Goal-directed Video Metrology

Ian Reid and Andrew Zisserman

Dept of Engineering Science, University of Oxford, Oxford, OX1 3PJ

Abstract. We investigate the general problem of accurate metrology from uncalibrated video sequences where only partial information is available. We show, via a specific example – plotting the position of a goal-bound soccer ball – that accurate measurements can be obtained, and that both qualitative and quantitative questions about the data can be answered.

From two video sequences of an incident captured from different viewpoints, we compute a novel (overhead) view using pairs of corresponding images. Using projective constructs we determine the point at which the vertical line through the ball pierces the ground plane *in each frame*.

Throughout we take care to consider possible sources of error and show how these may be eliminated, neglected, or we derive appropriate uncertainty measures which are propagated via a first-order analysis.

1 Introduction

The 1966 World Cup Final at Wembley Stadium, between England and West Germany, produced what is arguably the best known and most controversial goal in football history. In extra time, with the score at 2-2, Geoff Hurst the England number 10, received the ball from the right, turned, and struck a shot towards the German goal. With the goal-keeper beaten, the ball cannoned down from the crossbar, hit the ground and bounced back out into play (whence it was cleared by the German defence). English players claimed a goal – that the ball had passed completely over the line – and after consultation with his linesman, the referee concurred. England went on to win 4-2, but the controversy has never been satisfactorily resolved.

Here we resolve this controversy using video sequences of the goal. Two monocular sequences acquired from substantially different viewpoints are used for the analysis. Figure 1 shows a series of frames from each of the sequences. Using these, the question we wish to answer is *Did the ball cross the goal line?* And, if not, *How close did it come to crossing the goal line?*

The question is challenging because of the lack of available calibration:

1. The internal calibration of the cameras is unknown (and free — i.e. may well change during the sequence).
2. The motion of the cameras is unknown, and relative orientation (between stereo pairs) changes.
3. For many frames of the sequence there are few features available off the ground plane, other than moving objects — the players and the ball.

From an uncalibrated monocular sequence projective 3D measurements can be made [2, 4, 7] and upgraded to Euclidean (angles, lengths) [6, 12] for unchanging

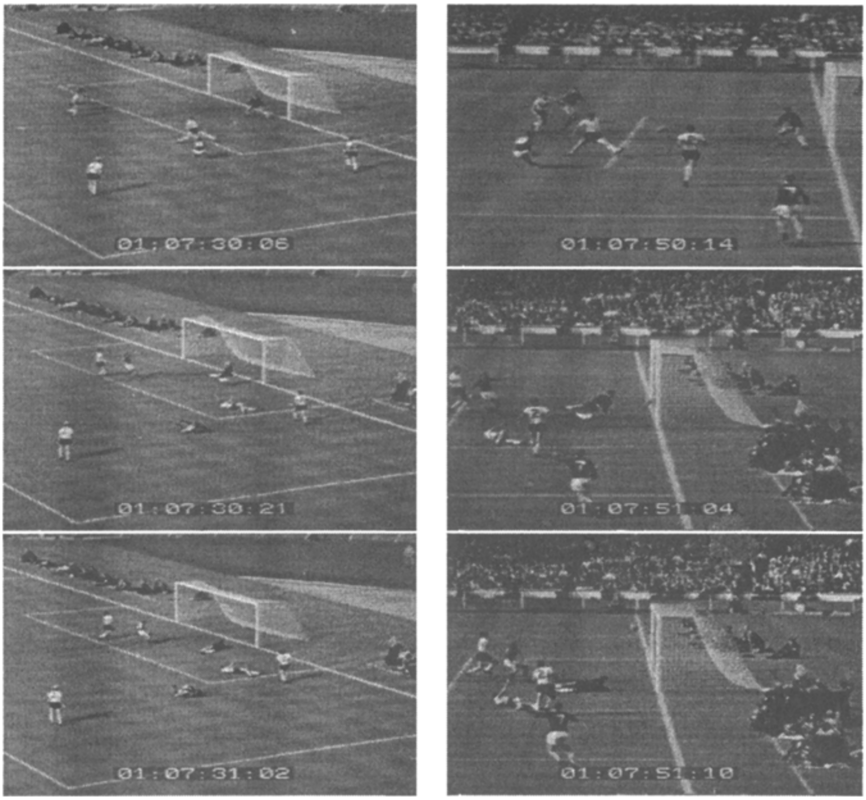

Fig. 1. Images from two available sequences of the incident.

internal parameters. These methods are not applicable for two reasons: first, because of the free internal parameters, and second because the object of interest (the ball) is moving relative to other objects in the image (in particular the ground). A binocular view (images acquired simultaneously) avoids the second problem, and, if the internal parameters and relative orientation of the cameras were also unchanging for a few frames, then 3D Euclidean structure could be recovered [1, 18, 19]. Again, the free parameters and changing relative motion prevent this. In the light of this paucity of information how can the question be answered?

It is answered by projecting the ball vertically onto the ground plane and charting the projected position and uncertainty in this position relative to the goal line. Vertical is defined by the goal posts, and metric information is provided by the dimensions of the ground plane markings. We employ the ground plane homography between views, together with vertical vanishing points, which can be computed even though point features off the ground plane are often not available. The technique is related to Quan and Mohr's [13] "shadow" algorithm for computing, from two images acquired from different viewpoints, the imaged intersection of a line with a plane. This algorithm was subsequently used to compute invariants of 3D objects [3], and for specifying points for robotic grasping [8].

The procedure here is a development on these papers in two ways: first, the intersection is not between an actual line and a plane, but between a virtual line constructed using the vertical direction vanishing point; and, second (and more importantly), a full error analysis is given for the projective transfer. The error analysis takes into account three sources of error: first, the localisation error of the points used to define the projective transformation; second, the localisation error of the computed vanishing point; and, third, the localisation error of the ball. This case study exemplifies the measurement of relative positions and their uncertainty, from uncalibrated image sequences, when there is insufficient information for a full 3D reconstruction.

We begin by describing details of the construction in section 2. The sources of error are outlined in section 3 and the implementational details, including treatment of errors, are given in section 4. Finally, results are presented in section 5.

2 Outline of method

We determine the vertical projection of the ball onto the ground plane from two images acquired simultaneously from different viewpoints. To visualise this, imagine dropping a (vertical) "plumb-line" from the ball to the ground [10]. We show that,

Given
1. two images acquired simultaneously from different viewpoints,
2. the vertical vanishing point in each image,
3. the homography (see below) between the images induced by the ground plane,
4. the images of a 3D point **B**.
then the intersection, **P**, *with the ground plane of a vertical line through the point* **B** *can be computed uniquely.*

In the following we denote world entities by upper case, and their corresponding images by lower case 3-vectors, e.g. $\mathbf{x}$ and $\mathbf{x}'$ for points, and $\mathbf{l}$ and $\mathbf{l}'$ for lines. Matrices are denoted by teletype capital letters. For homogeneous quantities, $=$ indicates equality up to a non-zero scale factor.

The geometry of the construction is illustrated in figure 2, where the line is defined by two points, **B** and **V**, in 3D. Actually, **V** is a point on the plane at infinity (an ideal point), and its images $\mathbf{v}, \mathbf{v}'$ are vanishing points, but this does not affect the projective construction. The plane projective transformation (homography) $\mathtt{T}$ between the two images via the ground plane provides a means of transferring lines between the two images. If $\mathbf{l}$ and $\mathbf{l}'$ are images of a line on the ground plane, then $\mathbf{l}' = \mathtt{T}^{-\top}\mathbf{l}$, where $\mathtt{T}$ is a 3×3 point transformation matrix: $\mathbf{x}' = \mathtt{T}\mathbf{x}$ for images of points on the ground plane.

There are four steps in the algorithm for computing $\mathbf{p}$, the image of **P**:

1. Compute the plane projective transformation, $\mathtt{T}$ between the two images from the correspondence of four lines (no three concurrent) i.e. $\mathbf{l}_i' = \mathtt{T}^{-\top}\mathbf{l}_i, i \in$

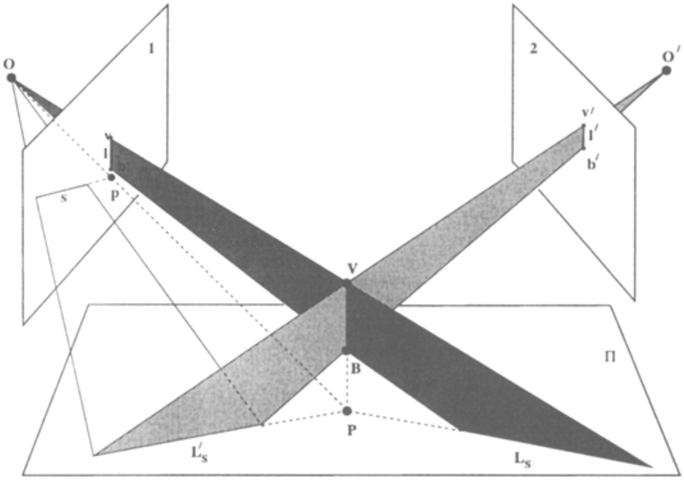

Fig. 2. **L** is a line passing through the points **V** (the vertical ideal point) and **B** (the ball), which intersects the plane π (the ground plane) in the point **P**. **l** and **l**$'$ are images of **L**, which does not lie on π. Equivalently, however, they are the images of lines $\mathbf{L}_s$ and $\mathbf{L}'_s$, respectively, which are on the ground plane — the line **L** casts a "shadow" $\mathbf{L}_s$ from view 1, where the plane π intersect the backprojection of the line **l** from the first image. A similar shadow, $\mathbf{L}'_s$, is generated from view 2. Since $\mathbf{L}'_s$ is on the ground plane, its image is $\mathtt{T}^{\mathsf{T}}\mathbf{l}'$ in image 1, where $\mathtt{T}$ is the point projective transformation between the images induced by the plane π. Since the lines $\mathbf{L}_s$ and $\mathbf{L}'_s$ intersect at **P**, their images **l** and $\mathbf{s} = \mathtt{T}^{\mathsf{T}}\mathbf{l}'$ respectively, intersect at the image **p** of **P**.

$\{1,..,4\}$. Details of this computation are given in section 4.

2. Compute the lines through the images of **V** and **B**. These lines are given by $\mathbf{l} = \mathbf{v} \times \mathbf{b}$, $\mathbf{l}' = \mathbf{v}' \times \mathbf{b}'$ in the first and second images respectively.
3. Transfer the line $\mathbf{l}'$ from the second onto the first image as $\mathbf{s} = \mathtt{T}^{\mathsf{T}}\mathbf{l}'$.
4. Then the image of the intersection point is $\mathbf{p} = \mathbf{s} \times \mathbf{l}$ in the first image. A similar construction determines $\mathbf{p}'$, the imaged intersection in the second image, as $\mathbf{p}' = (\mathtt{T}^{-\mathsf{T}}\mathbf{l}) \times \mathbf{l}'$.

This computation can also be transferred to a plan (rectified) view of the ground plane using the projective transformation between the points/lines on the ground plane and their images. In this case the six-yard goal markings are known (up to a plane Euclidean transformation), and these relative measurements provide metric calibration, and a 2D frame in which to evaluate uncertainty.

3 Sources of error

Potential sources of error are discussed in the following sub-sections. In each case it is demonstrated that these potential errors did not arise, or could be accounted for.

Synchronisation

The synchronisation of the two sequences used is an essential assumption of the method. The time difference between frames is assessed by computing the ground plane homography between the two images using features corresponding to *fixtures* (such as marked lines on the ground, and goal posts), and measuring the error when this transformation is applied to features corresponding to moving objects (player's shadows). The error is the pixel distance between the actual and transferred point. The synchronisation error could be up to 20ms between video frames (1/48 s between film frames).

Eight points are obtained to sub-pixel accuracy by intersecting straight lines fitted to the ground plane markings. The homography is then computed using a combination of linear and non-linear minimisation where the cost function is the transfer error.

The accuracy of the transformations is first assessed by measuring the error for fixtures not used in the computation of the transformation. Errors are typically less than two pixels (e.g. for the ground plane computation 8 matches are available, 6 are used to compute the homography, and the error measured on the remaining two). For moving objects the error between *corresponding* frames of sequences are similar to the fixture error, whilst for a *near corresponding* frame, the errors exceed 10 pixels. Figure 3 illustrates these cases. In summary, the ground plane homography is used to establish that the two sequences are "perfectly" synchronised.

Radial distortion

In order to take advantage of projective geometry we require that the image formation process be described accurately by a central projection model. This model is invalid if there is any significant lens distortion, the most common type of which is radial distortion. One manifestation of radial distortion is bending of straight lines near the periphery. We have therefore tested its effect by fitting lines to known straight features in the periphery of images in each sequence. Figure 4 shows two typical images and corresponding residuals after an orthogonal regression fit to putative straight edge data. The lines fitted are superimposed on the images. The side view shows no distortion (residuals are distributed evenly either side of the fitted line), while a small, but for our purposes insignificant, amount of distortion is apparent in the three quarter view (obtained with a wider angle lens).

Straightness of lines/planarity of ground "plane"

If an imaged line remains straight through a range of viewpoints then this is compelling evidence that the world line is straight. Similarly, the straightness of a number of transverse lines on a surface is evidence that the surface is planar. The image measured straightness of all lines of the six-yard markings throughout both sequences indicates the planarity of the ground.

Motion blur

One further potential source of error is motion blur. As the ball moves (while the camera is stationary) it is significantly blurred in the direction of motion.

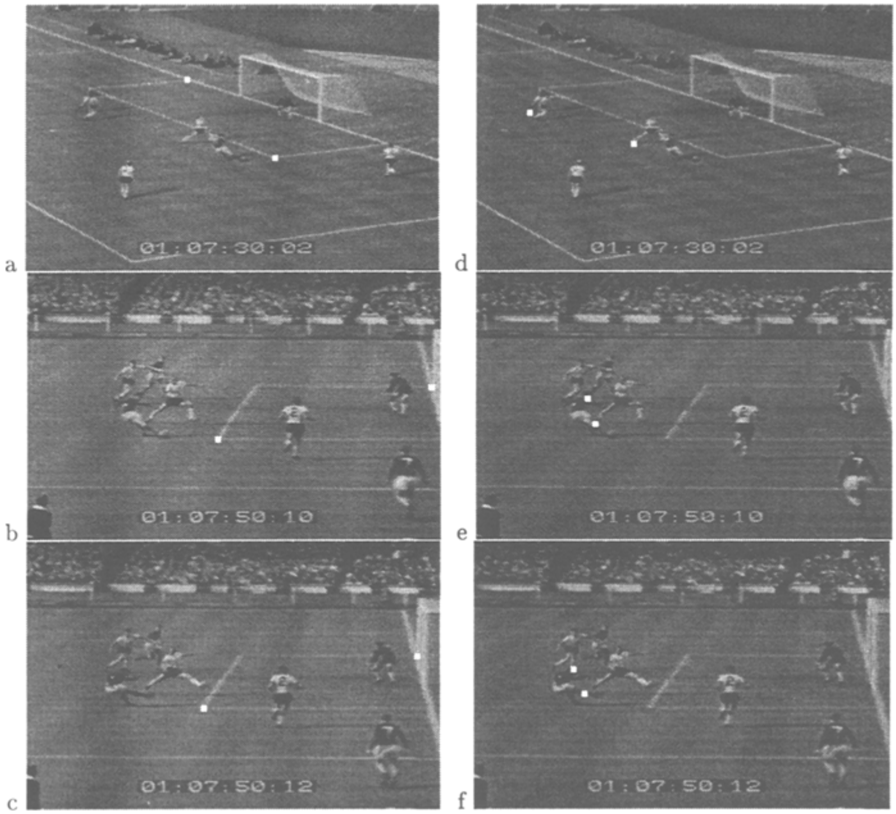

Fig. 3. Assessing the synchronisation of the sequences. (a) is a frame from sequence 1, and (b) (c) are near corresponding frames from sequence 2. The ground plane homographies between the frames (a & b, and a & c) are computed from imaged **fixtures**. Two points (fixtures) are marked on the ground plane of frame (a) and transferred to frames (b) and (c) using the appropriate homography. The disparity between transferred and actual position is negligible (i.e. less that a pixel), indicating the accuracy of the computed homography. (d) (e) (f) are the same frames with points corresponding to **moving** objects marked in (d). The points chosen are the left most point of the shadow of each player. The transferred points are superimposed on (e) and (f). In (e) the correspondence between transferred and actual position is again negligible, indicating that the frames are synchronised. However, in (f) there is a significant discrepancy (10 pixels) indicating that the frames are not synchronised.

Similarly, as the ball is tracked by the cameraman, the stationary features in the environment are observed to blur.

Fortunately during the crucial frames in which the ball is close to crossing the line, there is little blur due to camera motion. That which there is, is accounted for in the line uncertainty by inflating the line covariance appropriately. The significant blur is due almost entirely to the motion of the ball. In this case we take the blur into account by a greater uncertainty in the ball location in the direction of motion.

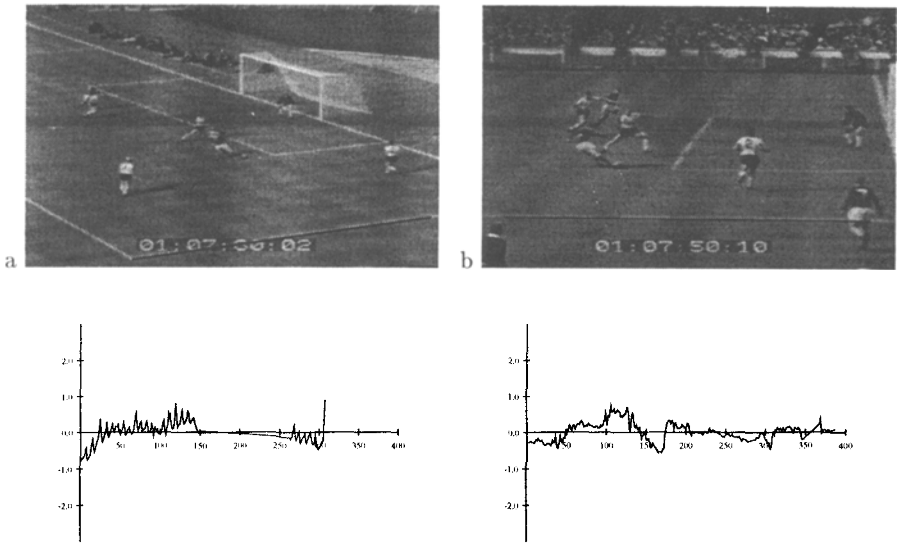

Fig. 4. Testing for radial distortion. (a) an image from the three-quarter view sequence with a detected line (shown black) near the periphery. The residuals of an orthogonal regression fit are shown in the graph below. The axes are in pixels. (b) is for a line from the side sequence.

4 Implementation and error analysis

In this section we discuss the representation, computation, and uncertainty of the geometric primitives and transformations required for the analysis. We follow the approach of [5, 11] computing uncertainty propagation via first-order approximations. We have verified the validity of the first-order model using Monte Carlo techniques. Full details are given in [15]. The covariance of a vector is denoted $\Lambda_{\mathbf{x}}$. The dimension of a matrix is indicated where necessary, in parentheses: e.g. $\Lambda_{\mathbf{x}}(2 \times 2)$.

Lines

Line segments are computed using orthogonal regression on a set of canny edge strings. Each (manually selected) set of edge strings is processed using the RANSAC algorithm [17] to enforce collinearity, adding greatly to the robustness of the line fitting by providing rigorous outlier rejection and by enabling multiple strings to contribute to one line segment.

Lines are represented both as homogeneous three vectors and by (inhomogeneous) two-parameter representations:

$$\mathbf{a} = [a, c]^{\top} \text{ such that } \begin{array}{ll} ax + y + c = 0, & \text{if the line is closer to horizontal} \\ x + ay + c = 0, & \text{if the line is closer to vertical} \end{array}$$

The uncertainty of a line is represented by a 2×2 covariance matrix, $\Lambda_{\mathbf{a}}$, which

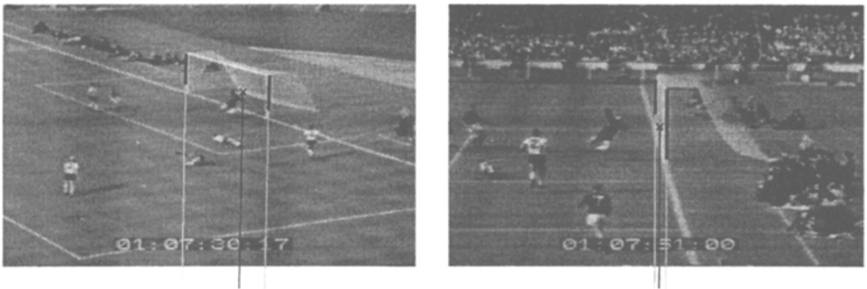

Fig. 5. Estimation of the vertical vanishing direction via the goal-posts. The × marks the ball, and the extended black line indicates the image of a vertical line through the ball (it joins the ball to the vanishing point computed from the goal-posts).

is computed from edgel uncertainties using the method described in Chapter 5 of [5], and where necessary, by a 3×3 form for this covariance, denoted $\Lambda_{\mathbf{a}}(3 \times 3)$.

Points

Points are generally localised by intersecting lines, and their covariance computed from the line uncertainty. Where this is not possible, e.g. the ball centre, the point is picked with a mouse, in which case the uncertainty is estimated as the mouse precision (about ±1 pixel in each direction). The uncertainty in a point's position is represented by the 2×2 covariance matrix $\Lambda_{\mathbf{x}}$.

The vertical vanishing point is obtained by intersecting lines computed from the obvious vertical cues in each image – the goal-posts. Figure 5 shows an example. Care must be exercised when intersecting lines to find vanishing points, since the final component of the homogeneous representation may be close to zero, rendering the computations of the inhomogeneous coordinates and covariance unstable. For the case of vertical vanishing points which are of special interest here, we derive inhomogeneous coordinates from $\mathbf{v} = [v_x, v_y, v_z]^\top$ as $[v_x/v_y, v_z/v_y]^\top$ and the covariance calculation is modified appropriately.

Intersections

The intersection of two lines is given by the cross-product of homogeneous lines, $\mathbf{v} = \mathbf{l}_1 \times \mathbf{l}_2$. Letting $\mathbf{x} = [v_x/v_z, v_y/v_z]^\top$, we obtain the uncertainty in the location of the intersection using a first-order error analysis:

$$\Lambda_{\mathbf{x}} = \mathtt{D} \begin{bmatrix} \Lambda_{a_1}(3 \times 3) & 0 \\ 0 & \Lambda_{a_2}(3 \times 3) \end{bmatrix} \mathtt{D}^\top, \quad \text{where} \quad \mathtt{D}(2 \times 6) = \left[\frac{\partial \mathbf{x}}{\partial \mathbf{l}_1} \mid \frac{\partial \mathbf{x}}{\partial \mathbf{l}_2} \right]$$

Ground plane homography

The homography between a camera view and the plan view, $\mathtt{T}$, is obtained from the corners of the six-yard area[1]. It is computed from image positions $\mathbf{p}$, which are estimated accurately by the intersection of extended lines, and

[1] We also compute the line transformation using four line correspondences, but have omitted the discussion here since it is analogous to that for points, although complicated by the need for two different parameterisations of lines.

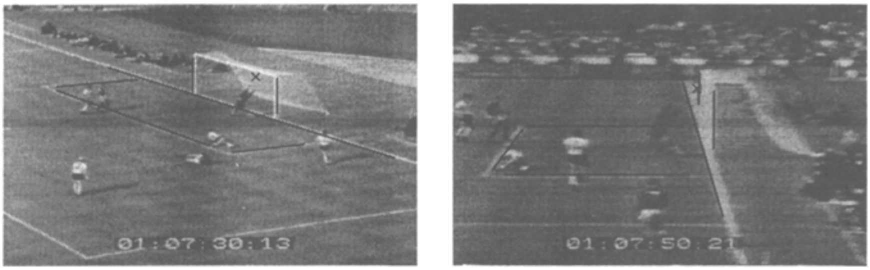

Fig. 6. The three-quarter and side views of the action just before the ball strikes the crossbar, showing the positions of lines used for registration.

corresponding plan view positions $\mathbf{P}$, which have known (exact) values. Each correspondence gives rise to two independent linear equations in the eight unknowns of $\mathtt{T}$, so from four such correspondences we can construct and solve the 8×8 matrix equation $\mathtt{A}\mathbf{t} = \mathbf{b}$, where $\mathbf{t}$ is an 8-vector of the free parameters of $\mathtt{T}$.

Defining the vector $\mathbf{z}$ to be the 8-vector containing the inhomogeneous coordinates of the image points $\mathbf{p}_{1\ldots4}$ (hence $\Lambda_{\mathbf{z}}$ is a block diagonal matrix consisting of the 2×2 submatrices $\Lambda_{\mathbf{x}_i}$ where $\mathbf{x} = [p_x/p_z, p_y/p_z]^{\top}$), it is straightforward to show that:

$$\Lambda_{\mathbf{t}} = \mathtt{D}\,\Lambda_{\mathbf{z}}\,\mathtt{D}^{\top} \qquad \text{where} \qquad \mathtt{D}(8 \times 8) = \frac{\partial \mathbf{t}}{\partial \mathbf{z}} = -\mathtt{A}^{-1}\frac{\partial \mathtt{A}}{\partial \mathbf{z}}\mathbf{t}$$

Transforming primitives

The final aspect of uncertainty which must be considered, is how to compute the uncertainty of a transformed primitive, when both the primitive and the transformation are uncertain. In the case of points (lines are analogous but once again, complicated slightly by the necessity for two different representations) the transformation is given by $\mathbf{P} = \mathtt{T}\mathbf{p}$. Thus the uncertainty in $\mathbf{X} = [P_x/P_z, P_y/P_z]^{\top}$ depends on $\Lambda_{\mathbf{x}}$ and $\Lambda_{\mathbf{t}}$ the covariances of the image position $\mathbf{x} = [p_x/p_z, p_y/p_z]^{\top}$ and homography $\mathtt{T}$ respectively, and is given by

$$\Lambda_{\mathbf{X}} = \mathtt{D}\begin{bmatrix} \Lambda_{\mathbf{x}} & 0 \\ 0 & \Lambda_{\mathbf{t}} \end{bmatrix}\mathtt{D}^{\top} \qquad \text{where} \qquad \mathtt{D}(2 \times 10) = \frac{\partial \mathbf{X}}{\partial\{\mathbf{x}, \mathbf{t}\}}$$

5 Results

As indicated previously, the four lines of the six-yard area are used to register each image with the plan view, and the goal-posts are used to determine the vertical vanishing direction. The centre of the ball is picked manually with a mouse (and its uncertainty set to reflect the error introduced by this process). Figure 6 shows one pair from the sequence with the lines and ball position superimposed.

The rectified six-yard area is shown in figure 7a, with the *centre* of the ball, and its covariance, for the frame before the ball strikes the crossbar. The ellipse represents the 3σ limit, meaning there is, practically speaking, no chance that the centre of the ball is outside this ellipse. The transferred vertical lines used

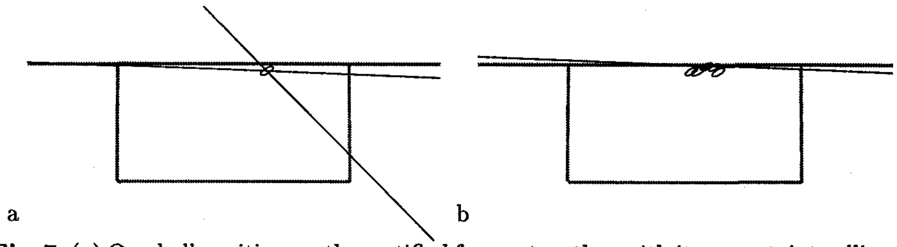

Fig. 7. (a) One ball position on the rectified frame, together with its uncertainty ellipse. The top line is the *front* of the (finite width) goal line; (b) Uncertainty ellipses and constraint lines for the crucial frames of the sequence.

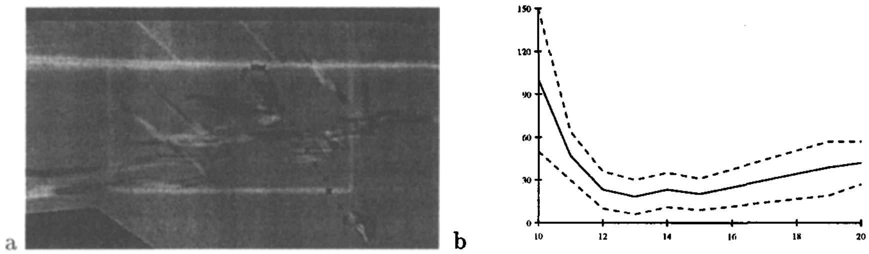

Fig. 8. (a) The computed position of the ball throughout the sequence; (b) distance away from being a goal (in cm) versus frame number for the crucial frames of the sequence.

to compute the position are shown, with the ball position being their point of intersection. The uncertainty ellipses for all of the computed ball positions from just before the ball struck the crossbar to the point where it reenters play can be seen in figure 7b. One point of note here is that there are four frames in the middle of this sequence when a ball position cannot be computed because the goal-keeper obscures the ball in the three quarter view. This does not affect our ability to decide the question of whether it crossed the line or not, since we still have one constraint on the ball position: The transferred vertical line which represents this constraint has been drawn in the figure for the "missing" frames in which the ball strikes the ground, clearly showing that wherever the ball is placed along this line, we can still say with certainty that it was not across the line.

The complete set of computed ball positions from the moment Hurst shot for goal, to the point where it hit the crossbar, then the ground, and finally back out into play is shown in figure 8(a), rendered with the rectified texture (from the last image pair). The answer to the question, "did the ball cross the line?" must also take the ball radius into account, and a more quantitative analysis is given in the graph of figure 8(b) which shows the distance of the ball from being a goal (taking its radius into account) plotted against frame number. The dotted lines indicate three standard deviations from the estimate, thus a conservative estimate has the ball still 6cm from being a goal.

6 Conclusions

While it has been known for some time that 3D structure can be computed from uncalibrated views of a scene given sufficient correspondences in general position, this has rarely been used to answer specific, metric questions about the data. The approach taken here has been to make use of plane projective homographies to compute an overhead view of the action from a sequence of disparate image pairs. An alternative approach might have used virtual parallax, as described in [16], but this could only give a qualitative answer and would have required that the whole of the goal-mouth be visible in all frames. Another, and more convenient method, would have been to use an affine approximation to the imaging geometry — since fewer features would have been required. However, this approximation was found to be insufficiently accurate for these sequences as the images exhibit non-negligible perspective effects. Thus, although in the past the use of affine structure has proved fruitful for various active vision tasks [8, 14] such as tracking or visual servoing, it is not well suited to tackling quantitative measurements tasks unless the projection model truly is affine.

The application we have presented is one of a wider class of problems (such as traffic monitoring) in which the ground plane trajectory of a target is desired, but in which the camera calibration is difficult or impossible to obtain with sufficient accuracy. Sporting domains are often ground-plane orientated, with well known regular marking which can be used for registration, and so are particularly suited to this analysis [9].

While at the same time providing a compelling example of the power of uncalibrated techniques, this work has made a tangible contribution in settling, once and for all, the argument over whether or not the ball crossed the line in the most famous goal of all. In describing the closing seconds of the match, commentator Kenneth Wolstenholme said of celebrating English fans: *They think it's all over. It is now!* Nearly thirty years on, his words are once again appropriate.

Acknowledgements

We are very grateful to Adam Jones and John Davison of the Sunday Times for supplying the video sequences, and for their suggestions together with those of Mike Brooks, Roberto Cipolla and Peter Sander. Financial support was provided by Esprit Project VIVA, ACTS Project VANGUARD and a Glasstone Fellowship.

References

1. P. A. Beardsley, I. D. Reid, A. Zisserman, and D. W. Murray. Active visual navigation using non-metric structure. In *Proc. 5th Int'l Conf. on Computer Vision, Boston*, pages 58–65. IEEE Computer Society Press, 1995.
2. P. A. Beardsley, A. Zisserman, and D. W. Murray. Navigation using affine structure from motion. In *Proc. 3rd European Conf. on Computer Vision, Stockholm*, volume 2, pages 85–96, 1994.
3. S. Demey, A. Zisserman, and P. Beardsley. Affine and projective structure from motion. In D. Hogg and R. Boyle, editors, *Proc. 3rd British Machine Vision Conf., Leeds*, pages 49–58. Springer-Verlag, September 1992.

4. O. D. Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig? In G. Sandini, editor, *Proc. 2nd European Conf. on Computer Vision, Santa Margharita Ligure, Italy*, pages 563–578. Springer-Verlag, 1992.
5. O.D. Faugeras. *Three-Dimensional Computer Vision.* MIT Press, 1993.
6. R. I. Hartley. Self-calibration from multiple views with a rotating camera. In *Proc. 3rd European Conf. on Computer Vision, Stockholm*, volume 1, pages 471–478, 1994.
7. R.I. Hartley, R. Gupta, and T. Chang. Stereo from uncalibrated cameras. In *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, pages 761–764, 1992.
8. N. Hollinghurst and R. Cipolla. Uncalibrated stereo hand/eye coordination. In J. Illingworth, editor, *Proc. 4th British Machine Vision Conf., Guildford*, pages 389–398. BMVA Press, 1993.
9. S. S. Intille and A. F. Bobick. Closed-world tracking. In *Proc. 5th Int'l Conf. on Computer Vision, Boston*, pages 672–678, 1995.
10. A. Jones and J. Davison. Sport: how science can end disputes. *Sunday Times*, 23 July, 1995.
11. K. Kanatani. *Statistical optimization for geometric computation: theory and practice.* AI Lab, Dept of Computer Science, Gunma University, Japan, 1995.
12. S.J. Maybank and O. Faugeras. A theory of self-calibration of a moving camera. *International Journal of Computer Vision*, 8(2):123–151, 1992.
13. L. Quan and R. Mohr. Towards structure from motion for linear features through reference points. In *Proc. IEEE Workshop on Visual Motion*, 1991.
14. I. D. Reid and D. W. Murray. Tracking foveated corner clusters using affine structure. In *Proc. 4th Int'l Conf. on Computer Vision, Berlin*, pages 76–83, Los Alamitos, CA, 1993. IEEE Computer Society Press.
15. I. D. Reid and A. Zisserman. Accurate metrology in uncalibrated video sequences. Technical report, Oxford University, Dept. of Engineering Science, 1996.
16. L. Robert and O. Faugeras. Relative 3d positioning and 3d convex hull computation from a weakly calibrated stereo pair. In *Proc. 4th Int'l Conf. on Computer Vision, Berlin*, pages 540–544, 1993.
17. P.H.S. Torr and D.W. Murray. Outlier detection and motion segmentation. In *Proc SPIE Sensor Fusion VI*, pages 432–443, Boston, September 1993.
18. Z. Zhang, Q.-T. Luong, and O. Faugeras. Motion of an uncalibrated stereo rig: self-calibration and metric reconstruction. Technical Report 2079, INRIA Sophia-Antipolis, October 1993.
19. A. Zisserman, P. A. Beardsley, and I. D. Reid. Metric calibration of a stereo rig. In *Proc. IEEE Workshop on Representations of Visual Scenes, Boston*, pages 93–100. IEEE Computer Society Press, 1995.