International Conference on Innovative Design and Manufacturing
DECEMBER 12-14, 2012, Taipei, Taiwan

INDUSTRIAL APPLICATION OF DHM FOR ERGONOMICS DRIVEN BY LOW COST MOCAP SOLUTIONS

Giorgio Colombo
Department of Mechanics,
Polytechnic of Milan
giorgio.colombo@polimi.it

Daniele Regazzoni
Department of Engineering,
University of Bergamo
daniele.regazzoni@unibg.it

Caterina Rizzi
Department of Engineering,
University of Bergamo
caterina.rizzi@unibg.it

Giordano De Vecchi
Department of Engineering,
University of Bergamo
giordano.devecchi@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 or the condition a man or woman may deal with. In particular, we refer to two different applications: pick and place operations of food items on a display unit and gait analysis. The paper describes two low-cost solutions that integrate two optical motion capture systems (one based on web-cam and another on MS Kinect sensor) with the main goal of determining the suitability of operators' working conditions and, eventually, providing a feedback to designers. 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 and the comparison with optical marker Mocap systems.

KEYWORDS

Human modelling, Motion capture, Virtual ergonomics, Pick and place.

1. INTRODUCTION

Shortening design cycle is a main issue enterprises and academia have been dealing with in the last years. A key strategy to reach this goal consists in the early identification of design faults to fix them before spending time and money in product development. In the case the object being designed is supposed to have an interaction with one or more operators ergonomics issues must be taken into account at the very first step of design as soon as product architecture is sketched.

Virtual ergonomics is the discipline that permits engineers to create and manipulate virtual humans to investigate the interactions between the worker, a generic user, and the product. For example, in product design, human factors such as positioning, visibility, reaching, grasping and lifting of weights can all be evaluated by using virtual humans, providing a feedback to designers without the need for physical prototyping.

This paper refers to this context and addresses to two different application: commercial refrigeration industry and gait analysis. The first one refers to the design of refrigerated display units to be used by different actors (e.g., customers or workers filling out shelves). The second one refers to the medical domain of lower limb prostheses design in which gait analysis is crucial to define setup parameters. In previous research activities we have experienced the use of virtual manikins specifically targeted for ergonomics to evaluate postures and movements according to requirements established by international standards to reduce health risks (Colombo et al, 2010). In this work 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 or patients performing specific activities. We compared two different acquisition solutions to ensure precise results for ergonomics while adopting low cost and highly portable solutions.

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

2. TECHNOLOGY AND APPLICATION DOMAINS

The implementation of the proposed virtual ergonomics platform requires the integration of Mocap and DHM. In the following we provide a brief overview of both techniques.

In literature we can find various DHM tools of different complexity depending on the target application. We can distinguish four main categories (Magnenat-Thalmann, 2004, Rizzi, 2011): entertainment (e.g., movies and videogames) (MassiveSoftware, 2012, Crowdit, 2012); clothing industry for virtual catwalks, virtual catalogues and to design garments (Volino et al, 2008, Li et al, 2011); ergonomics to analyse postures, simulate tasks and optimize working environments (Colombo and Cugini, 2005, Green and Hudson, 2011); and detailed biomechanical models diffused in the medical field (Abdel-Malek, 2009, Lifemodeler, 2012). For our purpose, we consider DHM tools belonging to the last category since it is the most appropriate to compare the results of the different Mocap solutions.

Mocap techniques 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 (Furniss, 2012): optical, mechanical, inertial, and magnetic. Optical system extrapolates the position of body joints by triangulation between images taken from different cameras. We can have systems with active or passive markers (Vicon, 2012). In the first case markers are placed in the key points of the body while in the second case a dedicated software module recognizes the human figure and its segments. Our choice goes to optical marker-less systems (Bray, 2012), such as webcams and Kinect sensors, since they are low cost and portable solutions.

In our work we consider techniques, developed for video games and entertainment, to verify their usability and performance in industrial context. Actually, such technologies benefit from a huge investment on research that leads to a rapid evolution but on the same time they keep affordable prices because of the target market they refer to. Moreover, the capillary diffusion ensures the ease of sharing of information on pros and cons, tips and unusual use researchers may find worldwide.

Due to the characteristics of the proposed solutions, integrated Mocap and DHM systems could finally enter in many domains where high costs and complex procedures prevented them to be used. In particular, in this paper we refer to two domains of application that have completely different requirements, constraints and require specific outcomes. Actually, the parameters we may want to extract from the virtual human acting like the real one vary depending on the domain of application.

2.1. CASE STUDIES

The first case study refers to a vertical refrigerator display unit to be installed in groceries or supermarket, with 4 or 5 shelves, optionally closed with doors (Figure 1).

Figure 1 - Vertical refrigerated display units (Image courtesy: www.costan.com)

At least three categories of people interact with a refrigerated display unit in a supermarket: customers picking up goods, maintenance technicians who need to access to some specific components and workers in charge of checking exposed products and filling out shelves with new ones. For each of this category some ergonomic aspects are more relevant than others: i.e., visibility and reachability of goods for customers, reachability of some display components for technicians and, the most important, posture and stress for operators who repeat the same task for several hours in a row and may occur in musculoskeletal disorders. Thus, it urges to assess the critical conditions potentially affecting workers' health and we simulate pick and place operations for the loading of a display unit. We have also 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 which is more efficient because some elements of the complete display unit (e.g. lateral walls) may interfere with the operator during motion capture.

The second case study refers to a completely different domain and so is the kind of output we want to achieve. Even though we applied the same methodology and same data exchange the pros and cons of the two systems may differ from the previous case study. Concerning this case we want to perform a gait analysis and the aim is to measure all distances, angles and times characterizing the movement of the body walking a set of steps. In this case the interactions are only with the floor but the precision requested is generally higher. The analysis we performed has the overall goal to determine whether a patient with leg prosthesis has reached a satisfactory quality of walking.

3. VIRTUAL TESTING SOLUTIONS

This paragraph shows the different solution of the two Mocap systems and the flow of data. Figure 2 shows the experimental platform.

We can distinguish three main components: the Mocap systems, a module for data exchange and DHM tool to create the human avatar and reproduce postures and movements. The first is dedicated to acquiring movements and postures from real operators. The data exchange module has been implemented since there is not a direct link between the selected Mocap systems and the human modelling tool. The third part consists in analysing by means of the Human modelling system (in our case LifeMOD) the data acquired and elaborated so that ergonomic outcomes can be gathered.

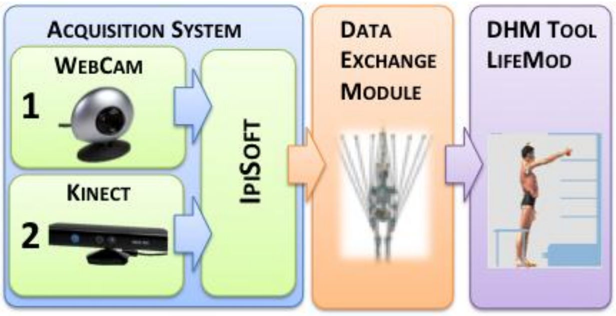


Figure 2 - Experimental platform

3.1. ACQUISITION TECHNIQUES

We considered two Mocap systems, both optical and marker-less: the former based on Sony Eye webcams and the latter on Microsoft Kinect sensor. 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.

Both solutions foresee the adoption of iPisoft Software, a non-real time marker-less system developed to work with Sony Eyes Webcams or Microsoft Kinect.

3.1.1 Webcam Solution

Figure 3 shows the acquisition system based on webcams. It is 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 recently also with Kinect. 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.

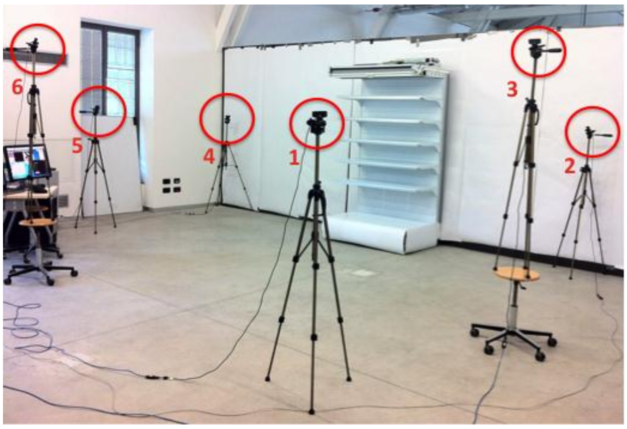


Figure 3 - Webcam solution

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. We decided to use 6 cams to be sure not to lose the tracking when some body segments are hidden by other boy segments of by any device. 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 DmC 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.

3.1.2 Kinect Solution

Similarly to the previous solution, this Mocap solution is composed by (Figure 4):

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

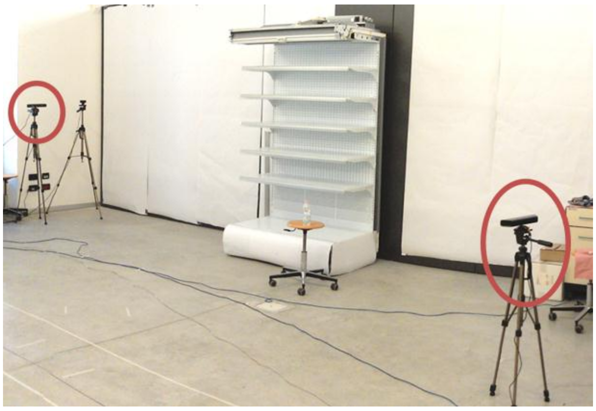


Figure 4 Kinect solution

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.

3.2. EXCHANGE MODULE AND HUMAN MODELLING

To reproduce the humans' movement we need to pass the data acquired from the Mocap system to the simulation environment. The data format of the two systems is not compatible; therefore, we developed an ad-hoc algorithm in Matlab, which translates the information relative to the joint hierarchy and to the motion contained in the BVH file to a SLF formatted file used by the commercial tool LifeMOD (Figure 5).

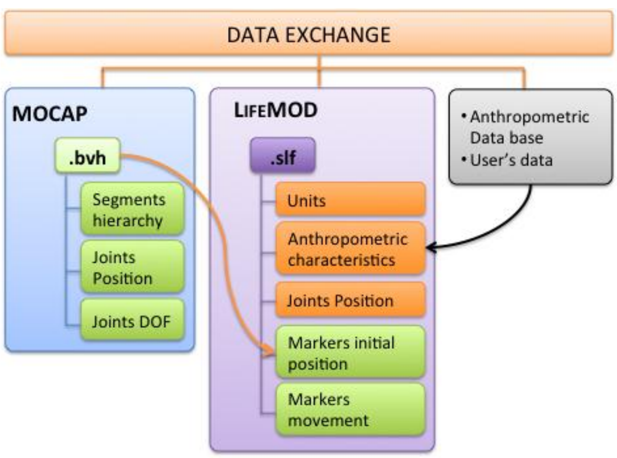


Figure 5. Data exchange and marker relocation

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 (Cheng, 1996), PeopleSize) that does not require any specific features or functions and can take into account the characteristics of major ethnic groups. However, before proceeding in reproducing the operators' tasks with virtual humans, another issue must be faced. The conversion performed in the previous step do not take into account that the information contained in the SLF file are related to the real position of the joints acquired subject while the biomechanical software uses external markers placed on the skin surface. In this case, we implemented a script to relocate parametrically the position of the markers in the virtual human modeller. To reduce the number of required operations necessary to run the analyses an automatic procedure has been created. A CMD script has been developed to generate the model segments and joints and to permit the adaptation of the data. This script, as well as the previous file format conversion script, is completely automatic and no user interaction is required because data relative to markers initial position and movements are produced converting this information directly from BVH file. As mentioned, we adopted LifeMOD for human modeling and analysis. It is a plug-in of ADAMS software, a multibody analysis system. The creation of a model starts with the generation of a basic set of connected human segments (e.g., bones and skeleton) 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. Each simulation starts with the virtual human in a specific initial position and each model can be active or passive. Passive models react to forces coming from the environment around them, for instance gravity and contact with the ground.

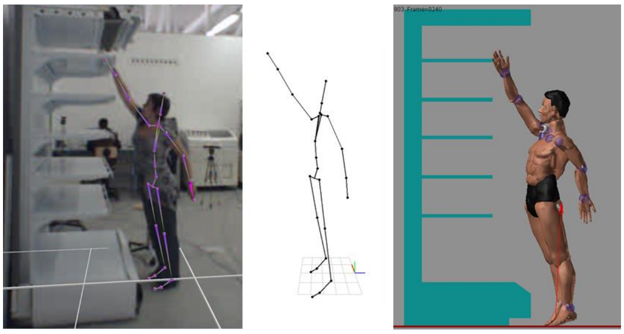
Active models produce reactions in the environment due to their movements. To obtain accurate simulations with the muscles and the articulations it's necessary to execute a first inverse dynamic simulation to drive the body with motion agents describing the movements to execute. Once that the movements are stored a direct dynamic simulation is run to calculate the forces created by the muscles and the stresses the body is subjected to. LifeMOD needs as input the Motion Capture data in SLF format; this file format can be used also to provide model anthropometric characteristics, initial position and markers positions.

The outcome provided by the system consists of forces and momenta acting on each joint in each time step of the analysis.

4. PRELIMINARY EXPERIMENTATION

In the followings we describe the case study adopted for the experimentation and the preliminary results. We acquired the tasks performed by the actors simultaneously with both Mocap systems.

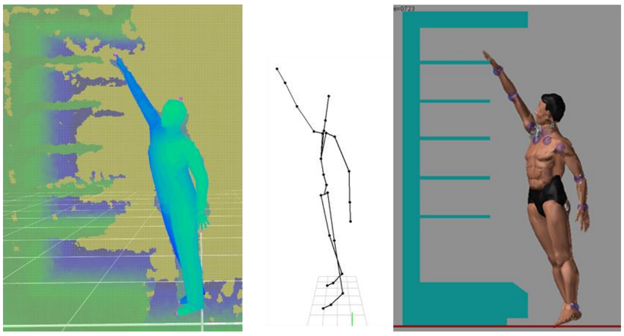
For what concerns the first case study actors were asked to perform as much as possible as if they were in the real environment and to follow a precise routine to produce comparable results with operators characterized by different anthropometric measures. The routine defines the initial and final position of each movement to be performed. For each solution, the loading of a bottle of water in each shelf has been reproduced. At first the solution based on webcam has been tested. This solution 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 analysed with LifeMOD. In Figure 6 one can see the representation of data related to the loading of the highest during the three main steps. The first one (Figures 6a) 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 6b). The third representation (Figures 6c) comes from the human modelling system where data have been converted, corrected and integrated with anthropometric databases.



(a) (b) (c)

Figure 6 - Webcam solution: loading highest shelf

Exactly the same performance has been captured simultaneously using the Kinect solution. The sensors were placed laterally to the actor. As an example, Figure 7 shows the same kind of images of Figure 6 in which 7(a) is the only one which differs being the depth map of the Kinect sensor.



(a) (b) (c)

Figure 7 - Kinect solution: highest shelf loading

In the second case study actors were asked to perform a straight walk looking ahead, starting before and stopping outside the area of capturing and keeping, as much as they could, the same speed.

First the solution based on webcam has been tested and as in the previous case study the system must be calibrated. Since the area to be acquired in this case is long and tight and we can access it from both sides the webcams are placed on the vertexes of a rectangle containing the area and corresponding to the mid points of the longest side. Several people of different height and size were asked to walk and their motions were captured with no major drawbacks since there were no devices interfering or creating occlusions.

Figure 8 shows as for the first case study the three main phases of the acquisition performed for each time frame.

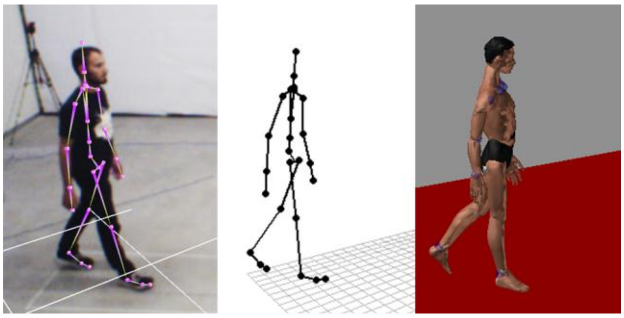


Figure 8 - Webcam solution: 1st step

The Kinect solution has been tested simultaneously on exactly the same walks with no need of calibration. The two sensors were placed on two opposite corners of the area to be captured so that they can both see the largest part of the path according to their limited range of use. The absence of any device occluding the scene ease the acquisition but to gather a good result both sensors must see the actor at any time. Figure 9 shows the three main phases starting with the Kinect depth map referred to the first complete step.

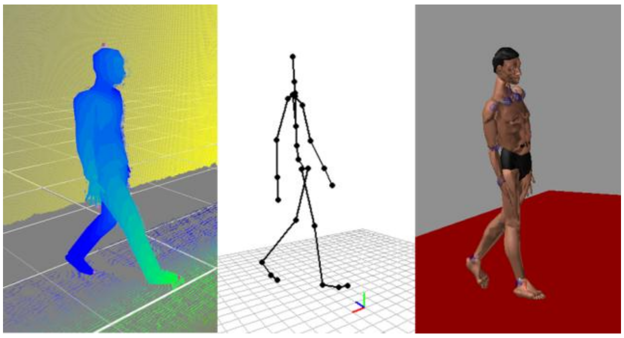


Figure 9 - Kinect solution: 1st step

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 perfectly match the subject posture, this is particularly true when body's area overlap. Head arms and hands rotations recognition is still quite poor and the area of acquisition is smaller if compared to the area covered by the webcams (this was particularly limiting for gait acquisition since only two or three complete steps were recorded at each run). On the other hand, Kinect solution does not require complex routine to calibrate sensors and, thus, the overall setup is much shorter. Kinect sensors are less sensible to light conditions and this contributes to a better portability in almost any working environment where the scene is small enough to be seen by both sensors.

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 in different domains of application.

Two solutions, based on two marker less Mocap systems have been tested considering the pick and place operations of a vertical display unit and the analysis of key parameters of gait. The results have considered promising and interesting; however a deeper analysis of tracking errors will performed in the near future. For the first case study 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. For the second case study a comparison with a more precise optical solution with marker (e.g., Vicon) will help in determining errors and to validate both low cost solutions also for medical application which were considered too complex to be addressed.

The results gathered so far techniques let us foresee a huge application in almost all industrial fields wherever a worker is asked to perform a manual operation and/or to interact with a machine.

REFERENCES

Abdel-Malek, K., et al., "A physics-based digital human model", Int. J. Vehicle Design, 51(3/4), pp. 324-340, 2009.

Bray, J., "Markerless Based Human Motion Capture: A Survey", Vision and VR Group, Dept Systems Engineering, Brunel University,

Cheng, H., "The development of the GEBOD program" Proceedings of the 1996 Fifteenth Southern Biomedical Engineering Conference, pp. 251-254, 1996.

Colombo, G. and Cugini,, U., "Virtual Humans and Prototypes to Evaluate Ergonomics and Safety", J Eng Design 16(2), pp. 195-203, 2005.

Colombo, G., De Ponti, G., Rizzi, C., "Ergonomic design of refrigerated display units", Virtual and Physical Prototyping, 3(5), pp. 139-152, 2010.

Furniss, M., "Motion capture", MIT Communications Forum.http://web.mit.edu/ commforum/papers/furniss.html. Accessed July 2012.

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.

http://crowdit.worldofpolygons.com. Accessed July 2012

http://www.lifemodeler.com. Accessed July 2012.

http://www.massivesoftware.com. Accessed July 2012

http://www.vicon.com. Accessed July 2012.

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.

Magnenat-Thalmann, N. and Thalmann, D., "Handbook of Virtual Humans", John Wiley & Sons, Chichester, UK, 2004.

Rizzi, C., "Digital Human Models within product development process", Innovation in Product Design, Springer Verlag, pp. 143-166, 2011.

Volino, P., Luible, C., Magnenat-Thalmann, N., "Virtual clothing", John Wiley and Son Ltd, 2008.