ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2015

Emotion Recognition-A Selected Review

Manoj Prabhakaran.K[1], Manoj Kumar Rajagopal[2]

Research Associate, School of Electronic Engineering, VIT University, Chennai Campus, Tamil Nadu, India [1]

Associate Professor, School of Electronic Engineering, VIT University, Chennai Campus, Tamil Nadu, India [2]

ABSTRACT: In Recent years, the research interest to automated analysis of human affective behaviour has attracted increasing attention in psychology, bio-medical, human computer interaction, linguistics, neuroscience, computer science and related disciplines. Since emotions plays an important role in the daily life of human being, the need and importance of automatic emotion recognition has grown with increasing role of human computer interaction applications. Emotion places major role of brain cognition for medically monitoring patients emotional states or stress levels, creating Animated movies, monitoring automobile drivers alertness / drowsiness levels, Human-Robot interaction, Advanced Gaming which enabling very young children to interact with computers, designing techniques for forensic identification, etc. This paper survey from the published work papers, since 2013 till date. The paper surveys about the six prototype expression in human through the various modes of extracting emotion. The paper detailed about that the face detection and tracking, features extraction mechanisms and the technique used for expression classification by the face modelling of state-of-the-art classification. The paper also discussed on multimodal expression classifier of human, i.e. from body and face. Notes also presented the tracking human through sensor and ends with the challenges and future with intention of helping students and researchers who are new to this field.

KEYWORDS: Modes of extracting emotion, expression recognition, face modelling, state-of-the-art, basic six emotion feature tracking, body tracking, and Microsoft Kinect sensor.

I.INTRODUCTION

In early Aristotelian era (4th century BC) study on the Physiognomy and Facial Expression. [1] Physiognomy is the assessment of a person personality or character from their outer appearance, especially face. But over the year, Physiognomy has been not interested, but facial expression continuously been active topic. From the foundational studies that formed the basis of today's research on facial expression can be tracked to 17th century. In 1649, John Bulwer in his book "Pathomyotmia" was given details of muscle movement of head and various expressions in human. [2] In 1667, Le Brun's lectures the physiognomy; later reproduce the book in 1734 [3]. In 18th century actors and artist referred to his book to achieve "the perfect imitation of 'genuine' facial expression". In 19th century, a facial expression analysis has direct relationship to the automatic face analysis and face expression recognition this was brought by Darwin. [4] In 1894, Darwin states that principles of expression in both humans and animals also groped the various kind of expression and cataloged the facial deformation. [5] In 1884, William James states "James Lange theory" the all emotion are derived from the presence of stimulus in body which evokes the physiological response.

[6]Another important study of facial expression analysis and human emotion is done by the Paul Ekman in 1967. He developed the facial expression recognizer, and analyst the facial expression muscular movement showed that different emotion in human through the photographic stimuli. In 1969, [7]Wallbott and Scherer determined emotion from the body movement and speech. In 1978, Suwa established the automatic recognition of facial expression system for analysis the facial expression from a sequence of image of tacking points. Although this system might not be clearly seen the tracking points till 1990s. Later 1992, Samal and Iyengar presented the Facial Feature and Expression analysis of tacking points in movie frames and they also states that the automatic recognition of facial expression requires robust face detection and face tracking system Early and late of 1990, led to the development of robotic face detection, face tracking, face expression analysis and at the same time become very popularity of fields are the Human –computer Interaction (HCI), Affective Computing, etc.,. Since 1990s, researchers are started to develop of interest on automatic facial expression recognition become very active. From 1990 to till now, researchers are mostly concentrating on automatic facial expression recognition and emotion in human using various modes of extraction

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2015

Therefore we will be concentrating in our work only on the published work from these bases of expression recognition. The remainder of the paper will be organized as follows: Section 2 mention the applications of automatic facial expression recognition systems, Section 3 covers the literature survey of the modes of extracting emotion, Section 4 covers the literature survey of automatic face analysis and Face model of state-of-the art, Section 5 covers the literature survey of Body tracking of expression, Section 6 mentions the Comparison of techniques and discussion, Section 7 mentions the Future work and Challenges and paper concludes with section 8.

II. APPLICATION

The uses of automatic face expression system can be found the application in several interesting field [8]. In Animation, expression recognition system mostly used to create 3d model which applying into a real time application for making animation movies like polar express, avatar, etc., and video gaming. In medical applications, the expression recognition system is used to monitor the emotion states of patient (like autism, epilepsy, and schizophrenic) with dementia to try to detect the depression case. In human robot interaction, the Sony's Aibo Robot and TR'S Robo Vie are the example of practically real time application which develops an animated character from the face expression recognition system that mirrors the expression of the user called CU Animate.

Apart from these applications, expression recognition system finds use in other domains like Telecommunication, Behavioral Science, Psychiatry, Automobile Safety, Affective sensitive music juke box and televisions, Educational Software, etc., These are applications and uses of expression recognition system becomes more real time and robust.

III.LITEATURE SURVEY OF THE MODES OF EXTRACTING EMOTION:

The modes of extracting emotion [9] in human through two methods are non-physiological signal and physiological signal. The non-physiological signal is observes emotion from their appearance and the physiological signal is observes the emotion from inside of human through physiological response and these methods are following bellow. The figure 1 represents the various modes of extracting emotion.

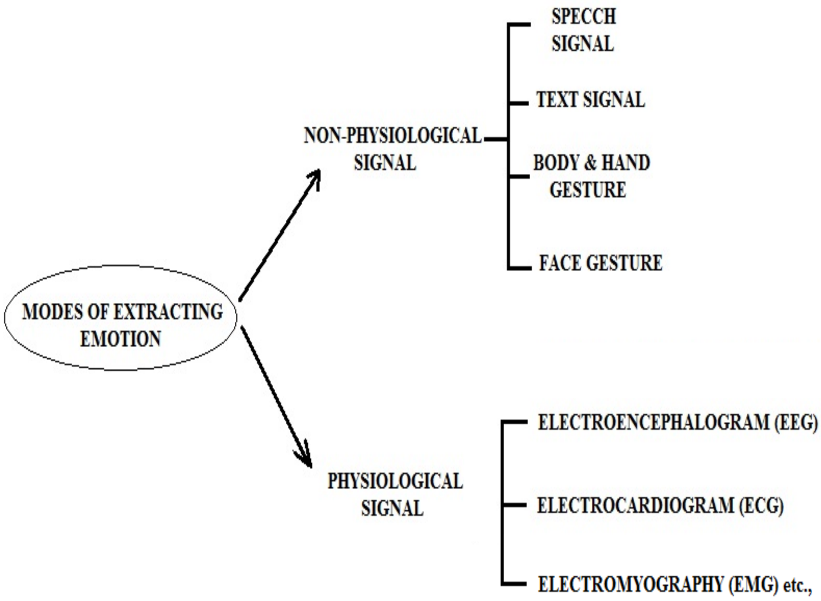

Figure 1: modes of extraction emotion [9].

1. NON-PHYSIOLOGICAL SIGNAL:

1.1. Speech signal:1.Tin Lay New, Say Wei Foo, Liyanage C. De Silva, "**Speech emotion recognition using hidden Markov models",** 2003 [10] , proposes an emotion recognition from the speech signal using methods of short time log frequency power coefficients(LFPC) to represent the speech signals and a discrete hidden Markov model (HMMs). The

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2015

Hidden Markov Model (HMMs) as the classifier of emotion from speech signals is compared with that of the linear prediction Cepstral coefficients (LPCC. Mel-frequency Cepstral coefficients (MFCC) feature parameters commonly used for speech recognition systems to compare the LFPC feature parameters performances. In the proposes system speech signal is preprocessing the signal and feature vector extraction, the selected feature based on training phase and recognition phase in LPFC and compared to LPCC and MFCC. For reduce computational used the vector quantization and finally speech recognition of six basic emotional states are estimated by using Hidden Markov model (HMM).

2.Nicholson, K. Takahashi and R. Nakatsu, **"Emotion Recognition in Speech Using Neural Networks",** 2000 [11], proposes an emotion recognition system by using the Neural Network of One-class-in-one network model (OCIO) form the speech signal. First, the speech processing part which the speech signal is calculated the utterance by speech features, then speech power extraction is compared with the predetermined power threshold which chosen the phonetic features and prosodic features of speech which extracted the speech power, pitch and the Linear predictive coding(LPC) is expressed the time variable feature of speech spectrum. The speech feature extraction is extracted from the voice utterance which compared the speech power spectrum of start point and end point of threshold. The Neural network architecture is a One-class-in-one network model (OCIO) is used for emotion recognition from speech signal. In this, the feature vector is input is given to the sub-neural network is the back propagation algorithm examined the training and testing epoch of utterance. Finally, the decision logic output is selected the best emotion recognition of speech signal.

1.2 Speech and Text signal: Ze-Jing Chuang and Chung-Hsien Wu **"Multi-Modal Emotion Recognition from Speech and Text", 2004**, [12], proposes emotion recognition is extracted from speech and text analysis. First, the acoustic feature vector is form based on Principal component analysis (PCA), which recognition emotion from the speech. In the Acoustic feature vector are estimated the four basic feature are pitch, energy, formant1 (F1) and zero crossing rate. According to the Hyper plane determined by the training corpuses for emotion from the acoustic feature vector is fed to the Support vector Machine (SVM). Support vector machine is detected the emotional state from acoustical and textual. The textual emotion recognition module is performed as an emotional descriptor which detects the emotional keyword appearances. The Front end process is the keyword spotting system used to transformed the speech signal to textual, by Hidden Markov model (HMm) are performed the keyword spotting and Mel-Frequency Cepstrum coefficient (MFCC) is extracted the an acoustic feature .the emotion modification value is defined for each emotion modification word in sentence. Finally, combining of these three output defined the multi-modal emotion from the acoustical and textual.

1.3 Body gesture: Garber Barron, Michael, and Mei Si, "**Using body movement and posture for emotion detection in non-acted scenarios**", 2012, [13] proposes the four emotion from the body movement and posture observed when people are playing video games. In this, estimates the body features consists of seven types of cues. Using database and recorded data determined the four emotion of human. Finally, they are applying a machine learning algorithm for to create the four features group for emotion recognition in human.

1.4. Face gesture: I. Kotsia and I. Pitas, **"Facial Expression Recognition in Image Sequences Using Geometric Deformation Features and Support Vector Machines"**, 2007, [14] proposes the two novels method for face expression using geometric deformable grid model. In this paper, the grid node information extraction is performed manually fit the face wireframe model on landmark of image sequence. For the face feature tracking using their Kanade-Lucas-Tomasi tracker (KLT) achieved robustness and accuracy. Finally, the grid node information classification is estimated the basic six emotion of human using by multiclass Support Vector Machine (SVM) and Face Action Coding System (FACS) of deformable grid node

1.5 Multimodal (combination of body, face, and speech): G. Caridakis, G. Castellano, L. Kessous, A. Raouzaiou, L. Malatesta, S. Asteriadis and K. Karpouzis, **"Multimodal emotion recognition from expressive faces, body gestures and speech"**, 2007, [15] analyze the emotion in human from multimodal fusion of face, body gesture and speech. In this examined unimodal systems to multimodal fusion. Initially, estimate the face feature point extraction and combination of the anthropometric, Facial Animation Parameter (FAPs) to determine the face expression recognition. In body gesture, the silhouette and hand blobs of human extract the expression action using Eyes Web Expressive Gesture Processing Library tracker. In speech, feature extraction is based on intensity, pitch, MFCC (Mel Frequency Cepstral Coefficient), Bark spectral bands, voiced segment characteristics and pause length. To fuse the three unimodal

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2015

to get the multimodal fusion based on two approaches implemented: feature level and decision level. In feature level three modalities are examined the single Bayesian classifier with feature and decision level, combine the output of classifiers of modality. Finally the probabilities of output of the decision level fusion recognizes the expression in human through multi fusion modal

2. PHYSIOLOGICAL SIGNAL:

2.1 Electroencephalogram (EEG): 1.Teresa K. W. Wong, Peter C. W. Fung, Siew E. Chua and Grainne M. McAlonan, **"Abnormal spatiotemporal processing of emotional facial expressions in childhood autism: dipole source analysis of event-related potentials"**,2008, [16] In Autism spectrum disorder children are previous studies investigated of the neural system for face perception and emotion recognition in adults and young children suggested abnormalities in anatomical ,development ,functioning and connectivity of distributed brain system involved in social cognition. But precise the spatial sequence and time course of rapid neural response (i.e. spatiotemporal neural system) specific to the processing of emotional facial expression of the young children and adult have not been examined. In the proposes, a spatiotemporal processing of emotional facial expression in childhood autism using dipole source analysis of event –related potentials and provides future development basic studies and compare with the typically developing control (TD). First, present the video pictures showing the six basic emotion or neutral facial expression to healthy adolescent and recorded 128-channel event related potential (ERPs) from the young and adult autism (aged 6-10 years), and parameters (age, sex, IQ) are matched with Typically developing during implicit and explicit processing of emotion. Second, the electroencephalogram (EEG) read the signal from autism and typically control where corrected the eye blink pattern by surrogate model patterns and rejected the head movement, muscle artifacts and reduced the slow channel drift to give the implicit baseline correction by the 1-Hz high –pass forward filter. Then the result of each individual EEG averaged was transformed to an average reference and for ERPs analysis and dipole source analysis (BESA) filtered by low pass filter at 30Hz.Selected electrode groups for Event related potentials (ERPs) analysis is measured P1, N170, P2 components from scalp regions corresponding to sites in high density face elicited ERPs studies and scalp region of interest are identified by individual data and visual inspection of grand average. ERPs score is analyzed the effect of age and extracted the individual peak amplitude is determined from the stimulus condition and grand average ERPs across groups and latencies. Whenever the sphericity assumption was violated is used by the Greenhouse-Geisser-corrected degree of freedom. The Dipole source analysis is also known as Brain Electrical Source Analysis (BESA) was used to localize the scalp ERPs cortical source and model their equivalent spatiotemporal current dipole with certain orientation and time-varying dipole moments. The majority of the variance in ERPs waveform explained sufficiently of three dipole source pair showed by Principal Component Analysis (PCA). Finally dipole source analysis (BESA) of high-density of Event related potential (ERPs) is examined both spatial sequence location and temporal profile of early electrical brain waves source activity in response to emotional facial salient stimuli.

2. Panagiotis C. Petrantonakis and Leontios J. Hadjileontiadis, **"Emotion Recognition from Brain Signals Using Hybrid Adaptive Filtering and Higher Order Crossings Analysis"**, 2010, [17], proposes a new feature extraction methods for a user–independent emotion system from electroencephalogram using Hybrid Adaptive Filtering and Higher Order Crossing Analysis. Initially, emotion elicitation process is based on Mirror Neuron system for face -expression image projection. The electroencephalogram (EEG) related the six basic emotions from the healthy subjects. Hybrid Adaptive Filtering (HAF) is a filtering procedure, for an efficient extraction of the emotion related human brain waves characteristic was developed by applying Empirical Model Decomposition with Intrinsic Mode Functions (EMD-IMFs) to the Genetic Algorithm (GA). The EMD-IMF extraction is used to selection of modes to automatic and signal-dependent time-variant filtering and Genetic algorithm (GA) was developed for optimized selection of modes and fitness function for GA selection using Energy-Based Fitness function (EFF) and One Fractal Dimension-Based Fitness (FDFF). The Hybrid Adaptive Filtering (HAF) output is selecting the IMFs produced reconstruction process of EEG signal (R-case) or without employing any reconstruction process (NR-case) is used to get input of Higher Order Crossing Analysis (HOC). The HOC analysis was estimate the feature extraction process of HAF-filtered signal. In Higher Order Crossing analysis is specified that spectrum related attitude of the signal and dependent on power of a certain frequency which analyzed time domain signal and employed without spectral transforms. By Dominant frequency principle, defined the relationship between the zero crossing and spectrum frequency in HOC-based perspective and HOC-Based feature vector. From the HOC feature vector extraction technique, it become robust and

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2015

efficient feature vector for emotion recognition. Finally, from HAF-HOC result to estimate a robust emotion recognition performance from the four different classifications are Quadratic Discriminant Analysis (QDA), k-Nearest Neighbor (k-NN), Mahalanobis Distance (MD) and Support Vector Machine (SVM).

3.Marini Othman and Abdul Wahab, **"Affective face processing analysis in autism using electroencephalogram"**, 2010, [18] they analyzes the affective face processing of Autism Spectrum Disorder people using by human brain waves activity. In the proposes, the pattern classifications are explaining the 2-dmensional emotion model, which defines the various human emotion in terms of pleasant/unpleasantness (valence) and intensity (arousal). The structure of emotional stimuli for ASD individual using movie clips which displayed in front of it, detect the signal from brain waves through electroencephalogram (EEG). The Feature extraction approach is employ by the Mel Frequency Cepstral Coefficient (MFCC). For pattern classification technique is based on Multilayer Perception (MLP) is employ for identified the emotion of the subject's. It's used a combination of several perceptions layer that interconnected to each other and exhibited a high degree of connectivity determined by the network synapses.

IV.LITERATURE SUVEY OF THE AUTOMATIC FACE ANALYSIS AND FACE MODEL

THE AUTOMATIC FACE ANALYSIS

The automatic face analysis [19] indeed as a complex and more difficult, reasons is that the individual face have different physiognomies and different appearance among individuals. The automatic face analysis system has three levels are represents in figure (2).

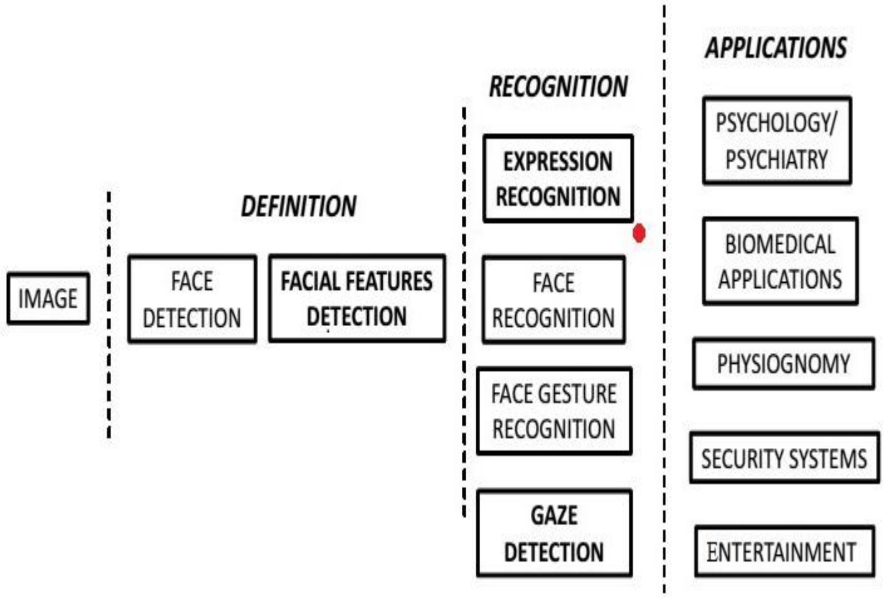

Figure (2): Automatic Face Analysis [19]

Levels of definition include the face detection and facial features detection. Levels of recognition include extracting information in form are expression, face recognition, face gesture and gaze detection. Levels of application are integration of domains of the Human–Computer applications.

FACE MODELLING-THE STATE-OF-TH-ART

From the levels of automatic face analysis, only concerned the face model for emotion model and expression recognition tasks. In computer graphics face model [19] is essential for facial synthesis and animation, computer vision is modeling the faces and its deformation of face analysis. The face model explore the state-of-the-art model in both domain are Analysis, Synthesis and Analysis-by-Synthesis.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2015

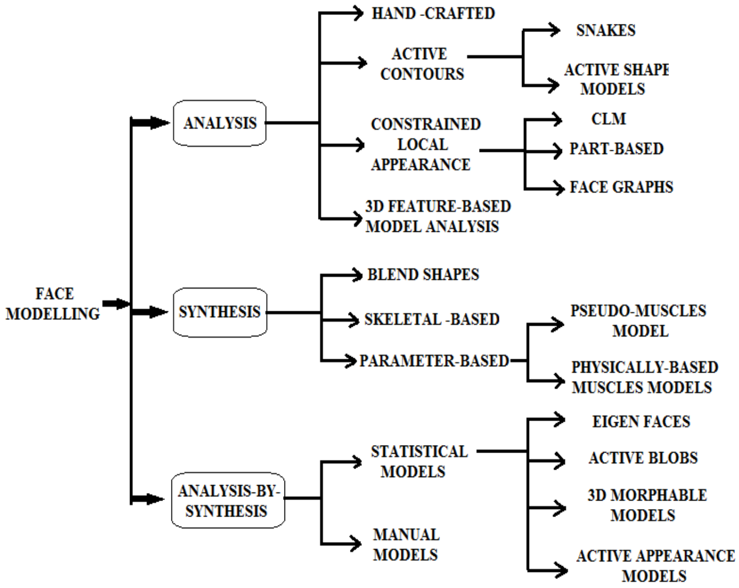

Figure (3): Face Modeling [19]

1. ANALYSIS:

It means analysis of the face in terms of shape and texture model in these four categories: hand crafted, active contours, constrained local appearance and 3D Feature based analysis

1.1 Hand crafted: In the hand crafted model, manually design a face feature using deformable template. [20] Build a parameterized deformable template of eyes and nose. First, design the feature of face using a deformable template and locating the feature of deformable template in face images. The extracting a face feature using a parameterized deformable template by an energy function which links edges, valley, peaks, weighted grads and weighted variance, etc.,. In this, define the mouth and eyes template model. [21] Proposes the mouth, eyes and nose template model, and [22] proposes the 3d head pose estimation using deformable templates. In the face model, interpolating B-splines cubic curves between the Face Animation Parameters (FAP) of MPEG4.

1.2Active contours:
1.2.1Snakes: Snakes defines for extraction of deformable face objects in computational bridges between the high level information and low level information Snakes are also called active contour models, its elastic, continuously flexible contour. [23] Proposes an energy minimization mechanism by external constrained forces and influenced by image forces drives its pull towards the image feature such as lines and edges. [24] Proposes a rubber snake for face feature segmentation by model based snake. In this, define the rubber snakes fundamentals for face feature segmentation. In order to segment a face feature by snakes, first localize an initial snake near to feature and extraction of face feature by template matching of horizontal segment and vertical segment. [25] This paper, proposes a new algorithm for the facial feature extraction part is “color snakes” applied to extract the facial feature points within the estimated face region. [26] Proposes a technique is 3d generic model of an extraction of face feature using by multiple snakes.

1.2.2Active shape models: Similar to Active Contour Model (ACM), but differ in those global constraints on the generated face. The Active shape model is defines the shape representation of invariant properties of class of shapes and provide a sacrifice model to specific a parametric description variability only deform in way found in the training set of data. [27] [28] Proposes a feature deformable template by a Point distribution model (PDM) build a landmark of face shape in training set. Using Principal component analysis (PCM), capturing the statistical of a set of aligning shapes for face. [29] Proposes for improve the robustness of calculating landmark using Adaboost histogram classifier

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2015

as local appearance model of landmark. [30] Estimated the non-linear boosted features training using Gentle Boost for improve computational and more accuracy for landmark detection

1.3Constrained Local Appearance

1.3.1 Constrained Local Model (CLM): The approach describes the template region surrounds the individual feature points instead of triangulation patches for appearance variation. [31] Proposes, first determine the feature of face model based on those local patches of landmark using their three types of detectors and built the joint modes of shape and texture variation of image by training sets to generate the set of features template. Then using normalized correlation for current feature template applied to search image for generates set of response surfaces. Finally, applying Nelder Meade simplex algorithm to drive the parameters of shape model in order maximize the response of image are obtained. [32] Proposes the Nearest Neighbor Template is updating a template instead of using fixed template during search engine. [33] Proposes a technique is Shape Optimized Search (SOS) for a feature detectors and response an each surface of feature.

1.3.2Part Based Model: [34] Proposes a Pictorial structure model also called Part based model is represents of an object in collection of part which are arranged in a deformable configuration. Separately a model of each part's of appearance, and deformable configuration is represented by spring-like connections between pairs of part. Most of the pictorial model restrict these connections to a tree structure, represents the spatial constraints between parts. [35] Propose a variation of Histogram of Oriented Gradients (HOG) descriptor is used to model an each part's appearance and the shape information is represented with a set of landmark points around the major facial features. For estimate the position of the landmark from each part's appearance by applying Multiple Linear Regression methods.

1.3.3 Face Graphs: [36] Proposes a design visual feature as a graph to track the face and pose variation. In this, set a salient feature for a face and each point corresponds to a nodes connected graph. The graph is labeled with Gabor filters to window around fiducial points and arches are formed by distance between the corresponding feature points. [37] Proposes a small set of sample face graph is constructed to stack-like structure called face bunch graph and for matching a new image graph by using Elastic Bunch Graph Matching.

1.4 3d Feature based model analysis: To analyze the facial deformation, 3d parametric model is creating for fit in the face describing in two steps: 2D facial feature tracking and extraction, followed by 3D parameters inference based on these points. Matching of 3D points of the model to the 2D feature by using an optimizations process can be needed. [38] Proposes a estimation of the features of an image using optical flow and applying displacement-based on Finite Element Models (FEM) to estimate the facial deformation. [39] Also track the feature point of image using optical flow. Using Levenberg-Marquardt (LM) optimization algorithm for Candide 3D model can be fit in face image which result in pose variations and template matching algorithm for Action parameter.

2. SYNTHESIS: The synthesis is model for create the face in animation parameter and represent the expression on face in order to recreate emotion respect to speed and sound that the animating face model in the synthesis fields are: blend-shape, skeleton shape and parameter based model.

2.1 Blend shape model: [40] Proposes a technique for to create a photorealistic 3d facial texture animation from photographs of subject. Then, they create a smooth transition between different facial expressions by morphing between these different models. First, from subjects recovers a different camera view corresponding to the 3d coordinate spaces of face. The estimates the feature points and facial features annotated by a set of curves in vary pose to form a generic face geometry (shaded surface rendering) using data interpolation technique. The recovers camera poses and face geometry creates the texture model of 3d face expression model. Finally create a realistic face animation model from a generate transition between these facial expressions, by using 3d shape morphing and blending the corresponding texture

2.2 Skeleton model: [41] In skeletal model, create a bones rig animation of facial expression in 3D mesh model. Like a real skeleton setup is formed an interconnected joint and nodes can be used to character bend into desired pose. Skinning is used for character's visual presentation associated with each bone rigs. Movement of vertex is simply done by the movement of skeleton model, which result the expression from object

2.3. Parameter-based model: The parameter based model can be used to create the different deformation corresponding to the parameter shape. The parameter based model has two approaches are:

1. Pseudo-muscles geometric models
2. Physically-based Muscle models

2.3.1 Pseudo-muscles geometric models: Using geometric approach for facial deformation through parameters, which a simulation of the visual effect of muscles is obtained.

2.3.1.1 Direct Parameterizations: Set of parameter which directly controls the vertex of the face. [42] [43] Proposes develop a parametric model for facial deformation. Set of parameter which controls the morphology of face and expression

2.3.1.2 Elementary deformations based model: The elementary movement of some predefined movement of face mesh. In this, define the various parameters for facial deformations are;

1. Face Action coding system (FACS) [44]
2. MPEG4 Facial Animation Parameters (FAP) [45] [46] [47] [48]
3. Abstract Muscle Actions (AMA) [49]
4. Minimal Perception Action [50]
5. Example-based deformation [51]

2.3.2 Physically-based Muscle models: Using physical muscle approach for facial deformation through parameter, which is a simulation of muscle action. [52] Proposes a mass-spring network model, modeled a skin behaviors according to muscle action.

3. ANALYSIS-BY-SYNTHESIS: In this model combine the both analysis and synthesis model serves the expression recognition and methods are:

1. Statistical model
    1. Eigen faces
    2. Active blobs
    3. 3D morphable model
    4. Active Appearance Model (AAM)
2. Manual model

3.1 Statistical model: It's built starting from real data and based on statistical tools

3.1.1 Eigen faces: [53] Proposes for face detection, track the face, and recognize the person by comparing of the faces to those of known individuals. Face images are projected onto feature image that coding and decoding the variation among known as "face images" is defined by Eigen face which are the Eigen vectors of set of faces.

3.1.2 Active blobs: [54] Proposes a region based approach (i.e. textured-mapped deformable 2D triangular mesh) to track the non-rigid motion of object. Shape is defined in terms of deformable triangular mesh captures the shape of an object, plus a color texture map that captures the appearance of an object and also modeled the photometric variation

3.1.3. 3d Morphable model: The model is 3d Morphable function is based on the linear combination of large number of 3D face scans. [55] Proposes a morph able texture 3D scans can either be generated automatically one or more photographs or through intuitive user interface directly. Starting from the laser scan, the set of 3D face model can derive the 3D morph able face model by transforming the shape and texture representation into a vector representation and defined to be sum of mean of the shape or texture plus the linear combination of set of prototypic basis. Using automated matching algorithm or guided manually, derives the 3D shape and texture reconstruction and illumination of corrected texture extraction derived a result of 3d face for expression analysis.

3.1.4 Active appearance model: Active appearance model (AAM) [56] is similar to the active blobs and morph able model but differs in examples, number of vertex, and number of parameters. The AAM model can be defined in two methods:

1. Combination of Eigen faces and Active shape models (ASM) [57]

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2015

2. Constrained local methods [31]

Combination of Eigen faces and Active shape Model [57]: In ASM modeling, the shape variation is to obtaining a set of shape parameters and in Eigen face modeling, the texture variation to get the set of texture parameters. Finally combining the shape and texture parameters is specific to AAM making the model of appearance and shape of face in the same vector.

3.2 Manual Model: The model is build based on the human knowledge and observation. [44] [45] [46] [47] [48] Candide model is 3D parameterization model of face shape with a Facial Action Coding System (FACS) and Face Animation Parameter (FAP). In shape modeling, using Action parameters, shape parameters and animation parameters are defined by the FACS and FAP and texture model is needed for fit Candide model in face images.

V.LITERATURE SURVEY OF BODY TRACKING OF EXPRESSION

1. Abhishek Kar, Dr.Amitabha Mukherjee and Dr.Prithwijit Guha, **"Skeleton tracking using Microsoft Kinect"**,2011, [58] In Human pose estimation has been a long standing problem in Computer vision application like motion capture, Human computer interaction and activity recognition, so they proposes a method to tackle the problem by skeleton tracking human body using the Microsoft Kinect sensor. From the Microsoft Kinect sensor, used cues from the RGB and depth stream to fit a stick skeleton model to human upper body for human body pose estimation. The foreground segmentation is extracted a depth threshold image from sensor. The used Viola-Jones algorithm is based on Haar Cascade detector to fizz the face and upper body position of the subject and also used Adaboost classifier for face detector. The hand position is detected by skin color segmentation on foreground segmented RGB image. The skeleton stick models fix the skeleton on subject from NASA Anthropometric data source. Finally computed with Extended Distance Transform Skeletonisation algorithm to estimate the skeleton parameters fitted to the upper human body.

2. David Antonio G´omezJ´auregui, Patrick Horain, Manoj Kumar Rajagopal, Senanayak Sesh Kumar Karri, **"Real-Time Particle Filtering with Heuristics for 3D Motion Capture by Monocular Vision"**,2010 [59] proposes a 3D human motion capture by without marker in real time by registering a 3D articulated model of the upper human body pose estimation. First, the monocular vision captured the upper human body to detected silhouette using a robust Back ground subtraction algorithm. The foreground human image is segmented in face, arms and clothes. Adabooster face detector is detected the face, then arm detection i.e. skin color sample is taken from the face region and a clothes samples also taken under the face. Each sample model used a simple Gaussian model in HSV color space. According to the pose in the vectors of parameter, the 3D model is projected by rendering the three class of region. Non-overlapping ratio is used for matching between the 3D model projection and segmented image. Particle filtering algorithm is used for robustly human body tracking at the cost of heavy computation in a high dimensional pose space, that particle filter is Sequential Monte Carlo (SMC) is estimated the posterior density of particle (pose). First, search the high dimensional space of 3D poses by generated new hypotheses with equivalent 2D projection by kinematic flipping. Second, based on local optimization used the semi-deterministic particle prediction. Third, determined resample the probability distribution for more efficient selection of particles. Principal component analyses (PCA) is reduced the parameter space dimensionality of integrating Gaussian prior pose probability. Finally, they demonstrated jointly a number of heuristics to improve robustness and achieved a real-time 3D human motion captured by monocular vision on GPU parallelized.

3. Raquel Urtasun, David J. Fleet, Pascal Fua, **"Monocular tracking of the golf swing" .2009,** [60] proposes an approach to including dynamic models into the human body tracking process of golf swing that yields full 3D reconstructions from monocular vision. In motion model, is represented the human body golf swing as a set of volumetric attached to an articulated 3D skeleton and identified the 4 key postures of human golf swing depicted by each motion and time warped at the same time. The sampled at regular time intervals using quaternion spherical interpolation. For tracking the 3D human pose of golf swing is formulated as the least-square minimization of an objective function with respect to motion model parameter and global motion of skeleton root node. The least- square of the principal component analyses are defined the observation of motion parameters and global motion of skeleton parameter. Initially, the image from monocular vision is segmented by the foreground and background segmentation for extracted rough binary mask of foreground and background observation function. The image projection constraint located the six joints to binary mask image and using WLS tracker is a robust motion-based 2D tracker to track only

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2015

joint constraint, for wrists used the club tracking algorithm. Using simple correlations –based algorithm used 2D point correspondences in pairs of consequence image projected the 3D model into first image of the pair, sample the projection and established correspondence of samples. Under the joint constraints tracking has occurs some problem in under constrained and a multiple set of solution are reduced by using correspondences. Finally, they formulated a tracking of the 3D monocular of human golf swing using a single-hypothesis hill-climbing approach as opposed to a computational complexity of multi-hypothesis algorithm.

4.Jungong Han, Ling Shao, Dong Xu, Jamie Shotton, "**Enhanced Computer Vision with Microsoft Kinect sensor: a Review**", 2013, [61] reviewed about that topic are Object Tracking and Recognition, Human Activity analysis, Hand gesture analysis and Indoor 3D mapping. In this paper, Microsoft Kinect sensor mechanisms, advantages are detailed. Compared the various type sensor and Microsoft Kinect sensor [62] [63] [64] [65] [66] [67]for detection and tracking

VI. THE COMPARISON OF TECHNIQUES AND DISCUSSION,

Comparison of various technique observations given in all references is discussed here.

Table 1: COMPARSION OF VARIOUS TECHNIQUE ANALYSES:

Sl.no	TECHNIQUE USED	EXPERIMENTAL RESULT
[10]	1. Principal Component Analysis (PCA). 2. Support Vector Machine (SVM).	Six basic emotion: 1. Based on acoustic: average result is 76.44%. 2. Based on textual: average result is 65.48%.
[11]	1. Short Time Log Frequency Power Coefficients (LFPC). 2. Linear prediction Cepstral coefficients (LPCC). 3. Mel-frequency Cepstral coefficients (MFCC). 4. Hidden Markov model (HMM)	Six basic emotion: 1. Average accuracy of 78%. 2. Best accuracy of 96%.
[12]	1. One-class-in-one network model (OCIO) 2. Linear predictive coding(LPC)	Eight emotion recognition rate approximately 50%
[13]	Seven cues: symmetry and limb alignment, head alignment and offset, body openness, average rate of change, relative movement, smooth-jerk, and location of activity	1. Body pose movement feature has accuracy of 66.5% differentiating between four emotion 2. Using the raw joint rotations, limb rotation movements, or posture features alone, are only able to achieve accuracy rates of 59%, 61%, and 62% respectively.
[14]	1. Kanade–Lucas–Tomasi (KLT) tracker 2. Multiclass Support Vector Machine (SVM). 3. Facial action UNITS (FAU) with SVM	Facial expression system Based on: Multiclass SVM has achieved 99.7% FAU with SVM has 95.1% No pose variation Low computational processing
[15]	1. MPEG4 FaciaL Animation Parameter (FAPS) 2. Eyes Web Expressive Gesture Processing Library 3. Intensity, pitch, MFCC (Mel Frequency Cepstral Coefficient), Bark spectral bands,	1. From face has achieved 48.6% of emotion 2. From body has achieved 67.1% 3. From speech 57.1% 4. Feature level has achieved 78.3%

	voiced segment characteristics and pause length 4. Bayesian classifier	5. Decision level has achieved 74.6% 6. In multimodal fusion accuracy is 85.1%
[16]	1. Surrogate model patterns 2. Greenhouse-Geisser-corrected 3. Brain Electrical Source Analysis (BESA) With Principal Component analysis (PCA)	Only Four emotion are recognize from the fusiform Cyrus (temporal lobe), medial frontal (frontal lobe)(occipital lobe) visual cortex are weaker and /or slower in autism spectrum disorder compared with typically control. More effortful compensatory analytical strategies from Parietal Somensatory cortices brain waves activity used to process of face gender and emotion in autism.
[17]	1. Hybrid Adaptive Filtering (HAF). 2. Higher Order Crossing Analysis (HOC).	Six basic emotion: highly classification rate up to 85.17 %
[18]	1. Mel Frequency Cepstral Coefficient (MFCC) 2. Multilayer Perception (MLP)	Three emotion only recognition of emotion which be calculated by spectral analysis.
[20,21,22]	1. parameterized deformable template 2. downhill simplex method 3. interpolating B-splines cubic curves 4. Facial Animation Parameter (FAPS)	1. Extract feature good and accurately 2. Good energy function 3. Not flexible for head pose and expression variation
[23,24,25,26]	1. Template matching 2. Segmentation 3. Rubber snake model 4. Color snake model	1. Feature extraction is accurately and good 2. No facial expression and head pose variation
[27,28,29,30]	1. Point distribution model (PDM) 2. Principal Component analysis (PCA) 3. Chord Length Distribution(CLD) 4. Adaboost	1. Good for shape model 2. No texture model 3. Accuracy is high compare to the snakes, and hand crafted
[31,32,33]	1. Constrained Local Model (CLM) 2. Template Selection Tracker (TST) 3. Shape Optimized Search (SOS)	1. Feature detection is accurate 2. Good for face expression analysis and head pose variation
[34,35]	1. Viterbi algorithm 2. Histogram of Oriented Gradients descriptor 3. Multiple linear regression models	1. Expression recognition is possible when using active appearance model 2. Part based model has good accuracy 3. Head pose variation is possible
[36,37]	1. Gabor wavelet transform 2. Gaussian envelope function 3. Taylor expansion	1. Good for head pose variation 2. Not flexible for expression variation 3. Good accuracy of feature extraction
[38,39]	1. Bayes Tangent Shape Model. 2. Geometrical Deformation Models(GDM) 3. Constrained Shape Model (CSM)	Good for Facial expression and head pose variation with 3d parameter

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2015

	4. Principle Component Analysis (PCA)	
[40]	1. Levenberg-Marquardt algorithm 2. Scattered data interpolation 3. Segmentation 4. Blend shape model	1. Good for facial expression and head pose variation but we create the new 3d animation setup for it 2. Need large number of database
[41]	1. Linear Blend Skinning (LBS) 2. Geometric model 3. Skeletal model	1. Good for body tracking 2. Not flexible for facial expression and pose variation
[42,43] [44] -- [52]	1. Direct Parameterizations 2. Face Action coding system (FACS) 3. MPEG4 Facial Animation Parameters (FAP) 4. Abstract Muscle Actions (AMA) 5. Minimal Perception Action 6. Example-based deformation 7. mass-spring network model	Parameter for facial expression and head pose variation is accurate
[53]	1. Eigen faces 2. Eigen vector 3. Principle Component Analysis (PCA)	1. Only face recognition 2. Not flexible for facial expression and pose variations.
[54]	3. Finite Element Mode 1. Eigen vectors 2. Active blobs 3. Marquardt-Levenberg method 4. Difference decomposition approach 5. Gaussian interpolate with finite support	1. Non-rigid motion of the face (motion of eye brow) 2. Not effective in head pose 3. Not tested the complex non – rigid motion of face (motion of mouth)
[55]	1. Morphing model 2. Segmentation 3. Automatic matching algorithm 4. Optical flow algorithm 5. Illumination-Corrected Texture Extraction	Good for face expression and head pose variation has accurate and very robustly, but need laser scan for create the 3d morphable model
[56,57,31]	1. Kanade–Lucas–Tomasi (KLT) tracker 2. Eigen faces 3. Active Shape model 4. Constrained Local appearance (CLM)	1. Face expression is accurate 2. Head pose variation is possible 3. Emotion recognizes done as accuracy of 81.1% 4. Low resolution ,need light intensities by using stereo camera
[44] -- [48]	1. Viola Jones algorithm 2. Face Action coding system (FACS) 3. MPEG4 Facial Animation Parameters (FAP) 4. Adaboost	1. Face expression done has a accuracy of 92.6% 2. Pose variation -35% to +35% 3. High computational processing 4. High resolution, no need light intensities ,also robustly skeleton tracking using Kinect sensor
[58]	Extended Distance Transform Skeletonisation algorithm by Haar Cascade detection is based on Viola Jones and Adaboost and skin segmentation	1. Achieved fair accuracy. 2. Running time at about 10fps. 3. Incorrect skeleton arm crossing

		and towards camera axis. 4. Reduced computational time by Template matching.
[59]	1. Particle filtering with heuristics (PFRTH) 2. Local Optimization (Downhill simplex) 3. Kinematic flipping 4. Principal Component analysis (PCA). 5. Gaussian prior probability.	1. Achieved higher robustness and accuracy in high dimensional space of monocular tracking in 3D motion. 2. Future work in more robustness in low dimensional space
[60]	1. Finger –Earth Mover's Distance (FEMD). 2. Template matching 3. Minimum Near-Convex decomposition(MNCD) 4. Threshold decomposition	1. More accuracy and efficient 2. Achieved 93.2% mean accuracy and runs in 0.075 s per frame

Discussion: The comparison table gives the various techniques used for estimate the face expression recognition observed in all references. From the [10] - [18] literature survey the facial expression system achieved from their various modes of extracting emotion. In this, face gesture has achieved more advantages such as high accuracy of emotion, head pose variation, very robustly compared to the other modes of extracting. From [20]- [57] reviewed that defines the automatic facial analysis and state-of-art model. From [20]- [39] defines face analysis model, retrieved such as a feature points, feature detection of face, shape and texture model, deformable model, energy function, face expression recognizer, head pose of face feature. From [40]- [52] reviewed that defines the facial animation parameter, retrieved such as morphable 3d face model, animation facial expression recognition also achieved the face expression but accuracy is low, no head pose variation, etc.,

From [53]- [57] and [44]- [48] defines the facial analysis by facial animation parameter for facial expression recognition system and head pose variation. Its achieved high accuracy of face expression and emotion recognition, head pose variations, very robustly (particularly these two models are Candide model and two methods of Active Appearance Model (AAM)) compared to the other face model In [58]- [61] reviewed the estimates a body tracking for emotion recognition using various sensor, particularly Microsoft Kinect sensor [62] has achieved high accuracy, very robustly skeletal tracking algorithm, also has a face expression classifier and head pose variation, compare to the various sensors. From these, we reviewed that the discussion and comparison of technique for automatic expression recognition.

VII.FUTURE WORK AND CHALLENGES

As discussed earlier, the major challenge that researchers are focusing on capturing spontaneous expression on images and video to get high accuracy, avoid the loss of authenticity, and also spontaneous head pose expression variation. Another challenging capturing spontaneous non basic expression compared to the capturing of basic expression and also achieves the same accuracy of basic six emotions. Another area, where more work is carried out on completely automatic expression system in real time application and expression synthesis. Researchers have also focus for multimodal fusion of expression recognition has achieved less accuracy when compared to the unimodal system. Also differences do exist in facial feature and facial expression between cultures, age groups is more challenging. In medical application, some patient likes autistic disorder, Asperger syndrome, epilepsy, etc. monitoring for the loss of facial expression analysis also challenging.

In skeletal body tracking, loss their authenticity when expression analysis is also challenging. I conclude here, these are future work and challenging of expression analysis

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2015

VIII. CONCLUSION

We reviewed the recent trends of facial expression analysis, automatic expression recognition system and associate areas. We have presentation in this paper our literature survey which will be easily understandable for new comers who are interested in this field but have no background knowledge on the same. We summarize from the literature survey of automatic expression recognizers system, reviewed that the basic of emotion, the various mode of extracting emotion analysis, defined the suitable face model for expression analysis of state-of-the-art. From these, analysis by synthesis face model is better for facial expression recognizes compared to the other model and skeletal tracking for expression recognition has achieved high accuracy using through Microsoft Kinect sensor compared to the other sensor. Finally, the future work and challenges are explained with their advantages which help the improvement of Human computer Interaction application (HCI).

REFERENCES

[1] Richard Wiseman and Rob Jenkins Roger Highfield, "How your looks betray your personality".: New Scientist, Feb 2009.

[2] [Online]. http://www.library.northwestern.edu/spec/hogarth/physiognomics1.html.

[3] J Montagu., "“The Expression of the Passions: The Origin and Influence of Charles Le Brun’s ‘Conférence sur l'expression générale et particulière,”," Yale University press, 1994.

[4] Charles Darwin, “The Expression of the Emotions in Man and Animals,”, 2nd ed., J Murray, Ed. London , 1904.

[5] William James, "What is Emotion," Mind, vol. 9, no. 34, pp. 188-205, 1884.

[6] Paul Ekman and Wallace.V. Friesen, "Pan-Cultural Elements In Facial Display Of Emotion," Science, pp. 86-88, 1969.

[7] Harald G., and Klaus R. Scherer Wallbott, ""Cues and channels in emotion recognition."," Journal of personality and social psychology, pp. 69-70, 1986.

[8] F., Malfaz, M., Sequeira, J., Gorostiza, J. F., & Salichs, M. A. Alonso-Martín, "A Multimodal Emotion Detection System during Human-Robot Interaction," Sensors, vol. 13(11), no. 15549., 2013.

[9] Stefanos Kollias and Kostas Karpouzis, "Multimodal Emotion Recognition And Expressivity Analysis," IEEE International Conference on Mulitimedia and Expo, pp. 779-783, 2005.

[10] Say Wei Foo,LiyanageC. De Silva Tin Lay New, "Speech emotion recognition using hidden Markov models," Speech communication , no. 41.4, pp. 603-623, 2003.

[11] K. Takahashi and R. Nakatsu Nicholson, "Emotion Recognition in Speech Using Neural Networks," Neural Computing & Applications , no. 9.4, pp. 290-296, 2000.

[12] Ze-Jing Chuang and Chung-Hsien Wu, "Multi-Modal Emotion Recognition from Speech and Text," International Journal of Computational Linguistics and Chinese Language Processing, no. 9.2 , pp. 1-18, 2004..

[13] Michael, and Mei Si, Garber Barron, "Using body movement and posture for emotion detection in non-acted scenarios. ," , 2012.

[14] I. Kotsia and I. Pitas, "Facial Expression Recognition in Image Sequences Using Geometric Deformation Features and Support Vector Machines," IEEE Trans. Image Processing, vol. 16, no. 1, pp. 172-187, 2007.

[15] George Caridakis, "Multimodal emotion recognition from expressive faces, body gestures and speech.", Artificial intelligence and innovations 2007: From theory to applications Springer US, pp. 375-388, 2007.

[16] Peter C. W. Fung, Siew E. Chua and Grainne M. McAlonan Teresa K. W. Wong, "Abnormal spatiotemporal processing of emotional facial expressions in childhood autism: dipole source analysis of event‐related potentials," European Journal of Neuroscience ():, no. 28.2, pp. 407-416., 2008.

[17] Panagiotis C. Petrantonakis and Leontios J. Hadjileontiadis, "Emotion Recognition from Brain Signals Using Hybrid Adaptive Filtering and Higher Order Crossings Analysis," 2010 Affective Computing, IEEE Transactions , no. 1.2, pp. 81-97, 2010.

[18] Marini, and Abdul Wahab Othman, "Affective face processing analysis in autism using electroencephalogram.," in Information and Communication Technology for the Muslim World (ICT4M), 2010 International Conference on. IEEE, 2010.

[19] Hanan Salam, "Multi-Object modeling of the face," Université Rennes 1, Signal and Image processing, Thesis 2013.

[20] Peter W.Hallinan and David S.Cohen Alan L Yulie, "Feature extraction from faces using deformable templates", International journal of computer vision, vol. 8(2), pp. 99–111, 1992.

[21] Zhang, and Ruan Qiuqi Baizhen, "Facial feature extraction using improved deformable templates.," in Signal Processing, 2006 8th International Conference on,Vol 4, 2006.

[22] Marius, and Francoise J. Preteux. Malciu, "Tracking facial features in video sequences using a deformable-model-based approach," in International Symposium on Optical Science and Technology. International Society for Optics and Photonics, 2000.

[23] A. Witkin, and D. Terzopoulos M. Kass, "Snakes: Active contour models," International journal of computer vision, vol. 1(4), pp. 321-331, 1988.

[24] Petia, and Eric Martí Radeva, "Facial features segmentation by model-based snakes," in International Conference on Computing Analysis and Image Processing, Prague, 1995.

[25] Won kim,Changmok Oh and Ju-Jang Lee Kap-ho Seo, "Face detection and facial feature extraction using color snake," in Industrial Electronics, 2002. ISIE 2002. Proceedings of the 2002 IEEE International Symposium, 2002.

[26] Rein-Lien, and Anil K. Jain Hsu, "Generating discriminating cartoon faces using interacting snakes," Pattern Analysis and Machine Intelligence, IEEE Transactions, vol. 25, no. 11, pp. 1388-1398, 2003.

[27] D.H. Cooper, C.J. Taylor, and J. Graha, T.F. Cootes, "Trainable method of parametric shape description," Image and Vision Computing, vol. 10, no. 5, pp. 289–294, 1992.

[28] C.J. Taylor, D.H. Cooper, and J. Graham T.F. Cootes, "Active shape models-their training and application," Computer vision and image understanding, vol. 61, no. 1, pp. 38–59, 1995.

[29] Yuanzhong, and Wataru Ito Li, "Shape parameter optimization for adaboosted active shape model.", in Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on vol 1, 2005.

[30] David, and Timothy F. Cootes Cristinacce, "Boosted Regression Active Shape Models," British Machine Vision Conference, 2007.

[31] David, and Timothy F. Cootes Cristinacce, "Feature Detection and Tracking with Constrained Local Models," British Machine Vision Conference, vol. 2, no. 5, 2006.

[32] David, and Timothy F. Cootes Cristinacce, "Facial feature detection and tracking with automatic template selection.," in Automatic Face and Gesture Recognition, 2006. FGR 2006. 7th International Conference on. IEEE, 2006.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2015

[33] D. Cristinacce and T.F. Cootes, "A comparison of shape constrained facial feature detectors," In Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on, vol. 375–380, 2004.
[34] P.F. Felzenszwalb and D.P. Huttenlocher., "Pictorial structures for object recognition," International Journal of Computer Vision, vol. 61, no. 1, pp. 55–79, 2005.
[35] Vahid, and Josephine Sullivan Kazemi, "Face alignment with part-based modeling," in Proceedings of the British Machine Vision Conference. BMVA Press, 2011.
[36] T. Maurer and C. Von der Malsburg, "Tracking and learning graphs and pose on image sequences of faces," in In Automatic Face and Gesture Recognition, 1996., Proceedings of the Second International Conference on, 1996, pp. 176–181.
[37] Jean-Marc Fellous,Norbert Kruger, Laurenz Wiskott, "Face recognition by elastic bunch graph matching," Pattern Analysis and Machine Intelligence, IEEE Transactions, vol. 19, no. 7, pp. 775-779, 1997.
[38] Ilse, and Hichem Sahli Ravyse, "A biomechanical model for image-based estimation of 3d face deformations," in Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on. IEEE, 2008.
[39] Ping Fan,Ilse Ravyse,Valentin Enescu,Rongchun Zhao and Hichem Sahli Yunshu Hou, "Smooth adaptive fitting of 3d face model for the estimation of rigid and nonrigid facial motion in video sequences," Signal Processing: Image Communication , vol. 26, no. 8, pp. 550-566, 2011.
[40] W.C. Tien, M. Desbrun, and F. Pighin Joshi, "Learning controls for blend shape based realistic facial animation," In ACM SIGGRAPH 2005 Courses, p. 8, 2005.
[41] Baran and J. Popovic, "Automatic rigging and animation of 3d characters," ACM Transactions on Graphics (TOG), vol. 26, no. 3, p. 72, 2007.
[42] Frederic I Parke, "A Parametric Model for Human Faces," University of Utah, PhD thesis 1974.
[43] Frederic I Parke, "Parameterized models for facial animation," Computer Graphics and Applications , vol. 2, no. 9, pp. 61-68., 1982).
[44] Paul Ekman and W.V. Friesen, "Facial action coding system," Stanford University, Palo Alto, Consulting Psychologists Press 1977.
[45] M. Rydfalk, "Candide, a parameterized face. Technical report," Dept. of Electrical Engineering, Linkoping University, Sweden, LiTH-ISY-I-866 Technical report 1987.
[46] J. Ahlberg, "Candide-3–an updated parameterized face," Dept. of Electrical Engineering, Linkoping University, Sweden, LiTH-ISY Technical report 2001.
[47] J. Ahlberg, "An active model for facial feature tracking," EURASIP Journal on Applied Signal Processing, pp. 566–571, 2002.
[48] Atlantic City, 1998), MPEG Video and SNHC, Text of ISO/IEC FDISDoc. Audio (MPEG Mtg, ISO/MPEG N2503. 14 496–3,.
[49] Nadia, E. Primeau, and Daniel Thalmann. Magnenat-Thalmann, "Abstract muscle action procedures for human face animation," The Visual Computer, vol. 3, no. 5, pp. 290-297., 1988.
[50] P. Kalra, "An Interactive Multimodal Facial Animation System," Ecole Polytechnique F´ed´erale de Lausanne, Switzerland, PhD thesis 1993.
[51] N. Stoiber, "Modeling Emotional Facial Expressions and their Dynamics for Realistic Interactive Facial Animation on Virtual Characters. ," Universit´e de Rennes, PhD thesis 2010.
[52] S.M. Platt and N.I. Badler., "Animating facial expressions," ACM SIGGRAPH computer graphics, vol. 15, no. 3, pp. 245–252, 1981.
[53] Mathew A.Turk and Alex P.Pentland, "Face recognition using eigenfaces," in Computer Vision and Pattern Recognition, 1991. Proceedings CVPR '91., IEEE Computer Society Conference on, 1991, pp. 586-591.
[54] S. Sclaroff and J. Isidoro., "Active blobs. ," in In Computer Vision,1998. Sixth International Conference , 1146–1153, p. 1998.
[55] V. Blanz and T. Vetter, "A morphable model for the synthesis of 3d faces," in In Proceedings of the 26th annual conference on Computer graphics and interactive techniques, 1999, pp. 187–194.
[56] G. Edwards, and C.J. Taylor T.F. Cootes, "A comparative evaluation of active appearance model algorithms," in In British Machine Vision Conference, vol 2, 1998, pp. 680–689.
[57] G.J. Edwards, and C.J. Taylor T.F. Cootes, "Active appearance models," in In IEEE European Conference on Computer Vision (ECCV '98), 1998, p. 484.
[58] Dr.Amitabha Mukherjee and Dr.Prithwijit Guha, Abhishek Kar, "Skeleton tracking using Microsoft Kinect," Springer , pp. 1-11, 2011.
[59] Patrick Horain, Manoj Kumar Rajagopal, Senanayak Sesh Kumar Karri David Antonio G´omezJ´auregui, "Real-Time Particle Filtering with Heuristics for 3D Motion Capture by Monocular Vision," Multimedia Signal Processing (MMSP), IEEE International Workshop., pp. 139 – 144, 2010.
[60] David J.Fleet, Pascal Fua Raquel Urtasun, "Monocular tracking of the golf swing," Computer Vision and Pattern Recognition, CVPR 2005. IEEE Computer Society Conference , pp. 932 - 938, 2005.
[61] Ling Shao,Dong Xu,Jamie Shotton Jungong Han, "Enhanced Computer Vision with Microsoft Kinect Sensor: A Review," IEEE Transactions On Cybernetics, vol. 43, no. 5, 2013.
[62] David Catuhe, "Programming with the Kinect for windows software development Kit," Microsoft Press , 2012.
[63] Henk Luinge, and Per Slycke Daniel Roetenberg, "Xsens MVN: Full 6DOF Human Motion Tracking," Xsens Motion Technologies BV, technical report 2009.
[64] Abhijit Jana, Kinect for Windows SDK Programming Guide., 2012.
[65] Grega Borenstein, MAKING THINGS SEE., 2012.
[66] [Online]. http://msdn.microsoft.com/en-us/library/hh438998.aspx.
[67] [Online]. http://www.ifixit.com/Teardown/Microsoft+Kinect+Teardown/4066