
AR Tamagotchi : Animate Everything Around Us

Byung-Hwa Park

i-Lab, Pohang University of Science and Technology
(POSTECH), Pohang, South Korea
pbh0616@postech.ac.kr

Se-Young Oh

Dept. of Electrical Engineering, Pohang University of Science and
Technology
(POSTECH), Pohang, South Korea
syoh@postech.ac.kr

Abstract

As image processing and computer vision(CV) technology become available in real time, Augmented reality(AR) technology enables us to more interact with Real world through the augmented data. As this paradigm, we can interact with everyday object through augmented data such as facial expression. This work describes use of CV and AR to interact with a everyday object as making them animated. For smart device application purpose to interact with everyday object, we developed touch event based object tracking algorithm using the camshift , grabcut and particle filter hybrid system. The Poisson image edit was used for facial expression blending to make look like the object had a facial expression inherent. To make the system fun to use, we adopted the Tamagotchi storytelling concept, which is a virtual pet simulator. For multi-platform game application purpose and computer vision programming, the AR programming environment was built using the cocos2d-x game engine and OpenCV. In experiments, users felt that AR made them feel that the everyday object was animated, had emotions and could interact with them.

Author Keywords

Augmented Reality, Computer Vision, Machine Learning, Game, Tangible Interaction, Facial Expression

ACM Classification Keywords

H.5.1 [Multimedia Information Systems]: Artificial, augmented, and virtual realities. I.2.10 [Vision and Scene Understanding]: Video analysis. K.8 [PERSONAL COMPUTING]: Games

Introduction

Anyone has his favorite object such as a gift from his lover, his friend, or his family. We have been living with those kind of everyday objects. However, because most of them are the static tangible objects, we just have been watched and held them to interact.

To interact with the everyday object better, many approaches have been proposed. Most of them were enabling them to make an action like making sound, move, or flash using electrical device in hardware perspective. But these hardware approaches were limited to the only the object.

As technology progressed rapidly, various AR applications were emerged. AR can support user to understand the scene better and propose other ways to interact with system to make the user use easier [1]. Therefore AR technology and its applications have been researched and developed to help user to interact with system in user oriented.

Meanwhile, a simple and an intuitive method have been used for a long time. That is using a facial expression. This approach can be found in various visual media. For instance, in Beauty and the Beast, Belle, heroine of the story, became a friend with the furniture of the Beast's castle which has emotion with facial expression [2]. In the annoying orange [3], many fruits, which had the facial expression, was appeared

and try to make fun. This concept have been widely used in many visual media to animate any object, but hard to find it in real world, because it was restricted to virtual media.

But, AR can make it possible using augmented data. In this concept, we can animate and interact with any everyday objects around us. In this work, we suggested the way to instil "life" into everyday objects and become their friend using CV, AR and Tamagotchi concept.

Design Concept

In an attempt to interact with and animate everyday object through Augmented Reality, the smart phone was chosen as a platform because many modern people are living their life with it everyday. To make the system more fun, we adopted the Tamagotchi storytelling concept which is virtual pet simulation game. So, cocos2d-x game engine was used for game development and multiplatform mobile development purpose. It is programmed on C++ environment and generates wrapper automatically to Android, or iOS. For simple computer vision algorithm, we combined OpenCV to Cocos2d-X.

When user touches his favorite everyday object on the smart device camera display, the object will be animated with facial expression and graphical effect, after that, it will be recorded on Server. User will interact with the object as the Tamagotchi scenario. For instance, the AR Tamagotchi will feeds and plays, and sends a message to the device about these activities. Then, the user watches the registered everyday object through a smart device, which identifies it through the recorded data on the server. Then user interacts with

the AR Tamagotchi and camera preview data is transferred to the Server, and the server analyzes the 3D structure of the object. Then, the server transfers a 3D model of the object to the device, which renders the AR Tamagotchi will be rendered on the display with some movement for realism (Figure 1).



Figure 1: Concept scenario of the system. User choose the target everyday object to animate on the smart device camera display, then it is animated and become the AR Tamagotchi.

Algorithm

Touch based Object Tracking

On the smart device environment, the most intuitive gesture for user to select the object on the camera preview is touch, because every user interface is based on the touch in the smart device. To animate the object through AR, the system should track the object and the user input to tracking algorithm should be the point coordinate, which is touch point on the camera preview.

But most of previous research with object tracking is not based on the touch event. The detection based object tracking doesn't need user input because it already trained the object, but it is limited to the only trained object. The selection based object tracking needs the object region which is rectangle generally, because the algorithm should know the information of the target object to track such as the whole shape and texture. So we developed the object tracking with image segmentation algorithm. We segmented the object from the image using GrabCut Algorithm [7], with the touched position as a strict-object-region and the border of the image as a probable-background-region.

After the object was segmented, the Continuously Adaptive Mean Shift(CAMShift)[3] algorithm was used to track the object. CAMShift is based on the Mean shift algorithm which tracks the Region of Interest based on finding the maxima of discrete data(Feature, Color etc.) cluster density. CAMShift changes its search window size adaptively while searching the maxima region of the data distribution. CAMShift also uses the Hue data distribution of the image to reduce the effects of lighting variations. CAMShift gives the pose information of the object, therefore we can blend the facial expression naturally. Moreover, CAMShift can track the selected object with point ROI, which the user indicated by touching the smart device display, as a starting search window, because CAMShift tracks the ROI by changing the search window.

Seamless Image Blending

We may be able to blend the facial expression to the object directly, but to make the object look like having the facial expression inherent visually, we blended the facial expression image with seamless image blending concept. For seamless facial expression blending, Poisson image edit [5] approach was used. Its approach is blending the Laplacian of the images not the images themselves because the human vision system understands the former not the latter according to psychophysical observation [5]. Figure 2 shows the example of the Poisson image edit blending.



Figure 2: Poisson Image Edit blending example, its basic concept is gradient blending, not image itself, so it shows seamless blending result

Poisson Linear Equation. For real time algorithm, we solved the equation using Fourier transform approach [6]. Also, to save the image processing time, we made the program to change the size of target image adaptively. Figure 3 shows the difference between direct (Figure 3(a)) and seamless (Figure 3(b)) image blending for two different everyday objects with same cartoon-like facial expression for comparison. The former shows more distinct blended effect than latter.



Figure 3: (a) Direct image blending and (b) Seamless image blending (Poisson image edit)

3D Modeling

During the tracking process, the object will be segmented from the scene using the Grabcut algorithm [6], and it will be sent to the server. As the data is accumulated, the server estimates the 3D model of the object from the SfM (Structure from Motion) [7]. When it is done, the 3D data will be rendered together on the device over the object with slight motion. It will make the user feel that the object is animated more.

Software Architecture

Cocos2d-x open source multi-platform game engine is used for the base platform of the project. In there, we programmed connection to OpenCV, because it doesn't support any CV programming. Using this combination, we made an AR programming environment. The benefit of this combination is that it enables AR programming on a game programming environment, also, it provides multi-platform programming based on native language, C++.

Tamagotchi Scenario

The scenario will follow the traditional Tamagotchi. There will be three basic interactions: feeding, healing, and playing around. In case of feeding and healing, the Tamagotchi will send a message about them to the device. These information will be recorded on the system and they will affect its vital and interaction score. In the case of playing around, we will make a simple AR game, and this interaction data will affect the interaction score on the system. The higher interaction score will make the AR Tamagotchi look lovelier.

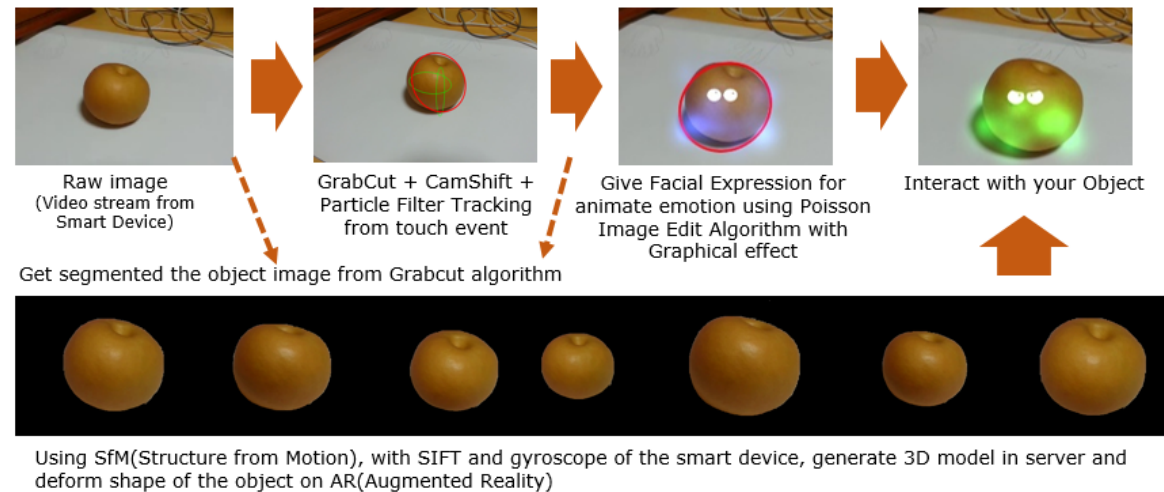


Figure 4: The overall procedure of the AR Tamagotchi system and current state of this work in progress

User Study

User Study Design

To quantitatively and objectively evaluate the system, we ran a user study. The goal of user study was to measure whether we achieved project goal, 'animate everyday object'. The participants were asked to respond to questions related with to project idea, prototype description power, and animation factor (Fig 1). The responses were recorded using a five-point Likert scale: strongly agree (5), agree (4), neither agree nor disagree (3), disagree (2), or strongly disagree (1). The participants was 17 and the experiment used a with-in group design.

User Study Analysis

Responses suggest that the participants they thought

that the project idea enabled them to feel that the everyday object is animated. But in the case of the prototype description power, the score of the same question- , 'This system may enable or enabled me to feel the object is animated', the was significantly lower ($p = 0.025$) when it was asked: after the demonstration ; than when it was asked before the demonstration, This difference means the that current prototype did not achieve the project idea well. Participants responded that the facial expression was the most important factor to animate the object, but that the graphical effect was not needed to achieve the project goal ($p=0.002$). So we can conclude that we can use augmented reality to animate any everyday object, and the user may be able to feel that it is natural. The most important factor to animate the object was facial expression.



Figure 6: Prototype which was used in formative evaluation and user study

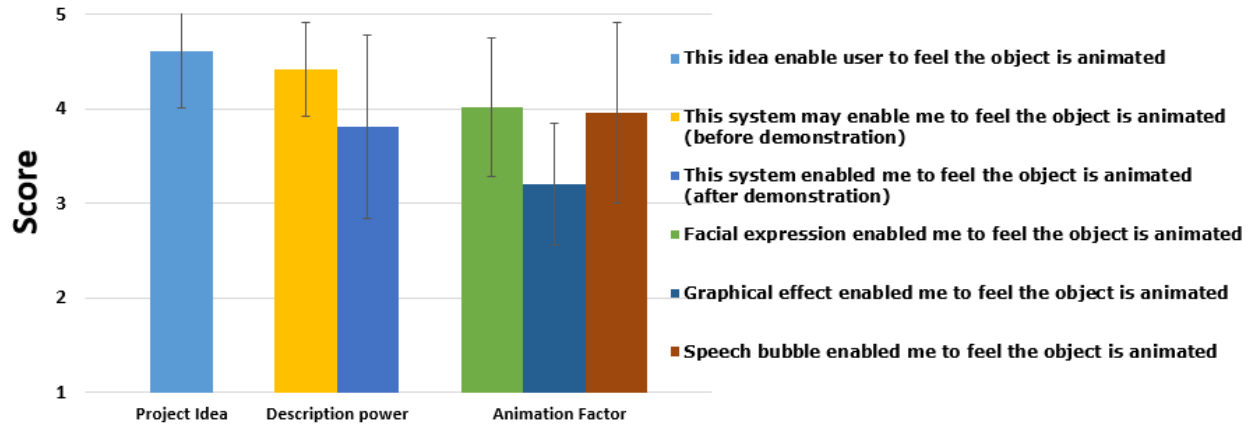


Figure 5: Mean user responses (scale of 1 – 5) to questions evaluating the study. Bars: ± 1 s.d., $n = 17$

Work in Progress Status

The computer vision part, which is touch based tracking, and image blending part was done. And the connection between Cocos2d-x and OpenCV was also done. We made the prototype and did formative evaluation with user study. From the result of user study, we are now developing the system better. Also we concentrate to construct the server which will communicate with the smart device. After that, we will move on to code a gaming contents of the AR Tamagotchi with 3D modeling for subtle movement.

Acknowledgements

"This research was supported by the MSIP(Ministry of Science, ICT and Future Planning), Korea, under the "IT Consilience Creative Program" (NIPA-2013-H0203-13-1001) supervised by the NIPA(National IT Industry Promotion Agency)

References

- [1] DWF. Van Krevelen, R. Poelman, A Survey of Augmented Reality Technologies, Applications and Limitations, The International Journal of Virtual Reality, 2010, 9(2):1-20
- [2] Kirk Wise, Gary Trousdale, Walt Disney Pictures, Beauty and the Beast(Film), 1991
- [3] Dane Boedigheimer, Gagfilms, The Annoying Orange(Comedy film), 2009
- [4] G. Bradski, "Computer Vision Face Tracking for Use in a Perceptual User Interface," Proc. IEEE Workshop Applications of Computer Vision, pp. 214-219, 1998.
- [5] Patrick Perez, Poisson Image Editing, SIGGRAPH, 2003
- [6] J.-M. Morel, A.B. Petro, Pattern Recognition Letters 33 (2012) 342-348, Fourier Implementation of Poisson image editing
- [7] Carsten Rother, Vladimir Kolmogorov, Andrew Blake, SIGGRAPH '04 ACM SIGGRAPH 2004 Papers Pages 309-314, "GrabCut": interactive foreground extraction using iterated graph cut