Real-time Tumor Tracking for Pencil Beam Scanning Proton Therapy

Sai Vemprala[1], Srikanth Saripalli[1], Carlos Vargas[2], Martin Bues[2], Yanle Hu[2], Jiajian Shen[2]

***Abstract*— In this paper, we describe the method and implementation of a real-time tumor tracking system for a pencil beam scanning (PBS) proton therapy system. PBS is an advanced cancer treatment system that can benefit from precise localization of the tumors through motion. We utilize techniques such as cross-correlation matching, correlation filters and small object saliency, creating an array of methods that can detect and track fiducial markers implanted in the cancer tumors. The final aim is to control the proton beam using real-time image guidance. Our technique works robustly on various types of markers such as ceramic/metallic fiducials, visicoil markers and surgical clips. Left and right views of an X-ray fluoroscopy system were utilized to also triangulate the marker positions in full 3D as they are tracked through normal breathing movement and organ motion. We have tested our detection system on data from several patients with different tumor locations both offline and in real-time and wish to implement it within a full treatment system soon. To the best of the authors knowledge, this is the first real time tracking system for PBS therapy that is applicable for various types of fiducials and tumor locations.**

I. Introduction

Proton beam therapy (PBT) is an advanced type of radiation therapy for cancer. In PBT, particle accelerators are used to energize a beam of protons, which is then used to target cancerous tumors and damage their DNA. PBT has several advantages of its own, mainly that the radiation stays focused on the tumor shape and only delivers low-dose side effects to the surrounding healthy tissue. A variant of proton beam therapy, known as pencil beam scanning proton therapy (PBS) is the most advanced form of PBT. In PBS, the protons from the accelerator are focused into narrow pencil-thin beams of 5 to 7 mm in width. These pencil beams are then steered laterally across the tumor and uniform tumor coverage in the depth direction is achieved by appropriately varying the kinetic energy (thereby the range) of the protons. PBS results in an even lower radiation dose to healthy tissue compared to older forms of PBT, while allowing sculpting of doses specific to the often complex shape of tumors.

One particular challenge in the delivery of PBS is organ motion. Natural organ motion can adversely impact any dynamic treatment technique, and particularly PBS. The enhanced ability of proton scanning to paint doses conformally increases the risk of target misses due to organ motion. In addition, structures moving in and out of the beam due to organ motion, may lead to unwanted radiation exposure to healthy tissues. Furthermore, interplay effects between organ motion and scanning motion of the pencil beam may give rise to undesirable hot and cold spots of radiation dose in the tumor. If a treatment system is used without careful attention to these characteristics, the treatment time may become considerably prolonged. These risks can be mitigated by image-guidance techniques that can help control the beam accordingly. At the same time, motion effects make it desirable to gate the proton beam in such a way that the proton beam is only delivered when the patient is in a particular phase of the breathing cycle: hence accurate beam gating requires accurate real-time image guidance and identification of the motion phase of the patient.

Proton beam therapy systems are usually installed in conjunction with X-ray systems for patient positioning: and these systems, in many cases, are operated in fluoroscopy mode. In fluoroscopy mode, X-ray images of the patient are obtained with low dosage and high acquisition rates, allowing for analysis of organ motion and accurate identification of the breathing phase. The tumors themselves often lack sufficient contrast under fluoroscopy to visualize their motion directly. Therefore, it is customary to implant radio-opaque fiducial markers in or near the tumor in order to allow for the tracking of the tumor under fluoroscopy.

In this paper, we describe a system that is capable of tracking cancerous tumors using these implanted fiducial markers. We utilize X-ray fluoroscopy images of the markers as the inputs and develop an approach to detect and track these markers through motion in real time. Our algorithm uses computer vision detection techniques and is capable of working on different types of markers and different locations of the tumor. Furthermore, we apply tracking on left and right fluoroscopy views, allowing for triangulation of the markers in 3D and identification of motion phases. Our system does not require a-priori knowledge of the shape of the markers, thus making it applicable for a variety of marker types. We have tested this algorithm on data from patients with lung, liver, prostate and mediastinal tumors and various types of fiducial markers. As of now, our system is capable of detecting and tracking tumor trackers in the fluoro images; actual beam control and gating are our future extensions. To the best of our knowledge, this is the first real time tumor tracking system that works on different types of markers and tumor locations.

II. Related Work and Significance

Real-time tumor tracking for radiotherapy purposes has been a topic of interest in the recent years. Shirato et al [1] present a real time tumor tracking system for gated

[1]Sai Vemprala and Srikanth Saripalli are with the Department of Mechanical Engineering, Texas A&M University, College Station, Texas, USA.

[2]Carlos Vargas, Martin Bues, Yanle Hu and Jiajian Shen are with the Department of Radiation Oncology, Mayo Clinic Arizona, Phoenix, Arizona, USA.

radiotherapy that works with 2mm gold markers and X-ray images of the implanted fiducials. While this is one of the earliest implementations of real time tracking, this work assumed a fixed template size and works for only one type of marker. Brewer et al [2] present another correlation-inspired method for 4D tumor tracking with both internal and external fiducials, but also limited to a single type of marker: surgical clips. Xu et al [3] use a method that combines optical flow and template matching to track lung tumors in fluoroscopic video. This method is robust for objects that exhibit only small displacements, but exhibits higher tracking error with more movement.

Other previous systems that were capable of real time tracking required implanted beacon transponders to localize the target [4]. The use of expensive specialized markers for guidance made this application unpractical. A different approach was devised by the mean of fluoroscopic guidance in a CyberKnife system [5] however, this system only allows for images to be acquired about every minute with the subsequent potential mismatch in-between. Much more recently, Hokkaido University has created a unique real time tracking system for true continuous fluoroscopic tracking [6]. However, their system can only be used with certain types and marker shapes. It is also only validated and used in certain clinical scenarios. It lacks the robustness and flexibility for implementation across sites and treatment systems.

In contrast to many works in the literature in this domain, our approach is suitable for a variety of tumor locations and marker types, requiring only 'region of interest' locations of the markers prior to operation. The system is also capable of recognizing variations in marker shape due to motion or deformation. This allows for the use of asymmetrical marker shapes such as surgical clips which are commonly left behind during surgical procedures. Our system can also recognize very sudden and large movements, which is a limitation of other systems that rely in marker visualization in a relatively small space. Our system is able to continue tracking after breath hold and during subsequent breath hold phases without realignment. Thus, this system is capable of providing a robust and flexible real time tracking system for different markers, clinical scenarios, and movements.

III. System Description

In order to implement and test our algorithm, we have obtained multiple datasets from the proton beam facility at Mayo Clinic Arizona, which is also our target application system. This facility operates a Hitachi PROBEAT-V proton beam therapy system (figure 1). All treatment rooms are equipped with 6 degrees-of-freedom robotic couches where the patients are positioned. Patient positioning is accomplished using two orthogonal X-ray systems angled at 45° and 315°, respectively. Each individual X-ray system includes a X-ray tube and a flat panel detector. The flat panel detector has an imaging area of 42.7×42.7 cm^2 and a native pixel matrix of 3072×3072; offering a spatial resolution of 0.139×0.139 mm^2 at the detector surface and 0.09×0.09 mm^2 at the iso-center. For fluoroscopic imaging, the X-ray system can be operated in either continuous or pulsed fluoroscopy modes. Both fluoro modes have a pixel matrix of 1024×1024 through 3×3 binning from the native pixel matrix. The frame rate is 25 fps for the continuous mode and 7.5 fps for the pulsed fluoro mode.

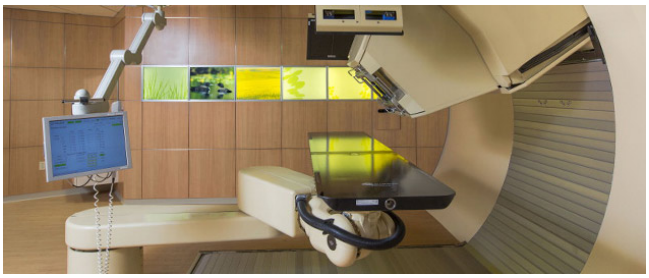

Fig. 1: A picture of a treatment room equipped with the Hitachi PROBEAT-V system.

For real-time image capture, we have installed a screen capture utility that records the X-ray fluoroscopy images from both the left and right views. These images are recorded at 60 Hz and a resolution of 480×480 for each view. An Intel NUC computer acts as the companion computer that processes these images and runs the detection and tracking algorithms: consisting of an i7-6770HQ processor clocked at 2.6 GHz and 32 GB of RAM. Our algorithms are capable of working on both real-time streaming of X-ray images or pre-recorded videos. As mentioned in the previous section, as of now the system only outputs marker positions; actual gating and control of the beam are part of our future work.

IV. Marker Detection and Tracking

In this section, we discuss the techniques we use for detecting and tracking the fiducial markers in the X-ray images. We have implemented multiple algorithms for our purpose and thus, the system contains an array of methods, from which the most applicable algorithm can be chosen based on factors such as tumor location, type of marker, patient characteristics etc. The biggest challenges encountered during detection and tracking in X-ray images are high amount of noise in the images, the small size of certain markers and a lack of sufficient contrast between the markers and the background. We attempt to enhance the images through a pre-processing step first and then apply our detection algorithm. For all of our techniques, we require an estimate of a region of interest where the marker(s) are present, prior to operation. We achieve this by letting an operator click on the marker locations prior to initializing the tracking.

A. Image preprocessing

The images obtained directly from X-ray fluoroscopy usually exhibit a significant amount of noise. In order to make the images more suitable for tracking and detection, we apply a denoising technique and a histogram equalization technique to the X-ray images. We use a fast non-local means

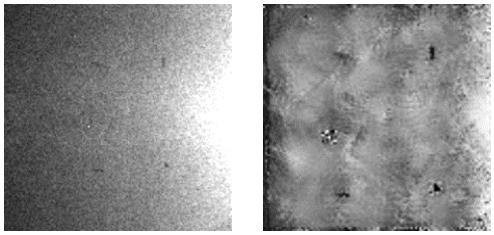

Fig. 2: X-ray fluoroscopy image with markers before (left) and after (right) pre-processing. Applying denoising and CLAHE results in enhanced clarity and easily distinguishable markers.

algorithm for denoising the image, which results in a lesser loss of detail compared to local means algorithms [7].

After performing denoising on the image, we attempt to improve the contrast of the markers against the background. Although the markers are made of radio-opaque material that is visible in the X-rays, often the markers are present behind layers of fat or between organs, thus would not be at a high contrast with the background. In our implementation, we use an algorithm known as contrast limited adaptive histogram equalization (CLAHE) [8]. CLAHE has a wide array of applications in the medical imaging area, and is an algorithm that improves upon conventional histogram equalization by modifying histograms for pixel neighborhoods, thus avoiding over-amplification of noise. In figure 2, we show the effects of this pre-processing step on a sample fluoroscopy image.

B. Template Matching

Given an initial estimate of the appearance of an object, a simple technique to detect it in any image is known as template matching. Template matching uses a brute-force search over an image to look for a predefined template, which in our case is a fiducial marker, in order to maximize an objective function. In our implementation we use the initial positions of the markers and for each position, we isolate a small window of pixels around these points as the templates. As subsequent images are streamed in from the fluoroscopy setup, the template matching algorithm attempts to detect these markers in every new frame. In our implementation, we use normalized cross correlation (NCC) [9] as our objective function to perform template matching. Essentially, this objective function attempts to minimize the difference between the template and the patch of image being considered. Applying this objective function results in a 'mask' image where the maximum pixel intensity indicates the estimated location of the template, and this process is repeated for all the markers.

C. Kernelized correlation filters

We here note that although the template matching is a simple technique to detect the location of a marker, it is prone to failure when marker appearance changes through scaling or rotation. In order to make template-matching based detection and tracking more robust, we couple it with a tracking algorithm, namely, kernelized correlation filters.

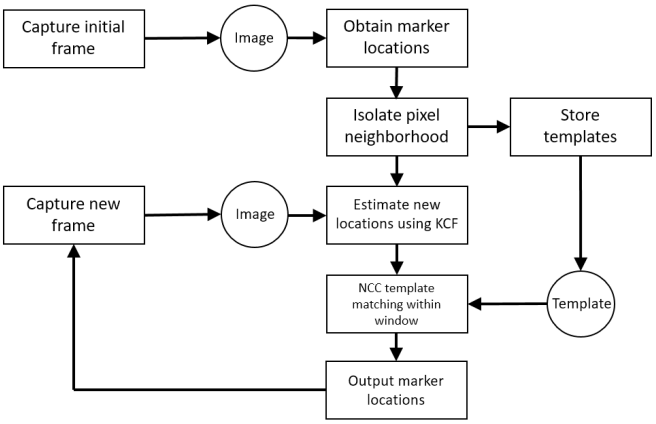

Fig. 3: Process flow of the template matching algorithm along with KCF tracking.

Kernelized correlation filters (KCF) were first proposed in [10], as a computationally efficient tracking method. Given a window of pixels to track, the KCF algorithm attempts to learn the characteristics of this area by computing a discriminative function α. First, the algorithm separates the image into two parts: regions within the targeted area are treated as positive samples and the regions outside as negative samples. The efficiency of the algorithm comes from the fact that all these samples are represented as circulant matrices: and standard algebraic operations of circulant matrices can be converted to cheap element-wise operations in the Fourier domain. Hence, the discriminative function is obtained through a discrete Fourier transform (between the vector and its complex conjugate). In the detection phase, a regression function is calculated for each image patch (candidate) z in the frame as

$$\hat{f}(z) = (\hat{k}^{xz})^* \odot \hat{\alpha} \tag{1}$$

$\hat{k}^{xz}$ is the kernel correlation between predefined patch x and the new patch z; $\hat{\alpha}$ is the discriminative function and $\odot$ denotes element-wise multiplication. The pixel position where this regression value is the highest is the tracked location. In our implementation, we utilize a Gaussian kernel. We note here that the output of the KCF is usually a window of pixels where the target could be located; and not precisely localized on the target itself. Hence, in order to track the markers, we combine this with the previous template matching technique. Once a region of interest is returned by the KCF, we refine the detection by applying NCC template matching (subsection A) to that window, finally outputting the marker locations. Similar to template matching, this also requires an estimate of the initial positions from which the first window is isolated. Unlike template matching, which keeps only the template from the first instant in memory, the window of pixels is updated every frame during the KCF.

D. Saliency based detection

In the previous subsections, we described a method of detection and tracking involving template matching and ker-

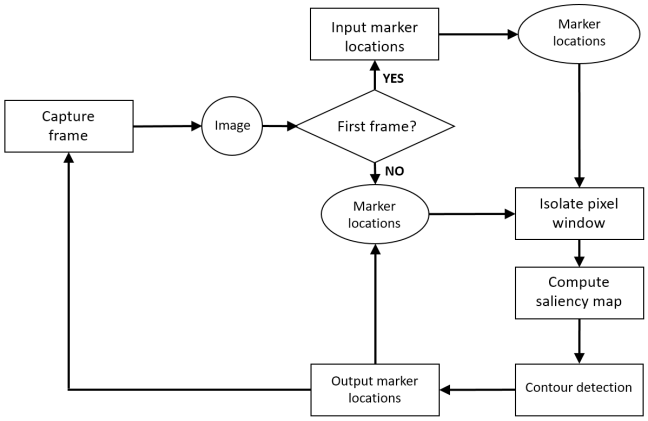

Fig. 4: Process flow of the salient object detection algorithm that uses minimum barrier salient object detection.

nelized cross correlation filters. For certain types of markers that undergo significant amounts of deformation within the body, the prior techniques would be inapplicable due to their continuously and unpredictably varying appearances. In order to handle such cases, we utilize the concept of 'salient object detection' in order to find the markers. The goal of salient object detection is to compute a representation of an image known as a saliency map, that highlights only the most important objects and suppresses the background in a scene. This difference between salient objects and the background is identified through the computation of an image distance transform. In our implementation, we use a technique named minimum barrier salient (MBS) object detection [11]. In MBS, this distance transform is obtained through the computation of a 'minimum barrier' distance (MBD) [12], represented as

$$\beta(t) = \max_{i=0}^{k} I(\pi(i)) - \min_{i=0}^{k} I(\pi(i)) \tag{2}$$

Here I denotes the image and π indicates a sequence of consecutive pixels. This distance is shown to be more robust to noise and blur than conventional methods such as geodesic distance, thus making it more suitable to the X-ray images that the real time tumor tracking algorithm deals with. The original method of computing the MBD has a complexity of $O(mnlog(n))$, where n is the number of pixels in the image and m being the number of distinct pixel values. In order to be applicable for a real-time method such as tumor tracking, saliency computation needs to be computationally efficient. Using the MBS algorithm also helps with this constraint, as the algorithm utilizes a fast and approximate iterative algorithm in order to compute the MBD, which can be regarded as having complexity $O(n)$.

In our implementation, we use the MBS algorithm for marker types and tumor locations that generally exhibit significant amount of motion and deformation, which could prevent template matching or KCF from working accurately. Similar to the previous techniques, the algorithm is initialized with an estimate of where the markers are at the first instant through user input. For the first frame, a neighborhood of pixels according to a predefined window size is considered around this estimate and the saliency analysis is performed on this pixel window in the next frame. Once saliency analysis is performed, the saliency map is thresholded into a binary image and contour detection is applied to it to compute precise locations of the markers. Centered around these new locations, a new pixel window is computed for the next frame. Unlike the previous techniques, there is no requirement to keep a template in memory: the algorithm only requires a position to continue tracking between frames and attempts to find the salient object(s) in every new pixel window.

V. 3D MOTION ESTIMATION

All of the techniques described above attempt to solve the detection and tracking problem in 2D, given images from X-ray fluoroscopy. For effective control of the proton beam, it is required to estimate the motion of the tumor in three dimensions. We recall that we have access to images from two X-ray detectors, providing a left and right view. In order to achieve 3D detection, we use the tracking algorithms on both views of the X-ray detectors and attempt to triangulate the obtained pixel locations to compute 3D positions of the markers.

In order to estimate the real world positions of the fiducial markers, the projective transformation between marker positions in the image and marker locations in real world should be known. We define 'fluoroscopic projection' matrices for the X-ray detectors, which encode this relationship, and estimate them through a calibration procedure. In this procedure, we used a 3D calibration target known as the MIMI Phantom. The MIMI Phantom is a cube of dimensions $140\times140\times140$ mm and contains thirty stainless steel spheres of 1 mm diameter implanted on five of the six faces and a high contrast sphere of diameter 6.4mm at the center of the cube (figure 5). This cube was installed on the patient platform and images were taken of the cube both in the continuous and pulsed types of fluoroscopy through the left and right detectors. In each image, we recorded the pixel positions of 13 of the spheres from the left and right views, whereas the real world positions of these spheres are already known.

We define two fluoroscopic projection matrices P_L and P_R, one for each detector. For each visible marker in the X-ray image of detector i, the corresponding matrix P_i encodes the equation of a straight line passing between the detector's iso-center and the location of the marker such that $[aP_i] = [b]$, where $[b]$ is the 2D marker position in the image expressed in homogeneous coordinates (i.e., the third coordinate is set to one), whereas $[a]$ is any position on the line in 3D space (also in homogeneous coordinates). P_i is a 4×3 matrix of rank 3. By recording the 3D-2D correspondences of 13 markers, we obtain 39 equations with 12 unknowns, which can be solved to output the parameters forming the fluoroscopic transformation matrix P_i. After obtaining these matrices, we attempted to reconstruction the

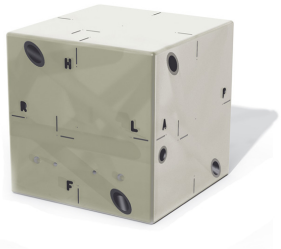

Fig. 5: Picture of a MIMI Phantom. For 13 of the 31 markers in the MIMI Phantom, we obtain 3D-2D correspondences from the left and right views of the fluoroscopy and thereby obtain fluoroscopic projection matrices for the two detectors.

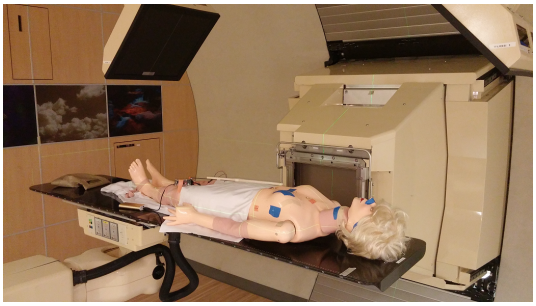

Fig. 6: Picture shows a 'phantom' model on the treatment table within the PROBEAT-V system at Mayo Clinic Arizona.

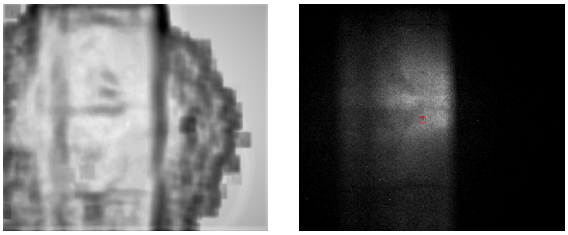

Fig. 7: NCC template matching attempts to detect a marker in a pelvic phantom with 10 cm of additional bolus. On the left, we show the template matching output: with the brightest pixel indicating the position. On the right, we show the successful detection, even in a noisy image and with a small marker size.

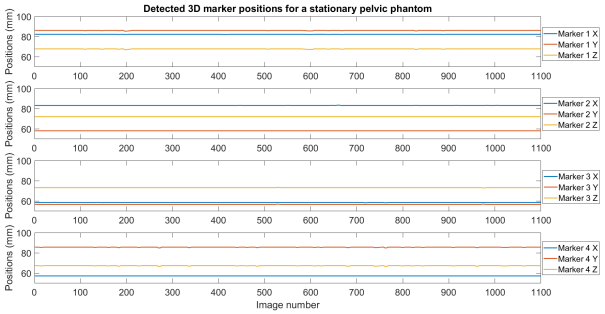

Fig. 8: NCC template matching applied to a stationary pelvic phantom with four markers imaged through pulsed fluoroscopy.

calibration cube using from both views, which we achieved with an average mean square error of 0.2 mm between the triangulated positions and ground truth, as well as a reprojection error of less than one pixel.

VI. Implementation and Results

In this section, we describe preliminary results obtained from the implementation of our algorithm. We have tested our algorithm with image data from both synthetic equipment as well as fluoroscopy datasets from clinical trials with real patients: with prostate, lung, liver and mediastinal tumors. Except for the one corresponding to the mediastinal tumor, the datasets were processed offline.

A. Pelvic phantom

Initially, we tested the algorithm on datasets of images recorded with a pelvic 'phantom': a realistic, tissue equivalent model of a human pelvis. We installed four metallic markers within the phantom and recorded multiple datasets under both continuous and pulsed fluoroscopy. For the datasets with the pelvic phantom, we attempted marker detection using the NCC template matching technique (figure 7).

In figure 8, we show the results of 3D triangulations of four markers installed in a stationary pelvic phantom that was imaged through pulsed X-ray fluoroscopy. The NCC algorithm was able to reliably track the markers through noisy images with an average error of < 0.1 mm.

We note here that the pelvic phantom is a representation of a very fit adult with low body fat. In order to mimic real patients, we also recorded images of the marker-implanted phantom along with additional bolus (representing layers of fat). The markers become less detectable with an increase in the amount of bolus. With our detection technique, we were able to reliably detect the markers even with additional bolus of 5, 10 and 15 cm.

B. Prostate tumor

The first dataset from a real patient that tumor tracking was tested with was for a prostate tumor. For this patient, carbon and ceramic fiducial markers were implanted in the tumor, and the left and right views of the prostate were recorded through pulsed fluoroscopy. Template matching was used to detect and track the markers in this video. Figure 9 shows sample results from the prostate tumor dataset, with three fiducial markers tracked in the fluoroscopy.

Fig. 9: Marker detection for a prostate tumor using template matching. Left: Fiducial markers visible in black through the X-ray image. Middle: NCC template matching output during marker detection. Right: Detected markers marked with red circles.

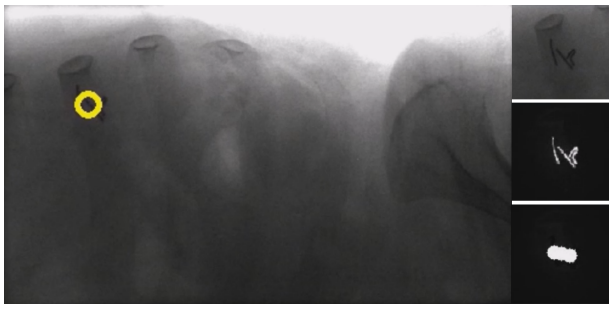

Fig. 10: Marker detection for a liver tumor using saliency analysis. On the right, we show three inset pictures: the region of interest around the markers (top), salient object detection (middle) and contour detection for marker positions (bottom). This information is used to output a final position (left).

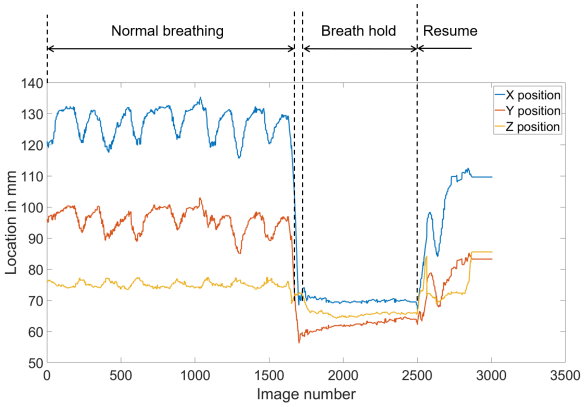

Fig. 11: Plots of the X, Y and Z positions of the markers for a liver tumor, showing different phases of the patient breathing cycle.

C. Liver tumor

As a next test case, we obtained fluoroscopic data of a patient with a liver tumor. In this case, linear gold fiducials known as 'Visicoil' markers were implanted in a small area around the tumor. In this dataset, we do observe significant motion in the markers coming from normal breathing phases as well as breath-holding and release as part of the treatment procedure. We chose salient object detection as the algorithm for this dataset. In figure 10, we show the working of saliency analysis: which detects the markers as the important objects in the scene, suppressing the background. Finally, we isolate the positions of the three markers through contour detection and consider the centroid of these points as the center of the region of interest. (In this case, as the markers were implanted close together, we considered the markers together as a point of interest). We performed this detection and tracking on both the left and right views, resulting in triangulated 3D positions of the marker positions. In figure 11, we show a plot of the X, Y and Z positions during the dataset: where the distinction between normal breathing, breath-hold and breath-release is clearly visible.

D. Lung tumor

Another dataset was obtained from a patient with a lung tumor and a single implanted visicoil marker. Again, given

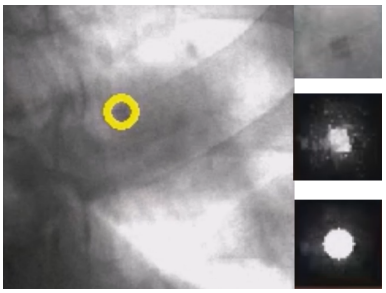

Fig. 12: Saliency analysis used for detecting a single visicoil marker in a lung tumor. This figure shows that saliency is able to capture the marker even through motion (and resultant blur). Similar to figure 10, right insets show the various phases of saliency analysis.

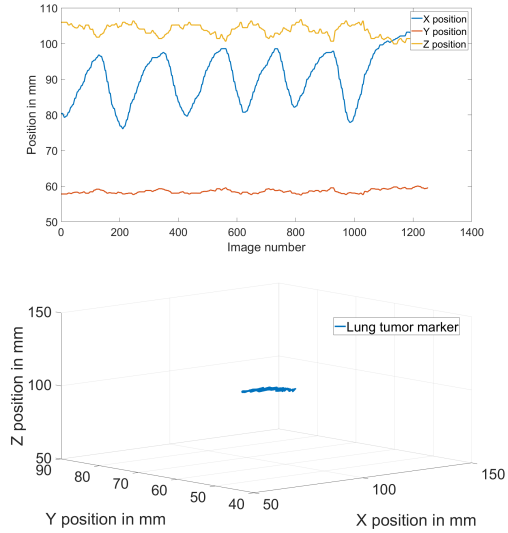

Fig. 13: Top: X, Y and Z positions of the visicoil marker in the lung tumor tracked through normal breathing motion. Bottom: Marker position and movement plotted in 3D space.

the location of the tumor, there was significant movement of the marker coming from normal breathing motion of the lungs. In this case also, we used saliency analysis as our preferred technique. In figure 12, we show a particularly challenging frame where the marker was in motion at a considerable amount of speed, thus resulting in blur in the X-ray image. The saliency analysis was still able to isolate this marker from the background, resulting in accurate positioning in 2D. Similar to the liver tumor, we triangulated the 3D position of this marker using both L and R views of the X-ray. We show the X,Y and Z positions of the marker over time and the 3D plot of its movement in figure 13. While we do not yet have a quantitative comparison of this detected movement with actual motion, these measurements correspond favorably to real life.

E. Mediastinal tumor

Mediastinal tumors are cancerous growths that form in the area of the chest between the lungs. For a patient with mediastinal tumor, titanium clips were implanted around the tumor, and during this clinical trial, we implemented

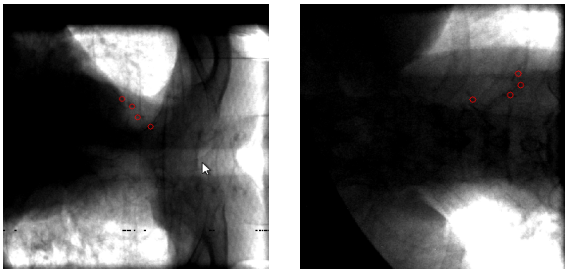

Fig. 14: Four titanium clips were implanted at a mediastinal tumor site and were tracked by our system in real-time. A sample frame showing the detection of the markers is seen here (detected marker locations in red).

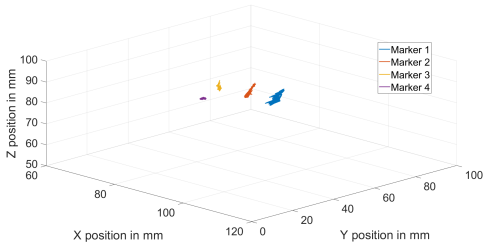

Fig. 15: 3D plot of four titanium markers implanted in a mediastinal tumor.

the algorithm in real time for detecting and tracking the tumor markers. Pulsed fluroscopy method was used to stream the images from the treatment room to the detection and tracking computer and we used the KCF + template matching technique to locate and track the markers robustly even through slight motion and rotation. Figure 14 shows a sample picture of output from our detection algorithm during real-time operation indicating pixel positions of four markers (clips) in both the left and right views. Figure 15 shows 3D plots of the four markers and their extent of motion.

VII. Conclusions

In this paper, we present a system for detection and tracking of tumors for a pencil beam scanning proton therapy system. Fiducial markers of various materials are implanted either in the tumors themselves or the surrounding areas and are subsequently imaged through an X-ray system operating in fluoroscopy mode. We utilize an array of algorithms such as cross-correlation template matching, kernelized correlation filters and saliency maps. Our algorithm requires an estimate of the initial location of the markers or an approximate region of interest at initialization and is independent of the size and shape of the markers. Our system is capable of working in real-time and provides multiple techniques to choose from based on marker/tumor types. We have applied our methods to both synthetic data and data from real patients, as well as in real-time. In this paper, we demonstrated robust detection and tracking of different types of fiducial markers in several tumor areas through movement.

We have identified a few areas for further improvement of our algorithms. Firstly, we wish to improve the working of the algorithm in case of significant blur by incorporating spatio-temporal information, in order to minimize the possibility of tracking loss. Secondly, we wish to couple the algorithms with a probabilistic filtering technique (e.g. Monte Carlo methods) such that in case tracking loss does occur, the system can output a probable region of where the marker could be and thus help in effective control of the proton beam system. At the same time, a robust filtering technique would have the ability to handle outliers efficiently. We also wish to investigate the possibility of estimating orientation of the tumors using information from multiple markers. In the future, we plan to integrate this system into the proton beam facility at Mayo Clinic Arizona to achieve beam control through real-time image guidance.

References

[1] Hiroki Shirato, Shinichi Shimizu, Tatsuya Kunieda, Kei Kitamura, Marcel van Herk, Kenji Kagei, Takeshi Nishioka, Seiko Hashimoto, Katsuhisa Fujita, Hidefumi Aoyama, et al. Physical aspects of a real-time tumor-tracking system for gated radiotherapy. *International Journal of Radiation Oncology Biology Physics*, 48(4):1187–1195, 2000.

[2] Johanna Brewer, Margrit Betke, David P Gierga, and George TY Chen. Real-time 4d tumor tracking and modeling from internal and external fiducials in fluoroscopy. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 594–601. Springer, 2004.

[3] Qianyi Xu, Russell J Hamilton, Robert A Schowengerdt, Brian Alexander, and Steve B Jiang. Lung tumor tracking in fluoroscopic video based on optical flow. *Medical physics*, 35(12):5351–5359, 2008.

[4] Twyla R Willoughby, Patrick A Kupelian, Jean Pouliot, Katsuto Shinohara, Michelle Aubin, Mack Roach, Lisa L Skrumeda, James M Balter, Dale W Litzenberg, Scott W Hadley, et al. Target localization and real-time tracking using the calypso 4d localization system in patients with localized prostate cancer. *International Journal of Radiation Oncology Biology Physics*, 65(2):528–534, 2006.

[5] Yvette Seppenwoolde, Ross I Berbeco, Seiko Nishioka, Hiroki Shirato, and Ben Heijmen. Accuracy of tumor motion compensation algorithm from a robotic respiratory tracking system: a simulation study. *Medical physics*, 34(7):2774–2784, 2007.

[6] Seishin Takao, Naoki Miyamoto, Taeko Matsuura, Rikiya Onimaru, Norio Katoh, Tetsuya Inoue, Kenneth Lee Sutherland, Ryusuke Suzuki, Hiroki Shirato, and Shinichi Shimizu. Intrafractional baseline shift or drift of lung tumor motion during gated radiation therapy with a real-time tumor-tracking system. *International Journal of Radiation Oncology Biology Physics*, 94(1):172–180, 2016.

[7] Venkateswarlu Karnati, Mithun Uliyar, and Sumit Dey. Fast non-local algorithm for image denoising. In *Image Processing (ICIP), 2009 16th IEEE International Conference on*, pages 3873–3876. IEEE, 2009.

[8] Stephen M Pizer, E Philip Amburn, John D Austin, Robert Cromartie, Ari Geselowitz, Trey Greer, Bart ter Haar Romeny, John B Zimmerman, and Karel Zuiderveld. Adaptive histogram equalization and its variations. *Computer vision, graphics, and image processing*, 39(3):355–368, 1987.

[9] Kai Briechle and Uwe D Hanebeck. Template matching using fast normalized cross correlation. In *Optical Pattern Recognition XII*, volume 4387, pages 95–103. International Society for Optics and Photonics, 2001.

[10] João F Henriques, Rui Caseiro, Pedro Martins, and Jorge Batista. High-speed tracking with kernelized correlation filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(3):583–596, 2015.

[11] Jianming Zhang, Stan Sclaroff, Zhe Lin, Xiaohui Shen, Brian Price, and Radomir Mech. Minimum barrier salient object detection at 80 fps. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1404–1412, 2015.

[12] Robin Strand, Krzysztof Chris Ciesielski, Filip Malmberg, and Punam K Saha. The minimum barrier distance. *Computer Vision and Image Understanding*, 117(4):429–437, 2013.