A Rapidly Adaptive Collaborative Ubiquitous Computing Environment to Allow Passive Detection of Marked Objects

Hannah Slay[1], Bruce Thomas[1],
Rudi Vernik[1,2], Wayne Piekarski[1]

[1] e-World Lab, School of Computer and Information Science, University of South Australia,
Mawson Lakes, SA 5095, Australia
{Hannah.Slay, Bruce.Thomas, Rudi.Vernik,
Wayne.Piekarski}@unisa.edu.au

[2] Defence Science Technology Organisation, PO Box 1500,
Edinburgh, SA 5111, Australia
Rudi.Vernik@dsto.defence.gov.au

Abstract. This paper presents a tool to support the rapid and adaptive deployment of a collaborative, ubiquitous computing environment. A key tool for the configuration and deployment of this environment is a calibration tool to quickly and efficiently calculate the positions of cameras in a dynamic environment. This tool has been incorporated into our current Passive Detection Framework. The paper describes the context where our rapidly adaptive collaborative ubiquitous computing environment would be deployed. The results of a study to test the accuracy of the calibration tool are also presented. This study found that the calibration tool can calculate the position of cameras to within 25 mm for all lighting conditions examined.

1. Introduction

Ubiquitous computing was coined by Weiser [1] in his seminal paper "The Computer of the 21st Century" where he described the notion of computing being woven into the backdrop of natural human interactions. One of the key steps in the move towards these computing environments is a shift from the use of traditional workstations as a primary computing interface, towards environmental or workspace interfaces. In these environments the traditional input devices of mice and keyboards will be replaced with more human-centric approaches.

We are applying the rapidly adaptive collaborative ubiquitous computing environment to allow passive detection of marked object technology to the intense collaborative domain, in particular the Augmented Synchronised Planning Spaces (AUSPLANS) project. AUSPLANS is a defence domain project that focuses on distributed synchronised planning in organisations through the rapid augmentation and enablement of physical workspaces using emerging ubiquitous computing and human interaction technologies. Two key aims of AUSPLANS are to create cost effective infrastructure to augment and enable physical spaces and to create components to have a short set up time and be adaptable to suit changing user and team needs. Our tracking infrastructure provides a cost effective

tracking platform that can be used to augment traditional meeting room environments. The application Calibrate Collaborative Environment (CalibrateCE) allows the tracking infrastructure to be quickly and easily reconfigured to suit the changing requirements of users in a collaborative, ubiquitous computing environment.

We foresee the research detailed in this paper to be applicable to any intense collaboration environment where users must work collaboratively to complete time critical tasks. In organizations such as the civil disaster relief departments, there are periods of high levels of activity. Rigid and fixed infrastructure may not be the most appropriate form for such intense collaborative environments. The shape of the environment may need to change depending on the number of collaborators. The location of the environment may not be known before the emergency, and the environment may have to be placed close to the emergency. In such environments, large, multi user, ubiquitous computing workspaces may need to be employed. The ability to replicate these environments is critical for such situations.

This paper presents a tool that we have created to support the rapid and adaptive deployment of a collaborative, ubiquitous computing environment. Central to making these systems rapidly adaptive is the ability to calibrate the resources available in such an environment for use with tracking of marked objects in that workspace. CalibrateCE allows users to calculate the six degrees of freedom (6DOF) pose (translation and orientation) of cameras in an environment. It uses cameras with image based pattern recognition to calculate the 6DOF pose of a camera relative to a fixed point in the environment. Once the poses of the cameras are known, the framework can be used to determine the real time physical world pose of objects such as laptops, furniture, computer input devices affixed with markers in the environment [2]. Given the known location of calibration points in a physical coordinate system, and a trained marker, the calibration process takes approximately one minute and a half for each calibration point (where a calibration point is a known position relative to the coordinate system). The user must simply place the trained marker on a calibration point, enter the x,y,z position of the marker on a computer, then start the calibration application on the computer to which the camera to be calibrated is attached.

1.1. Universal Interaction and Control in Multi Display Environments

This research is included in an overarching project focusing on universal interaction and control in multi display environments. These environments are likely to offer a variety of display and interaction modalities, each of which is best suited to specific classes of application. Gesture based interaction, for example may be more suited to an object manipulation application than a word processing application. This project aims to investigate appropriate interaction devices to control and manipulate information across a large range of display devices, applications and tasks in the context of command and control centres [3]. In particular, this project investigates the use of information visualization in such environments [4]. A key component we require in interacting with multiple displays in a collaborative environment is a tracker. This tracker can be used to calculate the position of devices, people and information in the environment.

Consider the scenario where a tidal wave hits low level parts of a small island country. A joint planning group would be formed to determine the actions to be taken to assist the

residents and return order to the country. A hall at a local school is transformed into temporary headquarters for the group. Laptops, projectors and other personal computing devices are taken to the hall to create both large public displays and private display areas. Cameras are attached to ceilings to track the information flow (physical documents) as well as position of mobile devices in the environment. CalibrateCE is used to rapidly calibrate the work environment so the information can be tracked.

1.2. Passive Detection Framework

The Passive Detection Framework (PDF) [2] was created as an infrastructure for physical meeting rooms that can be used to rapidly augment the space and transform it into a tracked environment. The user can track an object in the environment by attaching a marker card (or fiducial marker) to the object and moving the card around the workspace. The pose of the marker card is calculated using an image based recognition library called ARToolkit [5]. Once the pose is determined, it is placed on a shared location. These tasks are carried out passively on dedicated machines so that a wide range of devices (PDAs, tablet PCs, laptops, and traditional workstations) can utilise the infrastructure without draining the resources of the device. Cameras are mounted on the ceiling to provide the most complete view of the workspace, whilst still being discrete.
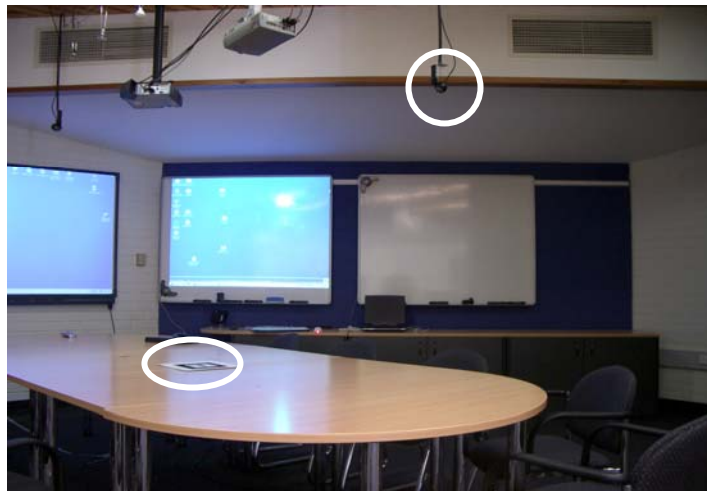

Fig. 1. Current LiveSpaces configuration. The camera and marker used in experiment are highlighted.

Unlike many tracking techniques, an advantage of the PDF is that the hardware components of the framework can be easily reconfigured to suit the requirements of the users of the workspace. Cameras can be repositioned in the environment using simple clamping mechanisms to attach to ceilings, desks etc, and computers can be relocated. For example, the default position of the cameras may be to spread them out over the entire meeting room, to provide the largest tracked volume possible. However, if a small group were to use the room they may want to reposition the cameras to give a more complete coverage

of a section of the workspace. To do this however, a mechanism must be created to allow users to quickly and efficiently calculate the pose of the cameras in the real world. CalibrateCE was created to allow users to perform this task.

Figure 1 shows the current configuration of the PDF in our environment at e-World Lab. All cameras are roof mounted using a simple clamping mechanism. This allows cameras to be moved both along the surface of the roof and moved further from / closer to the roof surface. Several of the cameras mounted in the e-World Lab can be seen in this image. There are two key steps in the process of transforming a traditional work area to a tracked working volume. Firstly the pose of cameras in the room must be calculated. Secondly using this information, the pose of marker cards can be calculated in physical world coordinates. In previous research [2], we have detailed the method in which we performed the latter of these two steps. This paper aims to describe the former step, workspace calibration.

1.3. Outline of Paper

This remainder of this paper is divided into four sections. Section 1 reviews existing research into the area of next generation work environments. This is then followed by a detailed description of CalibrateCE and its role in our adaptive collaborative ubiquitous computing environment. Section 3 section details an experiment performed to determine the accuracy of a workspace after calibration has been performed with CalibrateCE along with an analysis of the results. The final section contains conclusions and future directions for this research.

2. Related work

Several organisations and groups have been investigating the use of ubiquitous computing for future work environments. This research has two main focuses: embedding computational facilities into everyday furniture and appliances [6, 7], and creating interactive environments in which the environment can be controlled by a computer [8, 9]. We are concerned with the latter category, in particular the research being carried out by Stanford University's Interactive Workspaces (iWork) project, Massachusetts Institute of Technology's (MIT) Project Oxygen, and the University of South Australia's LiveSpaces project.

The key goal of Stanford University's iWork project is to create system infrastructure that provides fluid means of adjusting the environment in which users are working. Instead of automatically adjusting the environment of a workspace for the user (as is provided by Project Oxygen), iWork provides the user with the ability to smoothly and cleanly control their own environment[9].

The primary goal of Project Oxygen is to create a room that is able to react to users' behaviour. This is attempted by combining robotics, camera recognition, speech recognition and agent based architecture to provide intrinsic computational assistance to users in a workspace. This computation is designed to be available without the user of the computation having to shift their mode of thinking or interaction with people [8]. Project Oxygen has been investigating tracking the movement of people in the workspaces using a combi-

nation of vision based tracking[10, 11], and more hardware oriented systems such as pressure sensitive floor tiles[8]. We are not only interested in tracking the movement of users in a room, but in tracking the movement / placement of devices in the workspace.

LiveSpaces is the overarching project within e-World lab at the University of South Australia that is addressing how physical spaces such as meeting rooms can be augmented with a range of display technologies, personal information appliances, speech and natural language interfaces, interaction devices and contextual sensors to provide for future interactive/intelligent workspaces. Research is being undertaken to address how these future workspaces can be rapidly configured, adapted, and used to support a range of cooperative work activities in areas such as military command environments, large-scale software enterprises, and health systems [12]. AUSPLANS is an example of a LiveSpace as a military command environment.

For working indoors, a number of tracking technologies have been developed such as: the first mechanical tracker by Sutherland, ultrasonic trackers by InterSense, magnetic trackers by Ascension and Polhemus, and optical trackers such as the Hi Ball. These systems all rely on infrastructure to provide a reference and produce very robust and accurate results. The main limitation of most of these systems is that they do not expand over wide areas, as the infrastructure to deploy has limited range or is prohibitive in cost. Newman et al [13] describe the use of proprietary ultrasonic technology called Bats that can be used to cover large building spaces. The hybrid tracking technique described in Piekarski et al [14] operates using a number of input sources. Orientation tracking is performed continuously 3 DOF orientation sensor and indoor position tracking is performed using a fiducial marker system based on ARToolKit. The VIS-Tracker by Foxlin and Naimark [15] demonstrates the possibility of using dense fiducial markers over large indoor areas using small portable hardware. This system requires four or more markers to be within the camera's field of view for a 6DOF solution, compared to the single marker required by ARToolkit. The systems described by Newman, Piekarski, and Foxlin all require specialised hardware for each object to be tracked. Once the infrastructure of the PDF has been installed, each object only requires a new paper fiducial marker to be tracked. Each camera can track approximately twenty markers. Kato and Billinghurst's ARToolKit [5] produces reasonable results with the use of fiducial markers, and as mentioned is the underlaying tracking technology used for the PDF and CalibrateCE. This tracking does not drift over time and produces reasonably accurate results.

3. CalibrateCE

CalibrateCE is an application that allows users to quickly and easily create a tracked volume within their work environment, and then to easily reconfigure the infrastructure components (computers, hubs, cameras) to suit their changing requirements. In order to do this, CalibrateCE allows users to efficiently calibrate the features of their work environment (camera poses, room attributes) and store this data in a shared location for use by other applications at a later time. This data is accurate until a camera is moved in the work environment, at which point the workspace must be recalibrated. The output of CalibrateCE is used by the PDF to calculate the position of the markers in the physical world.

We calculate the physical world pose of the camera by calculating the pose of the camera with respect to the known location of a fiducial marker, and then factor in the pose of the marker in the physical world. Figure 2 shows the transformation of the camera in physical world coordinates. Arrows on axes show the positive direction of each dimension. To calculate the 3x4 transformation matrix C, the inverse of the transformation between the marker and the camera T, must be multiplied by the rotation in coordinate systems between the marker and the physical world R, and then multiplied by the transformation between the marker and the origin of the physical coordinate system M.
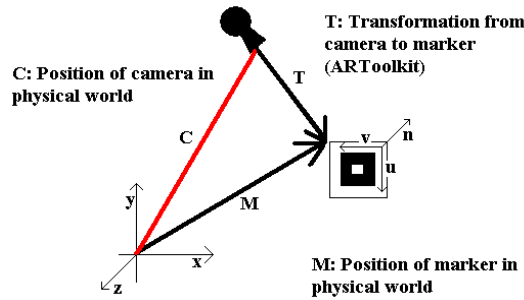

Fig. 2. Transformation of Camera in Physical World

CalibrateCE uses the PDFs distributed system consisting of two components: node computers and a combiner computer. Cameras are attached to the node computers and an ARToolkit based application [5] is executed per camera attached to the computer. The application calculates the transformation of the marker in the camera coordinate system and sends this 3x4 transformation matrix to the combiner computer. This calculation is performed and sent in a UDP packet 500 times to attempt to overcome factors such as uneven lighting. The combiner computer receives the UDP packets and calculates an average pose for each camera using the quaternion form of each rotation matrix. Having a distributed architecture makes CalibrateCE easy to extend to either track a larger volume or to provide a more complete coverage of the currently tracked volume. In both cases, an extra node computer can be added to the framework.

The output of CalibrateCE is a XML file containing user defined fields (such as a name for the environment), and an element for each computer that sent UDP packets to the combiner. Each computer element has sub-elements that detail the node number and pose of each camera. This XML file is placed in a shared location for use by other applications, such as PDF.

4.   Experimental Results

This section provides an overview of an experiment we have undertaken to determine the accuracy to which the CalibrateCE application can calculate the pose of a fixed camera in an environment. The work environment that the experiment was performed in has two separate lighting systems, fluorescent lights and halogen down lights. The fluorescent lights are divided into two sections, one section runs around the perimeter

of the ceiling, and the other runs in the middle of the ceiling. We refer to these as the outside and inside fluorescent lighting systems respectively. The inside fluorescent lighting system contains one quarter as many lights as the outside fluorescent lighting system. The halogen down lights are a group of 5 small directional lights, all of which are positioned to point vertically down from the ceiling to the floor. The position of the camera was calculated under four different lighting conditions: inside fluorescent lighting, outside fluorescent lighting, both fluorescent lighting, and down lights only. Each of the four tests involve the calculation of the position 500 times. Section 4.1 and 4.2 contain overviews of the positional and orientation results, along with Tables 1 and 2 which provide the results in a condensed format. Section 4.3 will provide an analysis of these results.

During these tests, four 150 mm square markers were trained but only one marker was ever visible in the video frame at one time. The four markers were trained so the program would have a large decision space to choose from. If only one marker was trained, this would result in multiple false positives in recognising markers. To reduce the effect of uneven lighting, all markers used were made of felt. The calculations for the position of the camera in all lighting conditions were performed with the marker in the same place for each test. This was done to attempt to minimize the human introduced errors in calculations (only one measurement of the position of the marker was taken and the marker remained in the same position throughout the execution of the experiment). The distance between the camera and the marker was 1645 mm.

4.1. Positional Results

Table 1 shows results obtained after performing the configuration under each of the four lighting conditions. Shown are the minimum, maximum, mean and standard deviation calculated over the 500 iterations for each lighting condition. All measurements are taken in room coordinates, where the centre of the transformation matrix space was in the corner of the room. The measured position of the camera is 3320, 2080, 1680 (x, y, z respectively). The minimum and maximum values for x, y and z were calculated separately. They therefore show the range of values calculated for each of the dimensions.

Table 1. Calculated positions of camera under four lighting conditions.

All	x (mm)	y (mm)	z (mm)	Out	x (mm)	y (mm)	z (mm)
Min	3270.41	2044.57	1618.69	Min	3264.09	2048.36	1612.73
Max	3315.53	2109.1	1737.79	Max	3313.26	2091.73	1756.15
Mean	3300.45	2065.77	1682.33	Mean	3299.18	2067.32	1687.74
St Dev	4.66	8.47	22.46	St Dev	4.48	7.96	24.98
Inside	x (mm)	y (mm)	z (mm)	Down	x (mm)	y (mm)	z (mm)
Min	3233.61	2037.01	1601.09	Min	3235.13	2032.62	1574.46
Max	3322.8	2207.35	1910.27	Max	3335.16	2201.51	1897.43
Mean	3300.74	2066.51	1686.61	Mean	3306.01	2062.15	1682.19
St Dev	5.91	11.34	27.59	St Dev	6.77	11.88	31.57

If we consider the top half of Table 1 (results for tests where all lights are on, and where only the outside lights are on), we can see that the fluctuations between minimum and maximum values are approximately 100 mm. However, if we compare these values to those obtained where the lighting was low and inconsistent (in the tests where only the inside lights or only the down lights were used), we can see that the fluctuations between minimum and maximum values have now risen to values of approximately 300 mm. However, by considering the mean values for each of the lighting conditions, we can see that the distance between the measured position and the mean positions are within 25mm of each other for all lighting conditions. We consider this error to be acceptable.

4.2. Orientation Results

Because the cameras can be rotated around three axes, we have found it difficult to measure the actual orientation of the cameras in the physical world. Instead of comparing the measured and calculated orientation of each camera, the accuracy of the orientation results will be discussed by comparing all results from all lighting conditions..

Table 2. Calculated Orientation of camera under four lighting conditions. The three values shown are Euler angles[1]. All measurements in degrees.

All	heading	bank	attitude	Out	heading	bank	Attitude
Min	-79.115	-89.993	-26.864	Min	-79.227	-89.981	-27.106
Max	-77.524	-85.376	-25.860	Max	-77.399	-85.127	-25.963
Mean	-78.219	-87.984	-26.399	Mean	-78.308	-88.206	-26.419
Std Dev	0.284	0.907	0.178	Std Dev	0.306	1.010	0.166
Inside	heading	bank	attitude	**Down**	heading	bank	Attitude
Min	-79.036	-89.992	-26.871	Min	-79.473	-89.966	-26.921
Max	-67.965	-70.427	-19.124	Max	-68.171	-69.869	-19.228
Mean	-78.270	-88.068	-26.375	Mean	-78.246	-87.808	-26.283
Std Dev	0.561	1.302	0.386	Std Dev	0.589	1.418	0.383

Table 2 shows the minimum, maximum, mean and standard deviation calculated over the 500 iterations for each of the four lighting conditions. The minimum and maximum values were calculated by comparing the heading, bank and attitude values separately. They therefore represent the range of values that each of the Euler angles took under each of the lighting conditions.

Before analysis was performed on the data shown in Table 2, outliers were removed. Six outliers were removed from data produced in the lighting situation where all lights were on, and where only the outside lights were on. 12 outliers were removed from the inside lighting tests, and 18 were removed from the down lights test. These outliers occurred only in the calculation of the bank, irrespective of the lighting condition being used.

---

[1] Euler angles were calculated using formulae described by Martin Baker in http://www.euclideanspace.com/maths/geometry/rotations/conversions/matrixToEuler/index.htm

By comparing the mean values for each of the lighting conditions, we can see that the largest fluctuation can be found in the mean bank value 0.081 radians (4.64 degrees) compared to a fluctuation of 0.002 radians (0.11 degrees) in heading and attitude.

4.3. Analysis

When considering the accuracy of the position of the cameras, we noticed that the x and y values are always more accurate than the z value in all lighting conditions. When the lighting conditions become uneven (in the cases where only the inside fluorescent lights or the down lights are used), the range that the x, y, and z values take becomes large. As an example, consider the range of values z takes in the case where all lights are used (a range of 119 mm), compared with the lighting condition where only the down lights are used (a range of 322 mm). Not surprisingly, the results we obtained when consistent lighting was used are comparable to those obtained in a study under similar conditions by Malbezin, Piekarski and Thomas [16].

For each of the lighting conditions, we discovered a number of outliers in the bank angle. Surprisingly there are no outliers for the heading or the attitude. The outliers found were all approximately 180 degrees away from the mean value. When the lighting conditions become more uneven, we found that the number of outliers increases. When only down lights were used, the number of outliers was triple those found in conditions of even lighting (all lights on, outside fluorescent lights on). We believe that poor lighting causes false detections, resulting in the outliers.

We also believe that part of the jitter in positional values can be attributed to the jitter in orientation. Because of the lever arm affect, a jitter of 1 degree at a distance of 1.645m will result in a movement of 28 mm.

5. Conclusions and Future Work

In this paper we have described tools that can be used to help rapidly configure a workspace to allow passive detection of marked objects. Fiducial markers are attached to objects such as laptops, furniture, or any other object to be tracked and their physical world pose can be calculated in real time. The application described in this paper, CalibrateCE, has become an integral part of the Passive Detection Framework as it allows users to quickly and efficiently recalibrate some of the features of the environment after the workspace has been reconfigured.

An experiment to determine the accuracy of CalibrateCE was also presented in this paper. We have shown that the accuracy of the calibration is not dependent on the lighting condition under which the calibration takes place. This is due primarily to the number of iterations undertaken in the calibration process, and the ability for the accurate results to cancel out any outliers.

The rapid deployment of the collaborative ubiquitous computing environment described in this paper may be achieved through the utilization of mobile computing and specialized equipment. The node and combiner computers are replaced by notebook computers. The firewire cameras are no longer placed on the ceiling of the room, but on

extendible aluminum tripods. The placement of cameras via the tripods is determined for the best coverage of the tracking region. The firewire hubs and cables are very portable. All these components would easily fit into a padded suitcase. We have not built this portable system, but there are no technology challenges to the construction of such a system.

6. References

1. Weiser, M., *The computer for the 21st century*. Scientific American, 1991. **265**(3): p. 66-75.
2. Slay, H., R. Vernik, and B. Thomas. *Using ARToolkit for Passive Tracking and Presentation in Ubiquitous Workspaces*. in *Second International IEEE ARToolkit Workshop*. 2003. Waseda University, Japan.
3. Slay, H. and B. Thomas. *An Interaction Model for Universal Interaction and Control in Multi Display Environments*. in *International Symposium on Information and Communication Technologies*. 2003. Trinity College Dublin, Ireland.
4. Slay, H., et al. *Interaction Modes for Augmented Reality Visualisation*. in *Australian Symposium on Information Visualisation*. 2001. Sydney, Australia.
5. Kato, H. and M. Billinghurst. *Marker Tracking and HMD Calibration for a Video-based Augmented Reality Conferencing System*. in *2nd IEEE and ACM International Workshop on Augmented Reality*. 1999. San Francisco USA.
6. Streitz, N.A., J. Geißler, and T. Holmer. *Roomware for Cooperative Buildings: Integrated Design of Architectural Spaces and Information Spaces*. in *Cooperative Buildings – Integrating Information, Organization and Architecture. Proceedings of the First International Workshop on Cooperative Buildings (CoBuild'98),*. 1998. Darmstadt, Germany.
7. Grønbæk, K., P.G. Krogh, and M. Kyng. *Intelligent Buildings and pervasive computing - research perspectives and discussions*. in *Proc. of Conference on Architectural Research and Information Technology*. 2001. Arhus.
8. Brooks, R.A., et al. *The Intelligent Room Project*. in *Second International Conference on Cognitive Technology Humanizing the Information Age*. 1997. Los Alamitos, CA, USA.: IEEE Computing Society.
9. Johanson, B., A. Fox, and T. Winograd, *The Interactive Workspaces Project: Experiences with Ubiquitous Computing Rooms*. IEEE Pervasive Computing, 2002. **1**(2): p. 67-74.
10. Morency, L.-P., et al. *Fast stereo-based head tracking for interactive environments*. in *Proceedings Fifth IEEE International Conference on Automatic Face Gesture Recognition, 20-21 May 2002*. 2002. Washington, DC, USA: IEEE.
11. Pentland, A., *Looking at people: sensing for ubiquitous and wearable computing*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000. **22**(1): p. 107-19.
12. Vernik, R., T. Blackburn, and D. Bright. *Extending Interactive Intelligent Workspace Architectures with Enterprise Services*. in *Evolve: Enterprise Information Integration*. 2003. Sydney, Australia.
13. Newman, J., D. Ingram, and A. Hopper. *Augmented Reality in a Wide Area Sentient Environment*. in *International Symposium on Augmented Reality*. 2001. New York, USA.
14. Piekarski, W., et al. *Hybrid Indoor and Outdoor Tracking for Mobile 3D Mixed Reality*. in *The Second IEEE and ACM International Symposium on Mixed and Augmented Reality*. 2003. Tokyo, Japan.
15. Foxlin, E. and N. Leonid. *VIS-Tracker: A Wearable Vision-Inertial Self-Tracker*. in *IEEE Virtual Reality*. 2003. Los Angeles, USA.
16. Malbezin, P., W. Piekarski, and B. Thomas. *Measuring ARTootKit accuracy in long distance tracking experiments*. in *Augmented Reality Toolkit, The First IEEE International Workshop*. 2002.