

Tracking and Control for Handheld Surgery Tools

Gontje C. Claasen and Philippe Martin
Centre Automatique et Systèmes, Mines ParisTech
60 bd St-Michel, 75006 Paris, France
Email: {caroline.claasen, philippe.martin}@mines-paristech.fr

Frederic Picard
Department of Orthopaedics
Golden Jubilee National Hospital
Beardmore Street, Glasgow, UK

Abstract—Handheld tools could be a valuable improvement to today's computer-assisted surgery systems. For tracking such a tool, we propose a high-bandwidth optical-inertial tracking system which is lightweight and low-cost. In a simulation, we show the impact of the bandwidth of the tracking system and the use of inertial sensors on the performance of the servo-control. We present an Extended Kalman Filter to fuse sensor data with a low-latency approach. A test with an experimental setup shows that the optical-inertial system does indeed follow human motion correctly and faster than an optical tracking system with a low bandwidth as found in commercially available systems.

I. INTRODUCTION

Many of the computer-assisted surgery systems available today use an image-free tracking system to acquire patient data intra-operatively. These are also called optical tracking systems and consist of two or more infrared cameras and markers which are fixed to the patient and tools. In orthopedic surgery systems, e.g. for knee replacement, the system acquires relevant patient anatomical landmarks and calculates appropriate prosthesis placement based on built up frame of reference. It then defines the desired cutting planes for the knee implant.

Cutting jigs are fixed to the patient's bone in accordance with the desired cutting planes. They guide the bone saw mechanically with good accuracy but mounting the jigs takes time and they have to be pinned to the bone. Using a handheld saw without any cutting jigs would have several advantages: the procedure would be less invasive, demand less surgical material and save time. Obviously, it would have to produce cuts with the same or even better accuracy to be a valuable improvement.

While a robotic system could achieve this task of cutting along a desired path, many surgeons wish to keep control over the cutting procedure. Therefore, an intelligent handheld tool should be used which combines the surgeon's skills with the accuracy, precision and speed of a computer-controlled system. Such a tool should be small and lightweight so as not to impede on the surgeon's work, compatible with existing computer-assisted surgery systems and relatively low-cost.

Controlling the tool position and keeping it along the desired cutting path necessitate the following steps: 1. define desired cutting plane relative to the patient, 2. track tool position and orientation relative to the patient and 3. compare desired and actual positions and correct the tool position accordingly. The first step is done by the surgery system and the second by a

tracking system. The handheld tool has to be able to carry out step 3.

Several handheld tools have been developed in recent years. In [1], the patient's bone and a handheld saw are tracked by an optical tracking system and the actual and desired cutting planes are shown on a screen. A robotic arm is used in [2] to maintain the tool orientation and correct deviations. So-called "intelligent" handheld tools which are able to correct deviations from the desired cutting plane automatically without additional material like robotic arms are presented in [3] and [4].

The handheld tool we consider here is supposed to be an extension for an image-free or image-based computer-assisted surgery system, hence it can make use of an optical tracking system but not of an active robotic arm. The tool is to be servo-controlled with motors which can change the blade position. We use a saw as an example but the same applies to drilling, pinning or burring tools.

The tracking system and particularly its bandwidth are a key for good performance of the servo-control. Firstly, the tracking system should be able to follow human motion and especially fast movements - these could be due to a sudden change of bone structure while cutting or to slipping of the surgeon's hand. These are the movements to be corrected. Secondly, it should be fast enough to let the servo-control make the necessary corrections. The faster the correction, the smaller the deviation will be. Finally, the servo-control should make the correction before the surgeon notices the error to avoid conflict between the control and the surgeon's reaction. We estimate the surgeon's perception time to be of 10ms which corresponds to a frequency of 100Hz and therefore consider the necessary bandwidth for the tracking system to be 200Hz. Commercially available optical tracking systems suitable for surgery have a bandwidth of only 60Hz.

We propose an optical-inertial tracking system which combines an optical tracking system with inertial sensors. These inertial sensors have a high bandwidth and are suitable for use in an operating room. In contrast to other systems using these sensors, we do not try to solve the line-of-sight problem. Our goal is a tracking system with a high bandwidth and low latency, i.e. the tracking values should be available with very little delay. This requires an algorithm which is particularly adapted to this problem and which reduces latency compared to similar systems as presented in [5] or [6] for example.

In Sect. II, we present a simulation of a simple model of

G.C. Claasen's work is funded by an AXA Research Fund doctoral grant.

a handheld tool and show the influence of different tracking systems on the correction of a deviation from the desired cutting plane. Section III presents the proposed optical-inertial tracking system and algorithm for a complete model for the motion of the tool. An experimental implementation and results can be found in Sect. IV.

II. SERVO-CONTROL FOR HANDHELD TOOLS

A handheld tool which can autonomously change the position of its tip (i.e. of the blade) does so with the help of servo-motors which control the motion of the tip relative to the tool. This technique is called servo-control. It can be used to correct small deviations from the desired position. The action of the servo-motors is determined by the error between the actual and the desired position. The latter is defined by the surgery system. The actual position is provided by a tracking system. The servo-control compares the desired and actual path and actuates the motors accordingly.

We are going to show the effect of a higher bandwidth in a Matlab/Simulink simulation using a simple model. The tool in Fig. 1 consists of a handle and a blade connected by a gearing mechanism which is actuated by a motor. The goal is to cut in y direction at a desired position z_r . The surgeon moves the tool in y direction at a speed of 0.5cm/s. A deviation from the desired z_r due to a change of bone structure is modeled by a disturbance D acting along z . In this simple model we assume the tool's motion to be constrained along z , except for the cutting motion which runs along the y axis (the blade oscillates along x). The blade position z is determined by $z = R\theta + z_0$ and $m\ddot{z} = F + D + mg$ where R is the radius of the gear wheel, θ the wheel's angular position, z_0 the handle position, F the force applied by the gear, m the mass of the subsystem carrying the blade and g is gravity. The motor is governed by $J\dot{\theta} = U - RF$ where J is the motor and gear inertia and U the control input. Combining these equations gives

$$\ddot{z} = \underbrace{\frac{U}{mR + J/R}}_u + \underbrace{\frac{D}{m + J/R^2} + \frac{\ddot{z}_0 - g}{1 + mR^2/J}}_d + g. \quad (1)$$

This yields the simplified model $\dot{z} = v$, $\dot{v} = u + d + g$. d includes the disturbance D due to bone structure as well as disturbances due to the surgeon motion (modeled by \ddot{z}_0). An optical tracking system measures the position $z_m = z$ with a frequency of $1/T = 50\text{Hz}$ at discrete instants $z_{m,k} = z_m(kT)$; an inertial sensor (accelerometer) measures $a_m = u + d + a_b$ where a_b is the accelerometer constant bias. The inertial measurements are considered continuous because their frequency is much higher than that of the optical ones. We do not take into account any measurement noise.

We now present three systems using different types of measurements in a standard servo-control design. In all cases h , L , l and K are appropriately calculated constant gains; d is modeled as a constant.

System 1 uses only optical measurements $z_{m,k}$. An observer estimates the state $x = [z, v, d + g]^T$:

$$\begin{aligned} \text{prediction: } \dot{\hat{x}}^- &= [\hat{v}, u + \widehat{d} + g, 0]^T \\ \text{correction: } \hat{x}_k &= \hat{x}_k^- + L(z_{m,k-1} - \hat{z}_{k-1}) \end{aligned} \quad (2)$$

where $\hat{x}_k^- = \int_{kT-T}^{kT} \dot{\hat{x}}^-(\tau) d\tau$ with $\hat{x}^-(kT - T) = \hat{x}_{k-1}$. The controller reads $u_k = -K\hat{x}_k + hz_r$. This system corresponds to the case where only an optical tracking system is used.

System 2 uses both optical and inertial data. A first observer with state $x = [z, v, a_b - g]^T$, measured input a_m and optical measurements $z_{m,k}$ reads:

$$\begin{aligned} \text{prediction: } \dot{\hat{x}}^- &= [\hat{v}, a_m - \widehat{a_b} - g, 0]^T \\ \text{correction: } \hat{x}_k &= \hat{x}_k^- + L(z_{m,k-1} - \hat{z}_{k-1}) \end{aligned} \quad (3)$$

where $\hat{x}_k^- = \int_{kT-T}^{kT} \dot{\hat{x}}^-(\tau) d\tau$ with $\hat{x}^-(kT - T) = \hat{x}_{k-1}$. This observer gives a continuous estimation $\hat{z}(t)$ which is used as a measurement $z_m(t)$ for a second observer with state $\tilde{x} = [\tilde{z}, \tilde{v}, \tilde{d} + g]^T$:

$$\dot{\hat{\tilde{x}}}(t) = [\hat{\tilde{v}}, \widehat{\tilde{d}} + g + u, 0]^T + L(z_m(t) - \hat{\tilde{z}}(t)) \quad (5)$$

The controller equation is $u = -K\hat{\tilde{x}} + hz_r$.

In system 3 which uses optical and inertial data we suppose that tracking and control are more tightly coupled than in system 2. A first observer is used to estimate the disturbance d with inertial measurements $a_m = u + d + a_b$:

$$\dot{\widehat{d + a_b}} = l(a_m - u - \widehat{d + a_b}) \quad (6)$$

This observer gives a continuous estimation $\widehat{d + a_b}(t)$ which is used as input for the second controller-observer. Its state is $x = [z, v, a_b - g]^T$ and it uses optical measurements $z_{m,k}$:

$$\text{prediction: } \dot{\hat{x}}^- = [\hat{v}, u + \widehat{d + a_b} - \widehat{a_b} - g, 0]^T \quad (7)$$

$$\text{correction: } \hat{x}_k = \hat{x}_k^- + L(z_{m,k-1} - \hat{z}_{k-1}) \quad (8)$$

where $\hat{x}_k^- = \int_{kT-T}^{kT} \dot{\hat{x}}^-(\tau) d\tau$ with $\hat{x}^-(kT - T) = \hat{x}_{k-1}$. The control input is $u_k = -K\hat{x}_k + hz_r - \widehat{d + a_b}$.

System 2 represents the setup we propose in Sect. III. System 3 employs the same hardware setup but a more complex algorithm using a model of the tool.

Figure 1 shows the simulated cuts for the three systems for a desired cutting position $z_r = 0\text{cm}$ and a disturbance d . This disturbance simulates a sudden change of bone structure or slipping of the surgeon's hand and occurs from $t = 2.002\text{s}$ to $t = 2.202\text{s}$, i.e. from $y = 0.5\text{cm/s} \cdot t = 1.001\text{cm}$ to $y = 1.101\text{cm}$.

The disturbance causes the largest and longest deviation in the first system. In system 2, the position deviation can be corrected much faster and its amplitude is much smaller. System 3 can correct the deviation even better. This simulation shows that using inertial sensors with a higher bandwidth allows the servo-control to correct a deviation caused by a disturbance much better than a system with a low bandwidth such as an optical tracking system.

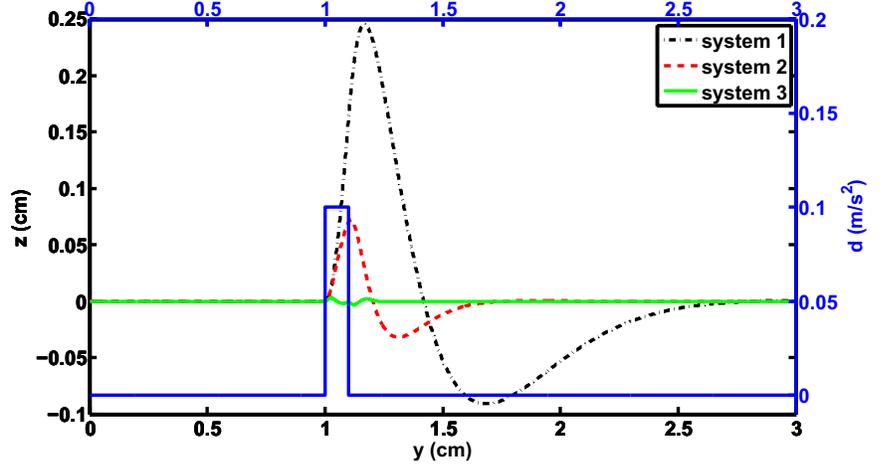
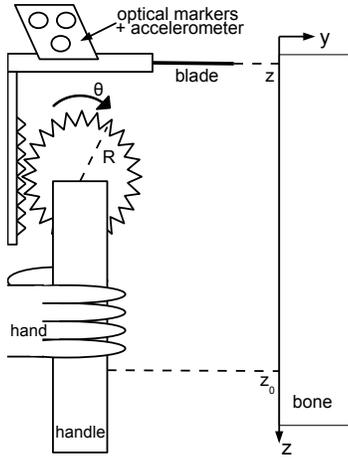


Fig. 1. Handheld tool (left) as used for the simple model and simulation results (right)

It is important to note that the controller-observer for system 1 cannot be tuned to reject the disturbance faster; the choice of K and L is constrained by the frequency of the optical measurements.

III. PROPOSED TRACKING SYSTEM

We have developed an optical-inertial tracking system which consists of a stationary stereo camera pair and a sensor unit. The sensor unit is made of an inertial measurement unit (IMU) and three optical markers and is attached to a handheld tool. We want to use this tracking system to calculate the position and orientation of the sensor unit relative to the cameras. This is achieved by a data fusion algorithm using a model of the system and the data from the cameras and the IMU.

The setup corresponds to the tracking used in system 2 in Sect. II but now we use a complete model for the tool/sensor unit motion in 3D. In this Section, we present the model we employ and our data fusion algorithm.

A. Mathematical Model

1) *Coordinate Systems*: The motion of the sensor unit will be expressed in camera coordinates which are denoted by C and are fixed to the right camera center. Their unit vectors are $E_1 = [1, 0, 0]^T$, $E_2 = [0, 1, 0]^T$ and $E_3 = [0, 0, 1]^T$. The camera's optical axis runs along E_1 . Image coordinates are expressed in the image sensor coordinate system S which is attached to one of the corners of the camera's image sensor. The left camera coordinate system is denoted by CL and the image sensor coordinate system by SL . The left camera unit vectors are \hat{E}_1 , \hat{E}_2 and \hat{E}_3 . Coordinates C and CL are related by a constant transformation (R_{St}, t_{St}) . The body coordinates, denoted by B , are fixed to the origin of the IMU frame. The world coordinate system is Earth-fixed and denoted by W .

2) *Dynamics and Output Model*: We consider the following state variables: sensor unit position ${}^C p$, velocity ${}^C v$ and quaternion ${}^{BC} q$ (representing orientation), accelerometer bias

${}^B a_b$ and gyroscope bias ${}^B \omega_b$. The dynamics equations read

$${}^C \dot{p} = {}^C v \quad (9)$$

$${}^C \dot{v} = {}^C G + {}^{BC} q * (a_m - \nu_a - {}^B a_b) * {}^{BC} q^{-1} \quad (10)$$

$${}^{BC} \dot{q} = \frac{1}{2} {}^{BC} q * (\omega_m - \nu_\omega - {}^B \omega_b) \quad (11)$$

and the biases obey

$${}^B \dot{a}_b = \nu_{ab} \quad (12)$$

$${}^B \dot{\omega}_b = \nu_{\omega b} \quad (13)$$

where ${}^C G = {}^W C q * {}^W G * {}^W C q^{-1}$ is the gravity vector expressed in camera coordinates. ${}^W G = [0, 0, g]^T$ is the gravity vector in the W frame and ${}^W C q$ describes the (constant) rotation from world to camera coordinates. Quaternion multiplication is denoted by $*$. ${}^B a_m$ and ${}^B \omega_m$ are the measured accelerations and angular velocities which are considered as the system's inputs. They are corrupted by noises ν_a and ν_ω assumed white and with constant biases ${}^B a_b$ and ${}^B \omega_b$.

The outputs are the marker images. We use a standard pin-hole model to project the marker positions to the cameras [7]. The measured output for the right camera for the i th marker in S coordinates reads

$$y_{im} = \frac{f_R}{\langle {}^C m_i, {}^C E_1 \rangle} \left[\frac{\langle {}^C m_i, {}^C E_2 \rangle}{\langle {}^C m_i, {}^C E_3 \rangle} \right] + s_{uR} + \eta_{yi} \quad (14)$$

where f_R is the focal distance and s_{uR} the principal point of the right camera. $\langle a, b \rangle$ denotes the scalar product of vectors a and b . The measurement is corrupted by noise η_{yi} . The position of marker i is calculated according to ${}^C m_i = {}^C p + {}^{BC} q * {}^B m_i * {}^{BC} q^{-1}$ using the known marker position ${}^B m_i$ in body coordinates. The output for the left camera is obtained analogously to (14) using the transformation ${}^{CL} m_i = R_{St} {}^C m_i + t_{St}$.

B. Data Fusion Algorithm

We propose to use an extended Kalman filter (EKF) to fuse inertial and optical data and obtain an estimation of the sensor

unit position and orientation. Since a quaternion has to be estimated, we have to make a modification to the standard EKF to preserve the unit norm. For the quaternion, we use the correction term $\hat{q}K_q(y-\hat{y})$ and an error quaternion $e_q = \hat{q}^{-1} * q$. This gives the so-called Multiplicative EKF (MEKF) [8]. The MEKF for our system (9)–(14) reads

$${}^C\dot{\hat{p}} = {}^C\hat{v} + K_p E \quad (15)$$

$${}^C\dot{\hat{v}} = {}^C G + {}^{BC}\hat{q} * (a_m - {}^B\hat{a}_b) * {}^{BC}\hat{q}^{-1} + K_v E \quad (16)$$

$${}^{BC}\dot{\hat{q}} = \frac{1}{2} {}^{BC}\hat{q} * (\omega_m - {}^B\hat{\omega}_b) + {}^{BC}\hat{q} * K_q E \quad (17)$$

$${}^B\dot{\hat{a}}_b = K_a E \quad (18)$$

$${}^B\dot{\hat{\omega}}_b = K_w E \quad (19)$$

with output error $E = y_m - \hat{y}$. We consider the state error $e = [\hat{p} - p, \hat{v} - v, {}^{BC}\hat{q}^{-1} * {}^{BC}q, \hat{a}_b - a_b, \hat{\omega}_b - \omega_b]$. The error system linearized around $\bar{e} = (0, 0, 1, 0, 0)$ satisfies

$$\Delta\dot{e} = (A - KC)\Delta e - M\nu + KN\eta \quad (20)$$

up to higher order terms where $K = [K_p, K_v, K_q, K_a, K_w]$, $\nu = [\nu_a, \nu_\omega, \nu_{ab}, \nu_{wb}]$, $\eta = \eta_y$ and A , C , M and N depend on the estimated state. This permits us to calculate the gain K as in a standard extended Kalman filter: $K = PC^T R^{-1}$ where P satisfies $\dot{P} = AP + PA^T + Q - PC^T R^{-1} CP$. $Q = M\tilde{Q}M^T$ and $R = N\tilde{R}N^T$ where \tilde{Q} and \tilde{R} contain white noise intensities.

The main difference of our approach compared to other systems fusing optical and inertial data lies in the choice of the output in the system model in Sect. III-A.2 and in the MEKF above. Other systems [5], [6] use marker positions ${}^C m_i$ which have been calculated by an optical tracking system as output measurements. Since pose estimation from optical data demands complex computations, this produces an important latency. Our approach using marker images directly as outputs reduces this latency.

IV. EXPERIMENTAL SETUP AND RESULTS

Our sensor unit is made of an ADIS16355 IMU and three infrared LEDs. The camera system consists of two Wiimote image sensors fixed in a stereo rig. The Wiimote is the remote control of Nintendo's game console. Its image sensor sees up to four luminous points. To use this sensor on its own, we unsoldered it [9]. Data from both sensors were acquired by an Atmega2560 microcontroller with a camera sample rate of 16.7Hz and an IMU sample rate of 250Hz. The data were processed offline with Matlab/Simulink.

In our experiment, the sensor unit was set on a horizontal surface and then moved quickly by hand mainly along the y axis. This experiment represents a small unintentional motion of a surgeon holding a handheld tool. The experimental data was fed to the MEKF which estimated the sensor unit position and orientation. To evaluate our results and compare them to optical tracking, we used only optical data from the same set to calculate the pose following [7] and [10].

Figure 2 shows the experimental results. Optical tracking alone detects the motion later than the optical-inertial system.

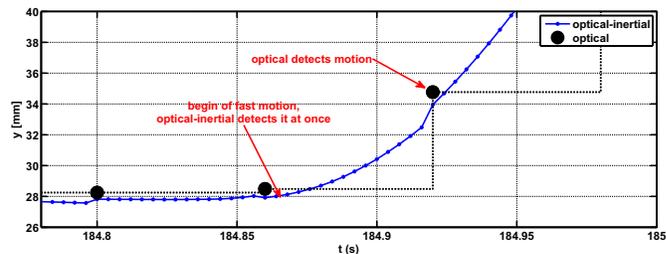


Fig. 2. y coordinate of the estimated position for optical tracking (dashed) and optical-inertial tracking (solid). Each circle on a line indicates a new estimated value.

This is in accordance with Sect. II where system 2 estimates (and rejects) the disturbance faster than system 1.

V. CONCLUSION

We considered the use of a servo-controlled handheld tool for cutting/drilling/burring in a computer-assisted surgery system which corrects errors relative to a desired path. In a simulation with a simplified model we showed that this kind of servo-control needs a tracking system with a high bandwidth and that optical tracking with its bandwidth of 50Hz has limited performance. For the purpose of tracking a handheld tool we proposed an optical-inertial tracking system which has a high bandwidth and low latency. Data obtained from our experimental setup shows that the optical-inertial system detects motion faster than a purely optical tracking system.

REFERENCES

- [1] H. Haider, O. A. Barrera, and K. L. Garvin, "Minimally invasive total knee arthroplasty surgery through navigated freehand bone cutting: Winner of the 2005 HAP Paul Award," *J Arthroplasty*, vol. 22, no. 4, pp. 535 – 542, 2007.
- [2] P. Knappe, I. Gross, S. Pieck, J. Wahrburg, S. Kuenzler, and F. Kerschbaumer, "Position control of a surgical robot by a navigation system," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems*, 2003, pp. 3350–3354.
- [3] G. Brisson, T. Kanade, A. M. D. Gioia, and B. Jaramaz, "Precision free-hand sculpting of bone," in *Medical Image Computing and Computer-Assisted Intervention*, ser. LNCS, vol. 3217. Springer, 2004, pp. 105 – 112.
- [4] G. Kane, G. Eggers, R. Boesecke, J. Raczkowski, H. Wörn, R. Marmulla, and J. Mühlhling, "System design of a hand-held mobile robot for craniotomy," in *Medical Image Computing and Computer-Assisted Intervention*, ser. LNCS, vol. 5761. Springer, 2009, pp. 402–409.
- [5] A. Tobergte, M. Pomarlan, and G. Hirzinger, "Robust multi sensor pose estimation for medical applications," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems*, 2009, pp. 492 – 497.
- [6] B. Hartmann, N. Link, and G. F. Trommer, "Indoor 3D position estimation using low-cost inertial sensors and marker-based video-tracking," in *IEEE/ION Position Location and Navigation Symposium*, 2010, pp. 319 – 326.
- [7] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge University Press, 2003.
- [8] J. L. Crassidis, F. L. Markley, and Y. Cheng, "Survey of nonlinear attitude estimation methods," *J Guid Control Dynam*, vol. 30, no. 1, pp. 12–28, 2007.
- [9] U. Jürss and W. Rudolph, "Sur la trace des hot-spots," *elektor*, vol. 11, pp. 64–69, 11 2008.
- [10] S. Umeyama, "Least-squares estimation of transformation parameters between two point patterns," *IEEE T Pattern Anal*, vol. 13, no. 4, pp. 376 – 380, 1991.