Real-Time 4D Tumor Tracking and Modeling from Internal and External Fiducials in Fluoroscopy

Johanna Brewer[1], Margrit Betke[1], David P. Gierga[2], and George T.Y. Chen[2]

[1] Computer Science Department, Boston University, Boston, MA 02215, USA
[2] Radiation Oncology, Massachusetts General Hospital, Boston, MA 02114, USA
`johannab@cs.bu.edu, www.cs.bu.edu/groups/ivc`

Abstract. Fluoroscopy is currently used in treatment planning for patients undergoing radiation therapy. Radiation oncologists would like to maximize the amount of dose the tumor receives and minimize the amount delivered to the surrounding tissues. During treatment, patients breathe freely and so the tumor location will not be fixed. This makes calculating the amount of dose delivered to the tumor, and verifying that the tumor actually receives that dose, difficult. We describe a correlation-based method of tracking the two-dimensional (2D) motion of internal markers (surgical clips) placed around the tumor. We established ground truth and evaluated the accuracy of the tracker for 10 data sets of 5 patients. The root mean squared error in estimating 2D marker position was 0.47 mm on average. We also developed a method to model the average and maximum three-dimensional (3D) motion of the clips given two orthogonal fluoroscopy videos of the same patients that were taken sequentially. On average, the error was 3.0 mm for four pairs of trajectories. If imaging is possible during treatment, such motion models may be used for beam guided radiation; otherwise, they may be correlated to a set of external markers for use in respiratory gating.

1 Introduction

Knowledge of tumor location is integral to radiation therapy. Fluoroscopy and computed tomography (CT) scans are typically used in the pre-treatment planning phase. Fluoroscopy is an imaging technique in which X-rays continually strike a fluorescent plate that is coupled to a video monitor. In a fluoroscopic image, tumors lack sufficient contrast with the surrounding tissue, so, in preparation for treatment, metal clips are often implanted around the tumor. Since these clips are radio opaque they are visible in fluoroscopy images and CT scans. In fluoroscopy, they provide a way to observe the tumor as it changes position due to various rigid and non-rigid body movements. For abdominal tumors, radiotherapy is particularly complicated by motion due to patient breathing. In order to compensate for this, the tissue volume that will be radiated, the planning target volume (PTV), is often expanded so that the tumor itself, the clinical target

volume (CTV), will receive sufficient dose. This leads to undesirable radiation of healthy tissue surrounding the tumor.

Breath-hold techniques [9] have been introduced to reduce tumor movement. These techniques are not always feasible for the patient [6]. Gating methods seek to compensate for the motion by activating the radiation beam only when the tumor is at a predetermined position. The indirect gating method relies on the correlation between external markers or lung air flow and internal motion [5, 15]. Studies attested to the correlation of external marker motion with the 2D internal motion of the diaphragm [15,7] or internal markers [2]. However, for a given external marker, the ratio of internal to external marker motion can be relatively large [2].

Tumor position can also be directly detected by means of online fluoroscopic imaging of internal fiducial markers [12,13,14]. Due to large patient volume at treatment clinics, there is pressure to decrease the time of a treatment session, but this comes at the cost of increasing the size of the gating window (the time when the radiation beam is on) and thus more healthy tissue is irradiated. This is even more problematic in the typically longer sessions of intensity modulated radiation therapy (IMRT). Beam-guided radiation therapy seeks to address the problems of gating by moving the radiation beam in synchronization with the tumor. This was first introduced in robotic radiosurgery [10], and later adopted for motion-adaptive radiotherapy [13,4,8]. To accomplish the synchronization of motion, the accuracy and reliability of the tracking method is of paramount importance. While there has been significant work in the computer vision community on tracking [1], in particular in medical image analysis in the areas of fMRI, cardiac motion, and blood flow analysis (e.g. [3,11]), there have been only rudimentary efforts in applying tracking technology to measure abdominal tumor motion. The contribution of this paper is to address the problem of tumor tracking in a rigorous manner and provide methods to (1) accurately track internal and external fiducial markers in fluoroscopy, (2) compute trajectories of the average and worst-case motion based on two orthogonal fluoroscopy videos, and (3) establish the correlation of clip motion with external markers. Results are presented that use 10 sets of data from 5 patients.

2   Materials and Methods

Fluoroscopy imaging provides a two-dimensional projection of the density values of the imaged body. The surgically implanted clips can be detected and tracked in the fluoroscopy images since metal is radio opaque and has a higher density than the surrounding tissue. The surrounding tissue, however, may also appear dark or contain edges due to high-density bone structures such as the spine, and the images can become dark during inhalation (Fig. 2).

During treatment planning, radiation oncologists typically request fluoroscopy from two views, the Anterior-Posterior and the Lateral views. Because these views are orthogonal, we can combine the 2D tracking data to recover 3D information about the motion of the clips 1. The setup of the imaging system is
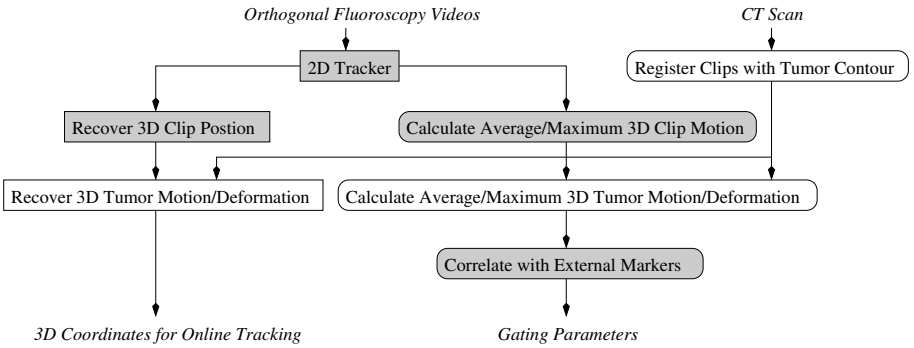
Fig. 1. Overview of method. Sharp-cornered boxes indicate online steps, and boxes with rounded corners offline steps. Methods in shaded boxes are implemented and results are presented here.
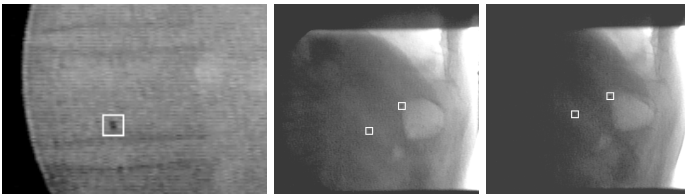

Fig. 2. Regions of fluoroscopic images containing surgical clips marked by white rectangles. Left: Spinal region with a clip and vertebrae edges with similar intensity values. Middle and Right: Two clips during inhalation and exhalation. The leftmost clip moved more significantly in the Cranio-Caudal direction than the rightmost clip. The image on the right is much darker.

such that the Anterior-Posterior and Lateral views share the y-axis of the respective images, which is the Cranio-Caudal axis of the patient, and the isocenter of the patient is at the center of the image.

To date, most hospitals do not have the capability to acquire two orthogonal views simultaneously. However, it is still possible to recover information about the tumor motion from two sequentially obtained orthogonal views. In simultaneous views, we can expect the Cranio-Caudal motion of a given internal marker to be the same in both images. In sequential views, this motion, although not exactly the same, is typically very similar because the fluoroscopy images are taken only minutes apart, during which the patient's breathing and anatomy do not change drastically. Because breathing is well described by a sinusoid, we use sine waves to approximate the average 3D motion and maximum range of the 3D motion in the two sequential views. These sine waves can be correlated with the motion of external markers for use in gating.

During treatment planning, tumors are also contoured manually on CT. This expert knowledge may be used to determine a relationship between the clips,

which are visible in the scan, and the contour. With a sufficient number of clips, the motion and deformation of the tumor in 3D may be extrapolated from the 3D motion of the clips.

2.1   2D Clip Tracking

Initialization. The tracking method is initiated by manual selection of a rectangular region r containing each clip in an initial fluoroscopic image I. In order to find a minimal rectangle containing each clip, the largest "dark" connected component in each region is found by first binarizing the image according to an automatically computed "p-tile threshold" for the region. The rectangle is used as a grayscale template T of the clip. The dimensions of T are denoted by w and h. The locations of the templates in I provide the starting coordinates for tracking the clips in subsequent fluoroscopy frames.

Tracking Algorithm. The normalized correlation coefficient is used to find the position of the clip in subsequent image frames. The tracking algorithm searches for the best match of the clip template T with a region of the image I. This is done by shifting the template T through the image to various points (x', y') and correlating it with each $(w \times h)$ sub-image of I. The value of the normalized correlation coefficient at position (x', y') is

$$R_{I,T}(x', y') = \sum_{x,y} (I(x' + x, y' + y) - \bar{I}) \, (T(x, y) - \bar{T})/(\sigma_I \sigma_T), \qquad (1)$$

where $\bar{T}$ and $\bar{I}$ are the respective mean intensities within the template and image window, and σ_I and σ_T the respective standard deviations. The location (x', y') which maximizes $R_{I,T}$ is taken to be the new clip location.

Searching over all positions (x', y') in I is computationally expensive. Because we know that the clips do not move much from frame to frame, it is possible to restrict the size of the sub-region of I which will be searched. The apparent velocity of the clip's movement in the image is calculated to predict the clip location in the next frame. Velocity $(u, v) = (\frac{dx}{dt}, \frac{dy}{dt})$ is approximated in terms of the rate of change in x and y from the past frame to the current frame: $(u_{t+1}, v_{t+1}) = (x_t - x_{t-1}, y_t - y_{t-1})$. We assume that the velocity is constant, and use (u, v) as an offset from the current position to determine the center of a 5×5 region of interest to search over in the subsequent frame.

2.2   Calculation of Average and Maximum Range of 3D Clip Motion

Our 2D tracking method produces a set of time-indexed 2D coordinates for each view. Given these 2D trajectories, the average and maximum range of 2D motion is calculated for each view in the Cranio-Caudal direction as well as the Anterior-Posterior or Left-Right direction. To determine the maximum range of 2D motion, trajectories are first smoothed using a 1D Gaussian kernel with a support of 10 frames (or 1/3 second) and a standard deviation of one frame. This smoothing operation helps to identify the global maximum and minimum of each trajectory uniquely.
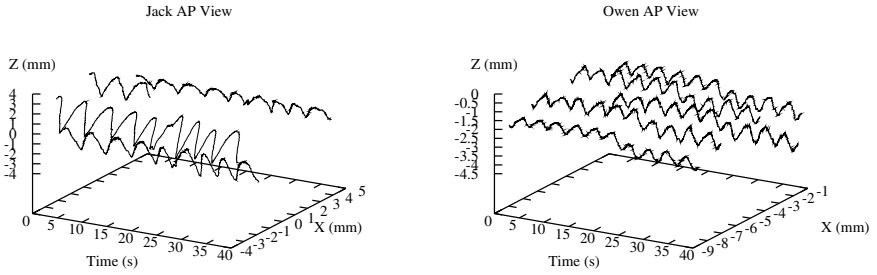
Fig. 3. 2D motion of clips of 2 patients in Anterior-Posterior Fluoroscopy.
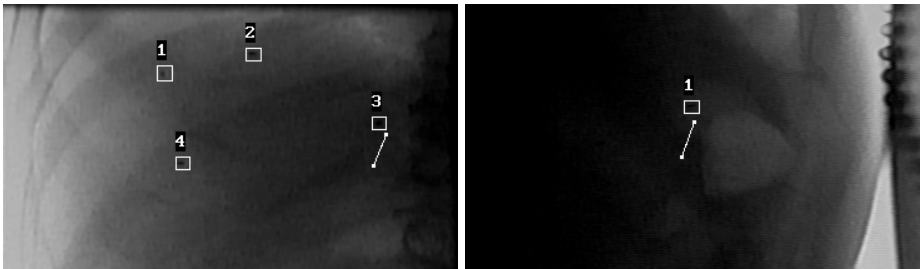

Fig. 4. Two subsections of fluoroscopic views (Anterior-Posterior on the left, Lateral on the right) taken during the peak of a breathing cycle. Clip 3 in the Lateral view and clip 1 in the AP view are the same clip. The white line is a projection of the 3D trajectory that models the average motion of this clip.

To compute an average 2D motion trajectory, first, the maxima and minima that correspond to full inspiration and expiration are located as follows. The difference in the Cranio-Caudal direction between the global maximum and minimum is scaled by 1.1. This scaled value is then used as a threshold during the search for the global maximum and minimum of each breathing cycle. In this search, if the most recently found optimum is a maximum, the distances between this maximum and subsequent points on the trajectory are aggregated until the threshold is reached. The global minimum of the breathing cycle is then determined among the set of points traversed. Similarly, the next maximum can be detected, and the process repeats until all breathing cycles are identified.

The set of minima and maxima is then used to compute amplitude a and frequency f of the 2D motion trajectory. The offset o is computed by averaging the values on the trajectory (1st coefficient of the Fourier Transform). We can then define the average 2D trajectory by sine wave $o + a\sin(2\pi ft)$, where t is time. Since two Cranio-Caudal motion trajectories are given, we average the parameters a, o, and f of these trajectories to obtain a single sine wave that describes average Cranio-Caudal motion. The average Anterior-Posterior

Table 1. Clip Localization Error

Patient	Direction of Motion	Root Mean Squared Error (RMS)		Max Error
		Mean (mm)	Std. Dev. (mm)	(mm)
Alice	Right-Left	0.61	0.58	2.31
Alice	Cranio-Caudal	1.21	0.58	2.00
Doug	Right-Left	0.40	0.15	0.74
Doug	Cranio-Caudal	0.39	0.38	1.50
Eve	Right-Left	0.25	0.25	0.74
Eve	Cranio-Caudal	0.23	0.26	1.00
Frank	Right-Left	0.25	0.23	0.74
Frank	Cranio-Caudal	0.41	0.33	1.00
Gary	Right-Left	0.48	0.32	1.48
Gary	Cranio-Caudal	0.47	0.24	1.00
Average		0.47	0.33	1.25

and Right-Left motions are described by sine waves with offset and amplitude parameters computed from the respective 2D trajectories. For convenience, the frequency parameter computed for the average Cranio-Caudal motion is also used to describe the Anterior-Posterior and Right-Left motions. The three sine waves together form a parametric description of the average motion of the clip in 3D.

3   Results

This study involved 7 patients (the names used here are fictitious). Fluoroscopy videos were collected with a Varian Ximatron radiotherapy simulator with a resolution of 640×480 pixels at 30 frames per second. The computer used to process the data was an Intel Xeon 1.7 GHz with 1 GB RAM and an ATI All-In-Wonder 9800 graphics card.

Ground truth was established for the motion of five clips, one per patient, using the Anterior-Posterior views. The centroid of each clip was manually recorded for every tenth frame of the video. The ground truth positions were then compared with the positions estimated by the proposed 2D-tracking method (Table 1). On average, the mean and standard deviation of the root mean squared (RMS) error in estimating the 10 motion trajectories are under 0.5 mm (< 1 pixel). We compared four pairs of trajectories to evaluate the accuracy of the average 3D motion trajectories in modeling clip motion (Table 2). The average RMS error is 3.0 mm. Fig. 4 shows a fluoroscopic image containing both internal and external markers. We found a high correlation between external and internal markers of two patients (0.88 on average).

Table 2. 3D Model Error

| Patient	Direction of Motion	Root Mean Squared Error (RMS)		Max Error (cm)
		Mean (cm)	Std. Dev. (cm)	
Jack Lateral View	y-direction	0.01	0.02	0.24
Jack Lateral View	z-direction	0.07	0.11	0.58
Jack AP View	x-direction	0.02	0.03	0.14
Jack AP View	z-direction	0.25	0.31	1.87
Nancy Lateral View	y-direction	0.33	0.27	1.23
Nancy Lateral View	z-direction	1.49	1.22	5.43
Nancy AP View	x-direction	0.01	0.01	0.06
Nancy AP View	z-direction	0.27	0.38	2.37
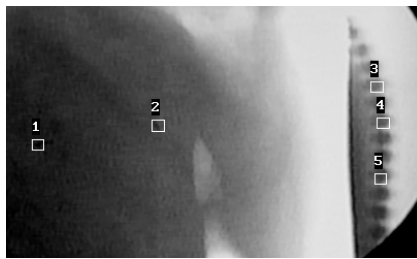
Fig. 5. Lateral fluoroscopic image with three external markers (beads), resting on the patient's abdomen (right), and two internal markers (clips).

4   Discussion and Conclusion

Estimating the location of a tumor accurately is crucial in providing effective and safe radiation therapy. The ground truth study of our tumor tracking system shows that the average RMS error is 0.47 mm, which is considerably smaller than the 1.5 mm error reported by Shirato et al. [14], given that the internal markers are typically only 5 mm long. Our method to model 3D tumor motion based on sequential fluoroscopy videos can be used in settings where the latest technology, i.e., simultaneous orthogonal fluoroscopy, is not available. Computing average tumor motion trajectories from fluoroscopy and CT data obtained in pre-treatment sessions may be helpful in estimating tumor position accurately during treatment.

References

1. D. A. Forsyth and J. Ponce. *Computer Vision, A Modern Approach*. Prentice Hall, NJ, 2003.
2. D. P. Gierga, G. Sharp, J. Brewer, M. Betke, C. Willett, and G. T. Y. Chen. Correlation between external and internal markers for abdominal tumors: Implications for respiratory gating. In *Proceedings of the 45th Annual Meeting of the American Society for Therapeutic Radiology and Oncology*, Salt Lake City, UT, October 2003. A journal version was submitted in May 2004.
3. I. Haber, D. N. Metaxas, and L. Axel. Using tagged MRI to reconstruct a 3D heartbeat. *Computing in Science and Engineering*, 2(5):18–30, 2000.
4. P. J. Keall, V. R. Kini, S. S. Vedam, and R. Mohan. Motion adaptive x-ray therapy: a feasibility study. *Phys Med Biol*, 46(1):1–10, 2001.
5. V. R. Kini, S. S. Vedam, P. J. Keall, S. Patil, C. Chen, and R. Mohan. Patient training in respiratory-gated radiotherapy. *Med Dosim*, 28(1):7–11, 2003.
6. H. D. Kubo, P. M. Len, S. Minohara, and H. Mostafavi. Breathing-synchronized radiotherapy program at the University of California Davis Cancer Center. *Med Phys*, 27(2):346–353, February 2000.
7. G. S. Mageras, E. Yorke, K. Rosenzweig, L. Braban, E. Keatley, E. Ford, S. Leibel, and C. Ling. Fluoroscopic evaluation of diaphragmatic motion reduction with a respiratory gated radiotherapy system. *J Appl Clin Med Phys*, 2(4):191–200, 2001.
8. T. Neicu, H. Shirato, Y. Seppenwoolde, and S. B. Jiang. Synchronized Moving Aperture Radiation Therapy (SMART): Average tumor trajectory for lung patients. *Phys Med Biol*, 48:587–598, March 2003.
9. K. E. Rosenzweig, J. Hanley, D. Mah, G. Mageras, M. Hunt, S. Toner, C. Burman, C. C. Ling, B. Mychalczak, Z. Fuks, and S. A. Leibel. The deep inspiration breath-hold technique in the treatment of inoperable non-small-cell lung cancer. *Int J Radiat Oncol Biol Phys*, 48(1):81–87, 2000.
10. A. Schweikard, G. Glosser, M. Bodduluri, M. J. Murphy, and J. R. Adler. Robotic motion compensation for respiratory movement during radiosurgery. *Comput Aided Surg*, 5(4):263–277, 2000.
11. A.M. Seifalian, D. J. Hawkes, C. R. Hardingham, A.C. Colchester, and J.F. Reidy. Validation of a quantitative radiographic technique to estimate pulsatile blood flow waveforms using digital subtraction angiographic data. *J Biomed Eng*, 3(3):225–233, May 1991.
12. Y. Seppenwoolde, H. Shirato, K. Kitamura, S. Shimizu, M. Van Herk, J. V. Lebesque, and K. Miyasaka. Precise and real-time measurement of 3d tumor motion in lung due to breathing and heartbeat, measured during radiotherapy. *Int J Radiat Oncol Biol Phys*, 53(4):822–834, 2002.
13. H. Shirato, S. Shimizu, K. Kitamura, T. Nishioka, K. Kagei, S. Hashimoto, H. Aoyama, T. Kunieda, N. Shinohara, H. Dosaka-Akita, and K. Miyasaka. Four-dimensional treatment planning and fluoroscopic real-time tumor tracking radiotherapy for moving tumor. *Int J Radiat Oncol Biol Phys*, 48(2):435–442, 2000.
14. H. Shirato, S. Shimizu, T. Kunieda, K. Kitamura, M. van Herk, K. Kagei, T. Nishioka, S. Hashimoto, K. Fujita, H. Aoyama, K. Tsuchiya, K. Kudo, and K. Miyasaka. Physical aspects of a real-time tumor-tracking system for gated radiotherapy. *Int J Radiat Oncol Biol Phys*, 48(4):1187–1195, November 2000.
15. S. S. Vedam, V. R. Kini, P. J. Keall, V. Ramakrishnan, H. Mostafavi, and R. Mohan. Quantifying the predictability of diaphragm motion during respiration with a noninvasive external marker. *Med Phys*, 30(4):505–513, April 2003.