A Real-time, Embedded Face-Annotation System

Shih-Wei Chu
Dept. of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106, USA
chushihwei@umail.ucsb.edu

Mei-Chen Yeh
Dept. of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106, USA
meichen@umail.ucsb.edu

Kwang-Ting Cheng
Dept. of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106, USA
timcheng@ece.ucsb.edu

ABSTRACT
Face detection and recognition have numerous multimedia applications of broad interest, one of which is automatic face annotation. There exist many robust algorithms tackling these problems but most of these algorithms are computationally demanding and have only been implemented in PC- or server-based environments. In this demonstration we show a real-time face-annotation system on a commercial PDA development platform. We examine the challenges faced in the design and development of a practical system that can achieve detection and recognition in real-time using limited memory and computational resources which are common constraints for embedded applications.

Categories and Subject Descriptors
C.3 [**Computer Systems Organization**]: Special-purpose and Application-based Systems – *Real-time and embedded systems.* I.5.4 [**Pattern Recognition**]: Applications – *Computer vision.* I.5.5 [**Pattern Recognition**]: Implementation – *Interactive systems.*

General Terms
Design, Experimentation.

Keywords
Face annotation, embedded system.

1. INTRODUCTION
Mobile telephony has become increasingly common and popular in recent years. With developments in handsets that have increased their storage capacity and expanded their communication capabilities, most mobile devices such as cellular phones and PDAs are capable of taking photographs and transferring images to the Internet, PC, or other mobile devices; however, given the increasingly complex image data captured and stored, these devices do not provide effective means of helping users manage their data.

In this demonstration we show a real-time face annotation system implemented on a cellular-phone-based embedded platform. In general, an image can be represented by various types of information. There are a number of potential applications of broad interest that could link visual data with other content. For example, the faces in a photograph could be linked to names listed in the user's address book or social-networking websites. Users could then browse or exchange their media collections based on these names.

A face annotation system contains two functions: face detection and face recognition. Both topics are very popular in computer vision. Many robust algorithms have been developed. Please see [3] and [4] for literature surveys. However, these algorithms are primarily designed for use on PC-like or server-based systems and, as such, achieving high accuracy is considered more important than execution efficiency.

Embedded systems have less computational resources available. This characteristic creates challenges for existing algorithms to be implemented on an embedded platform. For example, the Intel Open Source Computer Vision library (OpenCV) [5] provides many useful image processing libraries and computer vision algorithms. Porting OpenCV to embedded systems, however, will result in lower performance due to the huge amount of computation and memory usage required. Designing algorithms that can achieve both effectiveness and efficiency is the main concern in creating our face annotation system. More specifically, the design and implementation of an embedded face annotation system should take the following factors into account:

- **Low-quality inputs:** A mobile device generally captures images with a lower resolution and quality compared to a digital camera.
- **Limited computational resources:** Limited amount of memory and a less powerful processor are common characteristics of embedded designs. Algorithms that can achieve great detection/recognition accuracy using complex mathematic computations or floating-point calculations, unfortunately, may not be feasible due to these limited resources.
- **Limited space for display:** The display panel of a mobile device has a limited size and a relatively low resolution. The system thus requires a user interface that can effectively arrange all required functionalities in the compact display panel.

In the following we describe in detail our system design, including the face detection and recognition algorithms, and the demonstration setup.

2. SYSTEM OVERVIEW
We built our face annotation system on the Orchid Standard Carrier Board from the Toradex Corporation. This board consists of an Intel processor chip called the XScale PXA 320, and is based on the ARMv5TE architecture with a maximum clock rate

MM'08, October 26–31, 2008, Vancouver, British Columbia, Canada.
ACM 978-1-60558-303-7/08/10.

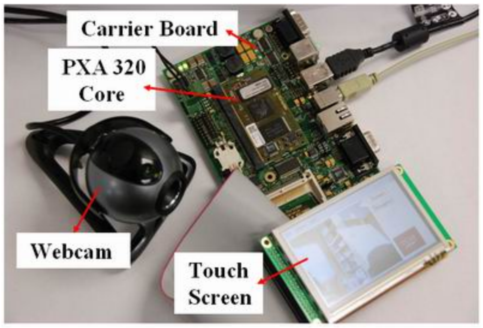

Figure 1. System hardware.

of 806 MHz. This clock rate is commonly used in mobile devices. Also, we used a touch screen as the I/O device, and a webcam to capture still images. Fig. 1 shows our system hardware.

Fig. 2 illustrates our system framework. Given an input image, we first locate all human faces through a face detector. Next, each localized face area is assigned a name by a face recognizer. During both processes, the user can be involved in this computing loop and provide labels. To deal with situations in which unknown people or false positives from the face detector appear, our system provides two additional labels—unknown and non-face—to achieve a better annotation quality.

2.1 Algorithm

We developed the face detector based on the OpenCV face detector [5]. This detection method, originally proposed by Viola and Jones [2], is one of the most widely used algorithms and can achieve a great detection rate with fast computation. This method applies a cascade classifier structure that consists of a series of stages. Each stage consists of several weak classifiers that can quickly reject non-face regions so more computational resources are reserved to examine more promising, face-like regions.

For recognizing faces, we applied the ARENA algorithm [1], which is a memory-based recognition method. It proposes to use $L_{0.5}$ distance rather than the Euclidean distance to calculate the distance between two images. We use ARENA for comparing all pixel values of the detected face with those of images in the training set. It assigns the query image the label of the training image that has the smallest distance in terms of the $L_{0.5}$ measurement. Compared to many transformation-based recognition methods, ARENA needs neither matrix computation nor any type of trained models. It also offers the user with the flexibility to add new names based on new, additional training images. We also performed pre-processing steps to normalize the image illumination before recognition.

Both detection and recognition methods require floating point calculations. However, emulating floating point operations in our target platform is very expensive—typically taking several hundred cycles for one floating point operation. Therefore, instead of performing floating point calculations, we implemented floating-point to fixed-point conversion followed by fixed-point calculations to improve the system performance.

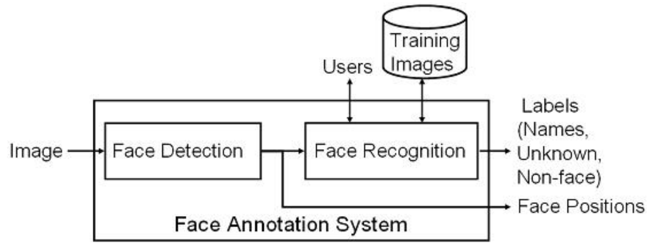

Figure 2. System framework.

Figure 3. User interface.

3. DEMONSTRATION SETUP

The demonstration has the form of a stand-alone application with a graphical user interface as shown in Fig. 3. The system streams the video and displays the image on the touch screen when the application is activated. The user can choose either the "detect" or the "recognize" button. When either button is pressed, the screen freezes on that frame and start the detection or recognition process. We will demonstrate the system's ability to execute real-time detection and recognition. We will also demonstrate an easy-to-use interface for annotating or modifying the results.

4. REFERENCES

[1] T. Sim, R. Sukthankar, M. Mullin, and S. Baluja. "Memory-based Face Recognition for Visitor Identification," In *Proceedings of the IEEE International Conference on Face and Gesture Recognition*, pp. 214-220, 2000.

[2] P. Viola and M. Jones. "Robust Real-Time Face Detection," *International Journal of Computer Vision*, 57(2), pp. 137-154, 2004.

[3] M. Yang, D. J. Kriegman, and N. Ahuja. "Detecting Faces in Images: A Survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(1), pp. 34-58, 2002.

[4] W. Zhao, R. Chellappa, P. Phillips, and A. Rosenfeld. "Face Recognition: A Literature Survey," *ACM Computing Surveys*, 35(4), 2004.

[5] Open Source Computer Vision Library. http://www.intel.com/technology/computing/opencv/