Available online at www.sciencedirect.com

ScienceDirect

Journal of Computational Design and Engineering 2 (2015) 26–37

www.elsevier.com/locate/jcde

A method for image-based shadow interaction with virtual objects

Hyunwoo Ha[a,b], Kwanghee Ko[a,b,*]

[a]School of Mechatronics, Gwangju Institute of Science and Technology, 123 Cheomdangwagiro, Bukgu, Gwangju 500-712, Republic of Korea
[b]Korea Culture Technology Institute, Gwangju Institute of Science and Technology, 123 Cheomdangwagiro, Bukgu, Gwangju 500-712, Republic of Korea

Received 22 August 2014; received in revised form 17 September 2014; accepted 18 September 2014
Available online 6 December 2014

Abstract

A lot of researchers have been investigating interactive portable projection systems such as a mini-projector. In addition, in exhibition halls and museums, there is a trend toward using interactive projection systems to make viewing more exciting and impressive. They can also be applied in the field of art, for example, in creating shadow plays. The key idea of the interactive portable projection systems is to recognize the user's gesture in real-time. In this paper, a vision-based shadow gesture recognition method is proposed for interactive projection systems. The gesture recognition method is based on the screen image obtained by a single web camera. The method separates only the shadow area by combining the binary image with an input image using a learning algorithm that isolates the background from the input image. The region of interest is recognized with labeling the shadow of separated regions, and then hand shadows are isolated using the defect, convex hull, and moment of each region. To distinguish hand gestures, Hu's invariant moment method is used. An optical flow algorithm is used for tracking the fingertip. Using this method, a few interactive applications are developed, which are presented in this paper.

Keywords: Shadow interaction; Hu moment; Gesture recognition; Interactive UI; Image processing

1. Introduction

There have been the increasing demands for a more active and interesting viewing experience, and interactive projection technology has been considered as a solution to this issue. For example, if you can flip pages with a gesture when you make a presentation, or write a sentence without any manual tools, then the presentations can be more immersive and attractive to the audiences. An interactive projection system also helps people to produce more attractive artistic exhibits, such as interactive walls and floors. Lately, a lot of attempts have been made to use human–computer interaction in plays and musical performances. Namely, if appropriate events occur when an actor performs on stage, a better reaction can be obtained from the audience because such events are well synchronized with the actor's performance. Using this concept, new applications with interesting interactions are possible such as the magic drawing board or virtual combat simulation.

From a technical standpoint, research on gesture recognition is a topic of interest in the field of computer vision. In particular recognizing gestures in real time is of paramount importance. Most research groups use the Kinect camera to recognize gestures precisely because the Kinect camera can discern both depth and color information. On the other hand, the Kinect cannot obtain depth and color information for shadows generated by light from behind the screen. The detection range of the Kinect is limited when applied to a large screen because the distance from the sensor to the screen is considerably large. Another method for gesture recognition is to recognize gestures from images. The image-based approach is less expensive than the Kinect-based method because it uses less hardware for gesture acquisition.

In this work, a vision-based interactive projection system is proposed, which recognizes shadow gestures with proper precision. The process consists of detection and recognition modules of shadow gestures in real time, which are the core parts of the

*Corresponding author at: School of Mechatronics, Gwangju Institute of Science and Technology, 123 Cheomdangwagiro, Bukgu, Gwangju 500-712, Republic of Korea. Tel.: +82 62 715 3225; fax: +82 62 715 2384.
E-mail address: khko@gist.ac.kr (K. Ko).
Peer review under responsibility of Society of CAD/CAM Engineers.

http://dx.doi.org/10.1016/j.jcde.2014.11.003
2288-4300/

proposed system. Next, several novel applications based on the proposed system are presented to demonstrate the potential of the proposed method for use in various applications.

Numerous studies have been conducted regarding interactive projection systems. In particular researchers are interested in generating events using hand gestures because the hand gestures can represent diverse shapes appropriate for recognition.

Mistry et al. [1] proposed a portable interactive projection system, SixthSense, based on natural hand gestures. It provides a wearable gestural interface that allows the user to interact with digital information augmented around the user. The system consists of a portable projector, a camera and a mobile wearable device, which shows digital information on physical objects in real time. Grønbæk et al. [2] introduced an interactive floor support system using a vision-based tracking method. The system consists of a 12 m^2 glass surface with a projector that projects the glass upward. Limbs of users (children) are tracked and recognized for various interactions, which provide learning environments for children. Wilson of *Microsoft Research* [3] reported the prototype of an interactive tabletop projection-vision system, called PlayAnywhere, which allows the user to interact with virtual objects projected on a flat surface. For this interaction, the shadow-based finger recognition, tracking, and various other image processing are incorporated to provide a convenient but flexible tabletop projection-vision system. It consists of off-the-shelf commodities such as a camera, projector, and a screen that do not require any detailed configurations or calibration. Berard [4] developed the "Magic Table" for meetings. It has a whiteboard on the surface. It was developed to overcome the limitation of the current whiteboard by providing various operations such as copy, paste, translation, and rotation of the drawn contents. It consists of a projector, two cameras and a white board. The pen stroke and the contents on the board are captured by the cameras. The captured images are then processed to extract the position and the contents using various image processing techniques. Practically, this system allows the user to interactively create and control the contents. In addition to the above-mentioned systems, various other projection systems have been developed worldwide [5–7].

This paper is structured as follows: Section 2 presents the overall process of the proposed algorithm. Section 3 explains the process of detection and separation of image data. In Section 4, the recognition process for distinguishing hand gestures is presented. In Section 5, the tracking process of shadows is presented. Section 6 shows the experimental results of the proposed algorithm. Finally, the conclusion of the paper is presented with future work in Section 7.

2. Overall process

Fig. 1 shows the entire system consisting of a beam projector, a screen, a web camera and a computer. If a user creates a gesture, the shadow is created on the screen, which is captured by the camera. The computer then performs calculations in order to recognize the gesture through image processing. Next, the computer controls the beam projector to create an event at the proper place in real-time.

The overall workflow of the proposed system is illustrated in Fig. 2. First, the computer receives an input image from the web camera. The image is processed to produce a binary image. Then, an AND operation is performed on the background and the binary image in order to remove the background. Shadows that are distinct from the background are detected using a labeling algorithm. The area of the hand can be obtained through curvature, a convex hull, and defect in each labeled area. The center of the hand can be recognized using the moment value. Invariant moments are used for gesture recognition. After the gesture is recognized, the shadow hand is traced by an optical flow algorithm. Finally, events corresponding to the gesture are generated and given to the user. In the subsequent sections, each module in the overall process is explained in detail.

3. Separation and detection process

In this section, the technical approaches for separation and detection are explained. Given an image, the shadow part is extracted using the background separation and shadow detection methods.

3.1. Background separation process

The background separation step segments the image into the background and objects. For this operation, an improved averaging background algorithm is employed.

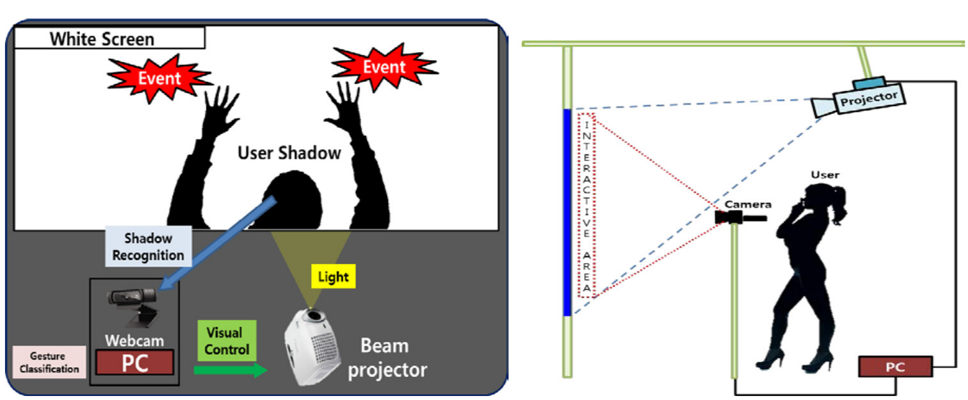

Fig. 1. Overview of the system.