

# Uncalibrated X-Ray Stereo Reconstruction

Dan Talmage, Alison Noble and Andrew Zisserman  
Robotics Research Group, Department of Engineering Science  
Oxford University, OX1 3PJ.

## Abstract

We describe a novel application of *uncalibrated* stereo reconstruction to Roentgen Stereophotogrammetry Analysis (RSA). In RSA, stereo X-ray images are taken of a bone containing a prosthesis (e.g. a replacement knee) and a number of metal markers. The aim is to recover the relative position of the prosthesis and markers in 3D.

Accuracy in previous RSA methods has been limited by two factors: manual feature selection and an assumption that camera calibration parameters are known to high precision - this is not the case in practice. Furthermore, the manual processing is slow and tedious.

We report progress towards developing a fully automatic RSA system. New algorithms are described for automatically localising marker points in X-ray images to sub-pixel accuracy, and using them to reconstruct accurate 3D positions using robust statistical methods. Preliminary experiments give excellent results.

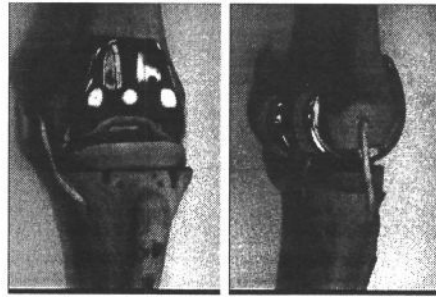
## 1 Introduction

Total knee replacement operations are one of the successes of medical surgery. The success rate for the operation is high, with implanted knee prostheses (figure 1) lasting an average of around ten years. A minority of prostheses need early revision due to pain from loosening of the prosthesis, which then migrates (moves) relative to the bone [12, 14]. It is important to examine these cases not only to predict early failure but also to provide manufacturers with information on which to base new prosthesis designs. A key challenge that this presents is that migration is small - typically of the order of 1mm/year.

One non-invasive method that has shown potential for measuring this migration is called Roentgen Stereophotogrammetry Analysis (RSA) [8, 13, 16]. For RSA, tantalum marker beads are implanted in the bones during a knee arthroplasty operation. These remain fixed and locate the positions of the bones. Post operatively, simultaneous stereo X-rays are taken of a calibration frame (figure 2) around the knee. The frame and the markers are then located in the images (figure 3), and photogrammetric methods are used to reconstruct the Euclidean positions of the bones and prosthesis. Repeated RSA examinations allow calculation of the prosthesis migration. To date, this approach has been based on conventional photogrammetric methods, using fiducial markers for camera calibration and calibrated reconstruction to recover position [4]. However, current

methods do not achieve the accuracy required in a robust and repeatable manner. This is due to reliance on camera parameters which may be poorly known, and due to manual feature extraction.

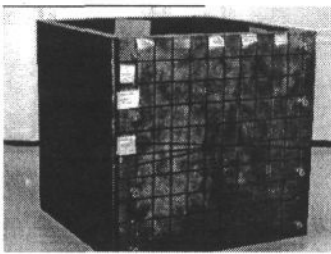
This paper describes a novel application of uncalibrated stereo reconstruction to RSA. Uncalibrated stereo allows projective reconstruction without knowledge of any camera parameters or of the relationship between the cameras. Euclidean reconstruction is then achieved by means of known 3D points.



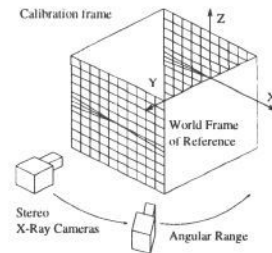
a

b

Figure 1: The experimental replica knee, showing the implanted prosthesis; (a) coronal, and (b) sagittal views.



a



b

Figure 2: The calibration frame; this is placed around the patient's knee during X-ray image acquisition. (a) a photograph of the frame, showing the etched lattice and the grid centre, formed by the diagonal lines. (b) The coordinate frame and camera motion. The world frame of reference ( $X, Y, Z$ ) is defined by the grid.

We focus on how to measure the relative migration of an artificial knee prosthesis relative to a bone. Namely, given two digital X-ray images of the bones, the prosthesis, fiducial markers in the bones and a calibration grid around the whole, determine the positions of the markers and of the prosthesis relative to a known frame of reference. An ultimate goal of our work is to achieve very high accuracy reconstruction to enable displacements of the order of  $100\mu m$  to be discerned. In

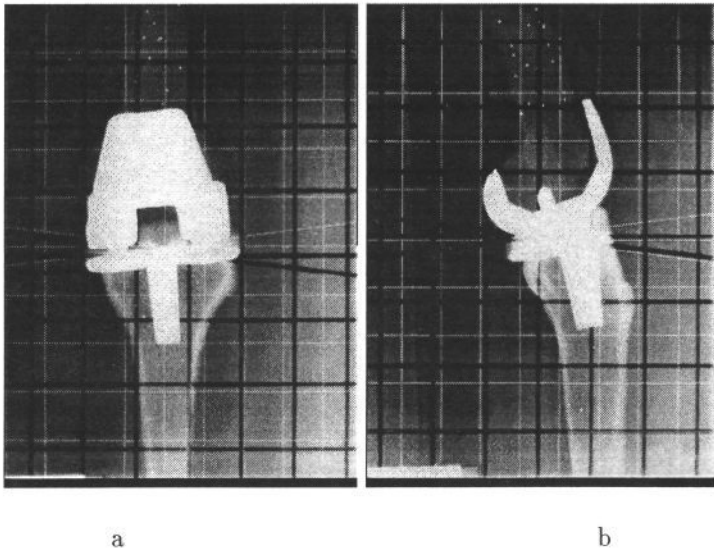


Figure 3: (a) Coronal and (b) sagittal X-rays of the replica knee, showing the calibration frame grid lines. The images are  $575 \times 772$  pixels.

this paper we describe progress towards developing a fully automated system to achieve this.

Our approach involves four steps:

1. *Feature localisation and matching:* A key problem in X-ray image segmentation is that conventional methods such as edge detection do not work well. Therefore a new parametric-model fitting algorithm has been developed to localise marker features to sub-pixel accuracy. This method is described in Section 2. Currently, features are manually corresponded across views. This step will be automated in future.
2. *Projective reconstruction:* We make use of the fundamental matrix and recent results in uncalibrated stereo [3, 5] to accurately recover projective structure. This step is described in section 3.
3. *Metric reconstruction:* Given a projective reconstruction we can transform it into a metric reconstruction using a number of known 3D locations in a Euclidean frame of reference. This step is described in section 4.
4. *Prosthesis registration:* The final step is to determine the prosthesis position in the same Euclidean reference frame as the bone.

Steps (1)-(3) are described in this paper. Step (4) is the subject of current work. The main contribution of the paper is to show that using automated feature extraction and uncalibrated stereo reconstruction, 3D locations can be recovered to within the accuracy required to measure prosthesis migration.

## 2 Automatic Marker Segmentation

In order to achieve the required accuracy for RSA, manual feature localisation is inadequate. Therefore a new method of segmentation has been developed to locate markers in an image to subpixel accuracy. The approach used is similar to that described in [7] but the model is quite different. Marker segmentation involves three steps:

**1. Finding seed points:** Localisation of markers involves first applying the morphological opening residue [11] to remove background gradations from the view followed by thresholding with a fixed grey level threshold value. This process detects regions which correspond to the markers, and others which correspond to spurious noise. The centre of gravity of each local region  $(x_0, y_0)$  is computed, and used to seed the parameterised model fit described in the next step.

**2. Fitting the parameterised model:** We model the X-ray projection of a marker as a 2-dimensional Gaussian, parameterised by its height  $A$ , shape parameters  $(\sigma_x, \sigma_y)$  and image location  $(x_c, y_c)$ . We assume that the marker is superimposed on an arbitrarily oriented planar background slope  $Q_0 + Q_1x + Q_2y$  where  $Q_i, i = 0, 1, 2$ , are constants. The complete marker intensity model  $I(x, y; \mathbf{a})$  therefore is described by 8 parameters and defined by,

$$I(x, y; \mathbf{a}) = Q_0 + Q_1x + Q_2y + A \exp - \left( \frac{(x - x_c)^2}{2\sigma_x^2} + \frac{(y - y_c)^2}{2\sigma_y^2} \right), \quad (1)$$

where  $\mathbf{a} = [A, \sigma_x, \sigma_y, x_c, y_c, Q_0, Q_1, Q_2]$ , and  $(x, y)$  is the image location.

The seed point  $(x_0, y_0)$  is used to initialise the model parameters  $x_c$  and  $y_c$ . We assume a nominal diameter  $d_m$  for a marker (in pixels). The parameters  $\sigma_x$  and  $\sigma_y$  are initialised as  $0.25d_m$ . An  $N \times N$  region-of-interest (ROI) is set around each seed point of height and width  $3d_m$ .  $Q_0$  is set equal to the average pixel value computed over the ROI.  $Q_1$  and  $Q_2$  are initially set to be zero. The initial value of  $A$  is set equal to  $I(x_c, y_c) - Q_0$ .

This initial model is refined to best-fit the image function using the Levenberg-Marquardt algorithm for non-linear least-squares parameter fitting [15]. This method iteratively refines the model parameters until the cost function  $C(\mathbf{a})$  defined by

$$C(\mathbf{a}) = \sum_{i=1}^N \sum_{j=1}^N [I_{i,j} - I(x_i, y_j; \mathbf{a})]^2, \quad (2)$$

is minimised. Here  $I_{i,j}$  is the actual intensity value at  $(x_i, y_j)$ ,  $I(x_i, y_j; \mathbf{a})$  is the value of the intensity model and the summation is taken over the ROI. The algorithm is iterated until this cost function decreases between iterations by less than 0.1%. Once the best-fit has been found the final values of  $x_c$  and  $y_c$  are taken as the marker centre.

Figure 4 shows a typical marker and the parameters that fit it. Extensive experimentation on many such markers has been carried out and has shown that the current initialisation method is good for cases where the marker approximately

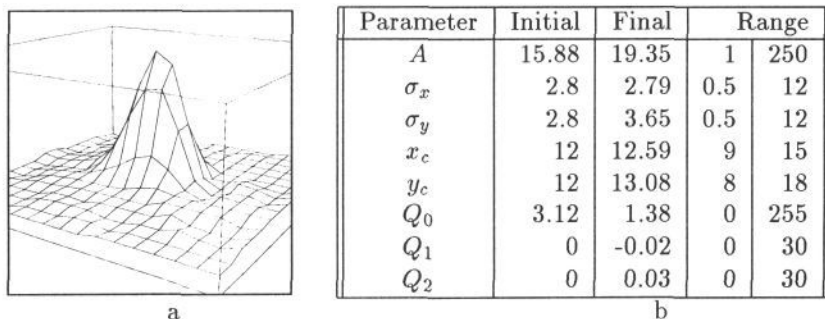


Figure 4: (a) the image surface in the region of a marker ball, with no grid lines included. (b) The table gives initial and final values for a successful fit. The last column shows the range of parameter values for which the fit converges, with the other parameters set at their initial values.

fits the model. The most critical thing is to provide good starting values for  $x_c$  and  $y_c$ . As a suggested rule of thumb, a seed point within  $\sigma_x$  and  $\sigma_y$  from the true centre works well except in those cases where the background is radically different from that expected by the model.

**3. Validating the model fit:** The final step in marker segmentation is to check that the fitted model parameters are valid. Currently the following checks are made.

- A. Marker centre:** If  $x_c - x_0$  or  $y_c - y_0 > d_m$  the fit is rejected. This condition rejects fits that lie too far away from the seed.
- B. Marker shape:** If  $\sigma_x$  or  $\sigma_y > 0.5d_m$  the fit is rejected. This test removes markers which have been confused by the background. This occurs, for example, when a marker bead aligns directly with a grid line.
- C. Marker size:** If  $\sigma_x$  or  $\sigma_y < 0.1d_m$  the fit is rejected. This removes fits on seed points which are associated with noise.
- D. Marker height:** If  $A < 5$  grey levels the fit is rejected. This condition rejects fits to areas of the image with a gentle image curvature. This condition is the least satisfactory; it occasionally rejects valid markers that have been poorly imaged.

These criteria were arrived at by fitting the model to around 140 markers in ten images. Validation rejects an average of 20% of the markers, depending on the image quality. Rejected points generally lie either behind the prosthesis, or on a grid line<sup>1</sup>. In the former case there is little which can be done. In the latter

<sup>1</sup>About half the markers are close enough to grid lines that the lines are included in the fit region.

case, one possible solution would be to enhance the intensity model described in Equation (1) to account for the shape of the grid line. This enhancement may be implemented in future work.

### 3 Projective Reconstruction

Given a set of marker locations identified in a number of X-ray image views the next step is to perform a projective reconstruction. The method we apply to do this is based on the fundamental matrix [3, 5].

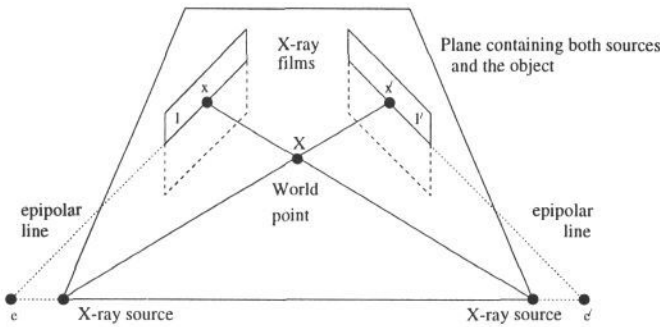


Figure 5: Stereo X-Ray Camera Geometry. An object  $X$  casts images  $x$  on the left film and  $x'$  on the right film. A point  $x$  in the left image lies on an epipolar line  $l$ , which is where the plane of the X-ray cameras cuts the image. The epipoles  $e, e'$  lie at the intersection of the film planes and the line joining the X-ray camera centres.

**Projective geometry** The imaging geometry of the X-ray image acquisition system is illustrated in figure 5. The X-ray cameras take perspective views of objects and therefore the standard model of stereo geometry under perspective projection is applicable. Stereo image pairs therefore satisfy the epipolar constraint, represented by the fundamental matrix.

**The fundamental matrix** Recall that the fundamental matrix is the  $3 \times 3$  matrix which satisfies,

$$\mathbf{x}'^T F \mathbf{x} = 0, \quad (3)$$

where  $\mathbf{x}$  is a point in the left image and  $\mathbf{x}'$  is a point in the right image. The epipoles  $e, e'$  are found by solving  $F e = 0$ ,  $F^T e' = 0$ . From these and  $F$ , a pair of projection matrices  $P$  and  $P'$  can be obtained which define the perspective projection from projective 3-space onto the image planes [1, 5].

The fundamental matrix can be computed in a number of ways. The method used here is to calculate  $F$  in a two-stage process, beginning with Longuet-Higgins eight-point algorithm [9] and utilising the result as the initial estimate for a robust non-linear method, which enforces rank two and rejects poor points [1, 10].

**Finding projective structure** The projection matrices  $P$  and  $P'$  define a point  $\mathbf{X}$  in projective space, by the equations  $\mathbf{x} = P\mathbf{X}$ ,  $\mathbf{x}' = P'\mathbf{X}$ .  $\mathbf{X}$  is found for each pair of image points  $\mathbf{x}$ ,  $\mathbf{x}'$  by Hartley's Iterative-Eigen method [6].

## 4 Metric reconstruction

The final stage is to find the  $4 \times 4$  matrix  $T_P^E$  that transforms the recovered projective structure into a known metric coordinate frame,

$$\mathbf{X}_i^E = T_P^E \mathbf{X}_i^P, \quad (4)$$

where  $\mathbf{X}^E$  is a metric point and  $\mathbf{X}^P$  is a projective point. This transformation has fifteen degrees of freedom and hence at least five known world points are required. These points are provided by grid intersections on the calibration frame. The matrix  $T_P^E$  is computed by minimising,

$$d_2 = \sum_i^N \|\mathbf{X}_i^E - T_P^E \mathbf{X}_i^P\|^2, \quad (5)$$

for  $N > 5$  points. In the following  $N$  typically equals forty. The method used for minimisation is singular value decomposition [15].

## 5 Reconstruction Experiments

This section describes three experiments to measure the accuracy of our reconstruction method. Ten X-ray images were acquired, at angles of ten degrees apart starting with the coronal view and progressing to the sagittal view, as in figure 3. The coronal view was successively matched to each of the others to give stereo pairs, on which feature extraction and reconstruction were performed using the methods described in sections (2)-(4).

**Experiment 1 - accuracy of grid reconstruction:** The first experiment examines how reconstruction accuracy of hand picked grid intersections varies with angular separation. For each stereo pair, the Euclidean positions of the grid intersections were found using the methods in sections (3)-(4). The root mean square (RMS) of the Euclidean distances between the reconstructed points and ground truth was then calculated. The results are shown in Figure 6a. Figure 6b shows a plot of the number of points used in the calculation for each stereo pair. The plot shows that the RMS error decreases with increasing angular separation, as expected. Beyond a separation of fifty degrees, there are no longer any common grid points between the images. This shows that there is a tradeoff between improving accuracy (for which a large angular separation is desirable) and finding many common points (which limits the maximum angular separation). Our experimentation shows that an angle of forty or fifty degrees gives best results. This angular separation was used in experiment 3.

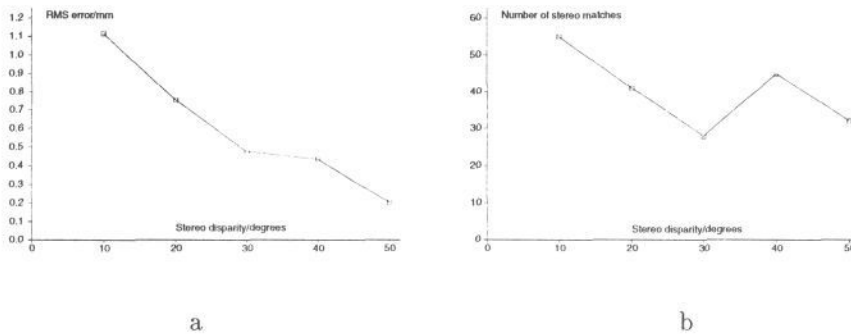


Figure 6: Results for experiment 1. (a) root mean square of the distance between grid intersections and calculated metric points, measured in mm, plotted against angular separation in degrees. (b) the number of matching grid intersections used for reconstruction.

**Experiment 2 - accuracy of feature extraction:** The second experiment examines the accuracy of automatic feature extraction versus hand picking for markers. The locations of the femoral markers were found in the ten views, using hand picking and automatic feature extraction. For each stereo pair, the fundamental matrix was calculated. For each point,  $F$  was used to find the corresponding epipolar line by  $\mathbf{l} = F\mathbf{x}$ .

The root mean square of the distances  $\sqrt{\sum_{i=1}^N \frac{d_i^2}{N}}$  between the lines and the points was found, where  $d = \frac{l_x X + l_y Y + l_z}{\sqrt{l_x^2 + l_y^2}}$ , the epipolar line is given by  $l_x x + l_y y + l_z = 0$  and the point lies at  $(X, Y)$ .

Figure 7a summarises the results. The main conclusion that we draw is that automatic segmentation gives a marked improvement in feature extraction over hand picking.

**Experiment 3 - accuracy of marker reconstruction:** The final experiment examines the accuracy of full marker reconstruction against ground truth. Marker features were extracted manually and automatically and reconstruction was done using two views 40 degrees apart. Ground truth was obtained by taking measurements on the experimental knee using a gauge with a precision of  $\pm \frac{1}{2} \text{ mm}$ . The table in 7b shows the distances between the actual marker balls and reconstructed points. These results show that good accuracy can be obtained using our method even for the hand picked points. However, the automatically extracted points gives superior accuracy. We conclude that the accuracy of the current algorithm is within the bounds needed for quantifying prosthesis migration.



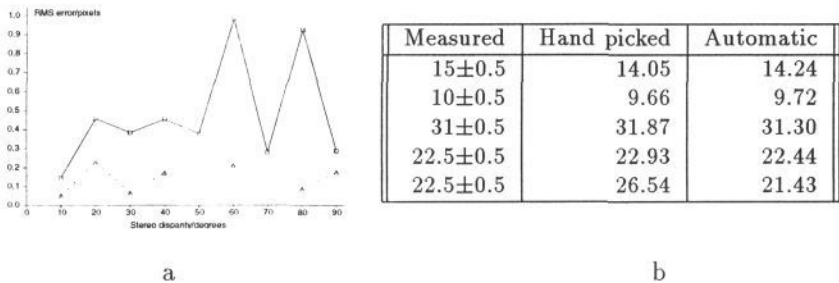


Figure 7: (a) Results for experiment 2. The x-axis shows the angular separation between views; the y-axis shows the root mean square of the distances between points and their epipolar lines. The solid line is for hand-picked points; the dashed line is for fitted points. (b) Results for experiment 3. Distances between markers, in millimetres; the first column shows measured distances, the second calculated distances based on hand picking, and the third calculated distances based on automatic extraction.

## 6 Discussion

We have presented a new approach to X-ray stereo reconstruction which is based on robust methods for automatic feature extraction and uncalibrated stereo. We have examined the accuracy and sensitivity of stages in the algorithm. Preliminary results are excellent and show that our method can achieve reconstruction accuracy to within the tolerance required of the application.

We are currently extending this work in two ways: first, assessing the advantage - principally an increase in accuracy, if multiple (more than two) stereo views are used, taking into account practical and clinical constraints. Second, we are investigating if the dependence on the calibration grid can be removed, with the prosthesis itself providing the information required for the conversion to metric structure.

**Acknowledgements:** This work has been done in collaboration with the Oxford Orthopaedic Engineering Centre. We wish to thank Dr John O'Connor and Dr David Murray for their continued enthusiasm about this project, Sue Stickney for helping to acquire the radio-graphs, and Dr Mark O'Neill for useful discussions. Thanks also to Dr Paul Beardsley for providing some of the reconstruction software. Financial support for this work has been provided by EPSRC, the Department of Health, and Esprit Project VIVA.

## References

[1] P.A. Beardsley, A.P. Zisserman, and D.W. Murray. Navigation using affine structure from motion. In *European Conf. Comp. Vision*, volume II, page 85ff. Springer-Verlag, 1994.

- [2] O.D. Faugeras. What can be seen in three dimensions with an uncalibrated camera? In *European Conf. Comp. Vision*, pages 563–578. Springer-Verlag, 1992.
- [3] B. Hallert. *X-ray photogrammetry. Basic geometry and quality*. Elsevier, Amsterdam, 1970.
- [4] R.I. Hartley, R. Gupta, and T. Chang. Stereo from uncalibrated camera. In *IEEE Conf. Comp. Vision. Patt. Recog.*, pages 761–764, 1992.
- [5] R.I. Hartley and P Sturm. Triangulation. In *Arpa Image Understanding Workshop*, pages 957–966, 1994.
- [6] J M Tobis K Kitamura and J Sklansky. Estimating the 3-d skeletons and transverse areas of coronary arteries from biplane angiograms. *IEEE Transactions on Medical Imaging*, 7(3):173–187, Sept 1988.
- [7] J Karrholm. Roentgen stereophotogrammetry - review of orthopædic applications. *Acta Orthopædica Scandinavica*, 60(4):491–503, 1989.
- [8] H.C. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. *Nature*, 293:133–135, 1981.
- [9] Q-T. Luong and O.D.Faugeras. A stability analysis for the fundamental matrix. In *European Conf. Comp. Vision*, pages 577–588, Stockholm, Sweden, 1994.
- [10] F. Meyer. Iterative Image Transformations for an Automatic Screening of Cervical Smears. *J. Histochemistry and Cytochemistry*, 27(1):128–135, 1979.
- [11] B Mjoberg. Fixation and loosening of hip prostheses. *Acta Orthopædica Scandinavica*, 62(5):500–508, 1991.
- [12] B Mjoberg. Theories of wear and loosening in hip prostheses. wear-induced loosening vs. loosening induced wear - a review. *Acta Orthopædica Scandinavica*, 65(3):361–371, 1994.
- [13] B Mjoberg, J Brismar, L Ingvar Hansson, Holgar Pettersson, Goran Selvik, and Rolf Onnerfalt. Definition of endoprosthetic loosening. *Acta Orthopædica Scandinavica*, 56:469–473, 1985.
- [14] W H Press. *Numerical Recipes in C*. Cambridge University Press, 2nd edition, 1992.
- [15] G Selvik. *Roentgen Stereophotogrammetry. A method for the study of the kinematics of the skeletal system*. PhD thesis, University of Lund, 1974.