

Concept of a mobile, cameraless VR-controller framework for working environments

Patrik Reizinger,¹ Ferenc Vajda¹

¹ Department of Control Engineering and Information Technology,
Budapest University of Technology and Economics, Budapest, Hungary

Abstract

In this paper the currently used approaches regarding VR-controllers are reviewed, and a concept based on low-cost inertial measurement units (IMUs) is described for determining orientation and position. Due to the lack of a widely-used supportive optical framework the discussed system has a higher degree of mobility although it implies several difficulties regarding position estimation. A prototype assembled along the described principles is presented, and the measurements are analyzed with proposals for future work.

Categories and Subject Descriptors (according to ACM CCS): [Human-computer interaction (HCI)]: Interaction devices

1. Introduction

In the recent years several different concepts were established for creating virtual reality controllers and frameworks, many of which are available as a commercial product, such as the HTC Vive or the Oculus Rift. The former uses the approach called *inside-out*, where cameras integrated into the HMD (*Head Mounted Display*) are responsible for position and orientation estimation whilst in the case of the Oculus Rift a separate camera unit serves these purposes, which is called *outside-in*.

It is clearly to see that outside-in systems have predestinably a less flexible design due to the presence of a detached processing unit. In the case of inside-out systems the situation is slightly better, but problems of the same origin still exist as in the case of inside-out solutions, namely, although the localization of the HMD is settled such as supposing the Microsoft HoloLens, but to solve the same problem for controllers, camera systems are still required simultaneously with IMUs (see HTC VIVE's Lighthouse technology).

The newest results show that with optical (generally using IR LEDs) systems a precision satisfying the needs of the most users can be achieved, as with the DodecaPen¹, an optical tracker developed by Facebook Research with sub-millimeter accuracy, used for 3D drawing applications.

Since the advantages of virtual, augmented and mixed reality are being discovered by a wide range of scientific/industrial professionals, implementing a new "all-round" device could be of benefit in accomplishing certain tasks in a more intuitive way than before.

Domains of the highest significance regarding VR-innovation are among others the medical industry. A unique approach for human-computer interaction is presented in ², the authors have realized an interaction device with the fusion of an IMU and two sEMG sensors. An application of high practical importance is the virtual reality framework for the manpower-intensive work of neuron tracing ³.

Furthermore, applications in the industrial field are getting more frequent, because these frameworks represent a source of information, a way for teleoperation and thus a gain of efficiency, such a solution is proposed in ⁴ for locating infrastructure in the field helping service operators.

Walks of life such as entertainment could also not being left untouched by the new possibilities, the way of evolution was quite straightforward, virtual reality served as an instrument for enhancing user experience in games, such as in ⁵ it is done for FPS (*First Person Shooter*) computer games, or in ⁶ for hybrid television based on augmented reality.

The appearance and fast spreading of smart mobile de-

vices have created new preferences on the market, which should be followed in the virtual reality industry, namely the mobility is a key factor recently, and the trends are showing that a significant increase standing before us is likely.

1.1. Inertial Motion Tracking

To accommodate these needs compromises should be made, because without a supportive framework the accuracy can not reach the same levels as with them - mostly due to the drift introduced by the low-cost MEMS *Microelectromechanical Systems* accelerometers, the compensation of which is more complicated than general signal processing problems, due to the fact that the motion sensed by the sensors does not show any signs of periodicity in general.

Depending on the application, one may use 6 or 9 DoF (*degrees of freedom*) IMUs, which consist of 3-axis accelerometers, magnetometers and gyroscopes. For determining absolute orientation and relative position, a 9 DoF sensor array is needed.

For determining absolute orientation several algorithms exist, considering the Kálmán-filter⁷ as an industry-level standard. Although the algorithm proposed by Rudolf Kálmán and its nonlinear successors provide satisfying results in many cases, they do not exactly fit in the low computational cost and low energy consumption environment the mobile devices of nowadays have created.

To revise the above mentioned disadvantages of the Kálmán-filter, several new algorithms were developed with the new guidelines. Mahony et al.⁸ have proposed a quaternion-based complimentary filter, which served as a basis for the algorithm proposed by Madgwick et al.⁹, which is also quaternion-based, and thus computationally less expensive, and because of the fact that it does not use a state space approach as in ⁷, it is considered more suitable for embedded systems.

These features have decided to use the Madgwick-algorithm in this project, the source code of which is freely available on GitHub.¹⁰

2. System architecture

This project is based on the results of a previous project of the Department of Control Engineering and Informatics (Budapest University of Technology and Economics), which have provided the hardware components extremely useful for experimenting¹¹.

2.1. Hardware

The realized system consists of two separate units, the base station and the input device called VR pen.

The reasons for choosing this kind of architecture are the

following. Because mobility and flexible use is of highest importance, the VR pen itself should have a minimal power consumption to maximize battery lifetime. The key components in the input device are an 8-bit PIC microcontroller, an SPI RF transceiver in the 433 MHz ISM frequency band from HOPE RF, a 9 DoF sensor array from ST Microelectronics and two pushbuttons for further interaction.

Thus the task of the VR pen is restricted to handling the sensor interface and communicating with the base station.

For the prototype a plastic case was designed and fabricated with 3D printing technology, the design of which resembles one of a pen and thus provides an intuitive way of handling for the users.



Figure 1: Prototype of the VR pen.

From the system's point of view the base station acts as an intermediary unit between the operating system of the PC and the VR pen, it provides a Virtual COM interface towards the computer. Although the task of the base station, which include data preprocessing regarding both orientation and position data, could have been replaced by the PC itself, and so an RF transceiver module (with a lightweight controller) would have been sufficient for the task, but the proposed architecture provides an encapsulation and a uniform interface. For accomplishing these tasks, the base station operates with a 16-bit PIC microcontroller, and also it has an RF transceiver for communication with the VR pen.

2.2. Software

As far as the software architecture is concerned, it reflects the complexity of the underlying hardware. Namely, the input device has a simple structure in software too. The communication occurs with the sensor array via I²C, and via SPI with the transceiver. Additionally, the handling of the buttons is done on the 8-bit PIC including debouncing.

The base station implements the communication with the PC and the VR pen (USB and RF, respectively), but the main

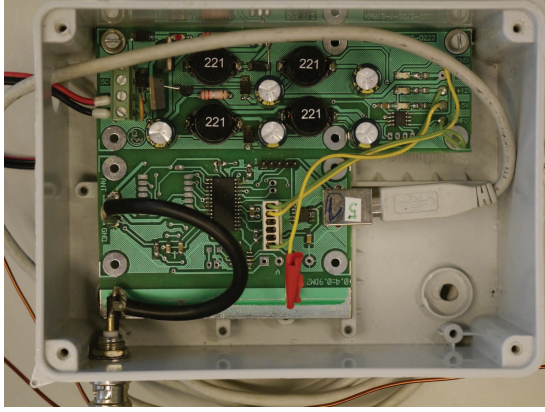


Figure 2: The base station used in the project.

software components are the algorithms responsible for orientation and position filtering. The former is the one proposed by Madgwick et al. (Figure 3), with minor improvements for extracting Earth acceleration, i.e. linear acceleration (which does not contain the contribution of gravity) in the Earth's reference frame. The main idea of the algorithm is to create a reference frame from the magnetometer and accelerometer measurements with the help of a cost function optimization problem, where the cost function describes the difference from the reference orientation.

$$\mathbf{C}() = {}^S_E \hat{q}^* \otimes {}^E \hat{d} \otimes {}^S_E \hat{q} - {}^S \hat{d}$$

The indices in ${}^1 E, S$ are referring to the Earth/sensor coordinate frame respectively, the q quaternions are describing the rotation of the ${}^E \hat{d}$ vector from Earth into the sensor frame, from which a reference vector ${}^S \hat{d}$ will be subtracted to get the value of the cost function. The factor calculated with the gradient method is used to compensate the drift of the gyroscope, which gives the estimated rate of change of the orientation (the estimated angular velocity), which is numerically integrated for the orientation estimate.

In the case of position estimation, an adaptive exponential averaging algorithm (Figure 4) was implemented based on the absolute value of the acceleration vector. A second level of filter is supervising the velocity values, for due to accelerometer noise the velocity values do not cancel out as the device gets to a motionless state. Additionally the filter has a built-in hysteresis, which basically helps to make more dynamic movements smoother, because the smaller λ value will be used in (Equation 1) which prevents that during more intense moves the filter parameter changes to 0.3.

$$a_{exp,t} = a_t + \lambda a_{exp,t-1}$$

The communication protocol of the system implements a

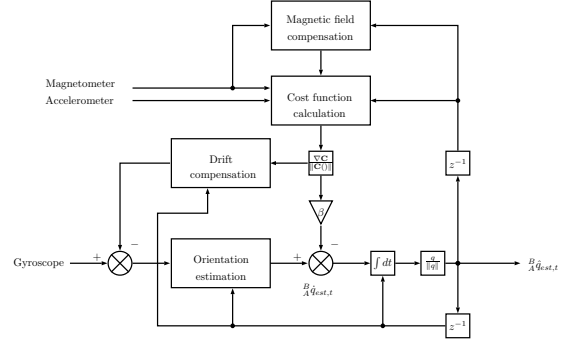


Figure 3: The structure of the Madgwick-algorithm.

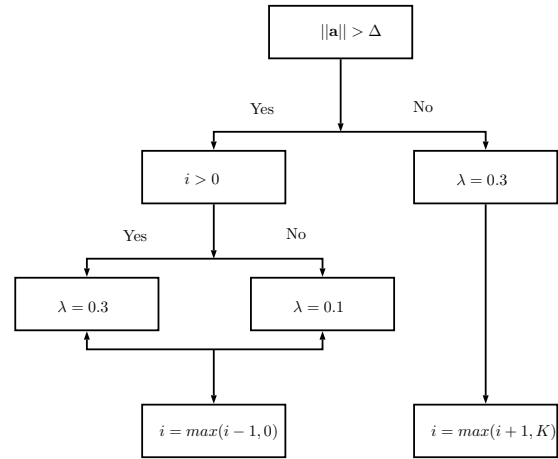


Figure 4: State machine of the position filter.

standard, stacked structure, with frames as the unit of communication. Transaction errors are detected via a one-byte checksum - although the used ISM frequency band has a quite intense traffic level, experiments with two-byte CRC error detection did not provide better results.

For the communication is two-directional, but regarding the data transmitted it is asymmetric and the applied transceivers make only half-duplex communication possible, the protocol does not contain acknowledgement signals or packets.

The problem of packet loss is handled in another way: first of all, due to the asymmetry mentioned above, a bulk transmission command has been introduced which enables the VR pen to transmit a configurable amount of frames. Right after the base station sends a bulk transmission packet for the VR pen, a timer is set up, conceptually a watchdog for packet loss. If a timeout occurs before the reception of the first packet, a command retransmission occurs immediately. In order to not to lock the base station into an infinite loop, the number of retransmissions is limited, on crossing the threshold an error message is sent for the PC letting know the user that the VR pen is unreachable.

The system consists of a visualization software for PC, which makes it possible to not only numerically evaluate the results. This part uses OpenGL as an engine and makes the orientation and position of the VR pen retraceable on the computer screen.

3. Results

Before evaluating the accuracy of the position and orientation filters, the magnetometer and the accelerometer were measured separately, and the raw data was corrected also: in the case of the magnetometer, offset-like errors termed hard-iron biases were compensated in software, whilst the white noise imposed on the accelerometer measurements could not have been compensated without the partial loss of information, because the useful signal is also wide-band. Following that, differentiated on amplitude, small signals were discarded, which made the measurements more reliable for quicker movements.

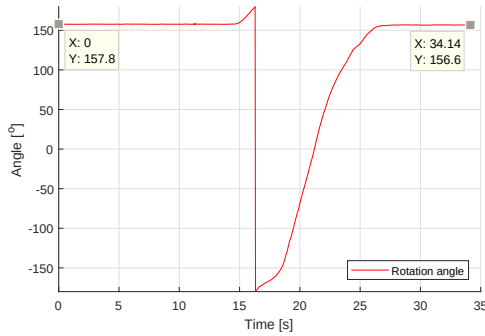


Figure 5: Angle measurements (Z-axis).

The former measurements have indicated that the z-axes have the weakest characteristics, so these were measured to obtain worst-case values. The Madgwick-algorithm was measured with the help of a rotating plate (Figure 5), a whole rotation was performed around the z-axis, which has yielded an error of about 1.2, including measurement error.

Concerning the position filter (Figure 6), the conclusion is the following: the algorithm has canceled out the accumulation of the sensor drift ($\sim 10\text{mm}/20\text{s}$), during the 1m movement in the direction of the $+z$ -axis the device shows an overshoot of 95mm (dynamic), and 39mm (static). But what seems to be a more serious issue is the fact that the end position does not equal the beginning in contrast with reality, namely it has an error of 87mm .

Although if we consider that the computer display provides visual feedback for the user, and thus a possibility to react, the problem seems in the context more manageable. Adding the fact that computer mice have in general a higher inaccuracy, one may conclude that regarding the conditions that the

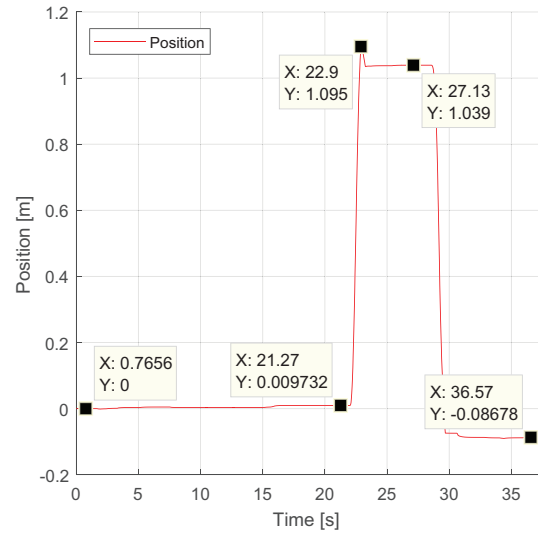


Figure 6: Position measurements (Z-axis).

realized system shows the possibility to create a 3D human-computer interaction system of full value for mobile environments.

4. Discussion and future work

In this paper the concept for a mobile VR-controller framework was evaluated, and with the help of a prototype it was stated that it is possible to fulfill the requirements towards that kind of input device.

The measurements carried out show that regarding orientation the currently existing solution are sufficient enough for most applications, while in the case of position estimation the only usage of accelerometers forces out a compromise. Although several methods exist for improving the characteristics of a MEMS accelerometer such as filtering with Wavelet-transformation¹², the current 16-bit architecture does not fit for the computational load introduced by these transformations.

Considering the factors mentioned above, the refactoring of the whole architecture is already in progress. For utilizing more complex algorithms in position estimation the core unit in the base station will be replaced with an also 16-bit microcontroller, but this time a dsPIC model will be used, to utilize its DSP-specific instruction set. The possibility will be examined whether using 8/16-bit Neural Networks could help in noise reduction and accuracy improvement.

Acknowledgments



This project was supported by the ÚNKP-17-1-I New National Excellence Program of the Ministry of Human Capacities.

References

1. P.-C. Wu, R. Wang, K. Kin, C. Twigg, Shangchen Han, M.-H. Yang, and S.-Y. Chien, "DodecaPen : Accurate 6DoF Tracking of a Passive Stylus," p. 10, 2017.
2. J. Kim, M. Kim, and K. Kim, "Development of a wearable HCI controller through sEMG & IMU sensor fusion," *2016 13th International Conference on Ubiquitous Robots and Ambient Intelligence, URAI 2016*, pp. 83–87, 2016.
3. W. Usher, P. Klacansky, F. Federer, P.-t. Bremer, A. Knoll, A. Angelucci, and V. Pascucci, "A Virtual Reality Visualization Tool for Neuron Tracing," vol. 2626, no. c, 2017.
4. A. Pena-Rios, H. Hagra, M. Gardner, and G. Owusu, "A fuzzy logic based system for geolocated augmented reality field service support," *2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pp. 1–6, 2017.
5. P. K. Krompiec and K. Park, "Enhanced player interaction using motion controllers for VR FPS," *2017 IEEE International Conference on Consumer Electronics, ICCE 2017*, pp. 19–20, 2017.
6. R. Sotelo, J. Joskowicz, and N. Rondan, "Experiences on hybrid television and augmented reality on ISDB-T," *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting, BMSB*, 2017.
7. R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Journal of Basic Engineering*, vol. 82, no. 1, p. 35, 1960.
8. R. Mahony, T. Hamel, and J. M. Pflimlin, "Complementary filter design on the special orthogonal group $SO(3)$," *Proceedings of the 44th IEEE Conference on Decision and Control, and the European Control Conference, CDC-ECC '05*, vol. 2005, no. 1, pp. 1477–1484, 2005.
9. S. O. Madgwick, A. J. Harrison, and R. Vaidyanathan, "Estimation of IMU and MARG orientation using a gradient descent algorithm," *IEEE International Conference on Rehabilitation Robotics*, 2011.
10. S. O. Madgwick, "https://github.com/xioTechnologies/Fusion."
11. G. Rác, G. Pilászy, and T. Horváth, "Tracking moving objects underwater in shallow lakes," in *Proceedings of the Workshop on the Advances of Information Technology: WAIT* (B. Kiss and L. Szirmay-Kalos, eds.), pp. 101–105, BME Irányítástechnika és Informatika Tanszék, 2015.
12. P. Wu, Y. Ge, S. Chen, and B. Xue, "De-noising algorithm based on compression of wavelet coefficient for MEMS accelerometer signal," *2010 IEEE International Conference on Information and Automation, ICIA 2010*, pp. 402–407, 2010.