Calibration of Video Cameras to the Coordinate System of a Radiation Therapy Treatment Machine

Scott W. Hadley, L. Scott Johnson, and Charles A. Pelizzari

University of Chicago
The Department of Radiation and Cellular Oncology
5841 South Maryland Ave. MC 1105
Chicago, IL 60637
`s-hadley@uchicago.edu`
`www.radonc.uchicago.edu`

Abstract. There has been recent interest in using video cameras along with computer vision and photogrammetric techniques to aid in the daily positioning of patients for radiation therapy. We describe here a method to calibrate video cameras to the beam coordinate system of a radiation therapy treatment machine. Standard camera calibration relates the three-dimensional coordinate system of a calibration phantom to the two-dimensional image coordinates. Using a calibration phantom designed for simultaneous X-ray and video imaging, both types of images can be calibrated to a single coordinate system. A series of X-ray images of the calibration phantom is taken using the motion of the treatment machine. Camera calibration parameters derived from these images are used to find a transformation from the coordinate system of the calibration phantom to the beam coordinate system of the treatment machine. This transformation is applied to the video camera calibration to provide a camera calibration directly to the beam coordinates of the machine.

1 Introduction

External beam radiation therapy uses a linear accelerator to produce megavoltage X-rays to deliver a prescribed dose to a tumor deep inside a patient. The linear accelerator is mounted on a gantry which rotates about a horizontal axis. The focal spot of the X-ray beam, for a typical machine, is 100cm from the axis of rotation and traces out a circle in a vertical plane as the gantry rotates. A patient lies on the treatment couch which rotates about a vertical axis. These two axes of rotation allow the beam to enter the patient's body from nearly any direction. Prohibited are those beam directions that would cause the gantry to intersect either the couch or the patient. The *isocenter* of the treatment machine is defined as the point in the room where the axes of rotation of the gantry and couch intersect. The standard convention is to define the isocenter as the origin, the Y-axis as the axis of rotation of the gantry and the Z-axis as the axis of rotation of the treatment couch, see figure 1.

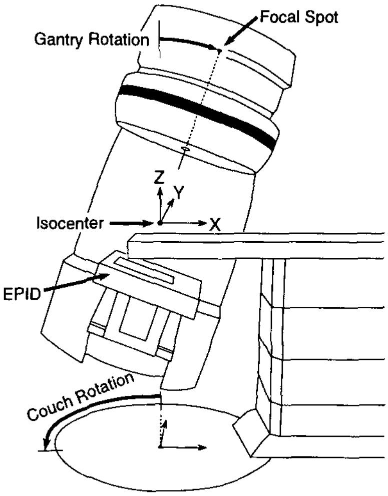

Fig. 1. Diagram of radiation therapy linear accelerator. The gantry houses the linear accelerator and rotates about a horizontal axis. The patient lies on the treatment couch which rotates about a vertical axis. These two axes of rotation intersect at the isocenter. Below the treatment couch is the electronic portal imaging device (EPID) which is attached to, and rotates with, the gantry.

The ideal treatment has the dose from the X-rays delivered to the tumor and little dose to the surrounding normal tissues and organs. Conformal radiation therapy seeks to minimize the dose to normal tissue surrounding the tumor by shaping the aperture of the X-ray beam to match the shape of the tumor in projection. Computer-based treatment planning using computed tomography (CT) images is used to decide how many beams will be used, what aperture they will have and what direction they will enter the body. A physician, on each slice of the CT image, outlines the tumor and any critical structures to be avoided. The 3-D tumor geometry is used to choose a point inside the patient that should be placed at the isocenter of the treatment machine during treatment. Typical conformal treatments use two to six different beam directions chosen to assure coverage of the tumor, exclude critical structures and spread dose to normal tissue over a large volume. The total dose to the tumor is divided into several fractions given once or twice a day. A typical treatment takes several weeks and the patient must be positioned 20 to 40 times.

The most common method of reproducing the position of the patient uses skin marks and isocentric lasers in both the CT room and the treatment room. At the time of image acquisition, the patient is placed on the couch of the CT machine in the pose required for treatment. Lasers mounted on the walls and ceiling trace out orthogonal planes to represent the treatment machine coordinates. A pen is used to trace where the lasers fall on the patient's skin surface. A calibration between the lasers and the CT machine is used to align the laser coordinates to the CT coordinates. A corresponding set of lasers in the treatment room is used to reproduce the position of the patient on the treatment couch.

There are two components to the errors in patient positioning for fractionated radiation therapy. Random errors reflect the inability of a positioning system along with human operators to reproduce the same position day after day. Systematic errors represent a deviation in the planned patient position from the average position. The goal of any system to position the patient is to minimize the random errors and have negligible systematic error. If it were possible to minimize the random error, then measurement of the systematic error could be more accurate and require fewer observations. Several research centers have been working on new techniques both to eliminate systematic errors and reduce random errors.

The position of the patient may be checked, typically once per week, by taking X-ray images of the patient at the time of treatment. An electronic portal imaging device (EPID) is a planar digital imager attached to the gantry to capture an X-ray image as the treatment beam exits the patient. EPID images can be compared to corresponding predicted images computed from the CT and treatment plan data. The three translation and three rotation parameters needed to place the patient in the proper position can be obtained through computer analysis of visible bony anatomy[1]. The major advantage of this method is that the X-ray beams used to treat the tumor also produce the images, which accurately represent the position of the patient relative to the beams. However, it is time consuming and contributes excess dose to normal tissue. Also it is not well suited to use in all sites of the body, for example in breast and lung treatments.

Recently there has been interest in using video cameras and photogrammetric techniques to aid in the daily positioning of patients. The German Cancer Research Center (DKFZ) demonstrated that photogrammetric techniques could track the motion of a patient's head with an accuracy of better than 0.1 mm[2]. The measurements were made relative to the initial position of the photogrammetric phantom and thus can not give an absolute measurement of the position in the treatment machine coordinates. This limits the technique to reducing the random errors in position. Without a calibration to the beam coordinates the systematic error in the patient's position remains unknown.

Recent work at the University of Chicago Department of Radiation Oncology has centered on using a pair of live video subtraction images to reproduce a patient's position in the treatment room[3]. Two video cameras mounted to the wall and ceiling of the treatment room are directed toward the isocenter. On the first day of treatment, images from the two cameras are saved for use as a reference for future positioning. The patient's position is reproduced by comparing live video with the reference image using a subtraction technique. Radiation therapists, while looking at the live video subtraction, interactively move the patient to produce a null image and thus reproduce the previous position. The technique has been shown to be fast and easy to implement in a busy radiation therapy clinic. Its use has facilitated a reduction in the random errors in patient positions. The current system dose not incorporate any type of geometrical calibration and thus cannot be used to measure the systematic positioning error of the patient relative to the beam coordinate system.

A set of video cameras calibrated to the beam coordinate system of the treatment machine could in principle be used to make absolute measurements of the patient position in beam coordinates before each treatment. Geometric information from CT images used for computer based treatment planning could then be directly related to live video images of patients in the treatment room. Such a system could help to reduce both systematic and random errors in patient positioning.

2 Method

Central to our method is a calibration phantom that can be used for both X-ray and video camera images. Our phantom consists of steel ball bearings in a known 3D configuration. The balls are painted white and mounted on black plastic surfaces and thus are easily visible in video images. In the X-ray images, the steel bearings appear as opaque circles, see figure 2. The center of a circle is found by matching a circular template with the edges of the circle in the image. The center of a circle is assumed to be the projection of the center of the spherical ball bearing. The extracted image points and corresponding 3D phantom coordinates are used to estimate the camera parameters using Tsai's method[4].

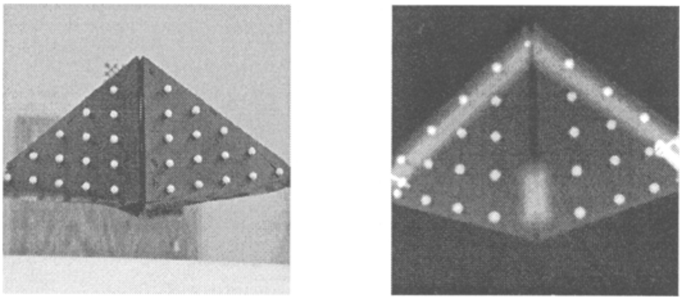

Fig. 2. Digital and X-ray image of the phantom used for camera calibration. Simultaneous images of the phantom tie the two imaging modalities to a common coordinate system.

Two sets of X-ray images are used to estimate the relationship between the calibration phantom and radiation beam coordinate system. The first image set is obtained while rotating the gantry around the stationary phantom. The camera parameters are used to estimate the plane in which the focal spot rotates and the center of rotation, i.e. the isocenter. The second set of images is obtained by rotating the couch, and thus the phantom, while holding the gantry and image plane fixed. This second set of camera calibration parameters is used to estimate the axis of rotation of the couch. One camera/phantom position is common to both sets and is used as a link between the two.

2.1 Camera Model and Calibration

We use Tsai's camera calibration method with both the video and X-ray images. The model includes 6 extrinsic parameters for the translations and rotations needed to express the calibration points in the camera coordinates. We express 3D coordinates with the homogeneous vector $\boldsymbol{v} = [x, y, z, 1]^{\mathrm{T}}$ and use a 4x4 matrix,

$$T_{\mathrm{p,c}} = \begin{bmatrix} R & \boldsymbol{t} \\ 0_3^{\mathrm{T}} & 1 \end{bmatrix}, \tag{1}$$

to represent the transformation from one coordinate system to another. Here, for example, $T_{\mathrm{p,c}}$ is a transformation from the phantom to the camera coordinates, where R is a 3x3 rotation matrix, $\boldsymbol{t} = [t_{\mathrm{x}}, t_{\mathrm{y}}, t_{\mathrm{z}}]^{\mathrm{T}}$ represents the translation between the origins and $0_3 = [0, 0, 0]^{\mathrm{T}}$. The camera model is such that the Z-axis of the camera coordinate system is perpendicular to the imaging plane, the X-axis is parallel to the rows of the pixel array, and the Y-axis is parallel to the columns of the pixel array. The 5 intrinsic camera parameters that are estimated include the focal length, f, the image center, $[C_{\mathrm{x}}, C_{\mathrm{y}}]$, horizontal scale factor, s_{x}, and the radial lens distortion, ρ.

After this first calibration the relationship, between the phantom and camera coordinates and the image plane is well known. At this point, the relationship between the beam coordinates and the phantom remain undetermined.

2.2 Beam Coordinate Estimation

The first step in finding the transformation from the calibration phantom to the beam coordinate system is to locate the isocenter of the treatment machine in the phantom coordinates. A graticule is placed between the image plane and the X-ray source and used to find the X-ray beam direction in the image plane. The image coordinate of the central spot of the graticule, in several calibrated images, is triangulated back to phantom coordinate system and considered the position of the isocenter, $\boldsymbol{t}_{\mathrm{iso}}$.

The set of camera calibration parameters, for the X-ray images, obtained with the calibration phantom held fixed is used to estimate the plane in which the gantry rotates. In the camera coordinate system, the focal spot is located at the origin. Thus let $\boldsymbol{s} = [0, 0, 0, 1]^{\mathrm{T}}$ be the focal spot of a treatment machine from a calibrated X-ray image. First, all focal spots are expressed in the coordinates of the calibration phantom and isocenter, $\boldsymbol{t}_{\mathrm{iso}}$, subtracted off. Using the inverse of the coordinate transformation in Equation 1

$$\boldsymbol{s}_{\mathrm{p}_i} = T^{-1}_{\mathrm{p,c}_i}\, \boldsymbol{s} - \boldsymbol{t}_{\mathrm{iso}}, \tag{2}$$

where the subscript i indicates the i^{th} camera position. The plane that best fits the focal spots is determined using single value decomposition (SVD). The result is 3 orthogonal eigenvectors used to form a 3x3 rotation matrix, R_{SVD}, where the eigenvector with the smallest eigenvalue is chosen to be the Y-axis of the new coordinate system. The X- and Z-axes are considered to be arbitrary. After expressing all the focal spots in the new coordinate system, the isocenter is further refined by subtracting off the average Y coordinate of the focal spots, $\boldsymbol{t}_{\text{y}} = [0, \bar{s}_{\text{gy}}, 0, 1]^{\text{T}}$. We denote the phantom-to-gantry transformation using the subscript g for gantry, and p for the phantom,

$$T_{\text{p,g}} = \begin{bmatrix} I_3 & -\boldsymbol{t}_{\text{y}} \\ 0_3^{\text{T}} & 1 \end{bmatrix} \begin{bmatrix} R_{\text{SVD}} & 0 \\ 0_3^{\text{T}} & 1 \end{bmatrix} \begin{bmatrix} I_3 & -\boldsymbol{t}_{\text{iso}} \\ 0_3^{\text{T}} & 1 \end{bmatrix}, \tag{3}$$

where I_3 is a 3x3 identity matrix. To obtain a transform from the gantry rotation plane to the camera coordinates, $T_{\text{g,c}}$, the inverse of this transformation is used. Next we need only to rotate about the Y-axis so that the axis of rotation of the treatment couch corresponds to the Z-axis.

To model the motion of the treatment couch, the second set of images taken with the X-ray camera held fixed and the calibration phantom rotated with the couch is used. The motion of the treatment couch is modeled by premultiplying the transformation from the home phantom position, p_{h}, to the camera by the inverse of the transform for the rotated phantom position, p_j, to the camera,

$$\begin{bmatrix} R_{\text{p}_\text{h},\text{p}_j} & \boldsymbol{t}_{\text{p}_\text{h},\text{p}_j} \\ 0_3^{\text{T}} & 1 \end{bmatrix} = T_{\text{p}_j,\text{c}}^{-1} T_{\text{p}_\text{h},\text{c}}. \tag{4}$$

In practice the translation $\boldsymbol{t}_{\text{p}_\text{h},\text{p}_j}$ is small, and we ignore it. To find the rotation about the Y-axis, we use the rotation model,

$$R_{\text{Y},\phi_j} \, R_{\text{Z},\theta_j} \, R_{\text{Y},\psi_j} = R_{\text{p}_\text{h},\text{p}_j}, \tag{5}$$

and solve the kinematic equations for the angles ϕ_j, θ_j and ψ_j[5]. If the couch rotation were perfect, θ_j would be the rotation angle of the j^{th} couch position and ψ_j would equal $-\phi_j$. Due to errors in the calibration and imperfections in the mechanical rotation in the couch, the rotation angles ϕ_j and ψ_j form a distribution. We use the average over all $-\phi_j$ and ψ_j to obtain a compromise rotation angle $\bar{\psi}$. The resulting transformation,

$$T_{\text{b,g}} = \begin{bmatrix} R_{\text{Y},\bar{\psi}} & 0_3 \\ 0_3^{\text{T}} & 1 \end{bmatrix}, \tag{6}$$

relates the machine's beam coordinates to the in-plane gantry coordinates. The transformation from machine coordinates to phantom coordinates is then

$$T_{\text{b,p}} = T_{\text{p,g}}^{-1} T_{\text{b,g}}. \tag{7}$$

This transformation is used to relate both X-ray and video camera calibrations to the machine's beam coordinates.

3 Experiments

We applied this technique to a pair of video cameras mounted to the wall and ceiling of the radiation therapy treatment room. The calibration phantom was placed on the treatment couch and x-ray imaged with 24 gantry angles and two couch rotations. Video image were acquired of the phantom. The phantom was then removed from the table and x-ray images using the graticule were taken from 12 gantry angles. The above technique was applied to the video cameras to obtain the transformation from the beam coordinate system to the video image.

A head phantom was used to test the ability of the system to project information from CT images into the treatment room. A CT image of the head phantom was taken with lead fiducial placed along the laser lines. The center of the lead fiducial were found in the CT image and an isocenter chosen. The head phantom was then carefully positioned on the treatment couch using the isocentric lasers. The position of the fiducial, as seen in the video images, and position as projected from the CT data were compared. Volume rendered images, from the perspective of the video cameras, of the head phantom CT data were prepared using the camera calibration data and compared to the video images.

The ability of the system to accurately track objects was tested as well. A target, similar to the calibration phantom but much smaller, was placed on a machinist translation stage, which has an accuracy of better than 0.02 mm, and translated over a 12.7 mm distance. The position of the target was found my minimizing the distance between the extracted points in the image and the projected points. The magnitude of the known translation was compared to the translation of the estimated position

4 Results

Figure 4 shows the CT image, video images and volume renderings of the head phantom. The fiducial landmarks located in the CT image have been projected into the video image as an "X". The standard deviation of the residual of the projected location of the fiducial and the location extracted from the image by hand was 0.4646 and 1.1419 pixels for the side and top views. The residual for the camera calibration of these two cameras was 0.4146 and 0.3541 pixels respectively. The volume rendered images represent the correct position of the head phantom on the treatment table.

Images of the tracking experiment are shown in figure 4. The magnitude of the known translation was compared to the magnitude of the translation estimated by the calibrated system using linear regression. The slope and intercept of the regressed line was 1.0067 ± 0.0041 and -0.0105 ± 0.0238 respectively.

5 Discussion

The head phantom experiment demonstrates the ability to project information from the CT image into the treatment room. Small differences in the projected

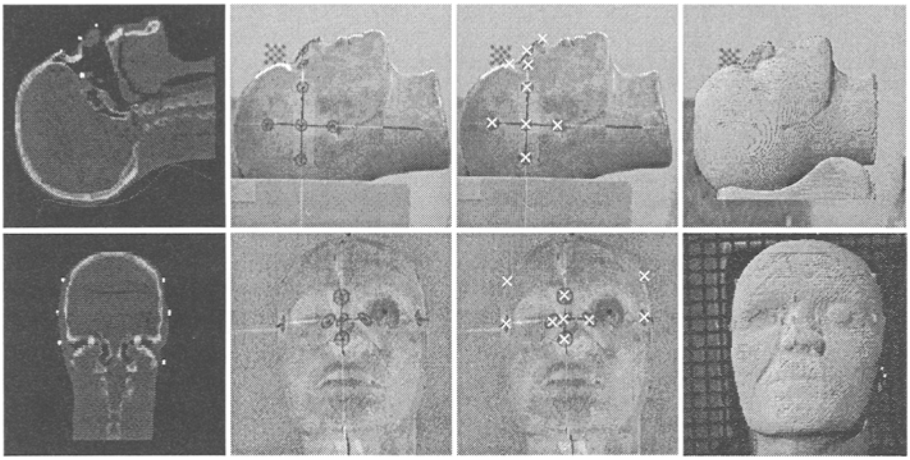

Fig. 3. Demonstration of projecting information from treatment planning CT scan into the treatment room using the calibrated video system. Fiducial landmarks were placed along the laser lines at the time of the CT image, first column. The the location of the fiducial, bright spots in CT image, were projected into the video image, second and third columns. The volume rendered images, fourth column, of the head phantom CT image demonstraits the ability of visualize the correct position, in the treatment room, before the first day of treatment.

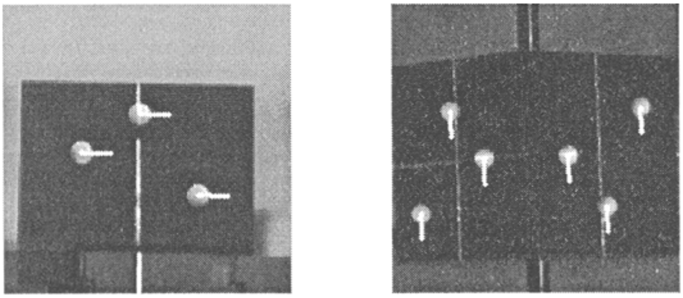

Fig. 4. The target for the tracking experiment consisted of six white spheres mounted to two orthogonal planes. The white lines on the images indicates the position of the spheres in subsequent images. Such a device could be incorporated into CT treatment planning images and, with the calibrated video system, aid in the positioning of patients for radiation therapy.

position of fiducials and their position in the video remain ambiguous. It is possible that the room lasers are not positioned exactly on the isocenter of the machine. Future experiments to determine the absolute accuracy of the video system are planned. The volume rendered images represent the position of the head phantom well. The rendered images do lack photorealism and do not represent how the patient will appear on the treatment couch.

The tracking experiment showed the systems ability to measure the targets position with an accuracy sufficient for radiation therapy. The goal we set for our system is patient positioning within 1 mm and 1 degree of rotation from the planned position. Using such a target incorporated into a bite block may be able to achieve this goal.

6 Conclusion

The potential for video cameras and photogrammetric techniques to aid with patient positioning has been previously demonstrated. This work has been limited to position measurements made relative to the coordinate system of a photogrammetric phantom or reproducing a position from a previous days treatment. These techniques can reduce the random positioning errors but cannot measure the systematic errors.

The technique presented here calibrates video cameras to the coordinate system of the radiation therapy linear accelerator. This would allow photogrammetric techniques to make absolute measurements and thus reduce both systematic and random errors in positioning. Information from CT data can also be accurately projected into live video images to aid human operators in daily patient positioning. Techniques such as these could reduce the systematic error in patient positioning and improve the accuracy of delivering radiation to the patient.

References

1. Kenneth G.A. Gilhuijs, *Automatic verification of radiation treatment geometry*, Ph.D. thesis, University of Amsterdam, 1995.
2. Markus Menke, Frank Hirschfeld, Thomas Mack, Otto Pastyr, Volker Sturm, Wolfgang Schegel , "Photgrammetirc accuracy measurements of a head holder system used for fractionated radiotherapy," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 29, no. 5, pp. 1147–1155, 1994.
3. B.D. Milliken, S.J. Rubin, R.J. Hamilton, L.S. Johnson, G.T.Y. Chen, "Performance of a video-image-subtraction-based patient positioning system," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 38, no. 4, pp. 855–866, 1997.
4. Roger Y. Tsai, "A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf TV camera and lenses," *IEEE Journal of Robotics and Automation*, vol. RA-3, no. 4, pp. 323–344, 1987.
5. Paul P. Richard, *Robot Manipulators*, MIT Press, 1984.