University of Dayton
eCommons

Computer Science Faculty Publications | Department of Computer Science

10-2016

MARIM: Mobile Augmented Reality for Interactive Manuals

Tam Nguyen
University of Dayton, tnguyen1@udayton.edu

Dorothy Tan
Singapore Polytechnic

Bilal Mirza
Singapore Polytechnic

Jose Sepulveda
National University of Singapore

Follow this and additional works at: http://ecommons.udayton.edu/cps_fac_pub


Part of the Graphics and Human Computer Interfaces Commons, Other Computer Sciences Commons, and the Theory and Algorithms Commons

eCommons Citation

Nguyen, Tam; Tan, Dorothy; Mirza, Bilal; and Sepulveda, Jose, "MARIM: Mobile Augmented Reality for Interactive Manuals" (2016). *Computer Science Faculty Publications.* Paper 72.
http://ecommons.udayton.edu/cps_fac_pub/72

This Conference Paper is brought to you for free and open access by the Department of Computer Science at eCommons. It has been accepted for inclusion in Computer Science Faculty Publications by an authorized administrator of eCommons. For more information, please contact frice1@udayton.edu, mschlangen1@udayton.edu.

MARIM: Mobile Augmented Reality for Interactive Manuals

Tam V. Nguyen
Department of Computer Science
University of Dayton
vantam@gmail.com

Dorothy Tan, Bilal Mirza, Jose Sepulveda
ARTIC Research Center, Singapore Polytechnic
bilal2@e.ntu.edu.sg, {dorothy_tan,
sepulveda_jose}@sp.edu.sg

ABSTRACT

In this work, we present a practical system which uses mobile devices for interactive manuals. In particular, there are two modes provided in the system, namely, expert/trainer and trainee modes. Given the expert/trainer editor, experts design the step-by-step interactive manuals. For each step, the experts capture the images by using phones/tablets and provide visual instructions such as interest regions, text, and action animations. In the trainee mode, the system utilizes the existing object detection and tracking algorithms to identify the step scene and retrieve the respective instruction to be displayed on the mobile device. The trainee then follows the displayed instruction. Once each step is performed, the trainee commands the devices to proceed to the next step.

Keywords

Augmented Reality, Authoring Tool, Interactive Manuals

1. INTRODUCTION

Today, thanks to the massive coverage of the Internet, people are able to access to useful do-it-yourself materials to help them in their work and daily lives. Situations where individuals need to get quick instructions can range from fixing consumer devices to cooking or baking. Some of these do-it-yourself materials can be found from popular video websites, *i.e.*, YouTube [1]. Others can be found from the websites such as Instructables [2] and wikiHow [3]. However, people are usually impatient to follow the full instruction videos and the instructions are not well organized which causes the users to get lost at some point. Also, there is a lack of interactive method to deliver the instructions to trainees in the conventional methods.

In recent years, the rapid proliferation of smartphones and tablets is profoundly changing the way that we behave, consume content and conduct commerce. Almost all of inexpensive smartphones are equipped with cameras, wireless network, gyroscope, and accelerometer. The improvements in imaging capabilities and computational power have given a rise to many exciting mobile applications. Augmented Reality (AR) is a technology applied for training purpose. Although there are many AR systems for learning, researchers do not usually use the most important feature of AR - showing an explicit relationship between the virtual learning content and real objects found in the natural environment.

Witnessing the demand of users and the growth of mobile devices, we would like to create a mobile AR-based system that serves as a platform where experts can create and upload manual instructions that can be used by the user to train themselves for the tasks. This new type of training will be based on mobile platforms tapping on all the capabilities. As a solution we propose a system named MARIM, known as Mobile Augmented Reality for Interactive Manuals, which guides the trainees to do things and at the same time ensures that they are doing it right. A short video about our system is available at https://www.youtube.com/watch?v=3LkwO41eXrM.

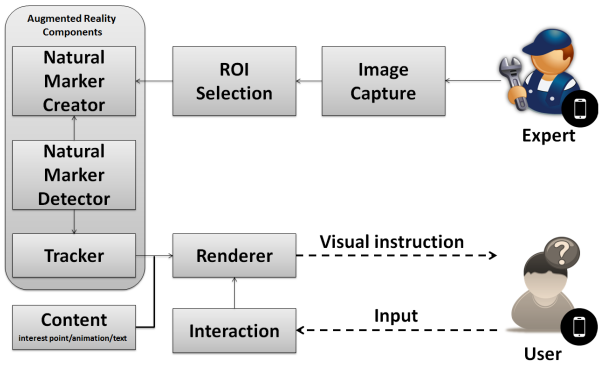


Figure 1: The flowchart of MARIM system.

2. TECHNICAL DETAILS

Figure 1 depicts the major components of our system. Our proposed system supports two modes, namely, trainer editor mode where the experts can design the manuals, and the trainee mode where the users can use the created manuals to solve their problems.

2.1 Expert/Trainer Editor

The actual flowchart of the editor mode is illustrated in Fig. 2. First, the image is retrieved from mobile camera. The expert chooses the interactive points on the image and

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

MM '16 October 15-19, 2016, Amsterdam, Netherlands

© 2016 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-3603-1/16/10.

DOI: http://dx.doi.org/10.1145/2964284.2973807

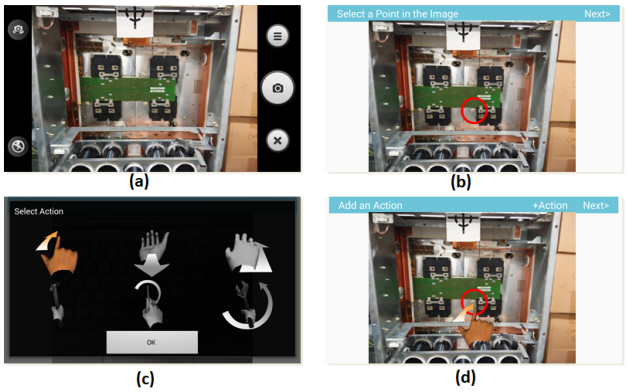


Figure 2: The screenshots of our expert/trainer editor mode. (a) input frame, (b) selecting interest point, (c) providing instruction, (d) the superimposition of instruction and interest point.

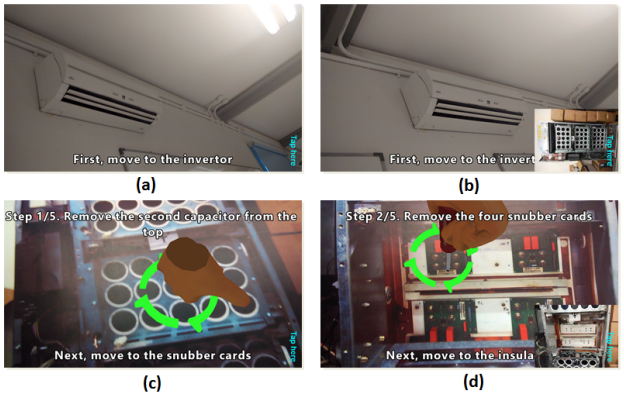


Figure 3: The screenshots of our user/trainee usage mode. (a) the instruction in the text form, (b) the reference photo to guide the trainee, (c) the superimposition of the action animation and text in the video frame, (d) the following step is displayed.

selects the corresponding instructions, *i.e.*, text and action animation. When the system receives these information, it finds the region of interests of the input image by incorporating its saliency map from the real-time AH method [6] and the interactive points (with applied Gaussian kernel) input by the experts. The regions with high salient values will be kept while a Guassian blur [7] is applied to the other regions. This is to ensure that the image matching and tracking functions focus more on the main object and the interactive points. Next, the feature extraction is applied to extract the keypoint descriptors, *i.e.*, SIFT [5], of the feature points. Note that the chosen descriptors are invariant to noise, scale, and rotation. Lastly, all the fine-tuned images, instruction information and keypoint descriptors are stored on the manual database.

2.2 User/Trainee Guidance

The trainees log into the system and select the appropriate manuals for their repairing tasks. As shown in Fig. 3, the trainee mode shows the instruction in the text form. In addition, the trainee mode shows the reference photo to guide the trainee point to the correct scene. Meanwhile, the video frame from the live streaming of mobile camera undergoes the feature extraction, namely, keypoint descriptors. Next, the keypoints of the current video frame are matched with the ones of the stored images in the scenario. Once the frame is matched, the system performs image tracking and compute the camera pose to the scene. Then, then instruction information associated with the current step is delivered back to the user's device, and the information is superimposed on top of the video image. Once each step is done, the trainee commands the devices to proceed to the next step via touch screen interface.

It is worth noting that the interactive manuals implicitly support reparation step verification. In other words, it means that the users will be reminded if they forget to perform any step. Our system is also supported on head mounted display [4], a mobile augmented reality device that provides an immersive augmented reality experience, and frees user's hands to do the task. The actual usage of the trainee mode is illustrated in Fig. 3.

3. DEMONSTRATION PLAN

For the demonstration, we will show our system with two modes, expert/editor and trainee modes. Demo presenters provide all necessary items for the demo, such as laptops, smart phones and tablets. Given the demo application, the participants are encouraged to generate interesting manuals in the editor mode and actually exploit the created scenarios to amend things. During the demonstration, we will also provide users guidance through steps by steps as shown in Fig. 2 and Fig. 3.

4. CONCLUSIONS

We have developed a user-friendly application creating interactive manuals from mobile devices. Given a reparation scenario, the MARIM system allows the expert to create the interactive manuals. In the meantime, the trainee can download the manuals to repair his/her own stuffs. Limited by the current performance of image matching and object tracking, some sudden camera movements may make the object mistracked. This issue can be further alleviated along with the development of state-of-the-art tracking methods. We believe this work can inspire more research in the future.

5. ACKNOWLEDGEMENTS

This work was supported by Singapore Ministry of Education under research Grant MOE2014-TIF-1-G-007.

6. REFERENCES

[1] YouTube. http://www.youtube.com/.
[2] Instructables. http://www.instructables.com/.
[3] WikiHow. http://www.wikihow.com/.
[4] Epson Moverio glasses. https://www.epson.com/cgi-bin/Store/jsp/Landing/moverio-bt-2000-smart-headset.do.
[5] D. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.
[6] T. V. Nguyen and J. Sepulveda. Salient object detection via augmented hypotheses. In *International Joint Conference on Artificial Intelligence*, pages 2176–2182, 2015.
[7] G. Stockman and L. G. Shapiro. *Computer Vision*. Prentice Hall PTR, Upper Saddle River, NJ, USA, 1st edition, 2001.