Establishing a Face Recognition Research Environment Using Open Source Software

Phillip I. Wilson and John D. Fernandez

Computing and Mathematical Sciences
Texas A&M University–Corpus Christi

Abstract

This paper discusses establishing a research environment for face detection and face recognition. This research environment provides real time facial detection and feature extraction from digital camera images. This environment allows research to be performed on various aspects of facial biometrics, including facial recognition, face tracking, and emotion-related feature extraction. The environment utilizes Intel's Open Computer Vision Library (OpenCV), an open source, platform independent library. This library provides the ability to analyze, extract, and modify information in real time for each frame of a digital camera or a video file. Tools are provided to create and use Haar classifier cascades to detect faces and facial features within an image or frame. OpenCV provides functions to perform principle component analysis (PCA), including covariance matrix calculations and eigenvectors calculation. The environment also utilizes the National Institute of Standards and Technology's (NIST) Facial Recognition Technology (FERET) database. This database provides hundreds of grayscale and color images of people in various lighting conditions and poses. The ultimate objective of this work is to develop an environment that can be used for multiple research initiatives related to usability and security.

Introduction

Biometrics, the science of reading measurable, biological characteristics of an individual in order to identify them to a computer or other electronic system[1], has become a very popular field of research. In recent years advancements in technology have made researching biometrics less expensive. Facial recognition, in particular, has become a field in which establishing a research environment has become much simpler. Current technology allows a real-time facial detection and recognition system to be created with very inexpensive hardware. Digital cameras with high resolution can be purchased for a marginal sum. Modern personal computers have enough processing capability to execute real-time detection algorithms.

The most difficult task to establishing the research environment is either creating the software or finding the software to perform the detection and recognition functions. Commercial off the shelf (COTS) software is often expensive and too inflexible for research. Intel has created a library, Open Computer Vision Library (OpenCV), to aid researchers in advancing the fields related to computer vision. Facial detection and recognition falls into this category. Intel's library provides various algorithms to perform object detection and recognition.

This library simplifies interfacing with digital cameras and allows a researcher the flexibility to modify standard facial recognition algorithms. It is open source and works on both Linux and Windows platforms.

Another difficulty with facial recognition is collecting enough images of human faces to acquire valid results from any experiments that are performed. The FERET database solves this problem by providing thousands of color and grayscale images to researchers. These images are composed of many individuals in various poses and lighting conditions.

Haar-Based Detection

One of the difficult tasks of face recognition is to differentiate the face from the background. Each frame from a real time image is a collection of color and/or light intensity values. Analyzing each of the pixels to determine where the human face is located can be very difficult due to the wide variation in pigmentation and shape of the human face. However, Haar-based detection simplifies the detection of objects within a digital image.

Haar Features

Instead of using intensity values, Haar detection uses changes in contrast values between adjacent rectangular groups of pixels. The contrast variances between the pixel groups are used to find relative light areas and dark areas. Areas with contrast variances form features, as shown in figure 1, which are used to detect the desired objects within the image[5]. These features can be easily scaled by increasing or decreasing the area of the pixels being examined. This allows features to be used to detect objects of various sizes.

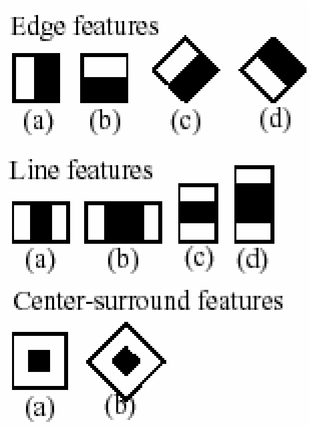

Figure 1. Haar features

The simple features of an image are calculated using an intermediate representation of an image, called the integral image[8]. The integral image is an array containing the sums of the pixel values to the left of a pixel and above a pixel at location (x, y) inclusive. So if **A[x,y]** is the original image and **AI[x,y]** is the integral image then:

$$AI[x,y] = \sum_{x' \le x,\, y' \le y} A(x', y')$$

The integral image array can be calculated with one pass through the original image. Using the integral image, only six to eight array references are required to compute a feature. Thus calculating a feature is extremely fast and efficient. It also means calculating a scaled feature requires the same effort as a non-scaled feature. Thus detection of various sizes of the same object requires the same amount of effort and time as objects of similar sizes.

Haar Classifiers

Although calculating features is efficient, computing the over 180,000 different rectangle features associated within a 24 × 24 sub-window is not feasible. However, only a small number of features are actually needed to determine if the sub-window contains the desired object[3, 8]. The goal is to select the one feature that distinguishes the desired object, like a face, from another object. If a sub-window does not have that feature, then it can be eliminated. All non-eliminated sub-windows are then analyzed for a more features, creating a slightly more complex classifier. This process continues until the desired detection rate is achieved. This method of identifying an object is fast and efficient. Viola and Jones[8] were able to achieve a 95% detection rate of a human face with only 200 features. It took less than one second to identify the faces within a 384 × 248 pixel image. It is possible to achieve a higher degree of accuracy with more features with only a minute increase in detection speed.

The open computer vision library (OpenCV) provides a library to utilize Haar classifiers. In addition to the ability to use the classifiers it also provides several classifiers for face detection as shown in Figure 2. This detection classifier is capable of detecting a face in real-time on a personal computer running a 1 gigahertz or better processor using a fifteen frames per second video stream with a resolution of 320 × 240 pixels. OpenCV also has a program that has the ability to create new classifiers with a desired accuracy by providing images that contain the desired object and images that lack the object. This allows one to easily expand or modify a detection routine to be more effective or more precise. One could build classifiers to identify facial features as well as a human face.

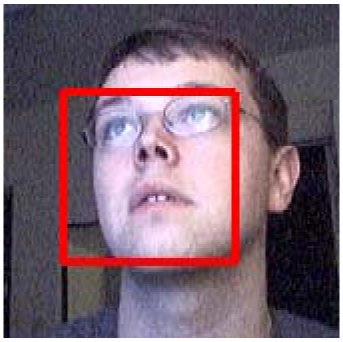

Figure 2. Detection of face in Web Camera Stream

Open Computer Vision Library

Intel's Open Computer Vision Library is a free open source set of C/C++ libraries for computer vision applications. OpenCV is designed to be used in conjunction with applications pertaining to the fields of human-computer interaction, robotics, biometrics, and other areas where visualization is important. The main focus for this paper is the possible applications of OpenCV in the field of facial detection and recognition. In addition to the previously mentioned support for Haar classifiers, there are many libraries that are useful for face detection and recognition[5].

OpenCV provides a simple application programmer interface (API) for interacting with various types of digital cameras, like a web camera. Many facets of this API are implemented on both Windows and Linux platforms. This allows for any programs to be easily ported from one operating system to another. OpenCV has the ability to utilize two cameras simultaneously. In addition it can synchronize the two camera images to create a stereo image. This functionality is useful if one wishes to explore the three dimensional techniques of facial recognition[4, 9].

Principal Component Analysis

One of the most common methods for facial recognition is principal component analysis (PCA)[2]. This method involves projecting a facial image into a subspace, called "eigen space". This subspace is formed from a set of basis vectors called eigen vectors or "eigen faces". The basis vectors are computed by calculating a covariance matrix of a dataset and then deriving the eigen values and eigen vectors. In this case the dataset would be a large set of facial images. By combining all of the eigen vectors one gets an image of a generalized human face. To perform facial recognition, one projects the desired image into face space and compares the resulting eigen vectors to another image's vectors. If a vast majority of the vectors are the same, then the two faces are the same[4].

One of the major advantages of OpenCV is the library to perform principal component analysis using images. This library provides the functions to calculate the covariance matrix, basis vectors, and perform the projections into eigen space using the image data structure used by all capture functions in OpenCV. This allows a researcher considerable time savings since all of the algorithms have already been designed and written.

FERET Database

In order to be able to create the covariance matrix, a large dataset of facial images is needed. The best approach to acquiring such a dataset is to request the Facial Recognition Technology (FERET) database from the National Institute of Standards and Technology (NIST). This database contains thousands of color and grayscale images that are used to test the accuracy of different facial recognition systems[7]. This database is free to researchers studying this field.

One advantage of the FERET Database is that it provides many different images of the same individual. These images include different angles, and lighting conditions. In addition, some of the images are grayscale, while others are color images. Since the accuracy of face recognition systems can be dependent on the changes of camera angles and lighting conditions of the captured image versus the comparison images[4], having various poses can test the versatility of a system. Also, the various poses allow the training set to be composed of images with the same angles and lighting conditions, producing the most accurate eigen faces to be used in comparisons.

The most important aspect to researching facial recognition is the results of the research. Prior to the FERET database many facial recognition systems boasted fantastic results[6]. However, many of the databases that these systems used contained very few images. This caused the results to be inaccurate. With the creation of the FERET database all researchers have access to a large quantity of images to perform the recognition process. This allows researchers to get more accurate statistics on the false positive and false negative rates for their algorithms and their recognition systems.

Future Plans

With an established research environment, the next goal is to research the ability for recognition of facial features that dictate general human emotions, like happy and sad. Recognition of human emotion would require detection and analysis of the various elements of a human face, like the brow and the mouth, to determine an individual's current expression. The expression can then be compared to what is considered to be the basic signs of an emotion. This research will be used in the field human-computer interaction to analyze the emotions one exhibits while interacting with a user interface.

References

1. M. Bishop. *Computer Security, Art and Science.* Massachusetts: Pearson Education Inc. 2003.
2. B. Draper, B. Kyungim, M.S. Bartlett, and J.R. Beveridge. "Recognizing faces with PCA and ICA," *Computer Vision and Image Understanding*, vol. 91. pp 115-137, July 2003.
3. P. Menezes, J.C. Barreto, and J. Dias. "Face Tracking Based on Haar-like Features and Eigenfaces," *5th IFAC Syposium on Intelligent Autonomous Vehicles*, Lisbon, Portugal, July 5-7, 2004.
4. N. Muller, L. Magaia, B.M. Herbst. "Singular Value Decomposition, Eigenfaces, and 3D Reconstructions," *SIAM Review*, Vol. 46 Issue 3, pp. 518 – 545. Dec. 2004.
5. *Open Computer Vision Library Reference Manual*. Intel Corporation, USA, 2001.
6. P. J. Phillips, P. Grother, R.J. Micheals, D.M. Blackburn, E. Tabassi, and J.M. Bone. "FRVT 2002: Overview and Summary," *Technical Report*, National Institute of Standards and Technology, March 2003.
7. P. J. Phillips, A. Martin, C.L. Wilson, and M. Przybocki. "An Introduction to Evaluating Biometric Systems," *IEEE Computer*, pp. 56-63, February, 2000.
8. B. Weyrauch, J. Huang, B. Heisele and V. Blanz "Component-Based Face recognition with 3D Morphable Models," *IEEE Workshop on Face processing in Video*, FPIV04, Washington, D.C., 2004.
9. P. Viola and M. Jones. "Rapid Object Detection Using Boosted Cascade of Simple Features," *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, 2001.

PHILLIP I WILSON

Phillip Wilson is a graduate research assistant for the Department of Computing and Mathematical Sciences. He is working towards his master's degree at Texas A&M University – Corpus Christi. He holds a Bachelor of Science degree in Computer Science and Mathematics. Phillip's research interests are computer security, biometrics, and cryptography.

JOHN D. FERNANDEZ

Dr. Fernandez is Assistant Professor of Computer Science in the Department of Computing and Mathematical Sciences. Having served 20 years in the U.S. Air Force and 10 years in private industry, Dr. Fernandez brings real-world experiences into the classroom for his students. His research interests are in HCI, information assurance, and software engineering.