A COMPLEX NON-CONTACT BIO-INSTRUMENTAL SYSTEM

Dan-Marius Dobrea (*) and Adriana Sîrbu

Technical University "Gh. Asachi", Faculty of Electronics & Telecommunications,
B-dul Carol I, no. 11, 700506- Iasi
ROMANIA
+4-0729188673, +4-0723076037
Fax ; +4-0232-217720
mdobrea@etc.tuiasi.ro, asirbu@etc.tuiasi.ro

INTRODUCTION

One of the major challenges that the human computer interface (HCI) faces nowadays is that of identifying a subject's state, in a real world environment, characterized mainly by: open-recorded, event-elicited and internal ***emotional state***-driven (Picard, Vyzas & Healey , 2001). The main requirement for such systems regards the noninvasive character of their working principle.

Subsequently, in order to improve communication in HCI systems or to asses the human state, the analysis of the ***body language*** could be a solution. Thus, a "sensitive computer" could use the body movements and the positions of the body in order to assess the state of a person (e.g. confusion, illness, nervousness, lack of attention, motor fatigue, agitation, etc.).

In the rehabilitation process, the measurements of the motion impairments are very important because they can quantify the patient's recovering between two consecutive medical sessions. Nowadays, this type of motion analysis is achieved by physicians through visual observation of the patient during some standard tests. As a result, the physician subjectivism is introduced and, much more, when different physicians evaluate the same patient, the reproducibility of the measurements becomes a difficult task.

To respond to the previously presented requirements in different application fields, a non-contact laser system was introduced by the authors (Dobrea, 2002), (Cracan, Teodoru & Dobrea, 2005), (Dobrea, Cracan & Teodoru, 2005), (Dobrea and Serban, 2005). Here, we present the implementation of an independent system constructed as a self-contained unit that can be further integrated in much more complex and intelligent structures, together with new possible applications.

This article proposes a new real-time, non-contact system able to:

- acquire and interpret the subject's ***body language***,
- recognize static ***hand signs***, and
- provide physicians with a quantitative tool to monitor the evolution of the ***Parkinson disease.***

BACKGROUND

The proposed bio-instrumental system (BIS) was designed to be used in the medical field, in applications such as: rehabilitation, functional movement analysis, evaluation of the cognitive deficits or motion and support offered to the vocally impaired subjects.

Nowadays, in order to evaluate and assess the severity of the ***Parkinson disease***, the physicians use different rating scales. The method used to assess the ***Parkinson disease*** is based on a questionnaire - Unified Parkinson’s Disease Rating Scale, (MDSTF, 2003). The most important disadvantage of the rating scales is the lack of results reproducibility. Different physicians obtain different results on the same patient due to different medical experience and the possibility to observe, at one moment, only one cross-section of the patient. The BIS presented in this article will be used in the quantitative analysis of the head tremor movements. Even if, for this application, other methods exist to acquire the movement (based on accelerometer sensors, (Keijsers, Horstink & Gielen, 2003), optical data flow and gyroscope, (Mayagoitia, Nene & Veltink, 2002)) no method has imposed yet as a standard. Recognition of the ***hand signs*** is a challenging task for the nowadays systems and it is very important for the vocally impaired people. Even if the research in this field fade in time, the first large recognized device for identifying the ***hand signs*** was developed by Dr G. Grimes (1983) at AT&T Bell Labs. This device was created for “alpha-numeric” characters communication by examining hand positions like an alternative tool to keyboards; it was also proved to be effective as a tool for allowing non-vocal users to “finger-spell” words and phrases. In order to understand the ***hand signs*** language the hand gesture must be acquired. Mainly, the ***hand signs*** are acquired using video cameras (Cui and Wenig, 1999), (Ho, Yamada & Umetani , 2005) or some devices that directly determine the position of the hand parts (such as gloves) (Hernandez-Rebollar, Kyriakopoulos & Lindeman, 2004).

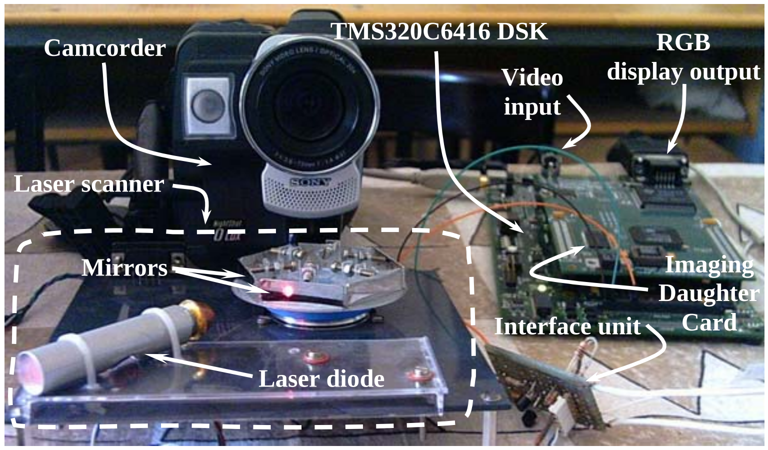
There are strong relations between psychological states and the body movements, confirmed by the theories of Kestenberg (1999) and Hunt (1968) or by the analyses realized in the field of the ***body language*** investigation (Pease, 1992). Moreover, these relationships make the subject of the somatic theory. The healthcare efficiency in the activity related to the human-computer interaction is directly dependent on both, the subject's state and the capability of the healthcare systems to recognize the specific needs of the user in order to change their response accordingly. Unfortunately, acquiring and interpreting this kind of information is very difficult and, as a consequence, all the actual systems have only a limited ability of communication. Current strategies for user's ***emotional state*** acquisition are either obtrusive (Picard et al. 2001) or the data captured by the systems consist in low level useful information.

A NEW TYPE OF NON-CONTACT BIO-INSTRUMENTAL SYSTEM

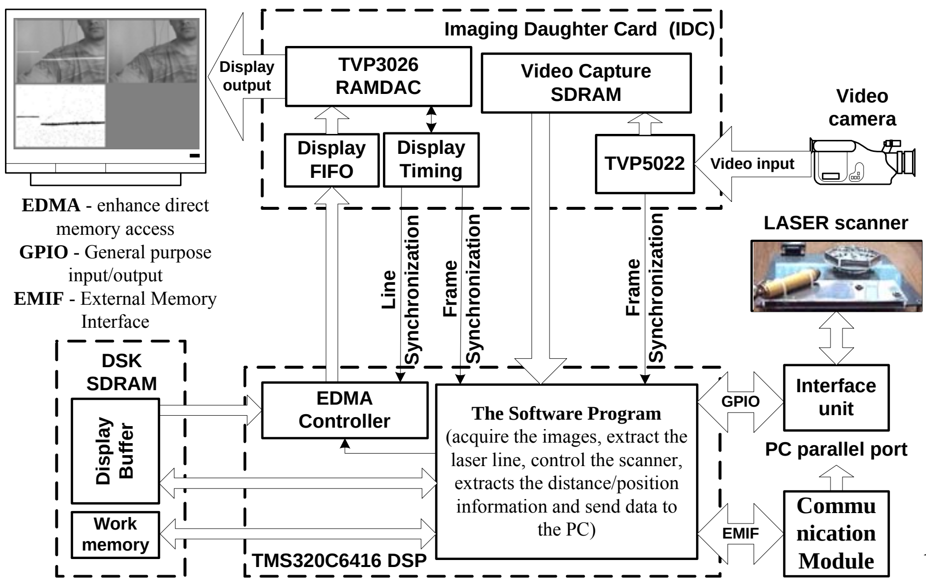
The new proposed BIS was designed to determine, in a fast way and without any physical contact with the subject, the movements, the position and the distance to an observation point. Using this information, the physiological and ***emotional states*** of the subject are estimated.

System's Architecture. Working Principle

The BIS is composed of a laser scanner, an interface unit, a video camera and a software program, running on a DSP platform that controls the scanner, acquires the images and extracts the distance/position information, as in Fig. 1a. The BIS schematics and the system data flow are presented in Fig. 1b.



a.



b.

Figure 1. The bio-instrumental system

a. View of the implementation

b. BIS schematics and the system data flow

The working principle of the whole system is based on a laser scanner that generates a laser plane at a constant angle from the horizontal plane. When the laser plane hits a target in the imaged area, a line of laser light appears on the body of the subject, see Fig. 2 – Img_{t+1} image. The video camera acquires two images: first, with the laser diode off, Img_t, and second, with the laser diode on, with a line of laser light that appears on the target, Img_{t+1}. Subtracting the two images we get only the laser line projected on the people's torso, *OutImg* – Fig. 2. In the ideal situation, all pixels for which $Img_{t+1}(x, y) \neq Img_t(x, y)$ describe the laser line which appears on the user's body torso. In real cases, the images are corrupted by noise. This problem was solved using an experimentally obtained noise model, σ. The criterion to extract the line of the laser light becomes now: $Img_{t+1}(x, y)–Img_t(x, y)>\sigma$. Other problems, such as shadows, slight body subject movements, light sources, video camera saturation, background changes, do not affect the reliability of the laser line feature extraction. This is happening because the time interval between the two images acquisition is less than 40 ms and the noise model presented above have been proven to be adequate. Based on this operating principle, the extraction of the laser line becomes a very fast task – a major advantage of this system.

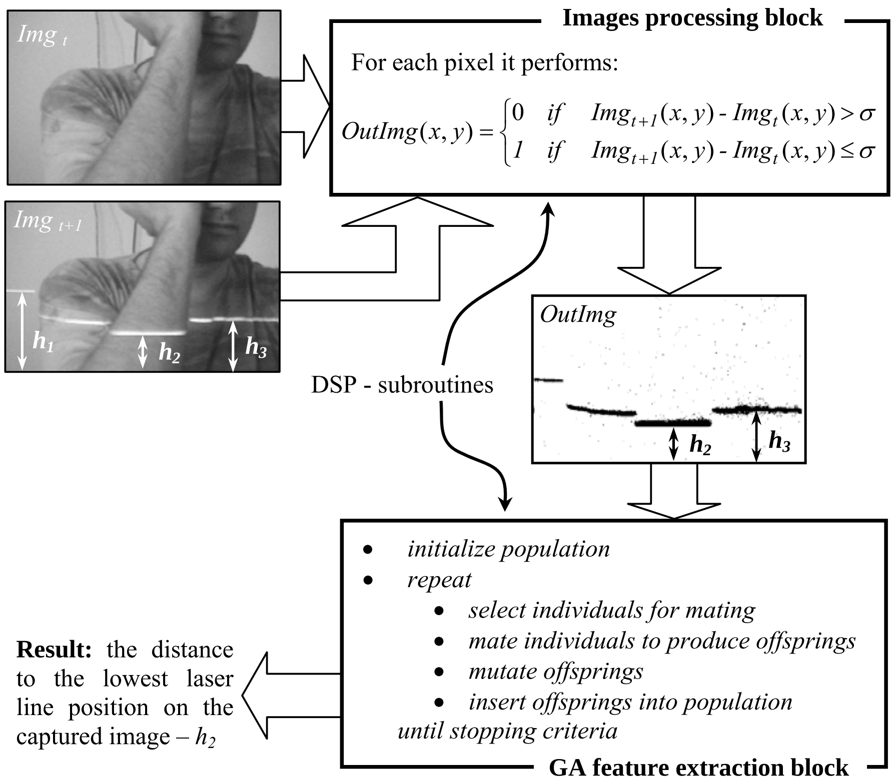


Figure 2. The data flow for the distance determination

If the object is far away, the extracted laser line will be farther from the bottom of the image, h_1. In the opposite situation, it will be closer to the bottom part of the resulting image, h_2. At this point, one knows the angle between the laser scanner and the horizontal plane, the position in space of the video camera and the extracted shape of the laser line on the subject body. The depth information of each point on the extracted laser line is calculated using some basic geometric formulae. Further on, having all these values, we exactly determine the real 3D subject body position with respect to the camera.

The hardware system has two components: the electro-mechanical scanner and the DSP system. The scanner has a low-power laser diode and a mechanical system with mirrors, Fig. 1a. (Dobrea, 2002). The plate with mirrors is attached to an engine shaft. The DSP system interfaces with the engine control system only through a single digital line that can start/stop the engine.

Since this application deals with images and all these type of applications are considered data and computing-intensive, the TMS320C6416 DSP was chosen due to its: high computing power, large on-chip memory and efficient data transfer mechanism.

In order to have a ***real time supervision*** of the system evolution, an output image, containing both the acquired image and the resulting one (*OutImg*), is formed and displayed on a RGB monitor (the image data are moved in background using for this the EDMA controller, Fig. 1b).

Movement-based subject state identification system

In order to test the BIS, we have developed an experiment intended to determine if there is a correlation between the ***emotional state*** of a person and the body torso movement of that person.

We admitted six subjects for this study. All of them were young healthy people (26.6±3 years, mean ± standard deviation) (Dobrea and Serban, 2005). But first, the ***emotional state*** must exist and must be manifested by the subjects. The emotion was induced by two films presented to the subjects: an action movie and a horror movie. At the end an analysis was done on the recorded body torso movements to characterize common behaviors of the subjects during the movies associated with special time moments of the films. In this way, the system was validated and analyzed.

The movement of the subject was characterized by the position of the subject torso, determined by means of the distance between the closest point of the chest situated on the laser line and the video camera. This distance is proportional to the distance from the lowest point of the extracted laser line (projected on the subject torso) to the bottom border of the image, h_3 on Fig. 2.

The distance determination

A special algorithm was designed in order to determine the distance between the subject and the laser diode. The algorithm use a standard genetic algorithm (GA), described by Goldberg (1989), Fig. 2. For each generation, an entirely new population is created by selecting individuals for mating from the previous population, according to a specified selection method. In order to implement the GA we have adopted a philosophy inspired by GALib, a C++ library for GA objects, (Wall, 2000).

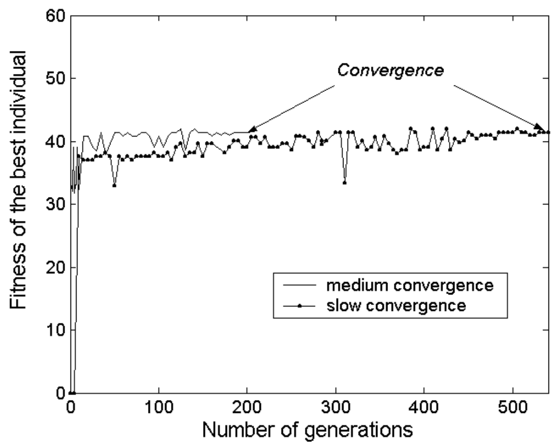
For our application, each chromosome has two genes, encoding the position of an image pixel (Dobrea, Sîrbu & Serban, 2004). We have implemented the binary string format for chromosomes, concatenating the binary representations of the coordinates, on the x and y axis. The fitness function was designed to maximize the number of image pixels in the vicinity of the selected image pixel and to minimize the distance between the pixel and the bottom of the screen. In this mode, the chromosome with the best fitness value will characterize a point belonging to the laser line segment which is closest to the bottom border of the image. To reduce the influence of noise, the pixels having in their vicinity less than a specified amount of adjacent pixels are ignored. The population was initialized randomly, uniformly distributed in the four quadrants of the image, to ensure the rapid convergence of the GA.

In Fig. 3a we present two evolutions of the GA for an extracted laser line, displaying the fitness of the best individual and stressing two behaviors: medium and slow algorithm convergence. For the tests we have run, the mean number of generation for convergence is 200, for a population of 100 individuals, with no elitism, 0.9 probability of crossover and 0.001 probability of mutation.

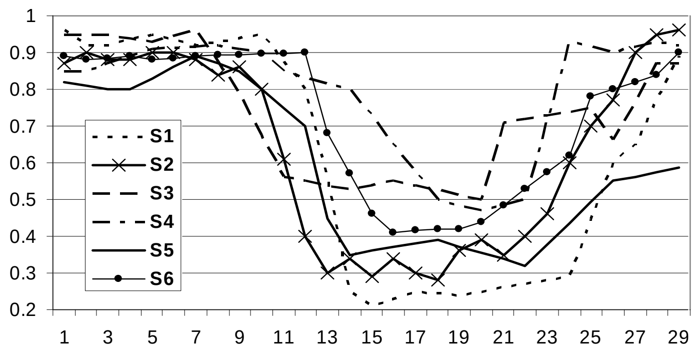
The emotional state detection

A problem this system must to deal with is given by the limbs movements that generate ***artifacts*** – e.g. in Fig. 2 the arm positioned in front of the body determines the GA to obtain the h_2 distance instead the correct distance h_3. These ***artifacts*** were removed using a special algorithm which takes into account the arm thickness.

The Pearson's Product-Moment ***Correlation coefficient*** was computed in order to characterize common behaviors of the subject's movements recorded during the movies and associated with special time events of the films. A time evolution of the distance between the view point (video camera) and the subject's chest position is presented in Fig. 3b.



a.



b.

	S1	S2	S3	S4	S5	S6
S1	**1**	**0.852**	**0.724**	**0.751**	**0.960**	**0.952**
S2	**0.852**	**1**	**0.871**	**0.781**	**0.807**	**0.865**
S3	**0.724**	**0.871**	**1**	**0.605**	**0.720**	**0.710**
S4	**0.751**	**0.781**	**0.605**	**1**	**0.553**	**0.864**
S5	**0.960**	**0.807**	**0.720**	**0.553**	**1**	**0.894**
S6	**0.952**	**0.865**	**0.710**	**0.864**	**0.894**	**1**

c.

Figure 3. Outcome of the algorithm which determines the distance between the subject and the laser diode

a. Two representative evolutions of the GA

b. The evolution of the torso position for the same segment of the movie for all the subjects

c. The correlation coefficient

The six subject's traces, representing distance evolutions, are presented and marked in this

figure with S1- S6. They correspond to a movie fragment able to impress the subjects.
The Pearson's ***correlation coefficients*** computed for all pairs of two time segments, shown in Fig. 3b, are presented in Fig. 3c.

The obtained results support and demonstrate the system's ability to evidence a common subject ***emotional state*** reflected by the body movements. The subjects' different behavior as a response to the same emotional state (through the movements of the body, hands, etc.) and the time delay required to manifest the emotional state determine the spread of the computed correlation values. For other time fragment of the movies, similar results were obtained.

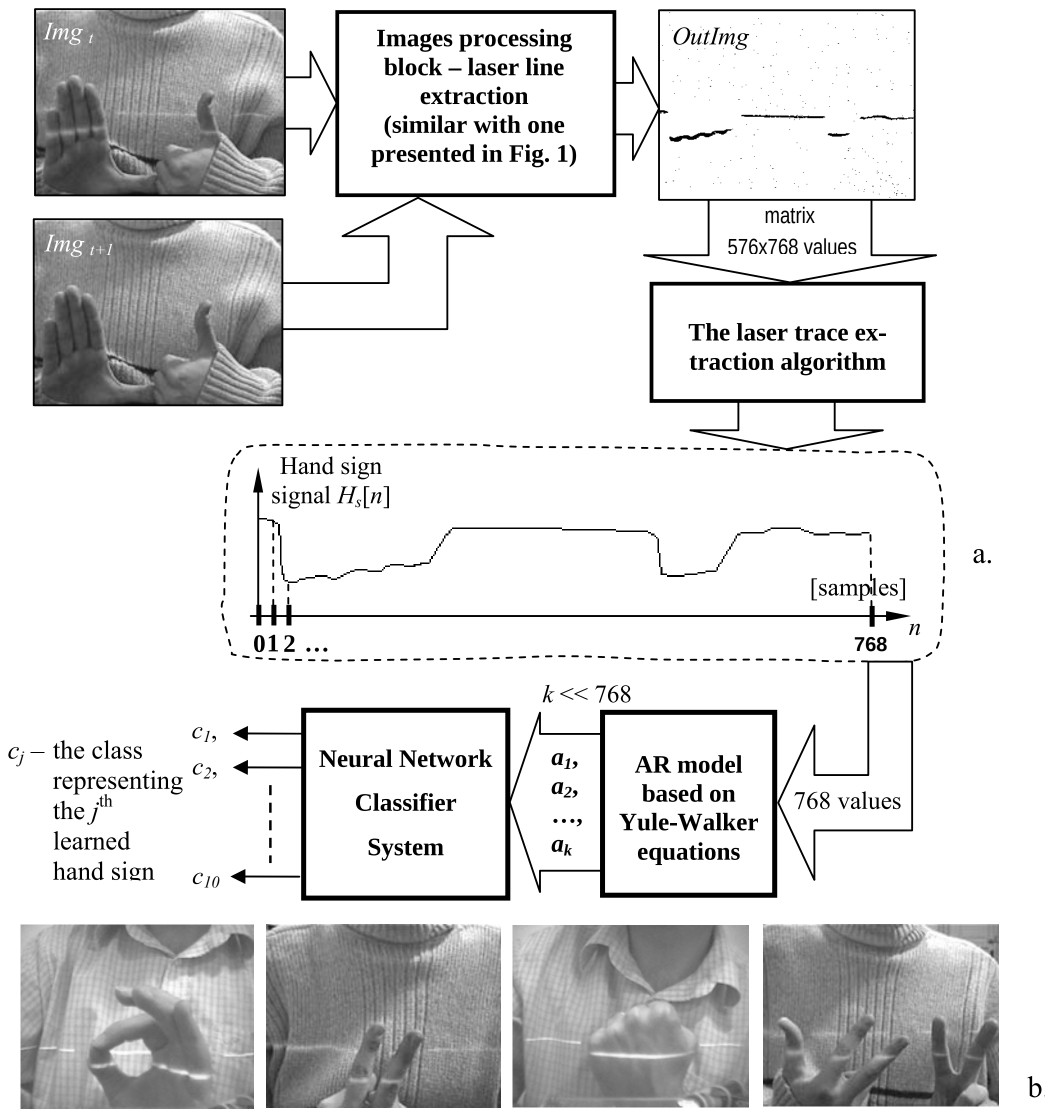


Figure 4. Demonstration of system's ability to evidence different hand signs

a. The system's flow chart and some of the partial results

b. Several hand signs recognized by the intelligent system

A hand sign recognition system

The system able to recognize the static ***hand sign*** we propose is a combination of two methods described in the literature: video and special device-based. The ***hand sign***s can be formed using one or both hands. Five of the ten used ***hand signs*** are presented in Fig 4a and b.

The algorithm used for the extractions of the projected laser line image is similar to the one presented above. The laser trace signal ($H_s\,[n]$ resulted from the laser extraction algorithm) is modeled using the coefficients of an auto-regressive (AR) filter, (Cracan et al., 2005), (Dobrea et al. 2005). The AR filter's coefficients are used to reduce the redundant input information passed to the classifier algorithm implemented on DSP.

A multilayer perceptron (MLP) neural network was used in the ***pattern recognition*** process (Cracan et al., 2005), (Dobrea et al. 2005). The correct recognition rates for all the ***hand signs*** were in the range of 0.823÷1. The necessary time between the first image acquisition and the end of the entire classification process was less than 1.5 seconds, adequate for ***real time supervision.***

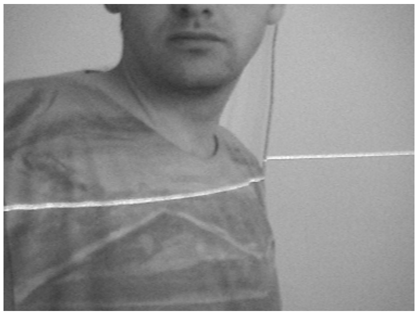
FUTURE TRENDS

From the subject's body language to emotional state identification

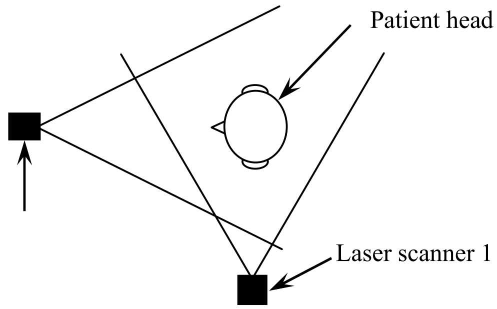
The identification of some particular postures like the arm position in front of the torso, as in Fig. 2, or the torso position, as in Fig. 5a, can be made by ***pattern recognition***, in the manner presented above. Each of these postures or body positions can be related to different internal subject states, (Pease, 1992), that can guide a system in order to improve the human computer interaction. For example, in Fig. 2, the subject posture can express boring if the subject keeps this posture for a long time.

Evaluation/analysis of Parkinson patients

Up to this moment there is no kind of standard method (either qualitative or quantitative) to evaluate the Parkinson symptoms. Moreover, in (MDSTF, 2003) one mentions a number of errors in the Unified Parkinson's Disease Rating Scale such as: some ambiguity in the text, inadequate instructions to rate some questioner rubrics, one deficiency in a unit of measurement and the lack of questions other than motor aspects of Parkinson’s disease.



a.



b.

Figure 5. Identification of some particular postures or movements

a. Image representing a particular body posture

b. The configuration of the laser systems in order to acquire the head position

Using two different laser scanner systems, the trajectory of the subject head (as in Fig. 5b) can be recorded and easily quantified in order to assess the patient rehabilitation. In this mode, the proposed system is able to quantitatively evaluate the severity and the progress of the Parkinson's disease and to offer a reproducibility of the obtained results. Thus, all the above presented drawbacks are eliminated.

CONCLUSIONS

In this article a DSP implementation of a new non-invasive BIS was presented. This project has a significant impact on the people’s life reflected in:

- the natural form of subject’s interaction and supervision by the healthcare systems, in order to determine the emotional and physiological state changing;
- the reproducibility of the evaluation and assessment of the severity in ***Parkinson disease*** – a way of helping the physicians to improve the quality of the medical act;
- the support offered to the vocally impaired subjects.

This system offers the possibility to use a new kind of information regarding the subject's emotional and physiological state, unexploited yet on the HCI systems, namely, the state of the subject expressed through ***body language***.

The system is inexpensive, easy to manufacture and, hence, attractive for practical applications. In the end, last but not least, the entire system is very fast, being adequate even for ***real time supervision***.

REFERENCES

Cracan, A., Teodoru, C. & Dobrea, D.M. (2005). Techniques to implement an embedded laser sensor for pattern recognition, *Proceedings of the International Conference on "Computer as a tool"*, Belgrade, Serbia & Montenegro, 21-24 November, 2, 417-1420

Cui Y. & Wenig J. (1999). A learning-based prediction-and-verification segmentation scheme for hand sign image sequence, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21, 798-804.

Dobrea, D.M. (2002). A New Type of Sensor to Monitor the Body Torso Movements Without Physical Contact, *Proceedings of Second European Medical and Biological Engineering Conference*. December 4-8, Vienna, Austria, 3, 810–811.

Dobrea, D.M., Sîrbu, A. & Serban, M.C. (2004). DSP Implementation of a New Type of Bioinstrumental Noncontact Sensor, *CD Proceedings of 4th European Symposium in Biomedical Engineering*, Patras, Greece, 25th-27th June

Dobrea, D.M., Cracan, A. & Teodoru, C. (2005). A Pattern Recognition System for a New Laser Sensor, *Proceedings of the 3rd European Medical and Biological Engineering Conference*, vol. 11, November 20-25, Prague, Czech Republic, 3011-3014

Dobrea, D.M. & Serban, M.C. (2005). From the movement to emotional state identification, *Proceedings of the 14th International Conference of Medical Physics*, Nuremberg, Germany, September 14–17, 776-777

Goldberg, D. (1989). *Genetic Algorithms in Search Optimization and Machine Learning*, Addison-Wesley.

Grimes, G.J. (1983). *Digital data entry glove interface device*. Bell Telephone Lab. Inc., Patent No: 4,414,537, United States, November 8.

Hernandez-Rebollar J.L., Kyriakopoulos, N. & Lindeman, R.W. (2004). A New Instrumented Approach for Translating American Sign Language into Sound and Text. *Proceedings of Sixth IEEE International Conference on Automatic Face and Gesture Recognition*, Seoul, Korea, 547 – 552.

Ho M.A.T., Yamada Y. & Umetani Y. (2005). An adaptive visual attentive tracker for human communicational behaviors using HMM-based TD learning with new State distinction capability. *IEEE Transactions on Robotics*, 3, 497 – 504.

Hunt, V. (1968). *The Biological Organization of Man to Move*. Impulse.

Keijsers, N.L.W., Horstink, M.W.I.M. & Gielen, S.C.A.M (2003). Online Monitoring of Dyskinesia in Patients with Parkinson's Disease. *IEEE Engineering in Medicine and Biology*, 22(3), 96-103.

Kestenberg-Amighi, J, Loman, S., Lewis, P. & Sossin, S. (1999). *The meaning of movement*, Gordon & Breach Publishers, 1999

Mayagoitia, R.E., Nene, A.V. & Veltink, P.H. (2002). Accelerometer and rate gyroscope measurement of kinematics: An inexpensive alternative to optical motion analysis systems. *Journal of Biomechanics*, 35(4), 537-542.

Movement Disorder Society Task Force on Rating Scales for Parkinson's Disease (2003). The Unified Parkinson's Disease Rating Scale (UPDRS): status and recommendations. Movement Disorders. 18(7), 738-750

Pease, A. (1992). *Body Language – How to read other's thoughts by their gesture*, Sheldon Press, London

Picard, R. W., Vyzas, E. & Healey, J. (2001). Toward Machine Emotional Intelligence: Analysis of Affective Physiological State, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(10), 1175-1191.

Wall, M. (n.d.). GALib from http://lancet.mit.edu/ga

TERMS AND DEFINITIONS

Human computer interface (HCI) – The process by which users interact with computers; based on it, one designs and implements human-centric interactive computer systems.

Digital Signal Processor (DSP) - A specialized, programmable computer processing unit that is able to perform high-speed mathematical processing. It refers to manipulating analogue information that has been converted into a digital (numerical) form.

Genetic algorithm (GA) - An optimization and search technique based on the principles of Darwin's theory of natural selection and Mendel's work in genetics on inheritance: the stronger individuals are likely to survive in a competing environment. It allows a population composed of many individuals (possible solutions) to evolve under specified selection rules to a state that maximizes their suitability for the specific application.

Fitness - A measure of the suitability of a potential solution for the given application. Each individual is an encoded representation of all the parameters that characterize the solution. It has an associated value (fitness) which is a measure of its performances.

Neural network - A system of programs and data structures that approximates the operation of the human brain.

Sensor - A device that measures or detects a real-world condition, such as an acoustic (microphone, hydrophone), electromagnetic (radar) or optical (camera) signal.

Biosensor – a device incorporating a biological sensing element (the recognition element is biological in nature).