



Multi-PIE

Ralph Gross^{a,*}, Iain Matthews^a, Jeffrey Cohn^b, Takeo Kanade^a, Simon Baker^c

^aRobotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213, United States

^bDepartment of Psychology, University of Pittsburgh, United States

^cMicrosoft Research, Microsoft Corporation, One Microsoft way, Redmond, WA 98052, United States

ARTICLE INFO

Article history:

Received 16 March 2009

Received in revised form 13 July 2009

Accepted 4 August 2009

Keywords:

Face database

Face recognition across pose

Face recognition across illumination

Face recognition across expression

ABSTRACT

A close relationship exists between the advancement of face recognition algorithms and the availability of face databases varying factors that affect facial appearance in a controlled manner. The CMU PIE database has been very influential in advancing research in face recognition across pose and illumination. Despite its success the PIE database has several shortcomings: a limited number of subjects, a single recording session and only few expressions captured. To address these issues we collected the CMU Multi-PIE database. It contains 337 subjects, imaged under 15 view points and 19 illumination conditions in up to four recording sessions. In this paper we introduce the database and describe the recording procedure. We furthermore present results from baseline experiments using PCA and LDA classifiers to highlight similarities and differences between PIE and Multi-PIE.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Facial appearance varies significantly with a number of factors, including identity, illumination, pose, and expression. To support the development and comparative evaluation of face recognition algorithms, the availability of facial image data spanning conditions of interest in a carefully controlled manner is important. Several face databases have been collected over the last decade for this reason, such as the FERET [1], AR [2], XM2VTS [3], Cohn–Kanade [4], and Yale B [5] databases. See [6] for a more comprehensive overview.

To support research for face recognition across pose and illumination the Pose, Illumination, and Expression (PIE) database was collected at CMU in the fall of 2000 [7]. To date more than 450 copies of PIE have been distributed to researchers throughout the world. Despite its success the PIE database has a number of shortcomings; in particular it only contains 68 subjects that were recorded in a single session, displaying a small range of expressions (neutral, smile, blink, and talk).

To address these issues we collected the Multi-PIE database. This new database improves upon the PIE database in a number of categories as shown in Fig. 1 and Table 1. Most notably a substantially larger number of subjects were imaged (337 vs. only 68 in PIE) in up to four recording sessions. In addition the recording environment of the Multi-PIE database has been improved in com-

parison to the PIE collection through usage of a uniform, static background and live monitors showing subjects during the recording, allowing for constant control of the head position.

This paper gives an overview of the Multi-PIE database and provides results of baseline face recognition experiments. Section 2 describes the hardware setup used during the collection. Section 3 explains the recording procedure and shows example images. We provide statistics on recording session attendance and the subject population in Section 4. Section 5 shows results of evaluations using PCA [8] and LDA [9] in experiments comparing PIE and Multi-PIE as well as in experiments only possible on Multi-PIE.

2. Collection setup

This section describes the physical setup and the hardware used to record the high resolution still images (Section 2.1), the multi-pose/illumination images (Section 2.2), and the calibration data (Section 2.3).

2.1. High resolution images

We recorded frontal images using a Canon EOS 10D (6.3-megapixel CMOS camera) with a Macro Ring Lite MR-14EX ring flash. As shown in Fig. 2, subjects were seated in front of a blue background in close proximity to the camera. The resulting images are 3072×2048 in size with the inter-pupil distance of the subjects typically exceeding 400 pixels.

* Corresponding author. Address: Robotics Institute, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213, United States.

E-mail address: rgross@cs.cmu.edu (R. Gross).

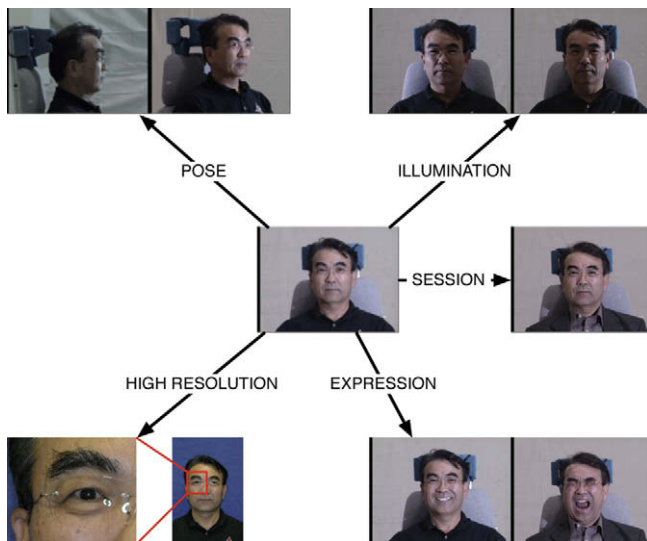


Fig. 1. Variation captured in the Multi-PIE face database.

Table 1
Comparison between the Multi-PIE and PIE databases.

	Multi-PIE	PIE
# Subjects	337	68
# Recording sessions	4	1
High resolution still images	Yes	No
Calibration images	Yes	No
# Expressions	6	4
# Cameras	15	13
# Flashes	18	21
Total # images	750,000+	41,000+
DB Size (GB)	305	40

2.2. Pose and illumination images

To systematically capture images with varying poses and illuminations we used a system of 15 cameras and 18 flashes connected to a set of Linux PCs. An additional computer was used as master to communicate with the independent recording clients running in parallel on the data capture PCs. This setup is similar to the one used for the CMU PIE database [7]. Thirteen cameras were located at head height, spaced in 15° intervals, and two additional cameras were located above the subject, simulating a typical surveillance-view. Fig. 3 illustrates the camera positions.¹ The majority of the cameras (11 out of 15) were produced by Sony, model DXC-9000, with the remaining four cameras (positions: 11_0, 08_1, 19_1, and 24_0) being Panasonic AW-E600Ps. Each camera had one flash (model: Minolta Auto 220X) attached to it; above for the 13 cameras located at head height and below for the two surveillance-view cameras. In addition, three more flashes were positioned above the subject between the surveillance-view cameras 08_1 and 19_1. See Fig. 4 for a panoramic image of the room with the locations of the cameras and flashes marked with yellow circles and white squares, respectively. All components of the system were hardware synchronized, replicating the system in [10]. All flashes were wired directly to a National Instruments digital I/O card (NI PCI-6503) and triggered in sync with the image capture. This setup was inspired by the system used in the Yale dome [5].

¹ The camera labels are derived from the names of the computers that they were attached to. We use the format (cc_h) with the camera number cc and the channel number h.



Fig. 2. Setup for the high resolution image capture. Subjects were seated in front of a blue background and recorded using a Canon EOS 10D camera with a Macro Ring Lite MR-14EX ring flash.

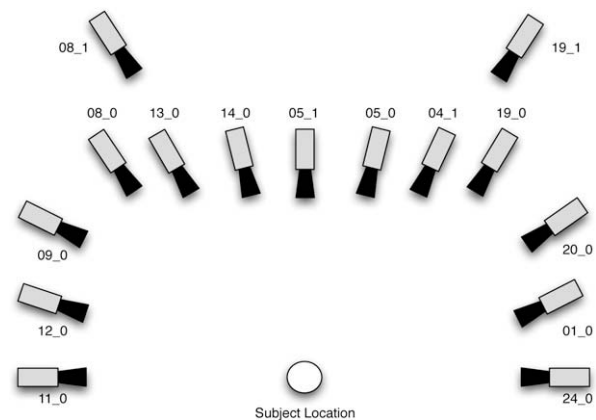


Fig. 3. Camera labels and approximate locations inside the collection room. There were 13 cameras located at head height, spaced in 15° intervals. Two additional cameras (08_1 and 19_1) were located above the subject, simulating a typical surveillance camera view. Each camera had one flash attached to it with three additional flashes being placed between cameras 08_1 and 19_1.

The settings for all cameras were manually adjusted so that the pixel value of the brightest pixel in an image recorded without flash illumination is around 128 to minimize the number of saturated pixels in the flash illuminated images. For the same reason we added diffusers in front of each flash. We also attempted to manually color-balance the cameras so that the resulting images look visually similar.

2.3. Calibration data

We recorded camera calibration images as well as color calibration images showing a GretagMacbeth ColorChecker chart. Calibration data was recorded after the conclusion of the data collection. During sessions 1–3, a number of flashes had to be replaced. As a consequence some camera and flash positions might be slightly different from what was measured during the collection of the calibration data. We furthermore determined the 3D locations of all cameras and flashes and the approximate location of the head of the subject using a theodolite (model: Leica TCA1105, see Fig. 5).

3. Data collection procedure

We recorded data during four sessions over the course of six months. During each session we recorded a single neutral high

Download English Version:

<https://daneshyari.com/en/article/526846>

Download Persian Version:

<https://daneshyari.com/article/526846>

[Daneshyari.com](https://daneshyari.com)