

Using a Wireless Motion Controller for 3D Medical Image Catheter Interactions

Dime Vitanovski, Dieter Hahn, Volker Daum, Joachim Hornegger
Chair of Pattern Recognition, Martensstr. 3, 91058 Erlangen, Germany

ABSTRACT

State-of-the-art morphological imaging techniques usually provide high resolution 3D images with a huge number of slices. In clinical practice, however, 2D slice-based examinations are still the method of choice even for these large amounts of data. Providing intuitive interaction methods for specific 3D medical visualization applications is therefore a critical feature for clinical imaging applications. For the domain of catheter navigation and surgery planning, it is crucial to assist the physician with appropriate visualization techniques, such as 3D segmentation maps, fly-through cameras or virtual interaction approaches. There has been an ongoing development and improvement for controllers that help to interact with 3D environments in the domain of computer games. These controllers are based on both motion and infrared sensors and are typically used to detect 3D position and orientation. We have investigated how a state-of-the-art wireless motion sensor controller (Wiimote), developed by Nintendo, can be used for catheter navigation and planning purposes. By default the Wiimote controller only measure rough acceleration over a range of $\pm 3g$ with 10% sensitivity and orientation. Therefore, a pose estimation algorithm was developed for computing accurate position and orientation in 3D space regarding 4 Infrared LEDs. Current results show that for the translation it is possible to obtain a mean error of (0.38cm, 0.41cm, 4.94cm) and for the rotation (0.16, 0.28) respectively. Within this paper we introduce a clinical prototype that allows steering of a virtual fly-through camera attached to the catheter tip by the Wii controller on basis of a segmented vessel tree.

Keywords: Calibration, Surgical Simulation, Therapy Planning, Treatment Planning, Wiimote Calibration, Wiimote Pose Estimation

1. INTRODUCTION

Interventional radiology (IR), under fluoroscopy, CT, ultrasound or MRI guidance, performs a minimally invasive therapy for endovascular treatment of vascular disease. In order to obtain an internal access to the diseased site, the interventional radiologist inserts a catheter into a blood vessel. Due to the complex anatomy of the vascular system and the significant risk associated with interventional radiology (IR) procedures, a high level of experience is required which can be obtained by efficient training of the interventional radiologists. Since learning on patients is becoming unacceptable,¹ different system simulations for training and planning purposes have been developed. One of the first interventional radiology (IR) simulators was developed by Dawson-Kaufman at HT Medical Inc.² Another, more complex simulator, was introduced by Xitact Inc. as a transportable table-top simulation platform for intravascular procedures, including cardiology, interventional radiology, and peripheral interventions.³ The Image Sciences Institutes, University Medical Center Utrecht has proposed a system for simulation of minimally invasive vascular interventions called MIVIS.⁴ The most advanced and popular PC based system for simulation of image-guided neuroradiology interventions is NeuroCath,^{5,4,6} Based on NeuroCath (see Figure 1), Biomedical Imaging Lab developed a new system for remote interventional radiology training, called "Tele-IR". Various training systems for cardiology intervention simulations were developed by major research institutes like: CIMIT,⁷ the Institute for Computational Medicine, University of Mannheim,⁸ and Barnes.⁹ To our best knowledge, there are no efforts aiming to develop simulators for wireless IR procedures training.

In the field of endoscopy, in order to reduce the risk and the costs, a virtual fly through inside the human organs is simulated, which makes the medical examination and treatment less invasive, safe and cost effective.

Further author information: (Send correspondence to Dime Vitanovski)
Dime Vitanovski: E-mail: vitanovski@gmail.com, Telephone: +49 179 4869157



Figure 1. System configuration for pretreatment planning showing monitors with both 2-D and 3-D views and a trainee manipulating a catheter through the instrument interface device.

Through virtual study on patient data the interventional radiologist can perform a diagnosis without physical operation on the patient, which reduces the patient discomfort.¹⁰ The goal of these systems is to help the physicians to improve their skills and their hand-eye coordination. Since the catheter navigation process requires a real-time data processing we have investigated how the controllers in the domain of computer game can be used for such purposes. With the ongoing development and improvement for controllers that help to interact with 3D environments in computer games, based on motion and infrared sensors, one can detect the 3D position and orientation of the human hand. One such controller is the Wiimote (see Figure 2), developed by the company Nintendo,¹¹ which can measure the earth acceleration and can track up to 4 Infrared LEDs (see Figure 3). In particular, it has a three-axes accelerometer, able to measure accelerations along three perpendicular axes over a range of $\pm 3g$ with 10% sensitivity and an additional monochrome 1024x768 Infrared camera, which tracks up to 4 Infrared LED sources within the field of view, is placed on the front of the Wiimote. The controller sends every 1ms over Bluetooth pre-processed data to the PC. By default the Wiimote controller only measures rough acceleration and orientation. Therefore, a pose estimation algorithm was developed for computing accurate position and orientation in 3D space. Although it was initially designed to work only with its main console hardware, it can be adapted for PC based navigation purposes. The communication between the Wiimote and PC is established via a Bluetooth wireless link. The Bluetooth controller is a Broadcom 2042 chip, which is designed to be used with devices that support the Bluetooth Human Interface Device (HID) protocol proposed by Tom van Delft of Microsoft.

A Human Interface Device, or HID, is a type of computer device that interacts directly with, and most often takes input from humans (via Output Reports) and provides output to humans (via Input Reports). The HID standard allows devices to be self-describing, using a HID descriptor block. This block includes an enumeration of reports understandable by the device. A report can be thought of as a network port assigned to a particular service. However, reports are unidirectional, and the HID descriptor lists for each port the direction (Input or Output) and the payload size for each port. "Output" refers to packets that are sent from the host to the Wiimote, and "Input" refers to packets that are sent from the Wiimote to the host.

Within this paper we introduce a new system for patient pre-treatment and training based on wireless navigation of catheter tip. The main benefit of the advance navigation, enabled by our wireless navigation system, includes real-time data processing of the steering inputs, high degree of freedom, and better orientation

2. METHODS

Since in the most of the simulators for IR, the human motion is tracked by series of electromechanical sensors, the emphasis in this work was how to estimate the pose of the device in 3D space based on 4 point correspondences. Pose estimation is a well-studied problem in the computer vision, which from known 3D/2D point correspondences, the orientation and the rotation (also called extrinsic camera parameters) of the camera can



Figure 2. Nintendo Wii Controller (Source: www.wiivii.tv)

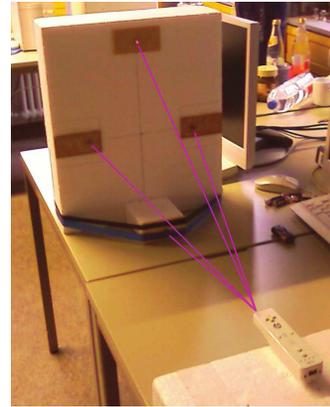


Figure 3. Infrared Beacon

be estimated. Pose estimation usually requires pre-calibrated camera. The traditional way to compute the intrinsic camera parameters is to use a known calibration object and as an input only information for 3D/2D point correspondences is required. Since the parameters are not depend on the position and orientation of the camera in space, they can be computed once by doing offline camera calibration, or if a sufficient number of point correspondences is available, they can be incorporated in the camera pose estimation. Within this paper we have applied two different technique for the pose estimation of the Wiimote IR camera: The Direct Linear Transformation method (see section 2.1) and Pose Estimation with Dual Quaternions from known 4 2D-3D Point Correspondences (see section 2.2).

2.1 Pose Estimation with Direct Linear Transformation Method from known 4 2D-3D Point Correspondences

In the case of the Direct Linear Transformation, an offline camera calibration is required. The intrinsic camera matrix (see Figure 4) is defined with five parameters: f_x, f_y, s, H_x, H_y . The real pixel coordinates are computed by multiplication with the camera matrix, where f_x, f_y is the focal length of the camera, s (skew) is the angle between the sensors and H_x, H_y is the principal point of the real coordinate system u/v . The Wiimote camera calibration was done by taking 300 different values for the extrinsic parameters and as a result, the following values for the intrinsic camera parameters are computed: the Focal length is $f_x = f_y = 1320$ pixels, the skew parameter $s = 0$ (the Wiimote sensor has a rectangle form) and the offset $H_x = 512, H_y = 384$.

$$K = \begin{bmatrix} f_x & s & H_x \\ 0 & f_y & H_y \\ 0 & 0 & 1 \end{bmatrix}$$

Figure 4. Intrinsic Camera Matrix

The Direct Linear Transformation (DLT) is one of the most commonly used approaches for camera pose estimation. With 6 point correspondences it can be rewritten in a closed form solution, which can be solved using SVD. Since the Wiimote IR camera can track up to 4 IR points, a non-linear optimization technique is required, in order to solve the problem of pose estimation. Hence, an offline camera calibration is necessary. The energy function for the camera pose estimation can be defined as the quadratic difference between the ground truth value and the projected values (see Formula 1):

$$f(R, T) = \underset{\hat{R}, \hat{T}}{argmin} \sum_{i=1}^4 \left(\begin{matrix} x_i^p (x_i^w p_{31} + y_i^w p_{32} + z_i^w p_{33} + t_x) - x_i^w p_{11} + y_i^w p_{12} + z_i^w p_{13} + t_x \\ y_i^p (x_i^w p_{31} + y_i^w p_{32} + z_i^w p_{33} + t_x) - x_i^w p_{21} + y_i^w p_{22} + z_i^w p_{23} + t_y \end{matrix} \right)^2 \quad (1)$$

2.2 Pose Estimation with Dual Quaternions from known 4 2D-3D Point Correspondences

In the previous section the pose estimation method was introduced which required offline camera calibration. This approach uses the properties of dual quaternions as parameterization for the rotation and the translation, and estimates the camera pose without calibration.

This method, first presented by¹² in 1995, uses some important properties of the quaternions:

$$\mathbf{r}\mathbf{q} = Q(\mathbf{r})\mathbf{q} = W(\mathbf{r})\mathbf{q} \quad (2)$$

$$Q(\mathbf{r})\mathbf{q} = Q(\mathbf{q})\mathbf{r} \quad (3)$$

where $Q(\mathbf{r})$ and $W(\mathbf{r})$ are 4×4 matrices associated with a quaternions:

$$Q(\mathbf{r}) = \begin{pmatrix} r_0 & -r_x & -r_y & -r_z \\ r_x & r_0 & -r_z & r_y \\ r_y & r_z & r_0 & -r_x \\ r_z & -r_y & r_x & r_0 \end{pmatrix} \quad (4)$$

$$W(\mathbf{r}) = \begin{pmatrix} r_0 & -r_x & -r_y & -r_z \\ r_x & r_0 & r_z & -r_y \\ r_y & -r_z & r_0 & r_x \\ r_z & r_y & -r_x & r_0 \end{pmatrix} \quad (5)$$

The main idea of this approach is to rewrite the equations 1 in the following form:

$$(1 \ 0 \ -x_p) R \begin{pmatrix} X_w \\ Y_w \\ Z_w \end{pmatrix} + (1 \ 0 \ -x_p) \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix} = 0 \quad (6)$$

$$(1 \ 0 \ -y_p) R \begin{pmatrix} X_w \\ Y_w \\ Z_w \end{pmatrix} + (1 \ 0 \ -y_p) \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix} = 0 \quad (7)$$

where the rotation matrix R and the translation vector \mathbf{t} are represented with dual quaternions

$$\begin{pmatrix} 1 & 0^T \\ 0 & R \end{pmatrix} = W(\mathbf{r})^T Q(\mathbf{r}) \quad (8)$$

$$\begin{pmatrix} 0 \\ \mathbf{t} \end{pmatrix} = 2W(\mathbf{r})^T \mathbf{s} \quad (9)$$

The 3D/2D point correspondences \mathbf{m}_{x_p} , \mathbf{m}_{y_p} and \mathbf{M}_w has to be written in a dual form notation:

$$\mathbf{M}_w = (0 \ X_w \ Y_w \ Z_w)^T$$

$$\mathbf{m}_{x_p} = (0 \ 1 \ 0 \ -x_p)^T$$

$$\mathbf{m}_{y_p} = (0 \ 0 \ 1 \ -y_p)^T$$

Equations 6 and 7 can be rewritten in the following form:

$$\begin{aligned}
\mathbf{m}_{x_p}^T W(\mathbf{r})^T Q(\mathbf{r}) \mathbf{M}_w + 2\mathbf{m}_{x_p}^T W(\mathbf{r})^T \mathbf{s} &= (W(\mathbf{r}) \mathbf{m}_{x_p})^T Q(\mathbf{r}) \mathbf{M}_w + 2(W(\mathbf{r}) \mathbf{m}_{x_p})^T \mathbf{s} \\
&= \mathbf{r}^T Q(\mathbf{m}_{x_p})^T W(\mathbf{M}_w) \mathbf{r} + 2\mathbf{r}^T Q(\mathbf{m}_{x_p})^T \mathbf{s} \\
&= \mathbf{r}^T A_i \mathbf{r} + 2\mathbf{r}^T B_i \mathbf{s}
\end{aligned} \tag{10}$$

$$\begin{aligned}
\mathbf{m}_{y_p}^T W(\mathbf{r})^T Q(\mathbf{r}) \mathbf{M}_i + 2\mathbf{m}_{y_p}^T W(\mathbf{r})^T \mathbf{s} &= (W(\mathbf{r}) \mathbf{m}_{y_p})^T Q(\mathbf{r}) \mathbf{M}_i + 2(W(\mathbf{r}) \mathbf{m}_{y_p})^T \mathbf{s} \\
&= \mathbf{r}^T Q(\mathbf{m}_{y_p})^T W(\mathbf{M}_i) \mathbf{r} + 2\mathbf{r}^T Q(\mathbf{m}_{y_p})^T \mathbf{s} \\
&= \mathbf{r}^T C_i \mathbf{r} + 2\mathbf{r}^T D_i \mathbf{s}
\end{aligned} \tag{11}$$

where the A_i, B_i, C_i, D_i are 4×4 matrices defined by:

$$\begin{aligned}
A_i &= Q(\mathbf{m}_{x_p})^T W(\mathbf{M}_w) \\
B_i &= Q(\mathbf{m}_{x_p})^T \mathbf{s} \\
C_i &= Q(\mathbf{m}_{y_p})^T W(\mathbf{M}_w) \\
D_i &= Q(\mathbf{m}_{y_p})^T \mathbf{s}
\end{aligned}$$

As explained in¹³ the dual quaternions have two constraints:

$$\begin{aligned}
\mathbf{r} \cdot \mathbf{r} &= 1 \\
\mathbf{r} \cdot \mathbf{s} &= 0
\end{aligned}$$

which has to be included in the objective function(12) with the Lagrange multipliers. By minimizing the objective function in respect to \mathbf{r}, \mathbf{s} one can obtain the rotation and the translation.

$$f(\mathbf{r}, \mathbf{s}) = \operatorname{argmin}_{\mathbf{r}, \mathbf{s}} \sum_{i=1}^4 \left((\mathbf{r}^T A_i^p \mathbf{r} + 2\mathbf{r}^T B_i^p \mathbf{s})^2 + (\mathbf{r}^T C_i^p \mathbf{r} + 2\mathbf{r}^T D_i^p \mathbf{s})^2 \right) + \lambda(\mathbf{r}^T \mathbf{r} - 1)^2 + \lambda(\mathbf{r}^T \mathbf{s})^2 \tag{12}$$

2.3 Point Tracking

One major problem with the described methods for camera pose estimation is that it relies on the correspondences between the target points in 3D world coordinates and their projections onto the 2D image plane. However, the Wiimote camera can track up to 4 IR Led's, it projects in its 2D image plane, with no particular order and no association to the Infrared LEDs (see Figure 3). Therefore, the first step is to determine the 2D/3D point correspondences, which Infrared LED corresponds to which projection, and to solve the non-linear equation. This problem is solved by first sorting the projected point in order to find the correspondent 3D points, afterwards a point tracking algorithm is applied. In order to overcome the measurements noise and estimation error, pose estimation is filtered using the standard Kalman model. The Kalman Filter is an optimal recursive data processing algorithm, which incorporates the estimated measurements from the previous step with the current measurement to compute an estimate for the current state. As stated in,¹⁴ the current values for the variables of interest are estimated based on all available measurement with the use of: knowledge of the system and measurement device dynamics,¹⁴ the statistical description of the system noises, measurement errors, and uncertainty in the dynamics models,¹⁴ available information about initial conditions of the variables of interest.¹⁴ An overview of the algorithm is given by Figure 5:

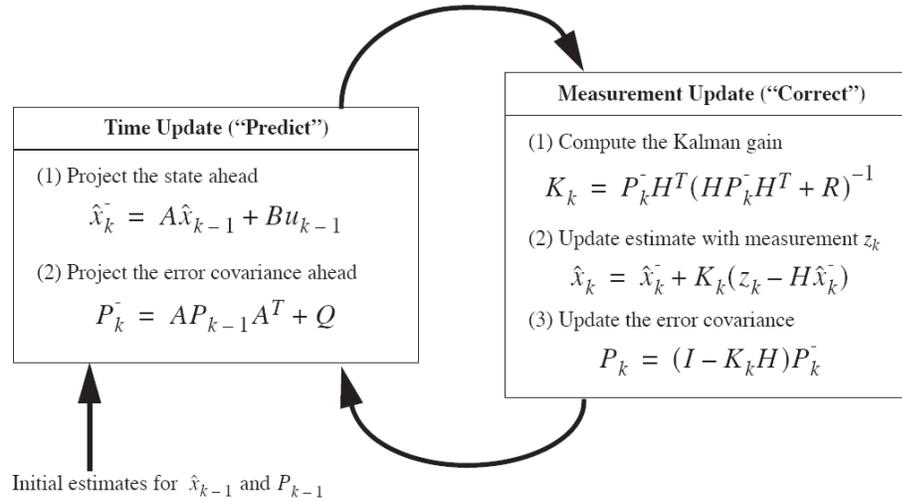


Figure 5. Kalman Filter Algorithm (Source: <http://www.cs.unc.edu/>)

3. RESULTS

The camera pose estimation approach described in methods was evaluated for several positions of the Wiimote attached to a robot arm. All tests were performed with a robot arm (Scorbot -ER VII), which allows precise shifts in all three dimensions with a precision of 0.1 mm. In the following the estimated position and orientation of the Wiimote is compared with the ideal data of the robot arm in order to compute the mean error and the standard deviation.

	R. around X axis	R. around Y axis	T. in X direction	T. in Y direction	T. in Z direction
Mean error	0.16212	0.28624	0.38446cm	0.41119cm	4.94116cm
Std.Deviation	0.40264	0.53502	0.62005cm	0.64124cm	2.22287cm

Table 1. Wiimote Accuracy Measurements (R = rotation, T = translation)

Due to the robot limitations, it was not possible to measure the mean error and the standard deviation of the device for rotation around the Y-axis. However, we believe that the Wiimote accuracy can be improved by using a professional Infrared beacon, infrared LEDs and more accurate robot arm.

4. CONCLUSION

Within our framework the exact rotation and translation of the Wiimote is determined by optimizing an energy function by means of Interior Point OPTimizer (IPOPT) algorithm.¹⁵ In order to overcome the measurements noise and estimation error, pose estimation is filtered using the standard Kalman model. The presented mathematical framework enables an accurate pose estimation computed in real-time, which satisfies catheter navigation requirements. In order to demonstrate the capabilities of our system under clinical conditions the InSpace environment is used to simulate virtual catheter navigation. With the Visualization Toolkit (VTK) a 3D volume was reconstructed from a CT brain series. The brain vessel tree was computed by applying a marching cube algorithm, which enables us a fly-through modus. Within this framework we developed a clinical prototype that allows steering a virtual fly-through camera attached to the catheter tip by the controller on basis of a segmented vessel tree (see Figures 6, 7, 8, 9).

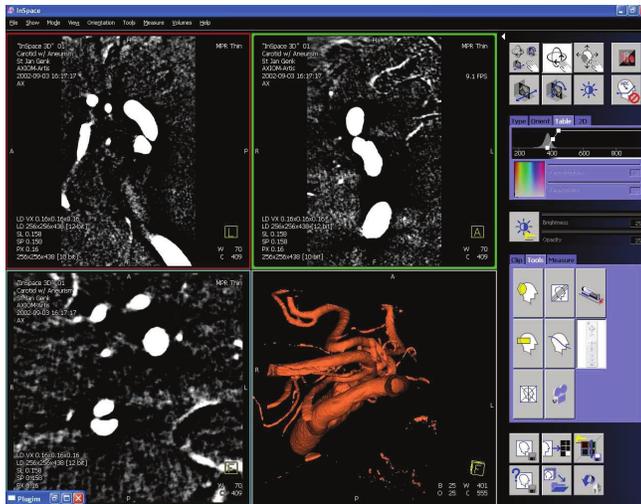


Figure 6. InSpace



Figure 7. Wiimote rotation



Figure 8. Wiimote translation



Figure 9. Wiimote translation

REFERENCES

- [1] Cotin, S., Luboz, V., Pergoraro, V., Neumann, P., Wu, X., and Dawson, S., "High-fidelity simulation of interventional neuroradiology procedures," *Proc. American Society of Neuroradiology ASNR* (May 2005).
- [2] Meglan, D., "Making surgical simulation real," *SIGGRAPH Comput. Graph.* **30**(4), 37–39 (1996).
- [3] Cotin, S., v Luboz, Pergoraro, P., Wu, X., and Dawson, S., *High-fidelity simulation of interventional neu-*

- roradiology procedures. Proc. American Society of Neuroradiology ASNR, Toronto, Canada: 443 (May 2005).
- [4] Alderliesten, T., Konings, M., and Niessen, W., *Simulation of minimally invasive vascular interventions for training purposes* (2004). Computer Aided Surgery.
 - [5] Wieslaw, L., Nowinski, and Chee-Kong, C., "Simulation of interventional neuroradiology procedures," *miar* **00**, 0087 (2001).
 - [6] Chui, C., Li, Z., Anderson, J., KMurphy, Venbrux, A., Ma, X., Wang, Z., Gailloud, P., Cai, Y., YWang, and Nowinski, W., "Training and planning of interventional neuroradiology procedures - initial clinical validation," *Medicine Meets Virtual Reality (MMVR)* **85**, 96–102 (Januar 2001).
 - [7] Cotin, S., Dawson, S., Meglan, D., Shaffer, D., Ferrell, M., Bardsley, R., Morgan, F., Nagano, T., Nikom, J., Sherman, P., Walterman, M., and Wendlandt, J., "Icts, an interventional cardiology training system," in [*Studies in Health Technology and Informatics*], **70**, 59–65, ISO (2000).
 - [8] Hofer, U., Langen, T., Nziki, J., Zertler, F., Hesser, J., Muller, U., Voelker, W., and Maenner, R., *CathI, catheter instrument system* (2002). Computer Assisted Radiology and Surgery CARS Paris.
 - [9] Barnes, S., Morr, D., Oggero, E., Pagnacco, G., and Berme, N., *The realization of a haptic (force feedback) interface device for the purpose of angioplasty surgery simulation*. Biomed Sci Instrum. (1997). 94(1-2):155-9.
 - [10] Dehmeshki, J. and Amin, H., "A smooth and stable fly-through scheme for virtual navigation of 3d ct colonography," in [*3DTV07*], 1–4 (2007).
 - [11] "<http://wii.com/>," Nintendo Wii Controller.
 - [12] Phong, T., Horaud, R. P., Yassine, A., and Tao, P., "Object pose from 2d to 3d point and line correspondences," *International Journal of Computer Vision* **15**, 225–243 (July 1995).
 - [13] Daniilidis, K., "Hand-eye calibration using dual quaternions," *The International Journal of Robotics Research* **18**, 286–298 (March 1999).
 - [14] Welch, G. and Bishop, G., "An introduction to the kalman filter," tech. rep., Chapel Hill, NC, USA (1995).
 - [15] Waechter, A. and Biegler, L., "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," 25–57 (2006).