International Archives of Photogrammetry and Remote Sensing, 34(5), pp. 363-368, Corfu (Greece), September 2002

IMAGE SEQUENCE ANALYSIS FOR HUMAN BODY RECONSTRUCTION

Fabio Remondino

Institute for Geodesy and Photogrammetry, ETH Zurich, Switzerland
E-mail: fabio@geod.baug.ethz.ch

Commission V, ICWG V/III

KEY WORDS: Camera Calibration, Least Squares Matching, Reconstruction

ABSTRACT

The generation of 3-D models from uncalibrated image sequences is a challenging problem that has been investigated in many research activities in the last decade. In particular, a topic of great interest is the modeling of realistic humans, for animation, manufacture or medicine purposes. Nowadays the common approaches try to reconstruct the human body using specialized hardware (laser scanners) resulting in high costs. In this paper a different method for the three-dimensional reconstruction of human bodies from image sequences acquired with a standard video-camera is presented. The core of the presented work describes the calibration and orientation of the images but the whole process includes also the extraction of correspondences on the body using least squares matching and the reconstruction of the 3-D body model.

1. INTRODUCTION

The actual interests in 3-D object reconstruction are motivated by a wide spectrum of applications, such as object recognition, city modeling, video games, animations, surveillance and visualization. In the last years, great progress in creating and visualizing 3-D models from images has been made, with particular attention to the visual quality of the results. The existing systems are often built around specialized hardware (e.g. laser scanner), often resulting in high costs. Other methods based on photogrammetry [Grün et al., 2001; Remondino, 2002] or computer vision [Pollefeys, 2000], can instead obtain 3-D models of objects with low cost acquisition systems, using photo or video cameras. Since many years, photogrammetry deals with high accuracy measurements from image sequences, including 3-D object tracking [Maas, 1991], deformation measurements or motion analysis [D'Apuzzo et al., 2000]; even if these applications require very precise calibration, automated and reliable procedures are available.

Concerning the reconstruction and modeling of human bodies, nowadays the demand for 3-D models has drastically increased. A complete model of a human consists of both the shape and the movements of the body. These two modeling processes are often considered as separate even if they are very close. A classical approach to build human shape models uses 3-D scanners [Cyberware, 2002, Vitus, 2002, Horiguchi, 1998]: they are expensive but simple to use and software are available to edit and model the obtained point cloud. Other techniques use structured light methods [Wolf, 1996], silhouette extraction [Zheng, 1994], multi-image photogrammetry [D'Apuzzo, 2002]. The human body models can be used in different fields, like animation, manufacturing or medicine. For animation purpose, only approximative measurements are necessary: the shape can be first defined (e.g. smoothing 3-D mesh with splines, attaching generalized cylinders or volumetric primitives to a skeleton) and then animated using motion capture data. For medical applications or in manufacture industries, digital surfaces are required for metric body information and design of clothes [McKenna, 1996]; therefore exact 3-D models of the body are needed and usually performed with scanning devices [Tailor, 2002].

In this paper a photogrammetric approach for the reconstruction of 3-D models of static humans from uncalibrated image sequences is described. The process consists of three parts:

1) Acquisition and analysis of the image sequence (section 2)
2) Calibration and orientation of the images (section 3)
3) Matching process on the human body surface and point cloud generation (section 4).

This work belongs to a project called Characters Animation and Understanding from SEquence of images (CAUSE). Its goal is the extraction of complete 3-D animation models of characters from old movies or video sequences, where no information about the cameras and the objects are available.

2. IMAGE ACQUISITION

The images can be acquired with a still-video camera or with a camcorder. A complete reconstruction of the human body requires a 360 degrees azimuth coverage, while, for the time being, only frames in front of the body are acquired. The acquisition lasts ca. 30 seconds and requires no movements of the person. This could be considered a limit of the procedure but also 3-D scanners need at least 15 seconds to acquire a full body model. Figure 1 shows three images (out of 6) of a sequence acquired with a Sony DSC-S70, with a resolution of 768x1024 pixels. During the acquisition, the camera constant was kept fixed not to deal with varying camera constant. If a video camera is used (section 5), the acquired video has to be digitalized and the artefacts created by interlace effects must be removed.

3. CALIBRATION AND ORIENTATION OF THE IMAGES

Camera calibration and image orientation are prerequisites for accurate and reliable results, in particular for those applications that rely on the extraction of precise 3-D information from imagery. The early theories and formulations of orientation procedures were developed in the first half of the 19th century and today a great number of procedures and algorithms is available. A fundamental criterion for grouping the orientation procedures is based on the used camera model, i.e. the projective camera model or the perspective camera one. Camera

models based on perspective collineation have high stability, require a minimum of three corresponding points per image and a stable optics. On the other hand, projective approaches can deal with variable focal length, but need more parameters, a minimum of six corresponding points and are quite instable (equations need normalization).
The calibration and orientation process is the core of the presented work and is based on a photogrammetric bundle-adjustment (section 3.3); the required tie points (image correspondences) are found automatically with the following steps:

- interest points extraction from each image;
- matching of potential feature pairs between adjacent images;
- false matches clearing using local filtering;
- epipolar geometry computation to refine the matching process and remove any outliers;
- correspondences tracking in all the image sequence.

In the following section these steps are described. The process is completely automated; it is similar to [Fitzigibbon et al., 1998] and [Roth et al., 2000], but additional changes and extensions to these algorithms are presented and discussed.

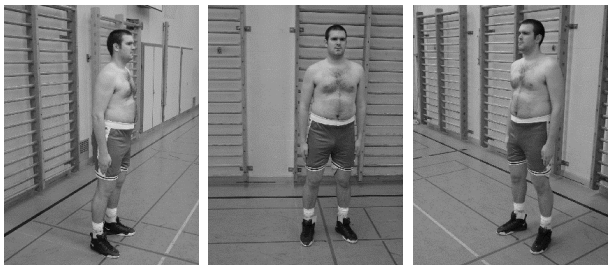

Fig.1: Three images (out of six) used for the reconstruction

3.1 Finding image correspondences

The first step is to find a set of interest points or corners in each image of the sequence. Harris corner detector is used. The threshold on the number of corners extracted is based on the image size. A good point distribution is assured by subdividing the images in small patches and keeping only the points with the highest interest value in those patches.
The next step is to match points between adjacent images. At first cross-correlation is used and then the results are refined using adaptive least square matching (ALSM) [Grün, 1985]. The cross-correlation process uses a small window around each point in the first image and tries to correlate it against all points that are inside a bigger window in the adjacent image. The point with biggest correlation coefficient is used as approximation for the matching process. The process returns the best match in the second image for each interest point in the first image. The final number of possible matches depends on the threshold parameters of the matching and on the disparity between image pairs; usually it is around 40% of the extracted points.
The found matched pairs always contain outliers, due to the unguided matching process. Therefore a filtering of false correspondences has to be performed. A process based on disparity gradient concept is used [Klette et al., 1998]. If $\mathbf{P}_{LEFT}$ and $\mathbf{P}_{RIGHT}$ as well as $\mathbf{Q}_{LEFT}$ and $\mathbf{Q}_{RIGHT}$ are corresponding points in the left and right image, the disparity gradient of two points ($\mathbf{P}$,$\mathbf{Q}$) is the *vector* G defined as:

$$G = \frac{|D(\mathbf{P}) - D(\mathbf{Q})|}{D_{CS}(\mathbf{P},\mathbf{Q})} \qquad [1]$$

where:

$D(\mathbf{P}) = (P_{LEFT,X} - P_{RIGHT,X}, P_{LEFT,Y} - P_{RIGHT,Y})$ is the parallax of $\mathbf{P}$, e.g. the pixel distance of $\mathbf{P}$ between the 2 images**;**
$D(\mathbf{Q}) = (Q_{LEFT,X} - Q_{RIGHT,X}, Q_{LEFT,Y} - Q_{RIGHT,Y})$ is the parallax of $\mathbf{Q}$, e.g. the pixel distance of $\mathbf{Q}$ between the 2 images**;**
$D_{CS} = [(P_{LEFT} + P_{RIGHT})/2, (Q_{LEFT} + Q_{RIGHT})/2]$ is the cyclopean separator, e.g. the difference between the two midpoints of the straight line segment connecting a point in the left image to the corresponding in the right one.

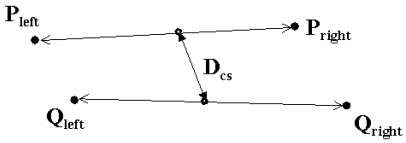

Figure 2: The disparity gradient between two correspondences (P and Q) in image left and right.

If $\mathbf{P}$ and $\mathbf{Q}$ are close together in both images, they should have a similar parallax (e.g. a small numerator in equation 1). Therefore, the smaller the disparity gradient G is, the more the two correspondences are in agreement. This filtering process is performed locally and not on the whole image, because the algorithm can achieve incorrect results due to very different disparity values and in presence of translation, rotation, shearing and scale between consecutive images. The sum of all disparity gradients G of each matched point relative to all other neighbourhood matches is computed. Those matches that have a disparity gradient sum greater than the median of the sums of G are removed. The process removes ca. 80% of the false correspondences. Other possible approaches to remove false matches are described in [Pilu, 1997] and [Zhang et al., 1994].
The next step performs a pairwise relative orientation and an outlier rejection using those matches that pass the previous filtering step. Based on the coplanarity condition, the process computes the projective singular correlation between two images [Niini, 1994], also called epipolar transformation (because it transforms an image point from the first image to an epipolar line in the second image) or fundamental matrix (in case the interior orientation parameters of both images are the same) [Faugeras et al., 1992]. The fundamental matrix $\mathbf{M_{12}}$ is defined by the equation:

$$\mathbf{p}_1^T \ \mathbf{M}_{12} \ \mathbf{p}_2 = 0 \text{ with } \mathbf{p}_i = [x_i \quad y_i \quad 1]^T \qquad [2]$$

for every pair of matching points $\mathbf{p_1}$, $\mathbf{p_2}$ (homogeneous vectors) in image 1 and 2. The epipoles of the images are defined as the right and left null-space of $\mathbf{M_{12}}$ and can be computed with singular value decomposition of $\mathbf{M_{12}}$. A point $\mathbf{p_2}$ in the second image lies on the epipolar line $\mathbf{l}_2$ defined as $\mathbf{l_2} = \mathbf{M_{12}p_1}$ and must satisfy the relation $\mathbf{p_2 l_2} = 0$. Similarly, $\mathbf{l_1} = \mathbf{M_{12}^T p_2}$ represents the epipolar line in the first image corresponding to $\mathbf{p}_2$ in the second image. The 3x3 singular matrix $\mathbf{M}$ can be computed just from image points and at least 8 correspondences are needed to compute it. Many solutions have been published to compute $\mathbf{M}$, but to cope with possible blunders, a robust method of estimation is required. In general least median estimators are very powerful in presence of outliers; so the Least Median of the Squares (LMedS) method is used to achieve a robust computation of the epipolar geometry and to reject possible outliers [Scaioni, 2001; Zhang et al., 1994]. LMedS estimators solve non-linear minimization problems and yield the smallest value for the median of the squared residuals computed for the data set. Therefore they are very robust in case of false matches or outliers due to false localisation.

The computed epipolar geometry is then used to refine the matching process, which is now performed as guided matching along the epipolar lines. A maximal distance from the epipolar line is set as threshold to accept a point as potential match or as outlier. Then the filtering process and the relative orientation are performed again to get rid of other possible blunders. However, while the computed epipolar geometry can be correct, not every correspondence that supports the relative orientation is necessarily valid. This because we are considering just the epipolar geometry between couple of images and a pair of correspondences can support the epipolar geometry by chance (e.g. a repeated pattern aligned with the epipolar line). These kinds of ambiguities and blunders can be reduced considering the epipolar geometry between three consecutive images. A linear representation for the relative orientation of three views is represented by the trifocal tensor **T** [Shashua, 1994]; it is represented by a set of three 3x3 matrices and is computed only with image correspondences without knowledge of the motion or calibration of the cameras. For every triplet of views (Figure 3), if $\mathbf{p}_1$, $\mathbf{p}_2$ and $\mathbf{p}_3$ are corresponding points in the images, then for every line $\mathbf{l}_2$ through $\mathbf{p}_2$ in image 2 and for every line $\mathbf{l}_3$ through $\mathbf{p}_3$ in image 3, the fundamental trifocal constraint states:

$$\mathbf{l}_2^T [\mathbf{T}\mathbf{p}_1] \mathbf{l}_3 = 0 \qquad [3]$$

where $[\mathbf{T}\mathbf{p}_1]$ is a 3x3 matrix whose (i,j) entry is

$$[\mathbf{T}\mathbf{p}_1]_{ij} = T_1^{ij} x_1 + T_2^{ij} y_1 + T_3^{ij} \qquad [4]$$

If we consider only the corresponding points, each triplet $\mathbf{p}_1$, $\mathbf{p}_2$ and $\mathbf{p}_3$ must satisfy the matrix equation:

$$[\mathbf{p}_2]_x [\mathbf{T}\mathbf{p}_1] [\mathbf{p}_3]_x = 0 \qquad [5]$$

with $[\mathbf{p}]_x$ the skew-symmetric matrix of an homogeneous vector, built as

$$[\mathbf{a}]_x = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \qquad [6]$$

where $\mathbf{a} = (a_1, a_2, a_3)^T$.
If a triplet of points $\mathbf{p}_1$, $\mathbf{p}_2$ and $\mathbf{p}_3$ satisfy equation (5), it means that the corresponding points *support* the tensor T_{123} (Figure 4).

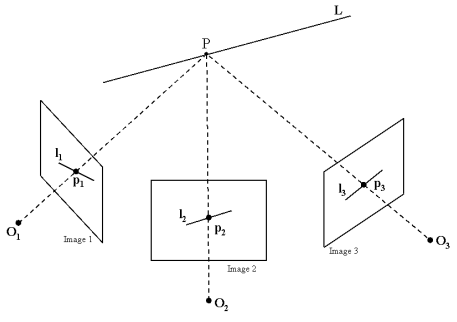

Figure 3: Three views geometry: correspondences $\mathbf{p}_i$ and $\mathbf{l}_i$ corresponding to point P and line L

Relation (5) can be used to verify whether image points (or lines) are correct corresponding features between different views. Moreover, with constraint (5), it is possible to *transfer* points, e.g. compute the image coordinates of a point in the third view, given the corresponding image positions in the first two images. The exterior orientation of the cameras is not required (as with collinearity equations) and only image measurements are needed. This transfer is very useful when in one view are not found many correspondences; calling $\mathbf{p}_1$ and $\mathbf{p}_2$ the point correspondences in the fist two images, the image coordinates of the corresponding point $\mathbf{p}_3$ in the third view are given (up to a non zero scalar factor) by:

$$\begin{aligned} \rho\mathbf{p}_3 &= [\mathbf{T}\mathbf{p}_1]_{1*}^T - x_2 [\mathbf{T}\mathbf{p}_1]_{3*}^T \quad \text{or} \\ \tau\mathbf{p}_3 &= [\mathbf{T}\mathbf{p}_1]_{2*}^T - y_2 [\mathbf{T}\mathbf{p}_1]_{3*}^T \end{aligned} \qquad [7]$$

where:
$[\mathbf{T}\mathbf{p}_1]_{i*}$ denotes the i^{th} row of $[\mathbf{T}\mathbf{p}_1]$;
$\mathbf{p}_1=[x_1, y_1, 1]^T$;
ρ,τ are non-zero scale factor.
In case of noise-free image measurement, both equations are equivalent. The same transfer can be done with lines. The point transfer can be solved also using the fundamental matrix, but the trifocal constraint can avoid ambiguities and remove blunders. Moreover, from the tensor it is possible to derive the fundamental matrices between the first and the third view; e.g., given 3 images, $\mathbf{M}_{13}$ between image 1 and 3 is given by:

$$\mathbf{M}_{31} = [\mathbf{e}_3]_x [T_1, \quad T_2, \quad T_3] \, \mathbf{e}_2 \qquad [8]$$

where:
$\mathbf{e}_i$ is the epipole of image i;
$[\mathbf{e}_i]_x$ is the skew-symmetric matrix (6) formed with $\mathbf{e}_i$.
Therefore, the transfer of $\mathbf{p}_3$ is expressed as:

$$\mathbf{p}_3 = (\mathbf{M}_{13}\mathbf{p}_1) \times (\mathbf{M}_{23}\mathbf{p}_2) \qquad [9]$$

e.g. the intersection of two epipolar lines in the third view.
The 27 unknowns of the tensor **T**, defined up to a scale factor, can be computed from at least 7 correspondences: using equation (5), each correspondence gives 9 equations, 4 of them linearly independent. In our process, for each triplet of images, the tensor **T** is computed with a RANSAC algorithm [Fischler and Bolles, 1981] using the correspondences that support two adjacent pair of images and their epipolar geometry. The RANSAC is a robust estimator, which fits a model (**T** tensor) to a data set (triplet of correspondences) starting from a minimal subset of the data. As result, for each triplet of images, a set of corresponding points, supporting a trilinear tensor, is available.
After the computation of a **T** tensor for every consecutive triplet of images, we consider all the overlapping tensors (T_{123}, T_{234}, T_{345},...) and we look for those correspondences which are present in consecutive tensors. That is, given two adjacent tensors T_{abc} and T_{bcd} with supporting points (x_a,y_a, x_b,y_b, x_c,y_c) and (x'_b,y'_b, x'_c,y'_c, x'_d,y'_d), if (x_b,y_b, x_c,y_c) in the first tensor is equal to (x'_b,y'_b, x'_c,y'_c) in the successive tensor, this means that the point in images a, b, c and d is the same and therefore this point must have the same identifier. Each point is tracked as long as possible in the sequence. The obtained correspondences are used as tie points for the successive bundle-adjustment.

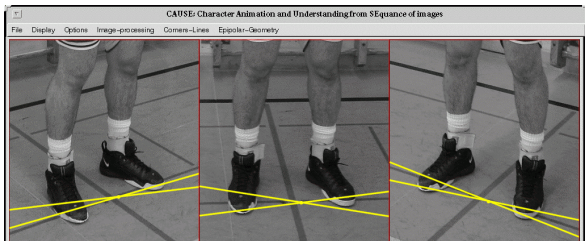

Figure 4: The relative geometry between a triplet of images

Figure.5: Extracted lines with Canny operator (a) and merged segment (b-c, d-e). Aggregated lines classified according to their direction (f,g,h). 4 control points measured manually on the body and used for the adjustment (i)

3.2 Initial approximation of the unknowns

Because of its non-linearity, the bundle-adjustment (section 3.3) needs initial approximations for the unknown interior and exterior orientations.
An approach based on vanishing point is used to compute the interior parameters of the camera (principal point and focal length). The vanishing point is the intersection of parallel lines in object space transformed to image space by a perspective transformation of the camera. Man-made objects are often present in the images, therefore geometric information of the captured scene can be derived from these features.
The semi-automatic process to determine the approximations of the interior parameters consists of:

- straight lines extraction with Canny operator (Figure 5, a)
- merging short segments taking into account segments slope and distance from the center of the image (Figure 5, b-c, d-e);
- interactive identification of three mutually orthogonal directions;
- classification of the extracted and aggregated lines according to their directions (Figure 5, f,g,h);
- computation of the three vanishing points for each direction [Collins, 1993]. Each line $\mathbf{l}_i$ is represented by its homogeneous coordinates (a_i, b_i, c_i); if there are only two lines, the cross product of them gives the coordinates of the vanishing point; if n lines $\mathbf{l}_1$, $\mathbf{l}_2$, ... $\mathbf{l}_n$ are involved, we get the "best fit" vanishing point forming the matrix $\mathbf{L}$ as:

$$\mathbf{L} = \sum_{i=1}^{n} \begin{bmatrix} a_i a_i & a_i b_i & a_i c_i \\ a_i b_i & b_i b_i & b_i c_i \\ a_i c_i & b_i c_i & c_i c_i \end{bmatrix} \qquad [10]$$

  and computing the vanishing point as the eigenvector associated with the smallest eigenvalue.
- determination of the principal point and the focal length of the camera [Caprile and Torre, 1990].

The approximations of the exterior orientation are instead computed using spatial resection. In our case, 4 object points measured on the human body (Figure 5, i) are used to compute the approximations of the positions of the cameras.

3.3 Bundle adjustment

Using the process described in section 3.1, a total of 148 correspondences are found in the images of Figure 1 and then imported in the adjustment. The points used for the space resection are imported as control points. Ten additional parameters [Brown, 1971] are used to model systematic errors: the camera constant correction, two principal point coordinate offsets, five parameters modelling the radial and tangential lens distortion and two parameters for a affine scale factor and shear [Beyer, 1992]. In our case, the principal point coordinate offsets, the parameter for the correction of the camera constant and the first term of the radial lens distortion turned out to be significant. The theoretical precision of the tie points is σ_X = 15.5 mm, σ_Y = 9.8 mm, σ_Z = 14.2 mm while the standard deviation of unit weight a posteriori is 1.8 micron (1/4 of the pixel size). The computed camera poses and 3-D coordinates of the tie points are shown in Figure 6.

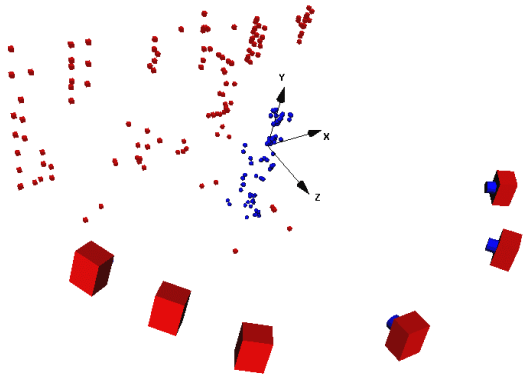

Figure 6: Recovered camera positions and object points

4. MATCHING PROCESS AND 3-D RECONSTRUCTION OF THE HUMAN BODY

In order to produce a dense and robust set of corresponding image points, an automated matching process is used [D'Apuzzo, 2002]. It establishes correspondences between triplets of images starting from few seed points and is based on the adaptive least squares method. One image serves as template and the others as search image. The matcher searches the corresponding points in the two search images independently and at the end of the process, the data sets are merged to become triplets of matched points. For the process, all consecutive triplets are used. The 3-D coordinates of each matched triplet are then computed by forward intersection using the results of the orientation process. At the end, all the points are joined together to create a unique point cloud. In order to reduce the noise in the 3-D data and get a more uniform density of the point cloud, a spatial filter is applied: the object space is divided into boxes and the points contained in each box are replaced by the center of gravity of the box. After the filtering process, a uniform 3-D point cloud is obtained, as shown in Figure 7. The generation of a surface model from unorganised 3-D point clouds requires non standard procedures which can be found in commercial packages. A standards 2.5 Delauney triangulation can not create a correct meshed surface from the obtained 3-D point shown in Figure 7.

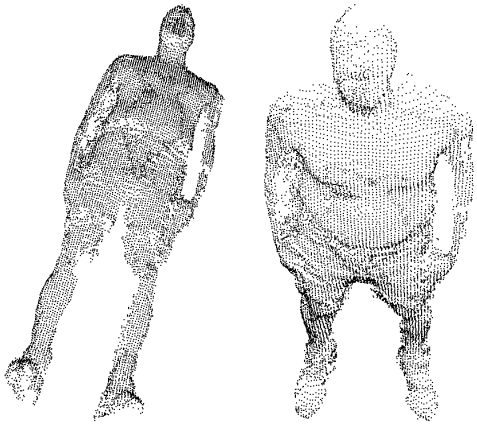

Figure 7: 3-D point cloud (8523) of the human body

Therefore, for realistic visualization of the results, each point of the point cloud is re-projected onto the central image of the sequence to get the related pixel colour. The results are presented in Figure 8.

Figure 8: Visualization of the point cloud with pixel intensity

5. ANOTHER EXAMPLE

The reconstruction of the human body is also performed using a sequence acquired with a video camera handycam Sony DCR-VX700E. The artefacts created by the interlacing during the digitization process have to be removed: therefore one field of the video is deleted and then the remaining lines are interpolated. A less smooth sequence is obtained as the resolution in vertical direction is reduced by 50 per cent. Instead of removing all the odd (even) lines, another possible approach could be to remove lines just in portions of the video where interlace artefacts are present.
For the process 12 frames are selected out of a one minute sequence around a standing person (Figure 9).

Figure 9: Three frames (out of 12) of the acquired video sequence. Images have a resolution of 576x720 pixel

The testfield in the background contains many similar targets (repeated pattern) and they are used just as features for the processing. No 3-D information is available. As the sequence is self-acquired, the parameters of the camera are known, but they are recovered as described in section 3.2. The process presented in section 3.1 is applied to find the required tie points for the orientation of the sequence; four control points measured on the body are used for the bundle. The final configuration of the system after the adjustment is shown in Figure 10.

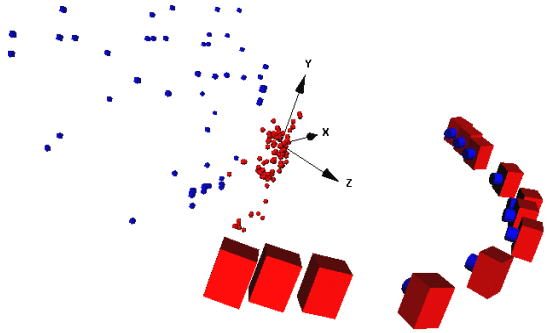

Figure 10: Recovered camera positions and object coordinates

Then the matching procedure is performed with all consecutive triplets of images and the obtained and filtered point cloud (5604 points) is shown in Figure 11 and 12.

Figure 11: 3-D point cloud of the human of Figure 9

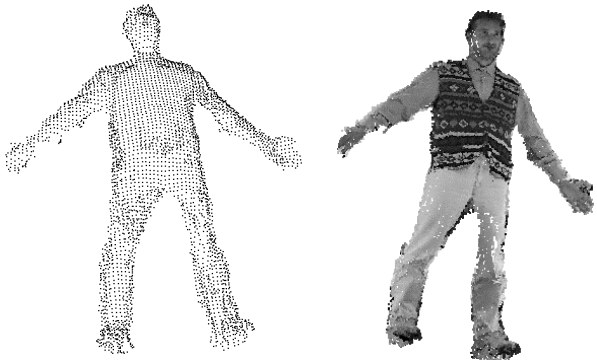

Figure 12: 3-D point cloud (left) and visualization with grey scale pixel intensity (right)

6. CONCLUSION

In this paper a method to create 3-D models of human bodies from uncalibrated image sequences has been presented. The process is automated and the recovered point cloud can be imported in commercial software for editing, surface modeling or animation purpose. As future work, the process for the identification of image correspondences described in [Pilu, 1997] could be included in our method, weighting the proximity matrix with the sigma of the ALSM process. The reconstruction of the body will also be extended to the back part, with a 360 degree coverage. Moreover, sequences where the camera is still and the person is moving or both camera and person are moving will be investigated.

ACKNOWLEDGMENT

I would like to thank G.Roth, of the Institute for Information Technology, NRC of Canada, for the useful talks about his PVT toolkit.

REFERENCES

Beyer, H.A., 1992: Geometric and Radiometric Analysis of CCD-Cameras. Based Photogrammetric Close-Range system. Ph.D. thesis 51, IGP ETH Zurich

Bhatia G., Smith K. E., et al., 1994: Design of a Multisensor Optical Surface Scanner. Sensor Fusion VII, SPIE Proc. 2355, pp. 262-273

Brown, D.C., 1971: Close-range Camera Calibration. PE&RS, Vol.37, No.8, pp. 855-866

Caprile B., Torre, V., 1990: Using vanishing point for camera calibration. International Journal of Computer Vision, Vol.4, No.2, pp. 127-139

Collins, R.T., 1993: Model Acquisition using stochastic projective geometry. PhD Thesis, Computer Science Dep., University of Massachusetts

D'Apuzzo N., Plänkers R, Fua, P., 2000: Least squares matching tracking for human body modeling. Int. Archives of Photogrammetry and Remote Sensing, Vol.33 (B5/1)

D'Apuzzo, N., 2002: Modeling human faces with multi-image photogrammetry. 3-Dimensional Image Capture and Applications V, SPIE Proc., Vol. 4661, pp. 191-197

Faugeras O., Luong Q.T., et al., 1992: Camera Self-calibration: Theory and Experiments. Lecture Notes in Computer Science 588, ECCV '92, Springer-Verlag, pp. 321-334

Fischler, M, Bolles, R, 1981: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Comm. Assoc. Comp. Mach., Vol. 24 (6), pp. 381-395

Fitzgibbon, A, Zisserman, A., 1998: Automatic 3D model acquisition and generation of new images from video sequences. Proceedings of European Signal Processing Conference, pp. 1261-1269

Grün A., 1985: Adaptive least squares correlation: a powerful image matching technique. South African Journal of Photogrammetry, Remote Sensing and Cartography Vol. 14, No.3, pp.175-187

Grün, A., Zhang, L., Visnovcova, J., 2001: Automatic Reconstruction and Visualization of a Complex Buddha Tower of Bayon, Angkor, Cambodia. Proc. 21 DGPF, pp. 289-301

Horiguchi C., 1998: Body Line Scanner. The development of a new 3-D measurement and Reconstruction system. Int. Archives of P&RS Vol.32, part B5, pp.421-429

Klette R., Schlüns, K., Koschan, A., 1998: Computer Vision: Three-dimensional data from images. Springer Press

McKenna P., 1996: Measuring Up. Magazine of America's Air Force, Vol. XL, No.2

Maas, H.G., 1991: Digital Photogrammetry for determination of tracer particle coordinates in turbulent flow research. PE&RS, Vol.57, No.12, pp. 1593-1597

Niini, I., 1994: Relative Orientation of Multiple images using projective singular correlation. Int. Archives of Photogrammetry and RS, Vol. 30, part 3/2, pp. 615-621

Pilu, M, 1997: Uncalibrated stereo correspondences by singular value decomposition. TR HPL-97-96, HP Bristol

Pollefeys, M., 2000: Tutorial on 3-D modeling from images. Tutorial at ECCV 2000

Remondino, F., 2002: 3-D reconstruction of articulated objects from uncalibrated images. 3-Dimensional Image Capture and Applications V, SPIE Proc., Vol. 4661, pp. 148-154

Roth, G., Whitehead, A., 2000: Using projective vision to find camera positions in an image sequence. 13th Vision Interface Conference

Scaioni, M., 2001: The use of least median squares for outlier rejection in automatic aerial triangulation. Proc. of 1st Int. Symposium on "Robust Statistics and Fuzzy Techniques in Geodesy and GIS", ETH Zurich, pp. 233-238.

Shashua, A., 1994: Trilinearity in visual recognition by aligment. In Ecklund, J.O.: ECCV, Lectures Notes in Computer Science, Vol.800, Springer-Verlag, pp.479-484

Wolf, H.G, 1996. Structured lighting for upgrading 2D-vision system to 3D. Proc. Int.Symposium on Laser, Optics and Vision for Productivity and Manufacturing I, pp. 10-14

Zhang, Z., Deriche, R, et al., 1994: A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. TR 2273, INRIA

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

Cyberware: http://www.cyberware.com [May 2002]
Taylor: http://www.taylor.com [May 2002]
Vitus: http://www.vitus.de/english/home_en.html [May 2002]