

Enhancement the Augmented Reality Framework Using Proposed Registration Method

Yossra Hussain Ali, Abdul Ameer Abdulla and Ikhlas Watan Ghindawi*
Department of Computer Science, University of Technology, Baghdad-Iraq.
*Corresponding Author: ikhlas_watan@yahoo.com

Abstract

The augmented reality field is a visualization technique using the computers, it provides an additional information to support the work of multiple applications. This research presents a proposed method to solve the registration problem in the augmented reality field. This method dealing with a special case of the registration process, when the virtual object registered with the real object, which sometime has an addition above it. It ensures the continuance of the registration process when there is an object exist on or pass through a real object. This method converting the image colors combining with using the interest features principles in order to implement an accurate registration process. This method work accurately to detect the right position and monitor the motion in the used video, depending on finding the important common features and geometric transformation principles. [DOI: [10.22401/JNUS.20.2.20](https://doi.org/10.22401/JNUS.20.2.20)]

Keywords: Augmented reality, Image registration, Virtual reality, Alignment techniques.

1. Introduction

Augmented Reality (AR) is an innovation which permits PC delivered virtual symbolism to exactly overlay physical articles constantly. Not at all like Virtual Reality (VR), where the client is completely splashed in a virtual area, AR allows the client to interface with the virtual pictures using genuine things in a consistent way [1]. Azuma [1] gives a typically recognized significance of AR as a development which are, Consolidates real and virtual symbolism, Is intelligent continuously and Registers the virtual symbolism with this real reality.

As such there are numerous conceivable areas that could profit by the utilization of AR innovation, for example, engineering, entertainment and education [2]. Augmented Reality (AR) makes an area in which virtual items are superimposed on client's point of view of the genuine environment. Augmented reality gets a lot of consideration as another technique for showing data or expanding the reality of virtual situations [3]. Recognizing is a helpful procedure of registration process, in which gives recognition applications a basic approach to assess the position and introduction of an item in a video succession [4].

To execute an augmented reality framework, we should take care of a few issues. Geometric registration is particularly

the most essential issue on the grounds that virtual items should be superimposed on the perfect position as if they really exist in this real reality. One of the major methodologies to the enlistment between the genuine and virtual universes is the vision-based technique [3,5].

Registration techniques are a standout amongst the most essential issues in the field of Augmented Reality (AR). AR systems overlay virtual articles onto this genuine reality to help their client do a development in this genuine reality. A great part of the time, AR systems need to know exact geometric relations between genuine things and clients' audit position to discover virtual articles onto the appropriate position of this genuine reality. Registration strategies determine these geometric relations. Different registration techniques have been produced in the field of AR, and it is imperative to choose the proper tracking strategy for the application necessities [6,7,8,9,10].

2. The Registration Problem

Augmented Reality upgrades the world we see with virtual items. The situation of the virtual articles with respect to the genuine ones, alluded to as the enlistment, is of real significance keeping in mind the end goal to make the consolidated world credible. Enrolling the frameworks and information in an Augmented Reality execution is not a

simple assignment. A few variables must be considered. One such variable is to guarantee the enlistment is as right as would be prudent keeping in mind the end goal to perform exact and safe methods. Another, and frequently impacting, consider is the speed which the enlistment should be performed [11]. The idleness could diminish or decimate the sentiment the virtual and genuine universes existing together. This conveys us to the meaning of the enlistment issue: Making this present reality and the virtual world seem to exist together, not as particular universes, but rather as one unit. Lamentably, enrolling virtual articles with this present reality is a difficult issue to fathom.

One reason for this, is the way that people are great at distinguishing little misregistrations. Even only a couple of pixels wrong are effectively seen [12].

3. The Proposed Method

This method focuses on a special case of the registration problem, which is when the real object is exposed to the presence of other objects like (lines, letters or any moving object) above it. This method uses two videos (Original video and Replacing video) to represent the real and virtual objects. The real object is an object within the Original video and the Replacing video is represents the virtual object, these two objects will be registered using the proposed method.

The proposed method suggests an idea of using the image colors combining with the interest features principles in the registration process. The interest and matched features principles used to discover the exact position of a real object within the video frames, and a conditional replacement process implemented to align the real and virtual objects in exact position, the conditional replacement process requires converting the real object colors to binary in all video frames. This framework is implemented in three phases: Real Object Detection phase, Convert Image Colors phase and Conditional Replacement phase. This method is characterized by its simplicity and ability to distinguish the specific object in the video frame and separating it from the other objects or any addition that exist above it.

Fig.(1) illustrates the block diagram of the proposed method.

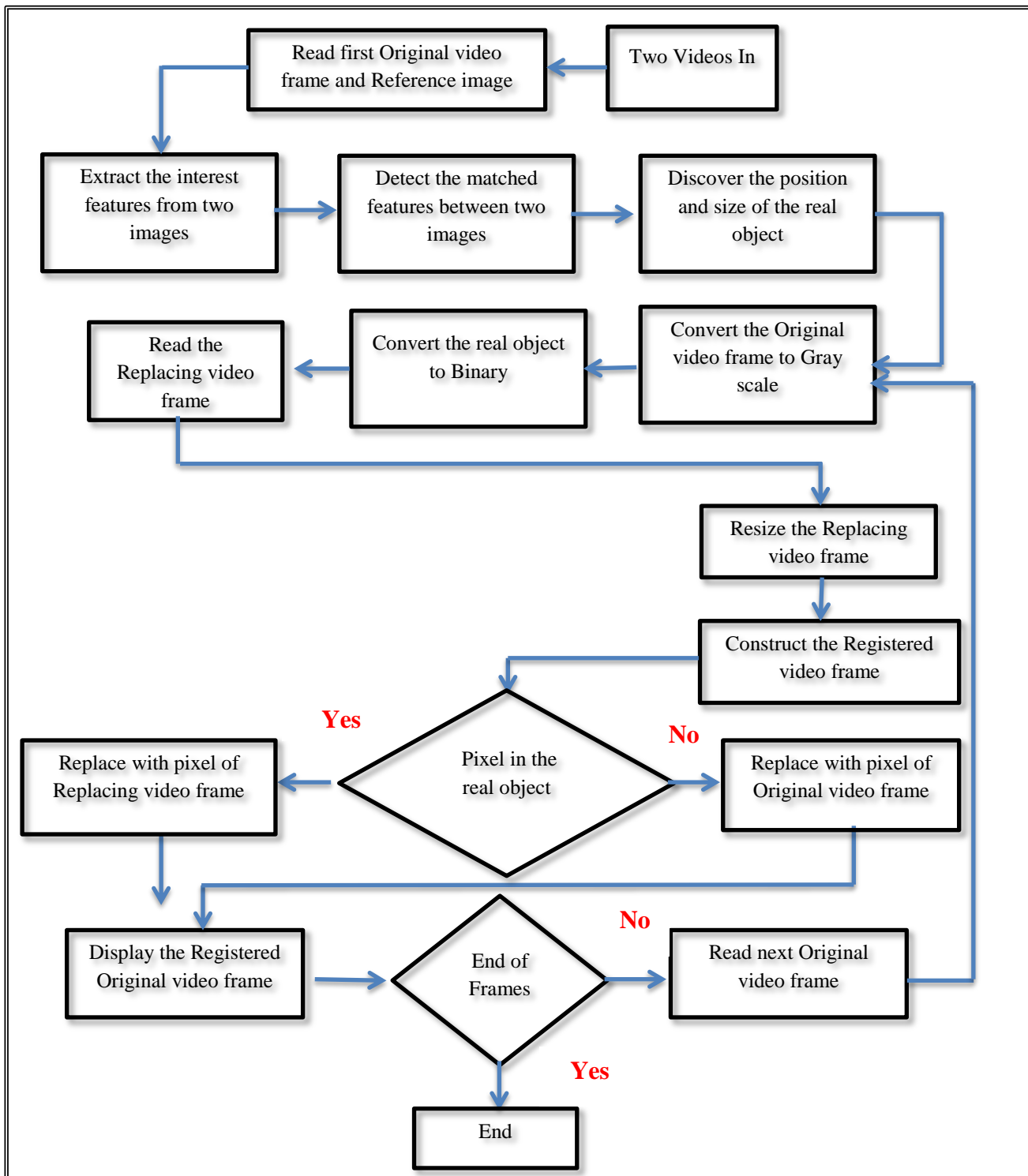


Fig.(3.1): The Block Diagram of the Proposed Registration Method.

3.1 Real Object Detection

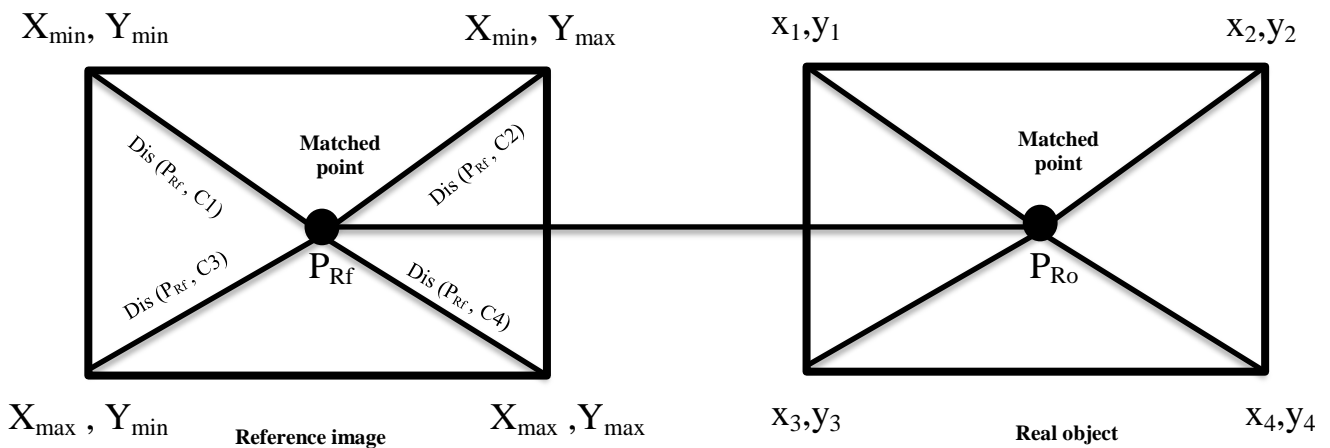
This phase uses the Harris Corner Detector to separate the interest features from the Reference image and the first Original video frame. The interest features which extracted from images are the corner features. The corner features are suitable for matching because it represent intersection point between two edges.

The RANSAC is used to find the matched features, these matched features used to discover the exact position of the real object. Discovering the position of the real object will be implemented once, because the real object is fixed in the same position on whole video frames.

After determine the exact position, this method identifies the four corners of the real

object (the top left (C1), the top right (C2), the lower left (C3) and the lower right (C4)), which leads to determine the object size. The proposed method identifies the size of the real object by selecting a pair of matched features, the distances between the matched point and the four corners of the Reference image will be measured. The four corners of the real object will be predicated using these distances. The main goal of using the four corners is to determine the size of the real object in video frames. The Original video frame and the reference image have common features which are matched features. In this step the matched point of the Reference image will be used to compute the distance between this point and the four corners of the Reference image. The process of determining the real object size used the first and fourth corner only, the first corner represents the minimum x, y and the fourth corner represent the maximum x, y of

the object area. In addition, the four corners found in order to predict the missing part of the object. Sometimes a part of the real object is covered by another static object, the four corners assist to identify the missed part. This phase identifies the position of the real object in the Original video frame by selecting a pair of matched features randomly from matched features set. The distances between the selected matched point and the four corners of the Reference image will be measured. The four corners of the real object will be calculated using these distances. The Original video frame and the Reference image have common features which are matched features. The selected matched feature has two points, P_{Rf} which is the matched point in the Reference image, and P_{Ro} , is the matched point in the real object.



The matched point of the Reference image (P_{Rf}) will be used to compute the distance between this point and the four corners of the Reference image.

$Dis(P_{Rf}, C1)$ is the distance between matched point $P_{Rf}(X_{Rf}, Y_{Rf})$ and the X_{min}, Y_{min} of the Reference image, which represents the top left corner of the Reference image.

$Dis(P_{Rf}, C2)$ is the distance between matched point $P_{Rf}(X_{Rf}, Y_{Rf})$ and the X_{min}, Y_{max} of the Reference image, which represents the top right corner of the Reference image. $Dis(P_{Rf}, C3)$ is the distance between matched point $P_{Rf}(X_{Rf}, Y_{Rf})$ and the X_{max}, Y_{min} of the Reference image, which represents the lower left corner of the Reference image.

$Dis(P_{Rf}, C4)$ is the distance between matched point $P_{Rf}(X_{Rf}, Y_{Rf})$ and the X_{max}, Y_{max}

of the Reference image, which represents the lower right corner of the Reference image.

The displacement from matched point (P_{Rf}) to four corners of the Reference image will be used to determine the four corners of the real object in the video frame using the matched point (P_{Ro}) of the real object.

$$C_{R1} = Dis(P_{Rf}, C1.X) - X_{Ro}, Dis(P_{Rf}, C1.Y) - Y_{Ro} \dots \dots \dots (1)$$

$$C_{R2} = Dis(P_{Rf}, C2.X) - X_{Ro}, Dis(P_{Rf}, C2.Y) + Y_{Ro} \dots \dots \dots (2)$$

$$C_{R3} = Dis(P_{Rf}, C3.X) + X_{Ro}, Dis(P_{Rf}, C3.Y) - Y_{Ro} \dots \dots \dots (3)$$

$$C_{R4} = Dis(P_{Rf}, C4.X) + X_{Ro}, Dis(P_{Rf}, C4.Y) + Y_{Ro} \dots \dots \dots (4)$$

Where the C_{R1} , C_{R2} , C_{R3} , C_{R4} are the four corners of the real object.

Using the four corners of the real object can be calculated the size of the real object. The process of determining the real object size uses the first and fourth corner only, the first corner represents the minimum x , y and the fourth corner represent the maximum x , y of the object area. The difference between x_1 and x_4 produces the X dimension of the real object and the difference between y_1 and y_4 produces the Y dimension of the real object.

$$X_{\text{real}} = x_4 - x_1. \dots\dots\dots (5)$$

$$Y_{\text{real}} = y_4 - y_1. \dots\dots\dots (6)$$

Where the x_1 is the x - axis of the C_{R1} and x_4 is the x - axis of the C_{R4} , and y_1 is the y - axis of the C_{R1} and y_4 is the y - axis of the C_{R4} .

3.2 Convert Image Colors

The area covered by the real object will be used in all Original video frames as a fixed area, this area represents the real object in all Original video frames. To deal with any object exist above the real object, the proposed method converts the real object area to binary. This method converts the Original video frame to gray scale as a preliminary step to convert the real object to a binary object. The threshold value, that used to convert the real object pixels from gray level to the binary, is the maximum value in gray level of the real object area pixels. Converting the real object to a binary object allows the possibility to identify any object or any addition is exist on the real object and provides the possibility to perform a conditional alignment process easily and effectively. The main idea is, just the pixels of the real object will take the value one and the other pixels will belong to other objects. All pixels of the real object area, which has no addition above it, will take the value one. When any addition existed above the real object, the pixel values of the real object area will take the values of ones or zeros, the value one represents the real object pixels and zero represent the addition.

3.3 Conditional Replacement

To register the Replacing video frame with the real object within the Original video frame, the Replacing video frame will be resized to

the size of the real object and the conditional replacement process will be used to align the real and virtual object. The pixels of the final registered video frame are selected conditionally from the Replacing video frame (virtual object) or the Original video frame. The pixel value of each location in the registered video frame depends on the pixel value of the real object in the same location. The pixel of the registered video frame will be take it's value from the Original video frame in the case of, the value of the real object pixel is zero. And in the case of, the value of the real object pixel is one, the pixel of the registered video frame will be take it's value from the Replacing video frame. This process is repeated at each Original video frame. Algorithm 1 shows the main steps of the proposed method.

Algorithm (1): The proposed Algorithm**Input : Original Video, Reference image, Replacing video.****Output: Registered video frames.**

Process:

Step 1: Read the first original video frame.

Step 2: Read the Reference image.

Step 3: Extract interest features using the Harris corner detector from two images.

Step 4: Find the matched features between two images using the **RANSAC** method.

Step 5: Find the four corners of the real object using the matched features.

Step 6: Calculate the size of the real object using the four corners .

Step 7: Transform the original video frame to gray scale.

Step 8: Transform the pixels of the real object to the binary.

Step 9: Read the replacing video frame.

Step 10: Scaled the replacing video frame to the real object size.

Step 11: Align the real object and the scaled replacing video frame

Step 12: Display the registered video frame.

Step 13: Read the next original video frame.

Step 14: While not end of original video frames Go to Step 9.

Step 15: End.

Fig.(2): Illustrates random five video frames after performing the registration process.



Fig.(2): Random video frames after registration process.

4. The Data and Results

The Fixed Objects - Binary Registration Method, has been implemented by four different videos with different number of frames and with resolution 720×1280. Various criterion measures have been used to evaluate the performance of the proposed method, the Target Registration Method (TRE), Root Mean Square Error (RMSE), and Structural Similarity Index for Measuring Image Quality (SSIM), The results of implementing the proposed method on four videos are summarized in Table (1).

Table (1)
The performance results of the proposed method.

Proposed Method	NO. Frames	TRE	RMSE	SSIM	Time/Sc
Video 1	963	0.00062465	0.00113241	0.904132	1.09691
Video 2	1865	0.000993115	0.00841811	0.865231	1.52143
Video 3	1880	0.000864631	0.00795421	0.894258	1.15113
Video 4	2327	0.00093210	0.00832534	0.902911	1.16221

This method deals with the video frame as a separate individual, where the registration process implementation in each frame without requiring any information from other frames. This characteristic is reflected positively on the accuracy of the registration and the time spent on implementation. Also, the fixed position of the real object contributed significantly to get the reliability of the results obtained. The Table 1 illustrates the results of four videos with different number of frames, the first video get the highest accuracy results and the shortest time, also the rest of videos satisfied a high level of accuracy compared to the different sizes of real objects and different colors in each one.

5. The Conclusion

In the Augmented Reality environment, the image registration is an essential issue in augmentation process. This research provides a solution to the problems facing the registration process. These proposed solutions have succeeded to achieve high reliability and accuracy, as well as solved the problem of delays in the registration process. Through implementation of the proposed methods can conclude the following:

- Relying on the interest features that forming the images, reflected positively on the precise of the registration process. By

using the important features principle, which provide a small work domain, the proposed methods succeeded in decrease the delay of the registration process.

- Dependence on the correct matched features leads to high level of accuracy and reliability, where the accuracy of the correct matches reflected in the accuracy of the detection process of the object position.
- Determining the exact position of an object in the registration process has great importance, any error in detecting the position would lead to visible misregistration and has an effect on the whole registration framework.
- The combination between the feature based detection technique and the converting process of the image colors, lead to more accurate registration process.
- The proposed method satisfied high level of accuracy, which is 0.00062465, and it's implemented in 1.09691 seconds.

6. References

- [1] R. T. Azuma, "A Survey of Augmented Reality," *Presence*, Vol. 6, No. 4, pp. 355–385, 1997.
- [2] K. Kiyokawa, "An introduction to head mounted displays for augmented reality". In M. Haller, M. Billinghurst, B. H. Thomas

- Eds, Emerging Technologies of Augmented Reality Interfaces and Design, pp. 43-63, Idea Group, Hershey, 2007.
- [3] S. Hare, A. Saffari, P.H.S. Torr, "Efficient Online Structure Output Learning for Keypoint-based Object Tracking", In Proc. Of CVPR, pp:1894-1901, 2012.
- [4] O. Ibrahim, Y. D. Bashi, "Face Recognition Based on PCA, LBP and SVM Techniques", Eng. & Tech. Journal, Vol.33, Part (B), No.3, 2015.
- [5] Y. Yokakohji, "Accurate Image Overlay on Video See-Through HMDs Using Vision and Accelerometers", Proc. of IEEE VR 2000, pp. 247-254, 2000.
- [6] K. Satoh, M. Anabuki, H. Yamamoto, H. Tamura, "A hybrid registration method for outdoor augmented reality". Proceedings of IEEE and ACM International Symposium on Augmented Reality, pp: 29-30, 2001.
- [7] J.D. Simmon, A. D. Cheok, "Augmented Reality Camera Tracking with Homographies", IEEE Computer Graphics and Applications, Vol. 22, No. 6, pp: 39-45, 2002.
- [8] T. Auer and A. Pinz, "Building a Hybrid Tracking System: Integration of Optical and Magnetic Tracking", Proc. Of IEEE International Workshop on Augmented Reality, pp.13-22, 1999.
- [9] W. K. Obeidy, H. Arshad, S.A. Chowdhury, B. Parhizkar, J. Huang, "Increasing the Tracking Efficiency of Mobile Augmented Reality Using a Hybrid Tracking Technique", Advances in visual Informatics, 2013.
- [10] X. Yang, K.T. Cheng, "Local Difference Binary for Ultrafast and Distinctive Feature Description", TPAMI, 36(1), pp: 188-194, 2014.
- [11] A. Fuhrmann, D. Schmalstieg, W. Purgathofer, 1999, "Fast Calibration for Augmented Reality", Proceedings of the ACM symposium on Virtual reality software and technology, <http://www.cg.tuwien.ac.at/TR/99/TR-186-2-99-16Paper.pdf>, 1999.
- [12] Gonzalez, R. C., Woods, R. E., Digital Image Processing, second edition, Prentice Hall, 2002.