TELKOMNIKA, Vol. 11, No. 7, July 2013, pp. 3568 ~ 3575
e-ISSN: 2087-278X

Mobile Camera as a Human Vision in Augmented Reality

Edmund Ng Gaip Weng*, Rehman Ullah Khan, Shahren Ahmad Zaidi Adruce, Oon Yin Bee
Faculty of Cognitive Sciences and Human Development (FSCSHD), Universiti Malaysia Sarawak (UNIMAS) 94300 Kota Samarahan, Sarawak, Malaysia
Tel: +60(82) 581491, +60(82) 581492, +60(82) 581493, Fax: +60(82) 581567
*Corresponding author, e-mail: nggiapweng@yahoo.com

Abstract

The real world objects can be recognized by using marker based and marker-less augmented reality systems. Mostly, the previous developers used markers based augmented reality systems. However, those systems actually hide the reality and it was also difficult to keep the markers everywhere. Furthermore, the previous marker-less approaches use client-server architecture, which is drastically affected by network latency. Smartphone camera is matured enough that it can recognize real world objects without markers. It can guide users about their location and the direction in a convenient way. The use of Smartphone is best suited for outdoor mobile augmented-reality applications. Therefore, a marker-less natural features based tracking system in mobile augmented reality was formulated. In the adapted framework, the state-of-the-art algorithm (speed up robust features) was modified for computing image features from live mobile camera image and compares with locally stored images features for recognition. Moreover, the local static database of location tagged image features using SQLite was implemented to bypass the server. The proposed system was tested in a mobile AR-prototype application using iPhone called UNIMAS Guide. It was found from the results that the adapted marker-less system could recognize the real world objects in speedy, easy and convenient way. This technology can be applied in tourism industry, surgery and educational fields.

Keywords: *augmented reality, marker-less, outdoor, image features, image recognition, static database*

Copyright © 2013 Universitas Ahmad Dahlan. All rights reserved.

1. Introduction

Augmented reality using emerging technologies such as global positioning system (GPS), accelerometer, gyroscope, compass and mobile vision, provides a best opportunity to Smartphone users to explore their surroundings. The real world objects can be recognized by using marker based and marker-less augmented reality systems. Mostly, the previous developers used markers based augmented reality systems. However, those systems actually hide the reality and it was also difficult to keep the markers everywhere. Furthermore, the previous marker-less approaches use client-server architecture, which is drastically affected by network latency. The markers-based augmented reality was applied in different fields like medical visualization, maintenance and repair, navigation and entertainment. However, the markers are not suitable for outdoor mobile augmented reality because markers hide the reality and need to keep everywhere [1]. Its range is also very limited and end-users often don't like them. Marker-less natural features based approache [2], can recognize real world objects, such as sights, buildings, and living beings and overcome these limitations. Robust feature descriptors such as SIFT [3], SURF [4], and GLOH [5] are most suitable for applications such as image recognition [6] and image registration [7]. These descriptors are stable under different viewpoints and lighting conditions. These descriptors are ideally suited for searching image databases because they are representing feature points as high-dimensional vectors. However, tracking from natural features is a complex problem and usually performed on a remote server [8], [9], [10]. It is therefore a challenging task to use natural feature tracking in mobile Augmented Reality applications. However, the mobile phones are very inexpensive, attractive targets for outdoor AR. The improvements in Smartphone capabilities and great potential of

Received January 6, 2013; Revised April 3, 2013; Accepted April 15, 2013

computer/mobile vision motivated us to implement marker-less natural feature based mobile augmented reality.

A recent example of an AR object tracking application is the Sudoku Grab [11]. This application can track a Sudoku puzzle and solve it, adding the missing numbers in the empty Sudoku slots. Georg Klein et al. created an application which can analyze the surroundings, making it possible to render different 3D characters look like they are sitting on the desk in the physical world [12]. Occipital developed Red Laser which can track its environment. It is a good example to scan and analyze the camera view by iPhone [13].

Nokia's MARA project by Kähäri and Murphy [14] does not perform any image analysis, instead it uses an external GPS for localization and an inertial sensor to provide orientation. Phone Guide [15] is client-server object recognition systems. The system employs a neural network trained to recognize normalized colour features and is used as a museum guide. Seifer et al. [16] used a mobile system based on a hand-held device, GPS sensor, and a camera for roadside sign detection and inventory. Their algorithm has good quality results in mobile settings.

For object detection and recognition, Fritz et al. [17] used a modified version of the SIFT algorithm. The system uses a client-server architecture, where a mobile phone client captures an image of an urban environment and sends it to the server for analysis. The SURF algorithm has been used successfully in a variety of applications, including an interactive museum guide [18]. Local descriptors have also been used for tracking. Skrypnyk and Lowe [19] use the SIFT features for recognition, tracking, and virtual object placement. Camera tracking is done by extracting SIFT features from a video frame, matching them against features in a database, and using the correspondences to compute the camera pose. Takacs et al. [20] applied SURF features using video coder motion vectors for mobile augmented reality applications.

Yeh et al. [21] proposed a system for determining a user's location from a mobile device via image matching. The authors first build a "bootstrap database" of images of landmarks and train a CBIR algorithm on it. Since the images in the bootstrap database are tagged with keywords, when a query image is matched against the bootstrap database, the associated keywords can be used to find more textually related images through a web search. Finally, the CBIR algorithm is applied to the images returned from the web search to produce only those images that are visually relevant.

The purpose of this research is to find solution for marker-less mobile augmented reality and solve the client-server problem. The proposed framework can discover the surroundings and provide information about different objects. This system can be used easily in different fields of life just by changing images features database. The system, uses modified version of SURF [4] for object tracking and recognition. In this research the researchers modified SURF because original SURF uses IPLImage and iPhone camera generates UIImage. The mobile memory was saved by reducing 64-element descriptor to 32-element without affecting its efficiency. The number of octaves was also reduced to two and number of intervals per octave to three. In the formulated approach, feature points were extracted from incoming video frames of iPhone camera at run-time and matched against a local database of feature points and GPS data stored in mobile. The GPS data was used as index and primary key in database design. The searching query was optimized by using this primary key. After successful matching a homography matrix and transformation matrix were calculated from matching points using computer vision techniques and mobile sensing technology such as accelerometer and compass [12], [23].

2. Methodology

iPhone 3GS and open surf library [24] were used for features extraction and creating the descriptors. Furthermore, Open CV 2.2 [25] was used for image conversion from UImage to IPLImage and vice versa.

2.1. Open SURF

Open-SURF library [24] uses the SURF [4] algorithm which is one of the best interest point detectors and descriptors currently available. It has best performance as compared to other interest points descriptors like [3] and GLOH as shown by Mikolajczyk [5]. As mobile has a limited resources so the open surf library was modified. The descriptor size was reduced to 32-

element, number of octaves to two and number of intervals per octave to three. Like many feature descriptor algorithms it also extracts regions of interest that tend to be repeatable and invariant under transformations such as brightness or perspective changes. An image is analyzed at several scales, so interest points can be extracted from across all possible scales. Additionally, the dominant orientation of each of the interest points is determined to support rotation-invariant matching. An example image and its detected interest points are shown in Figure 1.

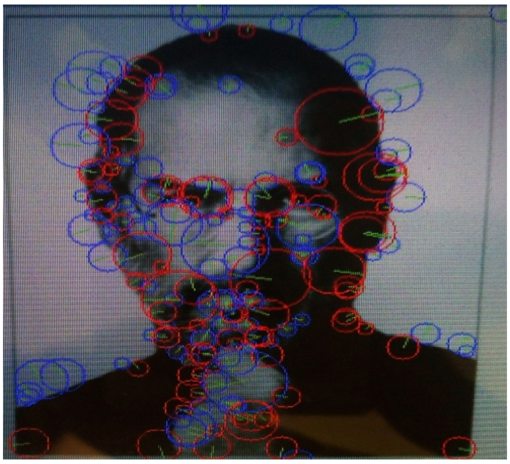

Figure 1. SURF interest points detection on iPhone 3GS, squares indicate found features

In the descriptor computation step, each extracted interest point defines a circular region from which one descriptor is computed. Each interest point is associated with a 32-element descriptor. It was found that the descriptor size in case of iPhone image thus ranges between 0KB and 190KB per image, with an average of roughly 35KB. In this way a database of more than one hundred images features can be loaded in one application.

2.2. Feature Database

A location tagged features database along with a framework given by [2] were used for getting GPS data and related information. The information was stored in a SQlite database. As SQLite is a server-less static database system so it was used in the resource folder. The latitude and longitude of location were employed as a primary key and index in the database design as shown in Figure 2.

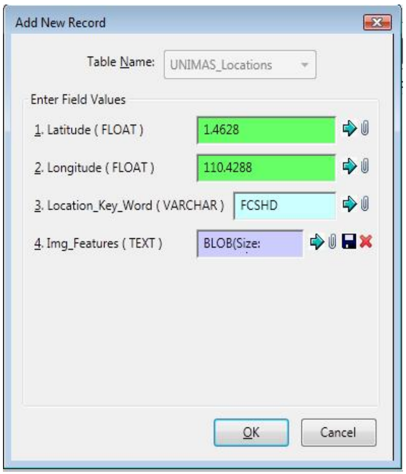

Figure 2. Record for storing image features

2.3. System Overview

The system is fully implemented on the mobile device, and runs at close to real-time, while maintaining excellent recognition performance. When the system starts it gets video frames from live camera. As iPhone uses UIImage class and OpenCV library uses the IplImage class to hold image data. OpenCV was used to create an IplImage image from a UIImage image, and a UIImage image from an IplImage image. Image features from live camera frames were extracted and at the same time the user location was calculated from GPS data. Next, these features are compared with the features stored in database. The query optimize was optimized by the using GPS position as a primary key and index. In this way the exact record was accessed without scanning the whole database. Then the surf descriptor comparison functionality was used to recognize the object. Once the object is recognized then the object's homgraphy matrix and transformation matrix are calculated from the matched points orientations. The conceptual diagram of proposed framework is shown in Figure 3.

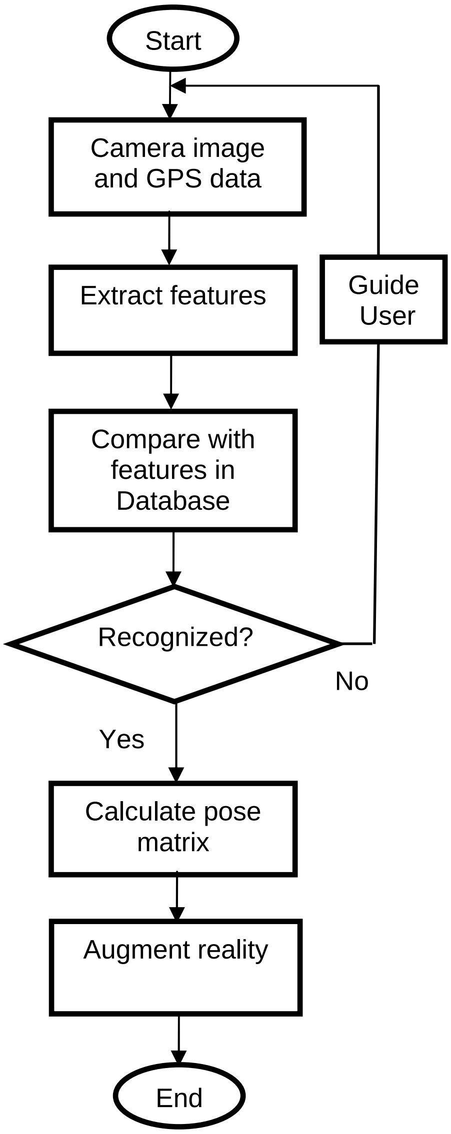

Figure 3. Conceptual diagram of proposed framework

The researchers intended that the whole process to be done directly on a mobile device for several reasons. The proposed framework significantly reduces the system latency as it is independent of server. Moreover, it can also work at any location of the user.

3. Results and Discussion

The Open SURF library was applied as a feature descriptor in the system for iPhone 3GS devices. Tests were conducted and results were recorded using iPhone 3GS. The scale level pyramid consists out of two octaves with three layers each. The system was tested outside the campus. The Steve Job's picture features were compared for recognition quality with Qualcomm SDK [26] and the proposed system. It was found that the proposed framework over performs the Qualcomm SDK [26] as shown in Figures 4 and 5.

From Figures 1 and 5 and from the analysis result it is obvious that Qualcomm can't track well this image. But the proposed sytstem can track this well even in different lighting conditions and from different viewpoints. In university campus the proposed system can recognize all those departments and centers which were included in the feature database from different viewpoints as shown in Figures 6 and 7.

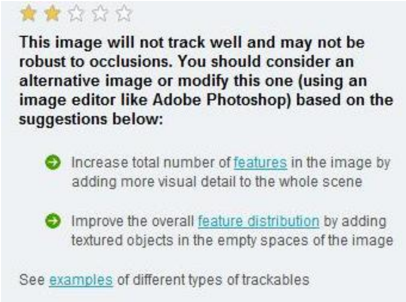

Figure 4. Analysis result of Qualcomm SDK

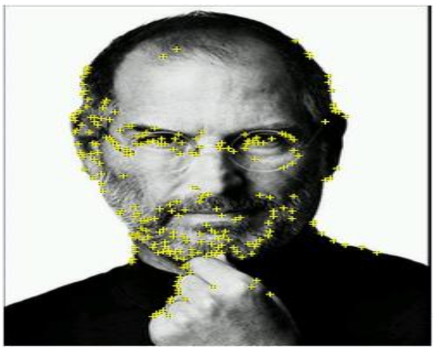

Figure 5. Features detected by Qualcomm SDK

Figure 6. Center of excellence monogram recognized by proposed framework in bad lighting condition

Figure 7. Recognition of Faculty logo by proposed framework

A survey of the state of the art AR mobile browsers was given by Butchart [27] in March 2011.To compare the proposed system with currently available browsers, selected columns of the survey table were evaluated.

Table 1. AR Browsers Comparison

Product	Marker based	Marker-less	Offline mode	Platform
Layar	No	No	Online only	iPhone, Android, Symbian
Junaio	Yes	Yes	Online only	iPhone, Android, Nokia (N8)
Wikitde API	No	No	Offline	iPhone, Android
Wikitde Worlds	No	No	cacheable	iPhone, Android, Symbian
Sekai Camera	No	No	Online only	IPhone, Android, iPad, iPodTouch
Libre Geosocial	source	plugin	Online only	Android

Two criteria such as marler-less and offline mode were used for evaluation of Table 1. It was discovered that, the most of the browser can't support marker-less. However, some of the browsers can support marker-less which depends on desktop powerful server using network. Since the previous browsers have few limitations such as network latency, uploading and downloading of contents. The proposed framework covered the above mentioned problems by using local feature database.

4. Conclusion

A marker-less vision-based AR system was presented that recognize and track real world objects in real-time without markers. Significant improvement in performance was achieved by using static database of image features. The proposed framework explored novel research directions. These were previously only possible with desktop computers and now can be executed with a mobile device. The developer and programmers can apply the proposed framework in tourism industry, games and educational fields.

References

[1] Reitmayr G, Schmalstieg D, editors. *Location based applications for mobile augmented reality*. 4th Australasian User Interface Conference. 2003. Adelaide, Australia: Australian Computer Society, Inc.

[2] Edmund Ng Gaip W, Rehman Ullah k, Shahren Ahmad Zaidi A, Oon YB. A framework for Outdoor Mobile Augmented Reality. *IJCSI International Journal of Computer Science Issues*. 2012; 9(2): 419-23.
[3] Lowe DG. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*. 2004; 60(2): 91-110.
[4] Bay H, Tuytelaars T, Van Gool L. Surf: Speeded up robust features. *Computer Vision–ECCV 2006*. 2006: 404-17.
[5] Mikolajczyk K, Schmid C. A performance evaluation of local descriptors. *IEEE transactions on pattern analysis and machine intelligence*. 2005; 27(10): 1615-30.
[6] Darrell KGaT, editor. *The Pyramid Match Kernel: Discriminative Classification with Sets of Image Features*. ICCV. 2005.
[7] Lowe MBaD. Automatic Panoramic Image Stitching Using Invariant Features. *International Journal of Computer Vision*. 2007; 74: 59–77.
[8] M Pielot NH, C Nickel, C Menke, S Samadi, and S Boll editor. Evaluation of Camera Phone Based Interaction to Access Information Related to Posters. *Mobile Interaction with the Real World*. 2008.
[9] Aydın B, Gensel J, Calabretto S, Tellez B. ARCAMA-3D–A Context-Aware Augmented Reality Mobile Platform for Environmental Discovery. *Web and Wireless Geographical Information Systems*. 2012: 17-26.
[10] Reitmayr G, Schmalstieg D, editors. Data management strategies for mobile augmented reality. STARS 2003; Tokyo, Japan.
[11] Schall G, Wagner D, Reitmayr G, Taichmann E, Wieser M, Schmalstieg D, Hofmann-Wellenhof B, editor. Global Pose Estimation using Multi-Sensor Fusion for Outdoor Augmented Reality. 8th *IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR 2009)*; 2009.
[12] Honkamaa P, Siltanen S, Jäppinen J, Woodward C, Korkalo O, editor. *Interactive outdoor mobile augmentation using markerless tracking and GPS*. Virtual Reality International Conference (VRIC). 2007; Laval, France.
[13] Greening C. iPhone Sudoku Grab. Japan 2008 [cited 2012 May 16 2012]; Available from: http://www.cmgresearch.com/sudokugrab/.
[14] Murray GKaD, editor. *Parallel tracking and mapping on a camera phone*. International Symposium on Mixed and Augmented Reality (ISMAR). 2009.
[15] Occipital L. Redlaser 2011 [cited 2012 May 16 2012]; Available from: http://redlaser.com/.
[16] Greene K. Hyperlinking reality via phones. *MIT Technology Review*, November/December. 2006.
[17] Bruns E, Bimber O. Adaptive training of video sets for image recognition on mobile phones. *Personal and Ubiquitous Computing*. 2009; 13(2): 165-78.
[18] Seifert C, Paletta L, Jeitler A, Hödl E, Andreu JP, Luley P, et al. Visual object detection for mobile road sign inventory. *Mobile Human-Computer Interaction–MobileHCI*. 2004: 587-90.
[19] Fritz G, Seifert C, Paletta L, editors. *A mobile vision system for urban detection with informative local descriptors*. ICVS '06: Fourth IEEE International Conference on Computer Vision Systems. 2006: IEEE.
[20] Bay H, Fasel B, Van Gool L, editors. *Interactive museum guide: Fast and robust recognition of museum objects*. 4th international conference on Adaptive multimedia retrieval: user, context, and feedback 2007; Springer-Verlag Berlin, Heidelberg: Citeseer.
[21] Skrypnyk I, Lowe DG, editors. *Scene modelling, recognition and tracking with invariant image features*. ISMAR '04:Third IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'04). 2004.
[22] Takacs G, Chandrasekhar V, Girod B, Grzeszczuk R, editors. *Feature tracking for mobile augmented reality using video coder motion vectors*. Sixth IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'07). 2007.
[23] Yeh T, Tollmar K, Darrell T, editors. *Searching the web with mobile images for location recognition*. IEEE. 2004.
[24] Evans C. Open SURF Computer Vision Library. [cited May 17, 2012]; Available from: http://www.chrisevansdev.com/computer-vision-opensurf.html.
[25] Open Source I. Open Computer Vision Library. sourceforge; [cited May 17, 2012]; Available from: http://sourceforge.net/projects/opencvlibrary/files/.
[26] Developers. Q. Qualcomm SDK. [cited May 24, 2012]; Available from: https://ar.qualcomm.at/qdevnet/sdk/ios.
[27] Ben Butchart E. Augmented Reality for Smartphones: A Guide for developers and content publishers Report. UK: JISC Observatory. May 2011. Report No: 1.1.