Journal of Theoretical and Applied Information Technology
15th November 2016. Vol.93. No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A NEW ILLUMINATION INVARIANT FEATURE BASED ON FREAK DESCRIPTOR IN RGB COLOR SPACE

[1]SIOK YEE TAN, [2]HASLINA ARSHAD, [3]AZIZI ABDULLAH
[1]Research Scholar, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Malaysia
[2]Associate Professor, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Malaysia
[3]Senior Lecturer, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Malaysia
E-mail: [1]p72765@siswa.ukm.edu.my, [2]haslinarshad@ukm.edu.my, [3]azizia@ukm.edu.my

ABSTRACT

Mobile markerless augmented reality (AR) requires a descriptor that can act in real time and at the same time robust to illumination invariance. Several color descriptors have been proposed to increase illumination invariance such as RGB-SIFT. According to the previous study, FREAK descriptor has been identified as the most suitable descriptor for mobile markerless AR due to FREAK reviews less computation time compared to other descriptors. However, color information presents important information has been largely ignored by FREAK. Hence, this research proposed a new illumination invariant feature based on FREAK descriptor in RGB color space for mobile markerless AR. Red, green and blue color channel will applied separately to FREAK descriptor. Hence, there are four types of descriptor been introduce; Red-FREAK, Green-FREAK, Blue-FREAK and RGB-FREAK. The proposed color FREAK descriptors are evaluated using images of different illumination color from Amsterdam Library of Object Images data set. The results show that the proposed color FREAK descriptors have enhanced the accuracy under illumination invariance for mobile markerless AR.

Keywords: *Markerless Augmented Reality, Feature Descriptor, RGB, Illumination Invariance, FREAK*

1. INTRODUCTION

The ultimate goal of AR is to enhance and improve user's perception of the surroundings by combining computing and display technology. Continuous tracking in AR applications is important for augmenting the digital content on top of the real world [1]. A common approach is to detect and extract features using detector and descriptor before matching process [2]. For this reason, the choice of feature descriptor is of great influence on the tracking accuracy. FREAK [3] descriptor has been proposed as the most suitable descriptor for mobile markerless AR application due to the less consuming time of feature extraction [4]. FREAK is implemented mainly for gray images. However, color channel like red, green and blue presents important information in feature description and matching tasks as in the human vision system. Color provides valuable information in feature extraction task. Many images can be unmatched if color contents are ignored [5]. Hence, this research proposed a new illumination invariant feature based on FREAK descriptor in RGB color space to improve the accuracy during illumination changes for mobile markerless AR.

Due to good result of the RGB-SIFT [6], this research proposed a new illumination invariant feature based on FREAK descriptor in RGB color space. A RGB image is split into three 8-bit channels (red, green and blue).

The RGB-FREAK applies FREAK directly to each of the three channels and each feature extracts three n-element. Each channel need to employ a normalization step to converts the ranges. This three n-element will then concatenate into one $3 \cdot n$-element feature.

RGB-SIFT had been proposed by [6, 7] which is invariant to illumination changes, but the processing time required for RGB-SIFT is too high for a mobile markerless AR application which required real time performance. Hence, binary descriptors come in handy [3, 8-10]. The most suitable descriptor for mobile markerless AR has been identified as FREAK descriptor because it

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

consumes the least computation time compared to other descriptors and also robust to scale and rotation changes [11]. However, color information which provides valuable information has been largely ignored by FREAK. Hence, this work proposed a new illumination invariant feature based on FREAK descriptor in RGB color space in mobile markerless AR to improve the tracking accuracy toward illumination changes.

2. RELATED WORK

This section presents the related works of this research such as markerless AR, FREAK, RGB color channel, RGB-SIFT and data set used in the evaluation of feature descriptor algorithm.

2.1 Markerless AR

Markerless AR systems attempt to track the features that naturally occur in the environment across consecutive video frames. By detecting, extracting and matching features across frames, the application can use a combination of triangulation, bundle adjustments and filtering to determine camera pose and follow by augmenting a 3D object on top of the natural environment.

Markerless AR tracking is a complex problem and usually demands high computation power. It is therefore difficult to use robust natural feature tracking in mobile applications of AR, which runs with limited computational resources such as on mobile phone or tablet [12]. Feature descriptor is one of the important component used in markerless AR to extract the natural features. Therefore performance of feature descriptor had been evaluated in term of speed and accuracy by [4]. There are five main requirement for an efficient mobile markerless AR; low computation time (act in real time), robust to illumination invariance, scale invariance, rotation invariance and view point change [1, 13-15]. To allow a mobile markerless AR to act in real time, the minimum frame rate is 10 frames per second. This is because [16] had introduced the first AR system was capable of running at 10 frames per second or 100ms per frame.

FREAK descriptor had been identify as the most suitable descriptor for mobile markerless AR because FREAK descriptor fulfill almost all the requirement of a mobile markerless AR application.

2.2 FREAK

Neuro-scientific research suggests that the human retina extracts details from an image using Difference of Gaussian at different sizes with the distribution of reception cells focused at the center, the authors [3] suggest using circular sampling points with higher density near the center. The closer to the center, the denser the points are. The descriptor is named as FREAK. To match the retina model, FREAK uses Gaussian kernels with different sizes to smooth the intensities of each sampling point. FREAK outperformed recent state-of-the-art feature descriptors while remaining simple and faster with lower memory load, hence suitable for real-time image matching performance. Generally, single bit of binary descriptor are calculated by comparing the intensity value of the point x in a sampling pair with the intensity value of the point y in the pair. A single bit of binary descriptor B on patch p can be calculated as equation (1) [8].

$$B(p;x,y) := \begin{cases} 1 & : I(p,x) < I(p,y) \\ 0 & : otherwise \end{cases}$$
(1)

where $I(p,x)$ was the pixel intensity at the point x of a sampling pair and $I(p,y)$ was the pixel intensity at the point y of a sampling pair. A binary feature descriptor can be formed by concatenating the bits formed by B as Equation (2).

$$\sum_{1\le i\le n} 2^{i-1}\ B(p;x,y) \qquad (2)$$

where the n value for FREAK is 512.

According to previous study, FREAK descriptor has been proposed as the most suitable descriptor for mobile markerless AR application due to the less consuming time of feature extraction [4]. However, an RGB image need to be converted to grayscale before using FREAK descriptor because FREAK descriptor only can extract grayscale features. Color information which provide valuable information had been largely ignored by FREAK. The importance of RGB color information will describe in the Section 2.2.

2.3 RGB Color Space

Color information had been largely ignored by FREAK descriptor. However color information like RGB color are able to provide valuable information

Journal of Theoretical and Applied Information Technology
15th November 2016. Vol.93. No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

or features that are important for a descriptor. The best examples are shown in Figure 1 (a) and Figure 1 (b). Both Figure 1 (a) and Figure 1 (b) shows the same content but using different color space; RGB and grayscale respectively. A person can easily recognize the number "8" that located in the middle of Figure 1 (a), but someone will find difficult to read the number in Figure 1 (b). Hence, researchers can conclude that descriptor which carries color information such as RGB is more valuable and robust compared to a descriptor which computes grayscale information only.

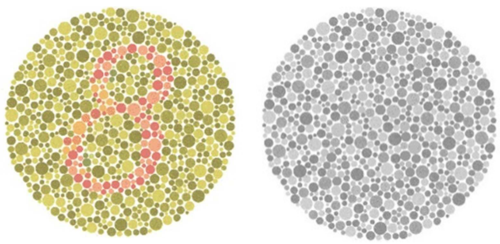

Figure 1 (a): Image carries RGB color information
Figure 1 (b): Image carries grayscale information

Color space is a set of colors which represented by using mathematics. The three most popular color space are RGB, YCbCr and YUV. RGB usually used in computer graphics, YCbCr or YUV usually used in video systems and CMYK usually used in color printing [17]. This research focus on color space used in computer graphics; RGB color space. Figure 2 show the model of RGB color space [18].

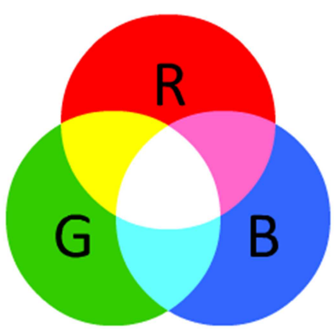

Figure 2: RGB color space

Red, green and blue (RGB) are the three primary additive colors. Red, green and blue light beams are added together in various way to form a desired color. Each color have a range of 0 to 255 as shown in RGB color bars (Table 1) [17].

Table 1: RGB color bars

	R	G	B
Nominal Range	0 - 255	0 - 255	0 - 255
White	255	255	255
Yellow	255	255	0
Cyan	0	255	255
Green	0	255	0
Magenta	255	0	255
Red	255	0	0
Blue	0	0	255
Black	0	0	0

According to the Table 1, when the amount of the three primary colors are at their minimum (0), black color will be produced; while the amount of the three primary colors are at their maximum (255), white color will be produced. The RGB color space is the most suitable choice for computer graphics because RGB color space use red, green and blue to create a desired color. Hence, RGB color space is the popular choice used in design a system or an algorithm. Section 2.3 will discuss about descriptor algorithm using RGB color space; RGB-SIFT.

2.4 RGB-SIFT

RGB-SIFT had been proposed by [6] to achieve better recognition rates for 2D ear recognition. They have extended SIFT descriptor to the RGB color channels to maximize the robustness of the SIFT descriptor. According to the researchers, there are two weakness of the current SIFT descriptor. Firstly, the SIFT descriptor is not invariant to color changes because the intensity channel is only using grayscale channel. Secondly, SIFT descriptor had neglects the color information which is important for recognizing objects.

A taxonomy of the SIFT descriptor's invariant properties under photometric changes has been proposed by [7] to solve the issues above. [7] introduced the diagonal model, diagonal-offset model and photometric analysis. Changes in the illumination can be modeled as Equation (3). Where f^c is the same image transformed under the reference light (canonical illuminant) , f^u is the image taken under an unknown light source and $D^{u,c}$ is a mapping matrix which maps color that are taken under an unknown light source u to the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

corresponding colors under the canonical illuminant c. Equation (4) shows the diagonal model. [19]

$$f^c = D^{u,c}, f^u \quad (3)$$

$$\begin{pmatrix} R^c \\ G^c \\ B^c \end{pmatrix} = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \begin{pmatrix} R^u \\ G^u \\ B^u \end{pmatrix} \quad (4)$$

The diagonal model is extended with an offset $(o_{1,}o_{2,}o_3)^T$ to include the diffuse light and resulting in the diagonal-offset model:

$$\begin{pmatrix} R^c \\ G^c \\ B^c \end{pmatrix} = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \begin{pmatrix} R^u \\ G^u \\ B^u \end{pmatrix} + \begin{pmatrix} o_1 \\ o_2 \\ o_3 \end{pmatrix} \quad (5)$$

SIFT descriptor is invariant to diffuse light under the extended diagonal model because the gradient, operating on the derivative and cancels out the offsets. SIFT descriptor under the extended diagonal model is normalized and the gradient magnitude changes have no effect on the final descriptor and make the SIFT descriptor invariant to light intensity change.

In the research of 2D ear recognition, [6] adapt the work of [7]. SIFT descriptors are extracted independently from R, G and B color channels to produce RGB-SIFT as Equation (6). Where $f(I)$ is the combined set of RGB-SIFT and $f(I_R)$, $f(I_G)$ and $f(I_B)$ are the sets of SIFT feature descriptors extracted from red, green and blue independently.

$$f(I) = f(I_R) \cup f(I_G) \cup f(I_B) = \{ f_1, f_2, \dots . f_n \} \quad (6)$$

R SIFT, G SIFT, B SIFT, RGB-SIFT and SIFT have been evaluated by [6]. RGB-SIFT have achieved the highest accuracy (86%) in the illumination experiment. Followed by G SIFT (75%), SIFT (69%), R SIFT (65%) and B SIFT (58%). Although RGB-SIFT is robust to illumination invariance but RGB-SIFT cannot act in real time because the computation time is too expensive. Hence, RGB-SIFT is not suitable for mobile markerless AR.

This research will proposed a descriptor by extracting R, G and B color channels independently using FREAK descriptor to improve the illumination invariance. FREAK descriptor is a suitable descriptor for mobile markerless AR because the short computation time. The proposed descriptor is an important algorithm for mobile markerless AR because the application required a descriptor that can act in real time and robust to illumination invariance.

3. RESEARCH METHODOLOGY

Based on the good result of the RGB-SIFT [6], this work proposed a new illumination invariant feature based on FREAK by extracting the red, green and blue color channel separately. As FREAK descriptor had been identified as the most suitable descriptor for mobile markerless AR [4]. However, the color information has been largely ignored by FREAK descriptor. The proposed RGB color extensions of FREAK (RGB-FREAK) is to improve the tracking accuracy for mobile markerless AR.

An RGB image is split into three 8-bit channels (red, green and blue). Let n denote the bits of the original FREAK descriptor (n is 512 bits). The original FREAK descriptor is calculated on 8-bit grayscale image during extraction, so an RGB image will need to transform to grayscale before the extraction process. A normalization step that converts the ranges of each channel is employed as discussed in Section 2.4. The original FREAK descriptor are then applied separately to each normalized channel and each feature extracts three n-element named R-FREAK, G-FREAK and B-FREAK. This is finally concatenated into one 3•n-element features, which is RGB-FREAK descriptor as shown in Figure 3.

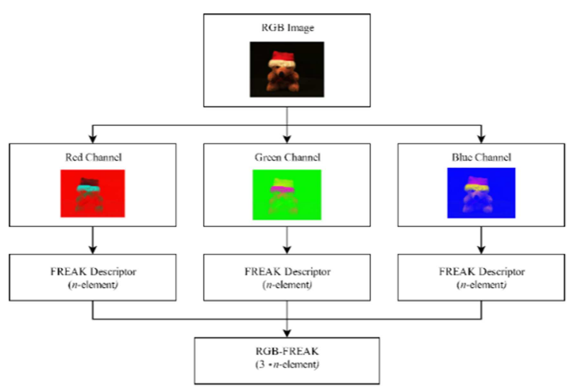

Figure 3: Block diagram of RGB-FREAK descriptor, where n denotes the dimension of the original FREAK descriptor

Under the extended diagonal model, the color FREAK descriptor is invariant to diffuse light since the gradient, operation on the derivative and cancels out the offsets. Color FREAK is also invariant to

Journal of Theoretical and Applied Information Technology
15th November 2016. Vol.93. No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

light intensity change, since the color FREAK is normalized and the gradient magnitude changes have no effect on the final descriptor.

Each descriptor is then applied in mobile markerless AR application for tracking process. Section below will discuss about the experimental results.

4. EXPERIMENTAL RESULTS

This section presents the experimental dataset used in this research and the tracking accuracy of the proposed new illumination invariant feature based on FREAK descriptor in mobile markerless AR. Computation time for the proposed descriptor is also listed down below for comparison purpose.

4.1 ALOI Dataset

ALOI dataset is a color image collection of 1,000 small objects. [20] had systematically varied illumination color, illumination angle, viewing angle and captured wide-based line stereo images to collect sensory variation in object recording. They had recorded over 100 images of each object (1000 objects) and yielding a total of 110,250 images for this ALOI data collection. This research is focused on the illumination invariance of the feature descriptor. Thus, the dataset that will be used is only with varied illumination color and illumination angle. The experimental setup for capturing the images are shown in Figure 4.

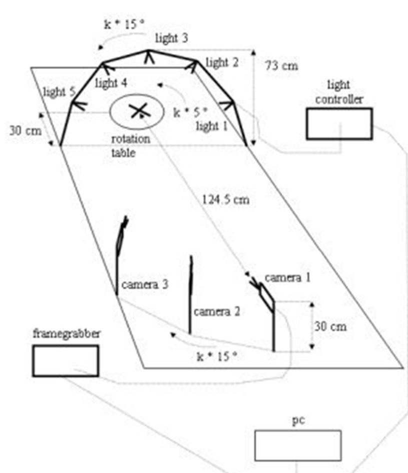

Figure 4: Experimental setup for capturing the images

Small objects with different illumination direction is varied in 24 configurations as Figure 5. Each object was recorded with only one out five lights turned on to produce five different illumination angles (conditions l1-l5). The illumination bow is virtually turned by 15 degrees for camera 2 and 30 degrees for camera 3 by switching the camera and turning the stage towards that camera (cameras c2 and c3). Thus, the aspect of the small objects viewed by each camera is identical, but the light direction has shifted by 15 degrees and 30 degrees in azimuth and results in 15 different illumination angles. Combination of turning on the lights were used to illuminate the small object too. Turning on two lights (light 1 and light 2) at the sides of the object produced an oblique illumination from right (condition l6). While turning on two lights (light 4 and light 5) at another side of the object produced an oblique illumination from left (condition l7). Turning on all lights yields a sort of hemispherical illumination (condition l8). After all the process of switching camera and turning on different light, a total of 24 different illumination condition were generated (c1, c2, c3 and l1, l2, l3, l4, l5, l6, l7, l8).

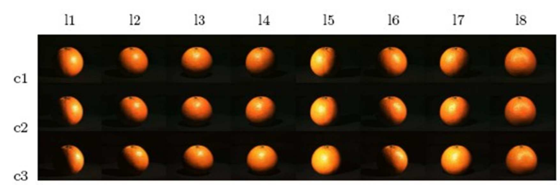

Figure 5: Object orange with different illumination direction is varied in 24 configurations

Small objects with different illumination color is varied in 12 configurations as Figure 6. Each object was recorded with all five lights turned on. Illumination color temperature is changed from 2175K to 3075K. Each small objects was illuminated under a reddish to white illumination color because camera were white balanced at 3075K (condition i110 - i250)

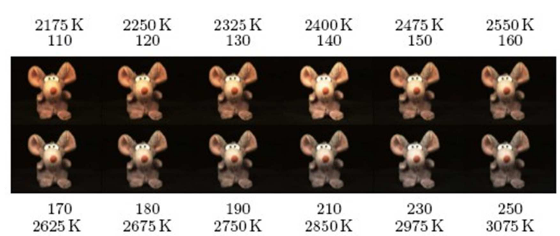

Figure 6: Object mouse with different illumination color is varied in 12 configurations

The robustness of the proposed feature descriptors are evaluated using this ALOI dataset with the two different illumination changes;

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

illumination direction and illumination color. The tracking accuracy of the feature descriptors are measure as Equation (6).

$$Accuracy(\%) = \frac{No. of\ Correct\ Matches\ (n)}{No. of\ Features\ Extracted\ (N)} x100\% \quad (6)$$

The computation times were recorded in milliseconds for every 500 features extracted. Let t_s denote as the starting time, t_e denote the ending time, and t_f denote the final computation time. If $x\ (a,b)$ is the extracted function of the total number of features, then the computation time for $x\ (a,b)$ is defined by Equation (7).

$$x\ (a,b) \rightarrow t_f(t_e - t_s) \quad (7)$$

4.2 Tracking Accuracy

The tracking accuracy for FREAK, RGB-FREAK, R-FREAK, G-FREAK and B-FREAK under light color changes and different lighting arrangement are shown in Figure 7 and Figure 8 respectively. The results are taking from the average of the accuracy values get from 1,000 images from ALOI dataset.

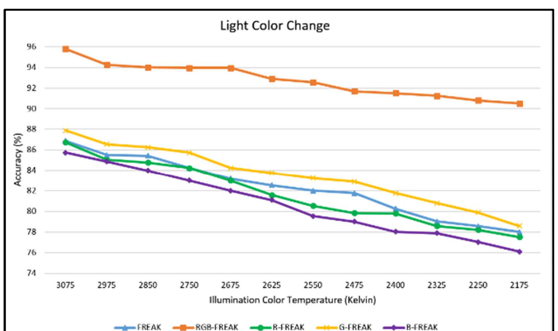

Figure 7: Accuracy for FREAK and color FREAK descriptor under light color changes, averaged over 1,000 images from the ALOI data set.

By observing the results with respect to light color changes, the combined RGB-FREAK descriptor have highest invariance to this property. RGB-FREAK achieve 92.76% of tracking accuracy followed by G-FREAK (83.47%), FREAK (82.29%), R-FREAK (81.66%) and B-FREAK (80.69%). There is a clear distinction in performance between RGB-FREAK and FREAK due to the used of RGB color extension.

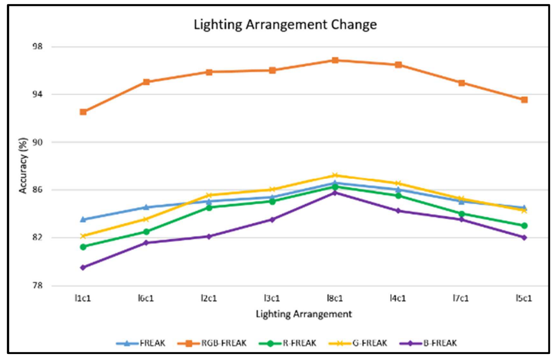

Figure 8: Accuracy for FREAK and color FREAK descriptor under different lighting arrangement, averaged over 1,000 images from the ALOI data set.

From the results shown in Figure 7, it can be seen that the RGB-FREAK still remain invariant and robust to lighting arrangement changes compared to others (FREAK, R-FREAK, G-FREAK and B-FREAK). RGB-FREAK achieved the highest accuracy in the lighting arrangement change with 95.18%, followed by FREAK and G-FREAK which achieve the same accuracy with 85.07%, R-FREAK with 84.01% and B-FREAK with 82.78%.

RGB-FREAK descriptor under the extended diagonal model is normalized and the gradient magnitude changes have no effect on the final descriptor and make the RGB-FREAK descriptor invariant to light changes. From the results in both Figure 6 and Figure 7, the theoretical invariance properties of RGB color extension descriptors are validated.

4.3 Computation Time

Computation time for each descriptor are listed down in Table 2. The computation time are measured for every 500 features extracted.

Table 2: Computation time in millisecond for every 500 features extracted

Descriptor	Computation time (ms)
FREAK	15
R-FREAK	14
G-FREAK	15
B-FREAK	13
RGB-FREAK	41

From the results shown in Table 2, computation time RGB-FREAK almost three times higher compare to other descriptors which is 41ms. This is because RGB-FREAK resulted in color descriptor

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

vector three times larger than the original descriptor, as explained in Section 3. All the descriptors (FREAK, R-FREAK, G-FREAK and B-FREAK) are 512-bits but RGB-FREAK is 1536-bits. Hence, theoretical computation time for RGB-FREAK is three times higher than the original FREAK. However, RGB-FREAK still consider as a good feature descriptor because the computation time is remain real time.

Overall, the performance of the RGB color extension FREAK still performs marginally better than FREAK algorithm. Since the tracking accuracy towards light color changes and light arrangement changes are more robust compared to FREAK descriptor.

5. CONCLUSION

In this research, a new illumination invariant feature based on FREAK descriptor in RGB color space for mobile AR application had been proposed. The results proved that the proposed RGB-FREAK are robust to illumination invariance compared to the existing algorithm, FREAK. FREAK descriptor is a fast descriptor and is robust to scale invariance, rotation invariance and view point change. Hence, the main advantages of RGB-FREAK is that it offer fast computation time and robust to scale invariance, rotation invariance, view point change and also illumination invariance. RGB-FREAK can be used in several image recognition application especially mobile markerless AR application because it fulfilled all the requirement of a mobile markerless AR application.

6. ACKNOWLEDGEMENT

The authors appreciate the anonymous reviewers for the critical review of this paper and those participated in this research as part of the project sponsored by Top Down Research Grant, 2016.

REFRENCES:

[1] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, B. MacIntyre, "Recent advances in augmented reality", *Computer Graphics and Applications, IEEE*, 21 (2001) 34-47.

[2] S.Y. Tan, H. Arshad, A. Abdullah, "Development of a PC-based markerless augmented reality", International Conference on Electrical Engineering and Informatics (ICEEI), *IEEE*2015, pp. 49-53.

[3] A. Alahi, R. Ortiz, P. Vandergheynst, "FREAK: Fast retina keypoint", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), *IEEE*2012, pp. 510-517.

[4] T. Siok Yee, H. Arshad, A. Azizi, "Evaluation on Binary Descriptor in Markerless Augmented Reality", The 3rd National Doctoral Seminar on Artificial Intelligence Technology, The Faculty of Information Science & Technology, UKM, 2014.

[5] M. Wafy, A. Madbouly, "Increase Efficiency of SURF using RGB Color Space", *International Journal of Advanced Computer Science and Applications (IJACSA)*, 6 (2015).

[6] J. Zhou, S. Cadavid, M. Abdel-Mottaleb, "Exploiting color sift features for 2d ear recognition", 2011 18th IEEE International Conference on Image Processing, *IEEE*2011, pp. 553-556.

[7] K.E. Van De Sande, T. Gevers, C.G. Snoek, "Evaluating color descriptors for object and scene recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32 (2010) 1582-1596.

[8] M. Calonder, V. Lepetit, C. Strecha, P. Fua, "BRIEF: Binary robust independent elementary features", *Computer Vision (ECCV)*, (2010) 778-792.

[9] E. Rublee, V. Rabaud, K. Konolige, G. Bradski, "ORB: an efficient alternative to SIFT or SURF", IEEE International Conference on Computer Vision (ICCV), *IEEE*2011, pp. 2564-2571.

[10] S. Leutenegger, M. Chli, R.Y. Siegwart, "BRISK: Binary robust invariant scalable keypoints", IEEE International Conference on Computer Vision (ICCV), *IEEE*2011, pp. 2548-2555.

[11] W.K. Obeidy, H. Arshad, S.Y. Tan, H. Rahman, "Developmental analysis of a markerless hybrid tracking technique for mobile augmented reality systems", Advances in Visual Informatics, *Springer*2015, pp. 99-110.

[12] D. Wagner, D. Schmalstieg, "Making augmented reality practical on mobile phones, part 1", *Computer Graphics and Applications, IEEE*, 29 (2009) 12-15.

[13] C. Van Wyk, "Markerless augmented reality on ubiquitous mobile devices with integrated sensors", *Stellenbosch: University of Stellenbosch*2011.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[14] M. Clothier, M. Bailey, "Overcoming Augmented Reality Tracking Difficulties In Changing Lighting Conditions", (2004).
[15] P. Fua, V. Lepetit, "Vision based 3D tracking and pose estimation for mixed reality", Emerging Technologies of Augmented Reality: Interfaces and Design, *IGI Global*2007, pp. 1-22.
[16] D. Wagner, T. Langlotz, D. Schmalstieg, "Robust and unobtrusive marker tracking on mobile phones", The 7th IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR), *IEEE*2008, pp. 121-124.
[17] K. Jack, "Chapter 3: Color Spaces", From Video Demystified: A Handbook for the Digital Engineer, *Newnes*2004.
[18] P. Nishad, "Various colour spaces and colour space conversion", *Journal of Global Research in Computer Science*, 4 (2013) 44-48.
[19] J. von Kries, "Influence of adaptation on the effects produced by luminous stimuli", *Sources of color vision*, (1970) 109-119.
[20] J.-M. Geusebroek, G.J. Burghouts, A.W. Smeulders, "The Amsterdam library of object images", *International Journal of Computer Vision*, 61 (2005) 103-112.