

Technical Characteristics of Database Collection

Project 12, eNTERFACE07

August 9, 2007

Contents

1	Introduction	2
2	Data Collection Room	2
3	Photos	2
4	Palm Scans	3
5	Audio/Video Data	3
5.1	Subject Reading fixed text	4
5.2	Subject performing four gestures	4
5.3	Subject moving head while speaking freely	5
6	Legal Issues	5

1 Introduction

One of the tasks of Project #12 “Benchmark for Multimodal Biometric Authentication”, which ran during the eNTERFACE’07 Summer Workshop in Istanbul, was to collect a multimodal database. We collected three different types of data of 92 subjects :

1. hand (palm) scans,
2. photos,
3. audio/video samples,

As explained below, the video samples include videos of gestures that are used by eNTERFACE Project #6 “Event Recognition for Meaningful Human-Computer Interaction in a Smart Environment” in the same workshop. This gesture data is not needed in Project # 12.

We report in this document the technical features of the database and the conditions in which the data was gathered, for reference of future users.

2 Data Collection Room

We took the data of the subjects in a small room of 10 square meters. Its height is approximately 3 meters and its walls are yellow colored. We set up a blue background canvas on the wall. The canvas was 1 meter long and 0.5 meters wide. We placed the canvas next to the only window in the room. The window shutters were closed to block sunlight. We attempted to illuminate the whole room uniformly by means of artificial light only. Because of the fact that we could not perfectly cover the whole window, the right side of each subject (left on the shots) is slightly more brightly lit, especially in shots taken between the times 12.00h and 15.00h. The room temperature was approximately 35° Celsius. This temperature and the associated humidity may have caused undesired bright patches in some of the photo shots.

3 Photos

A Cyber-shot camera, version DSC-W30 was used to take the photographs. Its resolution is 6.0 Mega Pixels. Shots were taken using a camera tripod placed 2 meters away from the background and at a height of 1.5 meters. We took the photos in full zoom. In spite of this, we could not prevent including some spots of yellow wall in the photos. We took six portraits of every subject:

1. Subject from front;
2. Subject from front, with glasses removed, if subject was wearing glasses in the first picture;
3. Subject with his/her head turned -90° about the zenithal axis;
4. Subject with his/her head turned -45° about the zenithal axis;

5. Subject with his/her head turned $+45^\circ$ about the zenithal axis;
6. Subject with his/her head turned $+90^\circ$ about the zenithal axis;

In the database, the image files are named as follows

`face_subject%n-%m.jpg`

where `%n` is a number identifying the particular subject, and `%m` is a number between 1 and 6 indicating which of the six pictures above is contained in the file. For example, `face_subject10-5.jpg` contains a picture of subject 10 with his/her head turned $+45^\circ$ about the zenithal axis;

4 Palm Scans

Palm scans of all subjects were taken by scanning the palms of their left hand on a HP Scanjet 4990 flatbed scanner, set to a resolution of 150dpi. Two left hand prints from each subject were taken with fingers spread. After taking the first left hand print we asked to the subject to lift and to rub his/her hand. Then the subject put his/her left hand on the scanner again and the second print was taken.

In the database, the image files are named as follows

`hand_subject%n-%m.jpg`

where `%n` is a number identifying the particular subject, and `%m` is either 1 or 2 indicating which of the two hand prints is contained in the file. For example, `hand_subject9-2.jpg` contains the second hand print taken of subject 9.

5 Audio/Video Data

Audio and video data using a camera attached to an Acer notebook computer, version Espire 5580 was gathered. Its built-in video camera is of VGA type with resolution 640*480 pixels. The notebook was placed on a table one meter away from the subject who stood in front of the blue background. A bidirectional microphone was used and subjects held the microphone under their chins, not covering their faces. Because the microphone was not of high quality, subjects read the text loudly. Even so, in some cases, the audio quality is low.

For each subject, three different videos were taken

1. Subject reading a fixed text;
2. Subject performing four gestures;
3. Subject moving head while speaking freely.

The first of these videos is relevant to Project #12, while the other two were collected for Project # 6. In the database, the video files are named as follows

`video_subject%n-%m.avi`

where %n is a number identifying the particular subject, and %m is either 1, 2 or 3 indicating which of the three videos is contained in the file. For example, `video_subject4-2.jpg` contains the video of subject 4 performing four gestures.

Furthermore, the audio portion of the first video type is extracted to a `wav` file, since only the audio is used by Project # 12. These audio files are named as follows

`audio_subject%n.wav`

where %n is the number identifying the particular subject.

5.1 Subject Reading fixed text

A fixed text was read by each subject. The text was hung on the table, so that the reader could see it. The text was as follows:

It is possible to authenticate individuals by means of the robust hashing or by means of the feature extraction functions. For instance, one may take a photo of the face of a person and robustly hash it in order to obtain a low dimensional descriptor. The same can be achieved by means of feature extraction functions. The identifiers thus obtained can be compared to preexisting ones in a database for a match. Systems based on multimodal robust hashing and feature extraction strategies combine two or more monomodal functions in order to increase security. One two three four five six seven eight nine ten.

The subjects were asked to read the whole text once and repeat the last line five times. They initially practised the text in order to become familiar with it. They were told to continue reading, even if they made errors in pronunciation.

5.2 Subject performing four gestures

A set of four hand gestures were videoed. The subject wore a green glove on the hand used to make the gesture. These gestures were as follows:

1. subject clicks the fingers of their hand, using thumb and middle finger;
2. subject moves hand with out-stretched finger across his/her neck;
3. subject holds hand out with palm facing upwards and moves hand upwards between lower and upper torso;
4. subject holds hand out with palm facing downwards and moves hand downwards between upper and lower torso.

In Project # 6, these gestures were used to represent the commands

1. switch on;
2. switch off;
3. volume up;
4. volume down.

5.3 Subject moving head while speaking freely

In this video, subjects were asked to speak freely in English while moving their head up and down and to the left and right.

6 Legal Issues

Every subject who gave data for the database was asked to sign a disclaimer allowing their data to be used in research projects. The text of the disclaimer is shown below.

CONSENT FORM for ENTERFACE Multimodal Biometric Database v0.1
— July–August 2007

By signing the present form, I understand and consent freely that my personal data, including biometric data (fingerprint images, face images, voice utterances, handwritten, hand images) and other personal data, including (name, contact details, age, sex, height, gender, wear glasses) will be collected, processed and used by the ENTERFACE Project no. 6 and 12, the data controller, in accordance with applicable laws.

I have been explained and informed that the collected database will be available to licensees worldwide, under restricted licensing conditions, for research purposes. Whenever the data acquired is transferred to licensees, only the biometric data will be released. The other personal data mentioned above will not be associated with the biometric data and will be kept confidential to the controller. In this way, the identities of the contributors will not be revealed to the licensees.

I have read the above and additional information provided to me and I understand that I am free to indicate whether or not I agree with the processing of my biometric and other personal data as described above. By signing the present form, I agree with the above stated.

Date :

Name :

Signature :

In case you don't agree that your biometric data can be published or displayed tick this box ☐