Robust Real-Time 3D Respiratory Motion Detection Using Time-of-Flight Cameras

Jochen Penne
Institute of Pattern Recognition, Friedrich-Alexander-University Erlangen, Nuremberg
Martensstr.3
91058 Erlangen, Germany Jochen.Penne@informatik.uni-erlangen.de

Christian Schaller
Institute of Pattern Recognition, Friedrich Alexander University Erlangen, Nuremberg
Martensstr.3
91058 Erlangen, Germany Christian.Schaller@informatik.uni-erlangen.de

Joachim Hornegger
Institute of Pattern Recognition, Friedrich Alexander University Erlangen, Nuremberg
Martensstr.3
91058 Erlangen, Germany Joachim.Hornegger@informatik.uni-erlangen.de

Abstract

Artifacts caused by respiratory motion of the patient during data acquisition in the field of emission or computed tomography require respiratory gating in order to track and correct these artifacts. In this paper, we present a system that uses Time-of-Flight (ToF) technology to compute a three-dimensional estimate of the respiratory motion of a patient. The work is characterized by three key contributions. The first is the employment of ToF sensors. Using ToF sensors it is feasible to acquire a 3D surface model of the chest and abdomen of the patient at frame-rates greater than 15 Hz. The second contribution is an algorithm to derive a surface representation which enables the estimation of the 3D respiratory motion of the patient. The proposed data-driven algorithm models the chest and abdomen three-dimensionally by fitting distinct planes to different regions of the torso of the patient. The third contribution is the possibility to derive a sub-millimeter accurate signal by observing the displacement of each plane. Our ToF modeling approach enables marker less, real-time, 3D tracking of patient respiratory motion with sub-millimeter accuracy.

1. Introduction

Physiological motion in emission or computed tomography leads to a reduction of overall image contrast and a loss of sensitivity. A reduction in lesion detection, loss of accuracy in functional volume determination and more difficult activity concentration recovery are the consequences of the associated blurring. Widely accepted approaches utilize respiratory gating: Only data acquired during a certain respiration state is used for the computation of 3D reconstructions. Upto eight respiration states have proven to deliver a good trade-off between temporal resolution and noise in the 3D reconstructions [11].
Furthermore, adaptive radiotherapy of lung cancer requires a breathing signal to enable the computation of 4D-CT models. Based on these data radiation delivery can be tracked or gated.

2. State of the Art

Current methods to acquire or compute a breathing signal are:

1. **Spirometry**: measurement of the amount (volume) and/or speed (flow) of air that can be inhaled and exhaled.

2. **Markers** are placed on the skin of the patient. The position of the markers is tracked [8, 3].

3. **Acquisition of stereoscopic images** of torso and computation of 3D torso surface points. Derivation of volumetric information from the surface [10].

4. **Hybrid techniques** [7].

Figure 1. Left: ToF camera MESA Imaging GmbH. Right: ToF camera PMDTec GmbH.

It is worth noting that method 3 is the only published method which is contactless and which enables the distinction between abdominal and thoracic breathing at a framerate of 5 Hz. The method is equivalent to spirometry and provides drift-free volume information, which is not the case for spirometry.

We propose an alternate method which is also contact-less and therefore more closely related to method 3 than to methods 1 or 2. We suggest a system based on the emerging ToF (Time-of-Flight) technology. ToF sensors provide a direct way for acquiring 3D surface information of objects with a single all-solid-state ToF camera by measuring the time of flight of an actively emitted optical reference signal in the infrared spectral range [2]. More recently, applications like obstacle detection [9], gesture recognition [5][6] and automotive passenger classification [4] are using ToF sensors. Currently, ToF cameras are also on their way to become a component of consumer electronics. Therefore, a decrease of production costs for ToF sensors due to mass production can be expected in the near future. As ToF sensors provide data at rates higher than 15 Hz, they are suitable for real-time imaging. Examples of available ToF cameras are shown in Figure 1. The 3D data available of a scene observed with a ToF camera is a 3D point cloud. As each 3D point corresponds to a pixel of the sensor matrix a triangulation of the 3D point can be derived trivially. The distance estimation of a point is accomplished by measuring the time of flight of an optical reference signal emitted by the camera and reflected by the scene. Besides the 3D information for each pixel an intensity value corresponding to the reflected amount of the reference signal is available. These intensity values are normally encoded as grey values and can be used to provide a texture for the 3D surface reconstruction. A schematic overview of the ToF principle and an example of the data available from ToF cameras is given in Figure 2.

We investigated the possibilities of extracting a breathing signal from these 3D data and to distinguish abdominal and thoracic breathing based on these information.

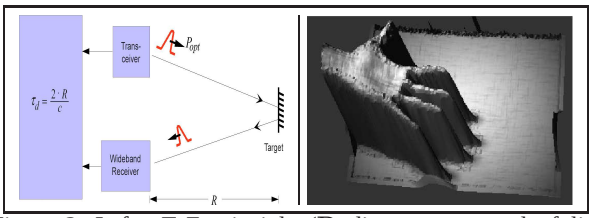

Figure 2. Left: ToF principle (**R**..distance, **c**..speed of light, τ_d..travel time of impulse $\mathbf{P}_{opt}$). Right: 3D closed surface reconstruction of human hand.

3. Methods

Our method requires the ToF camera to be rigidly mounted in a way which brings the full torso of the patient into the field of view of the camera. This implies a distance of approx. 60 cm–100 cm between ToF camera and patient. Figure 3 shows the examples of the acquired 3D reconstructions. The patient is assumed to be lying on an approximately planar table. The main steps of the method are the following:

1. **Calibration**: An image of the empty patient table is acquired. A best-fitting plane is computed for the surface of the patient table. This plane will be used for segmenting the torso of the patient in subsequent processing steps. In Figure 4 this is the plane with the number **1**. We will term this plane **table-plane**.

2. **Segmentation of the torso**: If the patient is lying on the table, his complete torso is segmented by rejecting all points which are are more far away from the ToF camera than the table-plane, i.e. which are behind the table-plane from the viewpoint of the ToF camera.

3. **Defining Regions-of-Interest (ROI)**: Two regions of interest are defined. One for the chest region and one for the abdomen. This step is necessary when setting up the camera the first time in the therapy room or possibly when patients of significantly differing size are treated (like kids and adults).

4. **Derivation of multidimensional breathing signal**: The 3D points of each ROI are processed with a mean filter of kernel size five to reduce the influence of noise and subsequently a best-fitting plane is computed for the processed 3D points of each ROI. Each best-fitting plane is not infinite but limited by a bounding polygon which is determined by the silhouette of the segmented torso. The Euclidean distance of each best-fitting plane to the table-plane constitutes one dimension of the breathing signal. Thus, for two ROIs (chest and abdomen in this case) a two-dimensional breathing signal is derived.

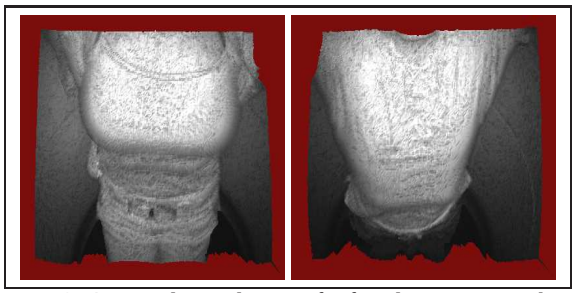

Figure 3. Acquired ToF-data. Left: female person. Right: male person.

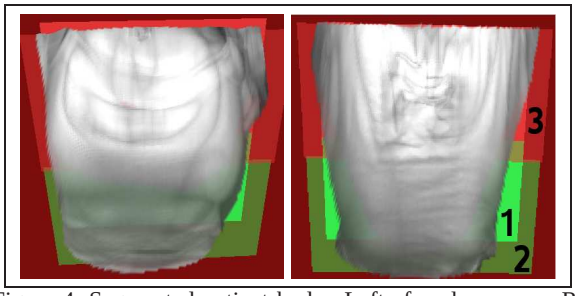

Figure 4. Segmented patient body. Left: female person. Right: male person. The colored planes are the clipping planes computed after segmenting the torso.

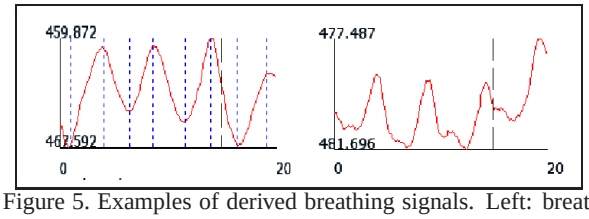

Figure 5. Examples of derived breathing signals. Left: breathing signal for the breast. Right: breathing signal for the chest. The vertical axis represents the time in seconds. The horizontal axis the distance in mm between table-plane and best-fitting plane for the corresponding ROI (chest, abdomen).

4. Evaluation

Available ToF cameras have a z-resolution of upto 1 mm. This accuracy is only reached under optimal lightning and illumination conditions. Our approach derives the breathing signal from best-fitting planes which are computed from probably noisy 3D points. Nevertheless, the usage of best-fitting planes increases the stability. To validate this fact a rigid, non-deformable model of the human torso was subject to the whole processing chain described in section 3. Thus, the breathing signal of a completely non-breathing patient was computed for varying viewing angles of the camera (from perpendicular viewing angle onto the patient table upto 30° viewing angle). The still observed breathing motion indicates the achievable z-resolution of the breathing signal. Considering a time-span of 20 seconds the standard deviation of the distance of best-fitting plane for the torso and the table-plane was computed. The computed values were always smaller than 0.1 mm, i.e. the z-resolution of the breathing signal is about a factor ten greater than the resolution of the original ToF camera data and respiratory motion can be detected in the sub-millimeter range.

To validate the correctness of the computed breathing signal we computed for 13 patients the correlation coefficient of the computed breathing signal with the breathing signal delivered by an ANZAI belt AZ-733V[1]. The patient was advised to breath with the chest when the ANZAI belt was attached to its chest and advised to breath with the abdomen when the ANZAI belt was attached to its abdomen. For the experiments a ToF camera SwissRanger SR3100 was used. The results are displayed in Figure 6. The smallest computed correlation value was 0.7.

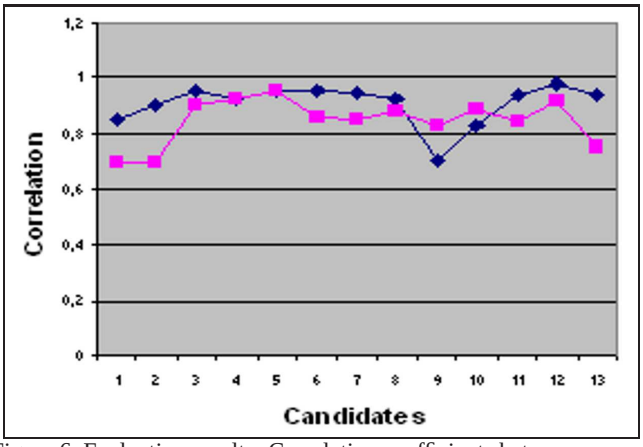

Figure 6. Evaluation results: Correlation coefficients between respiration signal computed from ToF camera data and respiration signal delivered by ANZAI belt.

5. Conclusion

We presented a non-invasive non-contact method for acquiring respiratory signals based on 3D point cloud surfaces which are acquired with a ToF camera. We have validated that the signal can be decomposed into abdominal and thoracic components and is significantly correlating with the reference breathing signal acquired with an ANZAI belt. The computational time for computing the respiratory signals does take approx. 25 ms on a 2.0GHz single core CPU. Thus, the proposed method is real-time capable. By deriving the respiratory signal of a certain anatomical region like chest or abdomen from best-fitting planes and not the original 3D point clouds a z-resolution of 0.1 mm is achieved.

References

[1] Anzai Medical Co.,Ltd. www.anzai-med.co.jp, 2007.

[2] B. Büttgen, T. Oggier, M. Lehmann, R. Kaufmann, and F. Lustenberger. CCD/CMOS Lock-In Pixel for Range Imag-

ing: Challenges, Limitations and State-of-the-Art. In *1st Range Imaging Day*, Zürich, Switzerland, June 2005.

[3] P.-C. M. Chi, P. Balter, D. Luo, R. Mohan, and T. Pan. Relation of external surface to internal tumor motion studied with cine CT. *Medical Physics*, 33(9):3116–3123, 2006.

[4] P. Devarakota, M. Castillo-Franco, R. Ginhoux, B. Mirbach, and B. Ottersten. Occupant Classification Using Range Images. *IEEE Transactions On Vehicular Technology*, 56(4):1983–1993, July 2007.

[5] M. Holte, T. Moeslund, and P. Fihl. View Invariant Gesture Recognition using the CSEM SwissRanges SR-2 Camera . In *Proceedings of the Workshop Dynamic 3D Imaging in conjunction with DAGM'07* , pages 53 – 60. DAGM e.V., 2007. Heidelberg, Germany.

[6] E. Kollorz and J. Hornegger. Gesture recognition with a time-of-flight camera . In *Proceedings of the Workshop Dynamic 3D Imaging in conjunction with DAGM'07* , pages 86 – 93. DAGM e.V., 2007. Heidelberg, Germany.

[7] D. A. Low, P. J. Parikh, W. Lu, J. F. Dempsey, S. H. Wahab, J. P. Hubenschmidt, M. M. Nystrom, M. Handoko, and J. D. Bradley. Novel breathing motion model for radiotherapy. *Int J Radiat Oncol Biol Phys*, 63(3):921–929, November 2005.

[8] E. Rietzel and G. T. Chen. Improving retrospective sorting of 4D computed tomography data. *Medical Physics*, 33(2):377–379, 2006.

[9] T. Schamm, S. Vacek, J. Schröder, M. Zöllner, and R. Dillmann. Obstacle detection with a PMD-camera in autonomous vehicles . In *Proceedings of the Workshop Dynamic 3D Imaging in conjunction with DAGM'07* , pages 70 – 77. DAGM e.V., 2007. Heidelberg, Germany.

[10] S. Tarte, J. McClelland, S. Hughes, J. Blackall, D. Landau, and D. Hawkes. A Non-Contact Method for the Acquisition of Breathing Signals that Enable Distinction Between Abdominal and Thoracic Breathing. *Radiotherapy and Oncology*, 81(1):209, October 2006.

[11] J. Wang, J. Byrne, J. Franquiz, and A. McGoron. Evaluation of amplitude-based sorting algorithm to reduce lung tumor blurring in PET images using 4D NCAT phantom. *Comput. Methods Prog. Biomed.*, 87(2):112–122, August 2007.