Image and Vision Computing 30 (2012) 713–727

Contents lists available at SciVerse ScienceDirect

Image and Vision Computing

journal homepage: www.elsevier.com/locate/imavis

Hi4D-ADSIP 3-D dynamic facial articulation database ☆

Bogdan J. Matuszewski [a,*], Wei Quan [a], Lik-Kwan Shark [a], Alison S. McLoughlin [b], Catherine E. Lightbody [b], Hedley C.A. Emsley [c], Caroline L. Watkins [b]

[a] Applied Digital Signal and Image Processing (ADSIP) Research Centre, School of Computing, Engineering and Physical Sciences, University of Central Lancashire; Preston PR1 2HE, United Kingdom
[b] School of Health, University of Central Lancashire, Preston PR1 2HE, United Kingdom
[c] Department of Neurology, Royal Preston Hospital, Preston PR2 9HT, United Kingdom

ARTICLE INFO

Article history:
Received 16 July 2011
Received in revised form 10 November 2011
Accepted 2 February 2012

Keywords:
Facial articulation database
Expression recognition
Facial
Dysfunctions
Facial expression validation

ABSTRACT

The face is an important medium used by humans to communicate, and facial articulation also reflects a person's emotional and awareness states, cognitive activity, personality or wellbeing. With the advances in 3-D imaging technology and ever increasing computing power, automatic analysis of facial articulation using 3-D sequences is becoming viable. This paper describes Hi4D-ADSIP – a comprehensive 3-D dynamic facial articulation database, containing scans with high spatial and temporal resolution. The database is designed not only to facilitate studies on facial expression analysis, but also to aid research into clinical diagnosis of facial dysfunctions. The database currently contains 3360 facial sequences captured from 80 healthy volunteers (control subjects) of various age, gender and ethnicity. The database has been validated using psychophysical experiments used to formally evaluate the accuracy of the recorded expressions. The results of baseline automatic facial expression recognition methods using Eigen- and Fisher-faces are also presented alongside some initial results obtained for clinical cases. This database is believed to be one of the most comprehensive repositories of facial 3-D dynamic articulations to date. The extension of this database is currently under construction aiming at building a comprehensive repository of representative facial dysfunctions exhibited by patients with stroke, Bell's palsy and Parkinson's disease.

1. Introduction

In the past two decades, automatic facial expression analysis and recognition have become an active research area in the computer vision community with significant efforts made in developing a wide range of applications, such as human–machine interaction [1], virtual reality [2], video conferencing [3] and biometrics [4].

Broadly speaking, algorithms developed for automatic facial data analysis can be categorised based on the modality of the data which can be 2-D or 3-D as well as static or dynamic. At present, the research concerned with facial articulations mostly utilises 2-D static images or 2-D dynamic videos due to their inherent acquisition simplicity. However, 2-D facial articulation analysis encounters difficulties in handling large amount of facial variations in images due to different head pose and illumination [5] and does not accurately reflect the 3-D nature of the face [6]. Using 3-D static data resolves some of these problems but lacks temporal cues. 3-D dynamic facial scans give the most accurate representation of facial articulations by including temporal information of dynamic facial movements, providing important cues at least for human observers, as confirmed by psychophysical experiments [7].

For the development and benchmarking of facial articulation analysis algorithms, a standardized testing resource is essential to enable meaningful comparisons. A large number of 2-D static and dynamic facial expression databases are currently available. Some of the most popular 2-D static databases include FERET [8], CMU-PIE [9], Multi-PIE [10], MMI [11], and JAFFE [12]. The commonly used 2-D dynamic databases are Cohn–Kanade AU-coded [13], MPI [14], DaFEx [15] and FG-NET [16]. Due to the requirement of a somewhat complex acquisition system, a rather limited number of 3-D static and dynamic databases are available to the research community. While the existing ones for the former include BU-3DFE [17], Bosphorus [18] and ZJU-3DFED [19], only a few have been recently developed for the latter, these are ADSIP [20], and BU-4DFE [21]. A list of available facial databases with their basic characteristics is given in Table 1. A more comprehensive review of these databases can be found in Refs. [22,23].

To support facial analysis research, Hi4D-ADSIP[1] has been created as a new comprehensive database of 3-D dynamic facial articulations [24]. This database currently contains 3360 3-D dynamic high-

☆ This paper has been recommended for acceptance by Lijun Yin.

* Corresponding author.
E-mail address: bmatuszewski1@uclan.ac.uk (B.J. Matuszewski).

[1] Hi4D-ADSIP stands for **Hi**gh Resolution **4** **D**imensional Database from the **A**pplied **D**igital **S**ignal and **I**mage **P**rocessing Research Centre.

0262-8856/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.imavis.2012.02.002

Table 1
Basic characteristics of the currently available and the newly created (bottom) facial articulation databases.

Databases	Format	Images/sequences	Expressions	Colour/gray	Resolution	Subjects	Year
FERET	2-D static	14,051	2	Gray	256×384	1199	1996
CMU-PIE	2-D static	41,367	4	Colour	384×286	68	2000
Multi-PIE	2-D static	750,000	4	Colour	3072×2048	337	2009
MMI	2-D static	200	6	Colour	720×576	52	2005
JAFFE	2-D static	213	7	Gray	256×256	10	1998
Cohn–Kanade	2-D dynamic	486	6	Gray	640×490	97	2000
MPI	2-D dynamic	60	4	Colour	450×400	8	2003
DaFEx	2-D dynamic	1,008	6	Colour	360×288	8	2005
FG-NET	2-D dynamic	399	6	Colour	320×240	18	2006
BU-3DFE	3-D static	2,500	7	Colour	1040×1329	100	2006
Bosphorus	3-D static	4,666	6	Colour	1600×1200	105	2008
ZJU-3DFE	3-D static	360	4	Colour	–	40	2006
ADSIP	3-D dynamic	210	7	Colour	601×549	10	2009
BU-4DFE	3-D dynamic	606	6	Colour	1024×681	101	2008
Hi4D-ADSIP	3-D dynamic	3360	14	Colour	2352×1728	80	2011

[1] The Hi4D-ADSIP database consists of seven basic facial expressions and further seven facial articulations.

resolution sequences from 80 subjects. It covers seven basic facial expressions shown in Fig. 1, as well as further seven facial articulations which are specifically included in the database to support the clinical research on facial dysfunctions. These additional articulations include mouth opening and closing, eyebrows raising and lowering and reading five predefined phrases. All these articulations are indexed in Table 2. The database has approximately 610,000 3-D individual face models. All the facial sequences have been captured at the

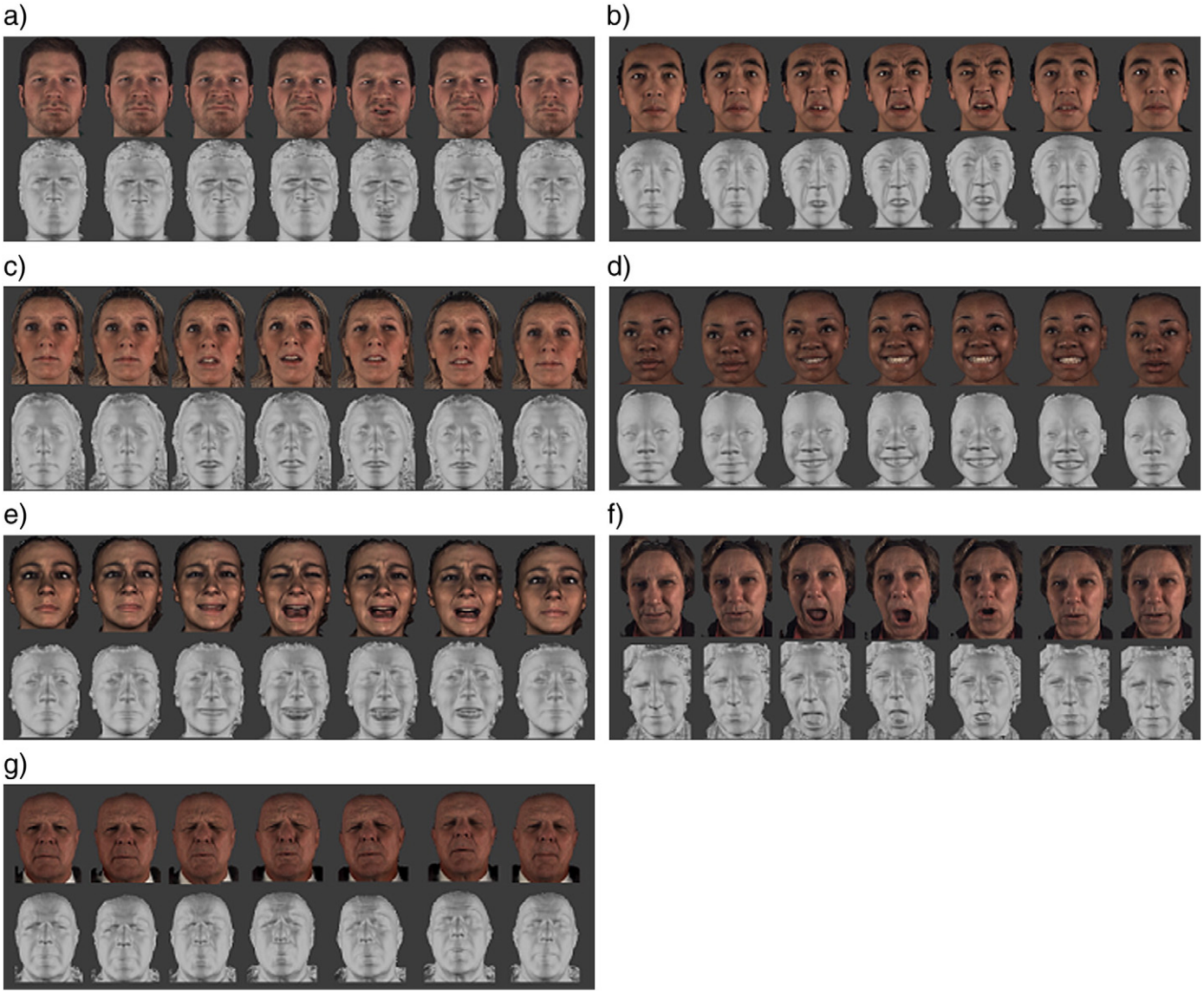

Fig. 1. Intermediate samples of facial expression: (a) anger, (b) disgust, (c) fear, (d) happiness, (e) sadness, (f) surprise, (g) pain.