MR COMPATIBLE OPTICAL MOTION TRACKING
Building an Optical Tracking System for Head Motion Compensation in MRI

Martin Hoßbach
Department Cognitive Computing and Medical Imaging
Fraunhofer Institute for Computer Graphics Research (IGD), Darmstadt, Germany

Keywords: Optical tracking, Head motion tracking, MRI-compatible.

Abstract: Magnetic Resonance Imaging (MRI), in spite of its potential in medical diagnosis, has one major drawback: Image acquisition is a slow process, requiring the patient to not move for several minutes. This renders MRI useless in a number of cases. In the case of MR Imaging of the Head, optical motion tracking can be used for motion compensation, thereby greatly improving image quality. In this paper, an MR-compatible approach of tracking the patient's head is presented which does not require his or her cooperation, based on stereo-optical marker tracking. It is adapted to work in the MRI scanner, does not influence the MR image acquisition and is easily integrated into clinical routine.

1 INTRODUCTION

Quite similar to the way a photograph gets blurred when the person in front of the camera moves, many medical imaging modalities suffer from motion artifacts when the patient is moving during image acquisition. This is normally dealt with by asking the patient not to move or by sedating him or her. In some cases, however, the patient will move, and sedating is impossible. When dealing with medical imaging of the human head (which is the focus of this paper), one solution is to track the patient's head motion, enabling compensation of the head motion for these cases.

Among the imaging modalities suffering from patient head movement, positron emission tomography (PET) is probably easiest to cope with: standard motion tracking equipment can be used since neither spatial nor environmental restrictions (like magnetic fields in MRI) apply. Langner tracked infrared markers mounted to a pair of ski goggles to track the head motion (Langner, 2008). Ma *et al.* reconstructed the patients head motion using a stereo camera system outside a PET scanner using SIFT feature tracking (Ma et al., 2008).

Dold *et al.* proved the concept of MRI head motion compensation by head motion tracking with standard IR tracking of a marker, held by the patient with his or her teeth (Dold et al., 2006). This marker is a major drawback of Dolds approach: Non-cooperative patients will not keep it between their teeth. However, Dold *et al.* showed that image quality can be greatly improved by motion compensation, even in the case of a cooperative patient.

This paper concentrates on head motion tracking for magnetic resonance imaging (MRI) and shows how to build a motion tracking system that is fully MR compatible, requiring no patient cooperation whatsoever. Particularly, the patient does not have to hold a marker between his or her teeth or wear marker-goggles. In section 2 the hardware setup is explained, as well as the difficulties resulting from that choice and how they were solved. In section 3 the system's performance is analyzed. Section 4 concludes the paper.

2 SYSTEM DESIGN

2.1 Hardware

The basic idea is to build an optical tracking system with two cameras that fits into the MR tomographs bore. To achieve MR compatibility, the system and its

Hoßbach M. (2010).
MR COMPATIBLE OPTICAL MOTION TRACKING - Building an Optical Tracking System for Head Motion Compensation in MRI.
In *Proceedings of the International Conference on Computer Vision Theory and Applications*, pages 453-456
DOI: 10.5220/0002830004530456

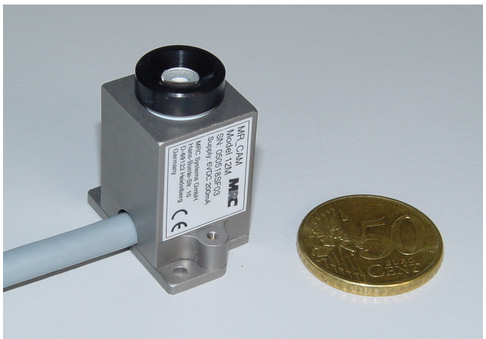

Figure 1: MR compatible camera built by MRC Systems, Heidelberg. Coin as size reference.

Figure 2: Camera mounting on top of Siemens Trio Head Coil.

components have to be designed such that they do not affect the MR image acquisition and that the tracking is not influenced by the MRI scanner. Spatial conditions in the bore and of the head coil need to be taken into consideration.

As a result, MR compatible cameras like those built by MRC Systems (Figure 1) are used, which achieve a decent image quality even under poor lighting conditions. The cameras provide a standard PAL video signal of 768×576 pixels at 25 fps. With a special mounting, two of these cameras are attached to the head coil with the patient's forehead in their field of view.

The camera mounting is designed so that it can be removed from the head coil and later be reattached without changing the position and orientation of the cameras in relation to the head coil, thus preserving camera and stereo camera system calibration (Figure 2).

The two cameras are then calibrated using a checkerboard calibration pattern and the stereo camera calibration functions found in the OpenCV library[1].

Figure 3: SURF-Features in an image acquired by the MR compatible cameras described in section 2.1. SURF-Features are found mainly in areas that can move independent from the skull, thus rendering them useless for head motion tracking.

2.2 Tracking Approach

Motion tracking can be performed using a markerless or a markerbased approach. Although integrating markerless tracking methods into the MRI scanning procedure is much easier, since only little operator intervention is necessary, a markerbased approach has been chosen: Feature tracking methods like SIFT, SURF or the "good features to track" work fine in various environments for the inverse problem of ego motion estimation (Se et al., 2001). However, these features performed poorly with the images provided by the MR compatible cameras mentioned in section 2.1 (Figure 3). Features are found mainly in facial areas that can move independently from the skull, thereby rendering the estimated motion unusable for head motion tracking.

In this scenario, it is essential that only the motion of the patients skull is tracked. For this reason, a different approach using fiducial markers has been chosen. However, a marker like the one used by Dold *et al.* can obviously not be used, because the patient has to hold it with his or her teeth. A non-cooperative patient will not do this, but nonetheless may be moving, thus prohibiting the acquisition of MR images.

Standard marker-based tracking techniques can still be used if markers of a special color are attached to the patients forehead. This color should normally not be found in the human face. Blue circular paper

[1]The OpenCV library: `http://opencv.willowgarage.com`

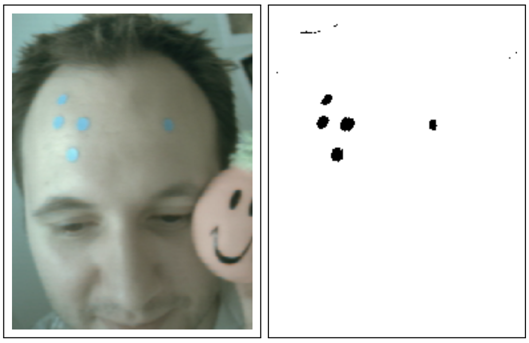

Figure 4: Image of volunteer with blue markers on his forehead (left), and the result of transforming the image to the HSV color model and thresholding the H-channel (right).

stickers, similar to those used in (Ohayon and Rivlin, 2006), turned out to work nicely: After transforming the image to the HSV color model, they can be segmented with simple image processing operations (see Figure 4). Segments are checked for plausibility. Then, the center of gravity is calculated for each segment.

2.3 Stereo Matching

Extracting the feature points (the centers of the blue markers) in each image results in a 2D feature list for each image. Now the corresponding feature points have to be found. Features are matched using two separate approaches, both relying only on the epipolar constraint: Let F be the fundamental matrix of the camera system, such that $x_l^{\mathrm{T}} F x_r = 0$ for one point in space X with image positions x_l and x_r. Let $f_l = \{x_{l,0}, x_{l,1}, x_{l,2}, \ldots, x_{l,n}\}$ be a set of feature points in the left camera image and $f_r = \{x_{r,0}, x_{r,1}, x_{r,2}, \ldots, x_{r,m}\}$ a set of feature points in the right camera image. A matching $M \subset f_l \times f_r$ needs to be found that contains each feature point once or not at all and that has a minimal epipolar error

$$c(M) = \sum_{(l,r)\in M} l^{\mathrm{T}} F r$$

Matching is basically a minimization problem. The Kuhn-Munkres-Algorithm (Munkres, 1957) can be used to accomplish this. It does, however, match all feature points, and therefore also matches points, like misdetected markers, that clearly should not be matched. This can be resolved by removing all matches with an epipolar distance exceeding a certain threshold.

During image acquisition (when the MRI scanner is actually working), mismatches must not occur. Therefore, a second, more conservative and therefore more robust matching approach has been developed. It requires knowledge of the feature points and the matches from the last set of images:

First, from all 2D feature points all points are matched that should obviously be matched because there are no other feature points on the same epipolar plane. Using the Kuhn-Munkres algorithm, the 2D feature points from the current set of images are matched to the 2D feature points from the previous set of images. Then we can tell if a certain pair of 2D feature points in the current set of images was matched in the last set of images. From the remaining 2D feature points, those that were matched in the previous set of images are also matched in the current set of images. All remaining 2D feature points are ignored.

Having matched the 2D feature points, the 3D position of the feature points are reconstructed in relation to the cameras by triangulation. An appropriate method can be found in (Hartley and Zisserman, 2000).

2.4 Model Tracking

The two matching strategies are used in different phases of tracking: The Kuhn-Munkres-Approach is used in preparation, when the operator is still able to change the marker setup, and ask the patient to move his head if mismatches occur. Once the matching looks fine, a click on a button ends preparation phase, and two things happen: First, the matching strategy is changed to the conservative strategy and secondly the current set of 2D feature points is triangulated. The resulting 3D positions are stored as reference for the tracking. From that point on, after matching of the feature points, the 3D positions of the respective markers is calculated.

Now the rigid transformation needs to be found that transforms the initial marker positions into the current ones. This is complicated assuming that both point clouds may be incomplete, noisy and contain outliers. A two-step-approach is used:

1. for each three element subset of the reference point cloud and each three element subset of the current point cloud, calculate the euclidean distances between the three points. If distances are similar enough (below a certain threshold), use these three correspondences to calculate a transformation. How many points of the two point clouds match if this transformation is applied? Find the transformation maximizing this number and therefore minimizing the average displacement.

2. Using the best transformation found in step 1, transform the reference point cloud, and calculate

a refined transformation, this time using not only three points but all matching points.

This approach is similar to the RANSAC algorithm. However, considering the small size of the point sets, an exhaustive search can be performed without any significant performance drawback.

The transformations are smoothed with a Kalman-filter: The head is considered a rigid body. The current position and rotation (using versor representation) are both 3D vectors $x, r \in \mathbb{R}^3$. The kalman state vector considers two derivatives, being therefore $(x, r, \dot{x}, \dot{r}, \ddot{x}, \ddot{r}) \in \mathbb{R}^{18}$.

Then, the transformations are passed on to the MRI scanner via ethernet.

3 RESULTS

The system described in section 2 has been implemented in C++. All image processing was done on the graphics board using the CUDA API. An overall frame rate of approximately 40 fps has been achieved (which is enough to analyze all frames coming from the cameras) and a latency of 0.05 s. The following experiments were made to illustrate the systems performance with MR compatible cameras, but under better lighting conditions than usually found inside the MR scanner.

3.1 Accuracy

One major problem in feature detection is that projective transformations transform circles into shapes similar to ellipses. In most cases, the circle center are not projected onto the center of gravity of the ellipse. A blue circular marker with a small black dot in its center was used to measure the distance. Under extreme angles the displacement can be as much as 10% of the circle radius.

To evaluate the accuracy of the center of gravity, a static scene has been constructed containing only a single circular marker. The experiment shows that under good lighting conditions, the center of gravity of a marker is detected with an accuracy of roughly 0.5 pixels.

To evaluate the accuracy of the triangulated 3D position of a single marker, the same setup has been used. Thus, stereo matching always works correctly and the triangulated 3D position should be exact. The reconstructed 3D positions have a standard deviation of 0.015 mm for the x- and y-axis, and 0.08 mm for the z-axis.

4 CONCLUSIONS

A method to build an optical tracking system that can be used for head motion compensation in MRI has been presented. To achieve MRI compatibility, certain drawbacks had to be accepted: MR compatible cameras are used that provide images with standard TV resolution. As a result, markerless tracking approaches cannot be used. Instead, an approach tracking circular blue markers sticking to the patients forehead has been chosen.

Because of the image resolution provided by the cameras, tracking accuracy of this approach is not as good as it would have been with standard industrial cameras with a much higher resolution and frame rate. Furthermore, the stereo feature matching and the model tracking algorithms had to be able to cope with noisy feature positions.

Thus, an MR compatible tracking system has been built that can be used with non-cooperative patients. In contrast to other approaches, the system is completely MR compatible. It uses markers that do not require the patients cooperation

REFERENCES

Dold, C., Zaitsev, M., Speck, O., Firle, E. A., Hennig, J., and Sakas, G. (2006). Advantages and limitations of prospective head motion compensation for mri using an optical motion tracking device. *Academic Radiology*, 13(9):1093–1103.

Hartley, R. and Zisserman, A. (2000). *Multiple View Geometry in Computer Vision*. Cambridge University Press, Cambridge.

Langner, J. (2008). *Event-Driven Motion Compensation in Positron Emission Tomography: Development of a Clinically Applicable Method*. PhD thesis, Faculty of Medicine Carl Gustav Carus, University of Technology Dresden, Germany.

Ma, W. P. T., Hamarneh, G., Mori, G., Dinelle, K., and Sossi, V. (2008). Motion estimation for functional medical imaging studies using a stereo video head pose tracking system. *Nuclear Science Symposium Conference Record, 2008. NSS '08. IEEE*, pages 4086–4090.

Munkres, J. (1957). Algorithms for the assignment and transportation problems. *Journal of the Society for Industrial and Applied Mathematics*, 5(1):32–38.

Ohayon, S. and Rivlin, E. (2006). Robust 3d head tracking using camera pose estimation. *International Conference on Pattern Recognition*, 1:1063–1066.

Se, S., Lowe, D., and Little, J. (2001). Vision-based mobile robot localization and mapping using scale-invariant features. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 2051–2058.