

## Low-Cost Garment-Based 3D Body Scanner

Nicolò BIASI, Francesco SETTI, Mattia TAVERNINI, Alberto FORNASER, Massimo LUNARDELLI,  
Mauro DA LIO, Mariolino DE CECCO

Department of Mechanical and Structural Engineering, University of Trento, Italy

### Abstract

While in the last decade Laser-based technologies became the reference for 3D body scanning, vision-based technologies became more and more important for motion capture. We are now proposing a mixed approach, based on passive vision technology and able to scan 3D human bodies, as well as capturing motion. In this paper we will present a low-cost multi-camera body scanning system based on a special wearable garment. The garment is a black suit with colored markers, about 3000, uniformly distributed on a square grid all over the body of the subject. Each colored marker is uniquely identified by an ID based on the color of the marker itself and its 8 neighbors. The sequence of 9 color was studied to be unique (no repetitions), independent to rotation (unique starting marker from which to read the color sequence) and non symmetrical (color sequence is unique read either in clockwise and anti-clockwise direction). This allows to deal with thousands of points with no matching outliers.

System is composed by 2 up to 12 RGB cameras in order to perform single limb or full body scanning. System calibration is performed in a single step semi-automatic procedure. Synchronous images acquisition is guaranteed by a triggering device, avoiding problems of subject moving during acquisition. Uniform illumination is assured by 6 neon lamps with high frequency electrical ballast. With this system is possible to achieve an accuracy of 1mm on a planar surface.

### 1. Introduction

As proved by several research papers [1,2] and commercial products [3,4], 3D body scanning is an important issue for clothes industry and an emerging research field. In the last years, two main techniques became leader for this application: Laser-based and Vision-based.

Laser-based systems consist in a laser emitter and one or more CCD sensor; this allows to perform triangulation for 3D reconstruction. This kind of scanners can be very accurate and they can provide a huge amount of data (points onto the scanning surface). The main drawback is due to the fact that, in order to scan a surface, laser emitter has to be rotated to project the beam onto the entire scanning area; this means the acquisition takes some seconds (typically 5 up to 20 seconds), allowing uncontrolled movements of the subject in the meantime, and the maintenance can be difficult. Most of commercial scanners use [3,4,5] this technology; anyway the very high cost of this systems, several hundreds of thousands of Euro, is a barrier to the widespread, leaving space on the market to different technologies.

State-of-the-Art methods for Vision-based scanners are mainly of two types: marker-based and markerless.

Markerless algorithms are based on silhouettes: using a high number of cameras all around the workspace (or a rotating camera or a camera and a rotating platform), images can be processed to extract 2D profile of the subject; these represent a set of visual cones, the intersection of them is a volumetric representation of the subject [6]. This method is very simple and set-up can be very cheap, accuracy and reliability are critical points. In particular, since the resulting shape can be very complex, the method can generate protrusions, and it is not able to handle cavities.

Marker-based (or photogrammetric) methods compute 3D position of a set of feature points (markers) by means of several synchronous images [7] reported the use of consumer cameras for close range surface measurement, comparing results obtained by using different cameras. Several proprietary software, such as Photomodeler by Eos [8], can be used for similar applications; studies as [9] and [10] report applications to humans and animals respectively, while [11] reported the use of this software for facial digitalization and compares results with a laser scanner. A new technology is now emerging in this field: structured light [12,13,14]. Differently from laser scanners, structured light works with one-shot images, avoiding problems related to movements of the subject; this technology is also cheap compared to laser. On the other hand, if the surface to scan is big or complex, several shots (or several projector-camera units) could be necessary.

A similar system has been proposed by [15], where a set of cameras and a special garment are used to perform 3D scanning. These families of products are cheaper than most of the other technologies, but accuracy and the number of reconstructed points are much lower.

Previous experiences have proven the quality of reconstruction of multi-stereo systems also with a higher markers density. The problem of multi-stereo is the merging of the different cloud of points corresponding to each stereo pair [16,17]. As natural evolution of these systems in this paper we present a multi-camera vision-based 3D body scanner. This system was developed in the framework of the EU FP7 project VERITAS for the estimation of the human kinematics parameters and ranges of motion. In this case we propose its use in a still condition.

Remarkable properties of this system are:

- good accuracy for most of the applications (~1mm)
- high number of points (~3000 markers)
- synchronized acquisition
- low acquisition time
- till to 10 Hz scan frequency
- possible automatic extraction of anthropometric parameters
- low-cost hardware
- low setup time
- portable

Synchronized acquisition is an important feature that allows to deal with undesired movement of the subject that, with other system (laser, structured light) will reduce the accuracy of reconstruction. This feature is particularly useful with certain categories of subject, like children and elderly, for which keeping a still position for few second can be difficult.

Low acquisition time associated with a high scan frequency can be used also to perform body scan while the subject is moving or performing a particular action, keeping track of anthropometric parameters during acquisition.

In section 2 we will provide a detailed description of the system, focus on the hardware (sub-section 2.1) chosen and the software realized (sub-section 2.2). Section 3 is dedicated to multi-camera calibration system developed for this particular application. Experimental results are reported in section 4.

## **2. Description of the system**

In this section the multi-camera marker-based 3D body scanning system will be described and analyzed in detail. In particular we will focus on the hardware that composes the system, describing the acquisition system and the suit worn by the subject, and the software to detect, label and triangulate the markers to reconstruct 3D subject body surface.

### **2.1 Hardware**

Hardware has been chosen in order to realize a 3D body scanner: modular, portable and low cost.

The hardware set-up is composed by:

- 2 up to 12 RGB cameras
- 1 trigger device
- 1 main PC for image acquisition, processing and user interface
- up to 5 secondary PCs for image acquisition (optional)
- a black dotted garment
- 6 neon lamps

A schematic representation of set-up is shown in figure 1.

For a single shot acquisition a single PC is needed. All the cameras are connected to this PC via USB and disposed all around the subject. Role of this PC is to initialize cameras and trigger device, which is a TTL function generator signal that assure that all cameras shot at same instant.

Once all cameras are ready the PC sends a signal to the trigger device that will mirror to all cameras a synchronized shot signal. Images captured are saved directly to PC's RAM, processed and then stored on the hard disk.

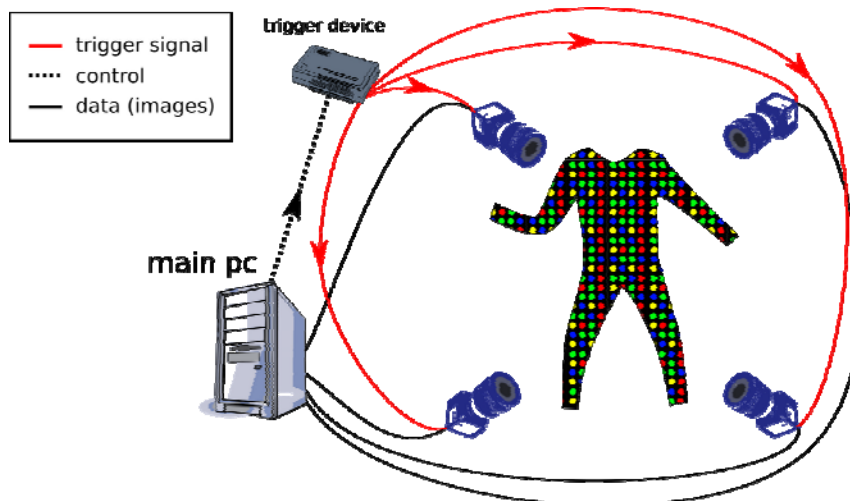


Fig. 1. Hardware connection of 3d body scanner system for single shot acquisition.

System is completely scalable using 2 up to 12 cameras in different configurations and dispositions. This is particularly useful because with the same system is possible to perform either a full or partial body scanning. Full body scanning requires the disposition of the cameras all around the subject, generating a working area of 2x2x2 meters. Body part or single limb scanning is possible positioning the cameras closer to the subject. System accuracy is affected by many elements, but in particular from camera-subject distance: focus on a body part and positioning the camera close to the subject means that accuracy will increase. Cameras employed are triggerable USB PointGray Chamaleon with a resolution of 1280x960 pixels.

It is also possible to perform multiple shots acquisition. Cameras employed indeed can capture images till a frequency of 10 Hz. This huge amount of data it is not manageable through USB bus on a single PC. Thus it is necessary to introduce other PCs. New connection scheme is shown in figure 2.

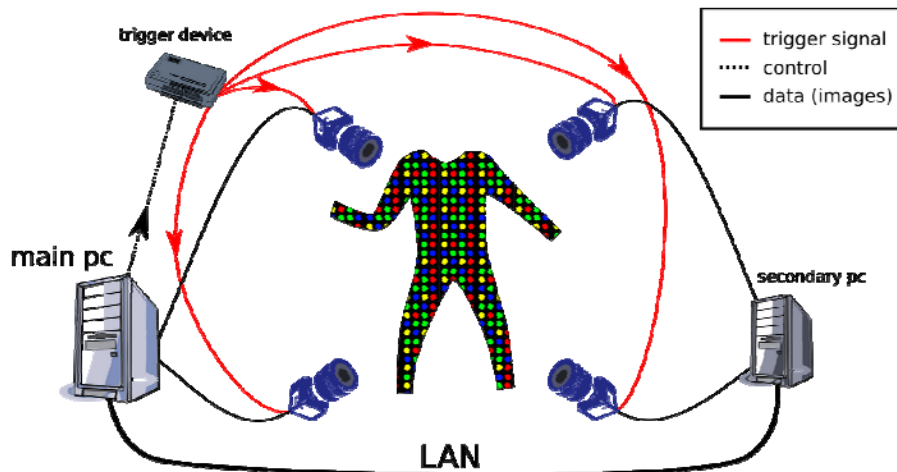


Fig. 2. Hardware connection of Body scanner system for multiple shot acquisition.

We employed one PC for each couple of cameras. These PCs, called "secondary", are managed by a PC, called "main", via LAN. Role of this main PC is to send a signal to secondary PC to initialize cameras and to manage trigger device. Once all secondary PCs are ready, main PC send a signal to the trigger device that will generate a TTL function, with a specific frequency, that synchronize the acquisition of all cameras.

A multiple shot acquisition with a frequency of 10 Hz can be either used to select the best reconstruction, mean the various acquisitions in one second or analyze motion of the subject.



Fig. 3. 3D Body scan system.



Fig. 4. Suit with colored marker worn by the subject of scansion.

The subject is requested to wear a special suit (figure 4); this is a black suit with colored markers (~3000) uniformly distributed on a square grid all over the body of the subject. Each marker and its 8 neighbors (figure 5) generate a color sequence, which, once converted to a numerical sequence, is the marker's ID.

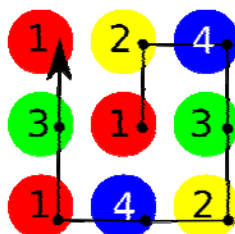


Fig. 5. Colored sequence of central marker and 8 neighbors.

We use 4 colors: red, yellow, green, blue; HSV components of colors have been chosen in order to maximize the reciprocal distance in color space. Disposition of the colors on the garment has been studied in order to create a sequence of color with specific features, in fact every ID is:

- **unique:** there are not two markers in the garment with the same ID
- **rotation invariant:** marker's ID is the same even if the colors of neighbors are all shifted forward (or backward) of  $N$  positions
- **non symmetric:** there is no colors symmetry along one of the axis of the marker (figure).

For all the markers seen by two or more cameras, it is possible to recover its 3D position through triangulation. Since everything is based on color coding, image acquisition must be achieved in controlled light condition. Uniform illumination of the scene is assured by 6 neon lamps with mercury vapor at low pressure and high frequency electrical ballast.

This system is considered low cost because employing standard hardware with a total cost of few thousands of Euros.

## 2.2 Software

The software developed has the following outline:

- marker detection (onto each image plane)
- ID assignment
- triangulation

Marker detection is a fully automated procedure based on Maximally Stable Extremal Regions (MSER) blob detection algorithm [18]. This detector allows to recover information on centroid coordinates, shape and color of the marker.

ID assignment is straightforward the detection of neighbors of each marker.

In some cases it is possible that one (some) neighbor of a marker are not detected in the image. This mainly happen because of occlusions, e.g. the arm of the subject partially occludes the torso (figure 6). In this example the 3 upper neighbors of the selected marker (red cross) are hidden by the subject's arm. In this case the central marker is identified and extracted from the image but will not receive an ID (because of the incompleteness of the color sequence).

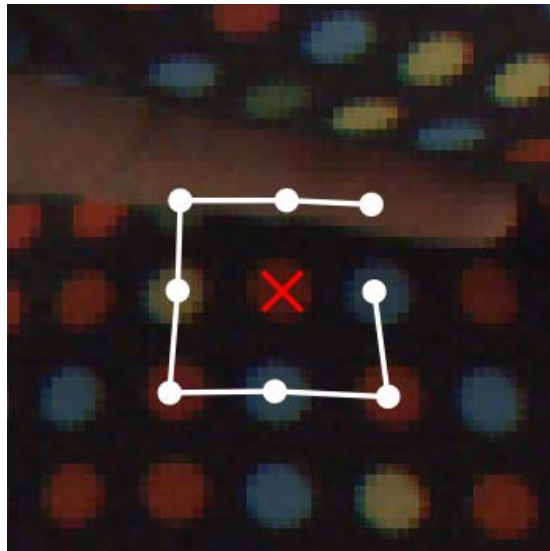


Fig. 6. Example of occlusion of 3 out of 8 neighbors (white dots) of a central marker (red cross).

To deal with occlusions, the grid printed on the garment has been mapped, recording relative position and ID of each marker. Mapped markers arrangement allows to recover the ID of markers with unassigned ID, once the ID of at least one of its neighbors is known. Prior mapping of the garment is also used to cross-check markers' IDs, looking for ambiguities and incompatibilities.

Once assigned an ID to each marker in the image, association of marker between images is straightforward. Note that uniqueness property of marker's ID allows us to match correctly regions with strong deformation and perspective effects.

Thus, after calibrating the system, it is possible to triangulate the marker recovering their 3D spatial position.

Marker triangulation procedure is divided in two steps:

- optical ray distance weighted mid-point triangulation
- Bundle Adjustment optimization

Despite two views are enough to perform triangulation, more views will give a strong improvement in accuracy. In order to triangulate markers seen from more than two cameras, we decided to compute all the possible coupled triangulation with mid-point method [19], and then fuse the results of triangulations with a weighted mean function. Weight of each triangulation is proportional to the square of the inverse of the distance of the optical rays for that particular triangulation. Triangulated points with an optical ray distance above a certain threshold ( $\sim 10$  pixel) are considered as outlier and thus deleted.

Position of 3D points thus calculated is then refined with a Bundle Adjustment [20], that minimizes the overall error of reprojection between the locations of observed and predicted image points.

Custom anthropometric measurements can be defined by an operator. Define these measurements is just as easy as create a list of ID of interest that are tracked during acquisition. In case we are interested in the length of a limb, it is sufficient to select the markers (and relative ID) at limb extremes, e.g. shoulder-elbow, hip-knee, knee-ankle. In the case of curve length, e.g. waist girth, it is necessary to define the list of makers, and relative ID, through which passes the curve. These examples are shown in figure 7.

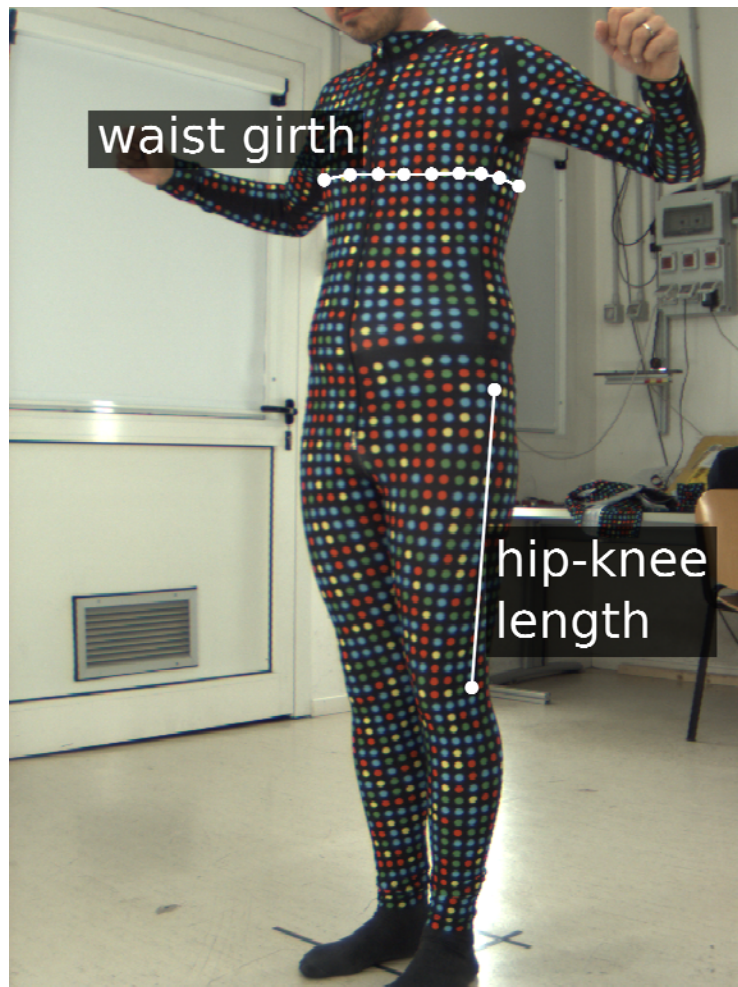


Fig. 7. Markers used to define custom anthropometric parameters.

### 3. System Calibration

Calibration is achieved in a two step semi-automatic procedure based on Bouguet [21] algorithm. Calibration procedure consists in:

- calculating intrinsic parameters for each camera and a first guess of extrinsic parameters
- optimize with Bundle Adjustment the only extrinsic parameters

Intrinsic parameters are: camera's focal length, skew factor, position of central point and lens distortions coefficients, while extrinsic parameters are rotation and translation of cameras' reference frame with respect to world reference frame.

This procedure is achieved capturing a series of about 100 images from every camera of a rectangular colored dotted grid plotted on a plane in different orientation and position (figure 8).

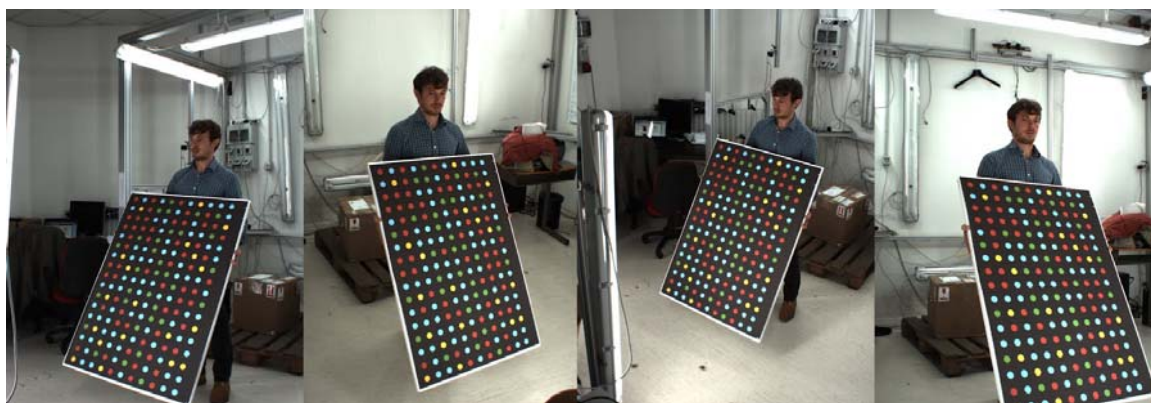


Fig. 8. Calibration grid captured by 4 different cameras.

Once the images are captured is possible to extract the marker position in image and thus recover the position and orientation of the grid in 3D space. Grid spatial position is used to calculate intrinsic parameters for each camera separately and a first guess of extrinsic parameters.

First guess of extrinsic is possible keeping track of the frames for which the grid is seen at the same time by two cameras, building a camera-connection matrix.

Extrinsic parameters are successively refined with a Bundle Adjustment technique, while intrinsic are kept fixed. After calibration, mean reprojection error is lower than 1 pixel.

Grid used to calibrate the system is generated with the same algorithm as the one used for the garment, i.e. each dot on the grid it is identified by a unique ID. This feature brings two main advantages compared with standard chessboard-based methods [21]:

- allow partial grid visibility, since we know relative position of each marker
- connection between cameras far apart, since we can handle matching with strong perspective deformation

### 4. Experimental Results

Employed garment has a total of ~3000 markers distributed on the whole body surface with a mean number of 30 markers in an area of 10x10cm. A finer reconstruction can be achieved reducing marker diameter and step grid on the suit and employing cameras with higher resolution and/or higher focal length.

Some results on full body scanning are shown in figure 9.

Precision obtained with this scanning system have been evaluated with the same planar grid used for the calibration. A new set of images of the grid has been acquired and 3D position error has been computed. The accuracy of the overall system has been estimated in 1mm for full body configuration. Better accuracy can be reached in case of partial body configuration of cameras.

### 5. Conclusions

The 3D body scanner presented in this paper confirms the usefulness of marker photogrammetry using color coded targets to digitize human bodies. 3D points on body surface of the subject are recovered with accuracy for most of the application and with a discrete density.

System proved to meet goals of portability, modularity and low cost, showing potential for extensive use in textile and entertainment industry.

Software realized allows a fully automated reconstruction of 3D points with minimal user intervention.

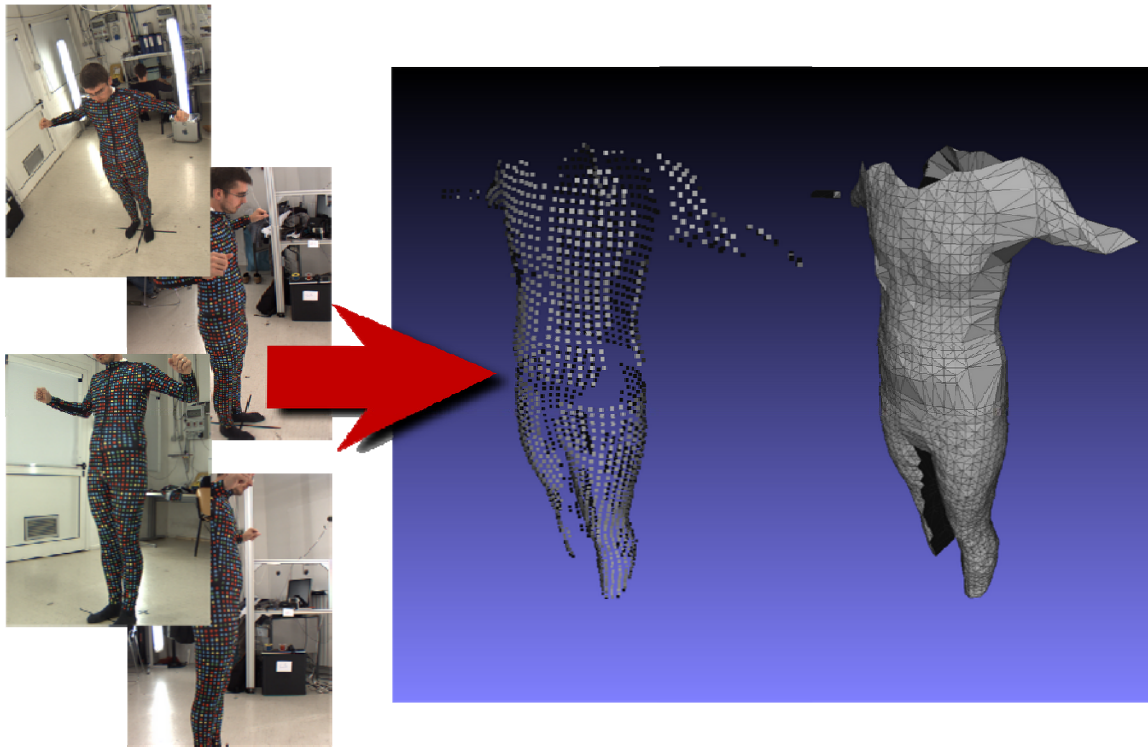
In particular introduction of colored coded sequence and ID to marker assignation is a very effective technique to facilitate matching procedure.

Mapping the suit has proven to be effective to deal with occlusions and cross-check of marker ID.

Introduction of these elements in multi-camera calibration system realized for this application improved also camera calibration reliability.

Better accuracies can easily be achieved by using a highly redundant sensor network, increasing the number of cameras, camera resolution or focus the attention only on a part of subject body.

Further studies will be focus on automatic body segmentation and anthropometric parameters calculation.



*Fig. 9. 3D points and mesh reconstruction of a single shot acquisition.*

## **Acknowledgment**

The work was supported by the Research Support Scheme FP7-ICT-2009.7.2, Accessible and Assistive ICT, grant no. 247765, IP project VERITAS (Virtual and Augmented Environments and Realistic User Interactions To achieve Embedded Accessibility DesignS).



## References

1. Faust, M.-E. and Carrier, S. (2009): "3D Body Scanning's Contribution to the Use of Apparel as an Identity Construction Tool", *Lecture Notes in Computer Science*, Vol.5620/2009, pp.19-28.
2. Bellemare, J. (2011): "The configurator as a part of a clothing mass customization program", *World Conference on Mass Customization, Personalization, and Co-Creation 2011*.
3. Whole body color 3D scanner "Model WBX" (accessed 2012): <http://www.cyberware.com>
4. 3D full body scanner "VITUS" (accessed 2012): <http://www.vitronic.de>
5. 3D bodyline scanner (accessed 2012): <http://www.hamamatsu.com>
6. Bottino, A. and Laurentini, A. (2001): "A Silhouette Based Technique for the Reconstruction of Human Movement", *Computer Vision and Image Understanding*, 83: 79–95.
7. Chandler, J. H., Fryer, J. G. and Jack, A. (2005): "Metric capabilities of low-cost digital cameras for close range surface measurement", *Photogrammetric Record*, 20(109): 12–26.
8. Photomodeler (accessed 2012): <http://www.photomodeler.com>
9. Larsen, P. K., Hansen, L., Simonsen, E. B. and Lynnerup, N. (2008): "Variability of bodily measures of normally dressed people using PhotoModeler Pro 5", *Journal of Forensic Science*, 53(6): 1393–1399.
10. Tasdemir, S., Yakar, M., Urkmez, A. and Inal, S. (2008): "Determination of body measurements of a cow by image analysis", *CompSysTech'08: Proceedings of the 9th International Conference on Computer Systems and Technologies and Workshop for PhD Students in Computing, Gabrovo, Bulgaria. V.8-1–V.8-7*.
11. Deli, R., Di Gioia, E., Galantucci, L. M. and Percoco, G. (2010): "Automated landmark extraction for orthodontic measurement of faces using the 3-camera photogrammetry methodology", *Journal of Craniofacial Surgery*, 21(1): 87–93.
12. Pribanic, T., Mrvos, S. and Salvi, J. (2010): "Efficient multiple phase shift patterns for dense 3D acquisition in structured light scanning", *Image and Vision Computing*, 28(8): 1255–1266.
13. Yang, R., Cheng, S. and Chen, Y. (2008): "Flexible and accurate implementation of a binocular structured light system", *Optics and Lasers in Engineering*, 46(5): 373–379.
14. Yu, W. and Xu, B. (2010): "A portable stereo vision system for whole body surface imaging", *Image and Vision Computing*, 28(4): 605–613.
15. Percoco, G. (2011): "Digital close range photogrammetry for 3D body scanning for custom-made garments", *The Photogrammetric Record*, 26: 73-90.
16. De Cecco, M., Pertile, M., Baglivo, L., Lunardelli, M., Setti, F. and Tavernini, M. (2010): "A Unified Framework for Uncertainty, Compatibility Analysis and Data Fusion for Multi-Stereo 3D Shape Estimation", *IEEE Transactions on Instrumentation and Measurements*, vol: 59 , Issue: 11, pp 2834 – 2842, 2010.
17. De Cecco, M., Pertile, M., Baglivo, L., Parzianello, G., Lunardelli, M. and Setti, F. (2009): "Uncertainty analysis for multi-stereo 3d shape estimation", *AMUEM International Workshop on Advanced Methods for Uncertainty Estimation in Measurement Proceedings*, Bucarest, 2009.
18. Matas, J., Chum, O., Urban, M. and Pajdla, T. (2004): "Robust wide-baseline stereo from maximally stable extremal regions", *Image and Vision Computing* 22.10 (2004): 761-767.
19. Hartley, R. I. and Sturm, P. (1995): "Triangulation." *Computer Analysis of Images and Patterns* 1995: 190-197.
20. Triggs, B., McLauchlan, P. F., Hartley, R. I. and Fitzgibbon, A. W. (1999): "Bundle Adjustment - A Modern Synthesis". In *Proceedings of the International Workshop on Vision Algorithms: Theory and Practice (ICCV '99)*. Springer-Verlag, London, UK, 298-372.
21. Bouguet, J.-Y. (2004): "Camera calibration toolbox for matlab". (accessed 2012): [http://www.vision.caltech.edu/bouguetj/calib\\_doc](http://www.vision.caltech.edu/bouguetj/calib_doc)
22. Ebert, A., Schaedlich, J., and Disch, A. (2003): "Innovative retexturing using cooperative patterns", *Proceeding* (396) *Visualization, Imaging, and Image Processing – 2003*.
23. Scholz, V., Stich, T., Keckeisen, M., Wacker, M. and Magnor, M. (2005): "Garment Motion Capture Using Color-Coded Patterns", *Computer Graphics Forum*, volume 24, pages 439–447. Wiley Online Library, 2005.