Information Fusion 32 (2016) 64–79

Contents lists available at ScienceDirect

Information Fusion

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

gb2sμMOD: A MUltiMODal biometric video database using visible and IR light

Belén Ríos-Sánchez*, Miguel F. Arriaga-Gómez, Javier Guerra-Casanova, Daniel de Santos-Sierra, Ignacio de Mendizábal-Vázquez, Gonzalo Bailador, Carmen Sánchez-Ávila

Group of Biometrics, Biosignals and Security, Technical University of Madrid, Spain

ARTICLE INFO

Article history:
Available online 28 September 2015

Keywords:
Multimodal
Biometrics
Database
Multisensor
Video
Realistic Environments
Touch-less

ABSTRACT

In spite of recent efforts in gathering multimodal databases containing a big number of traits, a huge amount of users and covering multiple realistic scenarios, there is still a lack of touch-less realistic samples, video recordings for some traits and the use of infrared cameras which allows, among others, to avoid lighting influence and test recently appeared biometric techniques such as hand vein recognition. For this reason, a new realistic multimodal database composed of 8,160 hand, iris and face videos has been captured. To this end, a total of 60 contributors have participated in three separated acquisition sessions in which three different cameras have been used, covering different ranges of the light spectrum: visible light and two different infrared wavelengths. To simulate real-world working conditions, the database has been recorded in an indoor environment with different lightings and backgrounds.

In addition, due to the relevance of performing evaluation experiments in such a way that a reliable comparison of the results can be accomplished, an evaluation protocol is provided at the end of this paper. Moreover, performance results are provided for several biometric traits in mono- and multi- modalities that can be used as a baseline.

1. Introduction

Combining multiple biometric modalities permits taking advantage of the strengths of each single modality, compensating some of the limitations derived from their intrinsic nature, the maturity of the technology and the capturing conditions. The exploitation of the complementarity between biometric modalities increases the overall accuracy, improves the performance, reduces the vulnerability and therefore results in more robust biometric systems which offer enhanced security against forgeries [1].

One of the most important steps in the development of any system is the testing process, which is crucial to determine its performance and guarantee a suitable response when used under real working conditions. The success of multimodal biometric systems is thus directly related to the availability of realistic multimodal databases to evaluate new algorithms. This is a really challenging task due to the fact that the capture of multimodal databases requires an extra-effort in relation to unimodal databases. Adding to those well-known difficulties related to the recruitment of a large enough number of people willing to donate their biometric traits, to supervise the acquiring procedure to ensure that valid samples are captured and to correct possible errors, as well as the legal issues about the storage and distribution of personal information, an extra effort to develop a specific software which facilitates the data acquisition process must be considered. Furthermore, it is important to point out that, naturally, the time required per contributor significantly increases, making the process tedious and really resource and time consuming, even more so if the design of the database includes multiple sessions, multiple sensors, several samples per session and sensor, and changes in the capture conditions such as different backgrounds or lighting.

Owing to these constraints, a first approach based on the combination of different unimodal databases in heterogeneous datasets was initially carried out, assuming independence between different traits given a person [2]. Nevertheless, despite the initial acceptance in the literature, the usage of synthetically generated users, also known as “chimeric users”, was discussed and compared with the use of emerging multimodal biometric databases [3–5], concluding that independence assumption does not guarantee the same performance when using “chimeric” than true multimodal users and that the results strongly depend on which traits are selected. According to these studies, a relatively significant number of multimodal biometric databases have appeared in a recent past aimed to provide reliable testing frameworks [6–23].

* Corresponding author. Tel.: +34914524900; fax: +34913364501.
E-mail address: brios@cedint.upm.es (B. Ríos-Sánchez).

http://dx.doi.org/10.1016/j.inffus.2015.09.003

Table 1
Abbreviations used for traits and traits captured using multiple sensors mean: Fav: Face video, Fai: Face image, Sp: Speech, HG: Hand Gesture, Fp: Fingerprints, Sg: Handwritten Signature, Ha: Hand (geometry and palmprint), Hw: Handwritting, Ir: Iris, Hr: Hand Reverse, FV: Finger Veins, Ga: Gait, Ks: Keystroking. Multiscenario column refers to different physical locations. Varied background and lighting include uniform, controlled and uncontrolled cases.

	Users	Traits	Sessions	Multisensor	MultiScenario	Background	Lighting
DAVID [6]	120	Fav, Sp	2[a]			Varied	Neutral
XM2VTSDB [7]	295	Fav, Sp	4			Uniform	Neutral
BANCA [8]					Controlled	Uniform	Neutral
	208[b]	Fav, Sp	12	Fav, Sp	Degraded	Office-like	Controlled
					Adverse	Uncontrolled	Uncontrolled
SmartKom [9]	45	Fav, Sp, HG	2	Fav, Sp, HG	Public, Home, Mobile	Varied	Varied
MCYT [10]	330	FP, Sg	1	Fp			
BIOMET [11]	91[c]	Fav, Fai(2D[d], 3D), Sp, Fp, Ha, Sg	3	Fp		Uniform	Neutral
MYIDEA [12]					Controlled	Uniform	Neutral
	104	Fav, Sp, Fp, Sg, Hw, Ha	3	Fav, Sp, Fp	Degraded	Office-like	Controlled
					Adverse	Uncontrolled	Uncontrolled
Biosec [13]	200	Fp, Fai, Ir, Sp	2	Fp, Sp		Varied	Uncontrolled
PDA [14]	60	Fav, Sp, Sg	3		Indoor	Varied	Varied
					Outdoor	Uncontrolled	Varied
M3 [15]	39[e]	Fav, Fai, Sp, Fp	3	Fav, Sp	Indoor	Uniform	Neutral
					Outdoor	Uncontrolled	4 directions
MBioID [16]	120	Fai(2D,3D), Fp, Ir, Sg, Sp	2			Varied	Varied
BMEC [17]	480	Fai, Fav, Fp, Sg	2			Varied	Varied
MBGC [18]	5,000,000+	Fai, Fav, Ir	-	Fai, Fav, Ir	Still Face Challenge	Uniform	Controlled
					Portal Challenge	Indoor	Controlled
					Video Challenge	Outdoor	Uncontrolled
POLYBIO [19]	45	Fp, Fav, Fai, Hr	1			Uncontrolled	Uncontrolled
SDUMLA-HMT [20]	106	Fai, FV, Ga, Ir, Fp	1	Fp		Uniform	Varied
BiosecurID [21]	400	Fav, Fai, Sp, Ir, Sg, Hw, fp, Ha, Ks	4	Fp		Office-like	Neutral
BMDB [22]	971,667,713[f]	Fav, Fai, Sp, Sg, Fp, Ha, Ir	2	Fav, Sp, Fp	Internet	Office-like	Neutral
					Desktop	Office-like	
					Mobile Indoor	Office-like	
					Mobile Outdoor	Uncontrolled	
MMU GASPFA [23]	82	Ga, Sp, Fai	1	Ga		Uniform	Uncontrolled

[a] Second session only for 31 people
[b] Plus 120 additional impostors
[c] Common to every session. Session 1 had 130 people and session 2 106 people
[d] Infrared
[e] Plus 108 impostor speakers
[f] Corresponding to the Internet, Desktop and Mobile Scenarios

First biometric multimodal databases were only composed by two traits captured using a single sensor under totally controlled lighting and background conditions [6,7]. Over time, more complete databases have been generated, including a bigger number of traits (up to 8 [21]), and/or a great increase of contributors, up to more than 5,000,000, as in the case of the largest publicly available multimodal database [18]. This way, some datasets made use of several sensors to record some traits, specially Face, Speech and Fingerprints [8–13,15,18,20–23]. Finally, aimed to cover different cases of use, some databases have been recorded in multiple scenarios [8,9,12,14,15,18,22], which permit testing specific applications besides evaluating different environmental conditions. This is a really important aspect in a database due to the fact that results can vary widely between laboratory and real world conditions. Moreover, some of these databases are compatible in sensors, making it possible to combine portions of them to increase the number of subjects [10,12,13,21,22], or have common contributors, enabling to increase the number of sessions for these people to study long-term variability [13,21,22]. A comparison between these previous works is detailed in Table 1.

In spite of this, a lot of work in multibiometric algorithms and the databases used to test them remains to be done. In recent years there has been a growing trend toward touch-less biometrics, strongly influenced by the widespread use of mobile devices, which can be used in any situation, highlighting the need of touch-less realistic databases. In addition, the availability of databases including certain traits, such as hand images, is quite low if compared with datasets including popular biometric data, such as face or speech. The lack of hand images in multimodal databases is a significant limitation not only because it is an important, non-intrusive technique with high acceptance, but also because a single image provides information about geometry, palmprint and even veins if infrared cameras are used, being multibiometric itself. Furthermore, the lack of hand images captured using portable devices is total, as well as the availability of hand videos, really important in the development of more robust algorithms which combine several frames to obtain biometric patterns.

For this reason, this paper describes a new collaborative multimodal database mainly designed to overcome the absence of hand videos and the lack of infrared images. The gb2sμMOD database is composed of 8160 videos of faces, irises and hands belonging to 60 contributors recorded in three separated sessions under varied environmental conditions. Sessions 1 and 2 were recorded under quasi optimum conditions to allow a better evaluation of the algorithms, and session 3 under realistic extreme conditions to test the performance of the system in real-world cases. Different small, low-cost and portable cameras were used to acquire the images covering diverse ranges of the light spectrum: visible and infrared at two different wavelengths. In this way, it is possible to study the influence of light and other environmental conditions over the performance of unimodal and multimodal biometric systems. The variety of traits and sensors allow the evaluation of typical biometric combinations, like hand geometry and palmprint or hand geometry and hand veins, but also novel fusions like hand veins and face to analyse their viability.

The amount of video sequences per subject also allows us to design different experiments in order to customize algorithm unitary tests depending on the final application in a biometric system.