Australian Journal of Basic and Applied Sciences, 9(21) Special 2015, Pages: 1-6

ISSN:1991-8178

Australian Journal of Basic and Applied Sciences

Journal home page: www.ajbasweb.com

A Novel Approach for Facial Expression Analysis in real time applications using SIFT flow and SVM

[1]K. Suganya Devi and [2]P. Srinivasan

[1]*Department of Computer Science and Engineering, University college of Engg Panruti, Panruti 607106, Tamilnadu, India*
[2]*Department of Physics, University college of Engg Panruti, Panruti 607106, Tamilnadu, India*

ARTICLE INFO

Article history:
Article Received : 12 January 2015
Revised: 1 May 2015
Accepted: 8 May 2015

Keywords:
Expression recognition, Facial region selection, Facial expression, Sparse learning technique, Scale Invariant Feature Transform flow, SVM

ABSTRACT

Facial expression is an effective way for humans to communicate since it contains critical and necessary information regarding human affective states. It is a critical part of affective computing systems that aim to recognize and therefore better respond to human emotions. Automatic recognition of facial expressions can be an important component in human-machine interfaces, human emotion analysis, and medical care. However, the task of automatically recognizing various facial expressions challenging. As a result, facial expression recognition has become a prominent research topic in human-computer interaction, as well as in the fields of image processing, pattern recognition, machine learning, and human recognition. A server sets up a database with training facial images from all expression classes. Since all images represent the face, it is necessary to extract discriminative features of these images that correspond to different expression classes in order to simplify the classification. A client requesting facial recognition service would supply a test image whose expression it desires to recognize. This test image would be encrypted in order to prevent the server from being able to gain access to its actual private contents. In this paper, we can evaluate the facial expressions using SIFT (Scale Invariant Feature Transform) projection and HOG (Histogram Oriented Gradient) approach. And classify the results using SVM (Support Vector Machine) approach. Our experimental result shows that good performance in real time applications.

To Cite This Article: K. Suganya Devi and P. Srinivasan., A Novel Approach for Facial Expression Analysis in real time applications using SIFT flow and SVM. ***Aust. J. Basic & Appl. Sci.****, 9(21): 1-6, 2015*

INTRODUCTION

Facial expression is a non-verbal communication which is voluntarily or non voluntarily adopted by humans. Through facial expression people express different emotions. Facial expression plays a significant role in human communication (Ekman, P. and W.V. Friesen, 1976). In each image the facial emotion is recognized by eye brow arcs, fiducially points such as nose tip, eye corner, onset, inset and apex. Eye contact is another major aspect of facial communication. A person's eye reveal much more about how they are feeling or what they are thinking. Automatic recognition (Michel, P. and R. El Kaliouby, 2003; Zheng, W., 2006) of facial expressions can be an important component of natural human-machine interfaces. Humans recognize facial expressions virtually without effort or delay, reliable expression recognition by machine is still a challenge.

Affective Computing as shown in Fig.1 is computing that relates to, arises from, or deliberately influences emotion or other affective phenomena. Emotion is fundamental to human experience, influencing cognition, perception, and everyday tasks such as learning, communication, and even rational decision-making. However, technologists have largely ignored emotion and created an often frustrating experience for people, in part because affect has been misunderstood and hard to measure. Our research develops new technologies and theories that advance basic understanding of affect and its role in human experience. We aim to restore a proper balance between emotion and cognition in the design of technologies for addressing human needs.

Our research work has contributed to: Designing new ways for people to communicate affective-cognitive states, especially through creation of novel wearable sensors and new machine learning algorithms that jointly analyze multimodal channels of information; Creating new techniques to assess frustration, stress, and mood indirectly, through natural interaction and conversation (Wang, Y., 2004) Showing how computers can be more

Corresponding Author: K. Suganya Devi, Department of Computer Science and Engineering, University college of Engg Panruti, Panruti 607106, Tamilnadu, India
E-mail: ssuganya.ucep@gmail.com

emotionally intelligent, especially responding to a person's frustration in a way that reduces negative feelings; (Zheng, W., 2006) Inventing personal technologies for improving self-awareness of affective state and its selective communication to others (Tian, Y.L., 2004) Increasing understanding of how affect influences personal health; and (Zeng, Z., 2009) Pioneering studies examining ethical issues in affective computing.

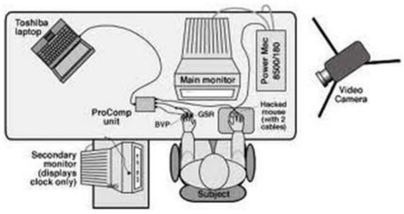

Fig. 1: Affective computing functions.

Affective Computing research combines engineering and computer science with psychology, cognitive science, neuroscience, sociology, education, psychophysiology, value-centered design, ethics, and more. We bring together individuals with a diversity of technical, artistic, and human abilities in a collaborative spirit to push the boundaries of what can be achieved to improve human affective experience with technology. Some of the areas of Affective Computing are:

- ***Detecting and Recognizing Emotional Information:***

Detecting emotional information begins with passive sensors which capture data about the user's physical state or behavior without interpreting the input. The data gathered is analogous to the cues humans use to perceive emotions in others. For example, a video camera might capture facial expressions, body posture and gestures, while a microphone might capture speech. Recognizing emotional information requires the extraction of meaningful patterns from the gathered data. This is done using machine learning techniques that process different modalities, such as speech recognition, natural language processing, or facial expression detection, and produce either labels or coordinates in a valence-arousal space.

- ***Emotion in Machines:***

Another area within affective computing is the design of computational devices proposed to exhibit either innate emotional capabilities or that are capable of convincingly simulating emotions. A more practical approach, based on current technological capabilities, is the simulation of emotions in conversational agents in order to enrich and facilitate interactivity between human and machine. While human emotions are often associated with surges in hormones and other neuropeptides, emotions in machines might be associated with abstract states associated with progress in autonomous learning systems. In this view, affective emotional states correspond to time-derivatives (perturbations) in the learning curve of an arbitrary learning system.
Some of the technologies of Affective Computing are:

- ***Emotional Speech:***

One can take advantage of the fact that changes in the autonomic nervous system indirectly alter speech, and use this information to produce systems capable of recognizing affect based on extracted features of speech.

For example, speech produced in a state of fear, anger or joy becomes faster, louder, precisely enunciated with a higher and wider pitch range. Other emotions such as tiredness, boredom or sadness, lead to slower, lower-pitched and slurred speech. Emotional speech processing recognizes the user's emotional state by analyzing speech patterns. Vocal parameters and prosody features such as pitch variables and speech rate are analyzed through pattern recognition.

Speech recognition is a great method of identifying affective state, having an average success rate reported in research of 63%. This result appears fairly satisfying when compared with humans' success rate at identifying emotions, but a little insufficient compared to other forms of emotion recognition (such as those which employ physiological states or facial processing). Furthermore, many speech characteristics are independent of semantics or culture, which makes this technique a very promising one to use.

- ***Facial Affect Detection:***

The detection and processing of facial expression is achieved through various methods such as optical flow, hidden Markov model, neural network processing or active appearance model. More than one modalities can be combined or fused to provide a more robust estimation of the subject's emotional state.

- ***Facial Action Coding System:***

Defining expressions in terms of muscle actions A system has been conceived in order to formally categorize the physical expression of emotions. The central concept of the Facial Action Coding System, or FACS, as created are Action Units (AU). They are, basically, a contraction or a relaxation of one or more muscles. However, as simple as this concept may seem, it is enough to form the base of a complex and devoid of interpretation emotional identification system.

With the rapid development of human–machine interaction, affective computing is currently gaining popularity in research and flourishing in the industry domain. It aims to equip computing devices with effortless and natural communication. The ability to recognize human affective state will empower the intelligent computer to interpret, understand, and respond to human emotions, moods, and possibly intentions. This is similar to the way that humans rely on their senses to assess each other's affective state. Many potential applications, such as intelligent automobile systems, game and entertainment industries, interactive video, indexing and retrieval of image or video databases, can benefit from this ability.

Facial expressions arise owing to a person's internal emotional states, intentions, or social communications. On the one hand, these facial changes present important challenges for face recognition algorithms, where researchers are proposing various expression-invariant face recognition algorithms. On the other hand, these facial changes are the best cues for recognizing facial expressions. Understanding the users' emotions is a fundamental requirement of Human-Computer Interaction systems (HCI) and facial expressions are important means of detecting emotions. Emotion recognition is the first and one of the most important issues in the affective computing field. It incorporates computers with the ability to interact with humans more naturally and in a friendly manner. Affective interaction can have maximal impact when emotion recognition and expression is available to all parties, human and computers. Most of the existing systems attempt to recognize the human prototypic emotions.

It is widely accepted from psychological theory that human emotions can be classified into six archetypal emotions: surprise, fear, disgust, anger, happiness, and sadness, which are so-called six basic emotions.

There are also several other emotions and many combinations of emotions that have been studied, but they are unconfirmed as universally distinguishable. Facial expression regulates face-to-face interactions, indicates reciprocity, interpersonal attraction or repulsion, and enables inter subjectivity between members of different cultures. The proposed method aims to make use of the facial features associated with all facial views in dealing with the expression recognition, in which the multi-view facial features are synthesized by the features of one facial view through Kernel Reduced-Rank Regression (KRRR) model and aims to simultaneously deal with the problems of facial region selection and expression recognition. To cope with the facial regions selection, the sparse learning technique that had been successfully used in many machine learning algorithms, such as Sparse Principal Component Analysis (SPCA) or Sparse Reduced-Rank Regression (SRRR), will be adopted. To this end, we propose a Group Sparse Reduced Rank Regression (GSRRR) model to describe the relationship between the multi-view facial feature vectors and the corresponding expression class label vectors.

I. Existing System:

A. Principal component Analysis (PCA):

Principal Component analysis is a mathematical procedure that is used to transform potentially correlated variables into uncorrelated variables. Suppose we have a data matrix of observations of N correlated variables X1,X2,. . . ,XN, PCA will transform the Xi variables into N new variables Yi that are uncorrelated. The variables Yi are called principal components. The first principal component is in the direction of the largest variance of the origin. The other principal components are orthogonal to each other and represent the largest residual variance. PCA can be used as a dimension reduction method to represent multidimensional, highly correlated data, with fewer variables. PCA is used for, e.g. information extraction, image compression, image reconstruction and image recognition.

B. Subspace Analysis:

In existing define distance measures (Hu, Y., 2008; Zheng, W., 2009) between two linear subspaces. They help us to study the similarity of two SISs and that between an SIS and the AIS. The subspace distance will play an important role in the prior distribution of SIS and make the MAP adaptation an eigenvalue problem. However, the discussion in this section is general and the subspace distance may have other applications such as the video based recognition. We also consider the distance between two nonlinear surfaces. In particular, we use the kernel trick, for which the surfaces are the pre-images of linear subspaces in high-dimensional feature space.

C. Linear Subspaces:

Although a subspace can be seen as a set of points, common distance measures (Tang, H. and T.S. Huang, 2008) of sets are not appropriate for subspaces. For example, the minimum distance between two point sets A and B in Euclidean space, defined as $d(A,B) = \min \|a-b\| \{a€A , b€B\}$ is always zero when A and B are subspaces, since the origin belongs to all subspaces. Another commonly used distance, the Hausdorff distance $d(A,B) = \max \min \|a-b\|$ is infinity if A $B\neq$, due to the unboundedness of linear subspace. Intuitively, the distance/similarity of two subspaces should reflect the difference between their "directions". That is, if two subspaces nearly coincide, they should have a small distance; if they are almost perpendicular to each other, they have a large distance. These intuitive ideas will be incorporated in definition of subspace distance.

II. Proposed System:

In this proposed work, we empirically study facial representation based on cascade features for person-independent facial expression recognition. SIFT features were proposed originally for texture analysis, and recently have been introduced to represent faces in facial images analysis. The most important properties of SIFT features are their tolerance against illumination changes and their computational simplicity. We examine different machine learning methods, including template matching; to reduce the facial feature points and create the SVM model to labeling the facial expression with class labels. We implement the SVM technique to find head poses from real time facial images. Finally, the model response corresponding to the expression class label vector is calculated and the expression category of the testing facial image can be obtained based on it. The following Fig.2 depicts the system architecture of the proposed work.

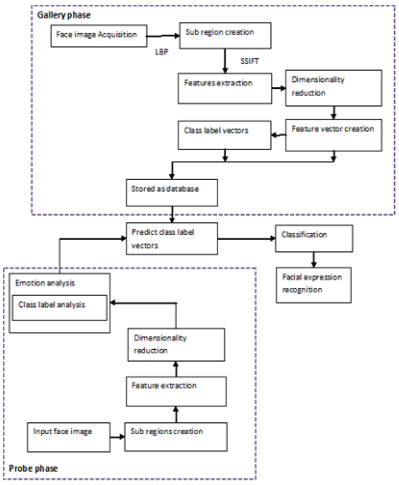

Fig. 2: System Architecture.

A. Face Image Acquisition:

In this module, we capture the face image or upload the datasets. The uploaded datasets contains 2D face images. In face registration we can identify the faces which are captured by web camera. Then web camera images known as 2D images. And we perform the preprocessing steps such as gray scale conversion, invert, and border analysis, detect edges and region identification. The Grayscale images are also called monochromatic, denoting the presence of only one (mono) color (chrome). The edge detection is used to analyze the connected curves that indicate the boundaries of objects, the boundaries of surface markings as well as curves that correspond to discontinuities in surface orientation.

B. Features Extraction:

In this module we can implement Local binary pattern technique to extract features from face image. The haar cascade feature vector, in its simplest form, is created in the following manner and it is shown in Fig.3:

- Divide the examined window into cells (e.g. 16x16 pixels for each cell).
- For each pixel in a cell, compare the pixel to each of its 8 neighbors (on its left-top, left-middle, left-bottom, right-top, etc.). Follow the pixels along a circle, i.e. clockwise or counter-clockwise.
- Where the center pixel's value is greater than the neighbor's value, write "1". Otherwise, write "0". This gives an 8-digit binary number (which is usually converted to decimal for convenience).
- Compute the histogram, over the cell, of the frequency of each "number" occurring (i.e., each combination of which pixels are smaller and which are greater than the center).
- Optionally normalize the histogram.
- Concatenate (normalized) histograms of all cells. This gives the feature vector for the window.

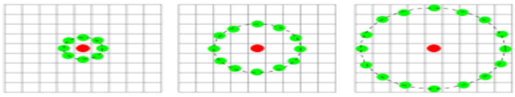

Fig. 3: The HAAR cascade feature vector.

C. Dimensionality Reduction:

In this module implement SIFT algorithm to extract the features points. The Fig.4 depicts the video recorded and frame conversion. Scale Invariant Feature Transform (SIFT) is an approach for detecting and extracting local feature descriptors that are reasonably invariant to changes in illumination, scaling, rotation, image noise and small changes in viewpoint.

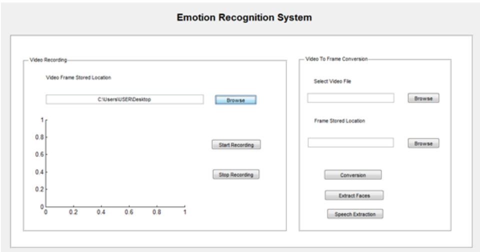

Fig. 4: Video is recorded and it is converted to frames for training.

Detection stages for SIFT features are as follows:

(1) Scale-space extrema detection: The first stage of computation searches over all scales and image locations. It is implemented efficiently by means of a difference of- Gaussian function to identify potential interest points that are invariant to orientation and scale.

(2) Key point localization: At each candidate location, a detailed model is fit to determine scale and location. Key points are selected on basis of measures of their stability and it is shown in Fig.5.

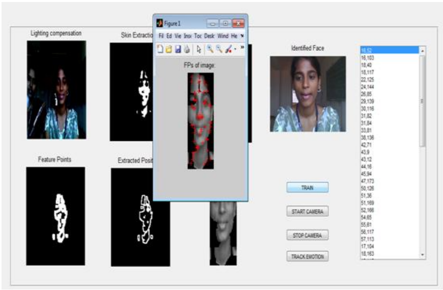

Fig. 5: Face is identified and compared with the trained images.

(3) Orientation assignment:

One or more orientations are assigned to each key point location on basis of local image gradient directions. All future operations are performed on image data that has been transformed relative to the assigned scale, orientation, and location for each feature, thereby providing invariance to these transformations.

(4) Generation of key point descriptors:

The local image gradients are measured at the selected scale in the region around each key point. These gradients are transformed into a representation which admits significant levels of local change in illumination and shape distortion.

D. ***Expression Analysis:***

In this module analyze on the expression recognition for testing facial images. For a testing facial image, we first extract the facial features and then perform the expression estimation, where SVM classifier is used for this purpose. After obtaining the expressions, we synthesize facial feature vectors based on testing facial feature vector and use them as the model predictors of the HOG model. Finally, the model response corresponding to the expression class label vector is calculated and the expression category of the testing facial image can be obtained based on it. This is shown in Fig.6.

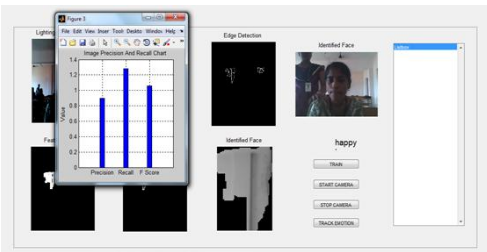

Fig. 6: Emotion is extracted.

III. Conclusion:

In this paper, we proposed SIFT based algorithm. Considering an expressive face as a superposition of a neutral face with expression component, we proposed an algorithm to decompose an expressive test face into its building components. For this purpose, we first generate grids for captured face using local binary patterns. Knowing that the face component of the test face has sparse representation in the face database and the expression part can be sparsely represented using the expression database; we decompose the test face into these feature vectors. The elements of the test face along with the vectors are then used for face and expression recognition. For this purpose, the separated components are sparsely decomposed using vectors while the grouping structures of the vectors are enforced into the sparse decomposition. The experimental results on both databases showed that the proposed method achieves competitive recognition performance compared with the state of the art methods under same experimental settings and same facial feature. As a future direction, we plan to model occlusions better, so that the overall performance of the system can be increased. We extend our work to less limited registration approach and Independent of nose visibility. Then Occlusion invariant recognition system has following aspects,

- Automatic occlusion detection and removal.
- Discriminative features other than depth information.
- And also implement this concept analyze expression in various illumination conditions.

REFERENCES

Ekman, P. and W.V. Friesen, 1976. Pictures of Facial Affect. Consulting Psychologists Press.

Michel, P. and R. El Kaliouby, 2003. “Real Time Facial Expression Recognition in Video Using Support Vector Machines,” Proc. Fifth Int’l Conf. Multimodal Interfaces, 258-264.

Wang, Y., H. Ai, B. Wu and C. Huang, 2004. “Real Time Facial Expression Recognition with Adaboost,” Proc. 17th Int’l Conf. Pattern Recognition (ICPR ’04), 926-929.

Zheng, W., X. Zhou, C. Zou and L. Zhao, 2006. “Facial Expression Recognition Using Kernel Canonical Correlation Analysis (KCCA),” IEEE Trans. Neural Networks, 17(1): 233-238.

Tian, Y.L., T. Kanade and J.F. Cohn, 2004.“Facial Expression Analysis,” Handbook of Facial Recognition, S.Z. Li and A.K. Jain, eds., Springer- Verlag.

Zeng, Z., M. Pantic, G.I. Roisman and T.S. Huang, 2009. “A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions,” IEEE Trans. Pattern Analysis and Machine Intelligence, 31(1): 39-58.

Hu, Y., Z. Zeng, L. Yin, X. Wei, J. Tu and T.S. Huang, 2008. “A Study of Non-Frontal-View Facial Expressions Recognition,” Proc. 19th Int’l Conf. Pattern Recognition (ICPR), 1-4.

Hu, Y., Z. Zeng, L. Yin, X. Wei, X. Zhou and T.S. Huang, 2008. “Multi- View Facial Expression

Recognition," Proc. Eighth IEEE Int'l Conf. Automatic Face and Gesture Recognition.

Zheng, W., H. Tang, Z. Lin and T.S. Huang, 2009. "A Novel Approach to Expression Recognition from Non-Frontal Face Images," Proc. IEEE 12th Int'l Conf. Computer Vision (ICCV), 1901-1908.

Zheng, W., H. Tang, Z. Lin and T.S. Huang, 2010. "Emotion Recognition from Arbitrary View Facial Images," Proc. 11th European Conf. Computer Vision (ECCV '10), 490-503.

Tang, H. and T.S. Huang, 2008. "3D Facial Expression Recognition Based on Automatically Selected Features," Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR '08) Workshops.

Tang, H. and T.S. Huang, 2008. "3D Facial Expression Recognition Based on Properties of Line Segments Connecting Facial Feature Points," Proc. Eighth IEEE Int'l Conf. Automatic Face and Gesture Recognition (FG '08).

Tang, H., M. Hasegawa-Johnson and T.S. Huang, 2010. "Non-Frontal View Facial Expression Recognition Based on Ergodic Hidden Markov Model Supervectors," Proc. IEEE Conf. Multimedia & Expo (ICME), 1202-1207.

Rudovic, O., I. Patras and M. Pantic, 2010. "Coupled Gaussian Process Regression for Pose-Invariant Facial Expression Recognition," Proc. 11th European Conf. Computer Vision: Part II (ECCV '10), 350-363.

Moore, S. and R. Bowden, 2011. "Local Binary Patterns for Multi-View Facial Expression Recognition," Computer Vision and Image Understanding, 115: 541-558.

Rudovic, O. and M. Pantic, 2011. "Shape-Constrained Gaussian Process Regression for Facial-Point-Based Head-Pose Normalization," Proc. IEEE Int'l Conf. Computer Vision (ICCV), 1495-1502.

Rudovic, O., M. Pantic and I. Patras, 2013. "Coupled Gaussian Processes for Pose-Invariant Facial Expression Recognition," IEEE Trans. Pattern Analysis and Machine Intelligence, 35(6): 1357-1369.

Yin, L., X. Wei, Y. Sun, J. Wang and M.J. Rosato, 2006. "A 3D Facial Expression Database for Facial Behavior Research," Proc. Seventh Int'l Conf. Automatic Face and Gesture Recognition, 211-216.

Lowe, D.G., 2004. "Distinctive Image Features from Scale-Invariant Keypoints," Int'l J. Computer Vision, 60(2): 91-110.

Ojala, T., M. Pietikainen and T. Maenpaa, 2002. "Multi-Resolution Gray- Scale and Rotation Invariant Texture Classification with Local Binary Patterns," IEEE Trans. Pattern Analysis and Machine Intelligence, 24(7): 971-987.