VIDEOMETRIC SYSTEM FOR SURFACE DETERMINATION IN MEDICAL APPLICATIONS

Josef JANSA[1], Balazs MELYKUTI[1], Christian ÖHRENEDER[1] [2]
[1] Vienna University of Technology, Austria
Institute of Photogrammetry and Remote Sensing (I.P.F.)
[2] Currently with Kapsch AG, Vienna, Austria
jj@ipf.tuwien.ac.at; bm@ipf.tuwien.ac.at; co@ipf.tuwien.ac.at

Working Group V-9

KEY WORDS: CCD, Real-time, Surface reconstruction

ABSTRACT

Efficient and robust three dimensional shape determination is still an important problem in computer vision in medical applications. Modeling of surfaces of human bodies is necessary for operation planning, operation supervision, and postoperative checking. For these tasks a videometric system has been developed based on low cost hardware elements, i.e. four standard video cameras, two pattern projectors and a PC with frame grabber card. Through a user friendly interface the system can be easily operated by a wide range of users without photogrammetric experience. All complicated procedures, such bundle block adjustment for the camera calibration and orientation and the matching software, run in background. The central element of the system is the matching algorithm for surface determination which uses a newly developed, reliable global approach leading to an accuracy that lies safely within 1 millimeter. Though not in real-time, the three dimensional model can be delivered shortly after image acquisition.

KURZFASSUNG

Effiziente und robuste Bestimmung der 3-dimensionalen Form ist immer noch ein wichtiges Problem auf dem Gebiet der Computer Vision im medizinischen Bereich. Die Modellierung von Teilen der Oberfläche des menschlichen Körpers wird gebraucht für Operationsplanung, für Operationsüberwachung und postoperative Überprüfung. Zu diesem Zweck wurde ein System entwickelt, das Niedrigpreiskomponenten verwendet, d.h. vier Standard-Videokameras, zwei Musterprojektoren und einen PC mit Frame Grabber Bord. Über ein benutzerfreundliches Interface kann das System leicht bedient werden, auch durch Personal, das keine photogrammetrische Erfahrung besitzt. Alle komplizierten Prozesse laufen im Hintergrund ab, wie etwa der Bündelblockausgleich für die Kalibrierung und Orientierung der Kameras oder die Matching Software. Ein zentrales Element des Systems bildet der neu entwickelte Matching Algorithmus, der auf einem verläßlichen globalen Ansatz aufbaut und mit dem eine Genauigkeit erreicht wird, die sicher innerhalb eines 1 mm liegt. Obwohl nicht in Echtzeit, so kann das dreidimensionale Modell doch kurz nach der Bilderfassung erhalten werden.

1   INTRODUCTION

In medical practice various technologies are employed for three dimensional measurement. Well-known are CT and MR technologies, although also traditional photogrammetric methods are still important due the fact that they do not use any harmful radiation nor do they need an expensive instrumentation (Deacon et al., 1991, D'Apuzzo, 1998). In all cases where the surface shape of a body is needed conventional photographic image acquisition is well suited. The drawback of the latter systems was, firstly, its long turn-around time between image capture and delivery of surface data, and secondly, the need for experienced personnel. In the meantime digital photogrammetry offers tools for quick, uncomplicated, inexpensive and accurate measurement.

The Department of Oral and Maxillofacial Surgery at the University Clinic located at the Vienna General Hospital was using a photogrammetric system several years ago for the determination of the shape of the facial surfaces for planning operations and checking their successfulness (Rasse et al., 1992). They were satisfied by the basic principle and wanted to move to a digital equipment. The I.P.F. developed a photogrammetric set-up that used state-of-the-art video surveillance cameras together with a software package, that allows accurate and reliable compilation within a reasonable time frame.

2   HARDWARE

The set-up of the system consists of four video cameras and two light projectors mounted on a frame in a square arrangement (Fig.1) thus forming a configuration of two stereopairs with vertical base lines. The two stereo models overlap in the centre and guarantee full coverage even in cases of complex surfaces.

Figure 1. Set-up with video cameras and projectors

Off-the-shelf monochrome video cameras with a CCD sensor matrix of 768 x 572 elements have been used. The sensor plate is 6,45 x 4,84 mm², each sensor element 8,6 μm x 8,3 μm. They deliver an analogue standard composite video signal. As optical system Pentax video lenses with a focal length of 16 mm have been employed in order to achieve a reasonable acquisition distance and depth of focus for face related applications. This type of video cameras have been widely used for surveillance purposes but also in various engineering applications (Beyer, 1992).

Taking pictures has to be quick in many medical applications as the living object, i.e. the patient, has to stand still for the time of data capture, otherwise time parallaxes must be expected which perturb the accuracy of the three dimensional compilation. Experience showed that a exposure time of less than one second is totally safe. In order to reduce acquisition time to a minimum the monochrome cameras have been plugged to the RGB-connector, with one camera on each colour channel. Since the frame grabber offers two multiplexed RGB sockets, one camera pair can be captured by one input channel perfectly synchronously. In this way one stereo model can be acquired within one video cycles, i.e. 1/25 second. The control routine switches automatically to the second colour input and grabs the second stereo model. In the optimum case the entire image acquisition can be finished within approximately 1/10 second. Figure 2 shows the basic configuration that offers the possibility to an extension from four to six cameras without any serious modification to the system just by plugging the fifth and sixth camera to the remaining two colour channels.
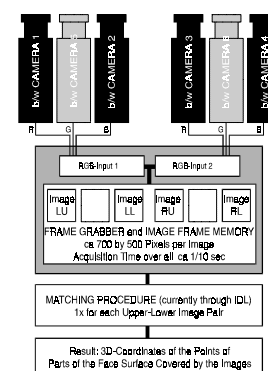

Figure 2. Block diagram of acquisition module

Two constantly illuminated pattern projectors are mounted on the frame, too.
They have to generate and project structured light onto the examined surface if natural texture does not exist. The slide that has been developed for that purpose is rich in contrast and consists of non-recurring random patterns in two frequency scales. One scale, the "low frequency pattern", is responsible for good convergence behaviour of the matching process, while the second scale, the "high frequency pattern" is to guarantee good matching accuracy. The whole frame can be moved and repositioned quickly without destroying the orientation of the instruments. The usual problem of a inadequate depth of field can be overcome by small iris settings.

3   SOFTWARE

The software design can be subdivided into three major parts:
1.   Software for grabbing has been provided as tool library that allows access to internal features of the grabber, such as image grabbing, channel switching and various possibilities of image display.

2. Software for camera calibration and orientation. I.P.F.'s bundle adjustment program ORIENT (Kager, 1989) has been implemented. Special macros allow running the program without user interaction. In addition, several utility programs have been written for control point detection and matching.
3. Software for surface determination. Öhreneder (1999) developed an algorithm for a global matching approach based on least squares in his PhD thesis. This software is currently running as a prototype version under IDL (Ineractive Data Language by Researh Systems Inc.). Conversion to C++ is planned.

3.1 Grabbing

A simple user interface facilitates the operation of the system. Very few input parameters are necessary, most of them are dedicated to image acquisition and camera calibration and orientation. During grabbing live images are displayed on screen. Though not absolutely necessary live images are very helpful when adjusting the camera positions and in particular the camera focus settings.

3.2 Calibration and Orientation

For the calibration and orientation of the cameras a three dimensional field of control points is essential. Here a two dimensional A4 size plate with a grid of circular targets has been used whose positions (i.e. xy coordinates, z=0) on the plate are known with high accuracy (Fig. 3). This plate serves as "model" in the bundle adjustment. It is tilted and rotated in front of the cameras so that each quadruple of captured images shows the same object (model) at different positions in three dimensional space. Geometric constraints apply that enable the determination of the parameters of the interior and relative orientation. By defining one of the plate positions as reference the complete outer orientation with respect to that plate can be obtained.

Figure 3. Control plate

The user of the system has to follow a simple recipe that describes all the (usually five different) positions of the plate in front of the camera set-up. After having taken all the images an automatic measurement procedure commences that is based on the weighted centroid algorithm (Trinder, 1989). On the plate there are 35 control points. By taking five shots with four cameras 175 points in the object space have to be determined from 700 image points. Four points of the control grid are specially indicated. The program is able to recognise and locate them in a first step. From their arrangement the plate orientation is determined allowing the program to automatically identify all other control points without the need to assign numbers manually or by reading a nearby code. The program warns and asks the operator for a new acquisition if problems occur during the matching process.
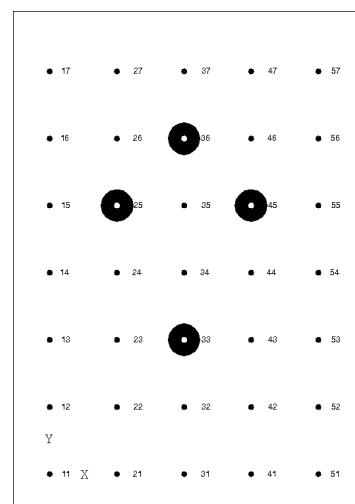
For correct utilisation of a bundle block adjustment program some photogrammetric experience is a prerequisite that must not be expected from users in hospital. Therefore, the bundle adjustment program ORIENT has been configured to run in background reading all necessary input from files that have been written by the control point matching procedure. The bundle block adjustment of the model is controlled by predefined macros, no interactive input is needed. Error messages and warnings are analysed automatically and passed to the user. The results of the adjustment are again written to a file for transfer to the surface matching program.

3.3 Matching

The method developed here (Öhreneder, 1999) can be called *global matching* because it tries to calculate at once a disparity function for the entire image or at least for great parts of the image rather than for individual points as local methods do. The matching algorithm is intensity based and follows the energy minimisation paradigm. The energy terms are integrals and the energy minimisation becomes a variation problem. The lower the energy, the higher the similarity. The algorithm uses the forward modelling principle, i.e. it is based on a model describing the dependency of the image data on the parameters of the surface model. The algorithm is open to multi-image matching, uses a hierarchical approach by employing image pyramids and is supported by the image orientation through epipolar geometric constraints.

For explaining the matching algorithm we treat only the normal case with two stereo images. The images are denoted by $f_L : F_L \mapsto \underline{R}$ and $f_R : F_R \mapsto \underline{R}$ ($\underline{R}$ denotes the set of real numbers). It is assumed that the images are monochrome and thus can be described by scalar valued functions. The aim of matching is to find corresponding points in the images. One way to setup a global description of the interrelation of the two images is to use a mapping between the domains $\mathbf{m} : F_L \mapsto F_R$. Alternatively we could establish the correspondence by relating both image domains to some third

domain. We will introduce such a domain called the reference domain G for modelling reasons. The geometrical connection between the F_R, F_L and G is given by $\mathbf{m}_L : G \mapsto F_L$ and $\mathbf{m}_R : G \mapsto F_R$ respectively. Possible choices for the reference domain are the object surface itself, the orthophoto domain, i.e. the *XY* plane of some object co-ordinate system, or whatever seems convenient.

In the following it is understood that the geometric connection as established by the $\mathbf{m}_R$ and $\mathbf{m}_L$ is properly parameterised. In a calibrated set-up a scalar parameterisation suffices to describe the geometric connection between the F_R, F_L and G. For the particular set-up using the disparity d as matching parameter is convenient. The maps can be written as

$$\mathbf{m}_L(\mathbf{x}) = \begin{pmatrix} x_1 + \dfrac{d(\mathbf{x})}{2} \\ x_2 \end{pmatrix} \quad \text{und} \quad \mathbf{m}_R(\mathbf{x}) = \begin{pmatrix} x_1 - \dfrac{d(\mathbf{x})}{2} \\ x_2 \end{pmatrix} \tag{1}$$

Besides the geometric aspects we need a radiometric model. Here we rely on the standard assumption of Lambertian reflection and ignore the details of how radiation quantities are related to greyvalues.

The matching problem is formulated as an optimisation problem. The free variable is the parameter of the geometric connection, the disparity d in the particular case. As we are treating the problem in a continuous setup, the optimization problem is a variational problem. The objective function is composed of energy values $E_{L,R}$ that indicate how well the images $f_{L,R}$ conform to the model given by $(\mathbf{m}_{L,R}, g)$. The notion *energy* or *image energy* is used because the quantity is the Gibbs energy of a random field. We also refer to this energy value as similarity measure.

Global matching methods often use the sum of squared differences as similarity measure. The new aspect we want to introduce stems from the fact that there are two alternatives to define such a quadratic measure. Let us assume, for the moment only, that we know the true texture g of the surface and consider a single image f. Many approaches (Amit,1994 and 1991, Heipke, 1990 and 1992) are based on the following similarity measure between an image f and the surface texture g,

$$E^r(\mathbf{m}) = \frac{1}{2} \int_G | f \circ \mathbf{m} - g |^2 \, \mathrm{d}A \tag{2}$$

dA denotes integration over a two dimensional domain (2). Evidently the above energy value is a valid similarity measure in that low energy values indicate that f and g are similar at corresponding positions. Energy values of this type are usually rationalized by assuming that the difference between f and g obeys a Gaussian distribution. The corresponding model is

$$g = f \circ \mathbf{m} + n_g \tag{3}$$

n_g denotes white Gaussian noise with standard deviation $\sigma_{ng} = 1$. Unfortunately equation (3) is not conform with our model assumption because the roles of f and g are interchanged. According to (3) the model quantity g differs from the image f by a random component. Obviously it should be the other way round. This change of roles is a consequence of the integral (2) which is performed in the reference domain. We refer to (2) as *reverse modeling similarity measure*. We want to correct the roles of image and texture model, i.e. we want the model to read

$$f = g \circ \mathbf{m}^{-1} + n_f \tag{4}$$

which leads to the following similarity measure

$$E^f(\mathbf{m}) = \frac{1}{2} \int_{F^M} | f - g \circ \mathbf{m}^{-1} |^2 \, \mathrm{d}A \tag{5}$$

This measure is based on integration in the domain F. The domain of integration comprises the modeled part $F^M = \mathbf{m}(G) \subseteq F$. It is referred to as *forward modeling similarity measure*.

A disadvantage of the energy term (5) is that it involves the inverse map $\mathbf{m}^{-1}$. Assuming bijectivity of $\mathbf{m}$, variable substitution can be employed and (5) written as an integral over G,

$$E(\mathbf{m}) = \frac{1}{2} \int_G | f \circ \mathbf{m} - g |^2 \det \mathbf{m}' \, \mathrm{d}A \tag{6}$$

The determinant $\det \mathbf{m}'$ is commonly referred to as Jacobian. In geodesy it is known as the area distortion. The form (6) is called the *reference space formulation* of the forward energy, whilst (5) is the *image space formulation*. It follows the principle to perform all calculations in the reference space. This computationally favorable structure can be achieved without sacrificing model assumptions.

Comparing (6) and (2) reveals that they differ by the weighting factor det **m**′. In (6) regions of G that are visible in F with bad spatial resolution receive a low weight and vice versa. It grants that regions of an image contribute to the similarity measure according to their size in image space and not according to their extent in the model domain. The benefits of the weighting factor become more clear when two energy terms are combined into a composite energy value:

$$E_{img}(d,g) = \frac{1}{2}\int_G |f_L \circ \mathbf{m}_L - g|^2 \det \mathbf{m}'_L \, \mathrm{d}A + \frac{1}{2}\int_G |f_R \circ \mathbf{m}_R - g|^2 \det \mathbf{m}'_R \, \mathrm{d}A \qquad (7)$$

The subscript 'img' is used because it comprises all image related terms.

The unknown texture g can be eliminated, which leads to the composite energy value

$$E_{img}(d) = E_{img}^f(d) = \frac{1}{4}\int_G |f_R \circ \mathbf{m}_R - f_L \circ \mathbf{m}_L|^2 \left(1 - \left(\frac{d'}{2}\right)^2\right) \mathrm{d}A \qquad (8)$$

Other authors (Belhumeur, 1996, Gennert, 1988, Horn,1991) have used a similar set-up but based on the reverse modeling similarity measure (2) which leads to:

$$E_{img}^r(d) = \frac{1}{4}\int_G |f_R \circ \mathbf{m}_R - f_L \circ \mathbf{m}_L|^2 \, \mathrm{d}A \qquad (9)$$

Again this differs by a weighting factor. The factor ensures that regions of the reference domain G which are occluded in either of the images – this is the case if the derivative of the disparity d′ reaches one of the values ±2 – do not account to the similarity measure. Mathematically and also practically the case of occlusions is a delicate matter. The characteristic of the model with respect to the limiting case of occlusions is however satisfying. A mathematical consequence of our forward modeling similarity measure is its continuity even in the case of occlusions. Some further generalizations are needed to handle the case of occlusions. These aspects are not further detailed in here. Another important aspect of the forward modeling similarity measure is its independence of the particular reference domain, which can be seen directly from (5).

4   ACCURACY ASSESSMENT

For the evaluation of the system both simple, geometrically well defined objects and a model of a human face were used. And in all cases artificial texture has been projected onto the surface by the pattern projector. The base line was set to 20 cm, the object distance was 1 m, yielding a base-to-height ratio of 1:5. The estimated accuracy according to the formulas for the normal case, assuming a parallax accuracy of 1 μm (i.e. approx. 1/10 of a pixel), would be ±0.3 mm.
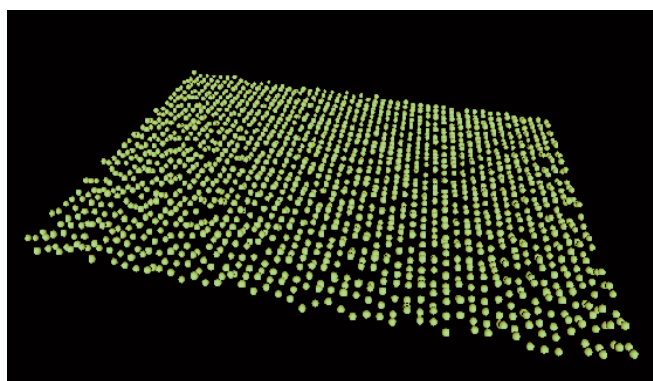

Figure 5. View of the matched points

As one of the simple object served a white, plane plate. Figure 5 shows a 3D view (realised in VRML) of the matched points. Each point is represented as a ball with a diameter of 1 mm. The r.m.s. deviation of the points from an mathematically ideal plane is $\sigma_{\Delta_z} = \pm 0.6$ mm.

A plaster model of a human face served as irregular surface and, of course, as surface that represents the object shape to be surveyed in practice, with the great advantage that exactly the same shape can be measured as many times as necessary. No reference measurement was available. Therefore, the accuracy check concentrated on the comparison of surface restitutions. One of the tests investigated the influence of the texture. With the calibrated and oriented set-up

two measurements were carried out. For the second the texture projectors have slightly been moved thus generating the pattern on different position on the plaster face. The results of the matching were again visualised through VRML (Figure 6). The matched points of one measurement are displayed as a green (1 mm) ball, those of the other measurement in red. Numerical difference assessment yielded an r.m.s. deviation of $\sigma_{\Delta_{xy}} = \pm 0.1$ mm and $\sigma_{\Delta_z} = \pm 0.5$ mm.
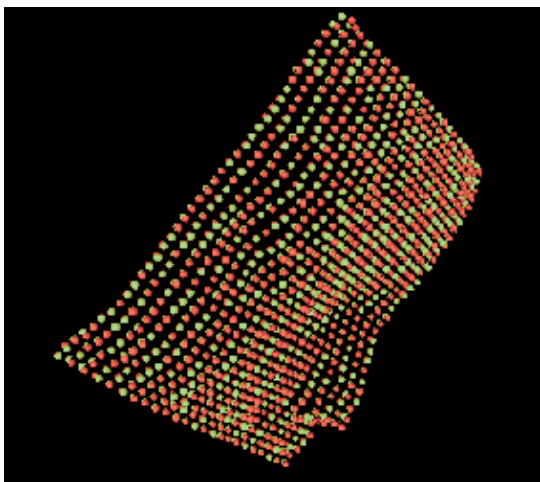

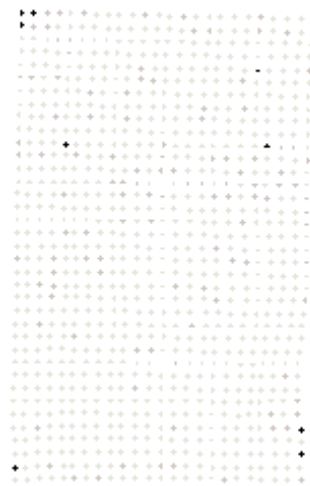Figure 6. Two measurements projected onto each other

Figure 7. The differences of the models

Figure 7 shows the differences again, projected onto the planimetric reference plane and grey coded. The lighter the less is the difference between the two independent measurements. The maximum Δz in this example is 2 mm and occurs along the boundary were the matching procedure works, not unexpectedly, less accurate.

5   CONCLUSION

This paper discussed the set-up of a digital close-range measurement system that may be used in a wide range of applications and not only for medical purposes as discussed above. The user interactions are kept at a minimum, making the system ideal in all those environments where photogrammetric experts are not present. The robust and reliable procedure delivers an accuracy that fulfils the requirements of many tasks even in engineering applications. Since a universal bundle adjustment program has been used for calibration and orientation, there is almost no limitation to extensions or modifications. The employed matching algorithm does not need careful considerations about provisional values, it is accurate and quick although running as non-optimised prototype version. Forthcoming improvements will include implementation under C++.

REFERENCES

Amit Y., 1994: A nonlinear variational problem for image matching. SIAM Journal of Scientific Computing, Vol.15(1), pp. 207-224.

Belhumeur P.N., 1996: A bayesian approach to binocular stereopsis. International Journal Computer Vision, Vol.19/3, pp 237-260.

Beyer H.A., 1992: Geometric and Radiometric Analysis of a CCD-camera based Photogrammetric Close-Range System. Dissertation No.9701, ETH, Zurich.

D'Apuzzo N., 1998. Automated Photogrammetric Measurement of Human Faces. International Archives of Photogrammetry and Remote Sensing, Hakodate, Japan, 1998, Vol. 32 (B5), pp 402-407.

Deacon A.T., Antony A.G., Bhatia N. and Müller J-P., 1991. Evaluation of a CCD-based facial measurement system. Med. Inform., 16(2): 213-228.

Gennert M.A., 1988: Brightness bases stereo matching. In: Proceeding of 2nd International Conference on Computer Vision, pp 139-143.

Heipke C., 1990. Integration von Bildzuordnung, Punktbestimmung, Oberflächenrekonstruktion und Orthoprojektion innerhalb der digitalen Photogrammetrie. Deutsche Geodätische Kommission, Reihe C, Heft 366, Bayerische Akademie der Wissenschaften, München.

Heipke C., 1992. A global approach for least-squares image matching und surface reconstruction in object space. Photogrammetric Engineering and Remote Sensing, Vol.58(3), pp. 317-323.

Horn B.K.P., 1991. Robot Vision. Electrical Engineering and Computer Science Series. MIT Press, Massachusetts, 7 edition.

Kager H., 1989. ORIENT: A Universal Photogrammetric Adjustment System. In Grün/Kahmen (Editors): Optical 3-D Measurement Techniques, Wichmann Verlag, Karlsruhe, S.447-455.

Öhreneder Chr., 1999. A Similarity Measure for Global Image Matching Based on the Forward Modeling Principle. Geowissenschaftliche Mitteilungen der Studienrichtung Vermessungswesen der TU Wien, Heft Nr.51, Institut für Photogrammetrie und Fernerkundung, Wien.

Rasse M., Forkert G., Waldhäusl P., 1992. Stereophotogrammetrische Untersuchungen der Gesichtsweichteile. Mitteilungen der Klinik für Kiefer und Gesichtschirurgie der Universität Wien ,36Seiten.

Trinder J., 1989. Precision of digital target location. Photogrammetric Engineering and Remote Sensing, Vol.55/6, pp.883 – 886.