

Visual Querying with Iconic Sketches for Face Image Retrieval

Gamhewage C. de Silva and Kiyoharu Aizawa

Department of Information and Communication Engineering, The University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
{chamds,aizawa}@hal.t.u-tokyo.ac.jp
<http://www.hal.t.u-tokyo.ac.jp>

Abstract. We propose a user interaction strategy for automated retrieval of face images. Face icons are placed on a rectangular drawing area to form an iconic sketch. This sketch is submitted as a visual query for image retrieval. The sizes and locations of face icons on the sketch are compared with results from automatic face detection on images, to find the best matching results. We design interactive user interfaces for creating the iconic sketches on multiple device platforms, in a simple and intuitive manner. We demonstrate two applications of the proposed interaction strategy; an image search engine, and a personal photo organizer.

Key words: Iconic sketch, face icons, visual querying, image retrieval, multimedia

1 Introduction

There has been a rapid growth of research in the area of Content Based Image Retrieval [1, 3]. The input to most of the existing image CBIR systems is a textual description about the contents of the image. Visual querying seems more prospective due to the difficulty in describing an image using only text. However, querying by detailed sketches has limited effectiveness due to the time consumed and sensitivity to color differences [10].

“Face” is an important category of content in photographs. Digital cameras utilize face detection to adjust camera settings, to ensure that the faces are in focus and well exposed. Social networks facilitate marking *faceboxes* in images for associating photos with people. Use of face detection for photo retrieval has recently received increased attention. Digital photo organizer programs such as *Picasa* [8] and *Apple iPhoto* [7], and search engines such as *Google Image Search* [5] use face detection and recognition for image retrieval. However, only a limited set of inputs, such as the presence/absence of faces and name of a person, are used as input in these systems.

This paper presents a user interaction strategy that utilizes faces as a cue for content based retrieval of digital photographs. It allows the user to describe a photo using an iconic sketch query that specifies the layout of faces in the photo,

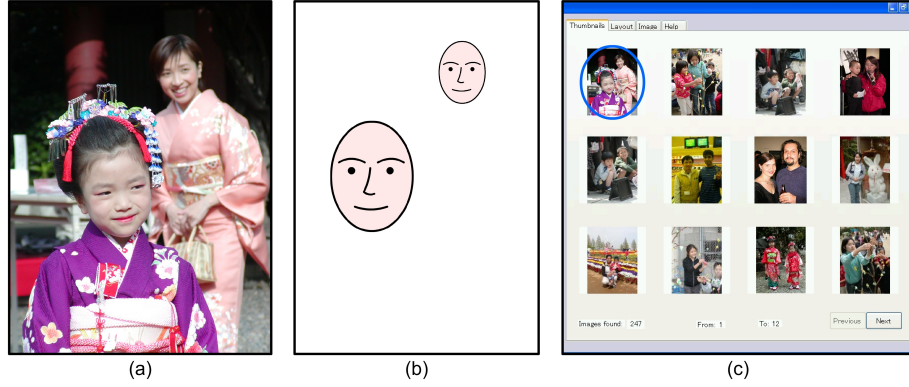


Fig. 1. The proposed user interaction strategy: (a) example “desired image” (b) iconic sketch for retrieval (c) images retrieved.

for more effective querying and retrieval. We implement interfaces using the proposed interaction strategy on two hardware platforms, and also demonstrate two applications.

2 Visual Query Formation

We propose an iconic sketch to represent the composition of an image with respect to faces. First, the orientation of the desired image is specified to initialize a canvas. The sketch is composed by placing face icons on the canvas and changing their sizes and locations. This sketch can be used as a visual query to represent a known image, or to specify the composition of images to be retrieved. Figure 1a shows an example image that a user might desire to retrieve. Figure 1b illustrates an iconic sketch representing this image.

Upon submission of the search parameters, attributes of the canvas and face icons are compared with metadata resulting from applying face detection to a collection of images, to retrieve similar images. The results are ordered according to their similarity to the face layout specified in the iconic sketch. Figure 1c shows a set of search results retrieved using the visual query in Figure 1b.

While the proposed iconic sketch notation seems simple and not scalable to represent classes of content other than faces, it is quite versatile when combined with user knowledge, context and other information. Recent researches have shown that more than 42% of image searches are for images of people [9, 11]. Combined with face recognition or manual annotation, the proposed interaction strategy can greatly enhance the search capabilities of both commercial photo databases, and personal photo organizers. For instance, the current version of *iPhoto* supports photo retrieval by only a single person name at a time. By combining the proposed interaction strategy with face recognition, we can create a system that can retrieve photos on which a number of known persons appear

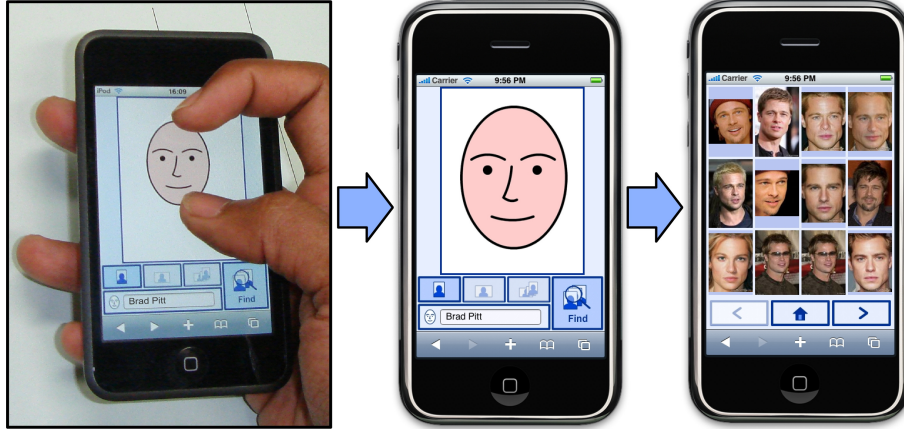


Fig. 2. Using the image search engine on an iPhone.

together. It is also possible to specify how the photo is composed, in case of searching for a known photo or having a preferred composition.

A significant advantage of the proposed notation is its ability to quickly express image composition. User studies by Westman et al. [11] show that image composition and layout has a very high level of significance in image search for professional work such as journalism. Commercial stock photo databases [2, 4] allow users to filter photos of persons by limited categories of composition such as “head-shot”, “waist-up” and “full-length”. Such categorizations are performed manually, and are therefore subjective. By changing the size and location of the face icon with respect to the canvas, a user can specify a much wider range of compositions and retrieve images with a higher precision.

3 User Interaction Design

We implemented the proposed interaction strategy on personal computer and *Apple iPhone/iPod Touch* platforms. For personal computers, the interface was designed such that only a pointing device is necessary for synthesizing the visual query. For iPhone/iPod Touch platforms, both single touch and multi-touch gestures, that are commonly used with other applications, were used.

For querying on personal computers, adding a face icon is performed by left-clicking the mouse on canvas. A face icon is selected by left-clicking the mouse on it. The user can drag an icon within the canvas to change its location. The size of a face can be changed either by rotating the mousewheel (if available) or by adjusting a slider, after selecting it. A face icon can be deleted by right-clicking the mouse on it.

For iPhone/iPod Touch devices, only a single face icon is allowed due to the limited screen size. The face icon can be moved by touching and dragging it with

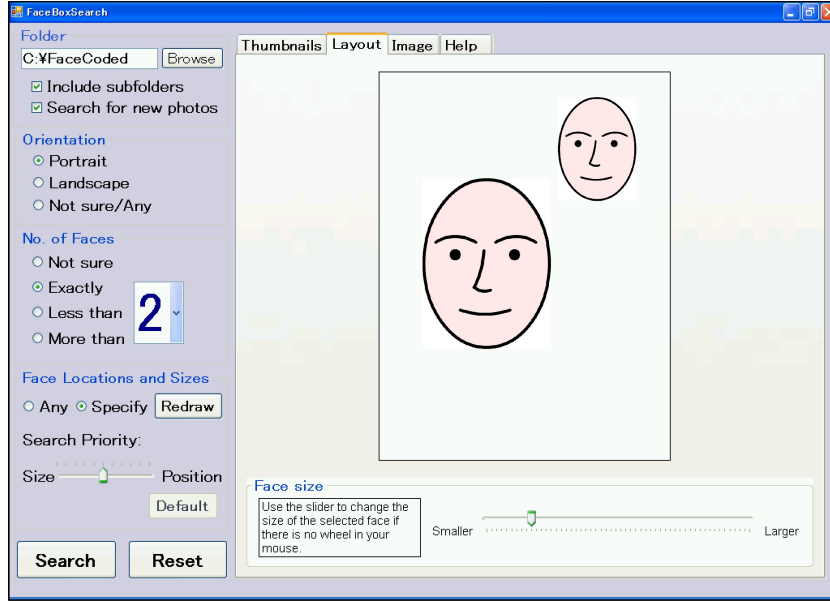


Fig. 3. Using the personal photo organizer for image retrieval.

the finger. The size of the icon is increased or decreased by placing two fingers on the icon and moving them away from or towards each other, respectively. The gesture for moving the face icon is same as that for rearranging program icons on the screen of these devices. The gestures for resizing the face icon are the standard gestures for resizing content on multi-touch enabled interfaces. Therefore, the selected gestures are very easy for users to learn.

4 Applications and Feedback

We demonstrate two applications of the proposed user interaction strategy. The first is a prototype search engine that retrieves photos of famous persons. The search index includes the locations and sizes of faces present in each image indexed by the search engine. The user can enter the name of the person together with a visual query, and retrieve images of the person with the specified layout. This search engine can be found at www.faceboxsearch.com, and is usable on both personal computers and iPhone/iPod Touch devices. Figure 2 illustrates how the application is used on an iPhone device.

The second application is a personal photo organizer. Face detection is carried out as a background process when photos are copied to the computer from a digital camera. The metadata are stored in the *comments* area of the JPEG image files. Given that the maximum length of a JPEG comment is 65536 bytes [6], it is easy to accommodate attributes of more than 1000 faces (sufficient for practical situations). Having metadata within the image files themselves make

them “ready to search” even when transferred to another computer. We demonstrate this system on a collection of 10,000 photos from a personal collection. Figure 3 shows the inputs for retrieving the image in Figure 1a.

We conducted a user study on the second system, and are in the process of analyzing the results. The users were able to learn to use the system quickly. The average search time for retrieving a desired image from the above collection was 32.3 seconds. This included the time taken for making the sketch, searching (by the system), and browsing the search results to find the desired image. The users commented that the interaction strategy is easy to learn, fast, and enjoyable to use.

5 Conclusion and Future Work

We presented a user interaction strategy based on iconic sketches for face photo retrieval. The visual queries based on iconic sketches provide an easy way to specify the composition of the desired image. User interfaces for both personal computer and iPhone/iPod platforms allow fast query formation. The interaction strategy was employed on two applications. The users found the interaction strategy intuitive, and easy to learn. We are working on extending the functionality of the personal photo organizer by incorporating face recognition.

References

1. Chang, S., Chen, W., Sundaram, H.: VideoQ: a fully automated video retrieval system using motion sketches Applications of Computer Vision, 1998. WACV '98. Proceedings., Fourth IEEE Workshop on 19-21, Page(s):270 - 271 (1998)
2. Stock Photography, Footage and Rights: Corbis, <http://www.corbis.com>
3. Datta, R., Joshi, D., Li, J., and Wang, J.Z.: Image retrieval: ideas, influences, and trends of the new age. *ACM Computing Surveys*, 40(2):5:1–5:59 (2008)
4. Getty Images: Stock Photography, Editorial Photos, Footage and Music, <http://www.gettyimages.com>
5. Google Image Search, <http://images.google.com>
6. Hamilton, E.: *JPEG File Interchange Format*. C-cube Microsystems, Milpitas, CA95035 (1992)
7. iPhoto: Faces, Places and Other Features, <http://www.apple.com/ilife/iphoto>
8. Picasa: the Easy Way to Find, Edit, and Share Your Photos, <http://picasa.google.com>
9. Rodden, K.: Evaluating similarity based visualizations as interfaces for image browsing. Technical Report UCAM-CL-TR-543, University of Cambridge (2002)
10. Rouw, M.: *Meaningful image-spaces and project photoindex*, Masters thesis. Utrecht schools of the arts, Hilversum, the Netherlands (2005)
11. Westman, S., and Oittinen, P.: Image retrieval by end-users and intermediaries in a journalistic work context. In *Proceedings of the 1st International Conference on Information Interaction in Context*, volume 176, pages 102–110 (2006)