



## Facial Expression Recognition – A Comprehensive Review

<sup>1</sup>Ebenezer Owusu\*, <sup>2</sup>Ebenezer Komla Gavua, <sup>3</sup>Yong-Zhao Zhan

<sup>1,2</sup>Koforidua Polytechnic, Ghana, <sup>3</sup>Jiangsu University, China

\*[kayowusueb@yahoo.com](mailto:kayowusueb@yahoo.com)

### Abstract

In this paper, we have provided a comprehensive review of modern facial expression recognition system. The history of the technology as well as the current status in terms of accomplishments and challenges has been emphasized. First, we highlighted some modern applications of the technology. The best methods of face detection, an essential component of automatic facial expression system, are also discussed. Facial Action Coding Systems- the cumulative database of research and development of micro expressions within the behavioral science are also enlightened. Then various facial expression databases and the types of recognitions are explained in detail. Finally, we provided the procedures of facial expression recognition from feature extraction to classifications, emphasizing on modern and best approaches. Then the challenges encountered when comparing results with others are highlighted and suggestions to alleviate the problems, provided.

Keywords: FACS; Expression recognition; spatial; spatio-temporal; expression classification

### 1. Introduction

The studies of facial expressions originally emerged as physiognomy, which is the general assessment of a person's character or personality from his outer appearance, especially, from the face (Highfield et al., 2009). In a general sense the term refers to features of the face, when these features are used to infer the relatively enduring character or temperament of an individual. Most of these facial features have their basis in the bony structure of the skull, on which the soft tissues lie. These features include the shapes and positions of the major areas and landmarks of the face, such as the forehead, eyebrows, nose, cheeks, and mouth.

In China and other Asian cultures, formal systems of face reading techniques developed in the first millennium B.C.E., and it was integrated with religious beliefs such as Confucianism (Highfield et al., 2009). Substantial confidence in such methods developed in these cultures, and physiognomic inferences included descriptions of character, suitability for certain positions, and predictions about life and death. In Western cultures, the association of

facial features with a person's characteristic was first noted in the writings of the ancient Greeks in 4 BC. Much later, several pseudo-scientific and cultish movements exploited the inference of character from physiognomic features. As time expired, physiognomy studies slowly and imperceptibly vanished and replaced by facial expression recognition.

The study of facial expressions had its root in the 17th century, when Le Brun officially gave a series of presentations on the subject in 1667. In his posthumously published treatise, "Méthode pour Manuscript apprendre à dessiner les passions" in 1698, he promoted the expression of the emotions in painting; which later had great influence on art theory for two centuries. Much of Le Brun's idea is found in "The Perfect Imitation of Genuine Facial Expression" (Montagu, 1994). However, the modern day study of facial expression recognition started with Charles Darwin in the nineteenth century. Charles Darwin established the implications of facial expressions in humans and animals and introduced the deformation patterns to facial expression recognition (Darwin, 1998). "Among his many extraordinary contributions Darwin gathered evidence that some emotions have a

universal facial expression. He cited examples and published pictures suggesting that emotions are evident in other animals, and proposed principles explaining why particular expressions occur for particular emotions – general principles he applied to all animals” (Ekman & Darwin, 2003). In the 20th century, Ekman & Friesen (1978) developed the Facial Action Coding System (FACS) which is the most widely used and versatile method for measuring and describing facial behaviors.

Today, facial expression studies have attracted researchers from human computer interaction (HCI), computer vision and pattern recognitions. The technology is applied in a wide variety of contexts, including robotics, digital signs, mobile applications and medicine. It is reported that “some robots can operate by first recognizing expressions” (Bruce, 1993). In behavioral sciences and medicine for instance, expression recognition is effectively applied for intensive care monitoring Morik et al., (1999). Currently there are developing systems that are capable of making routine examinations of facial behavior during pain in clinical backgrounds. The observers have been trained to perform real-time measurement of pain expression during physical examinations of low-back pain patients. In infants the Neonatal Facial Coding System (NFCS) has been employed for real-time assessment within 32 to 33 week post-conceptional age infants who are undergoing a heel lance. The technology is being used in more advanced settings to reduce accidents through the implementation of automated detection of driver drowsiness in public transports. These systems relay intermittent information of the drivers’ facial expressions or emotional states to observers for effective surveillance and necessary awareness. Another important application is the invention of wearable appliance-based glasses which sense facial muscle movements and recognizes significant expressions such as confusion or interest (Scheirer et al., 1999). The wearable glass uses piezoelectric sensors hidden in a visor extension to a pair of glasses that provides compact, user control, and anonymity. Fig. 1 shows a picture of the expression glass.



**Fig. 1.** The expression glass

The rest of the work is organized as follows: Section 2 discusses the universality of facial expressions; section 3 talks about face detection and pose estimation; section 4 discusses facial action units and facial action coding system; section 5 gives a note of facial expression types; section 6 provides types of expression databases; section 7 presents feature extraction; section 8 talks about feature selection and classification; section 9 discusses results of expression recognition systems; finally section 10 concludes the work.

## **2. The universality of facial expression recognition**

The arguments of universal facial expressions across cultures can be traced as far back as about 150 years ago. In an attempt to prove that expressions are universal Darwin (1872) conducted a research across English citizens living in Africa, America, Australia, Borneo, China, India, Malaysia and New Zealand and concluded that facial expressions are universal. However lack of substantive evidences and intercultural mixture among the respondents made the results inconclusive by many eminent social psychologists such as Klineberg (1940).

The claims of Darwin were later resurrected by Tomkins (1962, 1963) in about a century later. Tomkins suggested that emotion was the basis of human motivation and that the seat of emotion was in the face. He conducted the first study to demonstrate that facial expressions were reliably associated with certain emotional states (Tomkins & McCarter, 1964). Then, Ekman et al., (1969) conducted a research across several cultures to investigate the subject in detail. The results of the study revealed universal agreement, both within and across cultures for six emotional expressions – anger, disgust, fear, happiness, sadness, and surprise. Thus, systematic evidence for the universality of emotions and their expressions was reestablished. Since the finding of

Ekman and his associates, there have been many studies to confirm the results. For instance, Matsumoto (2001) demonstrated convincingly that a small set of facial expressions are universally. Recent studies conducted by Matsumoto et al., (2008) have

demonstrated that when emotions are aroused, the same facial expressions of emotion are reliably produced by people all around the world and from all walks of life. Fig. 2 shows a model of the six universal facial expressions.

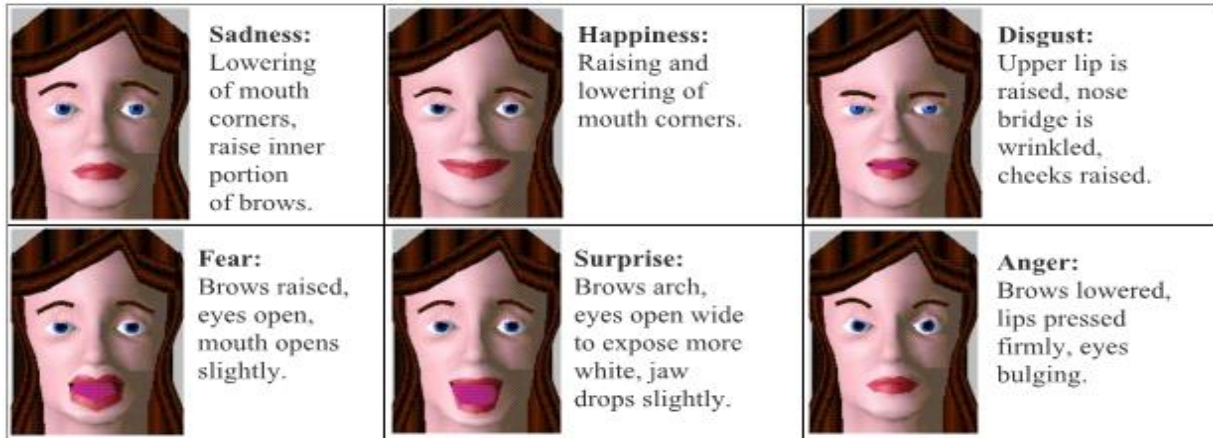


Fig. 2. A model of the six universal facial expressions.

### 3. Face detection and pose estimation

Face detection is a very important component in facial expression recognition systems. Faces in images must be detected before they are further processed for expression recognition. Face detection methods are grouped into feature-based and classification-based. The basic models of feature-based techniques largely lie on searching algorithms. The algorithms must effectively locate potential facial features such as eyes, nose and the mouth and classify them into faces according to their geometric relationships. As the performance of feature-based methods largely depends on the consistent location of facial features, its vulnerability to partial occlusion, unwarranted deformations, and inferior image quality is also devastating. The classification methods, however, utilize search windows over the input image and each local image in the window is classified as a face or non-face by a classifier. Classification methods have received considerable interest by many researchers because they are more effective, hence techniques such as support vector machine (SVM) (Heisele et al., 2003), neural networks (Bouzalmat et al., 2011; Khatun et al., 2011), local binary pattern (Lili et al., 2012) and Bayesian techniques (Liu, 2003) are utilized extensively.

### 4. Facial Action Units (FAUs) and Facial Action Coding System (FACS)

The human emotion is composed of thousands of expressions, although most of them differ in subtle changes of a few facial features. The spontaneous deformations of the facial muscles combine in groups to give a meaning to a particular facial expression. This involuntary or voluntary deformation of the facial muscles is referred in this sense as action units (AU). An AU represents the muscular activities that produce momentary changes in facial appearances. Stated another way, an action unit is a numeric code to describe the movements of facial muscles. The mapping between AUs and facial muscles is not necessarily 1:1; some AUs are composed of more than one muscle, and other AUs describe separate movements of the same muscle. Thus the term action unit is used because there is separation of more than one action from what most anatomists described as one muscle. The consistent exploration into action units (AUs) gave rise to Facial Action Units (FAU), which is a developed system to categorize the human facial expressions. The FAU system was first developed by Hjortsjo (1969) and later expounded by Ekman and Friesen (1978). Again years of continuing studies of AUs or FAUs resulted in the development of a database of AUs or FAUs known as Facial Action Coding System (FACS). The primary goal in developing the FACS was to develop a comprehensive

system which could distinguish all possible visually distinguishable facial movements. Scientifically, FACS is the cumulative database of research and development of micro-expressions within behavioral science. The last updated version of the system was in 2002 by Hagar, Ekman and Friesen (2002).

The FACS has become the most common manual for facial behavior analysis. Ekman and Friesen (1978) have defined 46 AUs to explain each autonomous face movement. Although they only defined a small number of idiosyncratic AUs, over 7000 dissimilar AU combinations have been observed so far (Ekman, 1982). FACS assesses all observable facial muscle movements, such as the mouth and eye movements, and not just those supposed to be relayed to emotion or any other condition. Recent studies conducted by Vick et al., (2007) show that FACS can be adapted to compare facial repertoires across similar species, such as humans and chimpanzees. According to the studies FACS can be modified by taking

differences in underlying morphology into account. “Such considerations enable a comparison of the FACS present in humans and chimpanzees, to show that the facial expressions of both species result from extremely notable appearance changes”. A cross-species analysis of facial expressions can help to answer the question of which emotions are uniquely human (Vick et al., 2007). The revised FACS coding manual affords a detailed description of all the appearance changes occurring with a given action unit. Fig. 3, Fig 4 and Fig. 5 show sample AUs. Fig. 5 shows sample combination of the 2002 FACS coding systems.








AU	Description	Facial Muscle	Image
1	Inner Brow Raiser	Frontalis, pars medialis	
2	Outer Brow Raiser	Frontalis, pars lateralis	
4	Brow Lowerer	Corrugator supercilii, Depressor supercilii	
5	Upper Lid Raiser	Levator palpebrae superioris	
6	Cheek Raiser	Orbicularis oculi, pars orbitalis	
7	Lid Tightener	Orbicularis oculi, pars palpebralis	
43	Eyes Closed	Relaxation of Levator palpebrae superioris; Orbicularis oculi, pars palpebralis	

Fig. 3. Upper Face Action Units







AU	Description	Facial Muscle	Image
10	Upper Lip Raiser	Levator labii superioris	
11	Nasolabial Depeener	Zygomaticus minor	
12	Lip Comer Puller	Zygomaticus major	
13	Cheek Puffer	Levator anguli oris (Caninus)	
18	Lip Puckerer	Incisivii labii superioris and Incisivii labii inferioris	
22	Lip Funneler	Orbicularis oris	

Fig. 4. Lower Face Action Units.











				
0	1+2	Video 1+2	4i	4ii
				
1+2+4i	1+2+4ii	j0	j1+4	j1+2+4

Fig. 5. Combinations of 2002 revised FACS.

### 5. Facial expression recognition types

A typical facial expression recognition system recognizes the expressions of the face, irrespective of the facial orientation. The earliest researches were focused on 2D orientations only and nearly all the accessible data collections of expressive faces were of restricted range full of only premeditated posed sentimental displays. These were mainly of the six

universal exemplary expressions consisting of anger, fear, disgust, surprise, sadness and happiness, and recorded under exceedingly constrained environment in terms of illumination and viewed angle. The problems with this single-viewed 2D analysis was that they could not totally exploit the information that is paraded by the face since 2D static image or 2D video recordings have problems in capturing the out-of-plane transformations of the facial surface. These

problems should not surprise anyone because the human face is 3D than 2D. Thus a purely 2D projection of the face is sensitive to changes in illumination and pose angle (Pantic & Rothkrantz, 2000). In the presence of these setbacks, the conditions that will make the 2D face data achieve a good performance in expression recognition are the abilities to normalize illumination and correct the head pose, which indeed is difficult to achieve (Quan et al., 2009). Interestingly, the mainstream of the current up-to-date facial expression recognition systems is footed on 2D facial images and videos, which only offer some appreciable accomplishment only for the data captured under the restricted environments already stated above. As a result, in recent times, there is a paradigm shift towards the application of 3D facial data since it is fundamentally invulnerable to changes in pose and illumination (Yin et al., 2006) and therefore generates better recognition accuracy. Again, since the facial data is the implicit supplier of information for the facial expression recognition assignment, it follows that the processing intricacy and the overall success of the built system soundly depends on the techniques which are employed to capture the data.

The current progress in structured light scanning, stereo photogrammetry and photometric stereo have made the acquisition of 3D facial structure and motion a viable task (Sandbach et al., 2012).

Currently we have a variety of gadgets and methods that are employed to capture 3D facial expression data. Samples include the use of single image reconstruction, structured light technologies, and stereo reconstruction algorithms. We have 3D imaging systems or scanners that can scan the expressive face and generate a geometric point cloud that corresponds to samples taken from the observed 3D facial data. Despite the fact that these scanners provide accurate surface measurements, they also require excessively long acquisition time. Non-contact scanners are much more suitable for 3D facial data acquisition. Again, notwithstanding these recent advancements, it must be stressed still that, the 3D face acquisition does not unravel all the problems; for example, it does not aid to alleviate the concerns associated with occlusions, where typical examples of facial occlusions include subjects wearing scuffs, make-ups, glasses, or having long hair that covers portions of the face. Thus occlusion is still a problem that must be unraveled to make facial expression

systems more attractive and dependable. For problems caused by pose variations, some researchers proposed that the use of multiple views of the face (Pantic & Rothkrantz, 2004) and deformable 3D models fitted on 2D images (Wen & Huang, 2003) or 3D images are tenable solution. No doubt, the success of the 3D face models may possibly advance view-independent facial expression recognition, which is very vital for spontaneous facial expression recognition, because the subject can be recorded in less controlled real-world scenery. Current attempts have also been directed to the recognition of multifaceted and impulsive emotional phenomena such as stress, frustration, depression and boredom rather than on the recognition of premeditated or the exhibited prototypical expressions of emotions (Gunes & Pantic, 2010; Zeng et al., 2009; Nicolaou et al., 2011; Vinciarelli et al., 2012).

## 6. Facial expression databases

One of the central contributors to the success of facial expression recognition researches is the development of 2D and recently 3D and 4D databases. The importance of the database is for the successful building of the system by aid of training, testing for the reliability and dependency of the system and for analysis and evaluation by means of making meaningful comparison with other systems. The choice of a particular database is based on the type of facial data used. The facial expression databases are cataloged into 4 main types - 2D static, 2D video, 3D static and 3D dynamic (also known as 4D). As facial expression recognition studies evolved from 2D datasets, the majority of present databases are found in this category. 3D and recently 4D datasets are not numerous. One would expect that, facial expression recognition; being such imperative research with immense applications, by now, would have had a de-facto standard dataset that would serve as a benchmark for measuring the success of all developed work; but till today, making such realization is just a dream. Failure to do so has left the comparisons of results a very difficult and challenging task. This problem is discussed in detail in section 9 of this work. Let us now move forward to discuss some of the influential databases that are publicly available for research. The descriptions are summarized in Table 1. The images of some of them are also displayed accordingly (See Fig. 6)

**Table 1**  
Description of face expression databases.

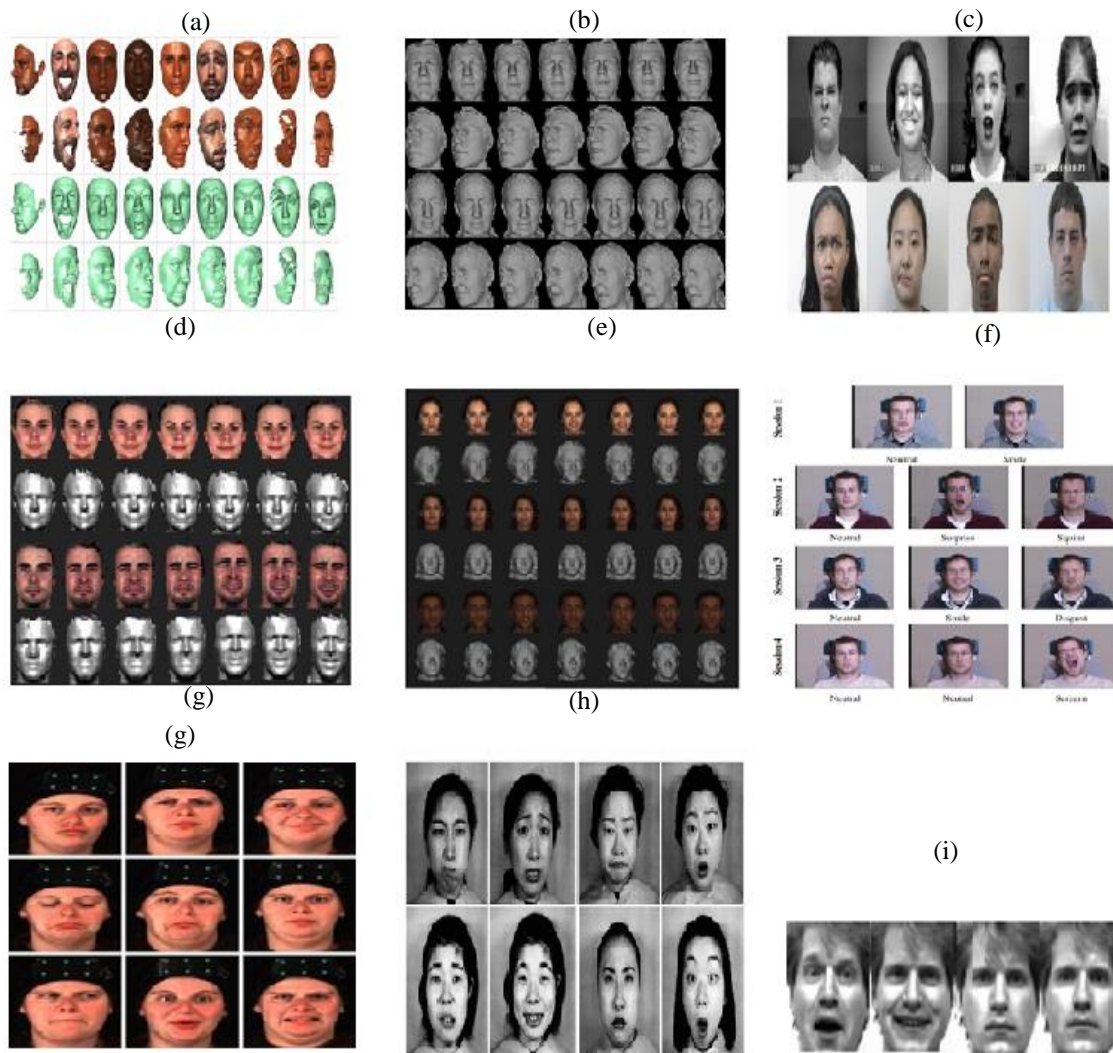
Database name	Type of database	Total images and subjects	Brief description
CMU-PIE	2D Pose (Static)	41368 images of 68 subjects	Each person has 13 different poses, 43 different illumination conditions, and 4 different expressions (Sim et al., 2002).
Multi-PIE	2D Pose	More than 750000 images of 337 subjects	Though the PIE database has been successful, it also has several deficiencies: small number of subjects, a single recording session and only few captured expression. To address these issues researchers at Carnegie Mellon University developed the Multi-PIE database. Images are captured under 15 view points and 19 illumination conditions in four recording sessions for the total number of the images (Gross et al., 2010; Lyons et al, 1999).
JAFFE	2D Pose (Static)	213 images of 10 subjects	The database has 7 facial expressions (6 basic facial expressions and 1 neutral) posed by Japanese female models. The database was assembled at the Psychology Department in Kyushu University (Lyons et al., 1999).
Yale (B)	2D Pose	165 grayscale images of 15 subjects	Grayscale images in GIF format. There are 11 images per subject, one per different facial expression: center-light, with glasses, happy, left-light, with no glasses, normal, right-light, sad, sleepy, surprised, and wink (Council et al., 2006).
Cohn-Kanade	2D video	500 image sequences from 100 subjects.	Very popular database. Available in 2 versions. Version 3 was planned to be released in spring 2013. They are AU coded database (Kanade et al., 2000). Subjects range in age from 18 to 30 years. 65% are female, 15% are African-American, and 3% are Asian or Latino. Each subject performs a series of 23 facial displays that included single action units and combinations of action units. Image sequences from neutral to target display were digitized into 640×480 or 490 pixel arrays with 8-bit precision for grayscale values. For a full description of version 2, also known as CK+, see (Lucey et al., 2010).
MPI	2D video	55 images of 19 subjects	MPI database was developed at the Max Planck Institute for Biological Cybernetic (Pilz et al., 2006). The database contains video sequences of four different expressions: anger, disgust, surprise and gratefulness. Each expression was record from five different views concurrently. All facial expressions are available in three repetitions, in two intensities, as well as from three different camera angles (Kaulard et al., 2012).
BU-3DFE	3D static	100 subjects	The database was developed at the Binghamton University for the purpose of 3D facial expression analysis (Quan et al., 2009). It contains 100 subjects, with ages within 18 to 70 years. It has a diversity of ethnic origins-Whites, Blacks, East-Asians, Middle-East Asians, Indians and Hispanics. Each subject displayed seven expressions, that consist of neutral and six prototypical facial expressions at four intensity levels. There are 25 3D facial scans containing different expressions for each subject. The total of the facial scans are 2,500. Each 3D facial scan in the database contains 13,000 to 21,000 polygons with 8,711 to 9,325 vertices (Sandbach et al., 2012).

**Table 1**

Description of face expression databases continued.

Database name	Type of database	Total images and subjects	Brief description
The EURECOM Kinect Face Dataset	3D	52 Subjects (14 females, 38 males)	The datasets are obtained with a Kinect sensor and captured in two sessions at different intervals of about 2 weeks. In each session, 9 facial images are collected from each person under different facial expressions, lighting and occlusion conditions: neutral, smile, open mouth, left profile, right profile, occluded eyes, occluded mouth, and side occlusion with a sheet of paper under light (Huynh, Min, & Dugelay, 2012).
Bosphorus	3D static	4666 images of 105 subjects	Images are captured with Inspeck Mega Capturor II 3D (Sandbach et al., 2010). The subjects consist of 60 men and 5 women, with the majority of the subjects being Caucasian, 27 of whom were professional actors, in various poses, expressions and occlusion conditions. Has rich repertoire of expressions: Up to 35 expressions per subject, FACS scoring includes intensity and asymmetry codes for each AU. There are systematic head poses (13 yaw and pitch rotations) and different kinds of face occlusions (beard and moustache, hair, hand, eyeglasses) (Savran, Sankur, & Bilge, 2010). Each subject expresses each of the 6 prototypical facial expressions and up to 24 AUs. The texture images are of resolution 1600×1200 pixels while the 3D faces consist of approximately 35000 vertices (Sandbach et al., 2012).
BU-4DFE	4D (3D dynamic)	606 images of 101 subjects	This database is an extension of BU-3DFE. It contains sequences of images captured at 25 frames per second (fps) of the six prototypical facial expressions with their temporal segments (onset, apex and offset) with each sequence lasting about 4s. (Sandbach et al., 2012; Matuszewski et al., 2011). The database provides no AU annotation.
ADSIP	4D (3D dynamic)	First edition contains 210 images of 10 subjects. Final edition will have 100 more added subjects	The first version (ADSIPmark1) was released in 2008. The participants were graduates from the School of Performing Arts in University of Central Lancashire (Frowd, 2009). Each subject displayed seven expressions: anger, disgust, happiness, fear, sadness, surprise and pain, at three intensity levels (mild, normal and extreme). Each sequence was captured at 24 fps and lasted within 3s. The final objective of the ADSIP database is to contain 3D dynamic facial data of over 100 control subjects and another 100 subjects with different facial dysfunctions. (Quan et al., 2010a).





**Fig. 6.** Sample facial expression databases: (a) Bosphorus 3D; (b) ADSIP final version with anger; (c) CK+ - The top images are from original CK database and the bottom ones are the representatives of the extended data; (d) ADSIPmark1; (e) BU-4DFE; (f) Multi-PIE; (g) MPI; (h) JAFFE; (i) Yale

### 7. Feature Extraction

Extraction of the expressive facial feature is an important step after the face is detected. Though there are many options, the choice of the method depends on the face representation and the kind of the input image – whether static or dynamic.

The facial representation features are classified into geometric and appearance. The geometric features embody the permanent features which in this

regard represent the shape as well as the locations of facial parts; the most significant ones are the eyes, eyebrows, nose and the mouth. The muscles round the eyes and the mouth are the major indicators for action unit recognition (Cohn & Zlochow,1995) The essential task involves accurate extraction of these features to represent the face geometry. The appearance features are rather the transient features of the face which constitutes the

skin texture and in effect, provides critical information for the recognition of certain AUs. Such features include wrinkles, bulges and furrows, which are extracted through the application of image filters to either the entire face or some explicit sections. Based on these two facial feature representation, in broad terms, there are two main feature extraction methods; holistic approaches where the face is processed as one piece, and local approaches, where attention is given to only a set of specific facial features or sections according to the targeted AU. Holistic feature extraction is the most extensively used technique. For this approach, each pixel of an image is considered valuable information. Many methods have been applied successfully with this approach. For instance, [Lanitis et al., 1997] presented an Active Appearance Model that was powered by principal component analysis (PCA) to locate facial features and recover 3D pose as well. Bartlett et al., (1997) convolve the entire face images by a set of Gabor wavelet kernels; the resulting Gabor wavelet magnitudes response was used as the input to a recognition engine. Koelstra et al., (2010) proposed an appearance dynamic model that detects AUs and the time phases; onset, apex and offset by using Free Form Deformations and Motion History Images as descriptors. The local feature class has also been very successful too. Valstar and Pantic (2011) utilized 20 facial points and applied a facial point tracker to track sparse set of facial points.

From the tracked points both static and dynamic features were computed to detect the temporal phases - onset, apex, and offset. Kakumanu employed local graph to track facial features (Kakumanu & Bourbakis, 2006). Chang proposed a method of toning the overlapping regions around the nose (Chang et al., 2006). Gundimada and Asari, (2009) extracted the local facial features by means of modular kernel eigen-spaces for multi-dimensional spaces. A third approach which is a hybrid of holistic and local has been proposed. Kimura and Yachida (1997) used Potential Net to fit a normalized image and computed the edges by utilizing differential filter via Gaussian filters. Tian et al., (2002, 2005) has indicated that the hybrid approach performs better than any of the traditional methods. Extensive discussion of the feature extraction method is beyond the scope of this material. For detail discussions, we refer interested readers to (Pantic & Rothkrantz, 2000; Shan, 2010; Tian et al., 2001; De la Torre & Cohn, 2011).

## 8. Expression feature selection and classification

The last step of expression recognition systems is classification. The expression classifier is trained with the extracted features to recognize expressions of unknown datasets also known as test datasets.

### 8.1 Feature selection

Sometimes it is necessary to select a few representational set of the extracted facial features to avoid misclassification. For instance Gabor and Haar wavelets are unnecessarily huge in number and contain redundant data that can inhibit their practical implementation. This is a potential cause to misclassification in training. Such huge features need to be processed by a feature selection tool that selects fewer but substantial representations to reduce feature dimensions to the classifier for processing. Quite a few feature selection algorithms have been developed. For instance, the AdaBoost algorithm originally proposed by Freund and Schapire (1995) has been very successful to select extracted facial features for classification (Shen & Bai, 2006; Sandbach et al., 2011; Zhou et al., 2006). The boosting algorithm boosts the performance of weak binary classifiers by strengthening training on misclassified sets. The scheme of the algorithm is to weight a set of weak classifiers in respect to a function of the classification error (Freund & Schapire, 1995). The final strong classifier  $H$  is a weighted combination of weak classifiers  $h_i$  followed by a threshold and it is denoted by:

$$h(x) = \begin{cases} 1: \sum_{i=1}^t \alpha_i h_i \geq \text{threshold} \\ 0: \text{otherwise} \end{cases}$$

where, the binary digits 1 and 0 represents the respective classified and misclassified sets. The principal component analysis (PCA) method is a traditional method that has also been successfully utilized for dimensionality reduction in a number of facial expression applications (Soyel & Demirel, 2009, 2010).

The selection system consists of searching amongst a certain number  $n$  of principal components,  $m$  such that ( $n < m$ ) are the best discriminant for the facial image expression recognition. Normally an iterative procedure is employed to select the components stepwise to assemble the optimal

components. The Linear Discriminant Analysis (LDA) (Soyel et al., 2010) has also been utilized to calculate the optimal subspace to create a certain number of dimensional discriminant subspaces. Under this procedure the LDA explores for the vectors in the core subspace that best categorize among sets. Lastly, Sha et al., (2011) proposed the normalized cutting based filter (NCBF) to select optimum features headed for pattern classification. The NCBF composes of two main parts; it works on the principle that the features with higher discriminative aptitude have stronger correlation with each other.

### 8.2 Expression classification

Classification is the final essential component of expression recognition system. The classifier has to be trained to recognize expressions of unknown datasets. The training process depends on the type of classifier and the type of the input image. For instance both spatial and special-temporal classification approaches have been employed to identify AUs for expression recognition.

The philosophy of a spatial classification is to locate the units on a spatial network and provide concurrent set of structured classes of these units which are compatible with the network. One of the traditional classifiers employed is discriminant analysis. For instance, Cohn et al., (2004) applied this approach for automatic facial expression recognition of video images. Their system focused on two action units; brow raising and brow lowering. The system achieved a recognition rate of 89% for two-state recognition and 76% for three-state recognition. Independent Component Analysis (ICA) (Bartlett et al., 2002) has also been used to recognize both facial characteristics and expression constructions. Neural network methods have also been very successful in expression recognitions (Bazzo & Lamar, 2004; Tian et. 2002; Ma & Khorasani, 2004). Ma and Khorasani (2004) employed DCT over the whole face image as a feature detector and used a neural network of one-hidden layer feed-forward neural network as an expression classifier for five expressions - smile, anger, sadness, and surprise. The best recognition rate on their system was 100% and 93.75% (without rejection). The problem with neural network classifiers is that, they find it rather difficult to train unconstrained facial behavior where there might be several thousands of AU combinations. The SVM

classifier is very popular in expression recognition systems. They have been very successful in recognition of several AUs (Gundimada & Asari, 2009; Valstar et al., 2005). Shan et al., (2009) extracted facial features using Gabor filters and boosted the features with AdaBoost technique; the selected features were input into LDA and SVM. The recognition in the SVM classifier showed superior performance.

The spatio-temporal methods have also been successful, though not as popular as the spatial methods. The commonest classification among these approaches is Hidden Markov Models (HMM). The power of HMM lies in their ability to model facial actions' dynamics (Otsuka & Ohya, 1998; Cohn et al., 1997). The HMM classification is executed by selecting AU or combination of AUs, that exploits the maximum likelihood of the extracted facial features. Outside the set of spatial and spatio-temporal classifiers, there are other ones which have also emerged strongly. For example, Bayesian Networks (BN) is a strong a force to be reckoned (Cohen et al., 2003; Zhang & Ji, 2005). Cohen et al., (2004) proposed a stochastic structure search (SSS) algorithm to train a BN classifier that recognizes facial expression from labeled and unlabeled datasets.

### 9. Comparing results of expression recognition systems

One disturbing problem that researchers of facial expression recognitions face is an attempt to make realistic comparisons of performances of proposed methods with the existing ones. There is lack of common benchmark to compare results hence various research groups use their own methods, but in doing so, cannot make direct comparisons since the choice of the methods and databases differs from one research group to the other (Shan et al., 2009; Lee, Huang & Shih, 2010). After all, upon what bases can one compare results if there are dissimilarities of testing? So far what researchers do is the traditional comparisons of just average results of different methods (see Table 2), or comparisons of performances of different methods the same group have performed with same datasets under similar conditions (see Table 3). As fellow researchers of the field, we think modern research groups can do better by publishing the results of their replicates such as detailed by Tong et al., (2007) so that other

researchers, when comparing performances can go a step further to conduct significant statistic testing techniques.

The use of statistical tests to assess the conclusions drawn in an experimental study is becoming a must both in the pattern recognition and soft computing communities. Researchers can

compare their results by using significant testing techniques such as the Wilcoxon test or paired T-Test.

**Table 2**

Comparing facial expression results of different methods

Author (s)	Database	Average Recognition rate (%)	Subjects	Expressions	Classifier/Method
Bindu, Gupta and Tiway (2007)	Cohn-Kanade	85.7	80 training set, 40 test set	Happy, anger, disgust, fear, neutral, sad, joy, surprise	Neural Network
Kotsia and Pitas (2007)	Cohn-Kanade	99.7	Whole database	Recognizes either the six basic facial expressions or a set of chosen AUs.	Multiclass SVM
Dornaika and Davoine (2008)	Non-commercial/Unknown	Not reported	Several video sequences	Six AUs: lower lip depressor, lip stretcher, lip corner depressor, upper lip raiser, eyebrow lowerer and outer eyebrow raiser. Surprise, joy disgust, anger, fear	Candide 3D face model and Second order Markov chains

**Table 2 continued.**

Author (s)	Database	Average Recognition rate (%)	Subjects	Expressions	Classifier/Method
Martin, Werner, and Gross (2008)	FEEDTUM mimic database	Lowest is 10, highest is 94.9	1438	Anger, happy, neutral, disgust, sad, surprise, fear, surprise	MLP and SVM. Highest detection recorded with SVM.
Soyel and Demirel, (2010)	BU-3DFE	93.72	420 with training set, 336 and test set 84	Neutral, anger, fear, happiness, surprise, disgust, sadness,	Probabilistic neural network (PNN)
Chin and Kim [90]	Unknown	On average, surprise receives 83, fear 55, disgust 72, anger 87, happiness 80, and sadness 89	50	Anger, surprise, sadness, fear, happiness, disgust, happiness	low-polygonal-based applications
Yun and Guan, (2010)	BU-3DFE	85.39	840 images, 560 for training, 280 for testing	Neutral, happy, fear, anger, surprise, disgust, sad.	improved kernel canonical correlation analysis (KCCA)
Valstar et al., ((2011)	GEMEP-FERA 2011 DATASET	Average value not reported.	1260 images. 158 portrayals (87 for training and 71 for testing) were selected for the AU sub-challenge. For emotion sub-challenge, a total of 289 portrayals were selected (155 for training and 134 for testing).	Several videos of AUs for anger, fear, joy, relief, sadness.	Local Binary Pattern features, Principal Component Analysis, and Support Vector Machine
Xiao, et al., (2011)	Cohn-Kanade and Feedtum databases.	On Cohn-Kanade database, the best performance of 89.35, 96.57 and 95.38 are achieved. For Feedtum database, over 70 is achieved.	1000 datasets.	Anger, happiness, surprise, disgust, sadness, fear, sadness,	Multi-manifold based classification, nearest neighbor classifier (NN) and support vector machine (SVM)

**Table 3**

Comparing performances of several methods for experiments performed with same datasets and under similar conditions.

Feature extraction methods	Recognition rates (%)		Testing speed per image (seconds)
	Cross Validation	Leave-one-out	
LDA+SVM	91.27	91.90	0.0367
2D-PCA+SVM	92.06	93.33	0.0357
ICA+SVM	93.35	93.81	0.0359
PCA+SVM	93.43	94.84	0.0353
2D-LDA+SVM	94.13	95.71	0.0357

Performance comparisons of facial expression recognition in JAFFE database, (Shih et al., 2008)

**Table 4**

Comparing performances of experiments performed under same conditions but different classifiers.

Expression	2D+3D	2D
Neutral	95.42	83.60
Disgust	82.63	70.72
Happy	90.84	81.21
Sad	73.97	61.85
Surprise	87.60	79.75

Tsalakanidou and Malassiotis (2010): Facial expression recognition rates (%) obtained for the proposed 2D and 3D classifier and the 2D appearance-based classifier under the per frame classification scenario.

## 10. Conclusion

Facial expression recognition has long history within the ancient studies of physiognomy as far as in the first millennia B.C.E. Years of successful studies yielded to a common agreement that expressions are universal across cultures. Since the facial data is the implicit supplier of information for the facial expression recognition assignment, the processing intricacy and the overall success of the automatic expression recognition system depends on the techniques which are employed to capture the data.

Face detection, which is an early step, used to have many challenges but today, the onset of good algorithms have helped improved the process extremely. Extraction of the expressive facial feature is an important step after the face is detected. The last component of expression recognition system is classification. The classifier has to be trained to recognize expressions of unknown datasets. The most successful classifier is SVM, when combined with AdaBoost as feature selection tool. The development of 2D, 3D and 4D databases and Facial Action Coding System (FACS) have also contributed immensely to the success of expression recognition.

Though a lot of the early problems faced with 2D data have been solved with the onset of 3D and 4D data acquisition, the 3D and 4D face acquisition does not unravel all the problems; for example facial occlusions, including subjects wearing scuffs, make-ups, glasses, or long hair that covers portions of the face are still major problems yet to be solved. For

problems caused by pose variations, it is proposed that the use of multiple views of the face and deformable 3D models fitted on 2D or 3D images are tenable solutions.

Finally, lack of standardize benchmark to compare results is a major challenge. Since various research groups use different methods, making direct comparisons with other findings is impossible. As a step to standardizations, we suggest that various research groups should publish their replicate results so as to make it possible for others to make significant statistical comparatives with theirs. The use of statistical tests to assess the conclusions drawn in an experimental study is becoming a must both in the pattern recognition and soft computing communities. Researchers can compare their results by using significant testing techniques such as the Wilcoxon test or paired T-Test.

## Acknowledgment

This paper is supported by the National Nature Science Foundation of China (No. 61272211, 61170126), the Natural Science Foundation of Jiangsu Province (No. BK2011521), and the Research Foundation for Talented Scholars of Jiangsu University (No. 10JDG065).

## References

- Bartlett, M. S., Movellan, J. R., & Sejnowski, T. J. (2002). Face recognition by independent component analysis. *IEEE Transactions on Neural Networks* 13, 1450-1464.
- Bartlett, M. S., Littlewort, G., Frank, M. G., Lainscsek C., Fasel, I., & Movellan, J. R. (2005). Recognizing Facial Expression: Machine Learning and Application to Spontaneous Behavior, in: *Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition*, pp. 568-573.
- Bazzo, J. J., & Lamar, M. V. (2004). Recognizing Facial Actions Using Gabor Wavelets with

- Neutral Face Average Difference, in: *Proc. Sixth IEEE Int'l Conf. Automatic Face and Gesture Recognition*, pp. 505-510.
- Bindu, M. H., Gupta, P., & Tiwary, U. S. (2007). Cognitive model-based emotion recognition from facial expressions for live human computer interaction, in: *IEEE Symposium on Computational Intelligence in Image and Signal Processing*, pp. 351-356.
- Bouzalmat, A., Belghini, N., Zarghili, A., Kharroubi, J., & Majda, A. (2001). *Face Recognition Using Neural Network Based Fourier Gabor Filters and Random Projection*, *International Journal of Computer Science and Security* 5, 376.
- Bruce, V. (1993). What the human face tells the human mind: Some challenges for the robot-human interface, *Advanced Robotics* 8, 341-355.
- Chang, K. I., Bowyer, W., & Flynn, P. J. (2006). Multiple Nose Region Matching for 3D Face Recognition under Varying Facial Expression, *IEEE Trans on Pattern Analysis and Machine Intelligence* 28(2006) 1695-1700.
- Chin, S., & Kim, K. Y. (2009). Emotional Intensity-Based Facial Expression Cloning for Low Polygonal Applications, *IEEE Trans on Systems, man, and Cybernetics, Part C: Applications and Reviews* 39, 315 -330.
- Cohn, J. F., & Zlochower, A. (1995). A Computerized Analysis of Facial Expression: Feasibility of Automated Discrimination, *Am. Psychological Soc., NY*.
- Cohn, J. F., Zlochower, A., Lien, J., Wu, Y., & Kanade, T. (1997). Automated face coding: a computer-vision based method of facial expression analysis, *Seventh European Conference on Facial Expression Measurement and Meaning, Salzburg, Austria*, pp. 329-333.
- Cohn, J. F., Reed, L. I., Ambadar, Z., Xiao, J., & Moriyama, T. (2004). Automatic Analysis and Recognition of Brow Actions and Head Motion in Spontaneous Facial Behavior, in: *Proc. IEEE Int'l Conf. Systems, Man, and Cybernetics*, pp. 610-616.
- Cohen, I., Sebe, N., Garg, A., Chen, L., & Huang, T. (2003). Facial Expression Recognition from Video Sequences: *Temporal and Static Modeling*, *Computer Vision and Image Understanding* 91, 160-187.
- Cohen, I., Cozman, F. G., Sebe, N., Cirelo, M. C., & Huang, T. S. (2004). Semisupervised Learning of Classifiers: Theory, Algorithms, and Their Application to Human-Computer Interaction, *IEEE Trans. Pattern Analysis and Machine Intelligence* 26 pp. 1553-1567.
- Darwin, C. (1998). Expression of the Emotions in Man and Animals, *Oxford University Press, USA*.
- De la Torre, F., & Cohn, J. F. (2011). Facial expression analysis, in: *Visual Analysis of Humans*, Springer, London, pp. 377-409.
- Dornaika, F., & Davoine, F. (2008). Simultaneous Facial Action Tracking and Expression Recognition in the Presence of Head Motion, *Int. J. Computer Vision* 76, 257-281.
- Ekman, P., & Darwin, C. (2003). Deception, and Facial Expression, *Annals New York Academy of Sciences*, 1000, 205-221.
- Ekman, P., & Friesen, W. V. (1978). Facial action coding system: A technique for the measurement of facial movement. Palo Alto, CA: *Consulting Psychologists Press. P. C. Ellsworth, C. A. Smith, from appraisal to emotion: Differences among unpleasant feelings. Motivation and Emotion*, 12, 271-302.
- Ekman, P., Sorenson, E. R., & Friesen, W. V. (1969). Pan-cultural elements in facial displays of emotion, *Science* 164, 86-88.
- Ekman, P. (1982). Methods for measuring facial action, *Handbook of methods in nonverbal behavior research*, pp. 45-90.
- Frowd, C. D., Matuszewski, B. J., Shark, L. K., & Quan, W. (2009). Towards a Comprehensive 3D Dynamic Facial Expression Database, in: I. Rudas, N. Mastorakis (Eds.), *WSEAS International Conference Proceedings, Recent Advances in Computer Engineering*.
- Freund, Y., Schapire, R. (1995). A decision-theoretic generalization of on-line learning and an application to boosting, *Computational Learning Theory: Eurocolt*, pp. 23-37.
- Gross, R., Matthews, I., Cohn, J., Kanade, T., & Baker, S. (2010). *Multi-PIE*, *Image and Vision Computing*, 28, 807-813.
- Gundimada, S., & Asari, V. (2009). Facial Recognition Using Multisensor Images Based on Localized Kernel Eigen Spaces, *IEEE Trans on Image Processing*, 18, 1314-1325.
- Gunes, H., & Pantic, M. (2010). Automatic, dimensional and continuous emotion recognition, *Int. J. Synthet. Emot.* 1, 68-99.

- Hager, J. C., Ekman, P., & Friesen, W. V. (2002). Facial action coding system. *Salt Lake City, UT: A Human Face*. ISBN 0-931835-01-1.
- Heisele, B., Serre, T., Prentice, S., & Poggio, T. (2003). Hierarchical classification and feature reduction for fast face detection with support vector machines, *Pattern Recognition*, 36, 2007-2017.
- Highfield, R., Wiseman, R., & Jenkins, R. (2009). How your looks betray your personality, *New Scientist*.
- Hjortsjö, C. H. (1969). Man's faced and mimic language, *Studen litteratur*.
- Huynh, T., Min, R., & Dugelay, J. L. (2012). An Efficient LBP-based Descriptor for Facial Depth Images applied to Gender Recognition using RGB-D Face Data, in: *ACCV Workshop on Computer Vision with Local Binary Pattern Variants, Daejeon, Korea*.
- Kanade, T., Cohn, J. F., & Tian, Y. (2000). Comprehensive database for facial expression analysis, in: *Proceedings of Automatic Face and Gesture Recognition*, pp. 46-53.
- Kakumanu, P., & Bourbakis, N. (2006). A Local-Global Graph Approach for Facial Expression Recognition, *Conference on Tools with Artificial Intelligence (ICTAI'06)*, pp. 685-692.
- Kaulard, K., Cunningham, D. W., Bulthoff, H. H., & Wallraven, C. (2012). The MPI Facial Expression Database - A Validated Database of Emotional and Conversational Facial Expressions, *PLoS one*, 7, e32321.
- Khatun A., & Bhuiyan, M. A. A. (2011). Neural Network based Face Recognition with Gabor Filters, *Int. J. Comput. Sci. Netw. Secur.* 11, 71-74.
- Kimura, S., & Yachida, M. (1997). Facial expression recognition and its degree estimation, in: *Proc. Computer Vision and Pattern Recognition*, pp. 295-300.
- Klineberg, O. (1940). *Social Psychology*, New York: Holt.
- Koelstra, S., Pantic, M., & Patras, I. (2010). A dynamic texture based approach to recognition of facial actions and their temporal models. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 32, 1940-1954.
- Kotsia, I., & Pitas, I. (2007). Facial Expression Recognition in Image Sequences Using Geometric Deformation Features and Support Vector Machines, *IEEE Trans. Image Processing* 16, 172-187.
- Lanitis, A., Taylor, C. J., & Cootes, T. F. (1997). Automatic Interpretation and Coding of Face Images Using Flexible Models, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 19, 743 - 756.
- Lee, C. C., Huang, S. S., & Shih C. Y. (2010). Facial Affect Recognition Using Regularized Discriminant Analysis Based Algorithms, *EURASIP Journal on Advances in Signal Processing*, 1:96.
- Li, H., Councill, I., Lee, W. C., & Giles, C. L. (2006, May). CiteSeerx: an architecture and web service design for an academic document search engine. In *Proceedings of the 15th international conference on World Wide Web* (pp. 883-884). ACM.
- Kanade, T., Cohn, J. F., & Tian, Y. (2000). *Comprehensive database for facial expression analysis*, in: *Proceedings of Automatic Face and Gesture Recognition*, pp. 46-53.
- Lili, L., Xiangbiao, L., & Wenhui, Z. (2012). Face Detection Algorithm Based on Haar and Local Binary Pattern Features, *JCIT* 7, 602-610.
- Liu, C., (2003). Bayesian discriminating features methods for face detection, *IEEE Trans. Pattern Analysis and Machine Intelligence* 25, 725-740.
- Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., & Matthews, I. (2010). The Extended Cohn-Kande dataset (CK+): A complete facial expression dataset for action unit and emotion -specified expression, in: *Computer Vision and Pattern Recognition Workshops (CVPRW)*, *IEEE Computer Society Conference*, pp. 94-101.
- Lyons, M. J., Budynek, J., & Akamatsu, S. (1999). Automatic classification of single facial images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12, 1357-1362.
- Ma, L., & Khorasani, K. (2004). Facial expression recognition using constructive feedforward neural, *IEEE Transactions on networks, Systems, Man, and Cybernetics, Part B* 34, 1588-1595.
- Martin, C., Werner, U., & Gross, H. M. (2008). A real-time facial expression recognition system based on active appearance models using gray images and edge images, in: *8th IEEE International Conference on Automatic Face & Gesture Recognition*, pp. 1-6.
- Matsumoto, D. (2001). *Culture and Emotion, The Handbook of Culture and Psychology*, Oxford University Press, New York, 171-194.
- Matsumoto, D., Keltner, D. D., Shiota, M. N., O'Sullivan, M., & Frank, M. (2008). Facial



- expressions of emotions, *Handbook of emotions* 3, 211-234.
- Matuszewski, B. J., Quan, W., & Shark, L. K. (2011). Facial Expression Recognition, Book Chapter in Biometrics Uniques and Diverse Applications in Nature. *Science, and Technology, INTECH Open*.
- Montagu, J. (1994). *The Expression of the Passions: The Origin and Influence of Charles Le Brun's Conférence sur l'expression générale et particulière*, Yale University Press.
- Morik, K., Brockhausen, P., & Joachims, T. (1999). Combining statistical learning with a knowledge-based approach - A case study in intensive care monitoring, in: *16th International Conference on Machine Learning (ICML-99)*, pp. 268-277.
- Nicolaou, M., Gunes, H., & Pantic, M. (2011). Continuous prediction of spontaneous affect from multiple cues and modalities in valence-arousal space, *IEEE Trans. Affective Comput.* 2, 92-105.
- Otsuka, T., & Ohya, J. (1998). Spotting segments displaying facial expression from image sequences using HMM, in: *IEEE Proceedings of the Second International Conference on Automatic Face and Gesture Recognition (FG'98)*, Nara, Japan, pp. 442-447.
- Pantic, M., & Rothkrantz, L. J. M. (2000). Automatic analysis of facial expressions: The state of art, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22, 1424-1445.
- Pantic, M., & Rothkrantz, L. J. (2004). Facial action recognition for facial expression analysis from static face images, *IEEE Transactions on Systems, Man, and Cybernetics, Part B* 34, 1449-1461.
- Pilz, K. S., Thornton, I. M., & Bulthoff, H. H. (2006). A search advantage for faces learned in motion, *Experimental Brain Research*, 171, 436-447.
- Quan, W., Matuszewski, B. J., Shark, L. K., & Ait-Boudaoud, D. (2009). 3-D facial expression biometrics using statistical shape models, special issue on recent advances in biometric systems: a signal processing perspective, *EURASIP Journal on Advances in Signal Processing*, 1-17.
- Quan, W., Matuszewski, B. J., Shark, L. K. (2010). A statistical shape model for deformable surface registration, *International Conference on Computer Vision Theory and Applications, Angers, France*.
- Quan, W., Matuszewski, B. J., Shark, L. K. (2010). Improved 3D facial representation through statistical shape model, in: *IEEE International Conference on Image Processing*, pp. 2433-2436.
- Sandbach, G., Zafeiriou, S., Pantic, M., & Rueckert, D. (2011). *A dynamic approach to the recognition of 3D facial expressions and their temporal models*, in: *IEEE International Conference on Automatic Face Gesture Recognition*, pp. 406-413.
- Sandbach, G., Zafeiriou, S., Pantic, M., & Yin, L. (2012). Static and dynamic 3D facial expression recognition: *A comprehensive survey*, *Image and Vision Computing* 30, 683-697.
- Sandbach, G., Zafeiriou, S., Pantic, M., & Yin, L. (2012). Static and dynamic 3D facial expression recognition: A comprehensive survey. *Image and Vision Computing*, 30(10), 683-697.
- Savran, A., Sankur, B., & Bilge, M. T. (2010). IEEE CVPR'10 Workshop on Human Communicative Behavior Analysis, *San Francisco, California, USA*.
- Scheirer, J. R., Fernandez, & Picard, R. W. (1999). Expression glasses: a wearable device for facial expression recognition, in: *CHI'99 extended abstracts on Human factors in computing systems*, ACM, pp. 262-263.
- Sha, T., Song, M., Bu, J., Chen, C., & Tao, D. (2011). Feature level analysis for 3D facial expression recognition, *Neurocomputing* 74, 2135-2141.
- Shan, C., Gong, S., & McOwan, P. W. (2009). Facial expression recognition based on Local Binary Patterns: *A comprehensive study*, *Image and Vision Computing* 27, 803-816.
- Shan, C., & Braspenning, R. (2010). Recognizing facial expressions automatically from video, in: *Handbook of ambient intelligence and smart environments*, Springer, USA, pp. 479-509.
- Shen, L., & Bai, L. (2006). Mutual Boost learning for selecting Gabor features for face recognition, *Pattern Recognition Letters* 27, 1758-1767.
- Shih, F. Y., Chuang, C. F., & Wang, P. S. (2008). Performance comparisons of facial expression recognition in JAFFE database, *International Journal of Pattern Recognition and Artificial Intelligence* 22, 445-459.
- Sim, T., Baker, S., & Bsat, M. (2002). The CMU pose, illumination, and expression (PIE) database, in: *5th IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 46-51.
- Soyel, H., & Demirel, H. (2009). Optimal feature selection for 3D facial expression recognition with geometrically localized facial features, in: *5th International Conference on Soft Computing, Computing*

- with Words and Perceptions in System Analysis, Decision and Control*, pp. 1-4.
- Soyel, H., & Demirel, H. (2010). Optimal feature selection for 3D facial expression recognition using coarse-to-fine classification, *Turk. J. Electr. Eng. Comput. Sci.* 18, 1031-1040.
- Tian, Y. L., Kanade, T., & Cohn, J. F. (2001). Recognizing action units for facial expression analysis, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23, 97-115.
- Tian, Y. L., Kanade, T., & Cohn, J. F. (2002). Evaluation of Gabor-wavelet-based facial action unit recognition in image sequences of increasing complexity, in: Proceedings of 5<sup>th</sup> IEEE International Conference on Automatic Face and Gesture Recognition, pp. 229-234.
- Tian, Y. L., Kanade, T., & Cohn, J. F. (2002). Evaluation of Gabor-Wavelet-Based Facial Action Unit Recognition in Image Sequences of Increasing Complexity, in: *Proc. 5<sup>th</sup> IEEE Int'l Conf. Automatic Face and Gesture Recognition*, pp. 218-223.
- Tian, Y. L., Kanade, T., & Cohn, J. F. (2005). Facial expression analysis, *Handbook of face recognition*, Springer, New York, pp. 247-275.
- Tomkins, S. S. (1962). Affect / imagery / consciousness, vol. 1: The positive effects, Springer, New York.
- Tomkins, S. S. (1963). Affect / imagery / consciousness, vol. 2: The negative effects, Springer, New York.
- Tomkins, S. S., & McCarter, R. (1964). What and where are the primary affects? Some evidence for a theory. *Perceptual and Motor Skills* 18, 119-158.
- Tong, Y., Liao, W., & Ji, Q. (2007). Facial action unit recognition by exploiting their dynamic and semantic relationships, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29, 1683-1699.
- Tsalakanidou, F. & Malassiotis, S. (2010). Real-time 2D+3D facial action and expression recognition, *Pattern Recognition* 43, 1763-1775.
- Valstar, M. F., Patras, I., & Pantic, M. (2005). Facial Action Unit Detection Using Probabilistic Actively Learned Support Vector Machines on Tracked Facial Point Data, in: *Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition, Workshop Vision for Human-Computer Interaction*, pp. 76-76.
- Valstar, M. F., Jiang, B., Mehu, M., Pantic, M., & Scherer, K. (2011). The first facial expression recognition and analysis challenge, in: *IEEE International Conference and Workshops on Automatic Face & Gesture Recognition (FG 2011)*, pp. 921-926.
- Valstar, M. F., & Pantic, M. (2012). Fully automatic recognition of the temporal phases of facial actions, *Systems, Man, and Cybernetics, Part B*, 42, 28-43.
- Vinciarelli, A., Pantic, M., Heylen, D., Pelachaud, C., Poggi, I., D'ericco, F., & Schroeder, M. (2012). Bridging the gap between social animal and unsocial machine: a survey of social signal processing, *IEEE Trans. Affective Comput.* 3, 69-87.
- Vick, S. J., Waller, B. M., Parr, L. A., Pasqualini, M. C. S., & Bard, K. A. (2007). A Cross-species Comparison of Facial Morphology and Movement in Humans and Chimpanzees Using the Facial Action Coding System (FACS), *Journal of Nonverbal Behavior* 31, 1-20.
- Wen, Z., Huang, T. (2003). Capturing subtle facial motions in 3D face tracking, in: *Proceedings of the IEEE International Conference on Computer Vision, vol. 2*, pp. 1343 -1350.
- Xiao, R., Zhao, Q., Zhang, D., & Shi, P. (2011). Facial expression recognition on multiple manifolds, *Pattern Recognition* 44, 107-116.
- Yun, T., & Guan, L. (2010). Human emotion recognition using real 3D visual features from Gabor library, in: *IEEE International Workshop on Multimedia Signal Processing*, pp. 505-510.
- Yin, L., Wei, X., Sun, Y., Wang, J., & Rosato, M. (2006). A 3D facial expression database for facial behavior research, in: *7th International Conference on Automatic Face and Gesture Recognition, (FGR 2006)*, pp. 211 -216.
- Zeng, Z., Pantic, M., Roisman, G., & Huang, T. (2009). A survey of affect recognition methods: audio, visual, and spontaneous expressions, *IEEE Trans. Pattern Anal. Mach. Intell.* 31, 39-58.
- Zhang, Y., Ji, Q. (2005). Active and Dynamic Information Fusion for Facial Expression Understanding from Image Sequences, *IEEE Trans. Pattern Analysis and Machine Intelligence* 27, 699-714.
- Zhou, M., Wei, H., & Maybank, S. (2006). Gabor wavelets and AdaBoost in feature selection for face verification, in: *Proceedings of Applications of Computer Vision workshop in conjunction with ECCV, Graz, Austria, 2006*, pp. 101-109.