Low-Cost Motion-Tracking for Computational Psychometrics Based on Virtual Reality

Pietro Cipresso[1(✉)], Silvia Serino[1], Irene Alice Chicchi Giglioli[1], Igor Giuliano[2], Davide Borra[3], Andrea Farina[4], and Giuseppe Riva[1,5]

[1]Applied Technology for Neuro-Psychology Lab, IRCCS Istituto Auxologico Italiano, Via L. Ariosto 13, 20145 Milano, MI, Italy
p.cipresso@auxologico.it
[2]Regola, Corso Turati 15/H, 10128 Torino, TO, Italy
[3]NoReal.it, Via Ugo Foscolo 4, 10126 Torino, TO, Italy
[4]Partner and Partners, Corso Brescia 9, 10152 Torino, TO, Italy
[5]Department of Psychology, Università Cattolica del Sacro Cuore, L.go Gemelli 1, Milano, MI, Italy

Abstract. Virtual Reality (VR) is a computer-based simulation designed to expose users to environments in order to replicate real world objects and events. In this framework, video games are one of the most popular forms of VR media all over the worlds. Their popularity has been fuelled by advancements in gaming technology and interactive devices at a low cost in home gaming market but also in clinical and research settings. In clinical and research virtual rehabilitation, the user should be able to interact (directly or indirectly) with the environment via a wide array of input technologies. These include activation of computer keyboard keys, a mouse or a joypad (indirect) and even by using special sensors or visual tracking (direct). For example, Microsoft Kinect provides low-cost motion tracking sensors, allowing to clinicians to interact with rehabilitation applications in the most natural and flexible way. This flexibility can be employed to tailor the user interaction to the specific rehabilitation user aims. According to this perspective, the paper aims to present a potential new platform, NeuroVirtual3D, which intends to develop a software interface for supporting assessment and rehabilitation of cognition function through several input/output devices, such as data gloves, joypad and Microsoft Kinect.

Keywords: Psychometrics · Virtual Reality · Biosensors · Psychophysiology

1 Introduction

Virtual Reality (VR) usually concerns the application of interactive simulations produced by computer hardware and software for engaging users in environments that are similar to the events and objects of the real world. [1]. Users interplay with virtual images and objects being able to perform several actions (such as handle and shift the objects) in a virtual environment, which elicits a sense of immersion and *presence*. In general terms, a first group of researchers defined presence as "Media Presence", by

L.T. De Paolis and A. Mongelli (Eds.): AVR 2014, LNCS 8853, pp. 137–148, 2014.
DOI: 10.1007/978-3-319-13969-2_11

focusing on the disappearance of the medium from the conscious attention of the individual during the human-computer interaction [1-7]. According to a second group of researcher, an individual is present in a space – real or virtual – when he/she can successfully acts in and performs his/her intentions. According to this perspective, presence can be envisaged as "Inner Presence", defined as the sensation to be in a sense external world around the self: it is a broad psychological phenomenon, not necessarily linked to the experience of a technology [8-10].

Advanced simulations are used to trigger broad empowerment processes induced by a strong sense of presence, leading to greater agency and control over one's actions and environment. In rehabilitation, advanced simulations can reveal in the user/patient what is defined as "transformation" of flow, which affects a person's ability to exploit the best (flow) experience for recognizing and utilizing new and unforeseen psychological abilities as source of engagement. [11]. A powerful sense of presence can be generate by the user through the interaction with virtual objects and events being, thus, able to perform several actions (such as handle and shift the objects) in the virtual environment. Virtual environment are typically experienced with special hardware and software for input (move information from the user to the system) and output (move information from the system to the user). The choice of appropriate hardware is essential since its characteristics may enormously condition the way users respond to a virtual environment [12].

Recent advances in video game technology have made available a large number of low-cost devices that can track the movement of the user, such as Nintendo Wiimote (TM), the Playstation Move (TM) and Microsoft Kinect (TM).

On these premises, the NeuroVirtual 3D platform aims at exploiting these potentiality by designing, developing and testing a low-cost integrated virtual reality solution for applications in clinical psychology and neuropsychological rehabilitation [13]. The platform has been developed expanding the features of the software NeuroVr 2.0 (http://www.neurovr.org/neurovr2/) [14] thanks to the development of a software plug-in allowing the interaction with external devices Microsoft Kinect (TM).

2 A New Platform for Assessment and Neurorehabilitation: NeuroVirtual 3D

Video games are one of the most popular forms of media all over the worlds. Their popularity has been fuelled by advancements in gaming technology, a persisting trend that is producing advanced gaming consoles such as such as Nintendo Wiimote (TM), the Playstation Move (TM) and Microsoft Kinect (TM). Thanks to this incredible spread on home gaming market, in the last years there are a wide series of interactive devices at a low cost also in clinical and research settings [15-18]. A key feature of these consoles is multimodal interaction, achieved through specialized low-cost hardware. In particular, Xbox One includes an advanced version of the motion sensing Kinect camera, updated with the ability to read biometric data and see using infrared. A PC version of the same camera is also available – Kinect for Windows sensor and the Kinect for Windows software development kit (SDK). This offers new opportunities for virtual neuropsychological assessment and rehabilitation. Sophisticated VR systems employ more than specialized visual displays. Engaging the user in the virtual environment may be enhanced via audio display, either ambient or directed to specific stimuli [19]. VR hardware that facilitates the input and output of information, in combination with programmed virtual environments provide the tools for designing tasks that enable users to perform in ways that help them achieve established assessment and rehabilitation goals. When creating a specific virtual rehabilitation tool the clinician and technical team face the challenge of choosing and integrating the software and hardware, and the input and the output methods. Likewise, in a virtual environment, for reaching an experience similar to the real one, is essential for the user be able to navigate and handle objects within it. Thus the user must be able to interact (directly or indirectly) with the environment via a wide array of input technologies. A first class consists of indirect ways for users to manipulate and navigate within a virtual environment, such as the start up of computer keyboard keys, a mouse or a joypad or even virtual buttons being present in the environment. [12]. A second class of input technologies may be considered as direct methods since users behave in a natural way, and the system tracks their actions and responds accordingly. Generally, this is achieved by using special sensors or by visual tracking. With the sensor approach, such as used by Intersense's (www.isense.com) InterTrax2, a three degree of freedom, inertial orientation tracker used to track pitch, roll and yaw movements, the user wears a tracking device that transmits position and orientation data to the VR system. With the visual tracking approach, such as used by VividGroup's video-capture VR system, the user's motion is recorded by video cameras, where special software processes the video image, extracts the user's figure from the background in real-time, and identifies any motion of the body. On this basis, the NeuroVirtual 3D platform is a low-cost integrated virtual reality solution for creating virtual environments useful for neuropsychological assessment and neurorehabilitation by interacting with real-like artifacts (Fig. 1). Actually, NeuroVirtual 3D platform includes two virtual environments: a VR-home for neuropsychological assessment and a VR-city for rehabilitation. VR-home is a two-story home with garden including a living room and kitchen on the ground floor and a bedroom, a bathroom and a studio on the first floor. Instead, VR-city is a small town made-up by 3x3 streets including a square, shops, and buildings. In this two virtual environments patients can carry out several ecological tasks: for example, in the VR-home participants can distribute playing cards, serve tea, or look for clothes in the wardrobe; in the VR-city, participant can look for stores, trash bin or people in the square. The main aim of NeuroVirtual 3D platform is to offer a wide range of interactive contents and rehabilitative exercises. Moreover, the platform now has a specific interface for interaction with external devices like Microsoft Kinect (TM), allowing users to interact in the most flexible with virtual environments. In the following paragraph, the technical features of Kinect-NeuroVirtual 3D interaction interface will be detailed.

Fig. 1. Through NeuroVirtual 3D, users are able to interact with real-like artifact. In this figure it is possible to see a kitchen containing many objects users can move and re-organize by the means of a Microsoft Kinect. Also further interaction and manipulation are developable.

2.1 Microsoft Kinect-NeuroVirtual 3D Interaction Interface: Technical Features

The interaction of the patient with the scene is detected analyzing movements and poses, which are defined relating to each peripheral to be interfaced to. These actions are processed and normalized before creating a new instance of the controller object, which is interpreted by the NeuroVR platform.

Acting on the configuration parameters is possible to perform an accurate tuning of the Microsoft Kinect detections with respect to user's biometric measures in the scene. Considering the data acquired from the Microsoft Kinect is possible to modify configuration parameters to proportionally bias the reactivity and the precision of movement, reported in the scene by the Avatar.

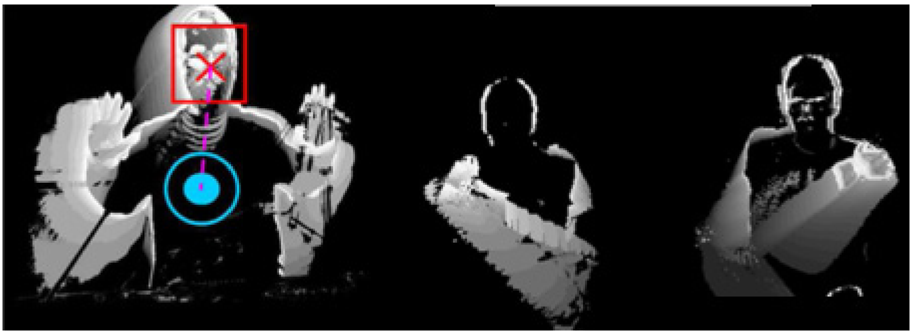

Fig. 2. Video signal processing from kinect

The Avatar's interaction in the virtual scene can be summed up through the following concepts:

Table 1. Avatar's interaction in the virtual scene

Translation	It defines the Avatar's spatial movement in the scene and it is featured by the following actions: go Ahead, go Back, go Right, and go Left.
Rotation	It defines the Avatar's point of view movement and it is featured by the following actions: Tilt Up, Tilt Down, Pan Left and Pan Right.
Command	It defines the recognition of the "poses" the Avatar takes, which are used for recognition of coded actions. e.g.: Left Click, Center Click, Right Click. This element is necessary to implement the Avatar's interaction with items in the scene, e.g.: take the pot, put down the pot.

It can be considered that the intensity of the movements performable in the scene can range from a minimum to a maximum (parametric) and that the movement speed be linearly proportional to the straight line (Fig.3), which links these two values. The Dead Zone (red), defined by the perception thresholds, is identified through the values (N_0, P_0) defined in the configuration. Values related to the upper (or lower) movement threshold are identified by N_0 e N_1 and are specularly acquired by Kinect in rest state. For movements performed within the Dead Zone no movement is notified, outside of it, the maximum value or a fraction of it.

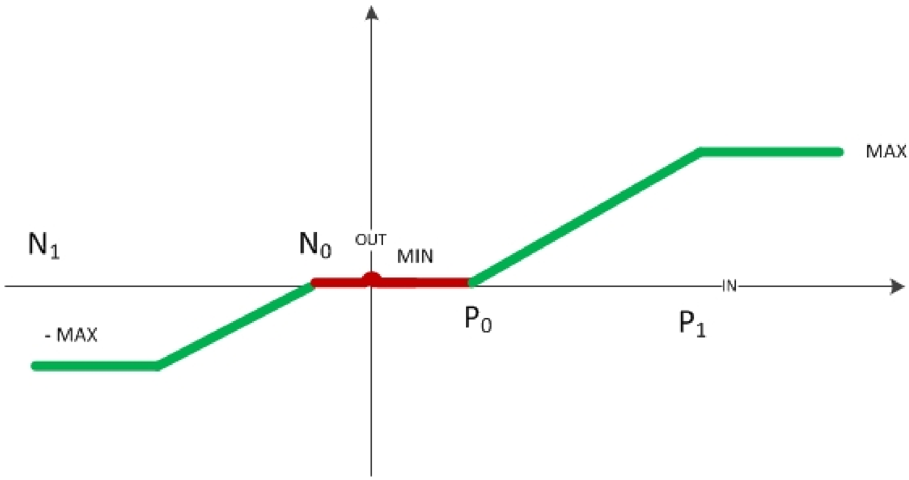

Fig. 3. The behaviour of the transfer function between the data acquired by Microsoft Kinect and the ones supplied to the platform can be made dependent from used parameters. When needed, it is possible to obtain a non-specular behavior with respect to the rest state.

Configuration parameters allow to fine tune the measurements acquired by Microsoft Kinect, when necessary, hence they have to be modified upon reading some biometric measurements of the patient in the studio. Microsoft Kinect reads these data on the user in rest state, and, in case of commands, in poses which the patient would consider natural and spontaneous. Acquired data constitute the threshold values beyond which the movements and the commands for that specific patient are taken into account. For each controllable axis some parameters have been defined, named in the form with the label Axis_ParameterName, with the following meaning for parameters:

- P0: positive value of lower threshold for the movement detection
- P1: positive value of upper threshold for the movement detection
- N0: negative value of lower threshold for the movement detection
- N1: negative value of upper threshold for the movement detection
- OutPut: maximum output value
- Bias: input correcting factor
- Gain: input multiplication factor

Furthermore, two different requirements have been considered to choose the suitable peripherals to be used in the project:

- Low cost and consequent widespread diffusion on the market
- Precise acquisition of patient's movements in the scene to reproduce them in the virtual scene

Microsoft Kinect, Windows version, has been selected, as it satisfies both requirements.

An additional requirement was to allow patients with impaired mobility to interact with virtual environment, as well. In this case, an eye tracker device, which allows patients to manage interaction just through their eyes movement, has been chosen.
The study on the devices to interface to, has considered Kinect, in the first place, due to its peculiar features and to its high diffusion on the market. The Microsoft Kinect-NeuroVirtual 3D Interaction interface has been created in different steps.

- **Sample Data Reading**: position data, as detected by Kinect, are supplied in the spatial form (X,Y,Z) and computed with respect to the device position. To reproduce walking in the virtual scene, values (both linear and angular) related to displacement, rotation and translation observed on the patient must be bound to the avatar corresponding parameters. For this to be accomplished, the spatial coordinates as supplied by Kinect must be processed to be translated in linear or angular measures. From this first requirement, the need to define some area (dead zones) has come out, were no movement is detected and from which displacement, rotation and translation are deduced by difference when going out of the zone itself.

  In the first place, dead zones with a fixed amplitude have been defined, but the use experience has pointed out that is not possible to reposition them when the patient stops moving, producing some "drift" effects, hence errors, in the patient position measurements, or that even distinguish between immobility and movement could be impossible. For such a reason, position and movement have to be determined in a relative fashion, correlating some patient's skeleton elements rather than setting fixed measure points in the deambulation environment (e.g.: marks on floor). The processing algorithm that has come out, allows recognizing the patient's moving will with respect to relative immobility positions.

- **Acquired Data Normalization**: as human body is always moving, even in an apparent rest position (for example the erect one), and since Kinect can detect even the smallest movement of the patient's skeleton, the need to "filter" acquired data has come out, so that the virtual scene representation would be as realistic as possible and not affected by the amplification of such oscillations. To accomplish this, the movement has to be analyzed to determine, for each sample to be acquired, some threshold values, above which the real movement could be detected.
- **Movement Analysis**: to detect the movements performed by the patient, the position she/he takes with respect to the device, and the quality of the process performed analyzing her/his posture, are taken into account. Some movement (particularly the patient's back movements) can be better detected by an accurate selection of the points marked on the patient's skeleton on which the elaboration is performed. To this end, by analyzing experimentally the movement, couples of measurement points that could lead to more accurate patient's position and movement measures, have been identified.

A further constraint, resulting from the movement thresholds setting, has come out, by evaluating the interaction of patients with different build. In these cases, the definition of preset thresholds causes a movement detection that could differ from reality, producing unwanted amplification or attenuation effects. At this end, and to limit these consequences, parametric thresholds, which are bound to patient's biometric features, have been introduced. In this way, a double result is obtained: first, reproducing the movement in the most accurate way, and second, exploiting these chances to introduce voluntarily some amplifications or attenuations on the detected movements (useful, for example, in case of reduced mobility patients).

In the movement analysis context, some studies have been conducted aimed to the actions sequence, or "gestures" recognition. These studies have pointed out big difficulties in this recognition, and, above all, in the univocal identification of gestures, since in a movements sequence many factors can affect negatively the recognition, although not visible at a first sight.

Examples are:

- Gesture execution speed
- Gesture execution amplitude
- Subject performing the gesture (child, elderly, adult, …)

In this regard, it seems more useful to concentrate on the recognition of "Poses", which are static movements, that is, not affected by the above-mentioned issues. Even in the Poses recognition, a thorough study on the patient position, so that all the defined ones do not introduce any ambiguity in the recognition, is necessary.

- **Transfer Functions**: at the conclusion of the studies and of the experimentations performed, through the definition of recognition thresholds, acquired data normalization and information filtering, real Transfer Functions have been defined, which are applied to the data coming from Kinect, when it detects the patient posture. The features of such functions will be discussed hereinafter.

To allow some abstraction on the used peripheral, an interface component (between the NeuroVR and the physical device) has been implemented, by which a generic object, on which the interaction is mapped and managed, has been defined. The communication between systems has been managed using the VRPN framework, by means of which, a device-independent architecture to interconnect peripherals in Virtual Reality applications, has been defined.
The two main devices selected for NeuroVirtual 3D are:

- **Kinect**: to be used in context where body movements can be used to interact with the NeuroVR platform.
- **Eyetracker**: to be used in context where body movements are not possible, due to patient impairments. In this case the interaction with the NeuroVR platform is carried out just moving eyes.

The implemented abstraction has defined a set of actions allowed in the scene, which are mapped on the VRPN object (controller) needed to perform the communication.

The use of the selected devices is always mutually exclusive and does not contemplate a concurrent use of both devices.

3 The Potentiality of NeuroVirtual 3D for Psychology

Sophisticated VR systems employ more than specialized visual displays. Engaging the user in the virtual environment may be enhanced via audio display, either ambient or directed to specific stimuli [19]. VR hardware that facilitates the input and output of information, in combination with programmed virtual environments provide the tools for designing tasks that enable users to perform in ways that help them achieve established assessment and rehabilitation goals. When creating a specific virtual rehabilitation tool the clinician and technical team face the challenge of choosing and integrating the software and hardware, and the input and the output methods. Likewise, in a virtual environment, for reaching an experience similar to the real one, is essential for the user be able to navigate and handle objects within it. Thus the user must be able to interact (directly or indirectly) with the environment via a wide array of input technologies. A first class consists of indirect ways for users to manipulate and navigate within a virtual environment, such as the start up of computer keyboard keys, a mouse or a joypad or even virtual buttons being present in the environment. [12]. A second class of input technologies may be considered as direct methods since users behave in a natural way, and the system tracks their actions and responds accordingly. Generally, this is achieved by using special sensors or by visual tracking. With the sensor approach, such as used by Intersense's (www.isense.com) InterTrax2, a three degree of freedom, inertial orientation tracker used to track pitch, roll and yaw movements, the user wears a tracking device that transmits position and orientation data to the VR system. With the visual tracking approach, such as used by VividGroup's video-capture VR system, the user's motion is recorded by video cameras, where special software processes the video image, extracts the user's figure from the background in real-time, and identifies any motion of the body.

Fig. 4. A virtual room designed for the object interactions with a Kinect

Fig. 5. A room, and in particular a wardrobe with the task of intercting with the virtual dresses using a Kinect

NeuroVirtual3D intends to develop a software interface for supporting assessment and rehabilitation of cognition function through several input/output devices, such as data gloves, joypad and Microsoft Kinect. Traditionally, the assessment and rehabilitation of impairments of cognitive functions (language, spatial perception, attention, and memory) have been carried out with pen-and-paper methods. Psychology has worked to develop several measures to effectively evaluate several cognitive abilities. On one side, traditional measures are generally reliable and have adequate validity. On the other side, however, a critical challenge for psychology has been to find new way to better evaluate and predict everyday abilities. In addition to a precise evaluation, indeed, there is a need for cognitive assessments that reflect real-world situations, in order to better assess functional impairment. Recently, the need for a more ecological and functional evaluation has been the focus of considerable research interest. In this perspective, NeuroVirtual 3D may be used in psychology to allow user to interact with immersed in a computer-generated environment in a naturalistic fashion. Furthermore, the full-body tracking allows selecting among various combination of gestures and limbs without wear any markers, carry any additional device or using specific limbs to interact with the system in an ecological environments. For example, Microsoft Kinets provides low-cost motion tracking sensors, allowing to clinicians to interact with rehabilitation applications in the most natural and flexible way. This flexibility can be employed to tailor the user interaction to the specific rehabilitation user aims [20].

Fig. 6. Using low-cost motion tracking sensors, like Kinect, it also possible to navigate in Virtual environment using body gesture (e.g., a foot ahed), making also this task the most natural and flexible possible

To sum, the future development of application of NeuroVirtual 3D-based therapy to rehabilitation could be very effective in many realms of medicine and psychology. Furthermore, NeuroVirtual 3D could be extremely effective to supply different modality of feedback in sensory deficits such as motion and tactile signals for patients with severe deficits through interactive virtual environments. In conclusion, the cost of tools is dropping, the software is more accessible than before and by now it is much easier to adopt this new technology. As mentioned above, VR systems are becoming very important in the assessment and intervention for cognitive-motor rehabilitation thanks to their specific and unique characteristics.

These characteristics make it highly suitable for the achievement of many rehabilitation aims including the encouragement of experiential, active learning, the provision of challenging but safe and ecologically valid environments, the flexibility of individualized and graded treatment protocols, the power to motivate patients to perform to their utmost capability and the capacity to record objective measures of performance.

Acknowledgements. This study was supported by the research project “NeuroVirtual 3D,” funded by Regione Piemonte (Grant No. FA 211-432C- 2012).

References

1. Sheridan, T.B.: Musings on telepresence and virtual presence. Presence-Teleop Virt. **1**, 120–126 (1992)
2. Schloerb, D.: A quantative Measure of Telepresence. Presence-Teleop Virt. **4**(1), 64–80 (1995)
3. Sadowski, W.J., Stanney, K.M.: Measuring and Managing Presence in Virtual Environments. In: Stanney, K.M. (ed.) Handbook of Virtual Environments Technology, pp. 791–806. Lawrence Erlbaum Associates, Mahwah, NJ (2002)
4. Ijsselsteijn, W., de Ridden, H., Freeman, J., Avons, S.E.: Presence: Concept Determinants and Measurement. In: P SOC PHOTO-OPT INS, San Jose, CA (2000)
5. Lombard, M., Ditton, T.: At the Heart of It All: The Concept of Presence. J. Comput-Mediat Comm. **3**(2) (1997)
6. Sheridan, T.B.: Further Musing on the Psychophysics of Presence. Presence-Teleop Virt. **5**, 241–246 (1996)
7. Marsh, T., Wright, P., Smith, S.: Evaluation for the Design of Experience in Virtual Environments: Modeling Breakdown of Interaction and Illusion. Cyberpsychol. Behav. **4**(2), 225–238 (2001)
8. Riva, G., Davide, F., IJsselsteijn, W.: A Being There: Concepts, Effects and Measurements of User Presence in Synthetic Environments. In: Riva, G., Davide, F. (eds.) Emerging Communication: Studies on New Technologies and Practices in Communication. Ios Press, Amsterdam (2003)
9. Riva, G., Waterworth, J.A., Waterworth, E.L., Mantovani, F.: From Intention to Action: The Role of Presence. New Ideas Psychol. **29**(1), 24–37 (2011)
10. Waterworth, J.A., Waterworth, E.L., Mantovani, F., Riva, G.: On Feeling (the) Present: An Evolutionary Account of the Sense of Presence in Physical and Electronically-Mediated Environments. J. Consciousness Stud. **17**(1–2), 167–178 (2010)
11. Riva, G., Castelnuovo, G., Mantovani, F.: Transformation of flow in rehabilitation: the role of advanced communication technologies. Behav. Res. Methods **38**(2), 237–244 (2006)
12. Rand, D., Kizony, R., Feintuch, N., Katz, N., Josman, N., Rizzo, A.A., Weiss, P.L.: Comparison of two VR platforms for rehabilitation: video capture versus HMD. Presence-Teleop Virt. **14** (2005)

13. Cipresso, P., Serino, S., Pallavicini, F., Gaggioli, A., Riva, G.: NeuroVirtual 3D: A Multi-platform 3D Simulation System for Application in Psychology and Neurorehabilitation. In: Ma, M. (ed.) Virtual, Augmented Reality and Serious Games for Healthcare 1, pp. 275–286. Springer (2014)
14. Riva, G., Gaggioli, A., Grassi, A., Raspelli, S., Cipresso, P., Pallavicini, F., Vigna, C., Gagliati, A., Gasco, S., Donvito, G.: NeuroVR 2–a free virtual reality platform for the assessment and treatment in behavioral health care. Stud. Health. Technol. Inform. **163**, 493–495 (2011)
15. Giakoumis, D., Drosou, A., Cipresso, P., Tzovaras, D., Hassapis, G., Gaggioli, A., Riva, G.: Using activity-related behavioural features towards more effective automatic stress detection. PloS One **7**(9) (2012b)
16. Giakoumis, D., Drosou, A., Cipresso, P., Tzovaras, D., Hassapis, G., Gaggioli, A., Riva, G.: Real-time monitoring of behavioural parameters releted to psychological stress. St Heal T. **181**, 287–291 (2012)
17. Lee, J., Chao, C., Thomaz, A.L., Bobick, A.F.: Adaptive Integration of Multiple Cues for Contingency Detection. In: Salah, A.A., Lepri, B. (eds.) HBU 2011. LNCS, vol. 7065, pp. 62–71. Springer, Heidelberg (2011)
18. Rigas, G., Tazallas, A.T., Tsalikakis, D.G., Konitsiotis, S., Fotiadis, D.I.: Real-time quantification of resting tremor in the Parkinson's disease. In: Conf. Proc. IEEE Eng. Med. Biol. Soc., pp. 1306–1309 (2009)
19. Vastfjall, D.: The subjective sense of presence, emotion recognition, and experienced emotions in auditory virtual environments. Cyberpsychol. Behav. **6**(2), 181–188 (2003)
20. Lange, B., Chang, C.Y., Suma, E., Newman, B., Rizzo, A.A., Bolas, M.: Development and evaluation of low cost game-based balance rehabilitation tool using the Microsoft Kinect sensor. In: Proc. IEEE Int. Conf. Eng. Med. Biol. Soc. (2011)