Geometric calibration between PET scanner and structured light scanner

Kjer, Hans Martin; Olesen, Oline Vinter; Paulsen, Rasmus Reinhold; Højgaard, Liselotte; Roed, Bjarne; Larsen, Rasmus

Published in:
Proceedings of the MICCAI workshop on Mesh Processing in Medical Image Analysis (MeshMed)

Publication date:
2011

Geometric calibration between PET scanner and structured light scanner

Martin Kjer[1], Oline V. Olesen[123], Rasmus R. Paulsen[1], Liselotte Højgaard[2], Bjarne Roed[3], and Rasmus Larsen[1]

[1] Informatics and Mathematical Modelling, Technical University of Denmark
Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, Denmark
http://imm.dtu.dk/
[2] Department of Clinical Physiology, Nuclear Medicine & PET, Rigshospitalet,
Copenhagen University Hospital, University of Copenhagen
[3] Siemens Healthcare, Siemens A/S, Denmark

Abstract. Head movements degrade the image quality of high resolution Positron Emission Tomography (PET) brain studies through blurring and artifacts. Manny image reconstruction methods allows for motion correction if the head position is tracked continuously during the study.
Our method for motion tracking is a structured light scanner placed just above the patient tunnel on the High Resolution Research Tomograph (HRRT, Siemens). It continuously registers point clouds of a part of the patient's face. The relative motion is estimated as the rigid transformation between frames.
A geometric calibration between the HRRT scanner and the tracking system is needed in order to reposition the PET listmode data or image frames in the HRRT scanner coordinate system. This paper presents a method where obtained transmission scan data is segmented in order to create a point cloud of the patient's head. The point clouds from both systems can then be aligned to each other using the Iterative Closest Point (ICP) algorithm.

Keywords: HRRT, PET, structured light, calibration, motion tracking, motion correction

I. Introduction

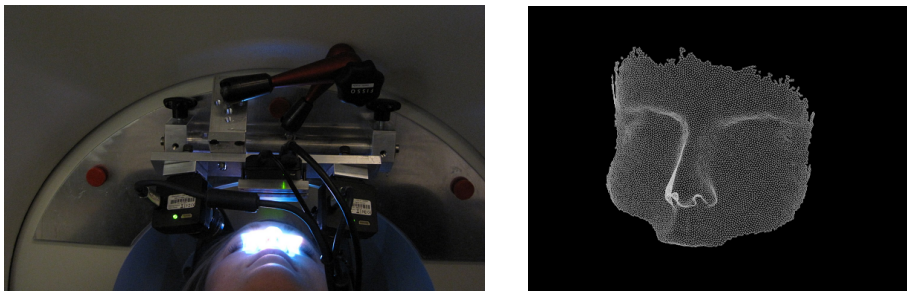
Technological improvement of the different medical imaging modalities leads to diagnostic images with increasing spatial resolution. As a consequence the techniques also become more vulnerable to the effects of patient motion during image acquisition.

In Positron Emmision Tomography (PET) patient movements can cause both artifacts and blurred images [1]. The increased spatial resolution gained by technological advancement is thus countered to a certain degree by the increased sensitivity to motion, unless patient fixation and motion correction is utilised.

Even with fixation methods such as vacuum cushions and head restaints motion still occurs albeit to a lesser degree [2]. The magnitude of motion tends to increase with the duration of the study, and acquisition times for PET images can be up to several hours. Typically the patient's head drifts slowly to one side, or at some point the patient repositions themself to lie more comfortably. The resolution of the Siemens High Resolution Research Tomograph (HRRT) is below 2 mm, and since the described movements can be larger motion correction becomes a necessity [3].

One approach for motion tracking is the Polaris Vicra system from *Northern Digital Incorporated*. It registers a tool with reflective markers attached to the patient's head. The position of the markers are relayed to a tool tracker through infrared light. The main issue with such system is to ensure that the tool stays attached and do not move relative to the patient's head. Further to maintain line of sight between tool and tool tracker, which is troublesome in the narrow patient tunnel of the scanner.

Our approach is a structured light system. Two cameras on both sides of a DLP pico projector from *Texas Instruments* are mounted on the HRRT scanner as shown in Figure 1(a). A series of cosine patterns are projected onto the object, and these patterns are imaged by the cameras. We use phase shift interferometry to obtain a 3D point cloud of the object - in this case a part of the patient's face as shown in Figure 1(b) [4]. The relative motion between image frames is estimated as the rigid transformation with six degrees of freedom that best aligns the two point clouds. The iterative closest point (ICP) algorithm can be used to find the transformation [11] [12]. In comparison to the tool tracking approach, this approach avoids the use of an optical tool, and can potentially be integrated into future scanners.



(a) The structured light scanner (b) Point cloud output

Fig. 1. The structured light scanner mounted on the HRRT scanner and an example of the 3D point cloud it produces.

The issue adressed in this paper is the geometric calibration between the HRRT scanner and the structured light system. Movements observed by the motion tracking system has to be translated into movements in the coordinate

system of the HRRT scanner. Then it is possible to reposition all the detected Line Of Responses (LOR) into the position of the head at the given time.

The method for calibration should not add any extra radiation dose to the patient. Furthermore it is undesirable to increase the total duration of the scanning session and to alter the normal workflow. Ideally the method should only employ the data that is already obtained - either the emission (EM) data or the transmission (TX) data.

Previous calibration methods use either a number of EM scans or TX scans of a calibration object and have both successfully been used for finding a calibration transformation [7] [8]. In both cases the motion tracking system was the Polaris Vicra system or a system very similar to it. In the EM approach a positron emitting point source added to the tracking tool allowed for measurements of the tracking tool position in both systems. Multiple independent measurements were required in order to determine all six degrees of freedom. In the TX scan approach retroreflective markers with a sufficient density allowed for detection of the tracking tool in both systems. Both of these methods find relatively few points with a known point-to-point correspondance from which a transformation can be estimated. The measurements must be performed in preparation of the patient scan, and the calibration is preserved as long as the tooltracker is not moved.

For the purpose of calibration with a structured light motion tracking system there is no markers to be detected with either EM scans or TX scans. From the TX data it is however possible to extract the iso-surface of the patient's head, thus producing a point cloud similar to what is obtained from the motion tracking system. The point correspondance is not known, however a large amount of points are available. The best rigid alignment between the two point clouds can be found with the ICP algorithm, and the transformation serves as the calibration. The measurements are a part of the normal scanning procedure, which is very advantageous in terms of time and simplicity. This also allows for adjustments of the motion tracking system or even completely detaching it from the HRRT scanner between scans.

A common approach for extracting iso-surfaces from volume data is the Marching Cubes Algorithm [5]. However, in our case, the iso-levels of the TX scans are not very well defined and would result in a noisy surface. We have therefore investigated an alternative approach to extract the interface between soft-tissue and air from the TX scans. This paper presents a segmentation method using path tracing on the reconstructed TX image.

II. Methods

A. Circular resampling

A typical slice from the TX image of a patient is seen in Figure 2(a). The border of the head has to be traced and the procedure repeated for each slice. The used path tracing algorithm finds the optimal path going from one edge of the image

to the other. However the boundary of the head in the TX image is located as a circular structure somewhere in the middle, and thus a reshaping is required before path tracing can be applied.

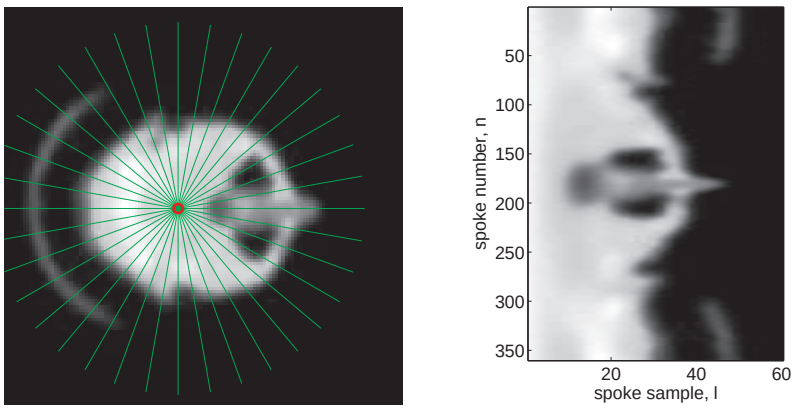
A point within the circular structure is chosen. The point (r_s, c_s) serves as a center from which N spokes of length L radially shoots out from, so that the end points are given as:

$$r_e = r_s + \cos\left(\frac{2\pi n}{N}\right) L \ , \ c_e = c_s + \sin\left(\frac{2\pi n}{N}\right) L$$

where $n = [0, 1, .., N-1]$. Each spoke sample L values of the underlying pixels using bi-linear interpolation. The sampled values are placed in a new image, as shown in Figure 2(b).

The center point is chosen as the centroid of the image. Using a fixed point could however be a viable strategy. Patients are always placed in the approximate radial center of the scanner tunnel, since this is where the spatial resolution is highest [6]. The placement is done manually so a slight inter-scan variation is expected. Using a fixed point saves a little computational time, however it is more likely to give a faulty resampling, if the spoke length is chosen too short.

The radial resolution - the amount of spokes - is also a consideration. The choice is dependent on the resolution of the TX image. Too few spokes corresponds to an undersampling and leads to loss of information about the true curvature and important small features primarily the nose and ears. Oversampling increases computational time without much extra information being gained.



(a) TX scan (b) TX scan resampled

Fig. 2. Circular resampling of a transversal TX slice using 360 spokes. Only every 10th spoke is displayed on Figure 2(a).

B. Path tracing

The chosen path tracing algorithm is a simplified version of Dijkstra's algorithm [9]. It is based on dynamic programming and designed to find optimal paths between the top and bottom of a grayscale image $\mathbf{I}(r, c)$ asillustrated in Figure 3.

Each pixel holds a cost value $C(r, c)$, and for the purpose of edge detection the first derivative or gradient of the image is used. The optimal path $\mathbb{P}$ is then defined as the list of pixels with the lowest (or highest) accumulated cost going from the top of the image to the bottom:

$$C_{tot} = \sum_{(r,c)\in\mathbb{P}} C(r, c)$$

It is possible to calculate the total cost for all paths, however the number of computations would quickly increase beyond reasonable with increasing image size. The algorithm is therefore limited, so that the path is only allowed to move down and up to two pixels to either side.

The algorithm operates with two images of the same dimension as the original image. A value in the accumulator image $\mathbf{A}(r, c)$ is the lowest possible (optimal) accumulated cost required to get to that pixel from the top. Path information is stored in the backtracing matrix $\mathbf{T}$. The value stored in $\mathbf{T}(r, c)$ is column index of the previous path entry, $\mathbb{P}\,(r-1, \mathbf{T}(r, c))$. The path is therefore read backwards, and the last path entry is the index that holds the lowest value in the last row of $\mathbf{A}$. This is illustrated in Figure 3.

An important aspect of the algorithm is the ability to wrap around small features. This is highly dependent on the restrictions imposed on the path. Allowing it to sidestep more pixels allows for more sharp features to be traced, although it could also lead to a much less smooth border. In this particular case an alternate possibility is to increase the radial resolution of the resampling, at the expense of more computations.

The traced edge points from each slice are transformed to points in the original image and combined to the resulting point cloud which is shown in Figure 4.

C. Transformations and ICP

The calibration between the motion tracking system and the HRRT scanner can be taken as the rigid transformation of the source point set $\mathbb{P} = (p_1, p_2, ..., p_N)$ that gives the best alignment to the target set $\mathbb{Q} = (q_1, q_2, ..., q_M)$. The mathematical measure of goodness of fit is the sum of squared errors after the transformation - which is the squared distance from points in $\mathbb{P}$ to their corresponding point in $\mathbb{Q}$:

$$E = \sum_{i=1}^{N} ||\mathbf{R}p_i + \boldsymbol{T} - q_i||^2 \tag{1}$$

where $\mathbf{R}$ is an ortogonal rotation matrix and $\boldsymbol{T}$ a translation vector.

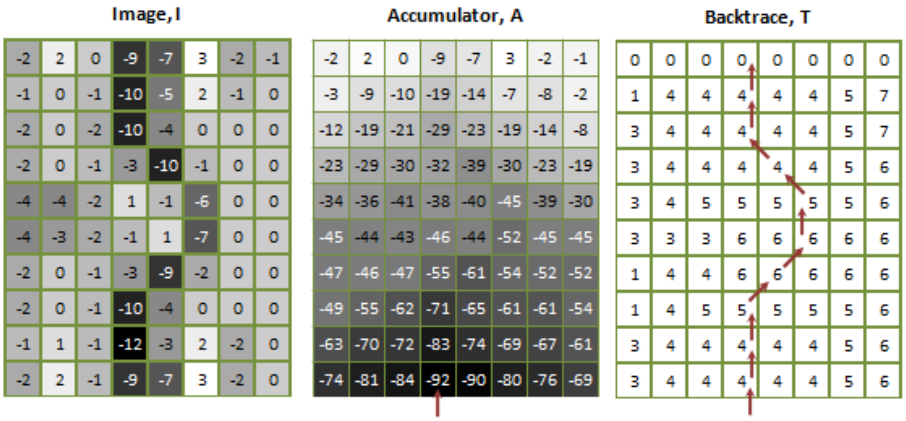


Fig. 3. Illustration of pathtracing on the resized version of Figure 2(b).

Minimization of Equation 1 requires knowledge of the point correspondance between the two point sets. When such information is not given, the problem can be solved with an iterative approach - the ICP. The algorithm iterates through the following steps:

- **Matching:** Points in $\mathbb{P}$ are matched with their nearest neighbor in $\mathbb{Q}$, and this is assumed to be the point correspondance.
- **Minimization:** Equation 1 is minimized using Singular Value Decomposition (SVD).
- **Transformation:** The transformation is applied to the points in $\mathbb{P}$, and the steps are repeated.

The algorithm can be modified further and improved by adding more steps, such as a selection of only a subset of the points or inclusion of a neighbor pair weighting [10].

When using the ICP there is a consideration of the designation of target and source point cloud. The point cloud from the TX image represents the entire head whereas the points from the motion tracking system only represents a subset of this. Many of the points from the TX image does therefore not have a meaningfull nearest neighbor correspondance, and consequently the TX point cloud is chosen as the target. Otherwise the ICP algorithm requires some kind of rejection scheme.

III. Results and Discussion

A. Segmentation

Transmission data from four patient studies was available and two of them were motion tracked with the structured light system. The settings and parameters

for segmentation of the TX data is based on the two untracked TX studies. The result is illustrated in Figure 4, and while it is recognisable as a face, it does have the issues to be addressed: Which slices to segment and edge tracing.

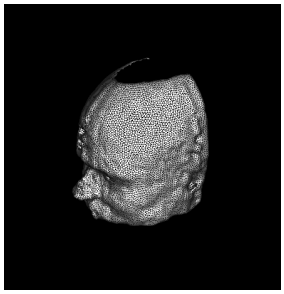


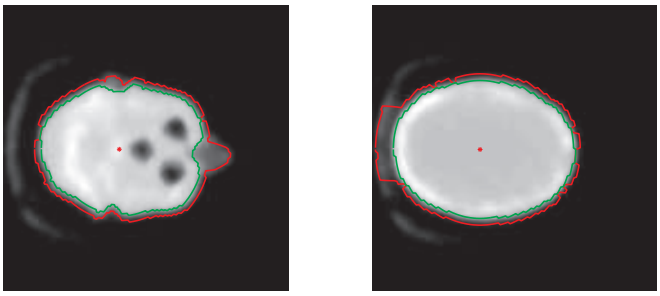
Fig. 4. The output from the TX segmentation. The point cloud was surface reconstructed [13] for the purpose of illustration. Notice that it is possible to see the start of the ear canal, however the ear is missing. Small indentations and bumps are present and the very top and the back-side of the head has been excluded.

Which slices to segment: The data contains 207 transversal slices, however the initial slices contains nothing except noise and attenuation from the headrest. The strategy with circular resampling and path tracing assumes that the image contains a circular structure, and otherwise the result is highly unpredictable. A simple threshold strategy was chosen as sanity check. The upper part of the skull is excluded, as seen in Figure 4. The loss is acceptable, since the excluded part is neither seen by the motion tracking system. More slices could be included with a more sophisticated sanity check that use connectivity and shape for instance.

The edge tracing: The cost values used to determine the optimal path is based on the derivative of the γ-ray attenuation. For the most part of the head there is a superficial thin layer of skin followed by the bone layer. Since bone is highly attenuating compared to soft tissues, there is a gradient going from air to skin, but an even steeper gradient going from skin to bone. The method thus favors segmenting the bone edges, which for the most part of the face is very similar to the skin surface. However it is paramount that the method segments important features - especially the nose. The upper part of the nose is quite dense and thus segmented, whereas the lower soft part is not. The resulting faulty segmentation is illustrated in Figure 5(a), and therefore it was neccesary to include a weighting of the cost values to improve the segmentation.

A strategy was chosen where all cost values outside the bone edge region were enhanced. This generally improved the skin and nose segmentation as shown in Figure 5(a). However the lower most part of the nose and the ears were still unsegmented. Further the enhancement introduced some artifacts. The gradient from the headrest attenuation was also amplified, resulting in partly segmentation of this as shown in Figure 5(b). The issue was most profound in slices

where head and headrest was in direct contact. Since the back of the head has no real interest it could be cut-off. Notice also in Figure 5(b) that even though the enhancement improved the tracing of the skin-border, the path occasionally went back to the bone-border resulting in bumps and indentations in the point cloud.



(a) TX image slice 169 (b) TX image slice 119

Fig. 5. Illustration of the effect of the added weighting scheme for segmentation. The green and the red lines show the result before and after application of the cost value weighting respectively.

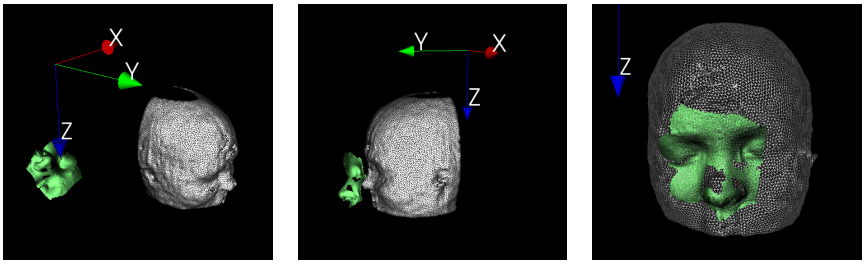
B. Calibration

With a point cloud measured from the HRRT scanner and one from the motion tracking system the calibration could be performed. The initial situation is seen in Figure 6(a). A fixed transformation was applied in order to bring the two point clouds into a decent starting position. This was chosen as a rotation of 180° around the y-axis, a rotation of −36° around the x-axis and a (x,y,z)-translation of 160 mm, 385 mm and 340 mm. The resulting situation before the application of the ICP is shown in Figure 6(b). The ICP converged to a minimum with less than 20 iterations, and the final total transformation is shown in Figure 6(c).

Visually it appears to be a decent alignment, however a more quantiative measure is required for validation of the result. One approach could be to use landmarks, however the error would then be correlated with the ability to place landmarks correctly. A more precise and valuable validation would be a comparison between PET images with and without motion correction, however the motion correction of PET images are beyond the scope of this paper.

C. Sources of error

The segmentation method assumes that the two point clouds are measured simultaneously. However the cameras' image capture is instantaneous compared



(a) Initial situation (b) Fixed transformation (c) Final transformation

Fig. 6. Illustration of the transformations. The point cloud from the TX scan is shown in white and has been surface reconstructed [13] for illustration purpose. The point cloud from the motion tracking system is shown in green.

to the time that the TX scan requires. The assumption is that the patient is motionless in this period. This is not likely completely true and could be a potential source of error.

Arguably a minor error for cooperative subjects, and the assumption must be generally accepted - since the TX data is deemed good enough to be used in the emission image reconstruction.

A great perspective for motion tracking is however to use it with patients that are less cooperative - children and patients suffering from disorders that affect motor-function. In that case the error in the TX data increases, and the chosen segmentation method for calibration would perform worse.

IV. Conclusion

A method for geometric calibration between the HRRT scanner and the structured light motion tracking system has been presented. The method exploits the data from the transmission scan, and thus it does not alter the scanning procedure or prolong the study. This is in key with the advantages of using the structured light system.

The presented approach segments the edges of the transmission scan slices using a path tracing algorithm, from which a point cloud of the patient's head is obtained. It was shown that the approach suffers from some difficulties due to the attenuation properties of the different tissues. An enhancement strategy for improved segmentation was proposed, and it was shown to give a better segmentation of the nose at the expense of more problematic segmentations in regions that has no interest. The method is not able to produce perfect point clouds of the entire head, however it can produce decent point clouds of the face, and thus the method seems suitable for the purpose of calibration with the structured light system.

The motion tracking system produces a similar point cloud, and it was shown that the two point clouds can be aligned to each other using the ICP algorithm with a visually good result.

References

1. Anton-Rodriguez, J.M. , Sibomana, M., Walker, M.D., Huisman, M.C., Matthews, J.C., Feldmann, M., Keller, S.H., Asselin, M.: Investigation of Motion Induced Errors in Scatter Correction for the HRRT Brain Scanner. IEEE Nuclear Science Symposium Conference Record (MIC), IEEE, (2010)
2. Green, M.V., Seidel, J., Stein, S.D., Tedder, T.E., Kempner, K.M., Kertzman, C., Zeffiro, T.A.: Head Movement in Normal Subjects During Simulated PET Brain Imaging with and without Head Restraint. Journal of Nuclear Medicine 35(9), pp. 1538–1546 (1994)
3. Olesen, O.V., Sibomana, M., Keller, S.H., Andersen, F., Jensen, J., Holm, S., Svarer, C., Højgaard, L.: Spatial resolution of the HRRT PET scanner using 3D-OSEM PSF reconstruction. Nuclear Science Symposium Conference Record 2009, pp. 3789–3790 (2009)
4. Olesen, O.V., Paulsen, R., Højgaard, L., Roed, B., Larsen, R.: Motion tracking in narrow spaces: a structured light approach. Medical image computing and computer-assisted intervention : MICCAI. International Conference on Medical Image Computing and Computer-Assisted Intervention, 13(Pt 3), pp. 253-60 (2010)
5. Lorensen, W.E., Cline, H.E.: Marching cubes: A high resolution 3D surface construction algorithm. SIGGRAPH Comput. Graph. 21(4), pp. 163-169 (1987)
6. Knöß, C.: PhD Thesis: Evaluation and Optimization of the High Resolution Research Tomograph (HRRT). RWTH Aachen Univeristy (2004)
7. Fulton, R.R., Meikle, S.R., Eberl, S., Pfeiffer, J., Constable, C.J. and Fulham, M.J.: Correction for head movements in positron emission tomography using an optical motion-tracking system. IEEE Transactions on Nuclear Science 49(1), pp. 116–123 (2002)
8. Bühler, P., Just, U., Will, E., Kotzerke, J. and Hoff, J.: An accurate method for correction of head movement in PET. IEEE Transactions on Medical Imaging 23(8), pp. 1176-1185 (2004)
9. Paulsen, R.R. and Moeslund, T.B.: Introduction to Medical Image Analysis. DTU Informatics (2011)
10. Rusinkiewicz, S. and Levoy, M.: Efficient Variants of the ICP Algorithm. Third International Conference on 3-D Digital Imaging and Modeling, pp. 145 (2001)
11. Chen, Y., Medioni, G.: Object Modeling by Registration of Multiple Range Images. 1991 IEEE International Conference on Robotics and Automation 3, pp. 2724–2729 (1991)
12. Besl, P.J., McKay, N.D.: A Method for Registration of 3-D Shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence 14(2), pp. 239–256 (1992)
13. Paulsen, R.R., Bærentzen, J.A., Larsen, R.: Markov Random Field Surface Reconstruction. IEEE Transactions on Visualization and Computer Graphics, pp. 636–646 (2009)