A Novel Marker Tracking Method Based on Extended Kalman Filter for Multi-camera Optical Tracking Systems

Li Liu, Bo Sun, Ning Wei, Chao Hu
Shenzhen Institute of Advanced Integration Technology (SIAIT),Chinese Academy of Sciences/Chinese University of Hong Kong
Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China
Key Lab for Biomedical Informatics and Health Engineering, CAS
{li.liu1, bo.sun, ning.wei, chao.hu}@siat.ac.cn

Max Q.-H. Meng
Department of Electronic Engineering
The Chinese University of Hong Kong
Shenzhen Institute of Advanced Integration Technology, CAS/CUHK
Shatin, N.T. Hong Kong
max@ee.cuhk.edu.hk

***Abstract*—In the Robotics Assisted Surgical System, the position and orientation of the surgical tools must be estimated precisely in real time. In order to decrease the possibility of the occlusion of line-of-sight, the multi-camera optical tracking system prototype is developed. Based on this hardware platform, a novel multi-marker tracking algorithm using Extended Kalman Filter (EKF) for multi-camera tracking systems is proposed, which makes full use of the redundant information of multi-camera. The marker target movement model can benefit from multiple views by performing a special EKF updating 3D state by 2D projection observations on the image planes. This method has the advantage to estimate the movement of a 3D point more accurately by directly using the actual observations from multiple views rather than the non-continuous and possibly error-prone 3D reconstructed point. The experimental results indicate that the presented tracking algorithm is able to track the 3D trajectories of multiple targets simultaneously, and the estimated 3D positions by EKF are in agreement with the actual measurements by stereo triangulation.**

Keywords- surgical navigation, multi-camera systems, optical tracking, extended Kalman filter, muti-target tracking

I. INTRODUCTION

During the surgical operation for the spinal injury or fractures, the surgeons have experienced the technical difficulties to locate the surgical tools precisely [1]. It is caused by the unstable human handling and the long time operation fatigue. on the other hand, this operation demands high safety and accurate positioning because the spinal area is distributed with major blood vessels and nerves. In order to help surgeons improve tool insertion accuracy in spinal surgery, it is of importance to take advantage of the emerging techniques such as positioning, navigation and robotics [2, 3]. The Robotics Assisted Surgical System (RASS) is a promising technique for the spinal surgery. In most procedures of the spinal surgery, there are many delicate operations involving inserting the tools accurately and precisely in a confined workspace. Although there are several positioning techniques, the optical tracking technique is a suitable choice to provide continuous and real-time position and orientation in the RASS. Additionally, the optical tracking system can also be applied to the rehabilitation robotic system to obtain the accurate motion information of stroke patient [4, 5].

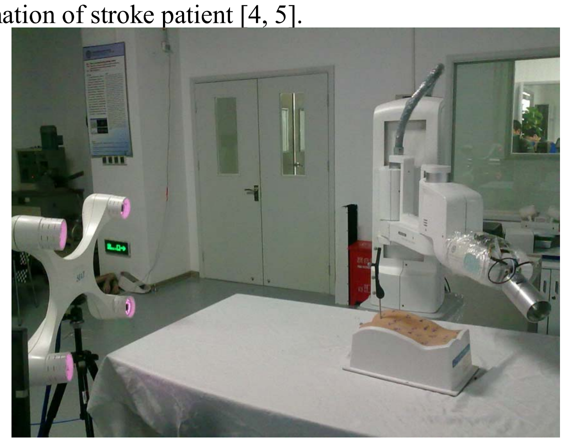

Fig.1. Spine surgical robot system

So far, commercially available optical tracking systems have been successfully used in medical spatial measurement with their satisfactory accuracy and flexibility, such as the Polaris Family of Optical Positioning System from Northern Digital Inc. [6] and the Firefly Motion capture system from Cybernet System Corporation [7]. However, these instruments might fail in crowded working space because the line-of-sight of the optical cameras is blocked.

To address this problem, a multi-camera optical tracking system prototype is designed in order to track precisely surgical tool position and attitude. There are a lot of advantages for multi-camera tracking systems, for instance the possibility of occlusions is decreasing with increasing number of cameras [8]. Occlusions will happen if the surgeon or other objects are obstructing the line-of-sight of one ore more cameras. Chen et al. presented in [8] a quality metric for multi-camera setups. The uncertainty in estimating the 3D position of a feature is computed geometrically based on the uncertainty in estimating the 2D position on the image planes of the used cameras. The metric for positional uncertainty is combined with metric considering occlusions to a overall metric. With the help of this overall metric the accuracy of the tracking systems can be determined based on the placement of the cameras and the 2D uncertainty of the cameras. Another advantage is that the

This work was supported in part by the grants from Key Lab of Robotics & Intelligent System, Guangdong Province (2009A060800016), the Guangdong/CAS Cooperation Project (2009B091300160), National Natural Sc. Foundation of China (60904031), Shenzhen Sc. & Tech. Research Funds, the Knowledge Innovation Eng. Funds of CAS, and the Funds of SRF for ROCS, SEM..

working volume can be increased by combining more cameras. In [9] Cerfontaine et al. presents a method for automated optimal camera alignment for an n-ocular optical tracking system. The alignment of the cameras is optimized for two goals, one is the widest possible working volume, and the other one is the maximum camera visibility.

In this article, a four-ocular optical tracking systems for spinal surgical robot are presented. Based on this hardware platform, a novel simultaneous multi-marker tracking method based on Extended Kalman Filter (EKF) is proposed. The method makes full use of the redundant information from multi-camera images. The target movement model can benefit from multiple views by performing a special EKF updating three-dimensional state by two-dimensional projection observations. The experimental results show that this tracking algorithm is able to track three-dimensional trajectory of multiple markers accurately and simultaneously, and the estimated three-dimensional positions by EKF are in agreement with the actual measurements by stereo triangulation, where the respective tracking accuracies in x-direction, y-direction and depth z-direction are acceptable.

The rest of this paper is organized as follows: In section II, we present our developed multi-camera optical tracking systems. In section III, we propose a novel multi-marker temporal tracking method for multi-camera systems based on EKF. In section IV, we demonstrate the experimental results and give some analysis, followed by the conclusion given in section V.

II. MULTI-CAMERA OPTICAL TRACKING SYSTEM ARCHITECTURE

The multi-camera optical tracking systems detect the projections of the features on the image plane and then estimate the 3D position by stereo triangulation [10]. The optical tracking system use usually IR retro-reflective spherical markers or IR LEDs to ease the process of detecting and estimating the position of the features in the taken image. When the image is captured by the camera, the features have to be segmented, identified and their positions on the image have to be determined. Because the measurements are always tainted with noise, the rays used for stereo triangulation will in general not intersect each other. This situation also occurs if using more than two cameras for the triangulation where the triangulation problem is over determined. The basic 2D to 3D stereo triangulation finds the optimal 3D position for two or more 2D camera views of the point, and is implemented using a linear least-square fit of the intersection of n rays defined by the 2D image points and 3D camera centers of each of the n cameras [10, 11]. It is not difficult to understand that the stereo triangulation will become the more precisely if the more cameras are used.

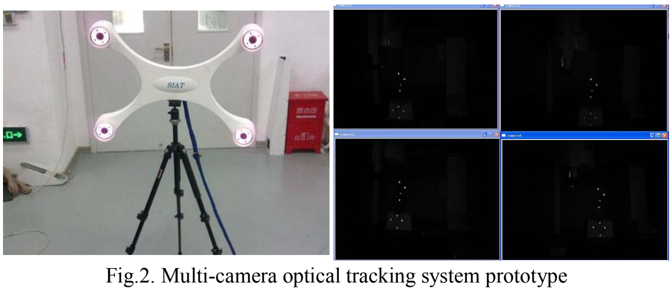

Fig.2. Multi-camera optical tracking system prototype

As is shown in Fig.2, our developed multi-camera optical tracking system prototype employs four cameras to locate the position and orientation of multiple surgical instruments simultaneously for robotics assisted spinal surgery.. In near future work, we hope that the amount and distribution arrangement of the cameras can be adjusted according to the specific requirement in order to obtain better tracking performance.

III. MULTI-MARKER TRACKING BASED ON MULTI-CAMERA OPTICAL TRACKING SYSTEMS

Aiming to track the multiple infrared markers attached to the surgical tools simultaneously, the following procedures must be settled:

- to determine the extrinsic parameters of all cameras with respective to the world coordinate system and the intrinsic parameters of each camera;
- to determine the 2D coordinates of the centroids of the blobs derived from the projections;
- to find the stereo correspondence of the blobs from multiple views;
- to determine 3D position of markers by stereo triangulation or real-time tracking by EKF.

The first procedure can be carried out offline using a standard calibration toolbox which is available in [12]. The second procedure can be performed by Circular Hough Transform in [13]. The third procedure can be realized via both epipolar constraint and neighborhood similarity constraint in [14]. This section focuses on a novel tracking algorithm based on EKF.

A. Extended Kalman Filter Overview

Kalman filter is statistically optimal tool for predicting linear movements based on filter state which captures information from all previous measurements [15, 16]. The filter has an internal state update function and an observation function.

State update function is used to model the changes in the state of the system:

$$s_t = As_{t-1} + Bu_{t-1} + w_{t-1} \tag{1}$$

Observation update function is used to model the measurement of the state:

$$y_t = Hs_t + v_t \tag{2}$$

Extended Kalman Filter (EKF) [15, 16] has a similar structure to Kalman filter except it allows non-linear state transition and observation functions. EKF works by linearizing the state update function and observation function using Jacobians. A Jacobian matrix consists of partial derivatives of a function with respect to all state parameters. Because only the observation process is nonlinear, the process dynamics are specified with (linear) matrix A.

The a priori estimation of these states based on the previous frame's posterior estimates are

$$\hat{s}_{t|t-1} = A\hat{s}_{t-1|t-1} \tag{3}$$

and

$$P_{t|t-1} = AP_{t-1|t-1}A^T + Q \tag{4}$$

In order to incorporate these observations, a gain term K is calculated to weight the innovation arising from difference between the a priori state estimate $\hat{s}_{t|t-1}$ and the observation y_t

$$K_t = \frac{P_{t|t-1} H_t^T}{H_t P_{t|t-1} H_t^T + R} \tag{5}$$

The observation matrix H_t is defined to be the Jacobian of the observation function evaluated at the expected state

$$H_t = \frac{\partial h}{\partial s}\bigg|_{\hat{s}_{t|t-1}} \tag{6}$$

And then the posterior estimations are

$$\hat{s}_{t|t} = \hat{s}_{t|t-1} + K_t(y_t - H_t \hat{s}_{t|t-1}) \tag{7}$$

and

$$P_{t|t} = (I - K_t H_t) P_{t|t-1} \tag{8}$$

B. Extended Kalman Filter For Multi-camera Optical Tracking

Based on the multi-camera optical tracking platform, the tracked marker movement model can benefit from multiple views by performing a special Kalman filter updating internal three-dimensional state by two-dimensional projection observations on the image planes. Extended Kalman filter observation function can be defined as a non-linear projection.

In the beginning of each tracking process, the Kalman filter predicts the positions of all interested markers. The priori predicted positions are then projected into all camera views and if close enough observations can be found (namely "Data association process"), the priori predicted 3D point is corrected to correspond to the observations from other views, and finally the posterior estimates of the 3D point can be obtained, the Fig.3 illustrates EKF tracking process.

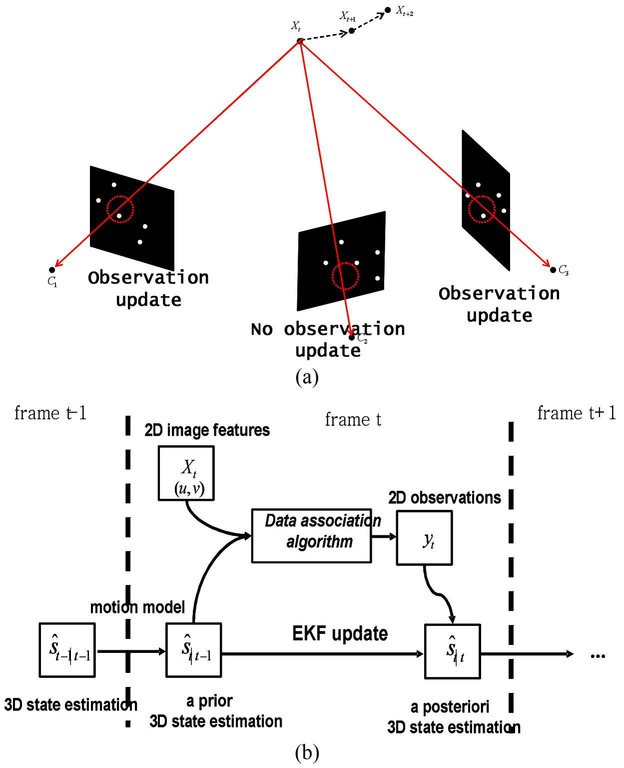

Fig.3. (a) One EKF iteration cycle illustrated. Each view may contribute one observation update to EKF; (b) EKF tracking flowchart

This tracking process allows inputting the filter with the actual projection observations rather than more error-prone 3D reconstructed points. All 2D projections of a 3D point can then be given to be the observation updates for the filter assuming that the observation function is changed to correspond to respective projection matrix of each view. Using only one view for update will eventually cause the filter to diverge because there is no restriction in the movement along the line of sight. When more than one view is used the estimated 3D point may stay in its optimal and filtered position. Each filtering cycle thus contains one prediction and one observation update per each view for a single 3D point. The benefit of this method is the ability to model the movement of a 3D point more accurately using redundant information from possible all view. What's more, the method directly uses the actual observations on the images rather than non-continuous and possibly error-prone 3D reconstructions by triangulation. Observation model for multi-camera Kalman filter has a projection matrix of each view as an additional parameter.

The EKF estimates state and its covariance based on a prior state estimate and incoming observations by using models of the state update process, the observation process, and estimates of the noise of each process.

We use a linear model for the dynamics of the system and a nonlinear model of the observation process.

Specifically, the time evolution of the system is modeled with the linear discrete stochastic model

$$s_t = A s_{t-1} + w \tag{9}$$

Here we treat the marker target as a simple orientation-free object, with state vector $s = (x, y, z, \dot{x}, \dot{y}, \dot{z})$ describing position and velocity in three-dimensional space.

In our case, the process update matrix A models the law of constant velocity of an orientation-free object.

$$A = \begin{bmatrix} 1 & 0 & 0 & \Delta t & 0 & 0 \\ 0 & 1 & 0 & 0 & \Delta t & 0 \\ 0 & 0 & 1 & 0 & 0 & \Delta t \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

with Δt being the time step. The random variable w represents this process update noise with a normal probability distribution with zero mean and the process covariance matrix Q. Despite the use of a simple constant velocity model, the more complex trajectory of the marker is accurately estimated by updating the internal state estimate with multi-view observations.

2D image point in homogeneous coordinates

$$x_i = (a_i, b_i, c_i), \tag{10}$$

We define the function ζ to convert from homogeneous to Cartesian coordinates, where $u_i = a_i / c_i$ and $v_i = b_i / c_i$

$$\zeta(x) = (u, v) = \left(\frac{a}{c}, \frac{b}{c}\right). \tag{11}$$

The 3×4 camera calibration matrix P_i models the projection from a 3D homogeneous point $\mathrm{X} = (X_1, X_2, X_3, X_4)$ representing the 3D point with inhomogeneous coordinates $(x,y,z)=(X_1/X_4, X_2/X_4, X_3/X_4)$ into the image point:

$$x_i = P_i X. \tag{12}$$

The observation vector $y=(u_1,v_1,u_2,v_2,...,u_n,v_n)$ is the vector of the image points from n cameras. The nonlinear observation function associates y_t with the internal state s_t by

$$y_t = h(s_t)+v \tag{13}$$

$$\begin{aligned} h(s) &= (h_1(s),...,h_n(s)) = (\bar{u}_1,\bar{v}_1,...,\bar{u}_n,\bar{v}_n) \\ &= \left(\frac{\bar{a}_1}{\bar{c}_1},\frac{\bar{b}_1}{\bar{c}_1},...,\frac{\bar{a}_n}{\bar{c}_n},\frac{\bar{b}_n}{\bar{c}_n}\right) = (\mathrm{H}(P_1X),...,\mathrm{H}(P_nX)) \end{aligned} \tag{14}$$

C. Data Association For Observation Update Process

The purpose of the data association step is to determine which image feature points will be used by which tracked target. Because multiple targets may be tracked simultaneously, correct detections must be associated with the respective Kalman model.

This section outlines how the data association step is applied to determine which blob point on the images to be the observation for a given marker target. The criterion for a point to be seen in a view is to have an observation closer than tolerance value to the ideal projection. Thus projection observations are searched inside a tolerance circle in all views, as shown in Fig4. In other words, the possible observation candidate, the detected blob location (u_j, v_j), must be within a threshold Euclidean distance from the estimated target location projected on the image. The Euclidean distance on the image plane is

$$D(u_j,v_j) = d_{euclid}\left(\begin{bmatrix} u_j \\ v_j \end{bmatrix}, \mathrm{H}(P_iX)\right), \tag{15}$$

where $\mathrm{H}(P_iX)$ finds the projected image coordinates of X, where X is the homogeneous form of the first three components of the state vector S, the estimated 3D position of the marker target.

Note that we set a Euclidean distance threshold on image plane, $Thresh_{dist2d}$, which is the radius of the tolerance circle mentioned above. EKF iteration works by doing a sequence of observation updates, once per each view. When there is at least one observation candidates inside the tolerance circle in some view, the blob corresponding to the minimum of $D(u_j,v_j)$ will be fed into kalman filter for observation update. If no blob inside the tolerance circle in some view, it means that this view can not provide observation update, as shown in Fig.4.

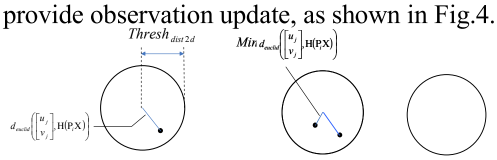

Fig.4. Tolerance cycle for data association, the radius of cycle is $Thresh_{dist2d}$

IV. Experiments and Discussion

Using our developed multi-camera tracking systems, we carried out several experiments including single target tracking and multi-target tracking simultaneously. The position of the marker in three-dimensional space can be calculated by stereo triangulation. However, when the systems need to locate multiple targets simultaneously, latency may be a challenge. Latency is the time delay between the moment in time motion occurs and the moment the position information is provided by the tracker. EKF tracking algorithm can overcome this problem to some extent.

We compare the estimated 3D positions by EKF with the 3D measurements by stereo triangulation. The experimental results prove that EKF tracking performance is acceptable.

Fig.5 (a) shows a single marker tracking result, where the blue curve represents the reconstructed trajectory using stereo triangulation while the red curve represents the estimated tracking trajectory using EKF.

Fig.5 (b)~(d) illustrates the tracking effect along x-direction, y-direction and depth z-direction. It is clear that the tracking error mainly comes from depth z-direction, which is relatively bigger than those along other directions

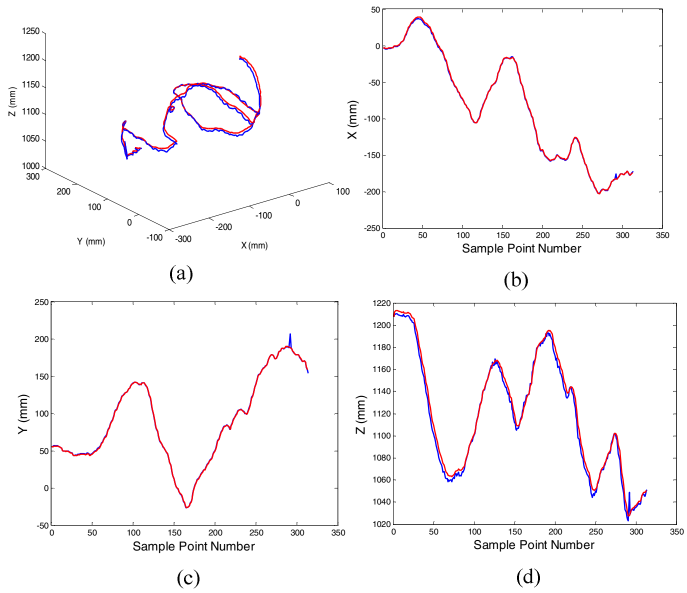

Fig.5. (a) Single target tracking trajectory, stereo triangulation reconstructed trajectory (blue curve), EKF tracking trajectory (red curve)
The tracking trajectory along the respective directions
(b) x-direction tracking trajectory, (c) y-direction tracking trajectory;
(d) depth z-direction tracking trajectory.

In the case of multi-target tracking simultaneously, the data association step plays an important role, which is able to ensure the correct observation update corresponding to the respective tracked target. Here we take two marker targets as an example, it is obviously seen that EKF tracking algorithm with multi-view obtains a good tracking trajectory shown in Fig.6 (a), which is almost in agreement with the measurements by stereo triangulation. Similarly, the tracking accuracy along the depth z-direction is lower than those along x- and y-direction, as shown in Fig.6 (b)~(d).

This situation is easy to understand. In general, a single projection can only correspond one 3D point assuming that there is only one 3D point in the line of sight. It means that projection observations are sometimes insensitive to some 3D points almost in the same line of sight with different depths. When more cameras with diverse position and attitude are integrated into the tracking systems which can supply more redundant projection information, their tracking accuracy along depth z-direction can be increased.

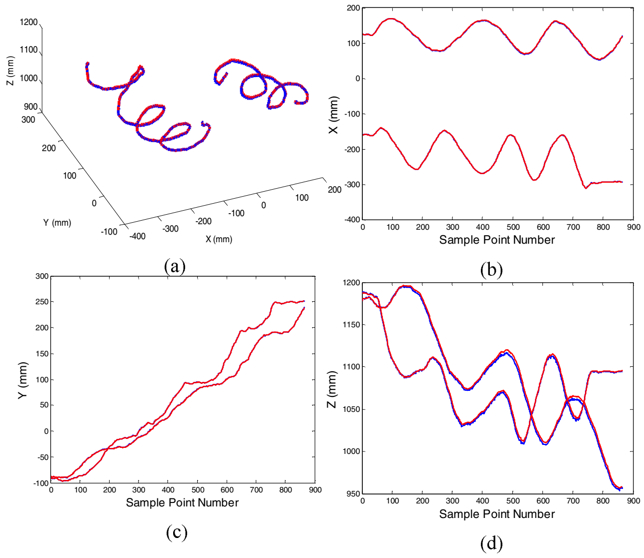

Fig.6. (a) Multi-target tracking trajectory, stereo triangulation reconstructed trajectory (blue curve), EKF tracking trajectory (red curve) Multi-target tracking trajectory in the respective directions (b) x-direction tracking trajectory; (c) y-direction tracking trajectory; (d) depth z-direction tracking trajectory.

V. CONCLUSION

This paper presented the multi-camera optical tracking systems for robotics assisted surgery in order to localize and track the position and orientation of the surgical tools accurately. Compared with the existing commercial available optical tracking products, our developed multi-camera optical tracking prototype has the ability to decrease the possibility of occlusion of line of sight to some degree.

Based on the multi-camera optical tracking platform, a novel multi-marker tracking algorithm based on Extended Kalman Filter is proposed, which makes full use of the redundant information supplied by multiple cameras. The marker target movement model can benefit from multiple views by performing a special EKF updating 3D state by 2D projection observations on the image planes. This method has the advantage to estimate the movement of a 3D point more accurately by directly using the actual observations from multiple views rather than the non-continuous and possibly error-prone 3D reconstructed point by stereo triangulation. The experimental results indicate that the presented tracking algorithm is able to track the 3D trajectories of multiple targets simultaneously, and the estimated 3D positions by EKF are in agreement with the actual measurements by stereo triangulation.

ACKNOWLEDGMENT

The work in this paper is supported by the grants from Key Lab of Robotics & Intelligent System, Guangdong Province (2009A060800016), the Guangdong/CAS Cooperation Project (2009B091300160), National Natural Sc. Foundation of China (60904031), Shenzhen Sc. & Tech. Research Funds, the Knowledge Innovation Eng. Funds of CAS, and the Funds of SRF for ROCS, SEM.

REFERENCES

[1] F. Langlotz and L.-P. Nolte, “Technical approaches to computer-assisted orthopedic surgery," European Journal of Trauma, vol. 30, pp. 1-11(11), February 2004.

[2] K. Leung, “New technology: Image-guided navigation in orthopaedic trauma surgeries," pp. 249-263, 2006.

[3] K. Peters, Terry; Cleary, Image-Guided Interventions: Technology and Applications. Springer, 2008.

[4] Y. Tao, H. Hu and H. Zhou, “Integration of vision and inertial sensors for 3d arm motion tracking in home-based rehabilitation”, Int. J. Robot. Res. 26 (2007), pp. 607–624.

[5] H. Zhou and H. Hu, “Inertial motion tracking of human arm movements in stroke rehabilitation,” in Proc. IEEE Int. Conf. Mechatronics Autom., Niagara Falls, ON, Canada, Jul. 29–Aug. 1, 2005, pp. 1306–1311.

[6] Northern Digital Inc. (2008) Polaris Family of Optical Tracking Systems.[Online]Available:http://www.ndigital.com/medical/polarisfamily.php

[7] Cybernet Systems Corporation. (2008) Firefly Motion Capture System. [Online].Available:http://www.cybernet.com/interactive/firefly/firefly.html

[8] X. Chen and J. Davis, “Camera Placement Considering Occlusion for Robust Motion Capture”, Tech. Rep. CS-TR-2000-07, Stanford University Computer Science, 2000.

[9] P. A. Cerfontaine, M. Schirski, D. Bundgens and T. Kujlen, “Automatic Multi-Camera Setup Optimization for Optical Tracking”, in Proceedings of IEEE Virtual Reality, 2006.

[10] R. I. Hartley and P. Sturm, “Triangulation”, Lecture Notes in Computer Science, 970(1995), pp. 190–197.

[11] R. I. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision, Cambridge University Press, June 2000.

[12] Camera Calibration Toolbox for Matlab, 2008. [Online]. Available:http://www.vision.caltech.edu/bouguetj/calib_doc/index.html

[13] C. Kimme, D. Balard, and J. Sklansky, “Finding circles by an array of accumulators”. Commun. ACM, 18(2):120–122, 1975.

[14] Van Rhijn, A., and Mulder, J. “Optical Tracking and Calibration of Tangible Interaction Devices”. Proceedings of 11th Eurographics Symposium on Virtual Environments, Aalborg, Denmark (2005), 41–50.

[15] Welch, G., and Bishop,” An Introduction to the Kalman Filter”, http://www.cs.unc.edu/~tracker/media/pdf/SIGGRAPH2001_CoursePack_08.pdf. SIGGRAPH 2001 Course 8 online material, Last visited 17.9.2005.

[16] Welch, G., and Bishop, “SCAAT: Incremental Tracking with Incomplete Information”. ACM Press/Addison-Wesley Publishing Co., 1997.