



---

## **Preliminary Analysis of Low-cost Motion Capture Techniques to Support Virtual Ergonomics**

---

**Giorgio Colombo**

Department of Mechanics, Polytechnic of Milan  
Milan, Italy

Tel: +39-02-23998259

Fax: +39-02-23998202

email: giorgio.colombo@polimi.it

**Daniele Regazzoni, Caterina Rizzi,  
Giordano De Vecchi**

Department of Engineering, University of Bergamo  
Dalmine (BG), Italy

Tel: +39-035-2052075

Fax: +39-035-2052077

email: [giordano.devecchi; daniele.regazzoni; caterina.rizzi]@unibg.it

**Abstract:** This paper concerns the development of a computer-aided platform to analyse workers' postures and movements and ergonomically validate the design of device a man or woman may deal with. In particular, we refer to pick and place operations of food items on the display unit shelves. The paper describes three low-cost solutions that integrate two optical motion capture techniques (one based on web-cam and another on MS Kinect sensor) and two human modelling systems (Jack and LifeMod) with the main goal of determining the suitability of operators' working conditions and, eventually, providing a feedback to the design step. The solutions have been tested considering a vertical display unit as case study. Preliminary results of the experimentation as well as main benefits and limits are presented. The results have been considered promising; however, we have planned to perform an acquisition campaign in the real environment, the supermarket.

**Keywords:** Human modelling, Motion capture, Virtual ergonomics, Pick and place.

---

### **1 Introduction**

Virtual ergonomics permits engineers to create and manipulate virtual humans to investigate the interactions between the worker or the consumer and the product. For example, in product design, human factors such as positioning, visibility, reaching, grasping and lifting of weights can all be evaluated, providing a feedback to designers in the early steps of product development.

This paper refers to this context and addresses commercial refrigeration industry; in particular those companies specialized in display units. It describes the use of virtual

ergonomics techniques in the design process of display units commonly used in supermarkets.

In previous research activities we have experienced the use of virtual manikins specifically targeted for ergonomics to evaluate postures and movements respect to requirements established by international standards to reduce health risks [1]. In this work we try to enhance that approach to provide more precise results. For example, when filling the shelves of a display unit, the workers act quite differently one from another. To face this issue we introduce the use of motion capture systems to reproduce with virtual humans the real way in which operators behave. In particular, we have evaluated the possibility to adopt and integrate Motion Capture (Mocap) and Digital Human Modelling (DHM) systems to perform ergonomic analysis relying on real movements performed by operators in everyday activities.

The paper, after a brief description of the state of the art of adopted techniques, describes three technical solutions based on low cost Mocap devices integrated with human modelling tool. Then, the case study and preliminary experimentation are presented.

## 2 Related works

The implementation of the proposed virtual ergonomics platform requires the integration of Mocap and DHM and system to acquire end-users' postures and movements to determine fatigue, stress and risk for workers' health. In the following we provide an overview of both techniques.

First DHM tools for ergonomics analysis appeared in late 60s, mainly in aeronautics and automotive industries. Nowadays, in literature we can find various tools of different complexity depending on the target application. We have classified them into 4 main categories [2-4]: *virtual human/actors for entertainment* used to populate scenes for movies and videogames production [5, 6]; *mannequins for clothing* used to create virtual catwalks, virtual catalogues, and virtual try-on show rooms and to design garments [7-8]; *virtual manikin for ergonomic analysis* used to define complex scenes, analyse postures, simulate tasks and optimize working environments [9-11]; and, finally, *detailed biomechanical models* [12, 13] whose applications concern ergonomics analysis, study of safety in transport, and human performance during sports activities, and medical device, etc. [14]. For our purpose, we consider DHM tools belonging to the last two categories.

Mocap techniques appeared in the 70's as a branch of imaging techniques applied to biomechanical processes and several applications have been developed for different contexts such as military, entertainment, sports, and medical applications. According to the working principle four main categories can be identified [15]: optical, mechanical, inertial, and magnetic that can be used for testing and validation also in a combined way [16].

We focus the attention on optical systems, considered the high-end technology. They extrapolate the position of body joints by triangulation between images taken from different cameras. We can have systems with active or passive markers [17] or marker-less [18]. In the first case markers are placed in the remarkable points of the body while in the second case a dedicated software module recognizes the human figure.

In literature, we can also find researches that integrate DHM either with Virtual Reality (VR) and Mocap systems to improve the level of interaction and realism within

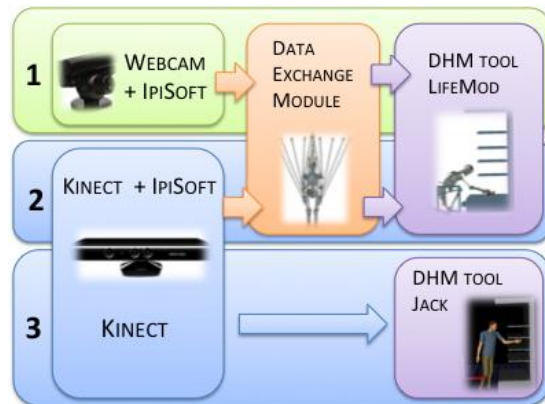
the virtual environment, to drive the virtual human and facilitate the evaluation of comfort and prediction of injuries that could arise when executing a task [19-20].

In our work we consider low cost techniques, developed for video games and entertainment, to verify their usability and performance in industrial context.

### 3 Technical Solutions

As mentioned before, main goal has been to experiment the integration of low cost Mocap system and DHM to evaluate postures and movements of workers filling the shelves of a display unit. We considered two systems, both optical and marker-less: the former based on Sony Eye webcams and the latter on Microsoft Kinect sensor. We have been using them to acquire the real movements and postures of operators so that the following simulation can be based on real data. Both solutions are not expensive and can be easily moved and, with some precautions, used also outside the lab in potentially any work environment we want to acquire.

Concerning human modelling, we adopted two different tools, Siemens Jack specifically targeted for ergonomics and LifeMod to generated detailed biomechanical model. Figure 1 shows the technical solutions adopted.



**Figure 1** Adopted technical solutions

#### 3.1 Webcam solution

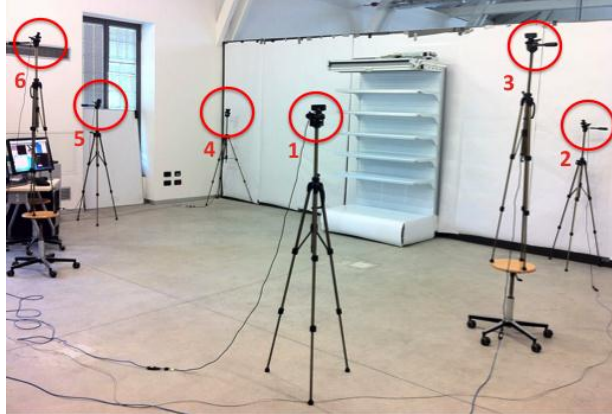
This solution includes three main components: webcam, a module for data exchange and LifeMod to create the human avatar and reproduce postures and movements. Figure 2 shows the acquisition system composed by:

- Six Sony Eye webcams with a resolution of 640x480 pixels at 60 Hz mounted on photographic tripods.
- Portable workstation Dell Precision M6500.
- iPi Desktop Motion Capture™ (Ipi DMC) software.

iPisoft Software is a marker-less system developed to work with Sony Eyes webcams and its main features are:

- Possibility to use from 3 to 6 webcams.

- A maximum acquisition area of 7 m x 7 m.
- Non real-time tracking.
- Input file format: MPEG.



**Figure 2** Webcam solution

The system acquires synchronized video sequences obtained with the cameras without having to apply any type of physical markers on the operator's skin. It automatically recognizes the different body segments and, then for each time step, calculates joints position and orientation. iPi DMC includes two main components: iPisoft Recorder for the motion capture phase that synchronizes images recorded from the six webcams and iPisoft Studio for the recognition and segmentation of body and tracking of movement. iPisoft Desktop Motion Capture output contains the recorded movement in BVH (Biovision Hierarchical Data) format. iPisoft adopts a skeleton made of 27 joints hierarchically organized, each characterized by proper d.o.f. and constraints. However, it has been necessary to develop a data exchange module to be integrated with the considered human modelling tool, namely LifeMod.

LifeMod is a biomechanical analysis software that permits to generate a complete biomechanical model of the human body. It is a plug-in of ADAMS software, a multibody analysis system. The creation of a model normally starts with the generation of a basic set of connected human segments based on the dimension contained in an anthropometric database; then, the joints, the muscles and the tendons are created and contact force with objects are defined. By the way, to reproduce the movement with the operators' avatar we need to exchange the data acquired from the Mocap system to the human modelling environment. This problem has been solved developing an ad-hoc conversion algorithm in Matlab. The algorithm translates the information relative to the joint hierarchy and to the motion contained in the BVH file to a SLF formatted file used by LifeMod. In order to complete the information required by a SLF file, anthropometric data are retrieved directly from user's data or from an anthropometric database (e.g., GeBOD, PeopleSize and US Army - Natick Database).

Before proceeding in reproducing the operators' tasks using LifeMod, another issue must be faced. Actually, the conversion performed in the previous step could not take into account the fact that the information contained in the SLF file are related to the real position of the joints acquired subject while the biomechanical software use external markers placed on the skin surface. Once again to fix this problem an ad-hoc script has

been created to relocate parametrically to model size the position of the markers. To reduce the number of operations necessary to run the analyses, an automatic procedure and a CMD script have been implemented. The script is completely automatic and no user interaction is required because data relative to the markers initial position and the marker movements are produced by converting this information directly from BVH file.

Once the model is defined, simulation phase can begin. To obtain accurate simulations with the muscles and the articulations it's necessary to execute a first inverse dynamic simulation to drive the body with motion agents describing the movements to execute. Once that the movements are stored a direct dynamic simulation is run to calculate the forces created by the muscles and the stresses the body is subjected to. The outcome provided by the system consists of forces and momenta acting on each joint in each time step of the analysis.

### *3.2 Kinect Solution #1*

Similarly to the previous solution, also this one comprehends three main parts. The main difference is the Mocap system, which is composed by (Figure 3):

- Microsoft Kinect with a resolution of 640x480 pixels at 30 fps, mounted on photographic tripod and connect via USB cable.
- Portable workstation Dell XPS.
- iPi Desktop Motion Capture software.



**Figure 3** Kinect solution #1 and # 2

As in the previous case iPi Recorder manages the recording of images and depth videos coming from Kinect, while the iPi studio performs environment calibration and video analysis. In addition, MS Kinect SDK libraries are required.

The other two components of this acquisition solution, namely Data exchange and DHM modules, are exactly the same of the webcam solution.

### *3.3 Kinect Solution #2*

This solution always uses a MS Kinect but, differently from the two previous one, Kinect sensor is fully integrated with the Jack, a well-known human modelling system (Figure 1, 3). It permits to define complex scenes with virtual manikins and objects, simulate many tasks and evaluate posture and ergonomics factors also using analysis tools such as RULA (Rapid Upper Limb Analysis) to investigate work-related upper limb disorders,

NIOSH (U.S. National Institute for Occupational Safety and Health) lifting equations to evaluate lifting and carrying tasks, and OWAS (Owako Working Posture Analysis System) to analyze postures during work. Therefore iPi DMC software is not necessary.

This solution requires:

- Mocap Toolkit, a module specifically developed by Siemens for Kinect sensor.
- MS Kinect SDK v 1.0 libraries.
- SkeletalViewer sw to transfer data streaming acquired with Kinect sensor to Jack.

The skeleton used in the transition from Mocap to Jack is made of 21 joints whose positions and movements are tracked and there is not a hierarchy among them. This skeleton is less complex than Jack's one and some details cannot be taken into account (i.e., head rotation, fingers).

## **4 Preliminary Experimentation**

In the following we describe the case study adopted for the experimentation and the preliminary results.

### *4.1 The case-study*

The case study refers to a vertical refrigerator display unit typically installed in groceries or supermarket, with 4 or 5 shelves that could be placed at different heights (Figure 4).

Different groups of people interact with the display unit: supermarket staff filling the shelves with goods, workers in charge of maintenance operations and customers picking up goods. In this paper we focus on the first category with the main goal of determining the suitability of operators' working condition.



**Figure 4** Vertical Refrigerator (Image courtesy: [www.costan.com](http://www.costan.com)) (left); Simplified display unit (right)

The task of loading a refrigerator is critical from the ergonomics point of view and may generate severe health disorders, mainly due to: repeated actions, holding and lifting loads, insufficient recovery times, uncomfortable postures, and uncomfortable environment (e.g., cold and humid).

To experiment the described solutions, we decide to focus the attention on pick and place operations of food items. We have conducted a preliminary study (interviews and direct observations) to analyse how operators really behave. This permitted to identify main operations to be reproduced in the lab and some occasional unacceptable practices such as assuming completely wrong and dangerous postures.

To acquire postures and movement we use a simplified version of the real unit (Figure 4 right) more efficient because some elements of the complete display unit (e.g. lateral walls) may interfere with the operator during motion capture.

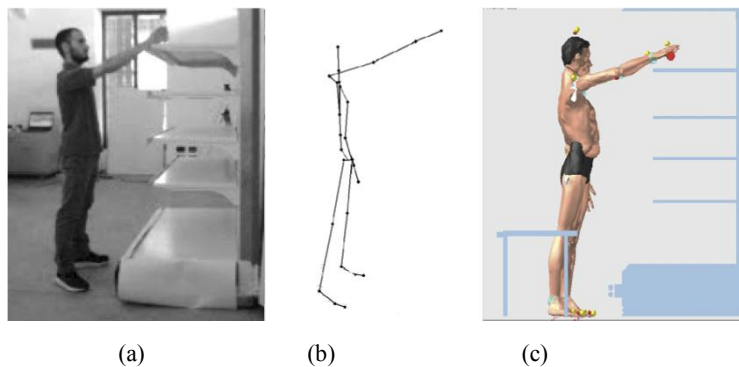
#### *4.2 Experimentation results*

Actors were asked to perform as much as possible as if they were in the real environment and to follow a precise routine to produce comparable results with operators characterized by different anthropometric measures. The routine defines the initial and final position of each movement to be performed. For each solution, the loading of a bottle of water in the four different shelves has been reproduced.

First the solution based on webcam has been tested. The system requires a first step of calibration that initializes the system and permits to correctly locate each camera in space.

A semicircle disposition of the webcams at different heights around the operator is the best choice. Once calibration is done, it is possible to realize the acquisitions taking care to avoid fast movements and carefully execute the required task. It has been conducted recording ten loading routines with subjects of both genders and of different heights to evaluate if the motion capture system is affected by any problem. In particular, the conversion algorithm has been tested as well as different ways of automatically changing the position of joints depending on height and structure of subjects.

After having tested the system the real motion capture campaign has been performed and the data of the movement have been converted and analyzed with LifeMod. In Figure 5 one can see the representation of data related to the loading of the highest during the three main steps. The first one (Figures 5a) refers to the environment in which the webcam images are captured and elaborated for each time step to gather joints positions shown in the second representation (Figures 5b). The third representation (Figures 5c) comes from the human modelling system where data have been converted, corrected and integrated with anthropometric databases.

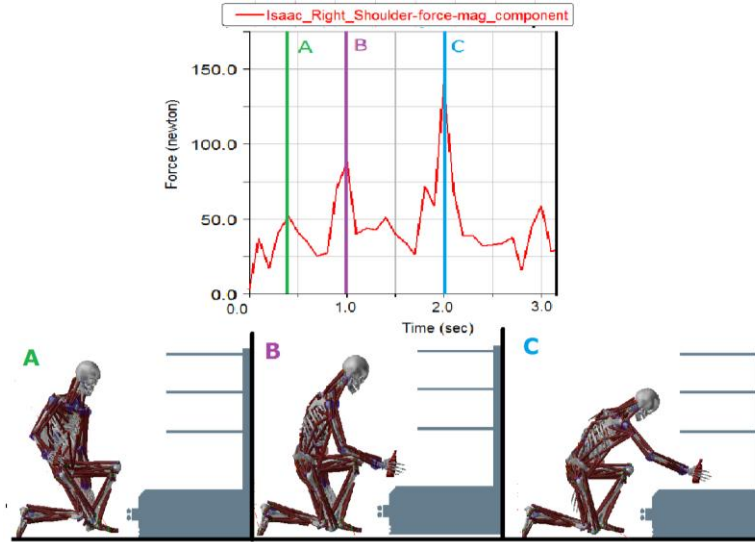


**Figure 5** Webcam solution: highest shelf loading

The DHM tool provides different kinds of output data relative to biomechanical analysis; in particular it is possible to obtain results such as center of mass position, velocity and acceleration of segments, forces, torque and power for joints and additional information for soft tissues and environment interaction. Figure 6 shows an example of analysis result for the virtual human with one knee on the ground while performing the loading of a bottle on the lowest shelf. As an example the resultant force on the shoulder

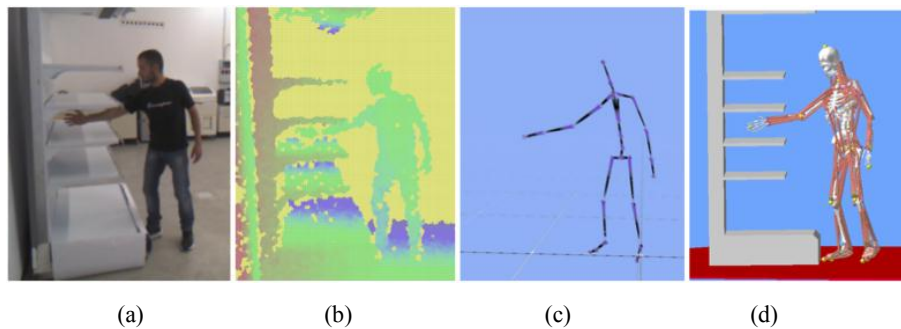


holding the bottle is plotted in the graphic and three screenshots corresponding to peaks are shown.



**Figure 6** Resultant force over time acting on right shoulder while loading the lowest shelf and three notable position

Regarding Kinect solution #1 we proceeded similarly to the previous one. Figure 7 shows the results related to the loading of the second shelf. Figures 7a refers to the environment in which the Kinect images are captured, Figure 7b shows the corresponding depth map, Figure 7c the reconstructed skeleton and Figure 7d the virtual avatar of the actor corrected and integrated with anthropometric databases. Results obtained were similar with those of the previous solution.



**Figure 7** Kinect solution #1: second shelf loading

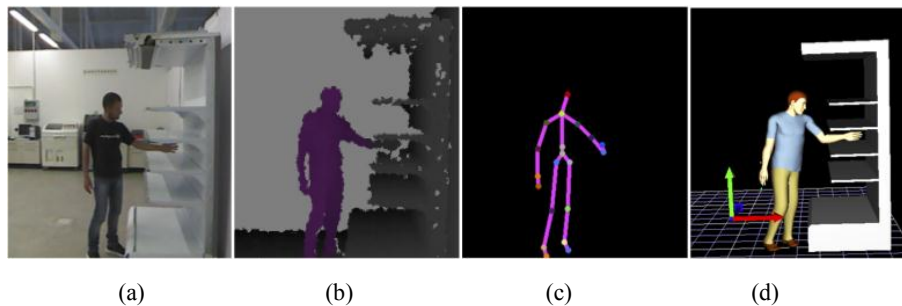
Finally we tested the Kinect solution #2, i.e. the Kinect sensor integrated with Jack. First, using Jack the virtual scene composed by the 3D model of the refrigerated unit and the operator avatar (in this case a virtual manikin 50th ile) is created; then, the Kinect plug-in and the SkeletalViewer software are launched and the acquisition session can start. Figure 8 shows the final step acquired during the loading of the second shelf. Precisely Figure 8a shows the image acquired, Figure 8b the corresponding depth map,



Figure 8c the corresponding reconstructed skeleton and Figure 8d the virtual avatar of the actor.

This solution has the advantage that the two environments, Mocap and the human modelling, are fully integrated; therefore Jack reproduces the operators' movements in real-time. For the other two solutions it is necessary to translate the data acquired with the Mocap system into a format readable by LifeMod and then reproduce the movements and postures of the actor/worker. Moreover Jack already includes dedicated modules that implement analysis tools such as RULA, NIOSH and OWAS to study and validate operator postures in order to avoid musculoskeletal disorders.

Kinect used to track human body constitutes an interesting novelty in the instrumentation available for Mocap and virtual ergonomics at industrial level. Its low cost, broad diffusion and availability of libraries let us foresee it will be commonly used for Mocap application in near future. By the way there are still some limitations affecting its performance when used with Jack. Actually, the skeleton does not always match the subject posture, this is particularly true when body's area overlap. It is also recommended to acquire the entire actor so that Jack can make the posturing more robust, but Kinect limited working area, if compared to webcam solution, can constitute a limit. Anyway, for most or ergonomic situation these limitations can be acceptable and successful simulations can be carried out.



**Figure 8** Kinect solution #2: second shelf loading

## 5 Conclusions

The goal of this work has been to verify the use of low-cost Mocap systems integrated with digital human modelling to evaluate ergonomics factors of industrial products, in our case refrigerated display units commonly used in supermarket.

Three solutions, based on two marker less Mocap systems and two different types of human modelling tools, have been preliminary tested considering the pick and place operations of a vertical display unit. The results have considered promising and interesting; however some limits have been highlighted and an acquisition campaign in a real environment, the supermarket, has been planned. In fact, starting from operators performing real tasks instead of standardized average tasks allows not only to be more precise in the final evaluation but also to consider unknown or incorrect postures or movements performed by operators in their everyday activities. In addition, data coming from acquisition campaign can be used to analyse performance variability and build a structured, domain dependant, database of real postures and movements to be applied to similar cases. This would allow benefiting from real data without the need of performing motion capture for similar cases.

## References

- 1 Colombo, G., De Ponti, G., Rizzi, C., "Ergonomic design of refrigerated display units", *Virtual and Physical Prototyping*, 3(5), pp. 139-152, 2010.
- 2 Sundin, A., Ortengren, R., "Digital Human modelling for CAE applications", *Handbook of Human Factors and Ergonomics*, 3rd Edition, John Wiley & Sons, New York, NY, USA, 2006.
- 3 Magnenat-Thalmann, N. and Thalmann, D., "Handbook of Virtual Humans", John Wiley & Sons, Chichester, UK, 2004.
- 4 Rizzi, C., "Digital Human Models within product development process", *Innovation in Product Design*, Springer Verlag, pp. 143-166, 2011.
- 5 <http://www.massivesoftware.com>. Accessed July 2012
- 6 <http://crowdit.worldofpolygons.com>. Accessed July 2012
- 7 Volino, P., Luible, C., Magnenat-Thalmann, N., "Virtual clothing", John Wiley and Son Ltd, 2008.
- 8 Li, S.S.M., Wang, C.C.L., and Kin-Chuen Hui, "Bending-invariant correspondence matching on 3D human bodies for feature point extraction", *IEEE Transactions on Automation Science and Engineering*, 8(4), pp. 805-814, 2011.
- 9 Colombo, G. and Cugini, U., "Virtual Humans and Prototypes to Evaluate Ergonomics and Safety", *J Eng Design* 16(2), pp. 195-203, 2005.
- 10 Mueller, A., Maier, T., "Vehicle Layout Conception Considering Vision Requirements - A Comparative Study within Manual Assembly of Automobiles", *Digital Human Modeling for Design and Engineering Conference and Exhibition*, 2009.
- 11 Green, R.F., Hudson, J.A., "A Method for Positioning Digital Human Models in Airplane Passenger Seats", *Advances in Applied Digital Human Modeling*, CRC Press, 2011.
- 12 Abdel-Malek, K., et al., "A physics-based digital human model", *Int. J. Vehicle Design*, 51(3/4), pp. 324-340, 2009.
- 13 <http://www.lifemodeler.com>. Accessed July 2012.
- 14 Bucca, G., Buzzolato, A., Bruni, S., "A Mechatronic Device for the Rehabilitation of Ankle Motor Function", *J Biomech Eng*, 2009.
- 15 Furniss, M., "Motion capture", MIT Communications Forum. <http://web.mit.edu/commforum/papers/furniss.html>. Accessed July 2012.
- 16 Schepers, H.M., "Ambulatory Assessment of Human Body Kinematics and Kinetics", PhD Thesis, Universiteit Twente, The Netherlands, 2009. Available at <http://www.xsens.com/images/stories/PDF/ThesisSchepers.pdf>.
- 17 <http://www.vicon.com>. Accessed July 2012.
- 18 Bray, J., "Markerless Based Human Motion Capture: A Survey", Vision and VR Group, Dept Systems Engineering, Brunel University, available at <http://visicast.co.uk/members/move/Partners/Papers/MarkerlessSurvey.pdf>. Accessed July 2012.
- 19 Colombo, G., De Angelis, F., Formentini, L., "Integration of virtual reality and haptics to carry out ergonomic tests on virtual control boards" *International Journal of Product Development*, 11(1/2), pp. 47-61, 2010.
- 20 Spada, S., Sessa, F., Corato, F., "Virtual reality tools for statistical analysis for human movement simulation. Application to ergonomics optimization of workcells in the automotive industry" *Work* 41, pp. 6120-6126, 2012