Remembering the Object Location with a Wearable Vision Interface

Takahiro Ueoka, Tatsuyuki Kawamura, Norimichi Ukita, Yasuyuki Kono, Masatsugu Kidode
{taka-ue, tatsu-k, ukita, kono, kidode}@is.aist-nara.ac.jp
Graduate School of Information Science, Nara Institute of Science and Technology

Abstract

We have developed a wearable vision interface system to aid users in remembering the location of objects used in daily life. Our system includes both registration and retrieval functions. Object identification is made through a novel camera device, which can easily extract only a region of a target object from an image. In addition, the identification of the object consists of two levels. The first level includes a rough identification which discriminates between similar objects. The second level includes identification with a subspace method from roughly identified candidate objects.

1 Introduction.

People often forget where they left an object. To solve this problem, we have developed a wearable object registration/retrieval system. A user of this system has to register objects to the system as targets of remembrance in advance. The system records an image in a user's viewpoint at all times by identifying an object in the image from the registered objects. The system then adds the index information (i.e. the name of the object) to the image. When the user chooses a registered object as the target object, the system retrieves the recorded images with the index and provides the last recorded image containing the target object. The user can remember the location of the object according to the perceived image including the target object.

2 Object registration/retrieval system

Figure 1 shows the system worn by a user. The system consists of a camera device (Figure 1(a)), a wearable PC (WPC), and a head-mounted display (HMD),(Figure 1(b)). The camera device takes images in the user's viewpoint, and transmits the images to the WPC. The WPC records input video sequentially, extracts the object region from the input image, and then makes output information for the user. The HMD gets information from the WPC, and displays the information to the user.

2.1 Combined camera device

We have developed a head-mounted camera device to take synchronized color and infrared images. The camera device includes an infrared light source, infrared and color CCD cameras, and a beam splitter (Figure 1(c)). The infrared light source diffuses infrared rays to an environment in front of the user. The infrared CCD camera takes an infrared image, which is a monochrome image lighted by the infrared light source. At the same time, the color CCD camera takes a color image. Due to the physical relationship between the two cameras and the beam splitter, color and infrared images are taken in the same ray line.

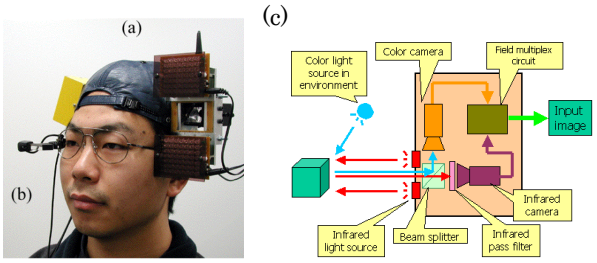

Figure 1: wearable system
(a) camera device
(b) head-mounted display
(c) structure of camera device

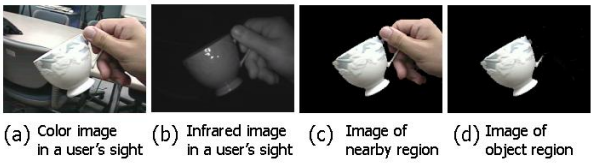

Figure 2: Extraction of object image

We use the combined camera device to extract only the image of the object held by a user [1]. The image in the user's viewpoint includes the background region, the object region, and the hand region holding the object (Figure 2(a)). To extract only the region of the object, we have to eliminate the background region and the region of the hand. In general, unless a strong infrared light source exists in the environment, the intensity of reflected infrared ray becomes darker as the further away the target object is from the camera device. In this way, we binarize the reflected infrared image (Figure 2(b)) by the threshold, and make a mask to eliminate the background region. Applying the background mask to the color image, we can get an image which includes only the nearby region that consists of the target object and user's hand (Figure 2(c)). Eliminating the region colored the same color as that of the user's skin from the nearby region, we can get an image which includes only the object region (Figure 2(d)).

2.2 Object registration and retrieval

In this section, we explain the functions of an object registration/retrieval system for remembering the location of an object. These functions utilize the images of the object region, which are extracted from the color and infrared images taken by the wearable combined camera.

First we describe the operation of object registration. A user wearing this system registers objects as targets of remembrance. In the registering operation, the user first inputs the name of the object into the system, and then holds the object and moves it into sight, including arbitrary rotations and movements (Figure 3(a)). The head-mounted camera device takes images of the object in a few seconds.

The video of an operation from the user's viewpoint includes object images in various shapes and various sizes. In the same shape set of the object image, a somewhat linear relationship exists between the size and the median luminance of the reflected infrared ray. As a feature value of the object, we use this relationship. This relationship shows an advantage in allowing us to easily discriminate between objects with relations of similarity, in other words, between those the same in shape and those different in size.

Next, we describe a method of registering an object's feature values. The system extracts the object regions from the source images by a circumscribed rectangle, and converts each object region image to a normalized image. Applying the eigenspace method [2] to the normalized image set, the system can make the manifold in the eigenspace of the object. Dividing the manifold into spatial regions, the system can make the same shape sets of the object image. The system derives the rough identification parameters that represent the relationship between the size and the median luminance of an infrared object image. The system then records the eigenspace vector and the rough identification parameters of the object as the object feature values (Figure 3(b)). In the next sentence, we describe two object retrieval functions in this system.

Observation object retrieval: Unless receiving any commands from the user, the system should extract object regions from input images and identify the observed object as one of the registered objects. In rough identification, the system has to identify the observed object by rough identification parameters to discriminate among the similar objects and to reduce the number of candidate objects. Among the roughly identified objects, the system identifies the observed object using the subspace method [3]. Finally the observed object is identified with the output of the subspace method.

When the identification succeeds, the system should associate the information of the identified object (i.e. the name of the object) with the input image, and record the image with that information. At the same time, the system should produce information about the object to the user as the result of retrieval. The information associated with the recorded image will be used in another object retrieval function described in the next sentence.

The assignment of object retrieval: When the user wants to remember the location of an object,

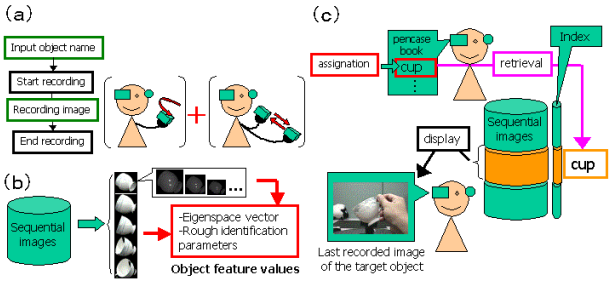

Figure 3: Object registration and retrieval
(a) user operations of the object registration
(b) feature extraction of the object
(c) the assignment of object retrieval

the user uses the function of the assignment of object retrieval. In this function, the system presents a name list of registered objects to be assigned the name of the target object by the user (figure 3(c)). When assigned, the system refers to the index of the names of the objects associated with the recorded images. Then the system retrieves the last images that include the target object. Finally, the system displays the retrieved images on the HMD.

3 Concluding Remarks

With this study we aim at a system that retrieves and recollects objects that we choose to remember in everyday life. We have made the wearable combined camera device on a trial basis, and have developed the object registration/retrieval system using images taken by the camera device.

To enhance the robustness of the system in an environment including a strong infrared ray, we will improve the infrared light source in a combined camera device and explore the method of extracting the object region. Lastly, we will continue work on reducing the weight and size of wearable devices.

Acknowledgements

This research is supported by the Core Research for Evolutional Science and Technology (CREST) Program and "Advanced Media Technology for Everyday Living" of the Japan Science and Technology Corporation (JST).

Reference

[1] Hiroshi Sasaki et al., "Hand-area extraction by sensor fusion using two cameras for input interface of wearable computers", *Scandinavian Conference on Image Analysis 2001*.

[2] Hiroshi Murase and Shree K. Nayar, "Illumination planning for object recognition using parametric eigenspace," *IEEE Trans. on PAMI*, Vol. 16, No. 12, pp.1219-1227, 1994.

[3] Y.Akiri, N.Ishikawa and Y.Sugiyama,, "Extraction and recognition of facial regions by subspace method," *ACCV 1995*, pp.III-738 - III-742, 1995.