

BoxFinder: Finding items in boxes using images and visual markers

Mizuho Komatsuzaki¹, Koji Tsukada¹ and Itiro Siio¹,

¹ Ochanomizu University, 2-1-1 Otsuka, Bunkyo-ku,
Tokyo, 112-8610, Japan
komatsuzaki.mizuho@is.ocha.ac.jp, {tsuka, siio}@acm.org

Abstract. We propose a novel search technique using images and visual markers to help users find commodities stored in boxes. Users capture images of the commodities stored in boxes with a visual marker, and the images are automatically uploaded onto an online database tagged with the box ID. Users can then browse these images using either an offsite Web browser or an onsite mobile computer. In the latter case, users superimpose images on the corresponding boxes using a camera-equipped portable computer. Moreover, while browsing the images, users can detect the location of the boxes using visual markers attached to each shelf. This information is used to locate a box in future quests for items.

Keywords: Search by image, visual marker, storage box

1 Introduction

People often find it difficult to remember and locate items stored in boxes. Thus, they often search an entire room and open many boxes before finding the target items. Classification is a method often used to avoid these problems; however, it is cumbersome to always categorize each item. We propose a novel search technique using images and visual markers to help users find unclassified items stored in boxes. There are various techniques to help people find items: a system employs active RFIDs and ultrasonic position detectors to locate a lost item [6]; a system utilizes a wearable camera to record everything around a user [7]; a system uses human detectors and speakers attached to objects to react approaching users by sound [8]. However, these systems require tedious setup process and huge cost, and are not practical in daily use.

2 BoxFinder

Our technique mainly consists of two steps: (1) registering images of items stored in boxes and (2) searching them on- and offsite (fig. 3). ARToolKit [1] visual markers are attached inside/outside boxes and to shelves, as shown in Figs. 2. The markers are both machine and human readable.

Adjunct Proceedings of Pervasive 2010 - Demos. The authors maintain the copyright of the paper.

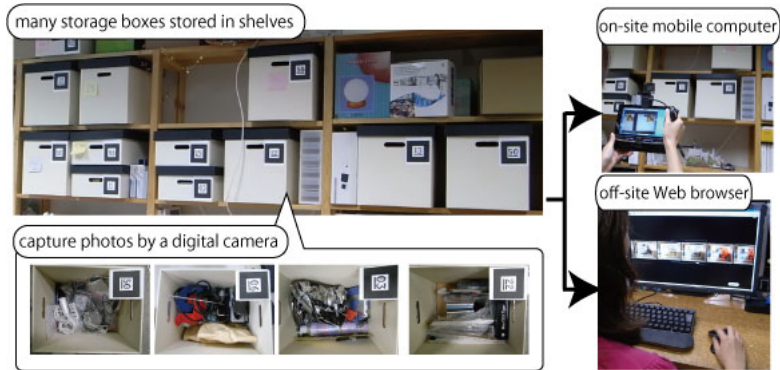


Fig. 1. The concept of BoxFinder. Users can easily capture photos of items stored in many boxes and search them on- and offsite.



Fig. 2. Example of a storage box with visual markers. ARToolKit visual markers are attached inside/outside boxes.

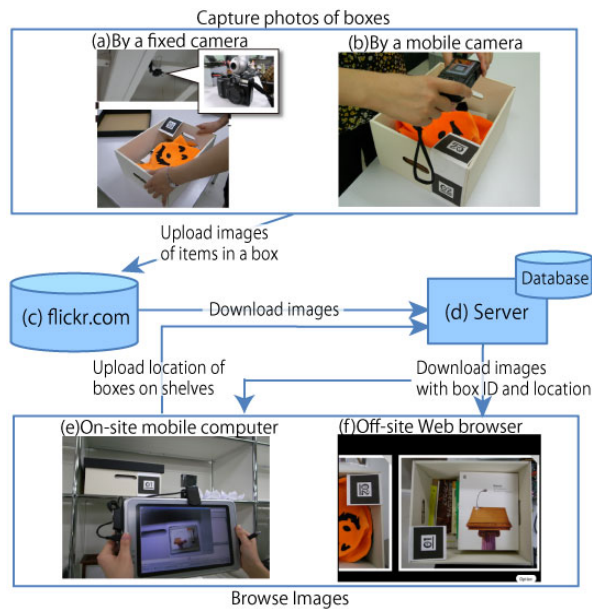


Fig. 3. System Architecture of BoxFinder

2.1 Capture photos of boxes

A user of the system takes a picture by a handheld digital camera when some contents of a box are updated (fig. 3b). To simplify this process, we have made an automated

camera that takes picture of a box when it is placed in a specified place such as a working table (fig. 3a and fig. 4). In this system, we have installed a regular resolution web camera (Logicool Qcam (QVP-30SV)) and a high-resolution digital still camera (Canon G10) on the ceiling above a working table in front of shelves where boxes are stored. The web camera monitors ARToolKit markers on the table. Once a marker is detected, the high-resolution digital camera starts to take pictures. The pictures are transferred to a PC through USB and uploaded to an online image storage server (e.g. Flickr [2], fig. 3c). Using this system, a user can take pictures in a box and store them to the server by simply opening the lid on the table, which is a common action in an item-locating task.

In some cases, a box is too heavy or located too far, and it is difficult to bring it to the working table. A handheld camera, mentioned above, will be used instead. Users capture a picture of the items stored in a box (including a visual marker) with a digital camera (fig. 3b), and the picture is uploaded onto Flickr (fig. 3c) via wireless network using an Eye-Fi card [3] equipped in the camera. The server (fig. 3d) is used to store the pictures, box IDs, and location of each box, if it is available. The server periodically downloads the pictures from Flickr and identifies and locates boxes by decoding the markers in the pictures.



Fig. 4. Capturing photos of boxes by a fixed camera. Users can easily capture an image of the items stored in a box only by putting the box on the photo space.

2.2 Browse images

Users can browse these images using either an onsite mobile computer (fig. 3e, fig. 5) or an offsite Web browser (fig. 3f, fig. 7). When users hold a camera-equipped mobile computer like NaviCam [4] in front of a box, they can browse the latest image of the contents of the box (fig. 3e, fig. 5). These images are superimposed on the visual markers attached outside the boxes. While users search for an item using the mobile computer, the relative distance between the box and the shelf markers is estimated (fig. 6). Thus, users detect the location of the boxes without any additional efforts.

To find an item, users can browse images from anywhere using the offsite Web browser (fig. 3f, fig. 7). Users can browse photos in the numerical order of boxes. The scale of photos is changed smoothly in proportion to scroll speeds [5]. Once users find the item in an image of a box, they can start locating the box by looking for the marker attached to the target box. The task of locating boxes will be much easier if the location is known previously using the mobile computer, as described above. A user can locate the shelf with the box on the Web browser screen (fig. 7).

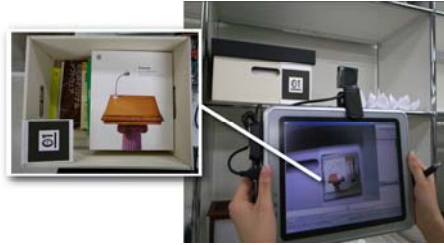


Fig. 5. Browsing images using an onsite mobile computer.

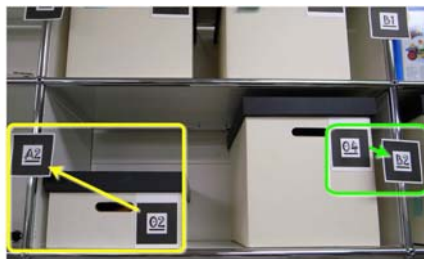


Fig. 6. Detecting the location of the boxes. The system relates shelf ID to box ID that the relative distance is the shortest.



Fig. 7. Browsing images using an off-site Web browser.

References

1. www.hitl.washington.edu/artoolkit.
2. www.flickr.com.
3. www.eyefi.com.
4. Rekimoto, J. NaviCam: A Palmtop Device Approach to Augmented Reality. *Fundamentals of Wearable Computers and Augmented Reality*, 2001.
5. Igarashi, T., Hinckley, K. Speed-dependent automatic zooming for browsing large documents. 13th Annual Symposium on User Interface Software and Technology, ACM UIST'00, San Diego, CA, November 5-8, 2000, 139-148.
6. Toyohisa Nakada, Hideaki Kanai, Susumu Kunifuji, "A Support System for Finding Lost Objects using Spotlight", 7th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI), 19 to 22 September 2005, University of Salzburg, Austria
7. T. Kawamura, T. Ueoka, Y. Kono, M. Kidode. Evaluation of View Angle for a First-person Video to Support an Object-finding Task, Proc. ICCS2006 (The 5th International Conference of the Cognitive Science), Vancouver, Canada
8. Shinnishi, Makoto; Iga, Soichiro; Higuchi, Fumito; Yasumura, Michiaki, "Hide and Seek: Physical Real Artifacts which Respond to the User", Proc. of SCI'99 (The 3rd. World Multiconference on Systemics, Cybernetics and Informatics) and ISAS'99 (The 5th. Int'l Conference on Information Systems Analysis and Synthesis), (1999 Aug)