A Survey: Augmented Reality for Guidance Using Various Techniques- Surf, Viola and Sift

ShilpiKumariMandal
M.Tech Student of C.S.E,
SVCST, Bhopal, M.P.-India

Mr. Pinkeshwar Mishra
H.O.D, C.S.E Dept., SVCST,
Bhopal, M.P.-India

Mr. KeshavTiwari,Assistant
Professor, C.S.E Dept.,
SVCST, Bhopal,M.P.-India

***Abstract*—These studies of augmented reality for object tracking & detecting. The core of our proposition concerns augmented reality solution. Our study for augmented reality is technically based on SURF, VIOLA JONES and SIFTS (Scale Invariant Feature Transform) features for localizationintegration of 3D models into video. These SIFT features are projected on a digital model of the building façades of the square to obtain 3D co-ordinates for each feature point. The algorithms implemented calculate the camera pose for frame of a video from 3D-2D point correspondences between features extracted in the current video frame and points in the reference dataset. The study shows the potential of SIFT, SURF and VIOLA JONES for purely image based object detection and tracking augmented reality applications. The end-user goal of our application is to integrate into a lightweight and multimedia platform the tools required to visit a place and to obtain additional information about this place.**

***Index Terms*—augmented reality. Sift, surf, viola jones,Object tracking and detection etc.**

Manuscript received Dec, 2015.

***Shilpikumarimandal**, C.S.E., S.V.C.S.T., Bhopal, India,7351138355.*
***Mr. Pinkeshwar Mishra**, C.S.E., S.V.C.S.T., Bhopal, India,7351138355.*
***Mr.KeshavTiwari**, C.S.E., S.V.C.S.T., Bhopal, India,7351138355.*

I. INTRODUCTION

Augmented Reality (AR) is an emerging technology appeared in the 1990s. It uses virtual world to augment people's perception of real world by combining computer-generated virtual information with real-world environment. AR technology has three main characteristics, including real-virtual combination, 3D registration and real-time interaction. 3D registration is a key technology which performs the position registration of virtual scene and the real world. It has two types, including technologies based on patterns and natural features [1].The AR technology based on patterns has developed maturely. Since the AR technology based on natural features can be achieved without placing man-made patterns in real scene, it has a wide range of applications.

Currently, research of AR is more concentrated on registration method based on natural features. L. G. Hua used Harris Corner detection algorithm to extract image corners firstly, then adopted a correlation matching method to match the Gray correlation value of corners and got the initial set of matching corners [2]. However, it had low positioning accuracy and only pixel level corners were detected. The Scale-invariant feature transform (SIFT) algorithm was adopted to replace the Harris algorithm [3, 4]. Its superiority is that feature descriptor is invariant to scale, orientation, affine distortion, and partially invariant to illumination changes. The detected features are robust to changes in noise and minor changes in viewpoint. However, its computationalefficiency is so low that reduces 3D registration efficiency. Kanade-Lucas-Tomasi (KLT) feature tracker can solve this problem, but the displacement between adjacent frames was limited in one pixel [5].

ISSN: 2278 – 7798

Since time-consuming and large errors of traditional 3D registration restrict the application of the AR technology, a real-time registration method based on natural features is presented. Because Speeded up Robust Features (SURF) features detection algorithm has the same advantages as SIFT but higher computing efficiency, we adopt SURF algorithm to replace traditional SIFT algorithm firstly. Since detecting image features from every frame reduce 3D registration efficiency, we use a Lucas-Kanade (L-K) optical flow algorithm based on image pyramid to trace the detected features. The accumulation of registration error is avoided by adopting features redetection strategy that maintains 3D registration stability.

A.Augmented reality

Augmented Reality (AR) is a variation of Virtual Environments(VE), or Virtual Reality as it is more commonly called. VE technologies completely immerse a user inside a synthetic environment. While immersed, the user cannot see the realworld around him/her. In contrast, AR allows the user to see the real world, with virtualobjects superimposed upon or composited with the real world. Therefore, AR supplements reality, rather than completely replacing it. Ideally, it would appear to the user that the virtual and real objects coexisted in the same space, similar to the effects achieved in the film "Who Framed Roger Rabbit?"

B. Augmented Reality equipment

An AR system requires a camera to capture the real environment, a processing unit and an output device. Advanced AR systems use also a microphone combined with speech recognition algorithms for human-machine interaction, a sound system to present information and or artificial light system for a better performance of the vision system.
The AR environment is generated by libraries that identify the marker on the scene and retrieve their pose to display the virtual contents associated to each marker, Figure 1. Those contents may be static or dynamic animations that will be superimposed on captured images and they are adjusted to the pose of the marker; if the marker is rotated the animation is shown from a different point of view.

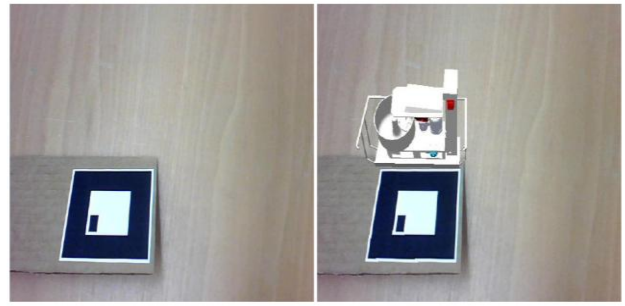

Figure 1. (Left) real environment; (right) virtual features on AR environment

All procedures to develop an AR application, since image acquisition until displaying the AR environment are available in AR libraries.

C. SIFT (Scale Invariant Feature Transform)

SIFT begins by taking a frame and creating several blurred version of it using a Gaussian blur, to increase the algorithms invariance to scale. Difference of Gaussiance (DoG) images are then created by subtracting adjacent images in the Gaussian blurring process, to approximate a Laplacian of Gaussian and produce an image that is invariant to image changes. The maxima and minima pixels are then found by iterating through all pixels and looking for pixels which are the maxima or minima of the 26 adjacent pixels (8 neighbors in the current DoG, 9 neighbors in the DoG produced by the previous level of blurring, and 9 neighbors in the DoG produced by the next level of blurring). These pixels are now considered key-points, and are further analyzed to filter out noisy keypoints. The contrast of the key-points with their neighbors is analyzed, and low contrast points are removed. Poorly localized points are then filtered out by removing points which have a high principle of curvature in one direction, but not in another, suggesting that it is lying on an edge of a surface. A 128 element vector key is calculated for each point by calculating a histogram of neighboring pixels.

ISSN: 2278 – 7798

Figure 2:Scale-Invariant Feature Transform

SIFT searches an image and finds unique and distinguishable key-points.

D. SURF (Speeded up Robust Features)

SURF algorithm is suitable especially for natural objects identification. However, many applications use this method to identify only a single object in an image. This is marker may appear in a scene. The most problematic part of the SURF marker identification is the matching of corresponding marker key points in both images. The key point's similarity is determined by gradient changes in the key point's neighborhoods (these are represented by feature descriptors). If the image contains areas with similar gradient changes, then such areas will be identified as the same or similar key points.

E.Viola Jones

With Viola and Jones Algorithm, it generates classifier by using Haar-like Feature and cascade based on Boost technology, and searches the results by using the Cascade structure. This algorithm is the Gray-based algorithm and constitutes the collection of very simple features to be studied. To extract these features, Haar-like Feature which uses the light-darkness distribution and has been proposed by Paul Viola et al[12], has been utilized as the feature collection for the purpose of face recognition.

II. LITERATURE SURVEY

Some of well-known algorithms for augmented reality presented in the literature are SIFT, SURF and so on. A survey of previous algorithms for AR has been discussed.

Monoslam: real-time single camera slam

This MonoSLAM [3]real-time algorithm which can recover the 3D trajectory of a monocular camera, moving rapidly through a previouslyunknown scene.Study system, which we dub MonoSLAM, is the first successful application of the SLAM methodology from mobile robotics to the "pure vision" domain of a single uncontrolled camera, achieving real time but drift-free performance inaccessible to Structure from Motion approaches. The core of the approach is the online creation of a sparse but persistent map of natural landmarks within a probabilistic framework. These add up to an extremely efficient and robust algorithm which runs at 30 Hz with standard PC and camera hardware. This work extends the range of robotic systems in which SLAM can be usefully applied, but also opens up new areas.

Henderson and Feiner (2009) [4] developed the field tested an AR system for maintenanceof armoured personnel carriers. The application, operational constraints andtasks supported were very similar to those required for the effective maintenanceof heavy mining equipment. The prototype was developed to support UnitedStates Marine Corps mechanics operating on LAV-25 armouredpersonnel carriers,which is a light-wheeled military vehicle. Like mining vehicles, the LAV-25 includes a large amount of electrical, mechanical, hydraulic and pneumatic infrastructure.

Wiedenmaier et al. (2003) [5] compared AR with paper instructions and expert guidance for a typical industrial assembling task and discovered that the assembling was completed in shortest time when the user was guided by an expert, followed by the use of AR support and in last place paper instructions.

*Nilsson and Johansson (2007) [6]*studied the use of AR as support in the assembling of a common medical device in a hospital and their analysis showed that the users were positive towards AR as a tool for instructions.

*Salonen and Sääski (2008) [7]*designed an AR system for the assembling of a 3D puzzle as well as a simulated industrial problem of assembling a tractor's accessory's power unit which used CAD models as graphical objects, but no user evaluations was performed as part of the study.

Robertson et al. (2008) [8] evaluated different AR solutions in the assembling of Lego blocks and found that a heads-up display with video see-through and context aware information was most efficient, but that the graphical information does not need to be located in the task area to be useful.

ISSN: 2278 – 7798

Radkowski and Stritzke (2012) [9] studied the use of AR for assembling by using virtual parts instead of real ones and they designed a solution based on a Microsoft Kinect video camera which was consider as intuitive to use according to the user tests.

*Hou et al. (2013) [10]*developed a prototype AR system for Lego assembling and found out that the system yielded shorter task completion times, less assembly errors, and lower total task compared to when using paper-based manuals.

Radowski et al. (2015) [11] investigated different visual features for AR-based assembly instructions and found out through user tests that the visual features used to explain a particular assembly operation must correspond to its relative difficulty level.

III. CHALLENGES

Beyond academic goals such as achieving precision and scalabilityof the researched algorithms, there is a set of practical concernsthat strongly affect the usability of AR experiences. These considerationsare only relevant for real-world applications of AR, and aretherefore not widely discussed in scientific literature. This may leadto the incorrect observation that these problems are not difficult ornot relevant for the success of AR. The following issues are relevantfor smartphones, but also for general purpose AR:

Real hardware development versus the"AR wish list":

The quality of cameras and other sensorsin current smartphone hardware is insufficient for high quality AR. Developers of AR applications would greatly benefit fromhardware advancements, such as stereo cameras, unified CPU/GPUmemory with random access, or Wi-Fi triangulation. Unfortunately, it is naıve to assume that mobile phones will be optimized for ARwithout a large established market. Any change in hardware configurationcosts millions of dollars in development, even more ifmarket expectations cannot be met afterwards.

*Dynamic scenery versus AR realism.*Current AR applicationsassume everything in a scene to be static. However, thereality is the exact opposite. Especially in outdoor scenarios, almosteverything is subject to change: people passing by, lightingand weather conditions, even buildings may be painted in differentcolors every few years. For localization, this causes a severe problem.

In dynamic scenes, basic assumptions that most algorithmsmake are violated right from the start. Assuming augmenting isa building facade while people pass by and partially occludeyour view. Due to missing occlusion reasoning, noticeable errorswill become apparent, no matter how good the visualization of augmentedcontent actually is and how powerful the hardware platformgets in the future. The lack of interaction between dynamic objectsand virtual content unconditionally harms realism in AR applications.Thus, the inclusion of object dynamic detection and trackingtechniques currently researched in CV are the key for high-qualityAR in the future.

Content creation versus registration: A large portion ofthe excitement about AR comes from the potential involvement ofend users in content creation. Personal content creation is a keyto actively integrate users rather than leaving them as passive observers.

However, basic mechanisms to facilitate this concept arestill missing. While interaction methods on mobile phones haveimproved greatly, the question how to conveniently and accuratelyregister even simple content in 6DOF using a 2D interface and noaccurate global model of the environment is still open. Assume thetask of augmenting a window on a building facade. Current methodsmight not even suffice for the task of simple tagging. There isno mechanism to input an arbitrary 3D position in open space, letalone specifying orientation. Current approaches typically use the(inaccurate) GPS position of the user, rather than of the object ofinterest to determine the tag.

IV. DISCUSSION

In above presented literature survey, we analyze regarding various or many existing research concept in terms of augmented reality, SLAM methodology, base mobile object tracking and detection approach, field tested with AR, comparison of AR with paper instruction, assembly of medical device in a hospital, assembly of a 3D puzzles, and so on which are given us to emerging method about solving the issue of AR operation.

V. CONCLUSION

In the offline mode, the detection with tracking work flow works quite well. However, on the mobile device, the detection is very slow, which directly affects accuracy of the tracking algorithm. Therefore, in the further improvement, a faster detection scheme shall be sued. SIFT is not suitable for the mobile platform and may be replaced by

ISSN: 2278 – 7798

FAST. From the homography of the image, we can get more information about the environment and some graph can be rendered in the images in real time to increase the interaction. Above all the reviews we have to use in our research Niblack+OpenSurf and Viola jones for solving the augmented reality issue.

REFERENCES

[1] Sidharth Bhatia, P. Vijayakumar "Augmented Reality Repair ugmented Guidance System" IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 4, No 3, July 2012ISSN (Online): 1694-0814.

[2] OhanOda, MenguSukan, Steven Feiner and Barbara Tversky "Poster: 3D Referencing for Remote Task Assistancein Augmented Reality" IEEE Symposium on 3D User Interfaces 201316-17 March, Orlando, FL, USA
978-1-4673-6098-2/13.

[3] A. J. Davison, I. D. Reid, N. D. Molton, and O. Stasse, "MonoSLAM: Real-time single camera SLAM," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 29, pp. 1052- 1067, 2007.

[4]Henderson, S, J, Feiner, S, 2009. Evaluating the benefits of augmented reality fortask localization in maintenance of an armored personnel carrier turret, in InternationalSymposium on Mixed and Augmented Reality, 2009 (ISMAR 2009). pp 135-144. Orlando, Florida.

[5]Wiedenmaier S, Oehme O, Schmidt L and Luczak H. Augmented Reality (AR) for Assembly Processes Design and Experimental Evaluation, International Journal of Human-Computer Interaction 2003; 16(3): 497-514.

[6] Nilsson S and Johansson B. Fun and Usable: Augmented Reality Instructions in a Hospital Setting. In: Proc. of 19th Australasian Conference on Computer-Human Interaction, 2007, pp. 123-130.

[7] Sääski J, Salonen T, Hakkarainen M, Siltanen S, Woodward C and Lempiäinen J. Integration of design and assembly using augmented reality. Micro Assembly Technologies and Applications 2008; 260: 395-404.

[8] Robertson CM, MacIntyre B and Walker BN. An Evaluation of Graphical Context when the Graphics are Outside of the Task Area. In Proc. of International Symposium on Mixed and Augmented Reality, 2008, pp. 73–76.

[9] Radkowski R and Stritzke C. Interactive Hand Gesture-based Assembly for Augmented Reality Applications. In: Proc. of the Fifth International Conference on Advances in Computer-Human Interactions, 2012.

[10] Hou L, Wang X, Bernold L and Love P. Using Animated Augmented Reality to Cognitively Guide Assembly. Journal of Computing in Civil Engineering 2013; 27: 439-451.

[11] Radkowski R, Herremaa J and Olivera J. Augmented reality based manual assembly support with visual features for different degrees-of-difficulty. International Journal of Human-Computer Interaction 2015.

[12] P. Viola, M. Jones, "Rapid object detection using a boosted cascade of simple features", IEEE Conf. on Computer Vision and Pattern Recognition, Kauai, Hawaii, USA, 2001.

ISSN: 2278 – 7798