A STEREOPHOTOGRAMMIC SYSTEM TO POSITION PATIENTS FOR PROTON THERAPY

Neil Muller
Department of Mathematical Sciences, University of Stellenbosch, Stellenbosch, South Africa

Evan de Kock, Ruby van Rooyen and Chris Trauernicht
Medical Radiation Group, iThemba Labs, Somerset West, South Africa

Keywords: Stereo vision, proton therapy, visually guided robotics.

Abstract: Proton therapy is a successful treatment for lesions that are hard to treat using conventional radiotherapy, as the radiation dose to nearby critical structures can be tightly controlled. To realise these advantages, the patient needs to be accurately positioned, and monitored during treatment to ensure that no motion occurs. iThemba LABS uses a fixed beam-line, and moves the patient using a suitable positioning device. In this paper, we discuss several aspects the stereo vision based system used to both determine the position of the patient in the room, and to monitor the patient during treatment.

1 INTRODUCTION

Proton therapy is a useful treatment method as the dose distribution properties of the proton beam allow higher doses to be delivered to the target volume with lower doses to the surrounding tissue. Due to the high cost associated of treatment, it is often reserved for lesions that are difficult to treat with conventional radiotherapy techniques, such as inter-cranial lesions or lesions close to critical structures (see for example (Webb, 1993)).

iThemba LABS[1] has been involved with proton therapy for over ten years. Due to cost restrictions, iThemba LABS uses a fixed beam-line to deliver the proton dose, and a motorised chair with 5 degrees of freedom to position the patient (see (Jones et al., 1995)). A second treatment vault is being built that uses a robot-controlled manipulator to position the patient. The position of the patient during setup and treatment is monitored by a number of cameras and stereo techniques are used to calculate the patient's position. A critical issue is the high accuracy required.

[1] Laboratory for Accelerator Based Science

2 PROTON VAULTS AT ITHEMBA LABS

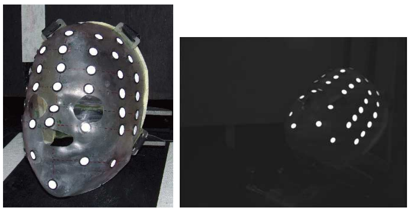
The treatment environment (see Figure 1), shows the treatment layout. We have 9 cameras observing the patient during treatment. During a treatment session, at least 3 cameras will be used to calculate the position of the markers on the mask (see Figure 2(a)). The cameras are positioned around the beam isocenter at an average distance of 2m from the isocenter. The positions were chosen to maximise the volume visible to the cameras while not obstructing activities in the vault.

The position of the target volume is determined from the CT scan used to plan the treatment. So that the position of the target volume can be related to the mask, the patient is scanned wearing the mask. When the patient is in the reference position for the scan, a stereo system, calibrated to CT scanners coordinate system (see (de Kock et al., 2002)), determines the position of the markers on the mask. Thus the relationship between the target volume and the markers on the mask is known.

The stereophotogrammetric (SPG) system is designed to be used in both vaults, and the control systems for the chair and the robot are designed to present a common interface to the SPG system. The

Muller N., de Kock E., van Rooyen R. and Trauernicht C. (2007).
A STEREOPHOTOGRAMMIC SYSTEM TO POSITION PATIENTS FOR PROTON THERAPY.
In *Proceedings of the Second International Conference on Computer Vision Theory and Applications - IU/MTSV*, pages 538-541

Figure 1: The treatment vaults..



(a) Treatment Mask. (b) The SPG System's View of the Treatment Mask.

Figure 2: The treatment mask.

SPG system is responsible both for positioning the patient correctly, and monitoring the patient during treatment to ensure that no excessive movement occurs.

3 THE STEREOPHOTOGRAMMETRIC SYSTEM

A central design aim of the current project is to achieve the required level of accuracy using commodity hardware. Thus we choose to use conventional PAL CCD cameras, and use frame-grabbers based off the venerable BT-878 chipset.

Since we need to monitor the patient position in real time and react to patient motion, we restrict ourselves to a subset of the cameras during a treatment session to minimise the computational load. We also try to ensure that the marker detection problem is as simple as possible, by using retro-reflective markers, illuminated by green light-emitting diodes (LEDs) mounted next to the cameras, and narrow band-pass filters on the cameras, to minimise the effect of the background on the image, and to ensure we have good contrast between the markers and the background. The markers are 10mm in diameter. A typical image of the mask is shown in Figure 2(b).

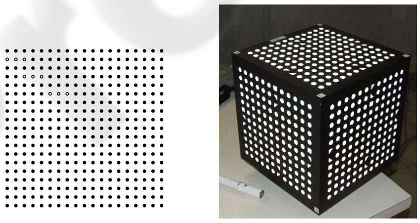
The markers used are circular, and, because of the lighting conditions, the contrast between the markers and the background is high. Thus the markers can be found by searching for bright regions, and the marker centroid can be found using ellipse fitting techniques. Each candidate marker is only accepted if the size is within acceptable limits, and the deviation of the observed shape of the marker from an ideal ellipse is sufficiently small (see (van Rooyen, 2003)).

4 CALIBRATION

Calibration of the SPG system needs to achieve two goals. We need to calibrate the parameters of the cameras, and their relative positions, so we can obtain accurate results from the stereo reconstruction. Since the SPG system must position the patient with respect to the beam, which is fixed in the room, calibration needs to establish the relationship between the room coordinate system and the coordinate system of the cameras.

4.1 Distortion

As we are using conventional zoom lenses, radial distortion is a factor. We use a two-stage approach to estimate the distortion. We determine an accurate base distortion model from observing a suitable pattern (shown in Figure 3(a)) from a known position.



(a) Distortion Correction Pattern. (b) The Calibration Cube.

Figure 3: Calibration patterns.

This is designed to be easy for our software to process. By observing the extent to which the lines of dots curve, the parameters of the distortion model can be obtained. We use a standard model for radial distortion, namely

$$(x_u, y_u) = (x_d, y_d) + (\bar{x}_d, \bar{y}_d)\delta(\kappa) \quad (1)$$

where

$$\bar{r}_d^2 = \bar{x}_d^2 + \bar{y}_d^2 = (x_d - c_{xr})^2 + (y_d - cyr)^2 \quad (2)$$

and

$$\delta(\kappa) = \kappa_1 \bar{r}_d^2 + \kappa_2 \bar{r}_d^4 + \dots \quad (3)$$

Since the relationship between the camera and the pattern is known, and the pattern is known to a high degree of accuracy, the only unknown parameters affecting the image are the distortion coefficients and the focal length. These can be estimated accurately using standard non-linear optimisation techniques.

We restrict the model to a fairly low-dimensional polynomial for numerical stability. While we do not directly model tangential distortion, we allow the centre of distortion to differ from the centre of the image, as this is a good approximation for the decentring distortion (see (Stein, 1997)).

This base distortion model is expected to change over time, as many events will require the cameras to be re-adjusted in the vault. Since such changes to the model should be small, and as removing the camera to recalculate the model in the fixed rig is inconvenient, we wish to update the model. Thus we mount the pattern in Figure 3(a) on a portable planar object. This is held so that it roughly fills the field on vision of the camera.

By extracting the lines from the pattern, we can test if the model is still valid. If the model is no longer valid, we can use our knowledge of the pattern to calculate an updated distortion model, (similar to (Tamaki et al., 2002)). Since the pattern is held so that it is approximately face-on to the camera, we can use that and current distortion model as a reasonable initial estimate, and the optimiser converges to the updated solution quickly.

From this updated distortion model, we can then decide whether the drift from the base model has become too large, in which case we remove the camera and recalculate the base distortion model, or else we simply use the updated model (see (van Rooyen and Muller, 2004) for more details).

4.2 Camera Calibration

Since we need to calibrate the cameras to the surveyed coordinate system in the room, we use a single cube to calibrate all the cameras. This cube is mounted on a specially constructed jig which can be reliably positioned at the reference position. The correspondence between the cube position and the room was obtained by surveying the room with the cube in position, and is known to a high degree of accuracy. The relationship between the cube and the beam-line is regularly checked using a theodolite mounted along the beam-line. We use different cubes for each vault, and the jig design is slightly different between each vault so there is no possibility of using the incorrect cube.

Each face of has a large number of circular markers, whose positions are known from the surveying results. Each face also as a distinctive pattern of squares and triangles which is used to automatically label all the points detected on the cube (see figure 3(b)). From the number of markers visible, we can determine if it has an adequate view of the volume around the beam isocenter.

The current distortion model is used to correct for distortion before calibration, and then calibration proceeds using Tsai's method (see (Tsai, 1987) for details).

5 POSITIONING THE PATIENT

For each target volume, several beams, each with different entry points are planned. To treat a beam, the patient is positioned so that the target volume is positioned at the beam isocenter, and the entry point lies along the beam axis, between the beam source and the target volume. A collimator is inserted into the beam line to shape the beam to the target volume, and, once the patient is in position, the collimator is rotated to match the profile of the target volume. In a single treatment session, the patient will be treated using multiple beams.

Before treatment, the patient is secured to the positioning system and moved to a standard reference position. In this position, most of the markers on the mask can be observed, and stereo vision techniques are used to determine the 3D position of a subset of the markers on the mask.

These markers are then registered to the marker positions in the reference position of the CT Scan. This determines the transformation between the CT scanner reference position and the current patient position. From this, the current position of the target volume and entry point can be determined. The required movement to position the patient is calculated and sent to the positioning system.

This procedure is repeated until the patient is in position. This final position is verified by comparing an X-ray taken along the beam path with the predicted X-ray view (see (van der Bijl, 2006)).

Positioning the patient for the next beam in a session follows the same procedure, except that, since the current position of the patient is known, there is no need to return to the reference position.

6 MONITORING THE PATIENT

Monitoring the patient is a subset of the positioning loop. Since the patient is monitored continually after being positioned, and the frame-rate is high in comparison to the speed of the movements of interest, we can a simple nearest neighbour comparison to track markers.

On each iteration, the current position of the target volume and the entry point are calculated and compared to the required positions. If the differences exceed the specified tolerances, the patient is declared to be out of position and the beam is interrupted.

Monitoring continues for some time after this to determine if this is a transient event. If the patient moves back into position after a short delay and remains in position for a suitable period, the treatment is resumed, otherwise the beam is aborted, and the patient is repositioned.

7 ACCURACY OF THE SYSTEM

To assess the accuracy of the SPG system in reasonably realistic circumstances, we ran a series of experiments using a dummy mask and using the motorised chair. We constructed 8 treatment beams for this mask, with the target point of each beam coincident with one of the markers on the mask. The accuracy was measured using a theodolite mounted along the beam-line.

The mask was positioned for each beam using the SPG system, and then the chair was moved until the marker chosen as the target point was centred in the view of the theodolite. The displacement required to centre the marker was recorded for each beam, and this was repeated for all the beams, using several different camera combinations for each beam. These experiments were repeated daily over 5 successive days.

Because the of position of the theodolite, the displacement along the beam-line, could not be observed, so we only measured the error orthogonal to the beam-line.

The average error was found to be 0.305mm with a standard deviation of 0.219mm. Since the slice thickness typically used for planning patients at iThemba LABS is 1mm, this is well within the required tolerances.

The results for each beam are listed in Table 1.

Table 1: Accuracy Test Results.

Beam	Avg Δd	Std Dev Δd
1	0.436	0.116
2	0.187	0.088
3	0.175	0.105
4	0.416	0.185
5	0.158	0.074
6	0.182	0.086
7	0.247	0.119
8	0.265	0.160
All	0.305	0.219

8 CONCLUSIONS

In this paper, we describe several aspects of the stereophotogrammetric system used for patient positioning at iThemba LABS. By exploiting the controlled natured of the environment, we can obtain good results using simple approaches. In addition to positioning the patient, we can monitor the patient and respond if the patient moves during treatment. The system achieves acceptable accuracy using commodity hardware.

REFERENCES

de Kock, E., O'Kennedy, B., and Muller, N. (2002). Calibrating a stereo rig and CT scanner with a single calibration object. In Greiner, G., Niemann, H., Ertl, T., Girod, B., and Seidel, H.-P., editors, *Vision, Modeling, and Visualization 2002*.

Jones, D. T. L., Schreuder, A. N., Symons, J. E., Rüther, H., van der Vlugt, G., Bennet, K. F., and Yates, A. D. B. (1995). Use of stereo-photgrammetry in proton radiotherapy. In Rüther, H., editor, *Proc. Int. FIG. Symp. on Photogrammetry in Enigineering Surveying*, pages 138–152.

Stein, G. (1997). Lens disortion calibration point correspondences. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 602–608.

Tamaki, T., Yamamura, T., and Ohnishi, N. (2002). Unified approach to image distortion. In Kasturi, R., Laurendeau, D., and Suen, C., editors, *16th International Conference on Pattern Recognition*, pages 584–587. IEEE, IEEE Computer Society.

Tsai, R. Y. (1987). A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal of Robotics and Automation*, RA-3(4):323–344.

van der Bijl, L. (2006). Verification of patient position for proton therapy uisng portal x-rays and digitally reconstructed radiographs. Master's thesis, University of Stellenbosch.

van Rooyen, R. (2003). Fast, robust detection of circular retro-reflective targets. In *Proceeding of the Fourteenth Annual Symposium of the Pattern Recognition Association of South Africa*, pages 21–26.

van Rooyen, R. and Muller, N. (2004). A method for efficiently re-estimating camera distortion parameters. In et al., F. N., editor, *Proceeding of the Fifteenth Annual Symposium of the Pattern Recognition Association of South Africa*. Pattern Recognition Association of South Africa.

Webb, S. (1993). *The physics of three-dimensional radiotherapy: Confromal radiotherapy, radiosurgery and treatment planning*. Insitute of Physics Publishing, Bristol and Philadelphia.