 2013).

http://www.worldwidelearn.com/online-education-guide/engineering/bioengineering-major.htm (Accesed April 2013).

Istrate, O. (2007). eLearning in Romania: the State of the Art, eLearning Papers, ISSN 1887-1542, No. 5.

Johnson, C.E., Hurtubise, L.C., Castrop, J. et.al. (2004). Learning management systems: technology to measure the medical knowledge competency of the ACGME, Medical Education, 38, pp. 599-608.

Jun, L., Hogrefe, D. and Jianrong T. (2005). Video image-based intelligent architecture for human motion capture, Graphics, Vision and Image Processing Journal, vol. 5, pp. 11-16.

Konstantinidis, S.T., and Bamidis, P.D. (2007). E-Learning Environments In Medical Education: How Pervasive Computing Can Influence The Educational Process, Bci, Sofia, Bulgaria, pp. 291-300.

Kybartaite, A. (2010). Impact of Modern Educational Technologies on Learning Outcomes: Application for e-Learning in Biomedical Engineering, Doctoral Thesis, Tampere University of Technology.

Lafuente, R., Belda, J.M., Sánchez Lacuesta, J., Soler, C., Poveda, R., and Prat, J. (2000). Quantitative assessment of gait deviation: contribution to the objective mesurement of disability. Gait and Posture, 11(3): 191 – 198.

Martínez Nova, A., Cuevas García, J.C., Pascual Huerta, J., Sánchez Rodríguez, R. (2007). Biofoot in-shoe system: Normal values and assessment of the reliability and repeatability. The Foot, 17:190-196.

Moeslund, T, and Granum, E. (2001). A Survey of Computer Vision-Based Human Motion Capture, Computer Vision and Image Understanding 81, pp. 231-268.

Popescu, E. (2010). Students' Acceptance of Web 2.0 Technologies in Higher Education: Findings from a Survey in a Romanian University, Proceedings DEXA Workshops, pp. 92-96.

Popescu, E. (2012). Project-Based Learning with eMUSE: An Experience Report, Proceedings ICWL, Lecture Notes in Computer Science, Springer, Vol. 7558, pp. 41-50.

Salerud, E.G., Petersson, H., and Ilias, M.A. (2006). EVICAB WP1: European Biomedical Engineering e-learning. Survey report of existing and planned BME distance courses in Europe, Linköping University Institute of Technology, EVICAB report.

Sánchez Zuriaga, D., López Pascual, J, Garrido Jaén, D., Peydro de Moya, M.F., Prat Pastor, J.M. (2011) Reliability and validity of a new objective tool for low back pain functional assessment. Spine, 36(16): 1279 – 1288.

Sezan, M.I., and Lagendijk, R.L. (1993). Motion Analysis and Image Sequence Processing, Springer.

*** Footscan plate documentation – Footscan User Manual.

*** SimiMotion manual.

Study on website development practices from a global market perspective

Ionuț Cristian Reșceanu*, Sabin Mihai Simionescu, Cristina Floriana Reșceanu***

* Department of Automation, Electronics and Mechatronics, Faculty of Automation, Computers and Electronics, University of Craiova, Romania (email: resceanu@robotics.ucv.ro)
** S.C. White Pyramid S.R.L., Online software solutions development department, Craiova, Romania (email: sabin@white-pyramid.com)

Abstract: This paper was written to summarize an analysis of current practices in multi-lingual website design. The analysis includes views on optimisation, speed, content distribution as well as age group, gender and local culture differences of the visitor impact on the usability and design of a website. Also analysed are techniques for information spanning across domain names and physical/virtual servers, SQL data distribution and methods of obtaining high quality translations for user generated content.

Keywords: Semantic Web, Web design, Uniform resource locators, Electronic publishing, Content distribution networks

1. INTRODUCTION

Information availability is something Internet users have become so accustomed to that the concept of not-available is considered a thing of the past. While information is widely available on the Internet, access to it can still be hindered for some users. Some users cannot afford the most recent in hardware, others have limited bandwidth with little possibilities for improvement, and some users have to access the Internet from library or school computers to do their research. After overcoming these situations, just having the information on the screen is not enough – the user has to be able to comprehend it. *The vast majority of websites are written in English, a language understood by over 3 billion people this day. Statistically, that leaves some hundreds of millions of Internet users often looking at information they do not understand.*

In the last decade, consumers have widely embraced the Internet. This has forced businesses to move a lot more focus in that direction, every year. While the Internet is a great place to find new clients and expand a business into new markets, connecting to each new market comes with new challenges. One of these challenges, one that affects any business, is the language barrier. Regulations in many countries require that product manuals be provided in the official languages for those countries, for each and every product. Customer support and warranties also have to be available in native languages.

The easiest and most cost-effective solution for this is having the company and/or product website created in more than the usual English – the website is available 24 hours a day, 7 days a week, and can handle an unlimited number of languages simultaneously. At the same time, the website source code can be written to take advantage of many pieces of information sent in the request headers, adapting the information provided on the web pages to the requirements of the visitor, often without having to ask the visitor any questions. On the pricing and costs perspective, the development of the website costs approximately the same as the recruitment + training process for a new employee, and the monthly running costs are similar to an employee's monthly salary – this makes having a website financially attractive to many businesses.

Multilingual websites are not a brand new concept. Since the introduction of DNS, HTML and the HTTP protocol, web designers could write websites in as many languages as they were requested to do. In the early days of the Internet, the major problem was that the HTML code had to be duplicated for every page. This would generate (number of pages) multiplied by (number of languages) .html files, creating a complex file structure which was hard to keep up to date. Also switching from language to language required a complex implementation, so the usual choice was to just link to the homepage in the destination language, effectively requiring the visitor to again navigate to the page they wanted to see.

Today, there are many server-side programming and scripting languages, reducing the workload for file and language management, and shifting the major effort to content creation and translation.

One of the biggest challenges in multilingual website development is keeping everything up-to-date, especially if the site is written in languages that the developer does not understand. In many cases, the site's base structure is written in English, and all the other languages are added on top of that. This makes developer collaboration more effective and less time-consuming. Also today's websites are more and more often crowd- or community-sourced, and every contributor only speaks a limited number of languages. Depending of the contents of the website, this challenge can be taken-on through different perspectives.

2. OTHER RELEVANT WORK

Most of the work done in this field is in the areas of real-time translation and cross-language search. These fields are still in their infancy, and while a lot of progress will be made and many developments are waiting around the corner, today search engines still need the users to input what they want, in the language that they want it. Even when cross-language search will become integrated in all major search engines, many users will still have problems with the results, because they will not understand the language the results are, even though the server did understand enough to consider them.

In the paper "OWL Rules: A Proposal and Prototype Implementation" (see Horrocks et al., 2005), the authors describe a way of using the OWL Web Ontology Language while reducing its limitations and proving meaning formal information to web pages with proper micro-formats coding.

In a paper about cross-language search engines (Chen and Bao, 2009) discussed methods of machine translation used by major search engines in an effort to find relevant content in other languages than the users type their queries. They also detailed a few digital libraries which implement multiple language content, like "Meeting of Frontiers", "International Children's Digital Library" and "The Perseus Digital Library". They concluded that bilingual users are the most likely users to benefit from content in multiple languages.

Also in the CLIR sector (cross-language information retrieval), in 2005, a business intelligence experiment was made by 4 researchers, Jialun Qin, Yilu Zhou, Michael Chau and Hsinchun Chen (see Qin et al., 2006). The experiment included web spider architectures, query translation, Query Expansion, Document Retrieval and Evaluation Methodologies. They concluded that translating the search query in each of the potential destination languages (the ones the system implements at a time) can significantly improve the search results quality. With today's technology, an improvement in indexing services would be to translate to a common language upon indexing, and query only against that to obtain results, while displaying the original text to the user at the end.

Live web-page translation gives users the ability to understand what that web-page is about. But it does not necessarily get new international visitors to the website. Today's search engine algorithms use keywords primarily to index web-pages. To index as actual semantic content, they would have to have a root concept-language and translate every indexed to and from that language. The main problem generated by such an approach is the web-page the user would see would be in the original writing language and they'd have to either understand that language or rely on automated translation solutions. While this method would be very useful for research and other high-level complex operations, ordinary everyday users would find this upsetting and most likely migrate to other search engines. For search engines to index a website's content in multiple languages that site would have to provide unique URLs for each page, in every language. Using cookies to store the current language chosen by the visitor and not differentiating the URLs will only get one language indexed, because search engine spiders will not use cookies or run JavaScript on the pages they process.

3. CONTENT TRANSLATION

3.1 General information

Big services providers, like Microsoft and Google, provide JavaScript APIs that automatically translates the HTML on any webpage, from one into another one of the available languages of the service. While implementing these APIs is very easy, the quality of these machine translations is acceptable for some users who want to get an idea on what the page is about, yet not enough for an official document or a research paper. Also these services are not always free – they might be free for a limited number of requests from a website, but nothing is free in unlimited form. Some web designers put language attributes on certain tags (like body, p, h1-h6 etc.) in the page, enabling web browsing software to provide the translation automatically to the user's chosen language, upon the users request.

The quality of human translations will continue to be better than the one of machine translation, even as developments in neural networks and machine interpretation make the difference less and less noticeable. Delivering the page directly in a language the visitor can easily understand can often make the difference between closing the sale or not.

While the multi-language feature can be added even after the website is live, designing the database optimized for storing the pages in multiple languages can increase the page-load speed significantly.

Keeping the pages in the database and responding to the browser's request with the proper language for every visitor does not fix the problem of actually obtaining translations for every page on the site. This problem can be solved by using human or machine+human translation, and translating page after page into each of the website's languages. This can be done manually, or by using an API from one of the professional translation websites. Creating an API connector for translation for the CMS of the website might seem unnecessary, but for large sites creating and implementing it can save a lot of time and money by automating the sending of translation requests of all new content as soon as the content is published in the first language.

The most important decision is whether to require all pages to be translated or not. For a company presentation site, and online stores, all pages should be available in all installed languages. Websites based on community data entry should have all the sections in all languages, yet not force users to input data in all installed languages. Users should be told that their entries may not appear in results lists in the languages they have not entered the data, and

that the website administrators have no obligation to translate the data if the users only provide the data in one language. Also creating a translations-centre page where contributors can see what translations are missing for which pages can significantly increase the speed of completion for language packs.

4. Implementation analysis

To make the analysis more conclusive, an implementation test was ran to study the impact of different programming languages and techniques would have on the performance and maintainability of a multi-lingual website.

The implementation started with a study to select which programming language and database software fit best for the task. Two product pairs were tested: PHP+MySQL and ASP.Net (C#) + MSSQL Server. The conclusion was that the integrated caching and multi-lingual resource files of ASP.Net out-weighed the open-source advantage of PHP.

4.1. Initial programming considerations

First, a layer of database abstraction is necessary, to speed up development and simplify implementation with a CMS. Since C# is object-based programming only, the solution was to implement a base-class, named DBStoredItem, which will handle standard CRUD operations between the application and the database. To operate with multiple objects at once, another base-class was written, to handle loading op multiple items at once: public class DBStoredItemList<T>: List<T> where T: DBStoredItem, new().

Actual loading from the database is done using the SQLDataAdapter class, and then using reflection the columns from the database are matched to properties in the destination object. The main problem was the ID column, usually a primary key in the database, which was a common between all tables. This made selections with JOINS problematic. The chosen solution was to implement a special attribute class, DBColumnAttribute, which was to be added to all properties that had their values stored in the database. Each class that inherits from DBStoredItem also has a custom attribute named DBTableAttribute attached, which tells the application the table name from which to select/update/insert into. The DBStoredItem class loads the data via a parameter of System.Data.DataRow type, matching the column names to properties in the current inheriting object. The function is mostly called by the DBStoredItemList class inheritors, which load the entire DataTable, instantiate new objects <T> for each row and call the objects LoadItemFromDataRow() method to copy the data from the DataTable to the destination object.

Data insertion is also covered, using the InsertIntoDB() method, which creates a new SqlCommand with INSERT syntax, adds all the necessary parameters via reflection, and then sets the ID of the object to the primary key auto-increment value that the database allocated to the new row.

Deletion is implemented at object-level, so database relations have to be managed before deletion – first the related objects, to avoid getting orphan entries, then the actual object. Child objects can be either deleted or moved to another parent, but they must not be left orphaned because there will be no way to access them from the site's interface afterwards.

Due to inheritance implementation limitations in C#, updating data could not be implemented in the base-class, so it was implemented using extensions.

```
public static class DBStoredItemExtension.UpdateColumn
<TValue> (this DBStoredItem item,
System.Linq.Expressions.Expression <Func< DBStoredItem,
TValue>> expression)
```

Updating data is called like:

```
model.UpdateColumn(item => model.Latitude);
```

One of the major advantages of this implementation is the synchronization between the application data-structure and the database table + table-columns structure. At application start-up, a call to the static function

```
DBSchemaManager.UpdateSchemaToLatestCode()
```

checks all the classes in the loaded assemblies list for the attribute DBTableAttribute and where found, matches the structure of the class against the structure of the database. In case of differences, the database structure is adapted (columns added or altered) to match the requirements of the application (data type can change from bool to int or string, and nvarchar columns sometimes need to change size to accommodate more information). Columns are not deleted though, to prevent production code from deleting columns added by the debug code. Tables that do not exist yet, as the classes that define them were not yet written at last application start-up, are created and a primary-key auto-increment identity(1) is created for each new table. This makes keeping the database structure synchronized to the application significantly automated, as the only operation the developer still has to do is create dedicated indexes for optimization.

While implementing this concept, another C# inheritance problem arose – C++ can have multiple base classes, while other languages can have 1 base class and implement as unlimited number of interfaces. To implement user accounts and roles, a switch had to be made, changing from *class DBStoredItem* to *interface DBStoredItem*. This allowed the following situation:

```
public class User : MembershipUser, DBStoredItem { … }
```

Thus, all members of the class DBStoredItem were moved to extension methods for the interface, and the interface itself only has one member: int ID. This allows all objects to use the unique primary key and link with each-other on a foreign-key basis, without having to always remember which column is the primary key. Also INNER JOIN loads are simplified by using different

names for the primary key columns in different tables, not all simply named ID and having to use aliases when executing SELECT statements.

The framework was built on the MVC architecture, using the MVC4 assemblies for ASP.Net.

Using System.Reflection and custom attributes, the actions in the controller classes are mapped to URLs for custom pages, after which the CMS part catches all unmatched requests and returns either the page (if found in the database) or a custom 404 error. The custom error page can be differentiated between debug-enabled users and regular visitors so that debuggers can get some extra info like processed page path and reason for 404 error:

Sitemaps are generated automatically as well. Webmasters will have to register the sitemap-index-[LOCALE].xml URLs in Google® and Bing® Webmaster Tools services, where [LOCALE] is replaced by the language-country pair for each enabled language on the site. The robots.txt file is also dynamic, listing all the sitemap index files, for each active language, and a sitemap.xml entry which is automatically redirected to the sitemap index for the default language for the current host.

4.2. Multi-lingual programming considerations

Another key interface is AppUrlProvider, which exposes the string TranslateUrl(int destinationLCID); method (LCID is the language code identifier, unique for every language in System.Globalization.CultureInfo). Classes that implement this interface are used mostly in custom applications, to provide the URL for the destination language so that the browser can take the visitors to the page they want to read, in the language they choose. Standard content management implementation for this will load the pages tree in the destination language, and then combine the URL parts of each page in the tree to generate an URL in the destination language.

Database normalization rules state that an application like this should split language data into separate tables. That poses a dilemma given the data-access-layer defined above. The chosen solution for this was:

```
public interface DBStoredMLO<T> where T : DBStoredItem {
      DBStoredMLOList<T> MLOs { get; set; }
      T GetMLO(int langID);
   }
public class DBStoredMLOList<T>:
System.Collections.Generic.Dictionary <int, T> where T:
DBStoredItem
   { }
public interface DBStoredMLOChild {
      int MLOOwnerID { get; set; }
      int LanguageID { get; set; }
      void LoadFromDBViaOwnerIDandLanguageID (int
OwnerID, int languageID);
   }
```

The solution uses composition to add the information required in multiple languages to the DBStoredItem inheritor, and there it can be accessed as properties indexed by LCID. Multi-language child objects, the classes that implement DBStoredMLO, also inherit DBStoredItem, as their loading is done by the same data-access-layer. This interface provides the link between the parent object (i.e. University – ID, street address, phone number …) and the child objects (UniversityMLO – name, description).

The HTML tag for a page that uses a header file template would look like

```
<html
lang="@System.Threading.Thread.CurrentThread.CurrentUI
Culture.TwoLetterISOLanguageName">
```

The generation of other-language URLs is done by a HtmlHelper like @Html.URLTranslation(1048), using:

```
public static class URLTranslationExtensions {
      public static string URLTranslation(this HtmlHelper
helper, int destinationLCID) {
                  return
((AppUrlProvider)(HttpContext.Current.Items
["TranslationProvider"])).TranslateUrl(destinationLCID);
      }
   }
```

4.3. Search engine optimization. Content and URLs

One of the most important features of a multi-language website is the ability to switch the current language and arrive on the same page content, just in the destination language. For search engine optimization reasons, some projects use different domains for different countries and languages. An example of this concept would be:

- 1) www.universities.com – international visitors, language EN-1033
- 2) www.universitati.ro – Romanian speaking visitors, language RO-1048
- 3) www.universitaten.de – German speaking visitors, language DE-1031

Doing this will maximize a web project's international reach, making visitors feel in the centre of attention. The default language on each TLD (top level domain, usually the part after the last dot) will be the language for that country/territory, and language attributes will be added to all relevant tags. In an interview for toprankblog.com, Andy Atkins-Krüger, CEO of UK based WebCertain, a specialist agency in multilingual search marketing, approached this subject at item 6, "Finding excuses to run with a dot com" (Atkins-Krüger, 2009).

Another potential problem, this one depending on the type of project, is keyword choosing. Keywords are critical for SEO, as a website is found more or less on the Internet depending on which keywords it was optimized for. Since universities share most of the keywords with each other, as they are in the same educational field, the difference will be made via TLD, optimizing for country first and keywords second. For a project to list all universities from all countries (project that goes above 200 individual pages per country), the most probable setup would be:

- 1) Use local TLD to display the universities in that country
- 2) Make every TLD in more than one language

- 3) Use www.domain.tld for the default language in that country
- 4) Use www.domain.tld/[lang-short-code]/, display the same content as www.domain.tld, only in another language
- 5) Lang-short-code is the 2 letter locale code for that country, most often the same 2 letters used in the TLD, defined in ISO 3166-1 alpha-2.
- 6) Practically, all the pages about universities in Romania will be available in RO, EN, DE, on the www.universitati.ro domain, all information about universities in Germany will be also in RO, EN, DE, just on www.universitaten.de and so on.

A small project would not benefit from the use of subdomains, as search engines consider subdomains individual web sites and thus each one would be very small and the website influence rank would be split between all the subdomains. The most common chosen solution is www.domain.tld/[lang-short-code]/. The goal is to keep all pages on the same domain, cumulating all content in one place, making the web site look and feel more complex to search engines. The major choice for this approach is whether to include a website homepage to allow the visitor to choose the language, or to redirect to the homepage URL for the default language, and allow the user to change the language from a menu.

Another strategy, one that needs significantly more content than the previous, is using only main domain, and create subdomain by pairing country and language in a syntax like: [country]-[language].domain.com. By doing this, the content is spread across many subdomains and thus the volume of content of each subdomain is only a very small part of the whole content published on the website. While costs are significantly reduced, as most TLDs are paid annually, this approach has worked for very few projects along the years. One example of success with this approach is Facebook®, which has grown not because of its search engine optimization, it has grown because people talk, read, write and print about it everywhere – an advertising method known as word-of-mouth.

Smaller projects usually do not usually require load balancing, so using one domain does not pose issues. Large projects, especially international/global projects, do require load balancing, and some projects grow so large that require more than one data centre to run. Medium projects, like a global list of universities, with faculty and alumni members, reviews, photo galleries and events catalogues, would benefit from using more average-powered servers than from using one very powerful server.

A good server structure for a project like that would be:

- One VPS/dedicated server in every country
- Real-time backup so if one of those fails, all traffic is transferred to the geographically closest server, or distributed between 2-3 closest servers.

A setup like this offers both fail-over protections from both hardware failure and natural disasters. The biggest advantage is the low latency provided by using local data centres, with links to local ISPs. This becomes very important if the website is multimedia rich, especially with many videos, and when no CDN is used to take the multimedia load off the main servers. For budget reasons, neighbouring countries can be grouped into one data centre, and split when traffic and loads are enough to justify another server.

Many elements from the graphical interface of the web sites will be similar – they all use the same CSS layout, same JS scripts, and mostly the same images. To reduce the number of DNS requests and because browsers only execute AJAX requests if the domain is identical to the current one, all common files, for all domains (and subdomains), for one server, will be placed in the same folder. ASP.Net comes with bundling and minification classes for optimizing the number of requests the browser needs to make to the server.

5. Other common industry practices

Today's web is governed by social media and social involvement, making web feedback an integrated part of the online ecosystem. Websites that do not offer visitors the ability to leave feedback start with a handicap and depending on the specifics of the website, risk losing visitors that feel more comfortable somewhere they can leave a mark. For multilingual websites, comments & feedback pose a new challenge: how to keep the integrity of the website and display both comments and content in the same language.

5.1. International audience analysis

Website audience = the complete collection of visitors (both humans and robots) that read one or more pages on a website, in a certain interval in time.

Young, old, male, female, English-speaking or not, with more or less computer skills, all these categories are part of the website audience. From a marketing perspective, it would be perfect if statistics like age, sex, spoken languages would be openly-available. Today, these statistics are only available through social networks. A big hit in marketing are web applications developed to be linked to social networks, these applications (mostly games) not only provide a good way to spend time online and learn about brands and products in a relaxing or enticing way, they also provide companies with demographical data about the users/players. Yet this information is not openly available by default – the users have to agree to allow the application access their social network accounts, and only some information is available to applications. Also, the more permissions an application requests from the user, the fewer users will agree to the terms and conditions and thus use the application.

By default, with no social network integration, today's web servers still receive some information about the visitor, yet mostly about the visitors' computer and not the person behind the screen. Because the most wide-spread protocol at this time is TCP-IP, the visitors IP address is sent to the server as a part of the request. The

IP address + the HTTP request headers provide the server with the following information:

- Device's location around the world
- Current browser
- Device's operating system
- Device's type (desktop /laptop or tablet, phone etc)

By correlating location with different calendars, the website skin can adapt to certain events taking place in some cities/countries, like "Earth Hour" which is global, to "National Day of […country name...]", which is different in every country. Doing this will make the visitor feel in the centre of attention, will make them feel more comfortable on the website and depending on the nature of the website, make them come back more often or even spend more time and/or money on the site.

5.2. Public relations: customer management

Newsletters are a big part of internet marketing today. Yet most companies only send newsletters in one language, even web portals with global reach. Delivering dedicated news, meaning both news of interest to the receiver, and news that the receiver can easily understand, in their native language, can strengthen the bond between the site and the user. CTR (Click-through-rate) can be significantly increased if the user relates to the material in the newsletter, and to do that first the user has to be able to understand the contents of that newsletter. Many web sites keep an online version of the newsletters, should the user not be able to read in their email client, and provide a link to the online version, in the email. These web versions can be improved with the ability to switch between languages. The email version can only be in one language, and should be in the language the user has chosen as default for the web site. Should the website not have a members-only section, and thus requirement for user registration and control panel, the native language can be identified by obtaining country and city information from the registration IP. For the benefit of the user, a header can be added to the email contents, with links all of the translated online versions. For websites with members sections, users should have the option to change the language the website sends messages to their email inboxes.

5.3. Traffic analysis and SEO considerations

Websites in multiple languages can use major service providers for web traffic analysis. The problem appears when detailed reports are required, with elements like a unique hits comparison on a single page, in all languages the page is written in. Obtaining this kind of information is not easy on common providers, as search engines lack the link between the pages that the resources table provides. For them, the same page in multiple languages looks like separate pages. The <link rel=canonical option does not help either, as it will only index 1 language, and the entire goal of a multiple language website is to be found and read in native languages of the visitors.

A solution for this is implementing a personal traffic analysis tool. This should be used in conjunction with a third-party service, for authentication of the numbers. Storing traffic information with language data can help webmasters identify pages with keywords that bring large traffic in some languages yet little in others, and thus know where to direct the keyword research efforts. Some languages will bring fewer visitors because people are accustomed to search for a certain product in certain languages, and in some countries the market for some products might be significantly smaller than in others. This should not discourage webmasters to continue providing the information in low-ranking languages, yet it should suggest that they make a very good initial assessment in which languages the website should be translated into.

5.4. Scaling for global reach and 24/7 availability

Location information (obtained via IP address) provides:

- Visitors country, county and city – perfect for dedicated advertising
- Country + city can also be processed to identify local language, and thus the visitors can be redirected to the section of the site in their native language
- ISP and network speed – redirect the browser to closer node in the CDN, to minimize loading time and network load
- Time zone – adapt "posted on" fields to the local time zone so "Monday at 13.45" is correlated to "10 minutes ago"

Browser information can be used to optimize CSS (and has been used for many years), but also to filter out bad web-bots and leeching. Also some features cannot be displayed on older browsers, and lack of certain plug-ins can sometimes make the page irrelevant. Because some manufacturers refuse to allow certain software on their devices, reading the browser and device information can help the server decide which page should it return, and if the website was properly coded, there should be a way to view any content on any device on the market at that time

Browser and location information can be used to identify what the device can display, and on what type of network the device is, and optimize the content so that the viewer gets the best experience on the website. Some providers also tax mobile usage per MB, so by reducing the number of images displayed, many of the visitors might save some money on their phone bill.

Depending on what the website is about and targeted for, it will be of interest for more or less people. A very large audience brings another problem: the bandwidth bottle-neck. While most web projects can be handled by one or a few dedicated servers running in the same datacentre, there are some that require more than one datacentre, on more than one continent. Depending on the project, using an existing CDN is sometimes a good way to go, but other times, as the project grows, traffic costs can be reduced by building or leasing datacentre space spread across the globe. An international audience also brings the problem of rush-hour spikes, as while people in some countries are asleep, others are at work 9-to-5, and others are resting or surfing the web after a hard day's work.

One approach is to identify the main audience, and optimize the location of the main cluster of servers for that audience, with the downside of increasing the distance and thus access times for the rest. Another approach is to cache data at various CDN hubs around the world, and deliver it from as close to the visitor as possible. This approach works best for data that is updated rarely, like large multimedia files, as it does not require caching of the entire website, leaving up-to-date data to be delivered from the central location.

A recently adopted approach is to optimize content delivery by location, not just large files through a CDN, but all the content, by identifying where it is most likely to be used most, and storing it as close to there as possible. An add-on to this approach is using local TLDs, allowing full optimization of content per country and language. While acquiring all domain names for all TLDs is a large yearly cost, it helps consolidate the business and brand name in every one of those countries.

For very large projects, combining the 3 approaches is usually the best way:

- one central, large installation, on the continent with the most users
- one or more other datacentres, located close to large internet hubs and distributed across continents
- leased CDN bandwidth for emerging markets, for cost optimization.

5.5. Special ecommerce considerations

Ecommerce websites benefit the most from having content in multiple languages. Still, between simply displaying the company presentation information (which rarely changes) and having to keep hundreds of product descriptions up to date there is a large difference in effort. The benefit of having more international clients is today shadowed by having to manually translate all products, offers and other information into another language, if not more than one. Ecommerce sites benefit the most from professional human translation APIs, as owners will only have to enter product information once, and they can have to option to display the product in search results on the site only if the product information has been translated X% into the visitors chosen language. The usefulness of a non 100% mandatory level depends on store type – a software programs e-store would have to translate more, while a clothing store would have far less per-product information to translate. A hardware e-store could have all the fields translated, and the operator/owner would just complete the fields' values for each product. This approach would reduce the time-to-market for many products, yet it might expose the store to some laws in some countries. Manuals and instructions should be available in the country-of-sale language – a user from Romania, buying from a store in the UK would get the manual in British English, as a mandatory option, while should that same user bought the product in a brick-and-mortar store in Romania, a product manual in Romanian would have been mandatory according to local laws.

Also support options should be offered in all the site's languages, whenever possible. If a user has bought a product (be it physical, virtual, immaterial or a service) while browsing the site in a certain language (which he or she supposedly understands enough to make the buy), when asking for support, be it via email, ticketing integrated solution or live-chat, the operator should answer in the same language as the buy was made. This would ensure a complete understanding of the users request and also comes with the highest chance that the user will understand the reply.

One of the most important parts of the ecommerce domain is the terms and conditions section. It is very important that this information is made available in all the site's languages. Also, it is considered good practice to keep previous versions available as well, in an archive (available to users or not – owners choice), so when a complaint is received, the legal department can analyze it in reference with the terms and conditions valid at the time it was received.

6. CONCLUSIONS

The central element on which a website should be build should be the visitor, the user. And not all users are literate in the native language of the company that owns the website. Writing a website in a single language in the 21st century should be limited to companies that only provide local services, to a well-defined type of clients, and even in this case, writing in multiple languages might help attract investors, partnerships and expertise exchanges. Considering the technology available today, users should have to translate the information on the Internet by themselves, it should be provided automatically in a language they can understand, or at least given an easy way to change the content language to one of their choice.

REFERENCES

Andy Atkins-Krüger (2009). CEO of WebCertain, in an interview for www.toprankblog.com/2009/10/top-10-pitfalls-of-international-seo/.

Chen, J. & Bao, Y. (2009). "Information access across languages on the Web: from search engines to digital libraries," *Online Proc. of 2009 Annual Conference of the American Society for Information Science and Technology*, Vancouver, British Columbia, Canada, Nov. 7-11, 2009.

Jialun Qin, Yilu Zhou, Michael Chau and Hsinchun Chen (2006). "Multilingual Web Retrieval", Journal of the American Society for Information Science and Technology, Vol. 57, Issue 5, pp. 671–683.

Ian Horrocks, Peter F. Patel-Schneider, Sean Bechhofer, Dmitry Tsarkov (2005). "OWL rules: A proposal and prototype implementation", *Journal Web Semantics: Science, Services and Agents on the World Wide Web archive*, Elsevier Science Publishers, Vol. 3 Issue 1, pp. 23-40.

Modeling and Hybrid Control of a SRSHR Manipulator

Viorel Stoian*, Daniel Strimbeanu**

*University of Craiova, Automation, Computers and Electronics Faculty,
Automation, Electronics and Mechatronics Department,
107, Decebal Street, 200440, Craiova, Romania
(Tel: +40-251-436.999; e-mail: stoian@robotics.ucv.ro).
**Technical High School "Nicolae Balcescu", Computers Department,
47, Nicolae Balcescu Street, 235100, Bals, Romania
(e-mail:strimbeanudaniel@yahoo.com)

Abstract: In this paper a new control algorithm for a hyper-redundant manipulator, which moves in a labyrinth and can touch the boundaries of the niche, is presented. The collisions or contacts are part of the task. The support-points from the walls help with the moving process. Initially, a study of the hyper-redundant locomotion by creeping and climbing is achieved. Also, the dynamic model by Lagrange method of the hyper-redundant robotic structure is established. Afterwards, a hybrid position/force controller for the hyper-redundant locomotion is proposed. Finally, the simulations of the hyper-redundant locomotion by climbing are presented.

Keywords: hybrid control algorithm, hyper-redundant manipulator, creeping and climbing movements, dynamic model.

1. INTRODUCTION

A Hyper – Redundant Manipulator (HRM) has the potential ability to perform non-conventional tasks, such as moving in highly constrained environment or grasping a various sizes and shapes of objects [Hirose, (1993)].

The term hyper-redundant refers to redundant manipulators with a very large or infinite number of degree of freedom. These manipulators are analogous in morphology to snakes, elephant trunks, and tentacles [Chirikjian and Burdick (1992)]. The term "wave" to describe the mechanism deformations associated with these hyper-redundant locomotion gaits (repetitive sequences of mechanism deformations that cause net motion of the robot) is used. Two highly idealized forms of biological locomotion are stationary wave locomotion and traveling wave locomotion.

The HRMs are categorized into 3 types of mechanisms: a large number of rigid segments in series (a hyper redundant chain of discrete links) - SRSHR manipulator, [Kobayashi and Ohtake (1995), Mochiyama et al. (1996), Paljug et al. (1995)], a cascade of modules composed of redundantly actuated parallel rigid platforms [Chirikjian and Burdick (1993), Zanganeh and Angeles (1995)] and truly flexible physical tentacle-like structures [Ivanescu and Stoian (1998)]. This paper presents a serial rigid link robot manipulator with hyper degrees of freedom, a SRSHR manipulator.

One of the cardinal goals of the robotic researchers is to provide control algorithms that allow robots to move in an unknown environment without collisions with the unstructured obstacles. Generally, in specific literature this problem is addressed as the collision avoidance problem [Chirikjian and Burdick (1993), Colbaugh et al. (1989), Khatib (1986), Zghal et al. (1992)]. In this paper the hyper-redundant manipulator is analysed in following situation: his job is to infiltrate, to creep through dropped walls after a cataclysm in a humanitarian action or for inspection in highly constrained environments.

The mechanics are for a snake-like structure but with serial rigid links. In this case, a compliant motion by touching of objects in many points of the manipulator like a human been in a very narrow cave which press with his shoulder the wall for his stability and with his hand push up a block (gravestone, manhole) or for climbing is necessary. The collisions or contacts are a part of the task. The support points from the walls help to the moving process. For example, in Fig. 1 (for simplicity, a planar structure) the robot fixes its i-link on the walls pressing the link's ends by F_i and F_{i+1} forces. In this situation the upper part (links i+1, i+2 … n) can be relaxed and inferior part (links 1, 2 …i-1) can be constricted. This process allows a creeping in horizontal plane or a climbing in vertical plane.

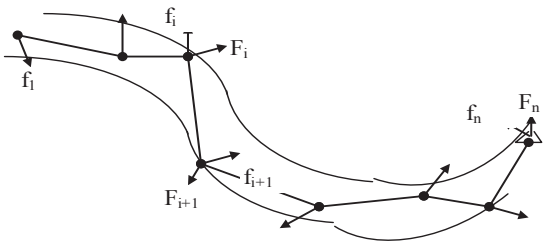


Fig. 1. The SRSHR manipulator in a labyrinth.

2. HYPER-REDUNDANT LOCOMOTION

Into a spatial and complex labyrinth two kinds of actions are necessary to combine: creeping and climbing. These actions, called "the hyper-redundant locomotion" are defined to mean motions resulting in net displacements of hyper-redundant robots due to the internally induced bending of the mechanism [Chirikjian and Burdick (1991)]. For simplicity, in this paper a planar structure is studied, but the ideas are generalized to the spatial case. Fig. 2 presents the hyper-redundant locomotion with amplitude varying wave that remains stationary with respect to body coordinates. ΔL represents the step for one cycle of motion.

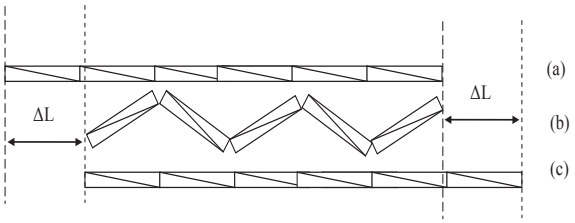


Fig. 2. The hyper-redundant locomotion.

In state (a) the robot there is in the initial position, in state (b) it has the upper end strongly fixed in terrain, the inferior end is discharged and it moves ahead. In state (c) the inferior end is strongly fixed in terrain and the upper end in discharged. Now, the hyper-redundant robot can relax ahead and it has again the initial position, but moved with the displacement ΔL.

In [Chirikjian and Burdick (1991)] the hyper-redundant locomotion with traveling wave with constant amplitude is presented. The two ends of the robot have the same functions like the ones from Fig. 2. Here there is only one wave that travels the length of the body from the bottom of the mechanism to the upper end.

In Fig. 3 a climbing process by the same hyper-redundant robotic structure is presented.

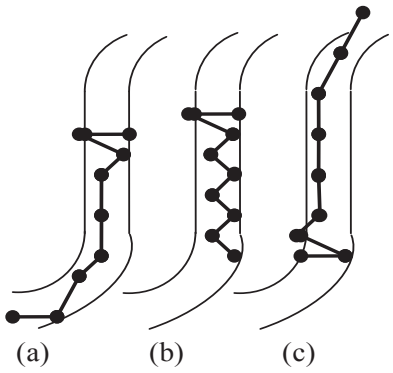


Fig. 3. The climbing process.

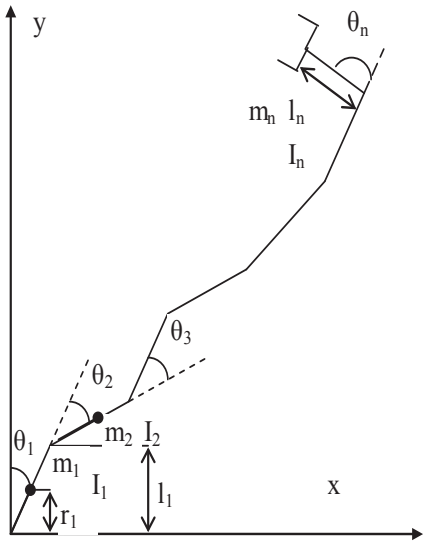


Fig. 4. The SRSHR manipulator structure.

3. DYNAMIC MODEL

In Fig. 4 a hyper-redundant robotic structure for establish a dynamic model by Lagrange method is presented. With the notations from that figure all phases of the algorithm are passed get through. The coordinates x(q) and y(q) are:

$$x_k = \sum_{i=1}^{n} l_i \cos\left(\sum_{j=1}^{i} \theta_j\right) \tag{1}$$

$$y_k = \sum_{i=1}^{n} l_i \sin\left(\sum_{j=1}^{i} \theta_j\right) \tag{2}$$

The velocities of the x and y are:

$$\dot{x}_k = -\sum_{i=1}^{k} \left\{ l_i \left[\sin\left(\sum_{j=1}^{i} \theta_j\right)\right]\left(\sum_{j=i}^{i} \dot{\theta}_j\right)\right\} \tag{3}$$

$$\dot{y}_k = \sum_{i=1}^{k} \left\{ l_i \left[\cos\left(\sum_{j=1}^{i} \theta_j\right)\right]\left(\sum_{j=i}^{i} \dot{\theta}_j\right)\right\} \tag{4}$$

The kinetic energy and the potential energy are:

$$E_c\left(q,\dot{q}\right) = \frac{1}{2}\dot{q}^T M(q)\dot{q} \tag{5}$$

$$E_p(q) = \sum_{i=1}^{n} E_{Pi} = g \sum_{i=1}^{n} mi \left[\sum_{j=1}^{i} l_j \sin\left(\sum_{k=1}^{j} \theta_k \right) \right] \quad (6)$$

where M(q) is the inertial matrix and

$$q = \left[\theta_1, \theta_2, ..., \theta_n \right]^T \quad (7)$$

The Lagrange function is:

$$L(q,\dot{q}) = E_c(q,\dot{q}) - E_p(q) = \frac{1}{2} \sum_{i,j=1}^{n} M_{ij}(q) \dot{q}_i \dot{q}_j - E_p(q) \quad (8)$$

Now, we can apply Lagrange's equation are:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} = T_i \quad (9)$$

The two terms of the Lagrange's equation are:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i} \right) = \frac{d}{dt}\left(\sum_{j=1}^{n} M_{ij} \dot{q}_j \right) = \sum_{j=1}^{n} \left(M_{ij} \ddot{q}_j + \dot{M}_{ij} \dot{q}_j \right) \quad (10)$$

$$\frac{\partial L}{\partial q_i} = \frac{1}{2} \sum_{j,k=1}^{n} \frac{\partial M_{kj}}{\partial q_i} \dot{q}_k \dot{q}_j - \frac{\partial E_p}{\partial q_i} \quad (11)$$

Now, the Lagrange's equation is:

$$\sum_{j=1}^{n} M_{ij}(q) \ddot{q}_j + \sum_{j,k=1}^{n} \Gamma_{ijk} \dot{q}_j \dot{q}_k + \frac{\partial E_p}{\partial q_i}(q) = T_i \quad (12)$$

where:

$$\Gamma_{ijk} = \frac{1}{2}\left(\frac{\partial M_{ij}(q)}{\partial q_k} + \frac{\partial M_{ik}(q)}{\partial q_j} + \frac{\partial M(q)}{\partial q_i} \right) \quad (13)$$

If we define the Coriolis matrix as:

$$C_{ij}(q,\dot{q}) = \sum_{k=1}^{n} \Gamma_{ijk} \dot{q}_k = \\ = \frac{1}{2} \sum_{k=1}^{n} \left(\frac{\partial M_{ij}(q)}{\partial q_k} + \frac{\partial M_{ik}(q)}{\partial q_j} + \frac{\partial M_{kj}(q)}{\partial q_i} \right) \dot{q}_k \quad (14)$$

Then $C(q,\dot{q})\dot{q}$ is the Coriolis and centrifugal forces vector. The third term is a position term representing loading due to gravity.

$$\frac{\partial E_p}{\partial q_i}(q) = G(q) \quad (15)$$

T_i represents the outer forces:

$$T_i = T_i - B(\dot{q}) - \tau \quad (16)$$

where T_i are the actuators forces (inner generalized forces), B represents the friction forces and τ represents the forces from joints due to the forces acting on the tool tip or on the another points from the robotic structure due to a load or contact with the environment. With the above-mentioned notations we have the dynamic model of the hyper-redundant robotic structure from Fig. 4:

$$M(q)\dot{q} + C(q,\dot{q})\dot{q} + B(\dot{q}) + G(q) + \tau = T \quad (17)$$

In this paper, by "position" we mean "position and orientation", and by "force" we mean "force and torque".

4. HYBRID CONTROL

During creeping in tunnels or labyrinths the robot comes in contact with environment; therefore, interaction forces develop between the robot and environment. Consequently, these interaction forces, as well as the position of the points from the robot, must be controlled. A number of methods for obtaining force information exist: motor currents may be measured or programmed, motor output torques may be measured and joint-mounted sensors may be used. There are two primary methods for producing compliant motion: a passive mechanical compliance built into the manipulator, or an active compliance implemented in the software control loop, force control. Passive compliance offers some performance advantages, but the force control method offers the advantage of programmability. This allows the manipulator to use the particular form of compliance necessary for the particular application. Throughout this paper is assumed that the desired velocity and force trajectories, which are commanded by the controllers, are consistent with the model of the environment [Lipkin and Duffy (1998)]. If this is not the case, it may be possible to modify the desired velocity and force trajectories to be consistent with the model of the environment. The compliant motion occurs when the position of a robot is constrained by the task. The so-called hybrid position/force controller [Chae et al. (1988), Raibert and Craig. (1981)] can be used for tracking position and force trajectories simultaneously. The basic concept of the hybrid position/force controller is to decouple the position and force/control problems into subtasks via a task space formulation because the task space formulation is valuable in determining which directions should be force or position controlled. Fig. 5 presents a hybrid position/force controller.

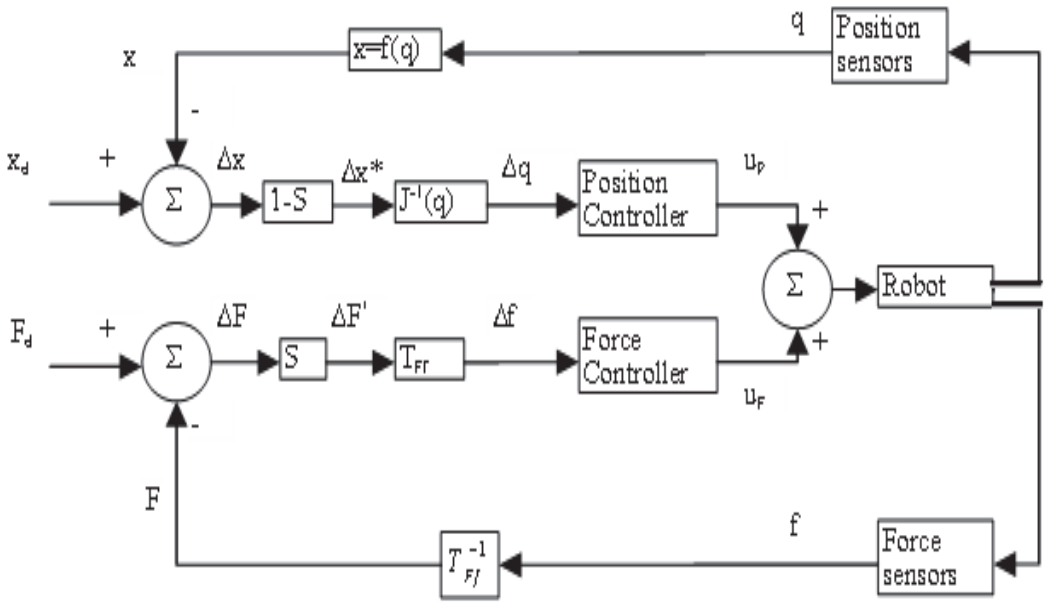


Fig. 5. The hybrid position/force controller.

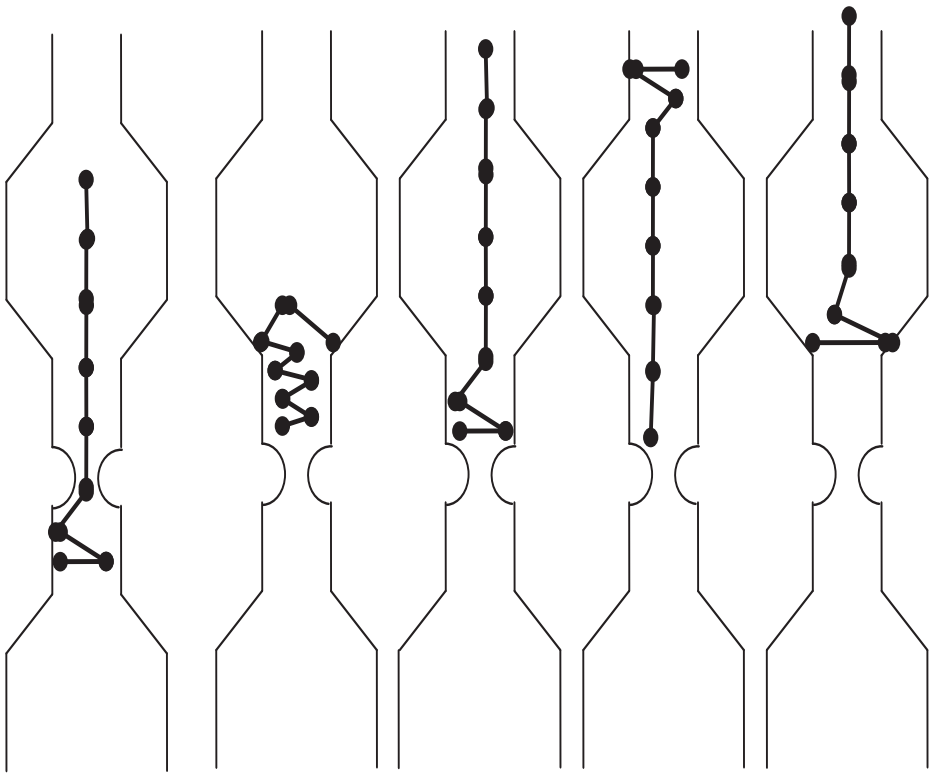


Fig. 6. Simulation for a climbing action.

$$x_d = \begin{bmatrix} x_d \\ y_d \\ z_d \end{bmatrix}; x = \begin{bmatrix} x \\ y \\ z \end{bmatrix}; \Delta x = x_d - x;$$
$$F_D = \begin{bmatrix} F_{xd} \\ F_{yd} \\ F_{zd} \end{bmatrix}; F = \begin{bmatrix} F_x \\ F_y \\ F_z \end{bmatrix}; \Delta F = F_D - F \quad (18)$$

$$q = \left[q_1, q_2, ...q_n\right]^T; f = \left[f_1, f_2, ...f_n\right]^T \quad (19)$$

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; S = \begin{bmatrix} S_x & 0 & 0 \\ 0 & S_y & 0 \\ 0 & 0 & S_z \end{bmatrix} \quad (20)$$

where S defines the compliance selection matrix.
S_i=0 for position control
S_i=1 for force control

$$T_{Ff} = J^T; T^{-1}{}_{Ff} = \left(J^{-1}\right)^T \quad (21)$$

where J(q) is the Jacobian matrix and T_{Ff} is a metrical operator which make the transformation between the forces F and the forces F. The relation x = f(q) is a relation which represent the forward kinematics. For position control, a PID controller is advantageously to use and for force control, a PD controller is advantageously to use.

$$u_p = k_1 \Delta q + k_2 \Delta \dot{q} + k_3 \int_0^t \Delta q dt \quad (22)$$

$$u_F = k_4 \Delta f + k_5 \Delta \dot{f} \quad (23)$$

The general control law u_i is:

$$u_i = \sum_j \left[\Psi_{ij} S_j \Delta f_i + \Xi_{ij} \left(1 - S_j\right) \Delta q_i \right] \quad (24)$$

where Ψ_{ij} is a force control law and Ξ_{ij} is a position control law. Fig. 6 presents a simulation for the climbing operation through a vertical labyrinth.

Any segments of the hyper-redundant robot are used for the anchorage of the structure on the walls of the vertical labyrinth while the others segments are moving up. The hyper-redundant locomotion by creeping is analogous.

5. CONCLUSIONS

A new control algorithm for a hyper-redundant manipulator, which moves in a labyrinth and can touch the boundaries of the niche, is presented in this paper. The collisions or contacts are part of the task. The support-points from the walls help with the moving process. Initially, a study of the hyper-redundant locomotion by creeping and climbing is achieved. Also, the dynamic model by Lagrange method of the hyper-redundant robotic structure is established. Afterwards, a hybrid position/force controller for the hyper-redundant locomotion is proposed. Finally, the simulations of the hyper-redundant locomotion by climbing are presented.

REFERENCES

Chae, A., Atkenson, C., and Hollerbach J. (1988). *Model-based control of a robot manipulator*, Cambridge, MA, MIT Press.

Chirikjian, G.S. and Burdick, J.W. (1991). Kinematics of hyper-redundant robot locomotion with applications to grasping, *Proc. of the 1991 IEEE international conference on robotics and automation (ICRA'91*), Sacramento, California, vol I, pp 720-725.

Chirikjian, G.S. and Burdick, J.W. (1992). A geometric approach to hyper-redundant manipulator. Obstacle avoidance. Transactions of the ASME, *Journal of mechanical design*, vol 114, pp 580-585.

Chirikjian, G.S. and Burdick, J.W. (1993). Kinematically optimal hyper-redundant manipulator configurations, *IEEE Transactions on robotics and automation, vol II*, No.6, pp. 794-805.

Colbaugh, R., Seraji, H., and Glass, K.L. (1989). Obstacle avoidance for redundant robots using configuration control. *Journal of robotic systems*, vol. 6, pp 72–44.

Hirose, S. (1993). *Biologically inspired robots: Snake-like locomotors and manipulators*. Oxford University Press, USA, ISBN-13: 978-0198562610.

Ivanescu, M. and Stoian, V. (1998). A control system for cooperating tentacle robots. *Proc. of the 1998 IEEE international conference on robotics and automation (ICRA'98)*, Katholieke Universiteit, Leuven, Belgium, vol 2, pp. 1540-1545.

Kobayashi, H. and Ohtake, S. (1995). Shape control of hyper-redundant manipulator. *Proc. of the 1995 IEEE international conference on robotics and automation, (ICRA'95)*, Nagoya, Japan, vol 2, pp. 2803-2808.

Khatib, O. (1986). Real-time obstacle avoidance for manipulators and mobile robots. *International journal of robotic research*, vol. 5, pp. 90–98.

Lipkin, H. and Duffy, J. (1988). Hybrid twist and wrench control for a robot manipulator, Transactions ASME. *Journal of mechanics, transmissions and automation design*, vol 110, pp. 138-144.

Mochiyama, H., Shimemura, E., and Kobayashi, H. (1996). Control of serial rigid link manipulators with hyper degrees of freedom: Shape control by a homogeneously decentralized scheme and its experiment. *Proc. of the 1995 IEEE inter. conference on robotics and automation (ICRA'96)*, Minneapolis, Minnesota, vol. 2, pp. 2877-2882.

Paljug, E., Ohm, T., and Hayati, S. (1995). The JPL serpentine robot: a 12 DOF system for inspection.

Proc. of the 1995 IEEE international conference on robotics and automation, (ICRA'95), Nagoya, Japan, vol. 3, pp. 3143-3148.

Raibert, M. and Craig, J. (1981). Hybrid position/force control of manipulators, *Journal of dynamic systems measurement and control*, vol. 102, pp. 126-133.

Zanganeh, K. E. and Angeles, J. (1995). The inverse kinematics of hyper-redundant manipulators using splines. *Proc. of the 1995 IEEE international conference on robotics and automation (ICRA'95),* Nagoya, Japan, vol 2, pp. 2797-2902.

Zghal, H., Dubey, R.V., and Euler, J.A. (1992). Collision avoidance of a multiple degree of redundancy manipulator operating through a window. Transaction of the ASME, *Journal of dynamic systems, measurement and control*, vol 114, pp. 717-721.

Author Index