RFID-based hybrid Camera Tracking in Virtual Studio

Po.Yang, Wenyan.Wu, M.Moniri and Claude C. Chibelushi
Faculty of Computing, Engineering and Technology
Staffordshire University
Stafford, UK
P.Yang@staffs.ac.uk, W.Wu@staffs.ac.uk

***Abstract*— this paper addresses the problem of Camera tracking in virtual studio environment. The traditional camera tracking methods can be classified into optical-based or electromechanical sensor-based. However, the electromechanical method is extensive time-consuming calibration procedures and cost too much; the optical method suffers from the error detection of references features and the chorma keying limitation in virtual studio. Therefore, in order to overcome those problems, we proposed a novel RFID-based hybrid camera tracking method in virtual studio application. Firstly, we designed a RFID passive tags based camera tracker. By using the triangular position algorithm, the accuracy could reach up to 5 centimeters. Secondly, we combined the optical based tracking method into RFID tracker with the aim to improve the orientation and position accuracy. Finally, the experiment results showed that this method could be a novel potential solution for camera tracking system in virtual studio applications.**

Keywords-RFID, camera tracking, chorma key, SLAM

I. INTRODUCTION

During the last few decades, the rapid development and continuous improvement in computing technology have led to significant revolution for the entertainment industry, such as video production. Virtual Studio (VS) technology is one of the most popular topics, which can be widely indorsed into the video industry for a variety of productions with special visual effects. Generally, a virtual studio system consists of mainly three modules, namely the camera tracking, the rendering and the compositing module. The camera tracking module is by all means the most crucial part of virtual studio system, since it determines the alignment of the live video with the available imagery.

The existing camera tracking systems in virtual studios are generally classified into the electromechanical and the optical ones [1] [2]. The electromechanical tracking [3] [4] methods are based on active sensors, which incorporate powered signal emitters and sensors placed in a prepared and calibrated environment. These approaches could be highly accurate but more expensive, also require time-consuming calibration procedures and suffer from random vibrations. The optical tracking methods are relied on the image processing schemes and computer vision technology. It could overpass the problem of time-costly calibration and vibrations of camera; however it easily failed when the references features are out of focus, occluded or even out of view.

In order to overcome those problems, our research work would aim to investigate novel techniques which can be put into practice to achieve a feasible camera tracking solution in virtual studio environments. In the last several decades, an RFID (Radio Frequency Identification) technique [5] has been of great interest for mobile computing and robotic researchers. It is widely applied into various industrial areas due to its low-cost features, such as supply chain optimization [6], counterfeit control [7], ubiquitous computing environments [8] and mobile robots navigation [9]. Moreover, to deal with the problem of mobile robots localization in a real and practical way, many researchers consider the system’s ease of use and range of applicability. Due to the unique ID of RFID device, some methods attempts to study the localization of mobile robot using RFID tags as landmarks in the environment [10] [11] [12]. Therefore, in our work, we would attempt to apply the RFID techniques into camera tracking for virtual studio. Since the problem of camera tracking in virtual studio could be treated as a problem of mobile robots localization indoor to some extend. However, the task of camera tracking requires much higher accuracy than the general robots localization. For this issue, we would combine the optical tracking technology with RFID based tracker to solve this problem.

In this paper, Section 2 gives a basic literature review of camera tracking in virtual studio and RFID for localization. Section 3 proposes a novel RFID-based hybrid camera tracking method in virtual studio environment. Section 4 describes the simulation results and analysis. Section 5 draws some conclusions and future work.

II. LITERATURE REVIEW

A. *Camera tracking in virtual studio*

Virtual studio [13] is a television studio that allows the combination of people or other real objects and computer generated environments and objects in a seamless, virtual reality-like manner. In general, a virtual studio system includes the following components: camera tracking system, rendering and compositing software, video mixer. Above these components, the task of camera tracking system is to accurately and robustly measure the position and orientation parameters of the foreground camera. Meanwhile, both real camera in virtual studio environment and virtual camera in rending software environment should have the same position and orientation parameters over time. Figure 1, shows the camera tracking system in virtual studio application.

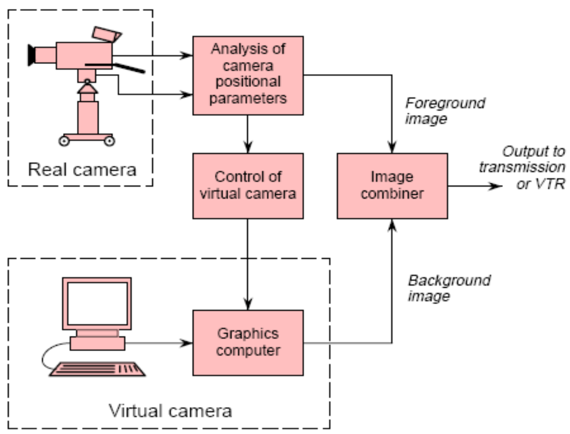

Fig 1 Camera tracking system for virtual studio

As we mentioned before, nowadays, camera tracking system in virtual studios can be generally classified into two broad types, namely the electromechanical (sensor-based) and the optical ones (vision-based). The electromechanical camera tracking system is to employ servo-control mechanisms to control the camera when 3-D camera motion is pre-determined, such as the mechanical sensor tracking system in ORAD [14]. This system is widely adopted in commercial broadcast applications since it can be highly accurate. However, the demerits of electromechanical camera tracking system are extensive time-consuming calibration procedures and expensive costs. The optical camera tracking system is based on the image processing and computer vision algorithms to extract camera motion on the basis of frames currently captured. To achieve this purpose, the single blue or green screen is extended to incorporate appropriate reference points [15], such as the pattern recognition camera tracking system in ORAD Hi-Ti [14], Hi-Ball system [18].

Although by using optical solutions, the problem of time-consuming calibration and vibrations of camera are overpassed, the optical systems are easy to fail when the referenced features are out of focus, occluded or even out of view. Therefore, some hybrid camera tracking systems are developed for virtual studios. BBC R&D proposed “Free-d” [16] camera tracking system by using bar-code patterns with retro-reflective material and electromechanical sensor. ORAD developed “InfraTrack” [17] camera tracking system, which use image processing software to recognize the light source image and calculate the camera locations by using infrared emitters and infrared tracking cameras. However, these hybrid camera tracking systems also suffer from expensive cost and complicated calibration and setup procedure.

B. RFID technique for localization

With the low-cost and massive utilities, RFID technology is a hot topic in industry, with simulating the desire of RFID-supported applications such as product tracking, supply chain optimization, asset and tool management. Besides, these “conventional” application areas, passive RFID tags are also suited to augmenting physical objects with virtual representations or computational functionality, providing a versatile technology for “bridging physical and virtual worlds” in ubiquitous computing environments.

Over the years, for the localization problem in indoor environment, various RF-based techniques have been developed. One of the well tried methods was using the known locations of reference stations, where the transponder position was reported by nearby stations [19]. However, this approach had the problem of cost and space with the number of stations and the interval between the stations affected the accuracy. Likewise, the signal strength-base distance estimation was often used with stations, but the accuracy deteriorated due to the scattering and reflecting of transmitting signal [20]. Another method uses the difference of arrival time between the RF signal and the ultrasonic pulse [21]. However, such system requires pre-organization of location of sensors or beacons, thus the setup and the running cost might limit their applicability in a large scale space. Also, the optical line of sight is needed for transmitting the ultrasonic pulse.

Although the range of application of the RFID technology has been widening including applications such as the localization of mobile robots, since it can perform recognition at high-speed and send data within distances appropriate for this application. Nevertheless, most of researchers did not use RFID techniques in tracking field since they think that RFID is just an identification standard and can not provide exact localization. Meanwhile, the anti-collision, material effect and false-reading could easily cause the error of tracking. Thus, most of researchers only focus on using RFID for localization of mobile robots indoor.

III. RFID-BASED HYBRID CAMERA TRACKING FOR VIRTUAL STUDIO

A. RFID tag infrastructures for camera tracking

In our research work, due to the localization feature of RFID technology, as a first step toward camera tracking system in virtual studio environment, we propose a simple and novel location camera system based on RFID localization technologies, shown as Fig.2.

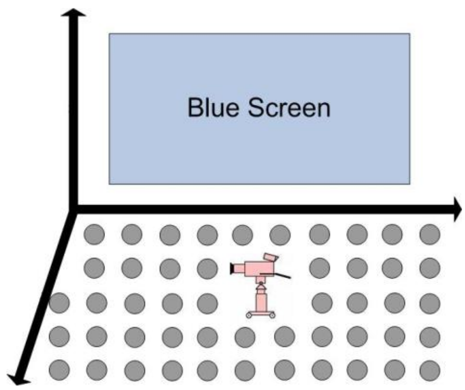

Fig.2 Camera Tracking in Virtual Studio by RFID passive Tags

In order to estimate the camera position by using the RFID system, RFID tags are arranged in a fixed pattern on the floor. Since the tag arrangement on the floor is pre-planned, each tag stores its absolute position data and sends them out when they are requested. The RFID reader installed on the bottom of the camera gathers the tag data, as shown in Fig 3. When the camera moves and stays on any tags, the RFID reader antenna forms en effective area. All the tags within the circle of radius r, which are under the effective area of the RFID antenna, are activated.

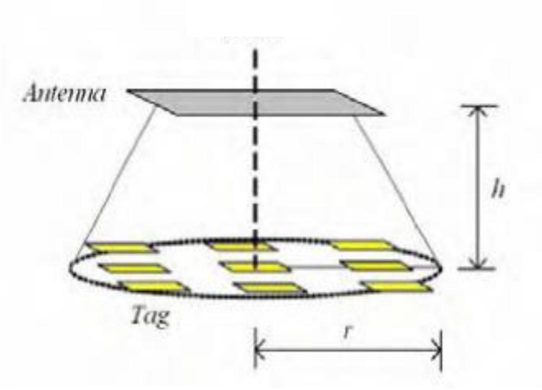

Fig 3 Recognition area model of RFID antenna.

B. RFID-based triangular camera localization

Under the RFID tag pattern mentioned before, the recognition area for the tags can be represented as a circle in Fig.4

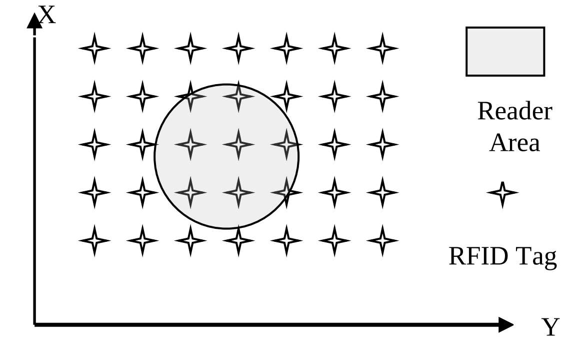

Fig.4. RFID reader recognition area

If we assume the camera position (x_c, y_c) that carries a reader antenna on the bottom can be obtained through the position data of the tags (x_T^N, y_T^N), where N represents the number of tags detected by the reader, and (x_T^N, y_T^N) represents the coordinate's information of the tags, then we can get:

$$x_c = \frac{\max\{x_T^1, x_T^2, \ldots x_T^N\} + \min\{x_T^1, x_T^2, \ldots x_T^N\}}{2} \qquad (1)$$

$$y_c = \frac{\max\{y_T^1, y_T^2, \ldots y_T^N\} + \min\{y_T^1, y_T^2, \ldots y_T^N\}}{2} \qquad (2)$$

In the camera tracking procedure, the estimation error always exists as shown in Fig.5.

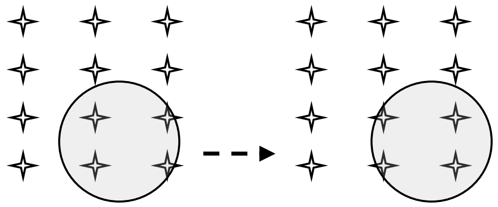

Fig.5. Estimation error in RFID Tag infrastructure.

The real position of camera in fig6 is changed, however, the estimation position of camera in RFID tag infrastructure is still the same. The estimation error represents the gap between the real and estimated position of camera. Therefore the estimation error is directly related to the gap between the distances of RFID tags. If we assume the distance between the tags is d , and the left boundary of the reader recognition area is denoted as M_1 and the right as M_2. The RFID reader can detect tags from (x_T^i, y_T^i) to (x_T^j, y_T^j) locate between M_1 and M_2, thus the estimation position of camera is (x_e, y_e) and the real position of camera is (x_c, y_c).

$$x_c = \frac{M_1 + M_2}{2} \qquad (3)$$

$$x_e = \frac{x_T^i + x_T^j}{2} \qquad (4)$$

Then the estimated error of X coordinate is defined as:

$$e_{err\_x} = |x_e - x_c| = \left| \frac{M_1 + M_2}{2} - \frac{x_T^i + x_T^j}{2} \right| \qquad (5)$$

Where the ranges of M_1 and M_2 can be described as

$$\begin{cases} x_T^{i-1} < M_1 < x_T^i \\ -d < M_1 - x_T^i < 0 \\ x_T^j < M_2 < x_T^{j+1} \\ 0 < M_2 - x_T^j < d \end{cases} \qquad (6)$$

Then from equation (5) and (6), now the range of estimation error can be represented as:

$$e_{err\_x} = \left| \frac{(M_1 - x_T^i) + (M_2 - x_T^j)}{2} \right| \le \frac{1}{2}|d| \qquad (7)$$

Therefore, the maximum estimation error in the X-Y Cartesian coordinates is represented as:

$$e_{err} = \sqrt{(\frac{d}{2})^2 + (\frac{d}{2})^2} \cong 0.707d \qquad (8)$$

From the equation (8), we can find that the accuracy of RFID tag infrastructure camera tracking system is directly related to the distance between passive RFID tags. However, the above estimation is based on the dense overlay square pattern. If we use the dense overlay triangular pattern, the error would be changed.

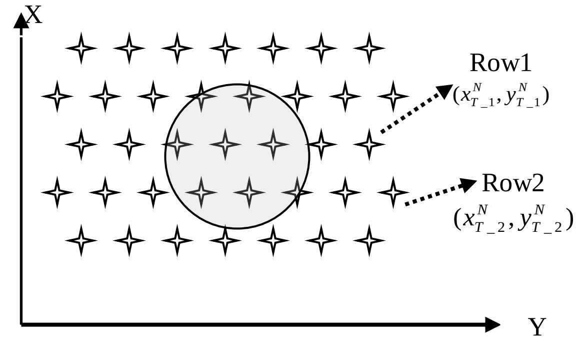

Fig.7. dense overlay triangular pattern

Based on this pattern, we should redefine two sets of RFID tag infrastructure $(x^N_{T\_1}, y^N_{T\_1})$, $(x^N_{T\_2}, y^N_{T\_2})$, which represent two different rows in this pattern. The RFID reader can detect tags from $(x^i_{T\_1}, y^i_{T\_1})$ and $(x^i_{T\_2}, y^i_{T\_2})$ to $(x^j_{T\_1}, y^j_{T\_1})$ and $(x^j_{T\_2}, y^j_{T\_2})$ locate between M_1 and M_2 , and the equation (6) would be defined as:

$$\begin{cases} x^{i-1}_{T\_2} < M_1 < x^i_{T\_1} \\ -\dfrac{d}{2} < M_1 - x^i_{T\_1} < 0 \\ x^j_{T\_2} < M_2 < x^{j+1}_{T\_1} \\ 0 < M_2 - x^j_{T\_2} < \dfrac{d}{2} \end{cases} \qquad (9)$$

Thus the estimation error in X-direction can be decreased as follows:

$$e_{err\_x} = |x_e - x_c| = \left| \frac{M_1 + M_2}{2} - \frac{x^i_{T\_1} + x^j_{T\_2}}{2} \right| \quad (10)$$

$$e_{err\_x} = \left| \frac{(M_1 - x^i_T) + (M_2 - x^j_T)}{2} \right| \le \frac{1}{4}|d| \quad (11)$$

$$e_{err} = \sqrt{(\frac{d}{4})^2 + (\frac{d}{4})^2} \cong 0.35d \quad (12)$$

Therefore, under RFID tag infrastructure, if we use the triangular pattern with anti-collision reader, the error would theatrically reduce.

C. Stereo vision algorithm for improving accuracy

In the section B, the RFID based tracker would achieve the task on approximate camera localization. However, the accuracy is limited by some issues, such as tag density, antenna collision. In order to improve the tracking accuracy, we would propose a stereo vision based algorithm. In this project, we would assume that there is a small camera sensor attached with RFID reader, which has the same field of view when RFID reader moving, as shown in Fig 8.

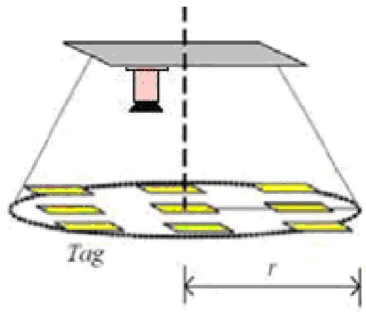

Fig 8 Recognition area of camera and RFID reader

Stereo vision [22] based algorithm is a very popular method in computer vision area. The core idea is to extract the camera parameters by using two or multiple continuous image sequences. The core idea of this method is to calculate the fundamental matrix. The fundamental matrix is the algebraic representation of epipolar geometry. The fundamental matrix from the mapping between a point and its epipolar line, and then specify the properties of the matrix.

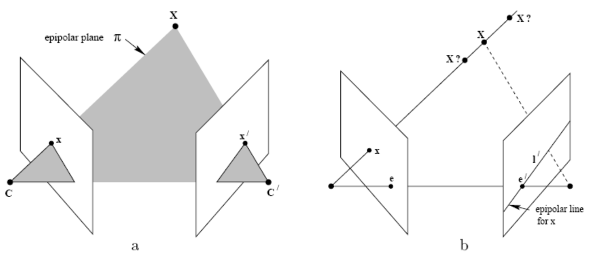

Fig 9. Epilolar geometry and fundamental matrix

Given a pair of images, it was seen in Fig 9 that to each point **x** in one image, there exists a corresponding epipolar line **l**' in the other image. Any point **x**' in th second image matching the point **x** must lie on the epipolar line **l**'. The epipolar line is the projection in the second image of the ray from the point **x** through the camera centre **C** of the first camera. Thus, there is a map

$$\mathbf{x} \mapsto \mathbf{l}'$$

from a point in one image to its corresponding epipolar line in the other image. It is the nature of this map that will now be explored. It will turn out that this mapping is a (singular) *correlation*, that is a projective mapping from points to lines, which is represented by a matrix F, the fundamental matrix. Therefore, if we assume the fixed feature points, we can extract the camera parameters by calculating the fundamental matrix F.

IV. SIMULIATION AND RESULTS.

A. RFID-Tracker

Based on the algorithm above, we can work some simulation in virtual 3D world to evaluate the camera tracking error impact on the 3D view of camera. According to the equation (12), due to the normal size of RFID reader antenna is 12cm*12cm, if we assume the distance

$$e_{err} = \sqrt{(\frac{d}{4})^2 + (\frac{d}{4})^2} \cong 0.35d \qquad (12)$$

between two passive tags on the floor are 3cm, 5cm, 10cm, the theoretical error should as Table 1:

Tag Distance	3 cm	5 cm	10 cm	>=12 cm
Error Value	1.05 cm	1.75cm	3.5 cm	>=12cm
Max Detected Tags	16	9	4	1

Table 1: Theoretical error for simulation.

In this case, we model the virtual 3D environment by using Maya, the avatar position is (0, 7.2, 0) and the camera position is (0, 7, 16). The angle is X axis -5 degree. Compare to the real virtual world environment, if the people are 170cm, according to the rate, the distance between camera and people is about 3.4 m, and the height of camera is about 1.6cm, as shown in Figure A. The image is captured by Maya 3D view, we can set it as the original view, and match the difference by difference errors.

In this virtual environment, according to the rate, we would evaluate the difference distance of passive tags for the tracking error. In our assumption rata, the tag distance on 12cm, 10cm, 5cm, 3cm separately match the error value on 0.51, 0.15, 0.075, and 0.0375. Therefore, according to this figure, we can assume that the camera position with error on X axis, and we would get the difference view B (0.51, 7, 16), C (0.15, 7, 16), D (0.075, 7, 16), E (0.0375, 7, 16).

From the results of images, we can easy tell the difference between A and B visually, the position of averter is different in the images. Thus, it shows that the RFID tag distance over 12 cm is not suitable for virtual studio environment.

However, compare A and C, also A and D, it is very difficult to tell the difference by human eyes. It means that the anti-collision RFID reader and density over-lay triangular pattern of RFID passive tags has potential application on camera tracking in virtual studio. Meanwhile, we did not show the E position camera view since the E position is too close to A position, and also RFID passive tag's distance is too small to widely use.

Fig10 A Original view, camera position (0, 7,16)

B Tag Distance (12cm), camera position (0.51, 7, 16)

C Tag Distance (10cm), Camera position (0.15, 7, 16)

. D Tag Distance (5cm), Camera position (0.075, 7, 16)

B. Stereo-based algortihm

Based on the results above, RFID tracker would possibly reach the camera tracking accuracy up to 1-3 centimeters. Besides it, we can work some simulation on stereo-based algorithm to improve accuracy. If it is assumed that 10 random points detection in a point (1,2,0.7) with a radius 1 centimeter. The camera sensor position is from [-0.2, -1, 0.2] to desired position [1.1, 0.5, 0]. Then we can get those error results between six camera parameters extraction estimation and desired camera position.

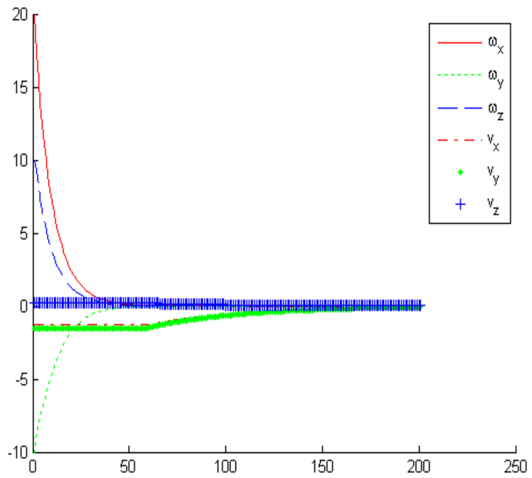

Fig 11 6DOF of camera extraction error

From the fig 11, it shows that the error of camera tracking by using this algorithm would be below 1 unit. In other words, the stereo-based algorithm would improve the RFID tracker accuracy below 1 cm, up to 2 mm, and orientation degree accuracy below 0.1 degree.

V. CONCLUSION AND FUTURE WORK

Virtual studios have long been used in commercial broadcasting and are starting to evolve into the next level with improved technology in image processing and computer graphics. In this paper, a RFID based hybrid camera tracking solution has been designed and implemented to solve the problem of camera tracking in virtual studio environment. The simulation results show that this method would achieve the camera position tracking up to 2 mm, and orientation degree accuracy below 0.1 degree. The future work will focus on the RFID tag pattern design, which gives a robust and reliable features rather than random points. I

REFERENCES

[1]. K. Cornelis, M. Pollefeys, and L. V. Gool. Tracking based structure and motion recovery for augmented video productions. In *Proc. ACM Symposium on Virtual Reality and Software Technology (VRST)*, Alberta, Canada, pp. 17-24. 2001.

[2]. M.I. Lourakis and A.A. Argyros, 2005. Efficient, causal camera tracking in unprepared environments. *Computer Vision and Image Understanding* 99, 2 (Aug. '05), pp. 259–290, 2005.

[3]. M. Livingston and A.State: Magnetic Tracker Calibration for Improved Augmented Reality Registration. Presence, vol. 6, pp. 532–546, 1997.

[4].F. Madritsh. Optical Beacon Tracking for Human-Computer Interfaces. PhD thesis, Technical University Graz, 1996. M.Lourakis, Egomotion estimation using quadruples of collinear image points, in: Proc. ECCV'00,vol.2, 2000,pp. 834-848.

[5] B, Glover., H, Bhatt, An Introduction to RFID, Chap. 1. RFID Essentials. 1. O'Reilly. 276. 2006

[6] J.R. Tuttle. Traditional and emerging technologies and applications in the radio frequency identification (RFID) industry. In *IEEE Radio Frequency Integrated Circuits (RFIC) Symposium*, pages 5–8, June 1997.

[7]. Sanjay E. Sarma. Towards the five-cent tag. Technical Report MIT-AUTOID-WH-006, MIT Auto-ID Center, 2001.

[8]. Roy Want, Kenneth P. Fishkin, Anuj Gujar, and Beverly L. Harrison. Bridging physical and virtual worlds with electronic tags. In *Proc. SIGCHI Conference on Human Factors in Computing Systems*, pages 370–377. ACM Press, 1999.

[9]. Kim M, Takeuchi T, Chong NY. Object location sensing using active RFID systems. Proc Int Symp Robot Automat 2004: 440–5.

[10]. D. Hahnel, W. Burgard, D. Fox, K. Fishkin, and M.Philipose, "Mapping and localization with RFID technology," Proc. IEEE Int. Conf on Robotics andAutomation, vol. 1, pp. 1015-1020, 2004.

[11]. J.R. Tuttle. Traditional and emerging technologies and applications in the radio frequency identification (RFID) industry. In *IEEE Radio Frequency Integrated Circuits (RFIC) Symposium*, pages 5–8, June 1997.

[12] Roy Want, Kenneth P. Fishkin, Anuj Gujar, and Beverly L. Harrison. Bridging physical and virtual worlds with electronic tags. In *Proc. SIGCHI Conference on Human Factors in Computing Systems*, pages 370–377. ACM Press, 1999.

[13]. S. Gibbs, C. Arapis, C. Breiteneder, V. Lalioti, S. Mostafawy, and J. Speier, "Virtual studios: An overview," *IEEE Multimedia*, vol. 5, No. 1, pp. 18–35, Jan.–Mar. 1998.

[14]. http://www.orad.co.il/en/page.asp?id=81 June 2008.

[15]. O. Faugeras, *Three-Dimensional Computer Vision*. New York: MIT Press, 1993.

[16]. G. A. Thomas, J. Jin, T. Niblett, and C. Urquhart. "A versatile camera position measurement system for virtual reality TV production." In *In Proceedings of International Broadcasting Convention, Amsterdam*, pages 284.189, 1997.

[17]. http://www.orad.co.il/en/page.asp?id=80 Aug 2007

[18]. Greg Welch, Gary Bishop, Leandra Vicci, Stephen Brumback, Kurtis Keller, and D'nardo Colucci. The hiball tracker: High-performance wide-area tracking for virtual and augmented environments. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, 1999.

[19] Werb J, Lanzl C. Designing a positioning system for finding things and people indoors. Spectrum, IEEE 1998;35(9):71–8.

[20] Bahl P, Padmanabhan Venkata N. RADAR: an in-building RF-based user location and tracking system. IEEE Infocom 2000;2:775–84.

[21] Fukuji Y, Minami M, Morikawa H, Aoyama T. DOLPHIN: an autonomous indoor positioning system in ubiquitous computing environment. In: Proceedings of the IEEE workshop on software technologies for future embedded systems (WSTFES03); 2003. p. 53–6.