ARgot: Text-Based Detection Systems In Real Time Using Augmented Reality For Media Translator Aceh-Indonesia With Android-Based Smartphones

Zalfie Ardian\textsuperscript{1a}, P. Insap Santoso\textsuperscript{2b}, and Bimo Sunarfrí Hantono\textsuperscript{3c}

\textsuperscript{1}Department of Information System, Faculty of Computer Science, Ubudiyah Indonesia University, Jalan Alue Naga, Desa Tibang, Banda Aceh, Indonesia.
\textsuperscript{2}Department of Electrical and Information Technology, Faculty of Engineering, Gadjah Mada University.
\textsuperscript{3}Department of Electrical and Information Technology, Faculty of Engineering, Gadjah Mada University.

\textsuperscript{a}zalfie.ardian@uui.ac.id, \textsuperscript{b}insap@mti.ugm.ac.id, \textsuperscript{c}bimo@te.ugm.ac.id

Abstract. Language is a media of communication which aims to understand each ethnic and certain nation. The understanding of language allows the increase of awareness and closeness value among certain individuals. Nevertheless, considerable number of Languages appears to be the causes of difficulty and misunderstanding on the meaning of a certain word. This research suggests one of the solutions in solving such problem. Hence it eases the process of understanding the meaning of the words of a language. Especially “bahasa Aceh”. This application called ARgot. ARgot is an application of translation of Aceh – Indonesia’s Language. This application is able to work as real time application without having a need of an input media such as a text which has to type through using a good quality of camera and having characteristic of autofocus in camera. This research emphasizes on the process of detection and tracer text which available within the society. This research is also intended to see how the technology Augmented Reality is able to be applied in markerless text and have different font of variety. The result of the research is a software for a android smartphone, as media translator in a real time, Hence it is able to be used almost in every kind of smartphone. Based on the trial and result of respondent, this application is able to attract the interest of user and it is also able to be applied in a real life.

1. Introduction
Language is an important element in the life of mankind as a communication tool to interact each other. That is why the language becomes one of the crucial factors in the life of society in the world. Indonesia is a diverse country that has a multi-ethnic group, race, religion, and language diverse region, then the language is a tool to unite the nation. Indonesia, which has population of hundreds of lives, has recorded more than 750 local languages, making Indonesia the country with the largest regional languages in the world. However, there are approximately 160 languages which are endangered due to the lack of depth of the local language.
There are two definitions of language. First definition declared that language is as a means of communication among society in the form of symbols of the sounds produced by the human vocal organs. Secondly, language is a communication system that uses vocal symbols (speech sounds) that are arbitrary [1]. Another definition, language is a form and not a state or an arbitrary sound symbol system, or also a system of many systems, a system of an order or an order of the systems [2]. Written text is one of the most common methods for conveying information in our daily life. However, when the text is found in a foreign language for some individuals, such information cannot be delivered [3]. So, to mitigate this problem, it requires an assisting device that acts as translators, ranging from a simple dictionary to electronic devices that simplify the translation process.

The application of AR technology in the text of the application of Augmented Reality technology that is rarely used in text-based markers. AR research largely focused on the use of marker, to display objects, and interactions using additional markers. There is still little use of AR with Markerless methods, especially the use of AR-based text, meaning that the image is detected to be converted into text form. Moreover, the use of AR by using the smartphone will add value to the higher mobility.

2. Methodology

Augmented Reality (AR) is a concept that combines real-world objects and objects generated virtual world of a computer system by adding information on the real object so that the boundary between the two became very thin. This system is closer to the real environment, because the reality is preferred in this system. With the help of AR technology, a real environment around us will be able to interact with the digital form (virtual). Information about the object and the environment around us can be added to the AR system then displays the information on a layer of real world in real-time so it seems that information be interactive and real. AR concept was first introduced by Thomas Caudell in 1990 when he worked at Boeing [13].

There are differences between the marker and Markerless methods, methods of using the marker is to use a box-shaped black and white image which allows applications to track and detect orientation and to adjust positions based on 3D objects. Meanwhile, according to Madden [16], Markerless is AR which is used to track objects that exist in the real world without a special marker. To perform object tracking, the system relies on the natural Markerless AR-feature and not a fiduciary marker. Moreover, no additional markers that must be made in advance which will be added to the environment. Another advantage is the possibility to issue Markerless information and characteristics of the environment for future use. However, the ability of tracking and registration techniques become more complex objects [17].

2.1 Materials

Vuforia uses Computer Vision technology to identify and track the planar images (Target Image) and simple 3D objects such as boxes, in real-time. Image registration capability allows developers to adjust the position and orientation of a virtual object, such as 3D models and other media, in connection with the real world image when it is viewed through a mobile device camera. The virtual object and then track the position and orientation of the image in real-time so that a user's perspective on objects in accordance with the perspective of the target image, then the virtual object can be displayed in the real world. Vuforia SDK supports various types of targets including 2D and 3D images Markerless, multi 3D targets, and forms a frame marker. Additional features of the SDK includes a local collision detection using virtual button, you can choose a target image in real-time, and the ability for computing at runtime. Vuforia provides Application Programming Interfaces (API) in C++, Java, and Objective-C.

2.2 Methods

There have been a lot of mobile applications for the translators which are available, but most are still using text input method in terms of its use. Augmented Reality technology offers something new, how
objects are detected in real time, this is what's behind what if as text object and do real-time detection and generate output translator translation resulting media-based Augmented Reality. Vuforia makes the detection and tracking of text elements by doing scanning of camera view called Region of Interest (ROI). This technology uses two different ROI, one for detection and one for tracking.

In terms of performance, the detection of the word too much in one tracking will lead to a longer detection time. When the text detection is enabled, Vuforia tries to match words that are found in the ROI with the words in a set of reference database called wordlist. A word database that contains words that are used to process the translator. As shown in Figure 1 below

![Figure 1. Text detection process.](image)

The test aims to determine the minimum requirement in recognizing the text as the target image. The test will use two types of smartphones with different capabilities and specifications. The test is performed with the determination of a distance of 10 cm as shown in Figure 2.

![Figure 2. Distance for detection.](image)
3. Results

Tests were carried out to look at the system's ability to detect the text, but this has not translated initially testing, this test is to look at the success rate of the system when it detects text, the testing process as shown in Figure 3.

Based on the tests performed, it can be concluded that the level of difficulty of the system if it detects the text with the four different conditions are like those described above. Then the result can be seen in Table 1.

Table 1. Testing outcomes.

| No | Test                           | Testing Outcome | Figure 1 | Figure 2 | Figure 3 |
|----|--------------------------------|-----------------|----------|----------|----------|
| 1  | Tests with different fonts     |                 | 3/5 (60%)| 5/5 (100%)| 5/5 (100%)|
| 2  | Test with a different background and color |             | 5/5 (100%)| 5/5 (100%)| 5/5 (100%)|
| 3  | Tests with different thickness levels |              | 5/5 (100%)| 5/5 (100%)| 3/5 (60%) |
| 4  | Tests with different spacing between letters |              | 5/5 (100%)| 5/5 (100%)| 0/5 (0%)  |

Based on the analysis performed on the test, the results can be concluded that the application ARgot does not have the problem of the text that is around as long as the text still has a pattern standard to a letter and this application also does not have a problem to the colors on the condition that the text color too much contrast with background so that the system is not difficult in terms of text detection. The test results summarized in the Table 2.

Table 2. Test result.

| Test   | Detection | Andromax U | Andromax U 2 |
|--------|-----------|------------|--------------|
| P 1    | Succeeded | Succeeded  |              |
| P 2    | Succeeded | Succeeded  |              |
| P 3    | Succeeded | Succeeded  |              |
| P 4    | Succeeded | Succeeded  |              |
| p1 + p3| Succeeded | Succeeded  |              |
| p1 + p2 + p3 | Succeeded | Succeeded |              |
| p1 + p3 + p4 | Succeeded | Succeeded |              |
Information:
P1 = Tests with different fonts
P2 = Test with a different background color and
P3 = Tests with different thickness levels
P4 = Tests with different spacing between letters

4. Discussion

Assets are the main components that will be used in displaying the AR. In the ARgot manufacture, it consists of several parts, they are asset Vuforia Word List (VWL) which serves as a database of text, and the text as an image asset targets. Vuforial Word List is a database of the vuforia package that contains the texts to be detected. In this application, use the file named "bahasa_aceh_utf8" containing words in Acehnese language, which is the word that will appear later to detect the text that will be translated. As a side note, if the word to be detected is not contained in this database, the word is not able to recognize that the process will fail. System applications are built on the Android-based smartphone consists of several major parts, such as import of assets, the implementation of a splash screen functions, initialization, rendering, and deinisialisation. In displaying text, text that serves as a target image to be imported first into the source code. This is because the program still capture text as an image before it is converted to text, the conversion process uses OCR (Optical Character Recognition), so that the letters in the image one by one match each database that is then displayed on the screen.

ARgot takes input text data based on words found on Vuforia wordlist (VWL), a database named Bahasa_aceh_utf8. After drawing terdetesi as text, then the system looks for the text in a database that will be displayed. Before displaying the text, first have to initialize the camera and render settings. First choose which camera to use, because there are smartphone has two cameras, then initialize the camera, turn on the camera and start tracking function. After initializing the camera, rendering initialization can be done. Initialization rendering aims to generate textures using OpenGL, make arrangements to be displayed using OpenGL and OpenGL settings for text display. After displaying text and completed applications are used, so that it is not stored in the buffer should be done deinisialisasi smartphone cameras to triangulate. This test uses the Android-based smartphone with a
wide range of specifications. This test uses four different types of smartphone specifications. Smartphones used has a significant hardware differences, in terms of performance processor and RAM capabilities so that it can be variations to test the system ARgot. Thus, it can be seen the influence of processor, RAM, GPU, the performance of the system ARgot.

5. Conclusion

From the results of the design, implementation, testing and the elaboration can be concluded as follows:
1. Applications ARgot is a translator for the English media applications Aceh are designed using the Java programming language that can be operated on Android devices, to display using Vuforia Augmented Reality SDK and rendering using OpenGL and OCR technology as an auxiliary reading of the image as text.
2. By applying the Augmented Reality uses a technique Markerless on ARgot, does not restrict users to detect certain text or images. So that users are free to detect any text as recorded on the database. With a variety of tests performed on the performance of ARgot ARgot has proven to work well and can be known limits of ARgot.
3. Augmented Reality can be innovation bari text in android-based translator media thus enhancing the quality of software for education, especially for studying Indonesian culture.
4. From a survey of 30 responders were obtained by distributing questionnaires obtained more than 80% expressed an interest and 85% of respondents stated ARgot really helpful if applied as a media translator. However the detection process that sometimes imperfect can create confusion for users.

6. References

[1] S. Keraf, Kadhipta, 1st ed. Jakarta: Balai Pustaka, 1991
[2] W. Mackey, Analisis Bahasa. Surabaya: Usaha Nasional, 1986
[3] V. Fragoso, S. Gauglitz, S. Zamora, J. Kleban, and M. Turk TranslatAR: A mobile augmented reality translator 2011 497–502
[4] M. Petter, V. Fragoso, M. Turk, and C. Baur, “Automatic text detection for mobile augmented reality translation,” 2011, pp. 48–55
[5] B. Parhizkar, K. Oteng, O. Ndaba, A. Lashkari, and Z. Gebril, “Ubiquitous Mobile Real Time Visual Translator Using Augmented Reality for Bahasa Language,” IACSIT Press, vol. 3, Apr. 2013
[6] M. Khan, M. Zaman, and T. Islam, “Application Of Augmented Reality : Mobile Camera Based Bangla Text Detection And Translation,” Brac University, 2012
[7] V. F. Bauset, J. M. Orduna, and P. Morillo, “Performance Characterization on Mobile Phones for Collaborative Augmented Reality (CAR) Applications,” in 2011 IEEE/ACM 15th International Symposium on Distributed Simulation and Real Time Applications (DS-RT), 2011, pp. 52–53
[8] C.-L. Lai and C.-L. Wang, “Mobile Edutainment with Interactive Augmented Reality Using Adaptive Marker Tracking,” in 2012 IEEE 18th International Conference on Parallel and Distributed Systems (ICPADS), 2012, pp. 124–131
[9] Wildan, Tata Bahasa Aceh untuk Madrasah Dasar dan Madrasah Lanjutan, vol. III. Banda Aceh: Global Educational Consultant Institute, 2002
[10] SIL, Bahasa-Bahasa di Indonesia (Language of Indonesia). Jakarta: SIL Internasional Cabang Indonesia, 2006
[11] M. Toha, Laporan Kekerabatan dan Pemetaan Bahasa-Bahasa di Propinsi Nanggroe Aceh Darussalam. Banda Aceh: Balai Bahasa, 2008
[12] B. Parhizkar, Z. M. Gebril, W. K. Obeidy, M. N. A. Ngan, S. A. Chowdhury, and A. H. Lashkari, “Android mobile augmented reality application based on different learning theories for primary school children,” presented at the 2012 International Conference on Multimedia Computing and Systems (ICMCS), 2012, pp. 404–408

[13] F. Miller, A. Vandome, and J. McBrewster, “Augmented Reality,” VDM Publ. House, 2009.

[14] J. Kim and H. Jun, “Implementation of image processing and augmented reality programs for smart mobile device,” in 2011 6th International Forum on Strategic Technology (IFOST), 2011, vol. 2, pp. 1070–1073

[15] O. Bimber and Raskar, “Spatial Augmented Reality,” K Peters, 2005

[16] L. Madden, Augmented Reality Browsers for Smartphones: Programming for JUNAIO, LAYAR, and WIKITUDE, First Edition., vol. 1. Wiley Publishing Inc., 2012

[17] “Markerless Augmented Reality.” [Online]. Available: http://www.arlab.com/blog/markerless-augmented-reality/. [Accessed: 23-Apr-2013]

[18] R. Azuma, “A Survey of Augmented Reality,” Eleoperators Virtual Environ., vol. 6, pp. 355–385, Aug. 1997

[19] J. Palenchar, “Survey: Samsung, Android Still On Top In U.S. Cellphone Market,” TWICE This Week Consum. Electron., vol. 27, no. 25, pp. 56–56, Dec. 2012

[20] “App Components | Android Developers.” [Online]. Available: http://developer.android.com/guide/components/index.html. [Accessed: 10-Dec-2013].

[21] “Auggie Awards,” Augmented World Expo. [Online]. Available: http://augmentedworldexpo.com/auggies/. [Accessed: 29-Nov-2013]

[22] “Vuforia SDK Architecture | Vuforia Developer Portal.” [Online]. Available: https://developer.vuforia.com/resources/dev-guide/vuforia-ar-architecture. [Accessed: 25-Apr-2013]

[23] Hanafi, “Sistem Aplikasi Mobile Advertising Berbasis Grafis dan Animasi (studi: PT. XL AXIATA TBK),” Thesis, Universitas Gadjah Mada, Yogyakarta, 2012

[24] Fred D Davis, “Technology Acceptance Model for Empirically Testing New End-User Information Systems: Theory and Results,” Massachusetts Inst. Technol. Boston, 1986