Chapter 10
Evaluation of the Intricacies of Emotional Facial Expression of Psychiatric Patients Using Computational Models

Swarup Poria, Ananya Mondal and Pritha Mukhopadhyay

10.1 Introduction

Over the past two decades, understanding of spontaneous human behavior has attracted a great deal of interests among the scientists across different domains of science, such as clinical and social psychology, cognitive science, computational science, and medicine, because of its possible wide range applications in the spheres of mental health, defense, and security. Individuals' own spontaneous facial expressions and their appraisal of facial expressions of others with little or no effort in their daily encounters are adjudged to be one example of spontaneous human behavior. Facial expression being such a fine index of one's inner experience could be considered as a very important tool for understanding the varied underlying emotional states of human mind. Facial expression classification can be effectively used in understanding prevailing affect and mood states of the individual under any pathological condition or in a deviant state of mind, particularly in the diagnosis of psychiatric disorders or in identifying the underlying affect and motives of criminals, alcoholics, and even in individuals with psychiatric disorders. Human facial expression is a prominent means that communicates one's underlying affective state and intentions (Keltner and Ekman 2000) almost instantaneously. Two main streams in the current research on automatic analysis of facial expressions are facial emotion detection and feature-based expression classifier. Facial action coding system (FACS; Ekman and Friesen 1978) is an objective measure used to identify facial configurations before interpreting displayed affect. Although different approaches such as facial motion

S. Poria
Department of Applied Mathematics, University of Calcutta, Kolkata, India

A. Mondal · P. Mukhopadhyay (✉)
Department of Psychology, University of Calcutta, Kolkata, India
e-mail: prithamukhopadhyay@gmail.com

(Essa and Pentland 1997) and deformation extraction (Huang and Huang 1997) have been adopted, the existing facial expression analysis methods still require considerable modification to increase their accuracy, speed, and reliability in identifying emotional expression, considering both the unconstrained environmental conditions and intra- and inter-personal variability of facial expressions. Existing behavioral studies indicate that there is a significant relevance of emotional facial expressions (EFEs) in psychiatric diagnosis in two ways:

(i)   The extent to which the patient is capable of expressing specific emotions and,
(ii) The extent to which they can accurately assess the facial emotional expression in others.

In spite of dramatic developments in the research on automatic facial expression with the advancement in techniques of image and video processing (Pantic and Bartlett 2007), application of computational model in psychiatric illness is a relatively new domain of research. Therefore, it is imperative to examine various aspects of EFE analysis from the perspective of psychiatric diagnosis through computational approach.

Computational techniques have proved to be powerful tools to complement the human evaluation of EFEs (Wehrle and Scherer 2001). In the early 60s, revolutionary work started with the Abelson's model of hot cognition (1963), Colby's model of neurotic defense (1963), Gullahorns' model of a homunculus (1963), and Toda's Fungus Eater (1962). However, these models were developed based on the appraisal theories of emotion (Wehrle and Scherer 2001). According to these authors, the limitation of neglecting the time dimension of models of appraisal theory could be overcome through the development of process theory, which considers multi-level processes than single-shot configurational predictions of labeled emotions. This process modeling is important as EFE constantly shows dynamic changes as the human brain in dynamic contexts receives continual input with an automatized demand on the changing of emotions.

In the backdrop of this increased understanding of the computational analysis of emotion, its application in the psychiatric diagnosis was considered to be important. Computer-based automatic facial action coding system (AFACS) enables researchers to process volumes of data in a short time in a more systematic manner with high accuracy, once the system secures the parameters for a model. Kanade (1973) was the first to implement computational modeling approach for automatic facial expression recognition.

The work presented here addresses how the database may be framed and what type of information could be stored in order to provide with concrete input for computational analysis. This could aid in development of computational model that infers diagnosis from facial expressions observed through real-time video to obtain accurate output through computational analysis.

The following will be the major objectives of this chapter:

• The importance of facial expression as a measure in psychiatric diagnosis.
• The effectiveness of computational models of automatic facial expression analysis (AFEA) to aid in diagnosis.
• Finally, the usefulness of computational approach in facial expression analysis as a measure of psychiatric diagnosis.

10.2 Importance of Facial Expression in Psychiatric Diagnosis

10.2.1 Emotional Facial Expressions

Emotional expression is a specific constellation of verbal and nonverbal behavior. Emotional messages mediated through nonverbal behaviors include facial expression, paralanguage, gesture, gaze, and posture. Face is thought to have primacy in signaling affective information since it portrays the significance of emotions to both basic survival and communication (Mandal and Ambady 2004). A schematic diagram of usefulness of facial expressions is shown in Fig. 10.1. Facial expressions owing to their deep phylogenetic origin are expected to have greater innate disposition to reflect true psychic reality through spontaneous and genuine emotional expressions, that is, EFE.

Darwin (1872) in "The Expression of the Emotions in Man and Animals" opined that human expressions of emotions have evolved from similar expressions in other animals, which are unlearned innate responses. He emphasized on genetically determined aspects of behavior to trace the animal origins of human characteristics, such as the contraction of the muscles around the eyes in anger or in putting efforts to memorize, or pursing of the lips when concentrating on any task. The study of Sackett (1966) on nonhuman primates also indicates that the evocation of emotional reactions to a threat display is controlled by "innate releasing mechanisms" that are indices of inborn emotions. Spontaneous facial muscle reactions that occur independent of conscious cognitive processes (Dimberg 1997; Ekman 1992a, b; Schneider and Shiffrin 1977; Zajonc 1980) are functions of biologically determined affect programs for facial expressions (Tomkins 1962). Wide range of cross-cultural research also confirmed the spontaneity (Dimberg 1982) and universality (Ekman and Friesen 1971; Ekman et al. 1969; Izard 1971) of facial expressions. More universality in EFE, owing to its phylogenetic underpinning, has possibility of being more accurately detected across the culture. But social display rules applied to basic emotions are culture bound. Ekman and Friesen (1975) proposed the existence of culture-specific display rules which govern the expressions of emotions. Consideration of the concept of these rules is relevant here to understand spontaneous facial expressions. Social display rules are informal norms of the social groups about expressing emotions.

Felt spontaneous emotion is also an expression of display rules when it is internalized and becomes apart of one's spontaneous emotional repertoire. In contrast to the felt emotion, when social display rules are not automatized, the nonfelt emotional expressions using these rules are posed emotions. In this context, it

Fig. 10.1 Usefulness of facial expression
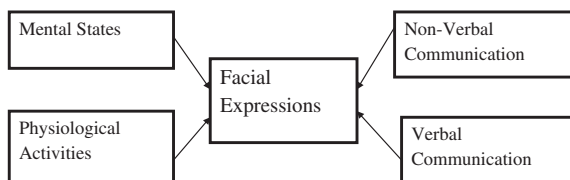
could be assumed that since patients with psychotic symptoms are not able to utilize environmental feedback immediately using the display rules, this in turn makes them unable to modify their spontaneous emotional expressions.

Automatized spontaneous emotion, which is involuntary in nature, may be mediated by subcortically initiated facial expression characterized by synchronized, smooth, symmetrical, consistent, and reflex-like facial muscle movements, in contrast to cortically initiated facial expressions those are subject to volitional real-time control, and tend to be less smooth, with more variable dynamics (Rinn 1984). This distinction between biologically disposed emotional expression and emotional expression that is modified in congruence with environmental feedback explains the possibility of spontaneous emotional expression in psychiatric patients who have poor reality-, self-, and source-monitoring capacity (Hermans et al. 2003; Radaelli et al. 2013). For example, in case of schizophrenia, one can structure diagnostic formulation by developing suitable computational model incorporating the database so far obtained from research outcome. It has been evident from neuropsychological studies that hemi-facial differences (Rinn 1984; Sackeim et al. 1978; Wolff 1943) give a direction of assessment of emotional expression in psychiatric population. Since, in psychiatric illness, spontaneous emotional expression is expected, it could be more rational to make computational program based on the information from the left (Wolff 1943) and lower regions of the face (Borod and Koff 1984), considering the degree of facial symmetry (Borod et al. 1997), which reflects the true inner emotional state of the patient that is beyond modification by social display rules. The more one is self-absorbed and less reality-oriented, the less will be the difference between the activation of the regions of the face responsible for involuntary expressions and voluntary modification of emotional expressions, in favor of involuntary expression. It makes the facial expression more symmetrical.

But the assessment of conscious control of emotional expression is also to be taken into account. It itself can be an index of one's ability to appraise social situations that directs one to modify the emotional facial expression, which, if detected through measures of facial affect in course of psychiatric treatment, could also be considered as a valid prognostic indicator of positive treatment outcome.

Since patients are reared up in a given culture and learn to express emotions following culture-specific display rules, so culture-specific facial expressions cannot be totally ignored in diagnostic assessment. At this juncture, it is important to consider whether the culturally learnt emotional expression is automatized in patients or has employed display rule to suppress genuine emotional response. Comprehensive database relating the emotional facial expressions with the hemifaces may help to solve this problem.

Spontaneous facial expression analysis as a diagnostic tool in psychiatric illness is important as it may be the true reflection of the patients' inner psychic reality. Since facial expressions help to better understand the present mental status of the patient, it could be an important marker for clinical diagnosis. In general, clinical diagnosis is being done by clinical observation of the patients using the observation method, using the case study method, and by using different psychometric

tools. Assessment of facial expressions through computational models definitely would have a significant contribution in confirming the diagnosis. Different types of facial expression measurements include observer's judgment, electrophysiological recording, and componential coding schemes.

In case of observer's judgment, observers view behavior and make judgments on the extent to which they see emotions in the target face. However, even though this method seems to be easy to execute, it is not free from its limitations. There is the possibility of subjective bias in judgment, which makes them neglect certain subtle morphological characteristics of facial expressions. The inevitability of instantaneous classification of facial expressions for diagnosis of different disorders is another major disadvantage of human judgment concerning facial expression.

Emotional facial expressions can also be measured through facial electromyography (EMG), which measures electrical potential from facial muscles in order to infer muscular contraction via the placement of surface electrodes on the skin of the face. But it may increase one's self-conscious behavior and minimize genuine emotional expressions. Another problem in EMG signal measure is "cross talk," that is, surrounding muscular contraction (potentials) interferes with the signal of a given muscle group. For emotional expressions, such changes in muscular contraction of adjacent muscle groups may misrepresent the picture of facial muscle movements, which may delude emotional interpretations (Tassinary et al. 1989).

By contrast, in componential coding schemes, coders follow some prescribed rule to detect subtle facial actions. The two most popular systems are the FACS, developed by Ekman and Friesen (1978), and the maximally discriminative facial movement coding system (MAX), developed by Izard (1979). MAX is a theoretically driven system where only that facial configuration can be coded, which corresponds to universally recognized facial expressions of emotions, whereas FACS is a comprehensive measurement system, which measures all the observable movements in the face. This system is not limited to the behaviors that are theoretically related to emotions. It allows discovering new configurations of movements.

With the advancement of technologies, recently FACS has become a more objective automatic facial expression componential coding analysis. This is an interesting and challenging area of research. It has important application in many areas such as emotional and paralinguistic communication, pain assessment, lie detection, and multimodal human computer interface. AFACS, in addition to processing volumes of data in a short time in a more systematic manner, performs with a higher accuracy than human coders who have large margins of error and may overlook important information. Acquisition of the data of parameters for a model and prediction of behavior (vs. simple detection and classification) using only a small sample are other advantages of automated facial detection technology when coupled with computational models.

To differentiate clinical population from normal controls, computational models identify differential features or action units involved in the patient's facial expression. Patients with schizophrenia, for example, those with impairment in source monitoring (Arguedas et al. 2012) as well as reality monitoring (Radaelli

et al. 2013), are not capable of appraising the situational demand to modify their felt emotional expressions, which are governed by the extrapyramidal tracts of subcortical origin (Van Gelder 1981, cited in Borod and Koff 1991). Thus, the identification of the more pronounced felt emotion in the hemi-face or symmetrical emotional expression between the two hemi-faces and information of the relative level of activation of the tracts of subcortical and cortical regions of the brain could be a strong database for computational analysis of facial expression in diagnosis of psychiatric illness.

Since accuracy of judgment of EFE increases with more pronounced emotion (in either directionless or more than usual expression), the efficiency of detection is presumed to be more in diagnosis of psychotic cases. In this context, Borod et al. (1997) suggested that asymmetry of facial action is not observed during spontaneous expression because subcortical structures innervate the face with bilateral fiber projections. This strong biological disposition suggests that facial symmetry for the felt emotion (governed by either automatized social display rules or innate disposition) is expected to characterize psychotic patients and may be considered as an important part of the database for computational analysis of EFE. In addition, patients with schizophrenia display more negative than positive emotions (Martin et al. 1990), exhibit expressions of contempt more frequently than other emotions (Steimer-Krause et al. 1990), and show a lower proportion of joyful expressions (Schneider et al. 1992; Walker et al. 1993) across the situations, which could be considered as idiosyncratic characteristics of schizophrenia and requires due attention for diagnosis during computational analysis. Gruber et al. (2008) in their research also suggested that participants at high risk of mania reported irritability and elevated positive emotion. Such information of disorder-specific emotional biasness can also be incorporated to distinguish different types of disorders.

Neuroanatomical research suggests relatively independent neuroanatomical pathways for posed or nonfelt and spontaneous emotional facial expression. These independent pathways produce facial asymmetry when one puts conscious effort in modifying felt emotion (Campbell 1978, 1979) in contrast to facial symmetry in case of spontaneous expression (Remillard et al. 1977). Posed emotional expressions are governed by the pyramidal tracts of the facial nerves that descend from the cortex (Van Gelder 1981, cited in Borod and Koff 1991), which suggests the voluntary cortical control over felt emotion. This information indicates the importance of considering the ratio of posed emotion in comparison with the felt affect in psychiatric cases to create the database. The greater the monitoring capacity, the less severe the disorder and the better the prognosis.

However, efficient monitoring capacity, in some psychiatric disorders characterized by manipulativeness, could be one of the diagnostic indicators of that disorder instead of being a positive quality of the individual. Consideration of this sort of overlapping of information in the database is also necessary for relevant differential diagnosis.

Even in anxiety disorders, in spite of putting effort to monitor emotional expressions according to the demand of reality and social display rules (Hermans et al. 2003), their intrinsic emotional force causes a veridical leakage in emotional

expression failing to totally suppress the ingenuity of emotional expression. In an early observation, Dunchenne, a nineteenth-century neurologist, claimed that genuinely happy smiles as opposed to false smiles involve contraction of a muscle near the eyes, the lateral part of the *orbicularis oculi*. According to him, zygomatic major muscle obeys the volition and its volitional disposition is so strong that the lateral part of the *orbicularis oculi* cannot be evoked by deceitful laugh (Dunchenne 1862/1990). Ekman confirmed this early observation (Ekman 1992a, b). Neurobiological disposition of the upper face is less voluntarily controlled than that of the lower face because the upper face has neural link with the motoric speech center (Rinn 1984). Again, it has been reported in early literature that being monitored by the respective opposite hemisphere, the left side of the face is more under unconscious control, expressing hidden emotional content (Wolff 1943), and the right side of the face is more under conscious control revealing interpersonally meaningful expressions (Wolff 1943, cited in Sackeim et al. 1978).

Like patients with anxiety disorder, depressive patients also show intact reality monitoring (Benedetti et al. 2005) but selective source monitoring (Ladouceur et al. 2006). The most prominent finding in depression so far has been the attenuation of smiles produced by zygomatic major activity (Girard et al. 2013). This led many researchers to conclude that depression is marked by attenuation of expressed positive affect.

It follows from the above discussion that bio-behavioral understanding of identification of such robust indices, such as facial asymmetry and facial areas in emotional expression, if properly utilized, may enhance the efficiency of the computational diagnostic tool. Not only facial asymmetry in general but the database of onset, latency of expression, apex duration (how long it remains), time distance between hemi-facial expressions, primarily occurred hemi-facial expression, and intensity dominance of the hemi-face also could be of immense diagnostic importance.

Understanding of one's inner state through the analysis of emotional expressions can be strengthened by corroborating analysis of facial expression of patients with the database of their appraisal of others' emotions.

10.2.2 Appraisal of EFE

As it is understood, another source to obtain information regarding the patient's understanding of emotion is to study the accuracy of their appraisal of others' emotional expression. The spontaneous emotional reaction, which is controlled by biologically driven affect program, also suggests that this facial affect program aids in spontaneous recognition of emotion in others. The reports (Benson 1999; Girard et al. 2013) of deficits in portraying emotional expressions in different groups of psychiatric patients have implied that they could have some experiential problems that get reflected when different groups of psychiatric patients appraise others' emotions.

It has been evident that schizophrenic patients are significantly inferior to normal controls in the ability to decode universally recognized facial expressions of emotions. The findings of impairment in schizophrenia, in their own emotional expressions and in recognition of emotions, imply that a person with impairment in one's own facial expressions finds it difficult to decode facial expressions of emotion in others. According to Hall et al. (2004), besides deficits in other aspects of face perception in schizophrenia, deficit in expression recognition performance is also an important index to be related to their social dysfunction (Hooker and Park 2002), thereby inducing research interest from the perspective of appraisal of emotional expression in others.

In a meta-analysis, Uljarevic and Hamilton (2013) also found that individuals with autism have trouble in classifying emotions. Disorder of social functioning in autism that is associated with impairment of automatic mimicry may be the reflection of impairment in the functioning of mirror cells in this disorder (McIntosh et al. 2006).

Skill deficits in the appraisal of facial expressions or in decoding EFE are also evident from the EFE recognition deficits in detoxified alcoholics but for different etiological factors. Deficit in detoxified alcoholics is correlated more with interpersonal problems compared to normal control (Kornreich et al. 2002). Thus, interpersonal difficulties could also serve as a mediator of EFE accuracy problems.

When such overlapping features are present in the appraisal of EFE by patients with different disorders, the profile of symptomatic and etiological features of the disorder and detailed clinical evaluation is to be incorporated for the accurate computational analysis for differential diagnosis in addition to database of EFE. This would give direction to diagnosis with inclusion of both the clinical and empirical approaches together to reach a final decision.

10.2.3 Differential Diagnosis

When the areas of facial expressions overlap across the diagnosis, differential diagnosis is required. For example, in case of anxiety disorder and affective disorder, signal value of facial expression of patients can produce equivocal results. Since patients with these disorders have reality contact and can modify their felt expression in accordance with situational demand, the signal value becomes overshaded by the attempt by the patients to mask the spontaneous EFE. When particular facial markers cannot differentiate the disorders, differential diagnosis is suggested based on strong theoretical groundwork of these disorders. For example, in depression, the patients' facial expression can be explained by the social risk (SR) hypothesis, which states how depressed mood minimizes social communication to restrict any negative outcome variability, which is not within the depressed persons' acceptable zone. They tend to show reservation in help-seeking behavior in reciprocal interactions even with their relatives and close ones who are likely to provide the requested help. In competitive contexts, however, their interaction is

dominated by submission and withdrawal (Girard et al. 2013), whereas for patients with anxiety disorder, researchers (Barlow et al. 1996; Mineka and Zinbarg 1996) hypothesize that the patients may perceive situations as unpredictable or uncontrollable. They always have a tendency to look for signs of threat; this hyper-vigilant activity can be reflected in their facial expressions. Thus, a detailed analysis of facial expression along with its assimilation with theoretical foundation and symptom manifestation can reduce the ambiguity of facial expression analysis among anxiety- and depression-related patients.

From the perspective of differential diagnosis, research studies claimed that acute schizophrenia showed greater emotion decoding impairment than the depressive and normal controls (Gessler et al. 1989). Paranoid schizophrenia subjects were more accurate than the nonparanoid subjects in judging facially displayed emotions (Kline et al. 1992; Lewis and Garver 1995). Anstadt and Krause (1989) used primary affects in portraits and concluded that schizophrenic patients were more impaired in terms of the quality and diversity of the action unit (AU) (the cluster of muscles in an area of the face during expression) drawn in facial expressions. Borod et al. (1993) suggested that these patients have difficulty specifically in comprehending facial emotional cues (facial expressions of emotions) but not in nonemotional facial cues. Domes et al. (2008) indicated that borderline individuals more or less accurately perceive others' emotions and show a tendency toward heightened sensitivity in recognizing specifically anger and fear in social context. Negative response bias in depressed patients may explain their tendency to attribute neutral faces as sad and happy faces as neutral (Stuhrmann et al. 2011). Their findings imply selective attentional bias in recognition of emotions among individuals with different disorders. Since decoding deficit of facial expression is evident across different diagnostic categories, computational analysis may help identify even subtle differential points necessary for accurate diagnosis and also for understanding diagnostic specificity, if any, in appraisal of emotions in others.

Thus, the specific ways in which an individual processes and attributes emotional information can be a strong determinant of psychiatric diagnosis, especially for affective and anxiety disorders. The accuracy with which a patient can assess the expression of emotion in others is another way to determine the severity of the psychiatric disorder, which can be utilized for framing the database for diagnostic computational modeling.

10.3 Computational Models for Automatic Facial Expression Analysis

Machine analysis of facial expressions attracted the interest of many researchers because of its importance in cognitive and medical sciences. Although humans detect and analyze faces and facial expressions with little effort, development of an automated system for this task is very difficult. Since 1970s, different approaches are proposed for facial expression analysis from either static facial
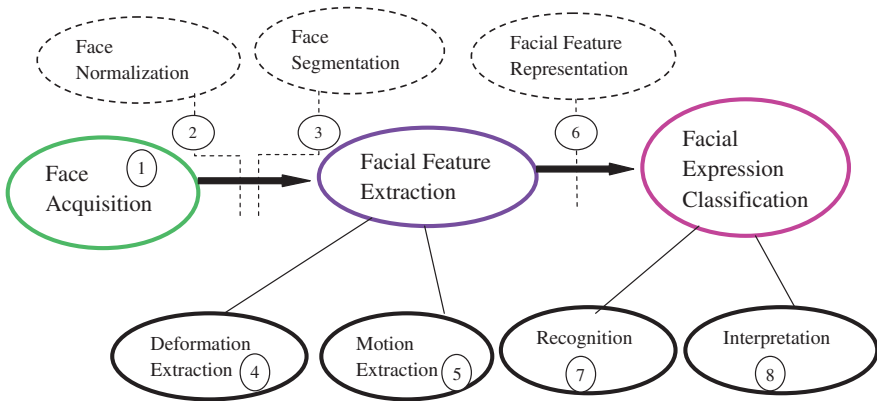
Fig. 10.2 A generic facial expression analysis framework

images or image sequences. AFEA is a complex task as physiognomies of faces vary from one individual to another considerably due to differences between age, ethnicity, gender, facial hair, cosmetic products, and occluding objects, such as glasses and hair.

According to Shenoy (2009), the first objective and scientific study of facial expression was done by Bell in (1844). Darwin (1872) argued that emotional expressions are universal and the same for all people based on his theory of evolution. Ekman has proposed the existence of six basic prototypical facial expressions (anger, disgust, happiness, sadness, surprise, and fear) that are universal. Though many facial expressions are universal in nature, but the way these are displayed depends upon culture and the upbringing. Kanade (1973) published the first work on automatic facial expression recognition. The first survey of the field was published by Samal and Iyengar in (1992) followed by others (Fasel and Luettin 2003; Pantic and Rothkrantz 2000a, b). An automatic face analysis (AFA) system was developed by Fasel to analyze facial expressions based on both permanent facial features (brows, eyes, and mouth) and transient facial features (deepening of facial furrows) in a nearly frontal-view face image sequence. He used Ekman and Friesen's FACS System to evaluate an expression. Many computational models have been developed for facial expression classification over the last few years. In general, any facial expression classification system would have the three basic units: face detection, feature extraction, and facial expression recognition. A generic facial expression analysis framework proposed by Fasel and Luettin (2003) is shown in Fig. 10.2.

Any computational model performs facial feature extraction and then uses dimensionality reduction techniques followed by a classification technique. A flowchart of computational models of facial expression analysis is shown in Fig. 10.3. Facial feature extraction consists of localizing the most characteristic face components such as eyes, nose, and mouth within images that depict human faces. This step is essential for the initialization of facial expression recognition or face recognition.
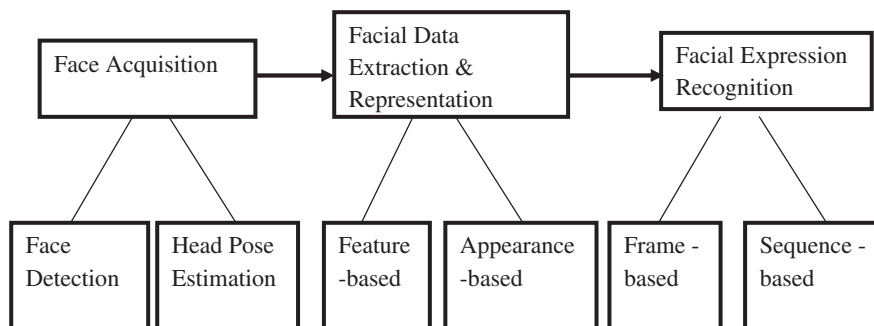
Fig. 10.3 Flowchart of computational models of facial expression analysis

There are many techniques for dimensionality reduction, such as principal component analysis (PCA) or singular value decomposition (SVD), independent component analysis, curvilinear component analysis (CCA), linear discriminant analysis (LDA), Fisher linear discriminant, multidimensional scaling, projection pursuit, discrete Fourier transform, discrete cosine transform, wavelets, partitioning in the time domain, random projections, multidimensional scaling, and fast map and its variants. PCA is widely being used for data analysis of varied areas, such as neuroscience, computational graphics, for extracting relevant information from confusing datasets because it is a simple, nonparametric method. PCA transforms higher-dimensional datasets into lower-dimensional uncorrelated outputs by capturing linear correlations among the data, and preserving input by output as much information as possible in the data. CCA is a nonlinear projection method that attempts to preserve distance relationships in both input and output spaces. CCA is a useful method for redundant and nonlinear data structure representation and can be used in dimensionality reduction. CCA is useful with highly nonlinear data, whereas PCA or any other linear methods fail to give suitable information. Fisher linear discriminant analysis (FLDA) has been successfully applied to face recognition, which is based on a linear projection from the image space to a low-dimensional space by maximizing the between-class scatter and minimizing the within-class scatter. It is most often used for classification. The main idea of the FLD is that it finds projection to a line so that samples from different classes are well separated. LDA is a special case of FLD in which both classes have the same variance. Belhuemer was the first to use the LDA on faces and used it for dimensionality reduction (Belhumeur et al. 1997). Different techniques have been proposed to classify facial expressions, such as neural network (NN), support vector machine (SVM), Bayesian network, and rule-based classifiers. SVMs introduced by Boser et al. in (1992) have become very popular for data classification. SVMs are most commonly applied to the problem of inductive inference, or making predictions based on previously seen examples. An SVM is a mathematical entity, an algorithm for maximizing a particular mathematical function with respect to a given collection of data. In the SVM classifier, there are a number of parameters

that must be chosen by the user. It is necessary to make the right choices of these parameters in order to yield the best possible performance. The basic concept underlying the SVM is quite simple and intuitive and involves separating our two classes of data from one another using a linear function that is the maximum possible distance from the data.

To understand the essence of SVM classification, one needs only four basic concepts: (1) the separating hyper-plane, (2) the maximum-margin hyper-plane, (3) the soft margin, and (4) the kernel function. But there exist free and easy-to-use software packages which allow one to obtain good results with a minimum of effort. Viola and Jones (2001) devised Haar Classifiers algorithm using AdaBoost classifier cascades that are based on Haar-like features and not pixels for rapid detection of objects including faces. Wavelet transform could extract both the time (spatial) and frequency information from a given signal, and the tunable kernel size allows it to perform multi-resolution analysis. Among different wavelet transforms, the Gabor wavelet transform has some impressive mathematical and biological properties and has been used frequently on researches of image processing. The Gabor wavelet is a linear filter where impulse response is defined by a harmonic function multiplied by a Gaussian function. This filter can be used to detect line endings and edge borders over multiple scales and with different orientations. Gabor wavelet is used for facial feature extraction in computational models of facial expression analysis.

Recently, Shan et al. (2009) studied facial representation based on local binary pattern (LBP) features for person-independent facial expression recognition. Recently LBP features have been introduced to represent faces in facial images analysis. The most important properties of LBP features are their tolerance against illumination changes and their computational simplicity.

Automatic pain recognition has received attention in the recent past because of its relevance in health care, ranging from monitoring patients to assessment of chronic lower back pain (Prkachin et al. 2002). Lucey et al. (2011, 2012) address AU and pain detection based on SVMs. Details of computational methods for AFEA are reported by Pantic and Rothkrantz (2000a, b). We discuss briefly the fundamentals of AFEA.

10.3.1 Face Detection

The first step in facial information processing is face detection. Determining the exact location of a face within a large background is a very difficult job for a computational system. An ideal face detection system should be capable of detecting faces within a noisy background and in complex scenes. Facial components, such as eyes, nose, eyebrows, are the prominent features of the face. In holistic face representation, face is represented as a whole, while on the other hand, in analytic face representation, face is represented as a set of facial features. Face can also be represented as a combination of these, and such a representation is called

hybrid representation. Many face detection methods have been developed to detect faces in an arbitrary scene (Li and Gu 2001; Pentland et al. 1994; Rowley et al. 1998; Schneiderman and Kanade 2000; Viola and Jones 2001). Most of them can detect only frontal and near-frontal views of faces. A neural-network-based system to detect frontal-view face has been developed by Rowley et al. (1998). Viola and Jones (2001) developed a robust real-time face detector based on a set of rectangle features. Modular eigenspace method for face detection was developed by Pentland et al. (1994). Schneiderman and Kanade (2000) developed a statistical method for 3D object detection that can reliably detect human faces. Li and Gu (2001) proposed an AdaBoost-like approach to detect faces with multiple views.

Huang and Huang (1997) apply a point distribution model (PDM) to represent the face. Huang and Huang utilize a Canny edge detector to obtain a rough estimate of the face location in the image. Pantic and Rothkrantz (2000b) detect the face as a whole unit—they use dual-view facial images. Kobayashi and Hara (1997) use a CCD camera in monochrome mode to obtain brightness distribution data of the human face. Yoneyama et al. (1997) use an analytic approach for face detection in which the outer corners of the eyes, the height of the eyes, and the height of the mouth are extracted automatically. Liu (2003) presents a Bayesian discriminating features (BDF) method. Kimura and Yachida (1997) use potential net for face representation.

In order to perform a real-time tracking of the head, Hong et al. (1998) utilized the person spotter system proposed by Steffens et al. (1998). Steffens et al. (1998) system performs well in the presence of background motion, but fails in the case of covered or too much rotated faces. To locate faces in an arbitrary scene, Essa and Pentland (1997) use the eigenspace method of Pentland et al. (1994). The method employs eigenfaces approximated using PCA on a sample of facial images.

10.3.2 Facial Expression Data Extraction

After face detection, the next step is to extract the features that may be relevant for facial expression analysis. In general, three types of face representation are used in facial expression analysis: holistic (Kato et al. 1992), analytic (Yuille et al. 1989), and hybrid (Lam and Yan 1998). For facial expression data extraction, a template-based or a feature-based method is applied. Template-based methods fit a holistic face model, and feature-based methods localize the features of an analytic face model in the input image or track them in the input sequence.

10.3.2.1 Static Images

Edwards et al. (1998) utilized a holistic face representation method. They used facial images coded with 122 points around facial features to develop a model known as active appearance model (AAM). Edwards et al. (1998) aligned training

images into a common coordinate frame and applied PCA to get a mean shape for generating statistical model of shape variation. The AAM search algorithm failed to converge to a satisfactory result (Cootes et al. 1998) in 19.2 % of the cases. The method works with images of faces without facial hair and glasses, which are hand-labeled with the landmark points beforehand approximated with the proposed AAM (Pantic and Rothkrantz 2000b). To represent the face, Hong et al. (1998) utilize a labeled graph. They defined two different labeled graphs known as big general face knowledge (GFK) and small GFK. A big GFK is a labeled graph with 50 nodes, and a small GFK is a labeled graph with 16 nodes. The small GFK is used to find the exact face location in an input facial image. On the other hand, the big GFK is used to localize the facial features. Hong et al. (1998) utilized the person spotter system, and the method of elastic graph matching proposed by Wiskott (1995) to fit the model graph to a surface image. Hong et al. (1998) utilize the person spotter system (Steffens et al. 1998) for facial expression analysis from static images.

Huang and Huang (1997) utilize a PDM developed by Cootes et al. (1998) to represent the face. The PDM is a model for representing the mean geometry of a shape and some statistical modes of geometric variation inferred from a training set of shapes. The mouth is included in the model by approximating the contour of the mouth with three parabolic curves. Success of the method is strongly constrained.

Padgett and Cottrell (1996) used a holistic face representation, but did not deal with information extraction through faces in an automatic way. They used the facial emotion database assembled by Ekman and Friesen. Yoneyama et al. (1997) used a hybrid approach for face representation. Their method will fail to recognize any facial appearance change that involves a horizontal movement of the facial features. For correct facial expression analysis using this method, the face should be without facial hair and glasses and no rigid head motion is allowed. Zhang et al. (1998) use a hybrid approach to face representation, but do not deal with facial expression information extraction in an automatic way. A similar face representation was recently used by Lyons et al. (1999) for expression classification into the six basic plus neutral emotions.

Kobayashi and Hara (1997) proposed a geometric face model of 30 facial characteristic points (FCPs). Later, they utilized a CCD camera in monochrome mode to obtain a set of brightness distributions of 13 vertical lines crossing the FCPs. A major drawback of the method is that the facial appearance changes encountered in a horizontal direction cannot be modeled. A real-time facial expression analysis system, developed by Kobayashi et al. (1995), works with online taken images of subjects with no facial hair or glasses facing the camera while sitting at approximately 1-m distance from it (Pantic and Rothkrantz 2000b). Pantic and Rothkrantz (2000b) use a point-based model composed of two 2D facial views for the frontal and the side view. They apply multiple feature detectors for each prominent facial feature (eyebrows, eyes, nose, mouth, and profile) to localize the contours of the prominent facial features and then extract the model features in an input dual view. The system cannot deal with minor inaccuracies of the extracted facial data, and it deals merely with images of faces without facial hair or glasses.

10.3.2.2   Image Sequences

Black and Yacoob (1995, 1997) used local parameterized models of image motion for facial expression analysis by using an affine, a planar, and an affine-plus-curvature flow model (Pantic and Rothkrantz 2000b). Otsuka and Ohya (1996) estimate the motion in the local facial areas of the right eye and the mouth by applying an adapted gradient-based optical flow algorithm (Black and Yacoob 1995). After the optical flow algorithm, a 2D Fourier transform is utilized to the horizontal and the vertical velocity field, and the lower-frequency coefficients are extracted as a 15D feature vector, which is used further for EFE classification (Pantic and Rothkrantz 2000b). This method is not sensitive to unilateral appearance changes of the left eye. Essa and Pentland (1997) used a hybrid approach to face representation. They applied the eigenspace method (Essa and Pentland 1997) to automatically track the face in the scene and extract the positions of the eyes, nose, and mouth. The method for extracting the prominent facial features employs eigenfeatures approximated using PCA. Essa and Pentland (1997) use the optical flow computation method proposed by Simoncelli (1993). This approach uses a multi-scale coarse-to-fine Kalman filter to obtain motion estimates and error-covariance information. The method used for frontal-view facial image sequences. Kimura and Yachida (1997) utilize a hybrid approach to face representation. The method seems suitable for facial action encoding. Wang et al. (1998) also use a hybrid approach to face representation. The face model used by Wang et al. (1998) represents a way of improving the labeled-graph-based models (e.g., Hong et al. 1998) to include intensity measurement of the encountered facial expressions based on the information stored in the links between the nodes.

10.3.3   Facial Expression Classification

The last step of facial expression analysis is to classify the facial features conveyed by the face. Many classifiers have been applied to expression recognition such as NN, SVMs, LDA, *K*-nearest neighbor, multi-nomial logistic ridge regression (MLR), hidden Markov models (HMM), tree augmented naive Bayes, and others. The surveyed facial expression analyzers classify the encountered expression as either a particular facial action or a particular basic emotion. Independent of the used classification categories, the mechanism of classification applied by a particular surveyed expression analyzer is either a template-based or a neural-network-based or a rule-based classification method.

10.3.3.1   Classification of Static Images

At first, automatic expression analysis from static images applies a template-based method for expression classification. The methods in this category perform expression classification into a single basic emotion category. Edwards et al. (1998)

introduce a template-based method for facial expression classification. The main aim of Edwards et al. (1998) is to identify the observed individual in a way which is invariant to confounding factors such as pose and facial expression. The achieved recognition rate for the six basic and neutral emotion categories was 74 %. Edwards et al. (1998) explain the low recognition rate by the limitations and unsuitability of the utilized linear classifier (Edwards et al. 1998). Success of the method for identifying expressions of an unknown subject is not known.

To achieve expression classification into one of the six basic plus neutral emotion categories, Hong et al. (1998) proposed another method. The achieved recognition rate was 89 % in the case of the familiar subjects and 73 % in the case of unknown persons. As indicated by Hong et al. (1998), the availability of the personalized galleries of more individuals would probably increase the system's performance. In order to perform emotional classification of the observed facial expression, Huang and Huang (1997) perform an intermediate step by calculating 10 action parameters. The achieved correct recognition ratio was 84.5 %. It is not known how the method will behave in the case of unknown subjects. Also, the descriptions of the emotional expressions, given in terms of facial actions, are incomplete. For example, an expression with lowered mouth corners and raised eyebrows will be classified as sadness. Lyons et al. (1999) report facial expression classification technique based on complex-valued Gabor transform. In general, the generalization rate is 92 %, whereas the generalization rate is 75 % for a novel subject. Yoneyama et al. (1997) extract 80 facial movement parameters, which describe the change between an expressionless face and the currently examined facial expression of the same subject. To recognize four types of expressions (sadness, surprise, anger, and happiness), they use 2 bits to represent the values of 80 parameters and two identical discrete Hopfield networks. The average recognition rate of the method is 92 %.

Now, we review methods for AFEA from static images applying a NN for facial expression classification. Except the method proposed by Zhang et al. (1998), the methods belonging to this category perform facial expression classification into a single basic emotion category. For classification of expression into one of the six basic emotion categories, Kobayashi and Hara (1992) used neural-network-based method. The average recognition rate was 85 %. For emotional classification of an input facial image into one of 6 basic plus neutral emotion categories, Padgett and Cottrell (1996) utilize a back-propagation NN. The average correct recognition rate achieved was 86 %. Zhang et al. (1998) employ a NN that consists of the geometric position of the 34 facial points and 18 Gabor wavelet coefficients sampled at each point. The achieved recognition rate was 90.1 %. The performance of the network is not tested for recognition of expression of a novel subject. Zhao and Kearney (1996) utilize a back-propagation NN for facial expression classification into one of the six basic emotion categories. The achieved recognition rate was 90.1 %. The performance of the network is not tested for recognition of the expression of a novel subject.

Just one of the surveyed methods for AFEA from static images applies a rule-based approach to expression classification. The method proposed by Pantic and

Rothkrantz (2000b) achieves automatic facial action coding from an input facial dual view in few steps. First, a multi-detector processing of the system performs automatic detection of the facial features in the examined facial image. From the localized contours of the facial features, the model features are extracted. Then, the difference is calculated between the currently detected model features and the same features detected in an expressionless face of the same person. Based on the knowledge acquired from FACS (Ekman and Friesen 1978), the production rules classify the calculated model deformation into the appropriate AUs-classes. The average recognition rate was 92 % for the upper face AUs and 86 % for the lower face AUs. Classification of an input facial dual view into multiple emotion categories is performed by comparing the AU-coded description of the shown facial expression to AU-coded descriptions of six basic emotional expressions, which have been acquired from the linguistic descriptions given by Ekman (1982). The classification into and, then, quantification of the resulting emotion labels are based on the assumption that each subexpression of a basic emotional expression has the same influence on scoring that emotion category. A correct recognition ratio of 91 % has been reported.

10.3.3.2 Classification from Image Sequences

The first category of the surveyed methods for AFEA from facial image sequences applies a template-based method for expression classification. The facial action recognition method proposed by Cohn et al. (1998) applies separate discriminant function analyses within facial regions of the eyebrows, eyes, and mouth. Predictors were facial point displacements between the initial and peak frames in an input image sequence. Separate group variance–covariance matrices were used for classification. The images have been recorded under constant illumination, using fixed light sources and none of the subjects wear glasses (Lien et al. 1998). Data were randomly divided into training and test sets of image sequences. They used two discriminant functions for three facial actions of the eyebrow region, two discriminant functions for three facial actions of the eye region, and five discriminant functions for nine facial actions of the nose and mouth region. The accuracy of the classification was 92 % for the eyebrow region, 88 % for the eye region, and 83 % for the nose/mouth region. The method proposed by Cohn et al. (1998) deals neither with image sequences containing several facial actions in a row, nor with inaccurate facial data, nor with facial action intensity (yet the concepts of the method makes it possible).

Essa and Pentland (1997) use a control-theoretical method to extract the spatio-temporal motion-energy representation of facial motion for an observed expression. By learning ideal 2D motion views for each expression category, they generated the spatio-temporal templates for six different expressions two facial actions (smile and raised eyebrows) and four emotional expressions (surprise, sadness, anger, and disgust). Each template has been delimited by averaging the patterns of motion generated by two subjects showing a certain expression. Correct frontal-view image sequences recognition rate of the method is 98 %. Kimura and Yachida (1997) fit a

potential net to each frame of the examined facial image sequence. The proposed method is unsuccessful for classification of image sequences of unknown subjects. Otsuka and Ohya (1996) match the temporal sequence of the 15D feature vector to the models of the six basic facial expressions by using a left-to-right hidden Markov model. The method was tested on image sequences shown by the same subjects. Therefore, it is not known how the method will behave in the case of an unknown expresser. Wang et al. utilize a 19-points labeled graph with weighted links to represent the face. The average recognition rate was 95 %.

Just one of the surveyed methods for AFEA from image sequences applies a rule-based approach to expression classification. Black and Yacoob (1995, 1997) utilized local parameterized models of image motion to represent rigid head motions and nonrigid facial motions within the local facial areas. The achieved recognition rate was 88 %. Lip biting is sometimes mistakenly identified as a smile (Black and Yacoob 1997).

10.4 Computational Method for Psychiatric Diagnosis

EFE is not restricted only to gross changes of facial expressions, but involves continual subtle changes in activation of facial muscles resulting in continuous modification in emotional expressions, in response to dynamic changes in the internal and external context of the individual. Thereby, the tool for understanding the psychiatric diagnosis from the EFE should have to be very sophisticated and sensitive to detect multi-level changes in emotion. Computational models for automatic face recognition could provide such improved methodology.

Computational method for understanding EFE is the development of the formulas or algorithms that are used to calculate the output on the basis of concrete input, given a concrete set of parameters (Wehrle and Scherer 2001). Any computational model, like those developed based on appraisal theory (Wehrle and Scherer 2001), requires determination of parameters to obtain input from all the relevant variables necessary for identification and digitization of location, intensity, and symmetry of hemi-faces in a diagnostic category to understand concrete affective response from EFE. Substantial database will finally help to formulate the digital version of the EFE and can predict the profile of affective response in a given psychiatric illness.

Though psychiatrists usually follow definite diagnostic procedure to diagnose psychiatric disorders by the method of interview, it could be one of the reasons for differences in opinion among them regarding diagnosis. By using facial images of a subject which dynamically changes over time during interview, we can acquire rich supplementary information from the changes in facial expressions that may immensely enhance the accuracy of the diagnosis. Kobayashi et al. (2000) also opined that conversation, behavior, and facial expressions are important in psychiatric diagnosis. They emphasized on the designing of automatic interview system in order to unify the contents and interpretations of psychiatric diagnosis. But as automatic

interview system may cause an artificial ambience, recording of video-based automated facial expression in natural ambience, subsequently translating it into digital version, appears to be the more justified approach for input of genuine data of EFE of the psychiatric patients for more accurate diagnosis. Wang et al. (2008) were the first to apply video-based automated facial expression analysis in neuropsychiatric research. They presented a computational framework that creates probabilistic expression profiles for video data and can potentially help to automatically quantify emotional expression differences between patients with neuropsychiatric disorders and healthy controls. Their results open the way for a video-based method for quantitative analysis of facial expressions in clinical research of disorders that cause affective deficits. The automated AU recognition system was also applied to spontaneous facial expressions of pain by Bartlett et al. (2006) where they used automated AU detector within AFACS to differentiate faked from real pain.

For automatic psychiatric diagnosis, video analysis of the EFE, as it happens during the period of interview, also could be incorporated to acquire the information of the change in facial expressions by using dynamically changing facial images of a subject. There are many methods for automatic extraction from different regions of face, such as eyes, eyebrows, and mouth. Frame-by-frame AU intensity analysis could make investigation of facial expression dynamics possible, following the design adopted in pain management program (Bartlett et al. 2006). In their study, coding of each frame with respect to 20 action units automatic detection of faces in the video stream was done by applying SVMs and AdaBoost, to texture-based image representations, where the output margin for the learned classifiers predicted AU intensity.

Next step in automatic psychiatric diagnosis is calculation of the correlations among different parts of facial muscles movements based on the information with respect to the changes in facial expressions. Then, we have to compare these correlations with the corresponding correlations observed in normal healthy subjects (Kobayashi et al. 2000). If significant deviation is observed between two groups with respect to these correlations, the difference may be considered as an index of expression of psychopathology.

Figure 10.4 may help design computation models for automatic identification of psychiatric disorder.

Objective coding of emotional expression in diagnosis of psychiatric population through computational modeling also may incorporate:

1. The database to understand emotional expression bias of each diagnostic category, for example, indifference or flat affect in schizophrenia, negative emotion bias (sadness) in depression, or positive emotion bias in manic patients.
2. From the view point of decoding emotion, impairment of decoding emotional expression in autism (McIntosh et al. 2006) or differential perceptual bias for negative emotions in borderline personality disorder (Domes et al. 2008) or in patients with anxiety disorder (Barlow et al. 1996); (Mineka and Zinberg 1996), or in schizophrenia (Borod et al. 1993) suggest not only emotional expression of the patients but decoding of emotional faces of others by patients can also provide an important basis of computational diagnosis.
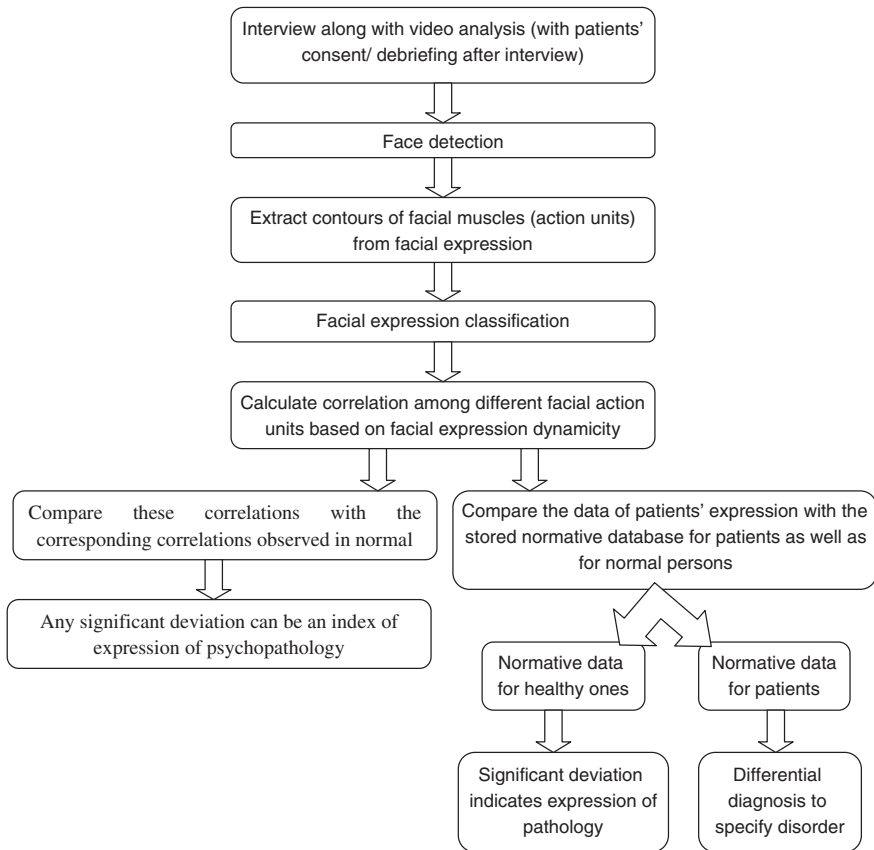
Fig. 10.4 Flowchart based on summarizing computational analytic procedures for psychiatric diagnosis

Following the above perspective, the data base can store three types of information regarding the appraisal of EFE of others by clinical population.

(i)   Nature of appraisal of facial expressions of the psychiatric patients of different psychiatric diagnostic categories under investigation is necessary to understand their emotional valance. These appraisals by the patients could be used as the diagnostic index for them.

(ii) Nature of appraisal of facial expressions of at least six basic emotions of normal population by the psychiatric patients could serve as a reference point, to help us in analyzing patients' vulnerability in appraising basic normative emotional facial expression.

(iii) Nature of appraisal of facial expression, blended with display rules of normal population by the psychiatric patients, helps us in detecting whether patients are able to decode display rules or not.

The development of these normative sets of data for the patients of different diagnostic categories is necessary to understand the unique constellation of characteristics of EFE of each diagnostic category and the differential point of diagnosis of a given psychiatric illness from the other. Consideration of efficiency for monitoring emotional expressions in congruence with the situational demand in the model could be used as an index of better prognosis. Another important issue to be considered is reliability of computational models. In case of computational analysis, the signal value for correct diagnosis is pronounced enough for patients with psychotic symptoms, those who widely deviate from normal pattern of EFE. However, distinguishing the patients with intact reality contact from normal could enhance the probability of error in computational diagnosis. In such cases, reliability analysis is imperative in order to reduce diagnostic errors. This methodological crisis could be overcome by checking the reliability of computational models through signal detection paradigm.

In the context of diagnosis of psychiatric disorders, real-time dynamic facial expression analysis techniques seem to be very helpful. Subtle differentiation is not possible without the assistance of sophisticated computer analysis and mathematical explanation. It is expected in near future that facial expression recognition will become very useful for early diagnosis of mental illness. LBP-based facial image analysis has been one of the most popular methods in recent years. Happy et al. (2012) propose a facial expression classification algorithm that uses Haar classifier for face detection purpose. LBP histogram of different block sizes of a face image as feature vectors classify various facial expressions using PCA. The algorithm is implemented in real time for expression classification since the computational complexity of the algorithm is small. Happy et al. (2012) noted the following algorithm for real-time facial expression analysis.

Training Algorithm (Proposed by Happy et al. 2012) (Fig. 10.5):

(i)   Detect face from the training image using Haar classifier and resize detected face image to N × M resolution.
(ii) Preprocess the face image to remove noise.
(iii) For each class (expression), obtain the feature vectors $\Gamma_{j,1}$, $\Gamma_{j,2}$, …, $\Gamma_{j,p}$ (jth class) of dimension $\left(\frac{M}{n} * \frac{M}{m} * b, 1\right)$ each.
   (a) Divide the face image to subimages of resolution n × m, find the LBP values, and calculate b—bin histogram for each block.
   (b) Concatenate the histograms of each block to get the feature vector $(\Gamma_{j,i})$ of size $\left(\frac{M}{n} * \frac{M}{m} * b, 1\right)$.
(iv) Compute the mean feature vector of individual class $\Psi_j = \frac{1}{P}\sum_{i=1}^{P} \Gamma_{j,1}$, $(j = 1,2,\ldots,6)$.
(v) Subtract mean feature vector from each feature vector $\Gamma_{j,i}$.

$$\varphi_{j,i} = \Gamma_{j,i} - \Psi_j, \quad (j = 1, 2, \ldots, 6).$$

(vi) Estimate the covariance matrix C for each class, given by

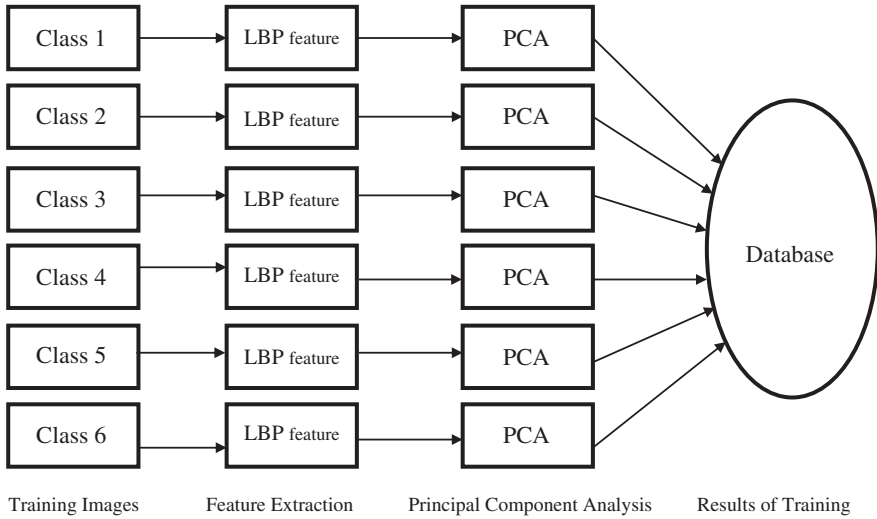$$C_j = \frac{1}{P}\sum_{i=1}^{P} \phi_{j,1}\phi_{j,i}^{T} = A_j A_j^{T} \quad (j = 1, 2, \ldots, 6)$$

Fig. 10.5 Flowchart for training phase proposed by Happy et al. (2012)

where $A_j = [\phi_{j,1}\phi_{j,2}, \ldots, \phi_{j,P}]$ of dimension $\left(\frac{M}{n} * \frac{M}{m} * b \times P\right)$ which is very large. Compute $A_j^T A_j (P \times P)$ instead as $P \ll \frac{M}{n} * \frac{M}{m} * b$.

(vii) Compute the eigenvectors $v_{j,i}$ of $A_j^T A_j$ using the equation

$$\sigma_{j,i}u_{j,i} = A_j v_{j,i} \quad (j = 1, 2, \ldots, 6).$$

(viii) Keep only K eigenvectors corresponding to the K-largest eigenvalues form each class (suppose, $U_j = [u_{j,1}, u_{j,2}, \ldots, u_{j,k}]$).

(ix) Normalize the K eigenvectors of each class.

Algorithm for facial expression detection (proposed by Happy et al. 2012) (Fig. 10.6):

(i) Detect face with Haar classifier algorithm and resize face image to $N \times M$ resolution.

(ii) Preprocess the face image to remove noise.

(iii) Find feature vector (Γ) for the resized face using similar methods as used in training phase (Step 3).

(iv) Subtract the mean feature vector of each class form (Γ)

$$\phi_j = \Gamma - \Psi_j \quad (j = 1, 2, \ldots, 6)$$

(v) Project the normalized test image onto the eigen directions of each class and obtain weight vector

$$W = [w_{j,1}, w_{j,2}, \ldots, w_{j,k}] = U_j^T \phi_j \quad (j = 1, 2, \ldots, 6)$$
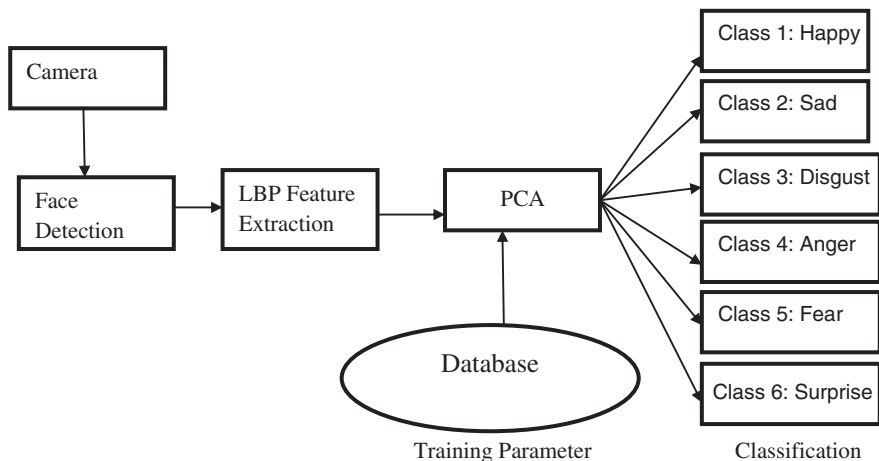
Fig. 10.6 Flowchart for testing phase proposed by Happy et al. (2012)

(vi) Compute $\cap\phi_j = \sum_1^k w_{j,i}u_{j,i}$ $(j = 1, 2, \ldots, 6)$
(vii) Compute error $e_j = \|\phi - \cap\phi_j\| j = 1, 2, \ldots, 6.$

The image is classified to the training set, to which it is closest (when the reconstruction error (e_j) is minimum).

It is envisioned that this algorithm will be very useful for diagnosis of psychiatric disorders.

10.5 Conclusion and Future Direction

Present review is an attempt to explore the possible methodologies of computational modeling of emotional facial expressions that may be finally developed into an objective psychiatric diagnostic tool in terms of fully automated facial action detection system of spontaneous facial expressions. The accuracy of automated facial expression measurement in spontaneous behavior may also be considerably improved by 3D alignment of faces. Bartlett et al. (2006) in his work with feature selection by AdaBoost though significantly enhanced both speed and accuracy of SVMs, but its application is still restricted to the field of recognition of basic emotions only. Expansion of its applicability in the task of AU detection in spontaneous expressions could be an important task in future. The computational analysis is capable of bringing about paradigmatic shifts in the field of psychiatric diagnosis by making facial expression more accessible as a behavioral measure and also may enrich the understanding of emotion, mood regulation, and social communication in the field of advanced cognitive neuroscience. We believe that a focused, interdisciplinary program directed toward computer understanding of human behavioral patterns (as shown by means of facial expressions and other modes of social interaction) should be established in order to achieve a major breakthrough.

References

Anstadt, T., & Krause, R. (1989). The expression of primary affects in portraits drawn by schizophrenics. *Psychiatry, 52*, 13–24.

Arguedas, D., Stevenson, R. J., & Langdon, R. (2012). Source monitoring and olfactory hallucinations in schizophrenia. *Journal of Abnormal Psychology, 121*(4), 936–943.

Barlow, D. H., Chorpita, B., & Turovsky, J. (1996). Fear, panic, anxiety, and disorders of emotion. In D. Hope (Ed.), *Perspectives on anxiety, panic and fear* (pp. 251–328). 43rd Annual Nebraska Symposium on Motivation. Lincoln: University of Nebraska Press.

Bartlett, M. S., Littlewort, G. C., Frank, M. G., Lainscsek, C., Fasel, I. R., & Movellan, J. R. (2006). Automatic recognition of facial actions in spontaneous expressions. *Journal of multimedia, 1*(6), 22–35.

Bell, C. (Ed.) (1844) *Anatomy and philosophy of expression as connected with the fine arts* (7th Rev. ed.). London: G. Bell & Sons.

Belhumeur, P. N., Hespanha, J. P., & Kriegman, D. J. (1997). Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 19*(7), 711–720.

Benedetti, F., Anselmetti, S., Florita, M., Radaelli, D., Cavallaro, R., Colombo, C., et al. (2005). Reality monitoring and paranoia in remitted delusional depression. *Clinical Neuropsychiatry, 2*(3), 199–205.

Benson, P. J. (1999). A means of measuring facial expressions and a method for predicting emotion categories in clinical disorders of affect. *Journal of Affective Disorders, 55*(2–3), 179–185.

Black, M. J., & Yacoob, Y. (1995). A tracking and recognizing rigid and non-rigid facial motions using local parametric models of image motions. In *Proceedings of International Conference on Computer Vision* (pp. 374–381).

Black, M. J., & Yacoob, Y. (1997). A recognizing facial expressions in image sequences using local parameterized models of image motion. *International Journal of Computer Vision, 25*(1), 23–48.

Borod, J. C., & Koff, E. (1984). Asymmetries in affective expression: Behaviour and anatomy. In N. Fox & R. Davidson (Eds.), *The psychobiology of affective development* (pp. 293–323). Hillsdale, NJ: Erlbaum.

Borod, J. C., & Koff, E. (1991). Facial asymmetry during posed and spontaneous expressions in stroke patients with unilateral lesions. *Pharmacopsychoecologia, 4*, 15–21.

Borod, J. C., Martin, C. C., Alpert, M., Brozgold, A., & Welkowitz, J. (1993). Perception of facial emotion in schizophrenics and right brain-damaged patients. *Journal of Nervous and Mental Disorders, 181*, 494–502.

Borod, J. C., Haywood, C. S., & Koff, E. (1997). Neuropsychological aspects of facial asymmetry during emotional expression: A review of the normal adult literature. *Neuropsychology Review, 7*, 41–60.

Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. In *COLT '92 Proceedings of the 5*th *annual Workshop on Computational learning theory* at NY, USA (pp. 144–152).

Campbell, R. (1978). Asymmetries in interpreting and expressing a posed facial expression. *Cortex, 14*, 327–342.

Campbell, R. (1979). Left handers' smiles: Asymmetries in the projection of a posed expression. *Cortex, 14*, 571–579.

Cohn, J. F., Zlochower, A. J., Lien, J. J., & Kanade, T. (1998). A feature-point tracking by optical flow discriminates subtle differences in facial expression. In *Proceedings of International Conference Automatic Face and Gesture Recognition* (pp. 396–401).

Cootes, T. F., Edwards, G. J., & Taylor, C. J. (1998). Active appearance models. In *Proceedings of the European Conference on Computer Vision* (Vol. 2, pp. 484–498).

Darwin, C. (1872/1998). *The expression of the emotions in man and animals*. New York: Oxford University Press.

Dimberg, U. (1982). Facial reactions to facial expressions. *Psychophysiology, 19*, 643–647.

Dimberg, U. (1997). Psychophysiological reactions to facial expressions. In U. Segerstrale & P. Molnar (Eds.), *Nonverbal communication: Where nature meets culture* (pp. 47–60). Mahwah, NJ: Erlbaum.

Domes, G., Czieschnek, D., Weidle, F., Berger, C., Fast, K., & Herpertz, S. C. (2008). Recognition of facial affect in borderline personality disorder. *Journal of Personality Disorders, 22*(2), 135–147.

Duchenne de Boulogne, G. B. (1862/1990). *The mechanism of human facial expression.* New York: Cambridge University Press.

Edwards, G. J., Cootes, T. F., & Taylor, C. J. (1998). A face recognition using active appearance models. In *Proceedings of the European Conference on Computer Vision* (Vol. 2, pp. 581–695).

Ekman, P. (1982). *Emotion in the human face.* Cambridge University Press.

Ekman, P. (1992a). Are there basic emotions? *Psychological Review, 99*, 550–553.

Ekman, P. (1992b). An argument for basic emotions. *Cognition and Emotion, 6*, 169–200.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in face and emotion. *Journal of Personality and Social Psychology, 17*, 124–129.

Ekman, P., & Friesen, W. V. (1975). *Unmasking the face*. Englewood Cliffs, N.J: Prentice-Hall.

Ekman, P., & Friesen, W. V. (1978). *Facial Action Coding System (FACS): Manual.* Palo Alto: Consulting Psychologists Press.

Ekman, P., Sorenson, E. R., & Friesen, W. V. (1969). Pancultural elements in facial displays of emotion. *Science, 164*, 86–88.

Essa, I., & Pentland, A. (1997). A coding, analysis interpretation, recognition of facial expressions. *IEEE Trans. Pattern Analysis and Machine Intelligence*, *19*(7), 757–763.

Fasel, B., & Luettin, J. (2003). Automatic facial expression analysis: A survey. *Pattern Recognition, 36*(1), 259–275.

Gessler, S., Cutting, J., Frith, C. D., & Weinman, J. (1989). Schizophrenics' inability to judge facial emotion: A controlled study. *British Journal of Clinical Psychology, 28*(1), 19–29.

Girard, J. M., Cohn, J. F., Mahoor, M. H., Mavadati, S., & Rosenwald, D. P. (2013). Social risk and depression: Evidence from manual and automatic facial expression analysis. In *IEEE International Conference on Automatic Face & Gesture Recognition*. Retrieved from http://www.pitt.edu/~jmg174/Girard%20et%20al%202013.pdf.

Gruber, J., Oveis, C., Keltner, D., & Johnson, S. L. (2008). Risk for mania and positive emotional responding: Too much of a good thing? *Emotion, 8*(1), 23–33.

Hall, J., Harris, J. M., Sprengelmeyer, R., Sprengelmeyer, A., Young, A. W., Santos, I. M., et al. (2004). Social cognition and face processing in schizophrenia. *British Journal of Psychiatry, 185*, 169–170.

Happy, S. L., George, A., & Routray, A. (2012). A real time facial expression classification system using local binary patterns. In *IEEE Proceedings of 4th International Conference on Intelligent Human Computer Interaction*, Kharagpur, India, December, 27–29.

Hermans, D., Martens, K., De Cort, K., Pieters, G., & Eelen, P. (2003). Reality monitoring and metacognitive beliefs related to cognitive confidence in obsessive-compulsive disorder. *Behaviour Research and Therapy, 41*(4), 383–401.

Hong, H., Neven, H., & von der Malsburg, C. (1998). An online facial expression recognition based on personalized galleries. In *Proceedings of the. International Conference on Automatic Face and Gesture Recognition* (pp. 354–359).

Hooker, C., & Park, S. (2002). Emotion processing and its relationship to social functioning in schizophrenia patients. *Psychiatry Research, 112*, 41–50.

Huang, C. L., & Huang, Y. M. (1997). A facial expression recognition using model-based feature extraction and action parameters classification. *Journal of Visual Communication and Image Representation, 8*(3), 278–290.

Izard, C. E. (1971). *The face of emotion*. New York: Appleton-Century-Crofts.

Izard, C. E. (1979). *The maximally discriminative facial movement coding system (MAX)*. Newark, Del: University of Delaware, Instructional Resource Center.

Kanade, T. (1973). *Picture processing system by computer complex and recognition of human faces*. Unpublished Ph.D. dissertation, Tokyo University, Tokyo, Japan.

Kato, M., So, I., Hishinuma, Y., Nakamura, O., & Minami, T. (1992). Description and synthesis of facial expression based on isodensity maps. In L. Tosiyasu (Ed.), *Visual computing* (pp. 39–56). Tokyo: Springer.

Keltner, D., & Ekman, P. (2000). Facial expression of emotion. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (pp. 236–249). New York, USA: Guilford Press.

Kimura, S., & Yachida, M. (1997). A facial expression recognition and its degree estimation. In *Proceedings of Computer Vision and Pattern Recognition* (pp. 295–300).

Kline, J. S., Smith, J. E., & Ellis, H. C. (1992). Paranoid and nonparanoid schizophrenic processing of facially displayed affect. *Journal of Psychiatric Research, 26*, 169–182.

Kobayashi, H., & Hara, F. (1992). Recognition of six basic facial expressions and their strength by neural network. *Proc. Int'l Workshop Robo, H.tand Human Comm.*, 381–386.

Kobayashi, H., & Hara, F. (1997). Facial interaction between animated 3D face robot and human beings. In *Proceedings of International Conference Systems, Man, Cybernetics* (Vol. 3, pp. 732–3,737).

Kobayashi, H., Takahashi, H., Kimura, T., Kikuchi, K., & Tazaki, M. (2000). Study on automatic diagnosis of psychiatric disorder by facial expression. In *26th Annual Conference of the IEEE on Industrial Electronics Society* at Nagoya, Japan (Vol. 1, pp. 487–492).

Kobayashi, H., Tange, K., & Hara, F. (1995). Realtime recognition of six basic facial expressions. In *Proceedings of the IEEE International Workshop on Robot and Human Communication*, Tokyo (pp. 179–186).

Kornreich, C., Philippot, P., Foisy, M. L., Blairy, S., Raynaud, E., Dan, B., et al. (2002). Impaired emotional facial expression recognition is associated with interpersonal problems in alcoholism. *Alcohol and Alcoholism, 37*(4), 394–400.

Ladouceur, C. D., Dahl, R. E., Birmaher, B., Axelson, D. A., & Ryan, N. D. (2006). Increased error-related negativity (ERN) in childhood anxiety disorders: ERP and source localization. *Journal of Child Psychology and Psychiatry, 47*(10), 1073–1082.

Lam, K. M., & Yan, H. (1998). An analytic-to-holistic approach for face recognition based on a single frontal view. *IEEE Trans. Pattern Analysis and Machine Intelligence*, *20*(7), 673–686.

Lewis, S. F., & Garver, D. L. (1995). Treatment and diagnostic subtype in facial affect recognition in schizophrenia. *Journal of Psychiatric Research, 29*, 5–11.

Li, S., & Gu, L. (2001). Real-time multi-view face detection, tracking, pose estimation, alignment, and recognition. In *IEEE Conference on Computer Vision and Pattern Recognition Demo Summary*.

Lien, J. J., Kanade, T., Cohn, J. F., & Li, C. C. (1998). A automated facial expression recognition based on FACS action units. In *Proceedings of International Conference Automatic Face and Gesture Recognition* (pp. 390–395).

Liu, C. (2003). A Bayesian discriminating features method for face detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 25*(6), 725–740.

Lucey, P., Cohn, J., Prkachin, K., Solomon, P., & Matthews, I. (2011). Painful data: The UNBC-McMaster shoulder pain expression archive database. In *International Conference on Automatic Face and Gesture Recognition and Workshops, IEEE* (pp. 57–64).

Lucey, P., Cohn, J. F., Prkachin, K. M., Solomon, P. E., Chew, S., & Matthews, I. (2012). Painful monitoring: Automatic pain monitoring using the UNBC-McMaster shoulder pain expression archive database. *Image and Vision Computing, 30*, 197–205.

Lyons, M. J., Budynek, J., & Akamatsu, S. (1999). Automatic classification of single facial images. *IEEE Trans. Pattern Analysis and Machine Intelligence*, *21*(12), 1357–1362.

Mandal, M. K., & Ambady, N. (2004). Laterality of facial expressions of emotion: Universal and culture-specific influences. *Behavioral Neurology, 15*, 23–34.

Martin, C. C., Borod, J. C., Alpert, M., Brozgold, A., & Welkowitz, J. (1990). Spontaneous expression of facial emotion in schizophrenic and right brain damaged patients. *Journal of Communication Disorder, 23*, 287–301.

McIntosh, D. N., Reichmann-Decker, A., Winkielman, P., & Wilbarger, J. L. (2006). When the social mirror breaks: Deficits in automatic, but not voluntary, mimicry of emotional facial expressions in autism. *Developmental Science, 9*, 295–302.

Mineka, S., & Zinbarg, R. (1996). Conditioning and ethological models of anxiety disorders: Stress-in-dynamic context anxiety models. In D. Hope (Ed.), *Perspectives on anxiety, panic and fear: Nebraska symposium on motivation*. Lincoln: University of Nebraska Press.

Otsuka, T., & Ohya, J. (1996). A recognition of facial expressions using HMM with continuous output probabilities. In *Proceedings of International Workshop Robot and Human Communication* (pp. 323–328).

Padgett, C., & Cottrell, G. W. (1996). A representing face images for emotion classification. In *Proceedings of Conference Advances in Neural Information Processing Systems* (pp. 894–900).

Pantic, M., & Bartlett, M. S. (2007). Machine analysis of facial expressions. In K. Delac & M. Grgic (Eds.), *Face Recognition* (pp. 377–416). Vienna, Austria: I-Tech education and publishing.

Pantic, M., & Rothkrantz, L. J. M. (2000a). Expert system for automatic analysis of facial expression. *Image and Vision Computing, 18*(11), 881–905.

Pantic, M., & Rothkrantz, L. J. M. (2000b). Automatic analysis of facial expressions: the state of the art. *IEEE Trans. of Pattern Analysis and Machine intelligence*, *22*(12), 1424–1445.

Pentland, A., Moghaddam, B., & Starner, T. (1994).View-based and modular eigenspaces for face recognition. In *Proceedings of IEEE Conference Computer Vision and Pattern Recognition* (pp. 84–91).

Prkachin, K. M., Hughes, E., Schultz, I., Joy, P., & Hunt, D. (2002). Real-time assessment of pain behavior during clinical assessment of low back pain patients. *Pain, 95*, 23–30.

Radaelli, D., Benedetti, F., Cavallaro, R., Colombo, C., & Smeraldi, E. (2013). The reality monitoring deficit as a common neuropsychological correlate of schizophrenic and affective psychosis. *Behavioural Sciences, 3*(2), 244–252.

Remillard, G. M., Andremann, F., Rhi-Suci, A., & Robbins, N. M. (1977). Facial asymmetry in patients with temporal lobe epilepsy. *Neurology, 27*, 109–114.

Rinn, W. E. (1984). The neuropsychology of facial expression: A review of the neurological and psychological mechanisms for producing facial expressions. *Psychological Bulletin, 95*, 52–77.

Rowley, H., Baluja, S., & Kanade, T. (1998). Neural network-based face detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 20*(1), 23–38.

Sackeim, H. A., Gur, R. C., & Saucy, M. C. (1978). Emotions are expressed more intensely on the left side of the face. *Science, 202*, 434–436.

Sackett, G. P. (1966). Monkeys reared in isolation with pictures as visual input: Evidence for an innate releasing mechanism. *Science, 154*, 1468–1473.

Samal, A., & Iyengar, P. A. (1992). Automatic recognition and analysis of human faces and facial expressions: A survey. *Pattern Recognition, 25*(1), 65–77.

Schneider, F., Koch, J. D., Mattes, R., & Heimann, H. (1992). Recognition of emotions in facial expression of visual half-field images by schizophrenic and depressed patients. *Nervenarzt, 63*, 545–550.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic information processing: 1. Detection, search and attention. *Psychological Review, 84*, 1–66.

Schneiderman, H., & Kanade, T. (2000). A statistical model for 3d object detection applied to faces and cars. In *IEEE Conference on Computer Vision and Pattern Recognition*.

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.

Shenoy, A. (2009). *Computational analysis of facial expressions.* Unpublished Ph.D. Thesis, School of Computer Science, University of Hertfordshire.

Simoncelli, E. (1993). *A distributed representation and analysis of visual motion*. Unpublished PhD thesis, Massachusetts Institute of Technology.

Steffens, J., Elagin, E., & Neven, H., (1998). A person spotter fast and robust system for human detection, tracking, and recognition. In *Proceedings of International Conference on Automatic Face and Gesture Recognition* (pp. 516–521).

Steimer-Krause, E., Krause, R., & Wagner, G. (1990). Interaction regulation used by schizophrenic and psychosomatic patients: Studies on facial behavior in dyadic interactions. *Psychiatry, 53*, 209–228.

Stuhrmann, A., Suslow, T., & Dannlowski, U. (2011). Facial emotion processing in major depression: A systematic review of neuroimaging findings. *Biology of Mood and Anxiety Disorders,*. doi:10.1186/2045-5380-1-10.

Tassinary, L. G., Cacioppo, J. T., & Geen, T. R. (1989). A psychometric study of surface electrode placement for facial electromyography recording: 1. the brow & cheek muscle regions. *Psychophysiology, 26*, 1–16.

Tomkins, S. S. (1962). *Affect, imagery and consciousness. The positive effects*. New York: Springer.

Uljarevic, M., & Hamilton, A. (2013). Recognition of emotions in autism: A formal meta-analysis. *Journal of Autism and Developmental Disorders, 43*(7), 1517–1526.

Van Gelder, R. (1981). *Neuroanatomy and facial asymmetry*. Internal publication of the ILO, Department of Psychology, Free University of Amsterdam.

Viola, P. & Jones, M. (2001). Rapid object detection using boosted cascade of simple features. *IEEE Conference on Computer Vision and Pattern Recognition*.

Walker, E. F., Grimes, K. E., Davis, D. M., & Smith, A. J. (1993). Childhood precursors of schizophrenia: Facial expressions of emotion. *American Journal of Psychiatry, 150*, 1654–1660.

Wang, M., Iwai, Y., & Yachida, M. (1998). A expression recognition from time-sequential facial images by use of expression change model. In *Proceedings of International Conference Automatic Face and Gesture Recognition* (pp. 324–329).

Wang, P., Barrett, F., Martin, E., Milonova, M., Gurd, R., Gurb, R., et al. (2008). Automated video-based facial expression analysis of neuropsychiatric disorders. *Journal of Neuroscience Methods, 168*, 224–238.

Wehrle, T., & Scherer, K. R. (2001). Towards computational modeling of appraisal theories. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotions: Theory, methods, research* (pp. 350–365). New York: Oxford University Press.

Wiskott, L. (1995). *Labelled graphs and dynamic link matching for face recognition and scene analysis*. Verlag Harri Deutsch, Thun-Frankfurt am Main, Reihe Physik, p. 53.

Wolff, W. (1943). *The expression of personality*. New York, NY: Harper.

Yoneyama, M., Iwano, Y., Ohtake, A., & Shirai, K. (1997). A facial expressions recognition using discrete hopfield neural networks. In *Proceedings of International Conference on Information Processing* (Vol. 3, pp. 117–120).

Yuille, A. L., Cohen, D. S., & Hallinan, P. W. (1989). A feature extraction from faces using deformable templates. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (pp. 104–109).

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35*, 151–175.

Zhang, Z., Lyons, M. J., Schuster, M., & Akamatsu, S. (1998). Comparison between geometry-based and Gabor-wavelets-based facial expression recognition using multi-layer perceptron. In *IEEE International Conference on Automatic Face & Gesture Recognition (FG)*, Nara, Japan, pp. 454–459.

Zhao, J., & Kearney, G. (1996). A classifying facial emotions by backpropagation neural networks with fuzzy inputs. In *Proceedings of the Conference on Neural Information Processing* (Vol. 1, pp. 454–457).