Surface measurement and tracking of human body parts from multi-image video sequences

Nicola D'Apuzzo *

Institute of Geodesy and Photogrammetry, Swiss Federal Institute of Technology, ETH-Hoenggerberg, 8093 Zurich, Switzerland

Abstract

This paper describes a method to measure and track moving surfaces of human body parts from multi-image video sequences acquired simultaneously by several cameras. The gained 3-D data can be of two different types: surface measurement of the visible parts of the human body at each time step of the sequence and surface tracking in the form of a vector field of 3-D trajectories (position, velocity and acceleration). The surface measurement process, which is based on multi-image photogrammetry, consists of five steps: calibration of the camera system, simultaneous acquisition of images from different views, establishment of corresponding points in the images, computation of their 3-D coordinates and, eventually, generation of a surface model. The high level of automation achieved in all the steps makes the processing of long image sequences possible. The tracking process is based on least squares matching techniques. The main idea is to track corresponding points in the multi-images through the sequence and compute their 3-D trajectories. When applied to all the points matched on the body, it results in a vector field of trajectories. Some key-points can be defined and tracked in the vector field, producing general 3-D information about the recorded movement. The main advantages of the presented method are: the capability to dynamically measure surface parts with high accuracy and the possibility to extract motion information from the acquired data without using markers. Two applications are presented to demonstrate the functionality of the proposed method: human face modelling and full body motion capture.
© 2002 Elsevier Science B.V. All rights reserved.

Keywords: surface measurement; surface tracking; video sequences; multi-image; 3-D modelling; human body; human face

1. Introduction

In recent years, surface and movement measurement of the human body has gained importance for medical applications. The interested disciplines are orthopedics (Ono, 1995; Hackenberg et al., 2000), orthodontics (Motegi et al., 1996; Hoeflinger, 1996), physiology (Tikuisis et al., 2001), plastic surgery (Okada, 2001) and biomechanics (Ronsky et al., 1999; Savatier et al., 2001). The common characteristic of the existing techniques applied in medical applications is the required high accuracy.

To measure the surface of human body parts, four different optical methods can be used: laser scanning, silhouette extraction, coded light and photogrammetric approaches. Laser scanning (Brunsman et al., 1997; Okada, 2001; Tikuisis et al., 2001) and structured light methods (Wolf, 1996; Hackenberg et al., 2000; Sitnik and Kujawinska, 2000) are widely used

* Tel.: +41-1-633-3054; fax: +41-1-633-1101.

E-mail address: nicola@geod.baug.ethz.ch (N. D'Apuzzo).

to measure objects in macro scale. They have common characteristics in the acquisition and processing. Their accuracy and simplicity in use and the wide range of software packages available for the processing, editing and modelling have made them the most widely used system for surface measurement. However, the acquisition time ranges from a couple of seconds to half minute, during which a person must remain stationary. This fact not only places a burden on the subject, but also makes it difficult to obtain stable measurement results because the person always moves slightly during the data acquisition (Daanen et al., 1997). Moreover, these techniques cannot be used for the measurement of moving surfaces, which is mostly the case in medical applications.

The silhouette extraction based methods (Zheng, 1994; Matsumoto et al., 1999; Savatier et al., 2001) are usually employed for full body measurements. Multiple cameras, a single moving camera or a single camera combined with a rotating platform are used to acquire a set of images around a person. The images are processed to extract the silhouettes, which are then combined to result in a 3-D volumetric model. The system based on multiple cameras can measure dynamic events, however, high accuracy cannot be achieved.

The photogrammetric solutions (Banda et al., 1992; Minakuchi et al., 1995; D'Apuzzo, 2002) offer instead accurate surface measurement by the simultaneous acquisition of images from different directions. A thorough calibration of the camera system and an accurate establishment of image correspondences are required to achieve high accuracy.

In this paper, a photogrammetric process for the surface measurement and tracking from multi-image video sequences is presented. Due to the simultaneous acquisition of all the required data, the proposed method offers the opportunity to measure dynamic events and to extract from the same data both 3-D shape and motion information.

At first, the system for data acquisition and the methods used for its calibration are described. Then the procedures for the measurement and tracking of human body parts are depicted. Finally, two applications are presented to prove the functionality of the proposed method.

2. Data acquisition and calibration

A sequence of images acquired simultaneously from different directions (*multi-images*) constitutes the input of the method presented in this paper. Different camera systems can be used to acquire multi-image sequences; however, at least three cameras and a precise electronic synchronisation are required for a robust and accurate measurement. Machine vision interlaced CCD cameras are the cheapest solution. A complete system can be built with three synchronised cameras and a single inexpensive frame grabber, acquiring image sequences using the three RGB-colour inputs. These cameras have, however, the disadvantage of acquiring odd and even lines in different times, resulting in the typical saw pattern effect if the camera or objects in the recorded scene move during the acquisition. To overcome this problem, only the odd (or even) lines have to be processed at the cost of reducing the resolution in vertical direction by 50%.

The highest accuracy is achieved using synchronised progressive scan CCD cameras. They differ from the cheaper interlaced CCD cameras in the way that a full frame is acquired in one pass, also offering full resolution of the sensor when the camera moves or when it records moving objects. Obviously, these cameras are more expensive and special frame grabbers are required.

For the development of our system, we used three synchronised progressive scan CCD cameras with a resolution of 640×480 pixels for full body tracking (see Section 5.2) and five interlaced CCD cameras with a resolution of 768×576 pixels for human face measurement (see Section 5.1).

One of two different methods can be chosen to calibrate the camera system: the reference bar method (Maas, 1998) is used, when the object space viewed by the camera is very large; and the reference field method, for the other cases. In the first method, a reference bar with two retroreflective target points is moved in the object space and at each location, multi-images are acquired. The image coordinates of the two target points are automatically measured and tracked during the sequence with a least squares matching based process (see Fig. 1); the multiple camera system could then be calibrated by self-calibrating bundle adjustment with the additional information of the known distance between the two points

Fig. 1. Automatically tracked and measured image coordinates of the two points on the reference bar.

at each location. The main advantage of this method is that it only requires a reference bar, while a disadvantage is the thorough data processing to calibrate the system.

For the reference field calibration method, a 3-D field with signalised points whose coordinates in space are known is used. In this case, the processing is simpler and full automation can be achieved using

Fig. 2. 3-D calibration frame with coded targets.

coded target points (Fig. 2), which are fully automatically recognised and measured in the images (Niederoest, 1996).

The results of both calibration processes are the exterior orientation of the cameras (position and rotations), the interior orientation parameters (camera constant, principle point, sensor size, pixel size), the parameters describing the symmetric radial and decentering distortion of the lenses and two additional parameters modelling differential scaling and shearing effects (Brown, 1971). A thorough determination of these parameters is required to achieve high accuracy in the measurement.

3. Surface measurement

The approach for surface measurement is based on multi-image photogrammetry using images acquired simultaneously. A dense set of corresponding points in the images is established by a matching process and the 3-D coordinates of the matched points are computed. The measured surface results, therefore, as a 3-D point cloud.

The multi-image matching process (D'Apuzzo, 2002) is based on the adaptive least squares method (Gruen, 1985) with the additional geometrical constraint of the matched point to lie on the epipolar line. The least squares matching algorithm (LSM) considers image patches around the selected points; one image is used as a template and the others as search images. The main idea of LSM is to leave the patches in the template image unchanged and to modify the patches in the search images by an affine transformation (translation, scale, rotation and shearing); a least

squares process determines the parameters of the affine transformation minimising the sum of the squares of the differences between the grey values in the patches. The additional epipolar constraint increases the robustness of the result. Fig. 3 shows an example: the black boxes represent the patches selected in the template image (left) and the affinely transformed in the search images (centre and right), the epipolar lines are drawn in white.

The least squares process cannot converge correctly if the contrast in the patches is not sufficient. In this case, enhancement algorithms are applied to increase the contrast of the original images (see Section 5.1). Or, for surface measurement applications, an artificial texture can be projected onto the surface (see Section 5.2); obviously, this method cannot be used for tracking applications because the projected texture is not linked with the surface and does not move with it.

The developed matching process based on LSM automatically produces a dense and robust set of corresponding points starting from a few seed points. Depending on the case, the seed points may be generated fully automatically or selected manually in one image. The manual mode is the most convenient for static surface measurements: it is fast but leaves the operator the choice of the seed points. In this mode, the seed points have to be selected manually in the template image; the corresponding points in the other images are then established automatically by searching for the best matching results along the epipolar line. Full automatic mode is useful for dynamic surface measurement and tracking processes, where the number of multi-image sets to be processed can be very large. In this case, Foerstner interest
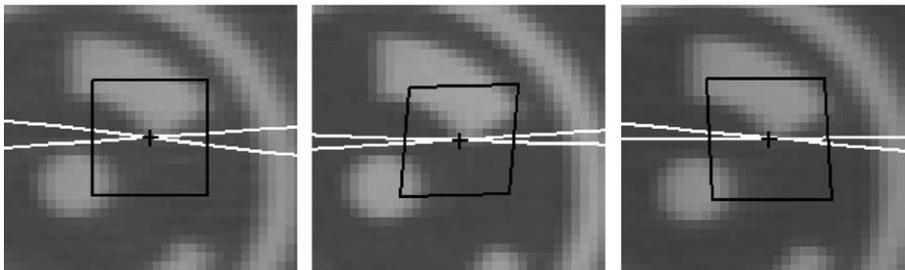

Fig. 3. Geometrical constrained LSM. The black boxes represent the patches selected in the template image (left) and the affinely transformed in the search images (centre and right), the epipolar lines are drawn in white.
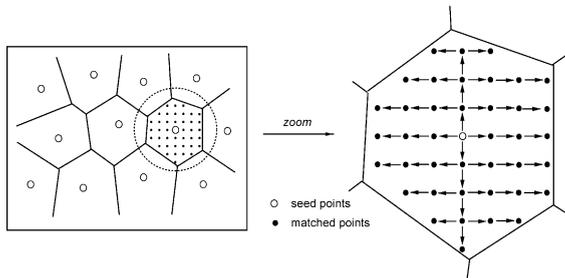
Fig. 4. Search strategy for the matching process. Left: Voronoi tessellation. Right: starting from the seed points, each region is covered by sequential horizontal and vertical shifts.

operator (Foerstner and Guelch, 1987) is applied on the template image to automatically determine marking points where the matching process may perform robustly; the corresponding points in the other images are then established by searching for the best matching results along the epipolar line.

After the definition of the seed points, the template image is divided into polygonal regions according to which of the seed point is the closest (Voronoi tessellation). Starting from the seed points, the automatic matcher produces a dense set of corresponding points in each region by sequential horizontal and vertical shifts (Fig. 4).

The process works adaptively at each shift, changing some parameters (e.g. smaller shift, bigger size of the patch) if the quality of the match is not satisfactory. Several indicators are used to define the match quality: a posteriori standard deviation of the least squares adjustment, standard deviation in x and y

directions, displacement from the start position in x and y direction and distance to the epipolar line. Thresholds for these values are defined according to the texture and the type of the images.

The coverage of the entire image is achieved by repeating the process for all the polygonal seed points regions. At the end of the process, holes of non-analysed areas can occur in the set of matched points. The process tries to match the missing points by searching from all directions around the holes.

With the proposed strategy, the time required by the matching process is short; to give an example, on a Pentium III 600 MHz machine, about 25,000 points are matched in a triplet in less than 10 min.

The 3-D coordinates of the matched points are then computed by forward ray intersection using the orientation and calibration data of the cameras. A simple centre of gravity filter is applied to reduce the noise in the data and to get a uniform density in the point cloud.

The surface measurement is a stand-alone procedure but it also constitutes the base for the surface tracking process described in the next section. In this case, the human body surface is measured at each time step of the acquired multi-image sequence.

4. Surface tracking

4.1. Tracking single points

The basic idea is to track corresponding points in the multi-images through the sequence and compute
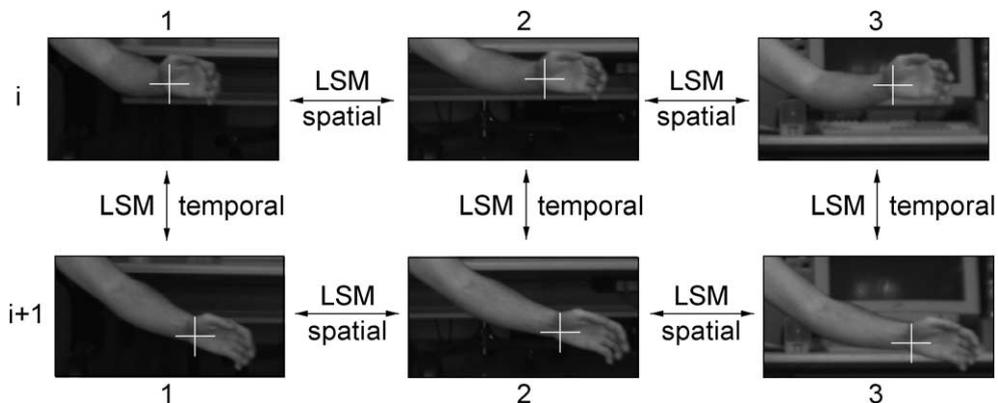

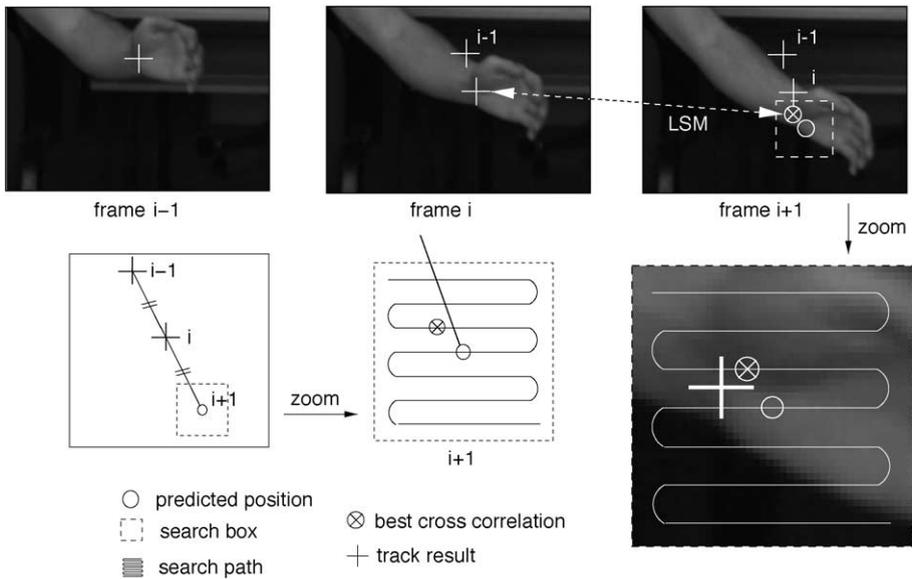Fig. 5. LSM tracking: temporal and spatial correspondences are established with LSM.

Fig. 6. Tracking in image space: temporal LSM is applied at the position of best cross-correlation.

their 3-D trajectories. The process is based on least squares matching techniques (LSM): the *spatial* correspondences between the images of the different views and the *temporal* correspondences between subsequent frames are computed using the least squares matching algorithm (Fig. 5).
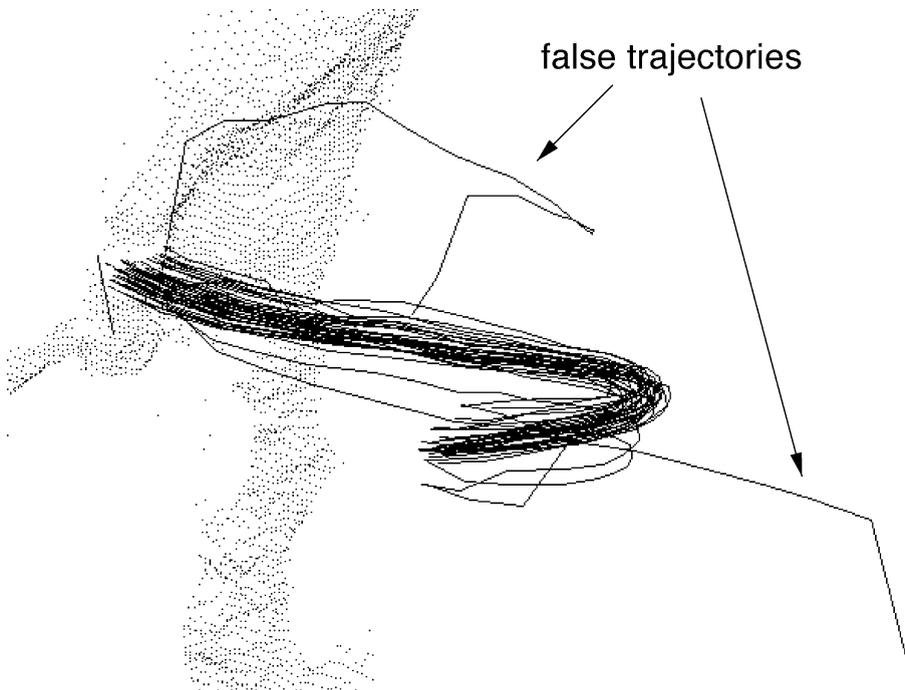

Fig. 7. Surface tracking: the false trajectories do not follow the uniform movement.
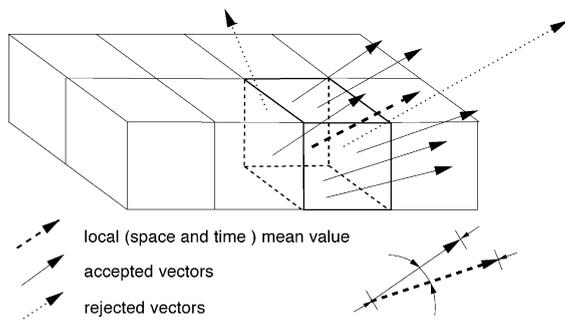
Fig. 8. Uniformity filter: the single vectors are compared in magnitude and direction with the mean vector.

The process starts matching corresponding points in the multi-images and continues with the tracking loop: (1) predict the position in the next frame, (2) search the position with the best cross-correlation value and (3) establish the point in the next frames using least squares matching (*temporal LSM*). Fig. 6 illustrates the process.

The position of the tracked point at time $i+1$ is linearly predicted from the two previous times $i-1$ and i (step 1). A search box is defined around this predicted position in the frame at time $i+1$ and is scanned for the position which has the best value of cross-correlation between the image of frame at time i
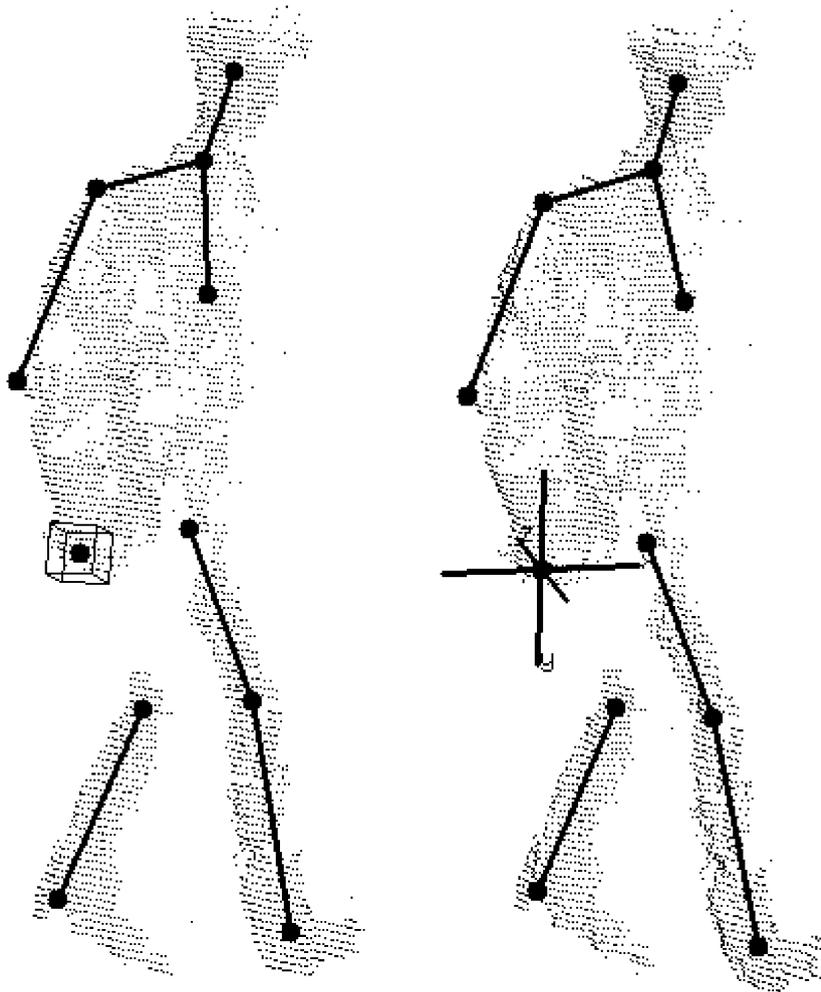

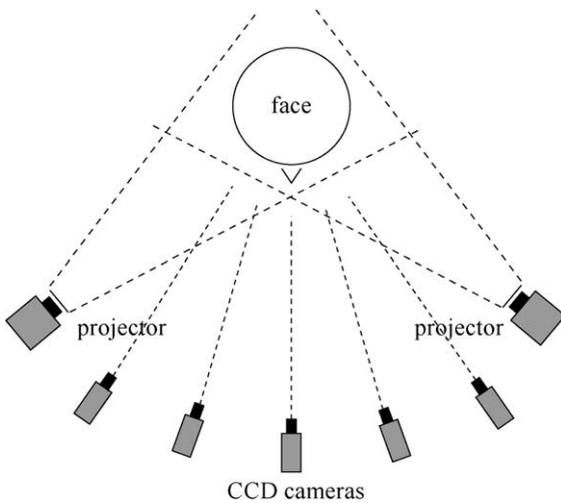Fig. 9. Details of a key-point definition. Left: definition of the size, right: definition of the position in space.

Fig. 10. Setup of cameras and projectors.

poral LSMs and if no significant differences occur between the two results, the point is considered tracked and the process can continue to the next time step. If the differences are too large, step (2) of the process is repeated by searching the value of best cross-correlation in a bigger region around the predicted position. If the new result is also rejected, the tracking process stops.

The result of the tracking process on a single point are its coordinates in the multi-images through the sequence, thus the 3-D coordinates of the point for each time step can be computed by forward ray intersection and its 3-D trajectory, velocity and acceleration is determined.

The proposed process can be used to track well-defined points on the human body. Trajectories of single points are however not sufficient to understand and record the motion and movement of a human or the surface changes of a body part. The next two subsections propose a method to extract more complex information from the image sequence.

4.2. Tracking surface parts

Requirement for the surface tracking process is the measurement of the surface parts of interest for each

and the image of frame at time $i+1$ (step 2). The least squares matching algorithm is applied at that position and the result can be considered the exact position of the tracked point in the new frame (step 3).

This process is performed in parallel for the different images. To test the individual results, *spatial LSM* is computed at the positions resulting from the tem-
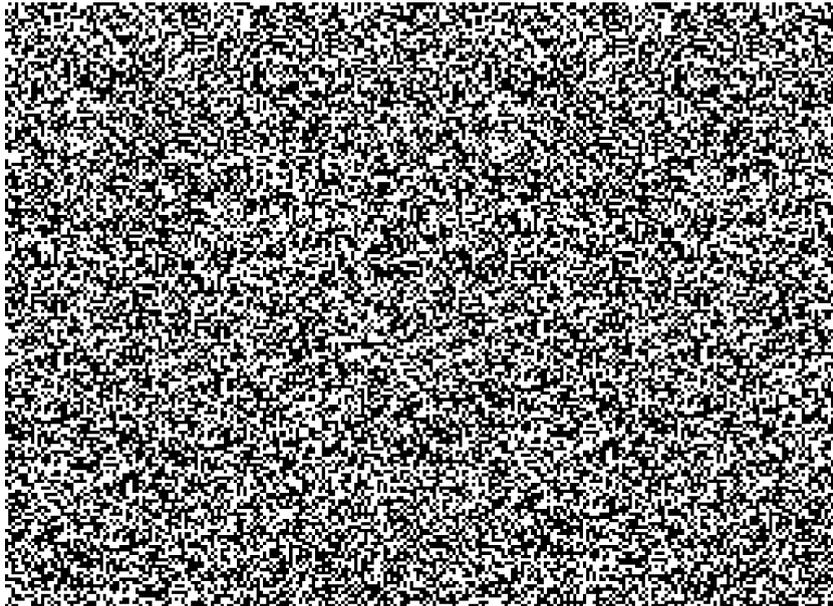

Fig. 11. Projected random pattern.

Fig. 12. Multi-images of a face with random pattern projection.

multi-image of the sequence. This can be achieved automatically with the process described in Section 3. The tracking process of Section 4.1 is then applied to all the points measured on the surface, resulting in a vector field of trajectories.

With this approach, a new problem has to be considered: during the sequence, some surface parts can get lost by occlusion and new parts of the surface can appear. For this reason, a new functionality has to be integrated in the tracking loop: before passing to the next time step, the density of the data resulted

from the tracking process is checked with a threshold. In the regions of low density, new points are imported from the previously computed data (surface measurement for each time step). In this way, new appearing surface parts or lost points are integrated in the tracking process.

In case of poor texture of the surface, the tracking process can produce false trajectories. An example is shown in Fig. 7, which displays the computed 3-D trajectories for a part of the arm of a walking person.
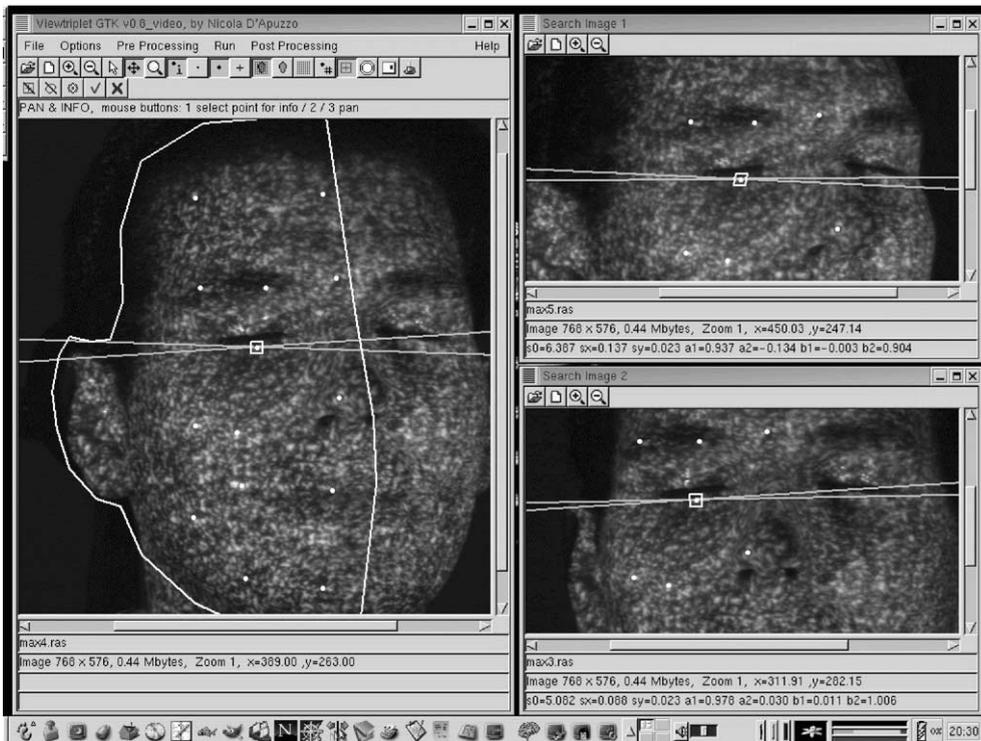

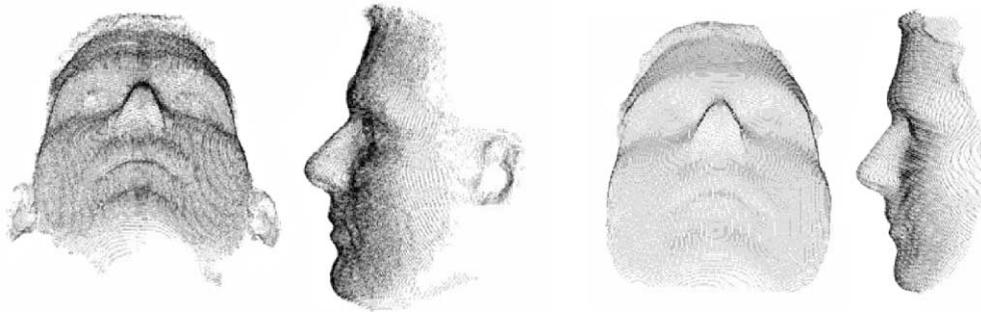Fig. 13. Graphical user interface of the face measurement software.

Fig. 14. Left: measured 3-D point cloud (45,000 points). Right: after filtering and thinning (10,000 points).

The false trajectories can easily be recognised because they do not follow the common movement of the majority. The vector field of trajectories can indeed be checked for local uniformity of the movement. Two filters are applied to remove or truncate false trajectories. The first one removes larger errors using thresholds of the velocity and acceleration. Depending on the movement performed, the two thresholds are defined at the beginning of the process and remain constant during the sequence. The second filter checks for the local uniformity of the movement both in space and time. To check this property, the space is divided in regular voxels and local mean values of the velocity vectors are computed in each voxel for each time step. The single trajectories are then compared to the local mean values and truncated or removed if the differences in magnitude and direction are too large (see Fig. 8).

After filtering the trajectories, a problem can still remain. Ideal trajectories start from the beginning of the sequence and last till the end. However, in some cases, depending on the quality of the image sequence

and on the type of surface, the result of the tracking process is a set of broken trajectories with varying length. This effect, particularly strong when measuring full human movements, is mainly caused by occlusion and lack of texture. To solve this problem, the concept of a *key-point* is introduced.

4.3. Tracking key-points

A key-point is a 3-D region defined in the vector field of trajectories, whose size can vary and whose position is defined by its centre of gravity. Key-points are interactively defined in a graphical user interface. They can be easily placed and moved in 3-D space (see Fig. 9).

The key-points are tracked in a simple way: the position in the next time step is established by the mean value of the displacement of all the trajectories inside its region. The size of the region to be tracked has an important role: if a small size is chosen, the key-point can be considered with good approximation to lie on the surface and to represent

Fig. 15. Photorealistic visualisation. Left: shaded surface model, texture image. Right: face model with texture.
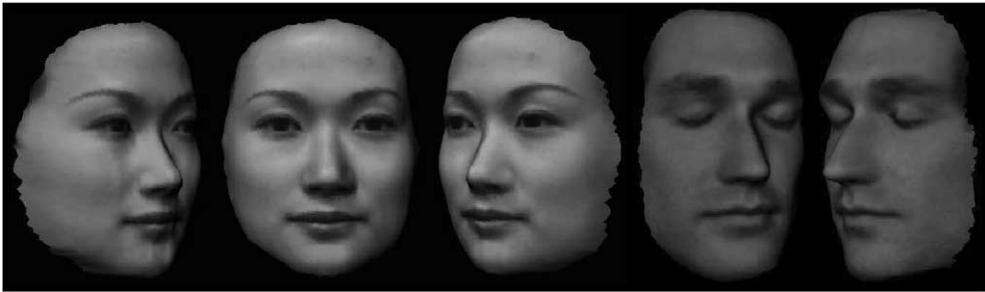
Fig. 16. Photorealistic visualisation: two other examples of face models.

fictitious markers on the human body. If, on the other hand, larger sizes are chosen, the key-point represents a more approximative position of, e.g. joints of the human body. An example is given in Section 5.2.

5. Applications

In this section, a static and a dynamic application of the described method are presented.

5.1. Static surface measurement: human face modelling

The extensively employed methods to produce 3-D computer models of the human face are laser scanning and coded light-based triangulation approach. The advantage of our method over these two techniques is the acquisition of the source data in a fraction of

seconds, allowing the measurement of human faces with higher accuracy and even the possibility to measure dynamic events like the speech of a person. Moreover, the developed software can be run on a normal PC, reducing the costs of the hardware.

In the past, we have collaborated in a project intended to evaluate the anatomical changes occurring with maxillo-facial and plastic surgery (Koch et al., 1996; D'Apuzzo, 1998). What we are developing now is a portable, cheap and accurate system for the measurement and modelling of human faces (D'Apuzzo, 2002).

Fig. 10 shows the setup of the acquisition system. It consists of five CCD cameras and two slide projectors arranged convergently in front of the subject. The cameras are connected to a frame grabber which digitises the five images at the resolution of 768×576 pixels with 8-bit quantisation. The system is calibrated using a 3-D reference frame with coded target points (see Section 2; Fig. 2).

Fig. 17. A frame of an acquired sequence (courtesy of the Computer Graphics Lab, EPF Lausanne). Left: original image, right: after Wallis filter contrast enhancement.

Fig. 18. A triplet of the processed multi-image sequence.

Since the natural texture of the human skin is relatively uniform, the projection of an artificial texture onto the face is required to perform robustly the matching process. A random pattern (see Fig. 11) is preferred over regular patterns to avoid possible mismatches and its resolution has to be fine enough to result in the images in structures the size of few pixels. The use of two projectors enables a focused texture even on the lateral sides of the face; Fig. 12 shows the five images acquired by the CCD cameras.

A colour image of the face without artificial texture is acquired by an additional colour video camera placed in front of the subject. It will be used for the realisation of a photorealistic visualisation.

A dedicated software with a user-friendly graphical interface (Fig. 13) was developed for the face measurement process. The intervention of the operator for the matching process is reduced to the interactive definition of the seed points (about 10) and the selection of a contour of the region to measure. This operation can be performed in a couple of minutes, then the process will continue completely automatically.

Since the human face is a steep surface and both sides of the face are not visible to the same camera, the five acquired images are divided into two triplets, one for each side of the face. They are processed separately and the results are then merged into a single data set.

The achieved mean accuracy of the computed 3-D points is about 0.3 mm in the sagittal direction and about 0.2 mm in the lateral direction.

As can be seen in Fig. 14 (left), the point cloud is very dense (45,000 points), with an overlapping area of the two data sets in the centre. To overcome the redundant data and remove the outliers, Gaussian filters (Borghese and Ferrari, 2000) and thinning processes are applied to the 3-D point cloud (Fig. 14 right).

A meshed surface is generated from the 3-D point cloud by 2.5-D Delauney triangulation. In order to achieve photorealistic visualisation, the natural texture acquired by the colour video camera is draped over the model of the face. Fig. 15 shows the surface model, the texture image and two views of the
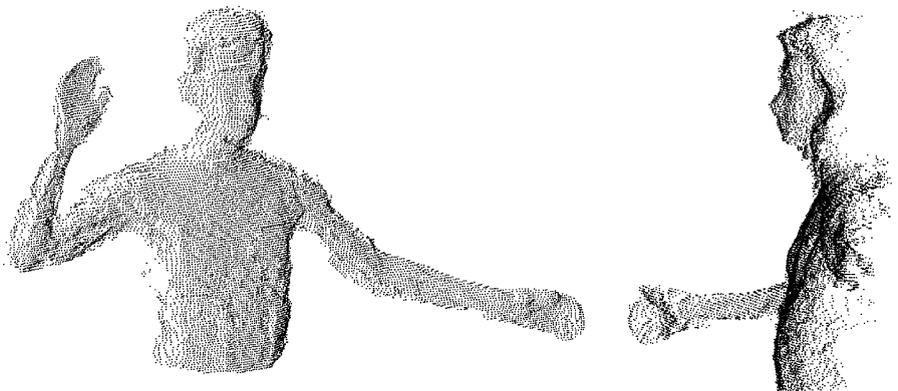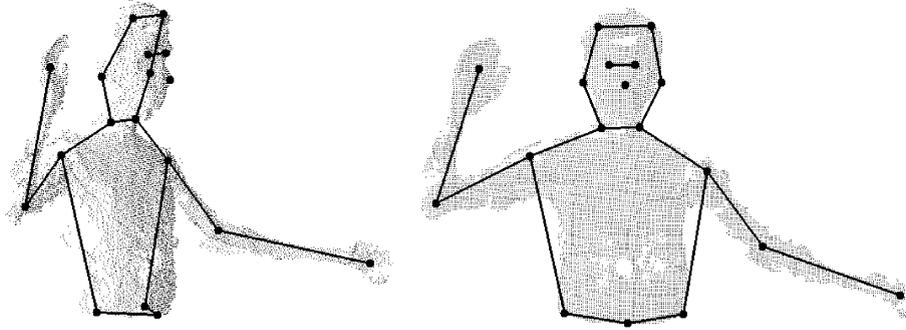

Fig. 19. Measured 3-D point cloud.

Fig. 20. Key-points tracked on the human body.

resulting face model with texture; Fig. 16 shows two other examples of face models.

The results achieved by the presented method are excellent and comparable with the computer models produced by the more popular but also more expensive laser scanning systems. The system implemented for PC on Linux platform is operational and we are evaluating applications in different fields.

5.2. Dynamic surface measurement/surface tracking: motion capture

In this section, the presented method is applied to extract, without using markers, 3-D full body motion information from multi-image sequences acquired with three synchronised CCD cameras. This work is part of a project aimed at developing a highly automated system to accurately model human bodies from video sequences (D'Apuzzo et al., 2000).

Three synchronised progressive scan CCD cameras in a triangular arrangement are used. A sequence of triplets is acquired with a frame grabber and the images are stored with 640×480 pixels resolution at 8-bit quantisation.

Fig. 17 shows a frame of a sequence where a person performs some complex movement of the upper part of the body; as can be seen, no special conditions nor signalised points are required. In order to increase the contrast, the images are enhanced using Wallis filtering (Wallis, 1976).

The surface of the visible parts of the human body is measured for each time step of the sequence in the
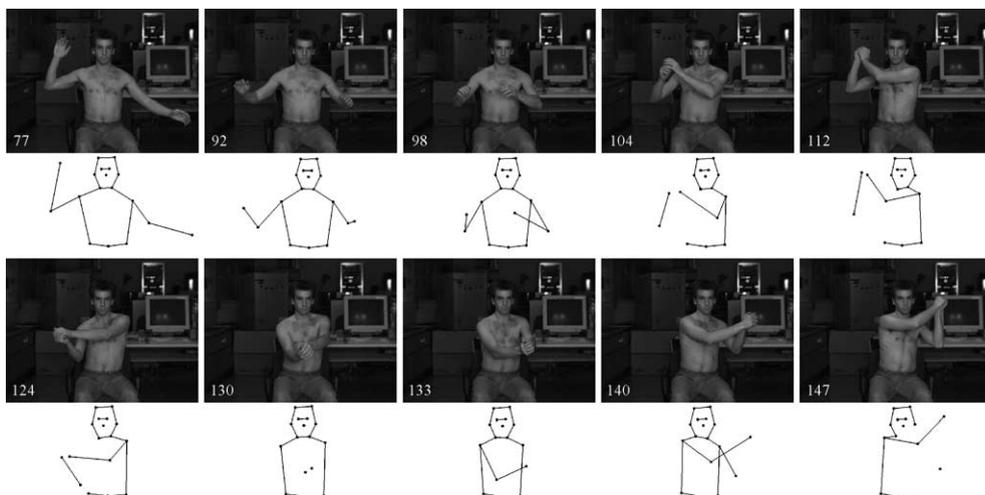

Fig. 21. Tracking key-points: some frames of the sequence and frontal view of the position of the tracked key-points.

way explained in Section 3. The amount of triplets to be processed is very large, therefore, full automatic mode is used for the definition of the seed points. The manual intervention of the operator is reduced to the definition of the contour of the region to measure. This procedure will be soon automated, extracting from the images the silhouette of the person (Haritaoglu et al., 1998; Nef, 2001).

The result of the surface measurement process is a 3-D point cloud for each time step; this data will be used in the successive tracking process. Fig. 18 shows a triplet of the sequence and Fig. 19 shows its computed 3-D point cloud.

The surface tracking process applied to the test sequence of Fig. 17 provides about 25,000 trajectories during 70 frames. Due to the poor natural texture and self-occlusion caused by the complex movement, the 3-D trajectories have an average length of nine frames, which is quite short. The process of key-point tracking is therefore required to extract general 3-D information of the performed movement.

The key-points are placed in such a way that their trajectories can describe the complex movement performed by the person. A small set is placed near the joints: hands, elbows, shoulders, sides of neck; three are placed on the bottom of the upper body part; three on the face (nose and eyes) and four on the head (see Fig. 20).

Because of the complexity of the performed movement of this sequence, the size of the region defining the key-points is chosen to be large enough to guarantee robustness. The key-points represent, in this case, an approximation for the joint trajectories. The final result of the tracking process is shown in Fig. 21 together with the corresponding frames; two views of the 3-D trajectories of the key-points are displayed in

Fig. 22 (note that the connections between key-points are intended only for a better understanding of the images).

As can be seen, e.g. in frame number 130 of Fig. 21, some key-points cannot be tracked during the entire sequence because of strong occlusions caused by the complex movement. A possible solution to this problem is the use of more than three cameras, acquiring multi-images all around the person.

6. Conclusions

A process for an automated extraction of 3-D data from multi-image sequences has been presented. The gained 3-D data can be of two different types: surface measurement of human body parts at each time step of the sequence or surface tracking in form of a vector field of 3-D trajectories (position, velocity and acceleration).

To demonstrate the functionality of the method, two applications were presented: high accuracy measurement of human faces and full body motion capture without using markers.

As far as the static surface measurement is concerned, the quality of the final results is excellent. Moreover, manual intervention is reduced to a minimum and the processing time is also short. Therefore, it can be concluded that the presented method is a valid alternative to laser scanning techniques and coded light-based triangulation approaches with an important reduction of the hardware costs.

The dynamic surface measurement and tracking procedure are part of a pilot project still under development. Its final goal is a fully automatic system for the realistic modelling of human bodies from video
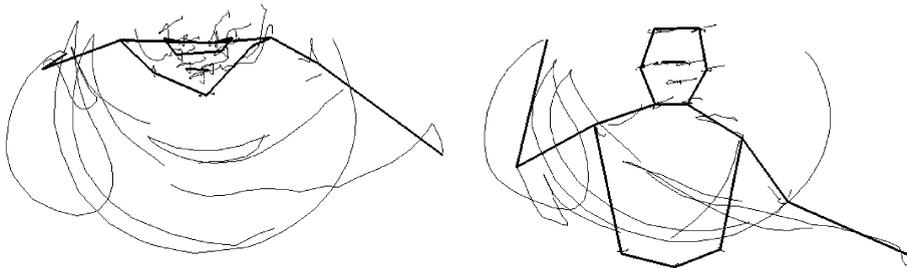

Fig. 22. View from the top (left) and from the front (right) of the 3-D trajectories of the key-points.

sequences. A lot of work still remains in the future to increase the automation level and to improve the quality of the extracted 3-D data. Moreover, the use of only three CCD cameras limits the measurement and tracking to only one side of the human body. For very complex movements, the body parts have to be imaged both in front as well as sideways and from the back, therefore the acquisition system has to be extended. Other important cues that could be determined from the vector field of trajectories resulting from the tracking process are all the rotations angles, which are required to precisely describe the movement of articulations such as shoulder–elbow–hand.

Acknowledgements

References

Banda, F.A.S., Muller, J.-P., Bhatia, S.N., Bukhary, M., 1992. Automatic generation of facial DEMs. International Archives of Photogrammetry and Remote Sensing 29 (B5), 893–896.

Borghese, A., Ferrari, S., 2000. A portable modular system for automatic acquisition of 3-D objects. IEEE Transactions on Instrumentation and Measurement 49 (5), 1128–1136.

Brown, D.C., 1971. Close-range camera calibration. Photogrammetric Engineering and Remote Sensing 37 (8), 855–866.

Brunsman, M.A., Daanen, H., Robinette, K.M., 1997. Optimal postures and positioning for human body scanning. Proc. Int. Conf. on Recent Advances in 3-D Digital Imaging and Modeling, 12–15 May, Ottawa, Canada, 266–273.

Daanen, H.A.M., Taylor, S.E., Brunsman, M.A., Nurre, J.H., 1997. Absolute accuracy of the Cyberware WB4 whole body scanner. In: Ellson, R.N., Nurre, J.H. (Eds.), Three-Dimensional Image Capture, Proceedings of SPIE, San Jose, USA, vol. 3023, pp. 6–12.

D'Apuzzo, N., 1998. Automated photogrammetric measurement of human faces. International Archives of Photogrammetry and Remote Sensing 32 (5), 402–407.

D'Apuzzo, N., 2002. Modeling human faces with multi-image photogrammetry. In: Corner, B.D., Pargas, R., Nurre, J.H. (Eds.), Three-Dimensional Image Capture and Applications V. Proceedings of SPIE, San Jose, USA, vol. 4661, pp. 191–197.

D'Apuzzo, N., Plaenkers, R., Fua, P., 2000. Least squares matching tracking algorithm for human body modeling. International Archives of Photogrammetry and Remote Sensing 33 (B5/2), 714–720.

Foerstner, W., Guelch, E., 1987. A fast operator for detection and precise location of distinct points, corners and centres of circular features. Proceedings of ISPRS Intercommission Conference on Fast Processing of Photogrammetric Data, June 2–4, Interlaken, Switzerland, 281–305.

Gruen, A., 1985. Adaptive least squares correlation: a powerful image matching technique. South African Journal of Photogrammetry, Remote Sensing and Cartography 14 (3), 175–187.

Hackenberg, L., Liljenqvist, U., Hierholzer, E., Halm, H., 2000. Scanning stereographic surface measurement in idiopathic scoliosis after VDS (ventral derotation spondylodesis). Zeitschrift fur Orthopadie und ihre Grenzgebiete 138 (4), 353–359.

Haritaoglu, I., Harwood, D., Davis, L.S., 1998. W4: Who? When? Where? What? A real time system for detecting and tracking people. Proceedings 3rd International Conference on Face and Gesture Recognition, April 14–16, Nara, Japan, pp. 222–227.

Hoeflinger, W., 1996. Digital stereophotogrammetric solutions for orthodontics. International Archives of Photogrammetry and Remote Sensing 31 (B5), 247–252.

Koch, R.M., Gross, M.H., Carls, F.R., von Bueren, D.F., Fankhauser, G., Parish, Y.I.H., 1996. Simulating facial surgery using finite element models. Proceedings SIGGRAPH '96, August 4–9, New Orleans, USA, pp. 421–428.

Maas, H.-G., 1998. Image sequence based automatic multi-camera system calibration techniques. International Archives of Photogrammetry and Remote Sensing 32 (5), 763–768.

Matsumoto, Y., Fujimura, K., Kitamura, T., 1999. CyberModeler: a compact 3D scanner based on monoscopic camera. In: Nurre, J.H., Corner, B. (Eds.), Three-Dimensional Image Capture and Applications II. Proceedings of SPIE, San Jose, USA, vol. 3640, pp. 3–10.

Minakuchi, S., Hirano, Y., Sekita, T., Suzuki, T., Kobayashi, K., Nagao, M., 1995. Three-dimensional analysis of lip movement by 3-D auto tracking system. International Archives of Photogrammetry and Remote Sensing 30 (5W1), 385–389.

Motegi, N., Tsutsumi, S., Wakatsuki, E., 1996. A facial growth analysis based on FEM employing three dimensional surface measurement by a rapid laser device. Okajimas Folia Anatomica Japonica 72 (6), 323–328.

Nef, M., 2001. Automatische Silhouetten-Extraktion aus Videosequenzen Semesterarbeit SS01, Institut fuer Geodaesie und Photogrammetrie, ETHZ, Zurich, Switzerland (in German).

Niederoest, M., 1996. Codierte Zielmarken in der digitalen Nahbereichsphotogrammetrie. Diplomarbeit SS96, Institut fuer Geodaesie und Photogrammetrie, ETHZ, Zurich, Switzerland (in German).

Okada, E., 2001. Three-dimensional facial simulations and measurements: changes of facial contour and units associated with facial expression. Journal of Craniofacial Surgery 12 (2), 167–174.

Ono, T., 1995. Trunk deformity in scoliosis studied by surface measurement. Nippon Seikei Geka Gakkai zasshi 69 (10), 915–926.

Ronsky, J.L., Boyd, S.K., Lichti, D.D., Chapman, M.A., Salkauskas,

K., 1999. Precise measurement of cat patellofemoral joint surface geometry with multistation digital photogrammetry. Journal of Biomechanical Engineering 121 (2), 196–205.

Savatier, X., Ertaud, J.Y., Dujardin, F.H., 2001. Three-dimensional multivision determination of human body segment characteristics. Proc. 18th Congress of the International Society of Biomechanics, 8–13 July, Zurich, Switzerland., CDROM paper 0011.

Sitnik, R., Kujawinska, M., 2000. Opto-numerical methods of data acquisition for computer graphics and animation systems. In: Corner, B.D., Nurre, J.H. (Eds.), Three-Dimensional Image Capture and Applications III, Proceedings of SPIE, San Jose, USA, vol. 3958, pp. 36–43.

Tikuisis, P., Meunier, P., Jubenville, C.E., 2001. Human body surface area: measurement and prediction using three dimensional body scans. European Journal of Applied Physiology and Occupational Physiology 85 (3–4), 264–271.

Wallis, R., 1976. An approach to the space variant restoration and enhancement of images. Proceedings of the Symposium on Current Mathematical Problems in Image Science. Naval Postgraduate School, Monterey, USA, pp. 329–340.

Wolf, H.G.E., 1996. Structured lighting for upgrading 2D-vision system to 3D. Proceedings International Symposium on Laser, Optics and Vision for Productivity and Manufacturing I, June 10–14, Besancon, France, pp. 10–14.

Zheng, J.Y., 1994. Acquiring 3-D models from sequences of contours. IEEE Transactions on Pattern Analysis and Machine Intelligence 16 (2), 163–178.