

## **A VERSATILE CAMERA POSITION MEASUREMENT SYSTEM FOR VIRTUAL REALITY TV PRODUCTION**

**G A Thomas (BBC)**

**J Jin, T Niblett, C Urquhart (The Turing Institute)**

### **ABSTRACT**

A key component of any virtual production system is a means of measuring the precise position and orientation of each studio camera, so that the virtual scene can be rendered from the appropriate viewpoint. There are a number of systems already available that provide camera position data, but none fully address the requirements of a hand-held camera operating in a very large studio. A system designed specifically for this application is therefore being developed. The system works optically, using a number of markers in known positions. The markers may either be placed in the scene and viewed by the main camera, or placed out of shot (for example, on the ceiling) and viewed by a separate camera mounted on the main one. The computational requirements are not excessive, making the system amenable to real-time implementation using either a general-purpose workstation with a frame-grabber, or a small amount of dedicated hardware.

### **INTRODUCTION**

The essential difference between a conventional blue-screen studio and a true 'virtual studio' is the ability to move the studio camera whilst maintaining the correct registration between the foreground and the virtual background. The method used to measure the camera position and orientation will have a significant effect on the ease of use and realism that can be achieved, and may also be a significant factor in the cost of the installation.

Ideally, a camera position measurement system should:

- allow unconstrained movement of a camera over an area up to about 800 square metres;
- work with a wide variety of camera mountings (including manual and robotic pedestals, cranes and hand-held);
- measure the position and orientation to a sufficient accuracy to introduce negligible drift or noise in the relative positions of real and virtual elements of the scene. For example, to maintain relative positions to an accuracy of  $\pm 0.5$  pixel, with a minimum field of view of 10 degrees or 50cm, requires the positional accuracy to be of the order of  $\pm 1$ mm and the angular accuracy to be about  $\pm 0.01$  degrees;
- measure the camera parameters with minimal delay. Long delays cause numerous operational problems, such as making navigation through the virtual world very difficult for the cameraman.
- place no significant constraints on either the scene content or the studio environment.

There are a number of systems already available on the market that provide camera position data, but none fulfil all these requirements. In particular, the use of hand-held cameras in a

very large studio poses significant problems. We therefore set out to develop a system specifically to address this scenario.

The remainder of this paper describes two approaches that we developed. The first one is best suited to temporary installations and pre-recorded programme production, whereas the second is better suited to live programmes, but requires more time to install in a studio. The second approach is discussed in more detail, since this one is best suited to typical TV production requirements.

## **APPROACHES TO THE PROBLEM**

### **First approach**

Our initial work was driven by an application requiring the free movement of a film camera throughout a very large studio. The requirement was for the camera to be mounted on a 'Steadicam' mounting, which precluded the use of any mechanical systems to measure its position. However, the camera lens was to be of fixed focal length, removing the need to measure changes in zoom.

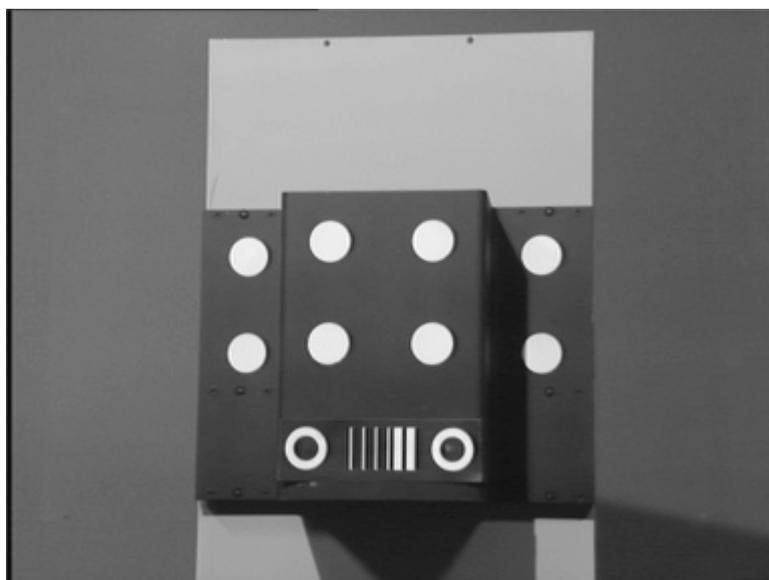
After considering a number of possibilities, an optical system based on markers placed in the scene appeared to be most suitable. The image is analysed to locate the exact position of each marker, from which the camera position and orientation (three position coordinates and three rotation angles) can be computed, given the known positions of the markers in the studio. During the shooting, the video signal from the electronic viewfinder (the 'video assist') is used as input to the image analysis processing. In post-production, the camera position may be re-computed by analysing the images on the film. This would provide more accurate measurements (due to the higher resolution of the image on the film) and allow the apparent motion introduced by effects such as film weave to be taken into account. The internal camera parameters (including focal length and pixel size) need to be known, but these can be determined by a separate initial calibration procedure, since they remain fixed for a given choice of lens.

The only constraint this system places on the positioning of markers is that at least one, and ideally a minimum of around three, need to be visible at all times. As many markers as are needed to cover the area of the action can be used, placed so as not to interfere with the movement of actors. This approach is inherently extendable to very large studios, and allows for scenes involving significant actor and camera movement, such as following actors running down long corridors.

Since the markers can be re-positioned to suit each scene being shot, it was important to include a method to automatically measure their positions to a high accuracy. This was achieved by capturing a number of images, each showing two or more markers. The identities of the markers in each image are determined automatically. A global least squares adjustment determines the relative position and orientation of all markers across the various images.

Each marker consists of eight circles in one shade of blue against a background in a slightly different shade, in order to allow them to be keyed out. The use of two-tone blue for camera position measurement has already been used effectively in earlier work, by Thomas [1] and more recently by Adler [2]. An example of a marker is shown in Fig. 1. The marker is designed so that the circles are not all co-planar; this allows the camera position to be computed unambiguously from an image containing a single marker, and also facilitates the automatic measurement of the marker positions. The marker shown in Fig. 1 has the central four circles mounted 10cm in front of the outer circles. The overall dimensions of this marker are 40x30x10cm. Each marker has a barcode to identify it uniquely.

The system was designed to run in real-time using a standard workstation equipped with a frame grabber. To achieve this, it was necessary to track the position of markers from one frame to the next, so that only a small fraction of the image (around the predicted marker positions) needed to be searched. Furthermore, this allowed the use of a simple algorithm to find the centre of each circle. Using this approach, it was possible to measure the camera position at a rate of 25Hz (when tracking a maximum of 20 circles) using a PC based on a 200MHz Pentium Pro. When the system is started, however, the entire image *does* need to be searched, which takes several seconds.



*Fig. 1 - A prototype marker attached to a stand.*

## **Second approach**

During the development of the system described above, the need arose for a system better suited to use in a live studio. The first approach relied on markers visible in-shot, which could be keyed out by virtue of being coloured in two shades of blue. In order to allow shadows to be extracted during the keying process, some care would have been needed during the post-production phase, to eliminate shadows caused by the markers themselves and by the variation of blue within them. Whilst this would have been acceptable for a production on film with a significant post-production budget, it was not possible with a live broadcast. Furthermore, the first system needed to rely on a tracking process to allow markers to be located; if the system ever lost track of the camera position (for example, if the camera lens was temporarily covered whilst the camera was moving), the system could not recover instantaneously, since a complete search of the image for markers was required. This could not be completed within a frame period using a conventional workstation. Although this shortcoming was not of major significance for the type of production for which it was designed, it was clearly not acceptable for live use.

Therefore an alternative system was designed. The in-shot markers were replaced with markers mounted out-of-shot, for example on the ceiling, and a small auxiliary camera (attached to the side of the main camera) was used to view them. A similar method has previously been used by Azuma and Bishop (3) to determine the position of a headset in an 'Augmented Reality' system. One disadvantage of this approach compared to the first is that the system takes some time to install (markers need to be fitted to the ceiling), whereas a system based on the first approach could be set up in any studio in a matter of minutes. Also, the need for an auxiliary camera increases the cost and complexity. However, there are also many advantages. Since the markers are no longer visible by the main camera, all constraints on scene composition (the need for several markers to be in-shot all the time, and the

presence of several shades of blue in the image) are removed. Indeed, the system can be used when there is no blue background in the scene, for example when placing a virtual object in a real scene. Furthermore, the markers can be of high contrast, rather than composed of two similar shades of blue, making it feasible to identify them without having to track them from frame-to-frame. A suitable way of achieving high contrast is to use retroreflective material for the markers, and to illuminate them by a light mounted adjacent to the camera. They may then be located using a relatively simple algorithm, even against a cluttered background containing, for example, studio lights.

Fig. 2 shows an example of the arrangement in a studio. Examination of some typical studio ceilings suggested that there would generally be sufficient space between existing obstructions, such as lighting hoists and scenery tracks, to mount a sufficient number of markers. Although the markers can be positioned in a coplanar manner, the accuracy of the computed camera position is significantly improved if the markers are placed at differing distances from the camera. For example, some markers could be mounted directly on the ceiling of a studio, with others mounted on girders or poles hanging a little way below.

The following sections describe this approach in more detail.

### Camera calibration

The auxiliary camera needs to have its internal parameters (such as focal length) accurately measured. This can be carried out when the system is installed, and only needs to be repeated periodically to ensure that the calibration has not changed. Images of a 'calibration object' (an object having a number of circles in precisely known positions, resembling the marker shown in Fig. 1) may be analyzed to determine these parameters. The relative orientation and position of the auxiliary camera and the main camera also need to be measured. The settings of zoom and focus of the main camera must be continuously measured using conventional mechanical sensors.

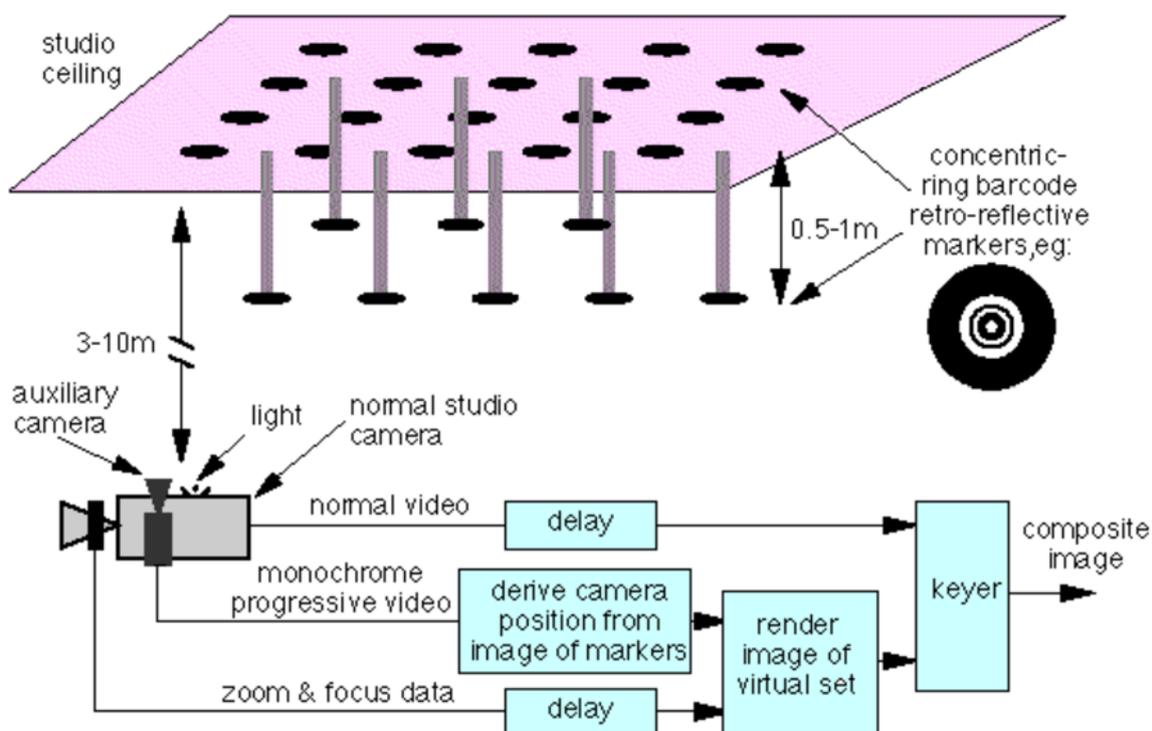


Fig. 2 - Arrangement used in the second approach to camera position measurement.

### Marker design

The marker design uses individual circles, rather than groups of eight, to simplify manufacture and mounting. The barcode, which identifies each marker, is incorporated within the circle itself, in the form of a set of concentric rings. The number of bits in the barcode

may be chosen to suit the size of the studio. For example, a 9-bit barcode is sufficient to identify uniquely every marker in a studio having a floor area up to about 50m<sup>2</sup>. For larger studios, more bits may be used. Alternatively, some codes may be allocated to several markers, and the system can deduce which marker is which by reference, for example, to the nearest uniquely-numbered marker.

The minimum size of the markers must be chosen to ensure that the barcode can be successfully read by the auxiliary camera. We have found that our barcode design and reading algorithm require a minimum spacing between rings corresponding to about 1.5 pixels on the camera sensor to allow reliable reading. Thus, if the markers were about 4m above the auxiliary camera, each contained a 9-bit barcode, and a conventional-resolution camera with a minimum field of view of 30 degrees was used, the minimum marker diameter is approximately 12cm.

### **Setting up the markers**

The markers need to be positioned so that the auxiliary camera can view a minimum of three, and ideally 10 or more, at any one time. Allowance must be made for some markers being obscured by lights, microphone booms, and so on. For example, consider a situation with markers at heights of 4m and 4.5m, a maximum working height of the camera of 2m and the minimum field of view of the auxiliary camera of 30 degrees. Positioning markers in a square lattice at 0.4m intervals would ensure that at least 12 were potentially visible, which should allow good performance even if half of these were obscured by obstacles such as lights.

The position of each marker needs to be known to within about 1mm in order to meet the overall accuracy requirements. This could be achieved by measuring the position of the markers using a theodolite, although this would be time-consuming. An alternative method has been developed, which operates in a similar manner to the automatic marker measurement system used in the first approach. First, the positions of the markers are measured roughly, to an accuracy of a few cm. A number of images of the markers are then analyzed to yield accurate positions.

### **System implementation**

A monochrome progressive-scan camera is used for the auxiliary camera, with a short integration time to minimise image blur when the camera moves. The processing has been implemented in software on an SGI 'O2' workstation, and can measure the camera position at a maximum rate of about 20 measurements per second. The majority of this time is spent in the initial marker location processing, which performs several simple thresholding and filtering processes on a subsampled image to detect regions of white surrounded by black. Additional camera position measurements may be interpolated to increase the effective measurement rate to 50 or 60 measurements per second as required.

Although the implementation based on a workstation can run just about fast enough for a practical application, there are several reasons why the use of a dedicated hardware unit to perform the processing would be preferable. These include a higher measurement rate, and a reduced processing delay (the processing of an image can start as soon as the first part of the image is captured, rather than having to pass through several layers of frame buffering as is commonly found in frame grabbers).

A hardware unit is therefore being designed. It will occupy a single printed-circuit board, and comprise some dedicated circuitry for implementing the filtering and thresholding used for marker detection, plus a digital signal processor to read the barcodes and perform the 3D camera position calculation. It is designed to compute the camera position within one frame period of the arrival of the last line of the video frame. It will be capable of making measurements at rates of up to 50 or 60Hz, although in practice the achievable rate is likely to be limited by the frame rate of the progressive camera. The camera we have chosen to use

runs at 30Hz, so temporal interpolation will be used to generate measurements at 50 or 60Hz as required.

### Experimental results

To evaluate the accuracy of the system, an experimental arrangement was set up in a studio. 24 markers were set up on one side of a studio; 16 were mounted directly on the wall, and the remaining 8 were mounted about 40cm in front of the wall. Their positions were first measured roughly, and then refined by analysing 14 images showing the markers viewed from various positions and angles. The auxiliary camera was then mounted on an optical bench, so that it could be moved in a straight line whilst maintaining a fixed angle. The arrangement is shown in Fig. 3.



*Fig. 3 - Experimental set-up for measuring the accuracy of the system.*

The random noise in the position and angle measurements was first assessed with the camera in a fixed position, with 15 markers in view. The standard deviations of the position and angle measurements computed from 100 consecutive images are given in Table 1. These are well within the accuracy requirements discussed earlier.

*Table 1:  
Standard deviations of measured position and angles for a stationary camera.*

	angle (degrees)	position (mm)
x	0.0020	0.073
y	0.0015	0.090
z	0.00072	0.030

The accuracy of the system was then assessed as the camera was moved along the optical bench. The bench was aligned so as to be parallel to the x axis (running left to right), and normal to the y (height) and z (depth) axes. Measurements were taken every 5cm. Fig. 4 shows the measured x position of the camera, plotted against its position read from the vernier scale on the optical bench. The measured position agreed with the actual position to within about 1mm over the whole length.

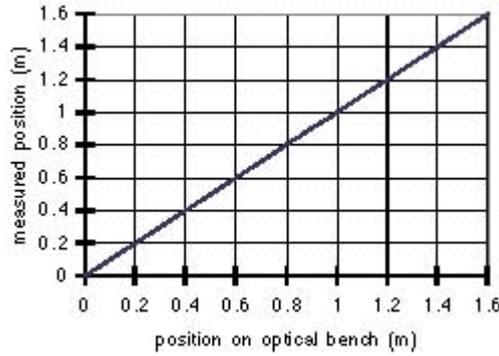


Fig. 4 - Measured horizontal camera position during movement along optical bench.

A further experiment was conducted to measure the noise in the system in the presence of camera movement. The camera was slowly pushed along the bench over a distance of about 0.8m, with about 500 measurements being taken during the move. Measurements were also taken before and after the move.

The position of the camera in the y and z directions as measured by the system are plotted in Fig. 5. For clarity, these are shown relative to an origin chosen to make the mean camera position zero. These values would be expected to remain constant, since the optical bench was normal to both axes. The positions were indeed constant to within about  $\pm 1$ mm overall, with the majority of this variation changing smoothly with camera position. Over small distances of the order of 10cm, the positions were constant to within about  $\pm 0.25$ mm.

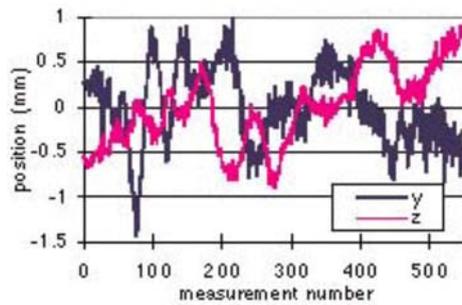


Fig. 5 - Measured y and z camera coordinates during the movement.

The variation in tilt (x), pan (y), and roll (z) angles measured during the move are plotted in Fig. 6. Over small distances, the angles varied by about  $\pm 0.01$  degrees. This is larger than the variations due to noise when the camera was stationary, but still meets the accuracy requirement discussed earlier.

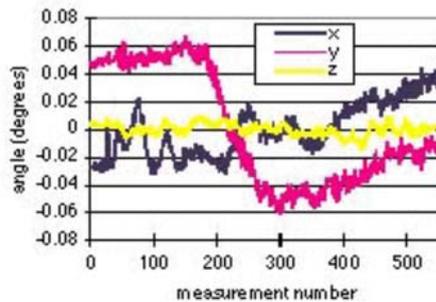
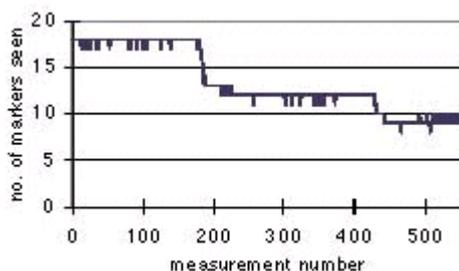


Fig. 6 - Measured camera angles during the movement.

Over longer distances the changes were slightly larger, up to about  $\pm 0.06$  degrees. It is unlikely that such changes in angle will be a problem, since they varied smoothly with camera position in a repeatable manner. These changes were probably due to the system seeing

different markers as the camera moved. Small inaccuracies in the measured positions of the markers will effectively give rise to changes in the position and orientation of the coordinate frame as different groups of markers are viewed.

The number of markers seen by the system during the movement is plotted in Fig. 7. By comparing Figs. 6 and 7 it can be seen that the largest change in  $\gamma$  angle started to happen at the point at which the number of visible markers dropped from 18 to 13. Note that the angular measurements did not change as rapidly as might have been expected by the sudden change in the number of visible markers. This was due to the action of a filter that the system uses to smooth the effect of 'switching attention' between markers, to ensure that any changes in measured parameters occur gradually.



*Fig. 7 - Number of markers seen during the movement.*

## CONCLUSION

Two methods for calculating the position of a camera for virtual studio applications have been described. Both methods are particularly suitable for hand-held cameras in large studios. The first method relies on markers placed in-shot, whereas the second uses markers out-of-shot, viewed by a second camera. The second method appears to be the most suitable for the majority of applications. A small stand-alone unit is being designed to implement this method, although implementation using a standard workstation equipped with a frame grabber is also possible. Measurements carried out on a prototype system implemented using a workstation have shown that the system has the required measurement accuracy.

The system should allow virtual production techniques to be used in large studios without imposing any restrictions on the type of camera mounting or the image composition. It could also be used for other applications that require the accurate measurement of the position and orientation of an object moving within a large volume.

## REFERENCES

1. THOMAS, G.A., 1994. Motion-compensated matting. Proceedings of the 1994 International Broadcasting Convention. pp. 651-655.
2. ADLER, A., 1996. The advantages of the pattern recognition approach in virtual sets. Proceedings of the 1996 International Broadcasting Convention. Late papers section.
3. AZUMA, R, and BISHOP, G., 1994. Improving Static and Dynamic Registration in an Optical See-through HMD. Computer Graphics Proceedings, Annual Conference Series, 1994, ACM SIGGRAPH, pp.197-204.