Photometric Estimation of 3-D Surface Motion Fields for Respiration Management

Sebastian Bauer[1], Jakob Wasza[1], Joachim Hornegger[1,2]

[1]Pattern Recognition Lab, Department of Computer Science
[2]Erlangen Graduate School in Advanced Optical Technologies (SAOT)
Friedrich-Alexander-Universität Erlangen-Nürnberg
sebastian.bauer@cs.fau.de

Abstract. In radiation therapy, the estimation of torso deformations due to respiratory motion is an essential component for real-time tumor tracking solutions. Using range imaging (RI) sensors for continuous monitoring during the treatment, the 3-D surface motion field is reconstructed by a non-rigid registration of the patient's instantaneous body surface to a reference. Typically, surface registration approaches rely on the pure topology of the target. However, for RI modalities that additionally capture photometric data, we expect the registration to benefit from incorporating this secondary source of information. Hence, in this paper, we propose a method for the estimation of 3-D surface motion fields using an optical flow framework in the 2-D photometric domain. In experiments on real data from healthy volunteers, our photometric method outperformed a geometry-driven surface registration by 6.5% and 22.5% for normal and deep thoracic breathing, respectively. Both the qualitative and quantitative results indicate that the incorporation of photometric information provides a more realistic deformation estimation regarding the human respiratory system.

1 Introduction

Respiratory motion management is an evolving field of research in radiation therapy (RT) and of particular importance for patients with thoracic, abdominal and pelvic tumors. Facing target locations in the upper torso, besides inaccuracies in patient setup and positioning, respiratory motion during treatment delivery induces a fundamental error source. To date, in clinical practice, the tumor is irradiated using RT gating techniques where the linear accelerator is triggered by an external 1-D respiration surrogate [1]. However, gating entails a low duty cycle, increasing the treatment time and hindering an efficient operation of the therapy facility. In contrast, real-time tumor tracking solutions [2,3,4] re-position the radiation beam dynamically to follow the tumor's changing position. Under ideal conditions, tracking can eliminate the need for a tumor-motion margin in the dose distribution while maintaining a 100% duty cycle for RT delivery [1].

In particular, methods that infer the internal tumor position from external torso deformations are expected to improve radiation therapy. Based on real-time range imaging (RI), the 3-D surface motion field of the patient's surface with respect to a reference is identified and related to a previously learned model correlating the torso deformation with the target position [2]. Estimating the displacement field using conventional surface registration techniques relies on the pure 3-D topology of the patient. Instead, using modern RI modalities that additionally capture photometric data, we expect the registration to benefit from this secondary information.

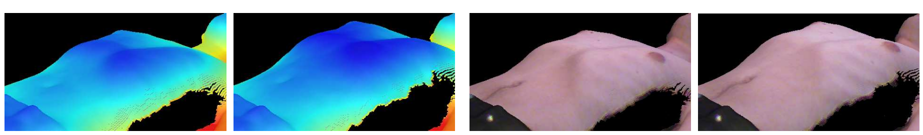
Hence, in this paper, we introduce a method for the identification of a dense 3-D surface motion field over non-rigidly moving surfaces observed by RI cameras. Instead of capitalizing on the acquired surface topology, we propose to estimate the optical flow in the 2-D photometric domain. Based on the known relation between the sensor domain and the corresponding surface in world coordinate space, we then deduce the 3-D surface motion field. In experiments on real data from Microsoft Kinect, we have investigated the surface motion fields estimated with our method compared to a purely geometry-driven registration.

2 Materials and Methods

The proposed method for estimation of a dense 3-D surface motion field of the patient's respiration state with respect to a reference relies on RI devices that deliver both photometric color and metric depth (RGB-D) information of the scene. Below, let $g(\boldsymbol{\zeta})$ and $f(\boldsymbol{\zeta})$ denote the geometric depth and photometric color measurements at a position $\boldsymbol{\zeta} = (\zeta_1, \zeta_2)^T$ in the 2-D sensor domain Ω. Indeed, based on the pinhole camera model, an orthogonal depth measurement $g(\boldsymbol{\zeta})$ describes a world coordinate position vector $\boldsymbol{x}(\boldsymbol{\zeta}) = (x, y, z)^T \in \mathbb{R}^3$. In homogeneous coordinates, this transformation can be denoted as:

$$\begin{pmatrix} \boldsymbol{x}(\boldsymbol{\zeta}) \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{g(\boldsymbol{\zeta})}{\beta_x} & 0 & 0 \\ 0 & \frac{g(\boldsymbol{\zeta})}{\beta_y} & 0 \\ 0 & 0 & g(\boldsymbol{\zeta}) \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \boldsymbol{\zeta} \\ 1 \end{pmatrix} , \tag{1}$$

where β_x, β_y denote the focal length. Using triangulation techniques, the point cloud $\boldsymbol{X} = \{\boldsymbol{x}\}$ can be interpreted as a 3-D surface $\mathcal{G}$. Fig. 1 illustrates the torso



$g(\boldsymbol{\zeta})$ (exhale) $g(\boldsymbol{\zeta})$ (inhale) $f(\boldsymbol{\zeta})$ (exhale) $f(\boldsymbol{\zeta})$ (inhale)

Fig. 1. RGB-D data at different respiration states (fully exhale/inhale). On the left, the measured orthogonal depth $g(\boldsymbol{\zeta})$ is color-coded. On the right, the additionally acquired photometric information $f(\boldsymbol{\zeta})$ is mapped onto the 3-D surface.

surface acquired from a male subject, textured with the color-coded orthogonal depth and photometric information, respectively. Below, we introduce our proposed method for photometric estimation of the 3-D deformation. In addition, we oppose a geometry-driven surface registration method. For the purpose of enhanced comparability and in regard to a potential combined formulation in future work, both approaches rely on a variational formulation.

2.1 Photometry-Driven Surface Registration

The proposed method for estimation of a dense 3-D displacement field from 2-D photometric information is based on a two-stage procedure: First, we interpret the acquired photometric information as a conventional planar image and compute a dense optical flow field. Second, we build on this 2-D deformation to extract a dense 3-D surface motion field.

In this work, we have used the combined local-global (CLG) method for optical flow computation proposed by Bruhn et al. [5]. Based on a variational formulation, it combines the advantages of two classical algorithms: the variational approach by Horn and Schunck [6] providing dense flow fields, and the local least-square technique of Lucas and Kanade [7] featuring robustness with respect to noise. The CLG method computes the 2-D photometric optical flow $\tilde{\boldsymbol{u}}_p(\boldsymbol{\zeta}) = (\tilde{u}_p(\boldsymbol{\zeta}), \tilde{v}_p(\boldsymbol{\zeta}), 1)^T$ as the minimizer of the energy functional:

$$\mathcal{E}[\tilde{\boldsymbol{u}}_p] = \int_{\Omega} \Big(\tilde{\boldsymbol{u}}_p(\boldsymbol{\zeta})^T J_\rho(\nabla_3 f) \tilde{\boldsymbol{u}}_p(\boldsymbol{\zeta}) + \alpha_p \| D\tilde{\boldsymbol{u}}_p(\boldsymbol{\zeta}) \|_F^2 \Big)\, d\boldsymbol{\zeta}\,, \tag{2}$$

where $\tilde{u}_p$ and $\tilde{v}_p$ denote the displacement in direction of ζ_1 and ζ_2, respectively. Further, using the original formulation [5], $\nabla_3 f = (f_x, f_y, f_t)^T$ denotes the spatiotemporal gradient, $J_\rho(\nabla_3 f)$ the structure tensor with some integration scale ρ, $D\tilde{\boldsymbol{u}}_p$ the Jacobian matrix of $\tilde{\boldsymbol{u}}_p$, $\|\cdot\|_F$ the Frobenius norm, and α_p a non-negative regularization weight. In our notation, a tilde placed on top of a variable denotes that it lives in 2-D pixel space, otherwise in 3-D metric real-world space. The numerical minimization of the energy functional $\mathcal{E}[\tilde{\boldsymbol{u}}_p]$ in Eq. 2 is performed by a conjugate gradient solver with a finite difference approximation for spatial discretization [8]. Based on the estimated flow field $\tilde{\boldsymbol{u}}_p$, we are now in the position to infer the 3-D surface motion field $\boldsymbol{u}_p = (u_p, v_p, w_p)^T$ between two respiration states t_1 and t_2:

$$\boldsymbol{u}_p(\boldsymbol{\zeta}) = \boldsymbol{x}_{t_2}\big(\boldsymbol{\zeta} + (\tilde{u}_p(\boldsymbol{\zeta}), \tilde{v}_p(\boldsymbol{\zeta}))^T\big) - \boldsymbol{x}_{t_1}(\boldsymbol{\zeta})\,, \tag{3}$$

using bilinear interpolation in the sensor domain Ω for computing the position $\boldsymbol{x}_{t_2}\big(\boldsymbol{\zeta} + (\tilde{u}_p(\boldsymbol{\zeta}), \tilde{v}_p(\boldsymbol{\zeta}))^T\big)$ on the surface $\mathcal{G}_{t_2}$.

2.2 Geometry-Driven Surface Registration

For evaluation of the proposed photometric approach, let us compare the estimated surface motion field $\boldsymbol{u}_p$ to a geometry-driven surface registration, based on [9]. Here, we represent the surface $\mathcal{G}_{t_2}$ at time t_2 by its corresponding signed distance function $d(\boldsymbol{x}) := \pm dist(\boldsymbol{x}, \mathcal{G}_{t_2})$, where the sign is positive outside the object domain bounded by $\mathcal{G}$ (outside the body) and negative inside. Furthermore, $\nabla d(\boldsymbol{x})$ is the outward pointing normal on $\mathcal{G}_{t_2}$ and $|\nabla d| = 1$. Based on $d(\boldsymbol{x})$, we can define the projection $P(\boldsymbol{x}) := \boldsymbol{x} - d(\boldsymbol{x})\nabla d(\boldsymbol{x})$ of a point $\boldsymbol{x}$ in a neighborhood of $\mathcal{G}_{t_2}$ onto the closest point on $\mathcal{G}_{t_2}$. Thus, let us quantify the closeness of a displaced template surface point $\phi(\boldsymbol{x}), \boldsymbol{x} \in \mathcal{G}_{t_1}$ to the reference $\mathcal{G}_{t_2}$ using

$$|P(\phi(\boldsymbol{x})) - \phi(\boldsymbol{x})| = |d(\phi(\boldsymbol{x}))\nabla d(\phi(\boldsymbol{x}))| = |d(\phi(\boldsymbol{x}))| \tag{4}$$

as a pointwise measure. Here, the deformation ϕ is represented by a displacement $\boldsymbol{u}_g = (u_g, v_g, w_g)^T$ defined on Ω, with $\phi(\boldsymbol{x}(\boldsymbol{\zeta})) = \boldsymbol{x}(\boldsymbol{\zeta}) + \boldsymbol{u}_g(\boldsymbol{\zeta})$, minimizing:

$$\mathcal{E}[\boldsymbol{u}_g] = \int_{\Omega} \Big(d\big(\boldsymbol{x}(\boldsymbol{\zeta}) + \boldsymbol{u}_g(\boldsymbol{\zeta})\big)^2 + \alpha_g \|D\boldsymbol{u}_g(\boldsymbol{\zeta})\|_F^2 \Big)\, d\boldsymbol{\zeta}\,, \tag{5}$$

where $D\boldsymbol{u}_g$ denotes the Jacobian matrix of $\boldsymbol{u}_g$ and α_g the regularization weight. For numerical minimization, we considered a conjugate gradient scheme again.

3 Experiments and Results

For experimental evaluation of the proposed method, we have acquired RI data from Microsoft Kinect (640×480 px, 30 Hz) for four healthy subjects S_1-S_4. Reclined on a treatment table, the subjects were asked to perform normal and

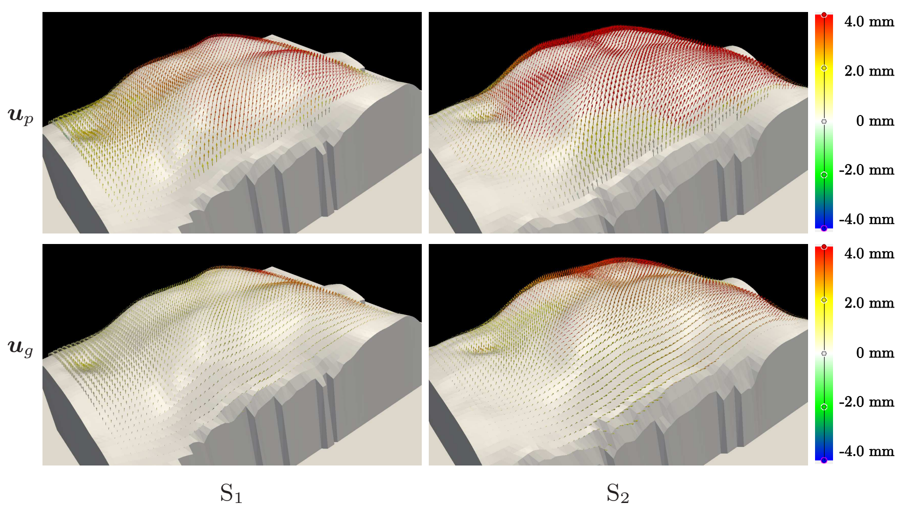


Fig. 2. Glyph visualization of the estimated 3-D surface motion fields $\boldsymbol{u}_p$ (upper row) and $\boldsymbol{u}_g$ (lower row), for subjects S_1 and S_2. The color of the displacement vectors encodes its magnitude in superior-inferior (SI) direction, according to the color bar.

Table 1. Photometry-driven vs. geometry-driven estimation of the 3-D surface motion field, for normal and deep thoracic breathing, for four subjects S_1-S_4. Given is the RMS distance before (ϵ_0) and after warping (ϵ_p, ϵ_g).

	normal breathing					deep inhale				
	S_1	S_2	S_3	S_4	**Mean**	S_1	S_2	S_3	S_4	**Mean**
ϵ_0	0.068	0.078	0.112	0.067	**0.082**	0.098	0.113	0.127	0.104	**0.110**
ϵ_p	0.051	0.054	0.081	0.047	**0.058**	0.061	0.073	0.058	0.060	**0.063**
ϵ_g	0.056	0.058	0.084	0.051	**0.062**	0.079	0.086	0.085	0.076	**0.081**

deep thoracic breathing, respectively. Prior to registration, the range data were preprocessed using edge-preserving denoising. The dataset is available from the authors for non-commercial research purposes.

Qualitative results of the photometry- and geometry-driven 3-D surface motion fields between the respiration states of fully inhale and exhale are illustrated in Fig. 2, using a suitable set of model parameters ($\alpha_p = 0.015, \alpha_g = 10^{-6}$). It can be observed that the photometry-driven surface motion field $\boldsymbol{u}_p$ in superior-inferior (SI) direction is more pronounced than the geometric variant $\boldsymbol{u}_g$. Clinical studies have shown that the SI direction is the prominent direction of human breathing [1]. Thus, let us interpret the results as an indication that even though both motion fields $\boldsymbol{u}_p$ and $\boldsymbol{u}_g$ are meaningful and valuable for application in tumor position correlation, the photometric variant is potentially a better choice for estimating the actual surface motion field regarding the human respiratory system.

Using real data for our experiments, the ground truth 3-D surface motion field is unknown. Hence, for quantitative evaluation, we projected $\boldsymbol{u}_g$ onto the 2-D sensor domain, transferred the deformation from metric real-world to pixel space and applied the resulting displacement $\tilde{\boldsymbol{u}}_g$ to the 2-D photometric data f_{t_1} at time t_1 (fully exhale). For evaluation, we then compared the warped images to the known reference photometric data f_{t_2} at time t_2 (fully inhale) over the patient's torso given by a mask $\mathcal{M}$. In particular, as a scalar distance measure, we computed the root mean square (RMS) photometric distance of the initial and warped data w.r.t. the reference f_{t_2}, respectively:

$$\epsilon_0 = \sqrt{\frac{1}{|\mathcal{M}|} \sum_{\boldsymbol{\zeta} \in \mathcal{M}} \| f_{t_1}(\boldsymbol{\zeta}) - f_{t_2}(\boldsymbol{\zeta}) \|_2^2}\,, \tag{6}$$

$$\epsilon_{p(g)} = \sqrt{\frac{1}{|\mathcal{M}|} \sum_{\boldsymbol{\zeta} \in \mathcal{M}} \| f_{t_1}\big(\boldsymbol{\zeta} + (\tilde{u}_{p(g)}(\boldsymbol{\zeta}), \tilde{v}_{p(g)}(\boldsymbol{\zeta}))^T\big) - f_{t_2}(\boldsymbol{\zeta}) \|_2^2}\,, \tag{7}$$

where ϵ_0 denotes the initial mismatch, $\|\cdot\|_2$ the Euclidean norm. The results on Microsoft Kinect RI data for thoracic respiration of the four subjects is given in Table 1. Note that the individual channels of f (RGB) lie in the range of $[0, 1]$. For normal and deep thoracic breathing, our photometric approach outperformed the geometric variant by $(\epsilon_g - \epsilon_p)/\epsilon_g = 6.5\%$ and $(\epsilon_g - \epsilon_p)/\epsilon_g = 22.5\%$ in average.

This underlines the observation that our photometry-driven registration provides a surface motion field that better resembles the actual torso deformation.

4 Discussion

We have presented a method for photometric reconstruction of a dense 3-D surface motion field over non-rigidly moving surfaces using RI sensors. In an experimental study for the application in RT motion management, we have investigated the performance of our photometry-driven method compared to a geometry-driven approach. Both rely on a variational formulation and are capable of providing dense surface motion fields for application in respiratory motion management. However, our results indicate that incorporating photometric information into the estimation of the torso deformation provides a more realistic surface motion field regarding the human respiratory system. Ongoing work investigates the fusion of both photometric and geometric registration within a joint framework.

Acknowledgments

S. Bauer and J. Wasza gratefully acknowledge the support by the European Regional Development Fund (ERDF) and the Bayerisches Staatsministerium für Wirtschaft, Infrastruktur, Verkehr und Technologie (StMWIVT), in the context of the R&D program IuK Bayern under Grant No. IUK338.

References

1. Keall PJ, Mageras GS, Balter JM, et al. The management of respiratory motion in radiation oncology, report of AAPM TG 76. Med Phys. 2006;33(10):3874–3900.
2. Fayad H, Pan T, Roux C, et al. A patient specific respiratory model based on 4D CT data and a time of flight camera. In: Proc IEEE NSS/MIC; 2009. p. 2594–2598.
3. Hoogeman M, Prévost JB, Nuyttens J, et al. Clinical accuracy of the respiratory tumor tracking system of the cyberknife: assessment by analysis of log files. Int J Radiat Oncol Biol Phys. 2009;74(1):297–303.
4. McClelland JR, Blackall JM, Tarte S, et al. A continuous 4D motion model from multiple respiratory cycles for use in lung radiotherapy. Med Phys. 2006;33(9):3348–3358.
5. Bruhn A, Weickert J, Schnörr C. Lucas/Kanade meets Horn/Schunck: combining local and global optic flow methods. Int J Computer Vis. 2005;61:211–231.
6. Horn BKP, Schunck BG. Determining optical flow. Artiff Intell. 1981;17(1-3):185–203.
7. Lucas BD, Kanade T. An iterative image registration technique with an application to stereo vision. In: Proc IJCAI; 1981. p. 674–679.
8. Liu C. Beyond Pixels: Exploring New Representations and Applications for Motion Analysis. Doctoral Thesis. MIT; 2009.
9. Bauer S, Berkels B, Hornegger J, Rumpf M. Joint ToF image denoising and registration with a CT surface in radiation therapy. In: Proc SSVM. vol. 6667 of LNCS. Springer; 2011. p. 98–109.