
Andrea Caroppo, Alessandro Leone, Pietro Siciliano
Institute for Microelectronics and Microsystems, Italian National Research Council (CNR), Via Monteroni, c/o Campus Università del Salento, Palazzina A3, Lecce, Italy

Abstract—Cognitive Rehabilitation (CR) is a relatively new approach to improve well-being for people with Alzheimer’s disease (AD). At present only preliminary evidence regarding efficacy is available but it is enough to suggest that this kind of rehabilitation has the potential to bring about changes in behavior, enhance well-being and maintain involvement in daily life. The present work presents a digital platform integrating Natural User Interface (NUI) for motor and cognitive rehabilitation of patients with different disease condition. It is made up of an embedded PC connected to a TV monitor with internet connection, a low-cost 3D sensor (we use Microsoft Kinect in order to allow wide diffusion of the proposed solution), and an optional e-shirt with textile electrodes for clinical signs detection. The main contribution of this work is the design and implementation of an information and communication technologies (ICT) platform, through a customized Virtual Personal Trainer (VPT) allowing the patients to perform the rehabilitation practice at home. Moreover, the system provides an audio/visual link with the medical center, so the physician can interact with the patient during the rehabilitation practice, increasing the compliance and the efficacy and making sure that the type and intensity of treatment are appropriate. Customized algorithms for calibration, people segmentation, body skeletonization and hands tracking through the Kinect sensor have been implemented in order to infer knowledge about the reaction of the end-user to the Graphical User Interface (GUI) designed for specific cognitive domains. For proper interaction, gestures of AD patients are acquired by the sensor in the nominal functioning range, allowing 100% hands detection rate, useful for an error free human-machine.

Index Terms—Alzheimer’s disease, Cognitive Rehabilitation, ICT Platform, Natural User Interface.

I. INTRODUCTION

In recent decades there has been significant progress in using ICT in the field of healthcare and more specifically in the case of both physical and CR practice. In this context, the development of a cost-effective home-care service seems to be very attractive, so a great effort has been addressed by the scientific community in order to provide specific enabling solutions. AD is a chronic neuro-degenerative disease (dementia) in which the first symptom is a slowly increasing memory loss. As the disease progresses, the brain deteriorates more rapidly with apparent cognitive limits. The importance of CR in the treatment of patients with dementia is underlined by recent scientific publications [1]-[2]-[3]-[4]. The development of a low-cost platform/home-care service with CR functionalities could be very useful in order to increase the chances of an appropriate medical therapy. Some preliminary studies show that ICT tools are well accepted by elderly people, although education and ICT skills level is often low. Moreover, it is scientifically proven that ICT technologies improving quality of life and increasing the permanence at home. Over the past years, new technologies have been developed, in the field of healthcare, for cognitive training and rehabilitation [5]-[6]-[7]. For example, virtual reality offers training environments in which human cognitive and functional performance can be accurately assessed and rehabilitated [8]-[9]. On the other hand, augmented reality provides safer and more intuitive interaction techniques allowing interaction with 3D objects in real world [10]-[11]. In this scenario social communication channels (natural speech, para-language, etc.) are not blocked, breaking down mental barriers applying such a technology to specific problems or disabilities. New solutions for cognitive assistance based on touch system have been implemented: in the field of CR commercial products (Nintendo’s Brain Age, Big Brain Academy, etc.) have been tuned as educational tools helping to slow the decline of AD [12]-[13]-[14]. More recently, the large diffusion of interaction devices enabling body movements to control systems have been investigated, with specific focus on ICT technologies for natural interaction. Microsoft Kinect is the state-of-the-art [15] as 3D device for body movements acquisition and gesture recognition and the effects of this kind of technology for rehabilitation purposes is widely investigated [16]-[17]. In this work, a NUI platform has been designed with the aim to support different kind of patients during the multi-domain rehabilitation practice without the presence of medical staff or caregiver. As AD patient may have troubles moving specific body parts, a new hands tracking filter has been implemented with the aim to overcome the well-known limitations of the royalties-free NUI middleware architecture used in the platform. The remainder of the paper is organized as follows. Section II presents some details of the platform with specific focus on the rehabilitation practice; moreover, a detailed description of a novel filter to improve human hand tracking for a better interaction with GUI is proposed. Section III presents the experimental results to show the performance of
II. MATERIAL AND METHODS

A. Platform Overview

The developed ICT platform provides a system for CR through a customized Virtual Personal Trainer (VPT) allowing the patients to perform the rehabilitation practice at home in an autonomous way. From the hardware point of view the platform is made up of a set-top-box connected to a TV monitor with Internet connection, a Microsoft Kinect RGB-D sensor for human body tracking and gesture recognition and a WWS system with textile electrodes for clinical signs monitoring (Fig. 1). The set-top-box (a commercial embedded pc) is the gateway able to automatically downloads sequences of exercises from a remote server through the Internet connection. The system provides a web-based platform that allows the physician to customize directly the therapy: this process is a highly innovative compared to existing systems [18]-[19] as the caregiver/physician defines a specific sequence of exercises (the therapeutic session) according to the residual abilities of the patient.

Moreover, the platform integrates streaming functionalities allowing visual/audio recording for post-verification or online feedback to the physician which is able to follow the execution of the exercises from a remote architecture. From this perspective, the physician/psychologist of the reference center could communicate to the patients through a remote connection and then monitor the progress or trouble in the execution of the different required tasks. At the end of the rehabilitation session, the central platform collects different kinds of data locally stored on the set-top-box. An ad-hoc multi-modal messaging procedure (e-mail, SMS, Mobile App for Android devices, etc.) is performed and relevant data are sent to the physician allowing instant verification of the performance through an easy-to-use GUI.

B. Multi-sensor devices as enabling technology

In order to approach the CR practice, the end-user needs a specific hardware device (Microsoft Kinect) which allows interacting intuitively with a GUI using their bodies (Fig. 2.a). The optical “eye” of the sensor detects the body parts according to the structured-light working paradigm. From the functioning principle point of view, the Kinect device is a RGB-D camera integrating both a high resolution RGB camera and an infrared depth sensor, able to output video at a frame rate of 9 Hz to 30 Hz depending on resolution.
remaining within an acceptable range for people body part detection. Another important feature of the platform is the continuous monitoring of physiological parameters during the execution of the therapy, in order to evaluate psycho-emotive stress of the patient involved in the CR practice. Physiological parameters are collected by a Wearable Wellness System (WWS) produced by Smartex [20]. The system includes a sensorized garment (Fig. 2b) and an electronic device (Fig. 2c). The sensorized garment is equipped with two textile electrodes directly connected to the device (named SEW). A jack connector links the sensorized garment with the electronic device. The e-shirt is available in both male and female version, with size ranging from S to XL. WWS is able to simultaneously acquire ECG, heart rate, breath rate and acceleration values along x-axis, y-axis and z-axis. Data acquired are sampled with different rates (breath-rate@0.2Hz, heart-rate@0.2Hz, 1-derived ECG channel@ 250Hz) and transmitted to the set-top-box via Bluetooth radio link.

C. Environment setup and rehabilitation practice details

The setup of the environment can be done without any specific help, but it is important to observe a few simple requirements for proper CR practice execution.

- Devices Positioning

For the best performance during the practice, Kinect must be placed allowing acquiring the whole body (see Fig. 3a). Some tips on how to place the sensor are listed in the following:

- sensor must be place near the edge of a flat, stable surface;
- sensor should be within 15 cm above or below a TV monitor, and between 0.6 meters and 1.8 meters from the floor;
- Avoid positioning the sensor in direct sunlight or within 0.3 meters of audio speakers.

- End-User Position

The end-user must be far from Kinect sensor at least 1.5 meters and never more than 3 meters (see Fig. 3b), assuring the proper functioning of 3D skeletonization procedure provided by Microsoft Kinect SDK. Some categories of exercises can be executed from a seated position; however the user must always respect the specific operative range of the platform.

- Interaction Procedure

In order to interact with the VPT, the patient must move the hand minimizing occlusion with other body parts. Hand tracking algorithms (as described in a subsequent section) have been implemented providing a customized level of movement sensitivity which is manually tuned by the physician (three different level of sensitivity have been implemented). For specific exercise (e.g. Personnel Guidance) gesture recognition algorithms have been developed in order to verify the correctness of the hand movement with respect to a previously recorded template.

D. Multi-Domain cognitive rehabilitation practice

The CR program is composed by sequences of exercises appropriately tuned by the physician or psychologist. Each exercise belongs to a category, bringing out specific cognitive activities according to guidelines of the state-of-the-art international evaluation scales for AD (e.g. Mini Mental State Examination [21]). An innovative feature of the platform deals with the opportunity to customize each exercise on the basis of the severity of cognitive impairment and residual skills of the target. For this purpose, during the setting procedure, few input parameters need to be defined a-priori (e.g. execution time, maximum numbers of allowed errors, movement sensitivity). From the taxonomic point of view, the following categories of exercises have been implemented: temporal orientation, personnel guidance, topographical memory, visual memory, hearing attention, visual attention, categorization and verbal fluency. Figure 4 shows the GUI of some CR exercises. The design of the therapy can be remotely performed, thanks to a web application that allows physician to configure all the exercises based on the patient's residual abilities and related performance. As a result, the interface of a specific exercise can be different for each patient. For example, the exercise “Topographical Memory” requires the following parameters: number of rows in the grid, number of columns in the grid, number of red dots (correct answers), whereas the exercise “Categorization” requires only the number of images to display. In addition to the GUI modelling, it is possible to establish the maximum length (in time unit) of every exercise and to set (only for a certain categories of exercises) the number of aids. The specific input required and the information about the presence or not of aids are reported in the next table:
In order to allow an appropriate display of every exercise and become independent from the specific output device (digital monitor, HD TV …) a software module for the best video rendering is implemented. The definition of graphics objects displayed on the GUI has been designed according to the principles of ergonomics, usability and acceptability as referred in ISO/IEC 2001a [22].

E. Improvement of hand gestures recognition for proper human computer interaction

As AD patient may have troubles moving the own hands, the procedure for hand tracking and gesture recognition provided by the Microsoft Kinect SDK [23] may be affected by critical issues, for example when hands and body torso are overlapped. In particular, AD patient may not be able to move hands in a spatially extensive environment, so that the interaction with objects belonging to the GUI could be hard and the CR practice could be dramatically affected. In this context graphical objects are codified as “hidden buttons” able to discover the hovering time for specific end-user choices. The skeletonization procedure provided by Microsoft Kinect SDK can be affected due to noise in the acquisition process so that the joint positions estimations can fail and the interaction with the GUI could be sometimes hard to handle. In order to overcome this kind of issue, a noise reduction filter has been designed removing as much as possible noise from raw data. The suggested filter operates as a smoothing filter that overcomes the performances of the Holt Double Exponential Smoothing Filter (HDESF) [24] built in SDK. The HDESF procedure reduces the jitters from skeletal joint data providing a smoothing effect with lower latency than other smoothing filter algorithms. The main issue with HDESF application is that Kinect sensor does not have sufficient resolution to ensure consistent accuracy of the tracked joints over time. Observing real data, the problem is more evident when different joints are overlapped. In this context, an improved smoothing algorithm has been designed. From the analysis of the state of the art, the Exponential Weighted Moving Average Filter (EWMAF) [25] appears as the best trade-off between smoothing effect and jitter control. The EWMA equals the present predicted value plus α times the present observed error of prediction. Thus,

\[
\hat{x}_{t+1} = \hat{x}_t + \alpha e_t
\]

and

\[
e_t = \hat{x}_t + \alpha(x_t - \hat{x}_t)
\]

where \(\hat{x}_{t+1}\) represents the predicted value at time \(t+1\) (the new EWMA), \(x_t\) is the observed value at time \(t\), \(\hat{x}_t\) is the predicted value at time \(t\) and \(e_t\) represents the observed error at time \(t\). Equation (2) can be written as:

\[
\hat{x}_{t+1} = \alpha x_t + (1 - \alpha)\hat{x}_t
\]

The coefficient \(\alpha\) is a constant smoothing factor between 0 and 1 and it represents the decreasing weighting degree of the filter. Values of \(\alpha\) close to 1 represents a lower smoothing effect, whereas values of \(\alpha\) closer to 0 make a greater smoothing effect.

III. RESULTS

The platform has been tested by 40 subjects (see Table II for details). The inclusion criteria were: 1) age ≥ 65 years, 2) diagnosis of AD mild cognitive impairment according to the criteria of the National Institute of Neurological and Communicative Disorders and Stroke-Alzheimer's Disease and Related Disorders Association Work Group (NINCDS-ADRDA) [26], and to the Diagnostic and Statistical Manual of Mental Disorders-Fourth Edition issues (DSM-IV) [27]. The exclusion criterion was the inability to acquire data due to the poor clinical condition of the patient. For each patient the cognitive status, which considers the psychological aspects that drive the behavior of individuals on the basis of personal resources, both emotionally and

TABLE I: Specific input required for each category of exercise (column in the middle); information about the presence or not of aids during the execution of the specific exercise (last column)

<table>
<thead>
<tr>
<th>Category</th>
<th>Specific Input</th>
<th>Help?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temporal Orientation</td>
<td>Year, month, day, day of the week</td>
<td>Yes</td>
</tr>
<tr>
<td>Personnel Guidance</td>
<td>Body part target</td>
<td>Yes</td>
</tr>
<tr>
<td>Topographical Memory</td>
<td>N° of answers, N° grid rows, N° grid columns</td>
<td>Yes</td>
</tr>
<tr>
<td>Visual Memory</td>
<td>N° of images</td>
<td>Yes</td>
</tr>
<tr>
<td>Hearing Attention</td>
<td>Text/Story, N° of words to identify</td>
<td>No</td>
</tr>
<tr>
<td>Visual Attention</td>
<td>N° of answers, N° grid rows, N° grid columns</td>
<td>No</td>
</tr>
<tr>
<td>Categorization</td>
<td>N° of images</td>
<td>Yes</td>
</tr>
<tr>
<td>Verbal Fluency</td>
<td>Target word, synonym/contrary word</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Fig 4 Examples of GUI for different categories of CR exercises
mentally, was assessed. To explore this aspect Mini Mental State Examination (MMSE) was used. It allows an evaluation of attention, space-time orientation, short and long term memory, calculation, executive ability, writing, and the appropriate use of language.

Since ground truth data for real-world image sequences is hard to obtain, according to the experimental section proposed in [28], the evaluation of hand tracking procedure was made analyzing the hand trajectory on different exercises performed by end-users. The analysis of the trajectories highlights the problems of “jittering” of the signal: in some critical situations (like, for example, the selection of a graphic element that is very close to another graphic element) HDESF could return unwanted choices/selections which doesn’t reflect the real intent of the end-user. The aforementioned issue may be negligible in gaming context, whereas in the considered CR scenario the consequences are apparent affecting the CR practice. On the other hand, the application of EWMAF allows to obtain a “smoothed” trajectory, which turns out to be more stable both in space and in time thanks to the filter application. A quantitative measurement of the performances can be carried out analyzing the difference between values calculated by a model (e.g. an ideal trajectory over specific graphics elements placed on the interfaces) and the tracked values recorded in the two previous cases (trajectory obtained with the application of HDESF and EWMAF). For this purpose, a frequently used measure is the Root-Mean-Square Deviation (RMSD) that represents the sample standard deviation of the difference between predicted and observed values.

### TABLE II: Gender and age distribution of trial participants, with specific focus on MMSE score

<table>
<thead>
<tr>
<th>Sex (%)</th>
<th>65y ≤ Age ≤ 70y</th>
<th>71y ≤ Age ≤ 75y</th>
<th>Age &gt; 75y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>10 (25%)</td>
<td>5 (12.5%)</td>
<td>3 (7.5%)</td>
</tr>
<tr>
<td>Female</td>
<td>13 (32.5%)</td>
<td>7 (17.5%)</td>
<td>2 (5%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MMSE (%)</th>
<th>21 ≤ score ≤ 26 (Mild Alzheimer)</th>
<th>7 (17.5%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>20 (50%)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>13 (32.5%)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MMSE (%)</th>
<th>10 ≤ score ≤ 21 (Moderate Alzheimer)</th>
<th>2 (5%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>7 (17.5%)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>5 (12.5%)</td>
<td></td>
</tr>
</tbody>
</table>

Table III shows the results obtained at seven different distances, selected in the operating range of Microsoft Kinect; the sample number used to estimate RMSD is approximately 500. RMSD values for EWMAF demonstrate the validity of the proposed solution: the average deviation never exceeds the value of 30 pixel, allowing greater precision in the selection of a graphic element. Instead, HDESF performs an average error always greater than 58 pixels, so that the gap in some cases affects the correct selection of a graphic element belonging to the GUI. EWMAF presents the issue of introducing a lag relative to the input data. In real time application latency is a critical factor, therefore it is essential to tune up in the right way the parameters of the filter, with the aim of obtaining the best performance with the lowest latency.

### TABLE III: RMSD values at different distance for HDESF and EWMAF on 500 samples

<table>
<thead>
<tr>
<th>Distance from Kinect</th>
<th>RMSD (HDESF, in pixel)</th>
<th>RMSD (EWMAF, in pixel)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,5 m.</td>
<td>58.76</td>
<td>21.95</td>
</tr>
<tr>
<td>1,8 m.</td>
<td>60.30</td>
<td>22.87</td>
</tr>
<tr>
<td>2 m.</td>
<td>60.95</td>
<td>23.11</td>
</tr>
<tr>
<td>2,4 m.</td>
<td>61.06</td>
<td>23.36</td>
</tr>
<tr>
<td>2,8 m.</td>
<td>68.12</td>
<td>25.67</td>
</tr>
</tbody>
</table>

### TABLE IV: RMSD values by evaluating EWMAF for different \(\alpha\) smoothing factor and amount of observations

<table>
<thead>
<tr>
<th># of observations</th>
<th>(\alpha = 0.1)</th>
<th>(\alpha = 0.3)</th>
<th>(\alpha = 0.5)</th>
<th>(\alpha = 0.7)</th>
<th>(\alpha = 0.9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>34.321</td>
<td>29.125</td>
<td>27.512</td>
<td>25.342</td>
<td>23.004</td>
</tr>
<tr>
<td>10</td>
<td>33.566</td>
<td>28.908</td>
<td>26.823</td>
<td>24.998</td>
<td>22.675</td>
</tr>
<tr>
<td>30</td>
<td>31.344</td>
<td>27.567</td>
<td>25.889</td>
<td>23.841</td>
<td>21.675</td>
</tr>
</tbody>
</table>

As mentioned in a previous section values of \(\alpha\) close to 1 presents a lower smoothing effect, since greater weight to recent changes in data is considered. A choice of \(\alpha = 0.9\) allows EWMAF to reduce significantly the RMSD compared to the result obtained with \(\alpha = 0.1\) and \(\alpha = 0.3\) (see Table IV). On the other hand, when the number of past observations grows in size results are better but a latency effect is introduced in the tracking procedure. For all considered scenarios, the best tradeoff is achieved for \(\alpha = 0.9\) and the amount of past observations equal to 10 (as 300 ms).

### IV. CONCLUSIONS

The major contribution of this work was to design and evaluate a platform for the execution of motor and cognitive rehabilitation of patients directly at home through an ICT platform integrating low-cost contact-less UBI optical device. As a patient with Alzheimer's disorder may have trouble moving the own hands, a new filter for hand tracking has been implemented, overcoming the performances of the built-in filter of Microsoft Kinect SDK. The improvement allows an easy and accurate interaction of the end-user and the platform, even in the presence of complex ad-hoc designed GUI. The proposed platform allows both the evaluation of the progress of the dementia (useful for the caregiver) and the cognitive stimulation of the end-user in several domains. Future works are addressed to deploy the platform to a wide class of...
Alzheimer’s disease patients in order to have the feedback (validation) during the technological tool usage. This will allow to tune up each component of the platform (GUI, knowledge discovery logic, filter parameters) in order to make it highly compliant with the needs and the requirements of the patients, according to the recent User Centered Design paradigm.

ACKNOWLEDGMENT

This work has been carried out within AL.TR.U.I.S.M. Project, funded by Apulia Region under the Health-care call for the promotion of regional partnerships for innovation. Authors thank the staff of the Geriatric Unit of the Hospital in San Giovanni Rotondo (Italy) for the validation of the platform.

REFERENCES

AUTHOR'S PROFILE

Andrea Caroppo received the Degree in Computer Science Engineering in 2004 from the University of Lecce. From 2004 to 2006 he was researcher fellow at the Italian National Research Council (CNR), Institute of Study of Intelligent Systems for Automation (ISSIA) in Bari (Italy). His research interests were in the area of image and video processing/coding, neural networks, motion estimation in video sequences and multidimensional signal processing. Since 2012 he is researcher fellow at the Italian National Research Council (CNR), Institute for Microelectronics and Microsystems (IMM) in Lecce (Italy). He is interested in signal and image processing, pattern recognition, computer vision and development of enabling technologies for healthcare with particular focus on the new Ambient Assisted Living (AAL) technologies. He is author of more than 20 papers in national and international journals and conference proceedings.

Alessandro Leone received the Degree in Computer Science Engineering in 2003 from the University of Lecce. Since 2003 he is researcher at the Italian National Research Council (CNR), Institute for Microelectronics and Microsystems (IMM) in Lecce (Italy). He is interested in Signal and Image Processing, Pattern Recognition, Computer Vision and Smart Multi-sensorial Systems with particular focus on the new Ambient Assisted Living technologies. Eng. Leone is the technical coordinator of the Signal & Image Processing Laboratory and he is mainly involved in development of enabling technologies for healthcare: fall detection & prevention, neurodegenerative cognitive rehabilitation, interoperability platforms and smart wearable sensors for vital signs monitoring. He is author of more than 60 papers in national and international journals and conference proceedings. Dr. Leone has been assistant lecture for the class master “Image Processing” at the University of Salento.

Pietro Siciliano Dr. Pietro Siciliano received his degree in physics in 1985 from the University of Lecce. He took the Ph.D. in Physics in 1989 at the University of Bari. He is currently a Director of Research at the Institute for Microelectronics and Microsystems (IMM-CNR) at CNR in Lecce, where he has been working from many years in the field of Sensors, MEMS, and Microsystems, being in charge of the Sensors and Microsystems group. He is author of about 350 scientific papers. Dr. Siciliano is referee and member of the advisory board of international journals. He has been responsible for several national (FISR, FIRB, PON) and international (V, VI and VII EU Framework) projects at IMM-CNR. He is member of the Steering Committee of AISEM, the Italian Association on Sensors and Microsystems. He has been chairman and member of the organizing committee of international Conferences and School in Sensors and Microsystems area. He is Director of the Section in Lecce of IMM-CNR. He is President of the Italian Association on “Ambient Assisted Living” (AiAAL), Responsible of “INNOVAAL”, the Public-Private Partnership on Active & Assisted Living and President of the National Technological Cluster for “Smart Living Technologies”.

DOI:10.17605/OSF.IO/8GJTV