

Pose Estimation Ambiguity Problem and Their Solution Approaches in Augmented Reality Registration

Ahmad Ali, Vinay Kumar Singh, Prabha S.Nair, Dr Manoj Wadhwa
School of computing science and engineering
Galgotias University, Gautam Budh Nagar, Greater Noida, U.P.

Abstract: *In augmented reality applications, tracking and registration of both cameras and objects is required because, augmented reality is a combine real and virtual object. In this paper we describe the pose estimation problem. In augmented reality registration problem is big problem. There are many problems in augmented reality system proposed by researchers. Such as, knowledge base-registration technology, tracking-based registration technology, and computer vision-based technology. The tracking based registration method we used camera sensors display etc. Here big problem is estimation problem mainly in location and orientation [1][2]. In this paper we described cause of pose estimation problem and their solution.*

Keywords-*augmented reality, pose estimation, estimation ambiguity.*

1. INTRODUCTION

Augmented reality is a system where real-world environments can interact with virtual objects and virtual objects will respond to real-world physical actions. Augmented reality is real time processing, data can be generated in real time. Augmented reality is imagery in 3D, graphics and models are created by designers. Current market AR has lot of application such as Medical, Manufacturing and repair, Annotation and visualization, Robot path planning, Entertainment, Military aircraft etc. But the problem in augmented reality is registration. Why we need registration such question may be occurred in our mind. Registration is very important in augmented reality. For example in some movie there is actor whose every character is real scene and actor has a partner that is a big animal which is computer generated virtual object. Now every character of animal we can see in real scene with actors. The main purpose of this example is to describe how the virtual object (animal) occluded with real object (actor) and we can see real scene in movie. This is augmented reality registration technique. There are various kinds of registration technology proposed by various researchers such

as knowledge-based registration, tracking-based registration and computer vision-based registration technology [3]. In tracking-based technology we used camera, display, sensors etc. pose estimation ambiguity is big problem here. There are mainly two types of pose estimation problems; one is orientation and another is location. In orientation estimation there are also two problems again; one is "degradation of orientation accuracy in frontal observation". Another is "pose ambiguity" [4][5] where the estimated orientation repeats switching between two values. In this paper we propose the solution of pose estimation ambiguity. In section II-we described pose estimation, and pose estimation ambiguity using perspective projection, section III-solution of pose estimation ambiguity in section IV-discussion and feature approaches in section V- conclusion and in section VI we described references.

II. POSE ESTIMATION

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

The camera pose is estimated using homography that is calculated using at least four keypoints correspondences as the result of the shape registration. The outliers from the keypoints are removed using the inverse homography. The camera pose is then optimized using Levenberg-Marquardt [6] by minimizing the re-projection error that is the distance between the projected keypoints from the shape database and the extracted keypoints in the captured frames. The camera pose is then refined by considering the keypoints correspondence to the detected shape in previous frame. These two optimizations produce a stable camera pose. We can describe using following figure.

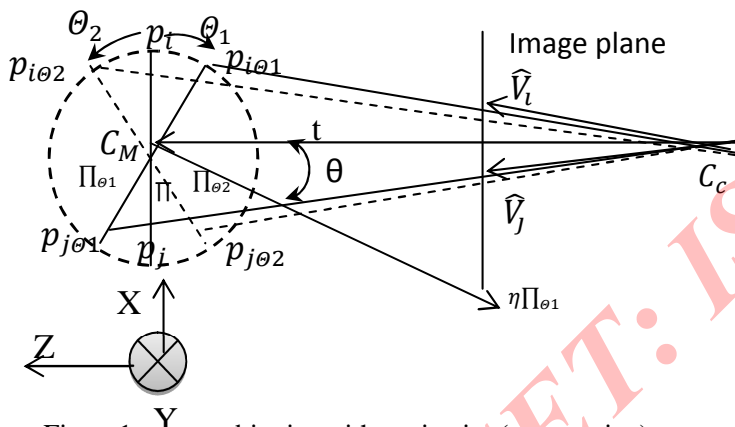


Figure 1: pose ambiguity with projection (perspective)

In figure-1 we consider $R=f(\theta_1, \theta_2, \theta)$ is orientation and $t=[t_x, t_y, t_z]^T$ is position between camera and scene coordinate system.

C_C is camera center. C_M is model center. p_i and p_j are the points of image plane model points are projected. Based on figure-1 we can write the following estimation algorithm [5].

- 1) Estimate the first pose $\hat{P}_1=(\hat{R}_1, \hat{t}_1)$, where \hat{P}_1 is local minimum of error function (E_{os} mention in section-III).
- 2) Transform the coordinate system according to (using $\tilde{v}_i=R_t \hat{v}_i \tilde{t}_1=R_t \hat{t}_1 \tilde{R}_1=R_t \hat{R}_1$ [5]) $\tilde{p}_1=(\tilde{R}_1 \tilde{t}_1)$ [7].
- 3) Estimate rotation matrix (\tilde{R}_2) [5].

- 4) Fixe $\theta=\tilde{\theta}_1$ and estimate all local minima of error function ($E_{os}(\theta, \theta_2, \tilde{t})$) for the parameter θ_2 , and \tilde{t} .
- 5) Undo the transformations of step 1 and 2 for all local minima to obtain poses \hat{p}_i .
- 6) Use all pose \hat{p}_i as a start value for the iterative pose estimation algorithm to get final poses p^*_i .
- 7) Decide the final and correct pose, which has the lowest error E_{os} .

Using above algorithm we can find the pose estimation.

Limitation

The proposed method is for replacing the sensor attached on the canvas object into a visual tracker. The visual tracker is beneficial since the sensor is no longer necessary for each canvas object. However, the visual tracking has some limitations. Visual tracking only relies on the camera for recognizing the object. Therefore, the disadvantage such as occlusions is unavoidable. As stated in the evaluation, our proposed method will fail when many occlusions occur. Even though, partial occlusions can be compensated. The accuracy of the system also decreases when the tracking fails. It occurs when the camera is in extreme position and orientation. In addition, rapid motion of the camera will create blur area on the image that makes the contour and keypoints become difficult to extract. Furthermore, when the outlines of the shape are not detected, the location information becomes inaccessible. On the other hand, on sensor-based system the sensor information is continuously transmitted. Therefore, even when the object is invisible by the camera, the location of the object is still accessible. However, in this system, we assume that the user always sees the canvas for painting, generally that makes the outlines of the shape always detectable.

III POSE AMBIGUITY SOLUTION

Here we discussed three approaches here. These are array mark, invalid estimation detection and orientation modification through orientation. these are in details as bellow.

III.1 ArrayMark

ArrayMark [8] consist of 4 reference points as shown figure 2 (a) and a plus symbol [1] (2D) is used to change angle direction along with virtual line. Here a stick is used to change the direction of plus sign as shown figure 2 (b). Here first detect the four refere-

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

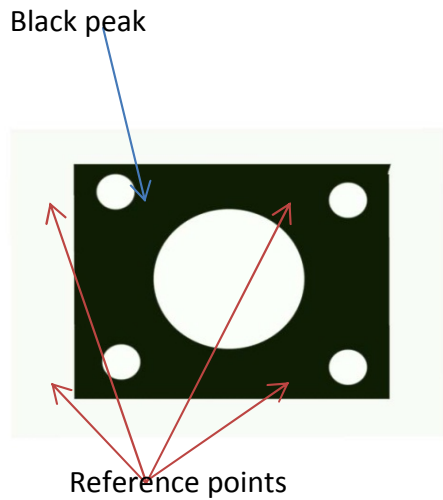


Figure2(a):array mark with 4 reference point

nce points and estimate the pose by perspective geometry method and then modify the orientation by using the angle information calculated by the position of the black peck as shown figure 2. If we obser- ved it from frontal direction, it is stable and accurate orientation estimation, where. estimation error less than one Drawback are that it is occurred with limited angle range.

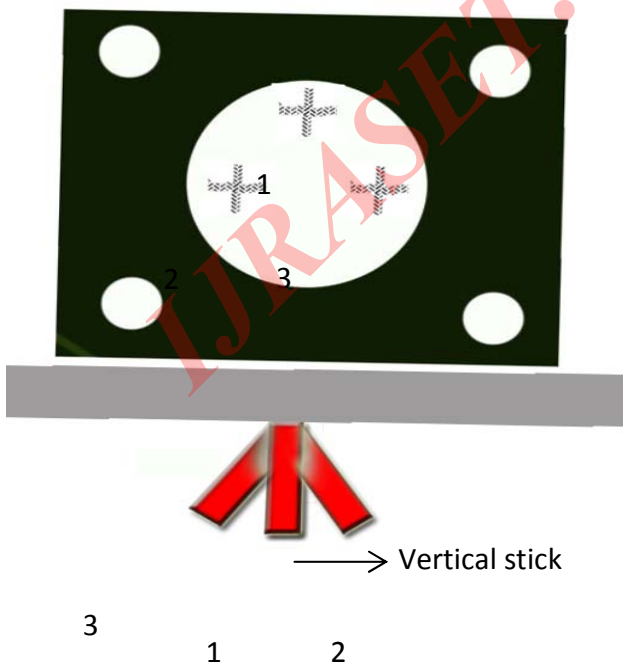


Figure2 (b): movement of plus sign with vertical stick.

III.2 Invalid estimation detection

Here we attached an additional reference point (ARP) [1]and estimate the ARP position by estimate pose and geometric model of marker. if we consider two estimated points P_1 and P_2 , they are located r distance from ARP, then we can find out estimation vector and detection vector as $V_e = P_{1estimated}P_{2estimated}$ and $V_d = P_{1estimated}P_{2detected}$. we can also find iner product $r = V_d \cdot V_e / (|V_d| |V_e|)$ where $-1 \leq d \leq 1$. when $d < -0.6$, invalid estimation occurred.

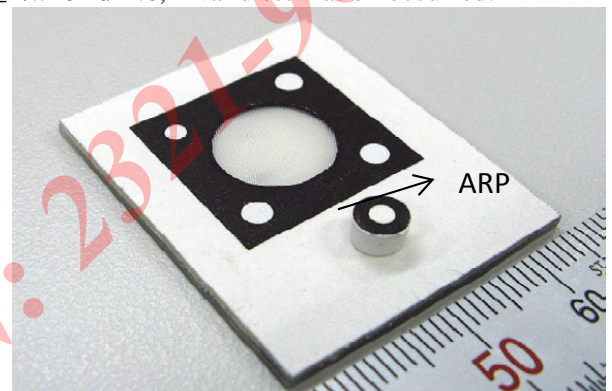


Figure3 (a): marker with additional reference point (ARP)

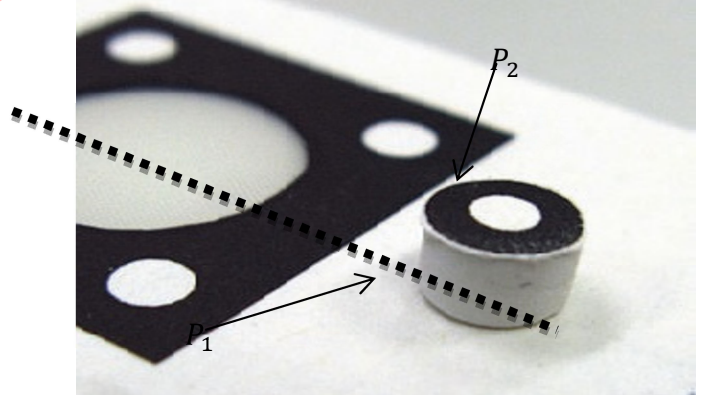


Figure3 (b) two points and with r distance from ARP P_1 P_2

figure3:New ArrayMark and detection of invalid estimation by ARP

III. Orientation modification by error model

We define a “visual-line angle” as a zenith-angle of visual line to the marker. It can be divided into θ_{vx} (around x-axis) and

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

θ_{vyc} (around y-axis) (Fig. 4(a)). We confirmed that the invalid orientation estimation is a phenomenon in which the signs of the visual-line angles are inverted. We call this geometrical model an “error model” of the invalid estimations caused by the pose ambiguity. When we detect an invalid estimation, we modify the orientation by rotation transform, based on the error model (Fig. 4(b)). This modification is done in the same way described in [8].

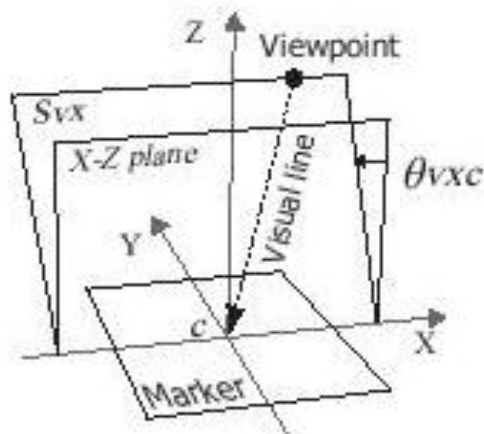


Figure 4(a): virtual line angle (θ_{vxc})

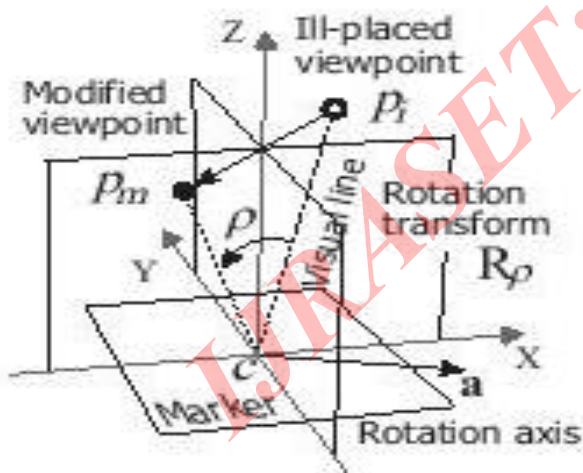


Figure 4(b): Orientation modification by inverting visual line angles.

IV. DISCUSSION AND FEATURE APPROACHES

However, the VMP method is not applicable in larger angle area, hence the invalid estimations still remain unresolved when the absolute value of either visual-line angle is more than about $10[\text{deg}]$. We can eliminate these residues by the ARP added in this study. In this way, we realized stable orientation estimation in wide angle range, and solved the big two problems of the conventional fiducial markers.

For reference, we also show the result of “Ref. Using ARP only”. Compared with the results by the VMP, the estimation of the circled parts is not so stable. It is because the error model in frontal direction is different from that of the pose ambiguity. Therefore, we need the VMP to modify the orientation errors in frontal direction.

We proposed a practical solution which works well on small markers and cheap cameras. Though we introduced a three-dimensional structure, the ratio of the height to the long side of the marker is only

about 0.1. This is a benefit brought by our strategy, in which we use the ARP not for pose estimation but for error detection. The error (invalid estimation) is easily detected even by using a noisy cheap camera and a not-so-accurate geometrical model of the ARP. If we use the ARP for pose estimation by using PnP techniques like [9], we need at least two ARPs placed on higher position. They will make shadows, hide other feature points, and need very accurate geometrical models. It is not practical. Bokode [10] is another high accuracy small marker having three-dimensional structure. Unlike the ArrayMark, it needs internal lighting, and cannot be recognized by a usual camera. Thus it is not suitable for practical use in AR. In the category of the planar marker, random dot markers [11] and Uniform Marker Fields [12] might contribute to stabilize the orientation estimation. However, they are designed to cover large areas in images, thus their methods are not applicable to smaller markers like the ArrayMark.

V CONCLUSION

We proposed a solution to the pose ambiguity problem of planar fiducial markers, which is also applicable to conventional markers. We used a three-dimensional reference point in order to detect invalid estimations. Although it has three-dimensional structure, it is so thin that we can use it as same as the conventional planar markers without impacting on the human-

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

living environment. The ArrayMark solved the big two unresolved problems of the conventional markers. It brings great benefits to many applications, e.g. augmented reality, robotics, and measurement. We will continue further development for more robust and accurate ArrayMark in real environments, working towards realization useful in practice.

VI. REFERENCES

- [1] Hideyuki Tanaka, Yasushi Sumi, Yoshio Matsumoto: "Farther stabilization of a microlens-array-based Fiducial marker" 1 - 4 October 2013, Adelaide, SA, Australia 978-1-4799-2869-9/13/\$31.00 © 2013 IEEE
- [2] Sandy Martedi, Maki Sugimoto, Hideo Saito, Mai Otsuki, Asako Kimura, Fumihisa Shibata: "A Tracking Method for 2D canvas in MR-based interactive painting system" 2013 International Conference on Signal-Image Technology & Internet-Based Systems, 978-1-4799-3211-5/13 \$31.00 © 2013 IEEE DOI 10.1109/SITIS.2013.128
- [3] Li Yi-bo, Kang Shao-peng, Qiao Zhi-hua, Zhu Qiong "Development Actuality and Application of Registration technology in Augmented Reality" 978-0-7695-3311-7/08 \$25.00 © 2008 IEEE DOI 10.1109/ISCID.2008.120.
- [4] R. Hartley and A. Zisserman. *Multiple View Geometry in computer-vision*. Cambridge University Press, 2004.
- [5] G. Schweighofer and A. Pinz. Robust pose estimation from a planar target. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 28(12), pp. 2024–2030, 2006.
- [6] Lourakis, M. "levmar: Levenberg-marquardt nonlinear leastsquares algorithms in C/C++." [web page] <http://www.ics.forth.gr/~lourakis/levmar/>, Jul. 2004. [Accessed on 23 Aug. 2013].
- [7] C. Lu, G. Hager, and E. Mjølness, Fast and globally convergent pose estimation from video images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 6, pp. 610-622, June 2000.
- [8] H. Tanaka, Y. Sumi, and Y. Matsumoto. A high-accuracy visual marker based on a microlens array. In *Proceedings of 2012 IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems*, pp. 4192–4197, 2012.
- [9] V. Lepetit, F. Moreno-Noguer, and P. Fua. EPnP: An accurate O(n) solution to the PnP problem. In *International Journal of Computer Vision*, Vol. 81(2), pp. 155–166, 2009.
- [10] A. Mohan, G. Woo, S. Hiura, Q. Smit-hwick, and R. Raskar. Bokode: imperceptible visual tags for camera based interaction from a distance. In *ACM Transactions on Graphics*, Vol. 28(3), No. 98, 2009.
- [11] H. Uchiyama and H. Saito. Random Dot markers. In *Proceedings of IEEE Virtual Reality*, pp. 35–38, 2011.
- [12] I. Szentandrási, M. Zacharias, J. Havel, A. Herout, M. Dubska, and R. Kajan. Uniform Marker Fields: Camera localization by orientable De Bruijn-Tori. *Proceedings of 2012 IEEE International Symposium on Mixed and Augmented Reality*, pp. 319–320, 2012.

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND
ENGINEERING TECHNOLOGY (IJRASET)

IJRASET: ISSN: 2321-9653