Depth Sensor-Based Realtime Tumor Tracking for Accurate Radiation Therapy

Björn Nutti[1,2] Åsa Kronander[1] Mattias Nilsing[1] Kristofer Maad[1] Cristina Svensson[1] Hao Li[3]

[1]C-RAD Positioning AB [2]AlgoritmFabriken AB [3]University of Southern California

Abstract

We present an image guided radiation therapy (IGRT) system for tracking tumors in realtime based on continuous structured light imaging. While an accurate positioning of the radiation isocenter to pre-imaged cancer cells is critical to minimize the risk of damaging healthy tissues, patients undergo involuntary motions such as breathing or unpredictable gestures during treatment. Moreover, multiple sessions are typically necessary and repositioning the patient accurately can be difficult. Our approach consists of determining the tumor position by densely tracking the deformation of a stream of 3D scans using a realtime variant of a state-of-the-art non-rigid registration algorithm and an FEM simulation on the interior body. We use interactive reprojection for visual guidance to adjust the posture of the patient and couch position, depending on the tumor location. Compared to existing techniques, our method uniquely estimates tumor deviations under body deformations. Our pipeline has been successfully commercialized as part of the C-RAD AB Catalyst™ product line and is already deployed in a number of hospitals.

1. Introduction

Since the advent of fractioned radiation therapy, image guided diagnostics have commonly been used to ensure the accurate placement of cancer patients for effective and safe treatment. Over the course of multiple sessions, the position of cancer cells typically move from its initial location due to the changes in pose and anatomy (e.g., weight loss). Image guided radiation therapy (IGRT) consists of an imaging process that updates the 3D position of cancerous tumors before each treatment. Another important use of IGRT is to minimize the exposure of healthy tissues to radiation while the patient is breathing by using a 4DCT imaging procedure and turning the beams on and off at the right respiratory cycle. While being reasonably accurate, computer tomography (CT)-based IGRT can by itself increase the risk of a secondary cancer and long term malignancy as the level of radiation can add up significantly when used too often.

To reduce the use of X-ray imaging, *optical tracking systems* are commonly employed to estimate tumor positions after an initial volumetric capture and to safely ensure proper alignment and monitoring the patient during treatment. Accurate and interactive 3D localization of the tumor is particularly challenging since the body of the patient is under constant deformation (e.g., breathing, involuntary gestures). Existing systems typically assume static or minimal deformations of the subject and determine the radiation isocenter based on global rigid surface alignment. Although marker-based motion capture techniques have been investigated [VKK*03], the setup times for each patient would be excessive for practical deployment.

We introduce a depth sensor-based tracking system for real-time and sub-millimeter accurate positioning of internal targets relative to the linear accelerator (linac) isocenter. Recent advances in 3D scanning technology and geometric capture of realistic human performances have allowed us to successfully adopt a state-of-the-art non-rigid registration algorithm [LAGP09] coupled with an efficient volumetric FEM simulation [MG04] to track tumors at interactive rates while the body can exhibit significant and arbitrary deformations. An integral part of the system is an interactive reprojection of visual guidance patterns directly onto the patient's body to adjust for misaligned postures during patient setup. Clinical studies have shown that our approach greatly improves the accuracy and safety of radiation while saving time for correct patient setup and pose adjustment.

Contribution.

- Comprehensive depth sensor-based tumor tracking framework that effectively combines non-rigid surface tracking, physical simulation, and interactive visual guidance for accurate radiation therapy.
- Real-time variant of a state-of-the-art non-rigid registration algorithm [LAGP09] for deformable surface tracking.

Background. While imaging based on either cone beam CT (CBCT) or electronic portal imaging device (EPID) is indispensable for patient alignment during IGRT, they are potentially dangerous due to extra level of radiation and the residual uncertainties caused by the slow and blurry volumetric imaging procedure [GRT*08]. Hence a wide range of 3D acquisition based tracking systems have been proposed

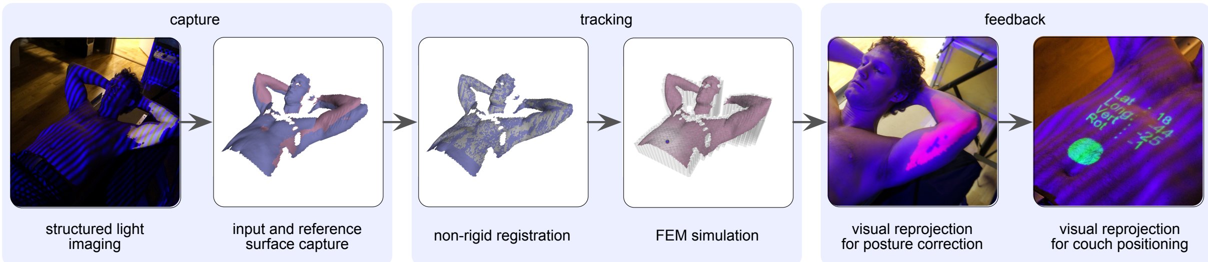

Figure 1: *Our pipeline begins with a real-time structured light-based surface capture. The tumor position is then tracked using a non-rigid registration and FEM simulation framework. Finally, visual guidance is projected onto the patient for accurate posture correction and couch positioning. Notice that the reference surface can either be the initial frame or a full CT capture.*

to provide a relative positioning between the *reference* CT structure set (mesh segmented from CT imaging) and the *live* surface acquisition. Existing patient positioning systems such as the laser-based Sentinel (C-RAD AB), the structured light imaging-based AlignRT [Vis09], or methods based on time-of-flight (TOF) [PSS*12] rely on static surface alignment techniques such as the iterative closest point algorithm (ICP) [RL01], even though these systems are free of fiducial markers and can acquire a reasonable amount of geometric details. Deformations such as inter-fraction posture variations, respiratory motion, and uncontrolled gestures can cause rigid surface alignment techniques to produce estimate tumor positioins inaccurately and negatively impact dose delivery.

More recently, a wide range of non-rigid alignment algorithms have been proposed, mostly motivated by the increasing use of real-time 3D scanning in computer graphics. An overview can be found in [CLM*10]. We adopted the deformable alignment method introduced in [LAGP09] due to its robustness to handle continuous surfaces that are poor in geometric features and scans that may be largely incomplete due to occlusions. The method has been successfully deployed in the visual effects industry for markerless facial performance capture and also used for high precision analysis of cardiac surface mechanics based on structured light imaging [LZL*12]. Since the method is formulated as a nonlinear optimization problem, the original technique could take several seconds to align a pair of scans. Through a simple axis-angle rotation-based deformation representation and an extensive use of multi-threading, we were able to achieve near interactive rates without sacrificing accuracy.

Since the real location of a tumor is described by the anatomy and complex internal organs of the patient, estimating its position in real-time from captured surface deformations can only be achieved through a simplified physical simulation. The typical assumption in interactive virtual surgeries is that soft body tissues behave in a plasto-elastic way. While a wide range of FEM-based physical simulation and model reduction techniques exist [ESHD05, SSB12], we adopted a variant of the widely spread unconditionally stable method of [MG04].

2. System Overview

Our system consists of a light weight rig that combines a depth sensor and a projector, which can be easily attached to the ceiling of any radiation therapy suite. The depth sensor uses near UV light patterns so that the synchronized reprojected visual guidance does not interfere with the capture process. Our framework then calculates the posture and setup error based on the tumor's position in the current frame relative to the reference. Figure 1 illustrates the pipeline.

Capture. Our reference geometry is either a skin geometry segmented from a CT image (ca. 30K vertices) or a surface scan (ca. 10K vertices). When the reference comes from CT, we apply isotropic remeshing [BK04] to improve the mesh quality for numerical stability during tracking. In case no skin geometry is available, the patient is positioned with information provided by a second IGRT system. When the patient is properly positioned, a surface scan is captured for future reference.The next step consists of aligning the surface of the reference 3D mesh to a stream of 3D scans using non-rigid registration.

Tracking. To ensure robust and drift-free surface tracking, we use an elastic as-rigid-as-possible deformation model [SSP07] to describe the transformation between the reference mesh and the captured surfaces. The reference mesh defines the rest state of the elasticity. A volumetric FEM simulator then uses the vertices of the tracked surface as constraints to determine the position of the isocenter in the new frame.

Feedback. The relative position between the reference tumor isocenter and the current one is used to monitor whether or not the tumor is moving out of place (e.g., due to gestures such as a tilting head or a moving arm). Since our surface tracking gives us a dense one-to-one correspondence between the reference mesh and the current range capture, we can precisely determine the optimal pose of the patient to match the tumor to the linac radiation beam. We project colored alignment patterns onto the patient's skin as a visual guidance to the clinician.

3. Deformable Surface Tracking

For every incoming input scan, we need to deform the reference mesh to the current frame in order to establish full correspondences of the recording. These correspondences are used as constraints for the body interior simulation to locate the tumor position and for the interactive visual guidance to

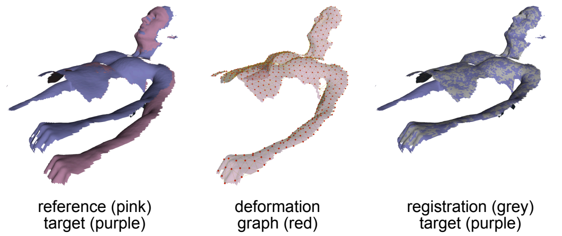

Figure 2: *Graph-based non rigid ICP for surface tracking.*

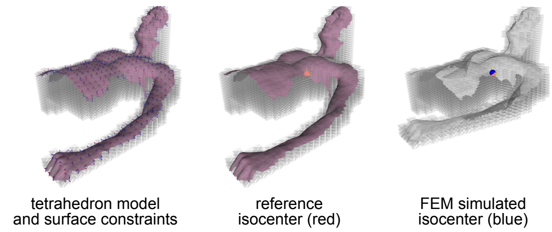

Figure 3: *FEM Simulation for isocenter positioning.*

correct misaligned postures. The patient can exhibit arbitrary deformations and because of occlusions, clothing, and devices on the subject the input scans can be largely incomplete during the treatment as opposed to the reference frame which is obtained by the volumetric imaging.

After a careful evaluation of existing techniques, we adopted the graph-based non-rigid ICP algorithm introduced in [LAGP09] due to its accuracy and robustness to outliers and incomplete data (c.f., Figure 2). The deformable ICP framework iterates between the closest point computation and a non-linear deformation optimization that maximizes local rigidity using a Gauss-Newton Solver and Sparse Cholesky Decomposition. Multiple iterations (up to 100) of a sparse linear solve are required for each deformation computation to ensure convergence, which is the largest bottleneck. As opposed to the original formulation, instead of reinitializing the optimization between consecutive input frames with decreasing stiffness parameters, we use the initial frame as a rest pose configuration and the registration is computed for incoming frames without changing the regularization energy weights to avoid drifts.

Speed up. To achieve interactive performance, we first modify the graph-based deformation model to implicitly describe local rigid motions in each node using an axis-angle rotation representation instead of a more general local affine transformation. Also, instead of adapting the graph node density to the deformation, we only compute it once with a uniform sub-sampling and use the same graph throughout the tracking. Our axis-angle rotation-based representation reduces the normal equation dimension in each Gauss-Newton iteration from $12n$ to $6n$ where n is the number graph nodes. Additionally, the regularization term $E_{\text{rigid}} = (\mathbf{a}_1^\top \mathbf{a}_2)^2 + (\mathbf{a}_1^\top \mathbf{a}_3)^2 + (\mathbf{a}_2^\top \mathbf{a}_3)^2 + (1-\mathbf{a}_1^\top \mathbf{a}_1)^2 + (1-\mathbf{a}_2^\top \mathbf{a}_2)^2 + (1-\mathbf{a}_3^\top \mathbf{a}_3)^2$ can be discarded since the axis-angle rotation matrix is orthonormal. We simply use Rodrigues' rotation formula to derive the axis-angle rotation. More details on the other energy terms can be found in [LAGP09]. The remaining performance improvements are obtained by exploiting the block structures of the sparse linear system where the normal equations can be assembled in parallel. Depending on the type of deformation, this implementation achieves in practice a speed up of factor 50 over the original version.

4. Body Interior Simulation

The 3D position of the tumor is continuously updated using a volumetric FEM simulator that uses the tracked surface vertices as displacement constraints. Our technique is based on the efficient linear model from [MG04] which is frequently used in virtual surgery. The interior body is first discretized into tetrahedrons using the reference geometry (c.f., Figure 3). In contrast to the original work, we apply displacements as boundary conditions rather than forces. To ensure robustness during the solve, we discard dynamic properties and treat the simulation as a static problem. Vertices of tetrahedrons that enclose the captured surface are added as constraints weighted through linear blend skinning. We then compute the deformation using a conjugate gradient solver with a successive symmetric-over relaxation (SSOR) pre-conditioner which reduces the number of iterations by a factor of five. The tumor isocenter is then obtained using barycentric coordinates. Since the method is linear, it can potentially yield undesirable stretches and not preserve volume properly. However, we found that our simulation were sufficiently accurate since the captured deformations and rotations are small.

5. Interactive Visual Guidance

Our unique interactive visual guidance system takes the information from the non-rigid surface tracking and FEM simulation to instruct the clinician on how to adjust for misaligned postures during patient setup. While existing tracking systems rely on immobilization devices and global positioning, our framework projects a pattern on the patient's skin and accurately highlights incorrect body postures and correction cues using colored patterns. While the body pose is being adjusted, our system also projects the exact displacement coordinates to position the couch.

6. Results

Performance. Our pipeline performs at up to 10 fps (Intel Xeon 6-Core at 3.2GHz). While 10 iterations of non-rigid ICP and 20 iterations of FEM solves are used, most computation is spent in the non-rigid registration. The system can also increase the number of iterations to handle potentially larger deformations. The interactive visual guidance is updated at 5 fps. The evaluations are based on an earlier version of the Catalyst™ system and do not reflect the current product.

Clinical Studies. The pilot studies have been conducted at several European radiation therapy centers including Herlev Hospital in Denmark, the University Medical Center in Mannheim and the Skåne University Hospital (SUS) [KÏ12].

Kügele [KÏ2] has compared our system to Varian's on-board imager (OBI) on 90 treatment fractions and observed an average tumor position difference of [-2.55mm(lat), 1.69mm(long), 2.09mm(vert)] and average deviation of 3.7mm. In the same study, our system was compared to a surface measurement system that comprises a room laser and markers that are applied to the skin. An average difference of [-0.05mm(lat), 0.93mm(long), -0.98mm(vert)] and average deviation of 1.35mm was measured. The deviations are due to the fact that setup variations on the surface do not entirely correspond to those of the spine and sternum. The results were encouraging as we could accurately predict internal tumor positions from dynamically captured surfaces.

Studies conducted at the University Medical Center in Mannheim [SWSL13] have shown very accurate non-rigid tracking for 224 treatment fractions in 13 patients for head/neck, thorax, and pelvis regions. The mean tracking deviations between the Catalyst™ and CBCT were all below [0.07 cm(lat), 0.13 cm(long), 0.15cm(vert), 0.43°(roll), 0.1° (pitch), 0.11° (yaw)]. These results demonstrate that our system is effective in reducing the use of CBCT imaging in patients where the tumor locations are correlated with the surface deformation.

breast	MVCT	Catalyst
x [mm]	1	-2.5
y [mm]	-9	-9.8
z [mm]	-6	-9.2

head/neck	OBI	Catalyst
x [mm]	-1.3	-0.4
y [mm]	1.3	0
z [mm]	8.1	9.1

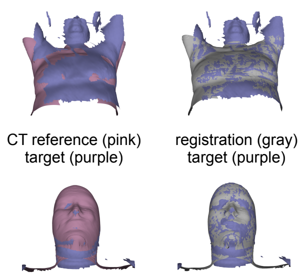

Figure 4: *Tumor tracking comparison between our surface-based approach and CT imaging (MVCT or OBI).*

With the assistance from clinics, we conducted a series of quantitative evaluations for different body deformations. We compared our tracked tumor positions to megavoltage CT (MVCT) and Varian's OBI. Figure 4 illustrates two common scenarios where postures might change substantially between different treatments:

- Head/Neck: any misalignment of the shoulder postures can displace the tumors that are located in the head and neck region.
- Breast: during the treatment of tumors located in the breast area, patients typically have their arms above the head with elbows pointing up. The reference pose can therefore be significantly different.

7. Conclusion & Future Work

Clinical studies of our depth sensor-based tumor tracking framework have generally shown an increase in accuracy and patient safety during the treatment of tumors that are located close to the skin such as breast and lung cancers. In particular, the use of our surface-based non-rigid registration technique was particularly effective in capturing very subtle skin deformations and handling arm motions. In the future, we plan to investigate more sophisticated methods for collision avoidance, the ability to distinguish the skin from the clothing using statistical human body shape priors [HSS*09], and bio-mechanics inspired models for more accurate physical simulation of the body interior.

References

[BK04] Botsch M., Kobbelt L.: A remeshing approach to multiresolution modeling. In *Proceedings of the 2004 Eurographics/ACM SIGGRAPH symposium on Geometry processing* (2004), SGP '04, pp. 185–192. 2

[CLM*10] Chang W., Li H., Mitra N. J., Pauly M., Wand M.: Geometric registration for deformable shapes. In *Eurographics 2010: Tutorial Notes* (2010). 2

[ESHD05] Erleben K., Sporring J., Henriksen K., Dohlman K.: *Physics-based Animation (Graphics Series)*. Charles River Media, Inc., Rockland, MA, USA, 2005. 2

[GRT*08] Gierga D., Riboldi M., Turcotte J., Sharp G., Jiang S., Taghian A., Chen G.: Comparison of target registration errors for multiple image-guided techniques in accelerated partial breast irradiation. *Int J Radiat Oncol Biol Phys* (2008). 1

[HSS*09] Hasler N., Stoll C., Sunkel M., Rosenhahn B., Seidel H.-P.: A statistical model of human pose and body shape. In *Computer Graphics Forum (Proc. Eurographics 2008)* (Mar. 2009), vol. 2. 4

[KÏ2] Kügele M.: *Evaluation of the Catalyst System for Patient Positioning during Breast Cancer Treatment*. Master's thesis, Lund University, 2012. 3, 4

[LAGP09] Li H., Adams B., Guibas L. J., Pauly M.: Robust single-view geometry and motion reconstruction. *ACM Transactions on Graphics (SIGGRAPH Asia) 28*, 5 (2009). 1, 2, 3

[LZL*12] Laughner J. I., Zhang S., Li H., Shao C. C., Efimov I. R.: Mapping cardiact surface mechanics with structured light imaging. *American Journal of Physiology, Heart and Circulatory Physiology* (2012). 2

[MG04] Müller M., Gross M.: Interactive virtual materials. In *Proceedings of Graphics Interface 2004* (School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada, 2004), GI '04, Canadian Human-Computer Communications Society, pp. 239–246. 1, 2, 3

[PSS*12] Placht S., Stancanello J., Schaller C., Balda M., Angelopoulou E.: Fast time-of-flight camera based surface registration for radiotherapy patient positioning. *Medical Physics 39*, 1 (2012). 2

[RL01] Rusinkiewicz S., Levoy M.: Efficient variants of the icp algorithm. In *International Conference on 3-D Digital Imaging and Modeling* (2001). 2

[SSB12] Sin F. S., Schroeder D., Barbič J.: Vega: Nonlinear fem deformable object simulator. In *Computer Graphics Forum, to appear* (2012). 2

[SSP07] Sumner R. W., Schmid J., Pauly M.: Embedded deformation for shape manipulation. In *ACM SIGGRAPH 2007 papers* (New York, NY, USA, 2007), SIGGRAPH '07. 2

[SWSL13] Stieler F., Wenz F., Shi M., Lohr F.: A novel surface imaging system for patient positioning and surveillance during radiotherapy. *Strahlentherapie und Onkologie 189*, 11 (2013), 938–944. 4

[Vis09] VisionRT: http://www.visionrt.com/, 2009. 2

[VKK*03] Vedam S. S., Kini V. R., Keall P. J., Ramakrishnan V., Mostafavi H., Mohan R.: Quantifying the predictability of diaphragm motion during respiration with a noninvasive external marker. *Medical Physics 30*, 4 (2003). 1