Information Systems 48 (2015) 296–300

Contents lists available at ScienceDirect

Information Systems

journal homepage: www.elsevier.com/locate/infosys

An interactive image clipping system using hand motion recognition

Kowoon Lee, Kyungha Min *

Department of Media Software, Sangmyung University, Seoul, Republic of Korea

ARTICLE INFO

Available online 14 July 2014

Keywords:
Image clipping
Hand recognition
Region recognition
Kinect

ABSTRACT

We present an efficient hand recognition algorithm for an interactive image clipping system, which is widely used for environments such as public facilities and security environments where personal capturing devices including mobile phones are not allowed. User-friendly interface and accurate image capturing function are required for an image clipping system. We build the system by combining Microsoft Kinect, HD webcam and projector. The Kinect and webcam are used to capture the motions of users' hand and project is to display the user-selected area from the capturing material. Hand recognition is composed of three steps: (i) the region occupied by users' hand is extracted from an image, (ii) the fingertips of the extracted hand region are analyzed using *k*-curvature algorithm, and (iii) the height of the fingertip is estimated using the depth image from Kinect. The height of the fingertip informs whether users' finger touched the surface of the target. The region captured by the fingertip is clipped from the image and stored as the target image. The excellence of our hand recognition algorithm is proved through a user test.

1. Introduction

Collecting information from various media is an important process in many fields. Conventionally, most of the collection operations such as copying notable pages are executed physically. Recently, the progress of digital devices makes the collection operations to be digital. Since the digital operations can save the resources wasted in physical operations and improves the efficiency of the operations, digital information collection is widely used in many applications. Image clipping, one of the most widely used information collection operation, is defined as a series of operations that extract a whole photograph or a part of a photograph from printed materials such as books or papers. The image clipping presents lots of benefits than the conventional physical copy operations. It reduces the waste of papers and it enhances the efficiency of image collection operations. Many applications programs available on smart phones are developed for image clipping operations. Unfortunately, there are several environments where smart phones are not allowed for security. In this paper, we present an image clipping system composed of a projector, an HD webcam and a Microsoft Kinect (see Fig. 1). We also present a hand gesture-recognition interface through which users can control the whole image clipping process. In our system the webcam located on the top clips the image on the paper and captures the users' hand motions. The Kinect is to measure the depth of the hand for correct clipping and control. The project highlights the area to clip on the paper.

The advantages of our system are listed as follows:

(1) We present a robust image clipping environment where the depth of the papers or books where images

* Corresponding author.
E-mail addresses: kwlee@smu.ac.kr (K. Lee), minkh@smu.ac.kr (K. Min).

http://dx.doi.org/10.1016/j.is.2014.05.011
0306-4379/

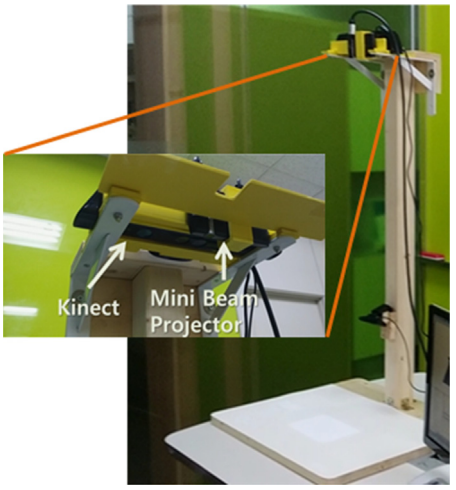

Fig. 1. The structure of our image clipping system.

are clipped and the color of the images to clip do not prohibit the image clipping operations. With the aid of Kinect, the image whose color is similar to that of hands is robustly clipped.

(2) The interface of the system, which is based on the philosophy of natural user interface (NUI), allows an environment where users are not required to be trained or exercised for the image clipping operations. Our interface captures the users' hand motions and analyzes them to control the image clipping process.

The rest of this paper is organized as follows: In Section 2 we review related work on image clipping techniques. In Section 3 we suggest the overview of our system. We show how to clip images using hand motion recognition in Section 4. We present our results in Section 5 and draw conclusions as well as suggest directions for future work in Section 6.

2. Related work

Cerezo [1] presented a real-time tracking system for hands and fingers using the depth image captured Kinect. Trigo and Pellegrino [2] developed a hand gesture analysis is algorithm based on geometric shape of the hand. Their algorithm presents a more excellent performance than the existing HCI systems. Premaratne and Nguyen [3] proposed a system that recognizes a set of pre-defined hand motions on several devices. Their system works in real-time and presents very precise recognition result. Girshicky et al. [4] presented a novel approach for recognizing human gestures by using depth information. They extended the existing recognition algorithms through reasoning and learning process. Sato et al. [5] developed a system that tracks the center of hands and fingertips. After extracting hands they apply a template matching strategy to track the fingertips. An interesting approach is that they used infar-red images to track users' hand motions on desktop environment. Oikonomidis et al. [6] implemented Kinect-based 3D hand joint tracking system using both RGB and depth images. Keskin et al. [7] applied hidden Markov model to track hands in real-time and developed human–computer interaction interface based on 3D gesture recognition. Using a pair of web cameras, they trace 3D positions and directions of hands without markers and gloves. Chen [8] developed a color-based hand motion recognition system. Segen and Kumar [9] presented a hand gesture recognition system using a lighting device and a camera. Their system recognizes four gestures: Point, Reach, Click and Ground. Na et al. [10] exploited depth information to segment hand regions and finally extracts hand centers and fingertips. The learning system based on the extracted information recognizes the pre-defined hand gestures.

3. Overview of the system

The overview of our system is illustrated in Fig. 2. Our system is composed of four modules as follows:

(1) *Motion Recognition Module*: A module that recognizes users' hand gesture using Microsoft Kinect.
(2) *Projection Module*: A module that presents interactive user interface using a small beam projector.
(3) *Capture Module*: A module that captures the region specified by users' motion.
(4) *Data Manage Module*: A module that collects the clipped images.

Since these modules operate in their own coordinate system, we devise Co-ordinate Convert Module that unifies the coordinates of the modules. We present a hand motion recognition scheme for Motion Recognition Module to specify the region to clip.

4. Recognizing hand motion

4.1. Extracting hand region

The first step of hand motion recognition is to extract hand region. Our strategy is to use depth information collected using Microsoft Kinect. For the depth map from Kinect, we determine the smallest value from the depth

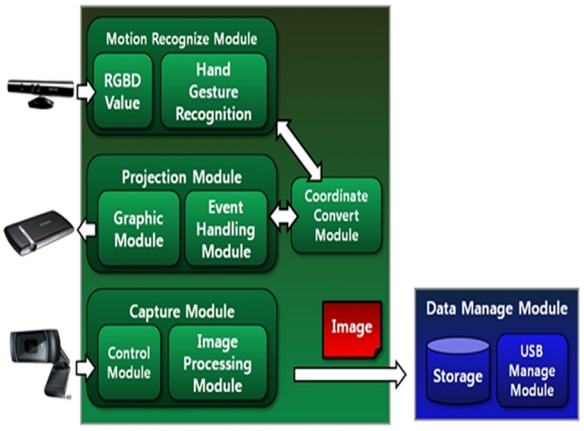

Fig. 2. The overview of the algorithm.