IJCSE International Journal of Computer Sciences and Engineering Open Access
Review Paper Volume-6, Issue-1 E-ISSN: 2347-2693

A Review of Movement Exaggeration Techniques to Enhance the Precision Identification for Minute Facial Feelings

M. Jyothirmai[1*], Y.V. Sree Chandana[2], C.Vishnu Vardhan[3]

[1*]Department of ECE, St. Peter's Engineering College, Medchal, India
[2] Department of ECE, St. Peter's Engineering College, Medchal, Hyderabad, Telangana, India
[3] Department of ECE, St. Peter's Engineering College, Medchal, Hyderabad, Telangana, India

[*]Corresponding Author: jyothirmai@stpetershyd.com Tel.: +91-9949053929

Available online at: www.ijcseonline.org

Received: 03/Jan/2018, Revised: 06/Jan/2018, Accepted: 18/Jan/2018, Published: 31/Jan/2018

Abstract— Acknowledgment of regular feelings from human countenances is an interesting point with a number of potential applications like human-framework connection, computerized frameworks, image and video recovery and similar development platforms. Much research has already been done in this area and there is scope for further improvement. Comparison was done for four different algorithms based on accuracy of recognition rate. The goal is to achieve improvement compared to previous algorithms. By using PCA-SIFT the accuracy was improved between 6%-18%.

Keywords— Extreme learning machine, Spatio-temporal descriptor, Binary decision tree, Scale invariant feature transform

I. INTRODUCTION

Identifying facial feelings in uncontrolled situations is an extremely difficult undertaking because of vast intra-class varieties brought on by elements, for example, light and stance changes, impediment, and head development. The precision of a facial feeling acknowledgment framework for the most part relies on upon two central points: (i) extraction of facial components which are vigorous under intra-class varieties (e.g. posture changes), yet are unmistakable for different emotions. (ii) plan of a classifier that is equipped for recognizing diverse facial feelings in light of uproarious and defective information (e.g., enlightenment changes and impediment). Facial feature identification is one of the basic parts of face investigation and comprehension with face restriction and face acknowledgment. [1][2]

It turns into a more finished area which is used as a part of an expansive number of utilizations, in the midst of which we discover security, new correspondence limit, biometrics and numerous others. Face discovery is meant to identify human faces in still pictures or recordings, in various circumstances. Numerous calculations execute the face-recognition errand as a paired example characterization assignment.

That is, the placate of a given picture is changed into prepared classifier or concentrates the facial component, after which a prepared classifier faces choose whether that specific part of the area of the image is face or not. As often as possible, a window-sliding technique is in work. That is, the ELM classifier is utilized to sort typically square or rectangular bits of a picture, at all areas and scales, as additionally faces or non-faces.

The paper is organised as follows. Section I gives the introduction of identifying facial feelings using various techniques, Section II shows the related work on various algorithms used to find the recognition rate of facial feelings, Section III gives the methodology of PCA-SIFT, Section IV contains the experimental work performed for the extraction of facial feelings with block diagram, Section V gives the simulation results obtained after performing the experiments, Section VI gives the table for the comparison of different algorithms and the accuracy of recognition rate obtained, Section VII concludes the research work.

II. RELATED WORK

The existing facial feeling identifying frameworks are accessed in research controlled environment in certifiable applications. An exhaustive writing overview of the same can be found.

2.1 Extreme Learning Machine

The learning speed of the feed forward networks is slower than required and it has been a major bottleneck in many applications. The essence of ELM is that the learning parameters of hidden nodes including input biases are randomly assigned and need not be tuned while the output weights can be analytically determined by the simple generalized inverse operation[3].

2.2 Spatio-Temporal Descriptor

Spatio-Temporal appearance (STA) descriptors are settled length descriptors that speak to a progression of transiently related points. Two variations exist: STA descriptors of the principal arrange (STA1) and STA descriptors of the second request (STA2). In STA the accessible area is partitioned into a normal lattice of rectangular patches. Then a self assertive picture work is computed and lattice vectors are averaged which is a satisfactory representation of easier spatio-worldly features.[4]

2.3 Binary decision tree

A choice tree is a basic representation for ordering illustrations. It is one of the best methods of directed arrangement learning.

For this area, we consider that the majority of the elements have limited discrete spaces, and there is a solitary target include called the characterization. Every component of the space of the grouping is known as a class. A choice tree or a grouping tree is a tree in which each inner (non-leaf) hub is marked with an element.

For every element experienced in the tree, the circular segment relating to the estimation of the case for that component is taken thereafter. At the point when a leaf is achieved, the grouping related to that leaf is returned.

2.4 Scale invariant feature transform

Scale-invariant component change (or SIFT) is a calculation in PC vision to distinguish and depict nearby elements in pictures.

Applications incorporate protest acknowledgment, mechanical mapping and route, picture sewing, 3D demonstrating, signal acknowledgment, video following, singular recognizable proof of untamed life and match moving. For any question in a picture, fascinating focuses on the protest can be removed to give an "element depiction" of the question.

This portrayal, removed from a preparation picture, can then be utilized to distinguish the emotions when endeavoring to find the question in a test picture containing numerous different articles. To perform solid acknowledgment, it is critical that the components extricated from the preparation picture be noticeable even under changes in picture scale, commotion and brightening. Such focuses more often than not lie on high-differentiate locales of the picture like edges. [5]

Another critical normal for these components is that the relative positions between them in the first scene shouldn't change starting with one picture then onto the next. For instance, if just the four corners of an entryway were utilized as elements, they would work showing small concern to the entry way's position; yet in the event that focuses in the edge were additionally utilized, the acknowledgment would come up short if the entryway is opened or shut.

Such parameters would ordinarily not work if any adjustment in their inward geometry happens between two pictures in the set being prepared. Be that as it may, by and by SIFT distinguishes and utilizes a much bigger number of components from the pictures, which lessens the commitment of the mistakes brought about by these nearby elements.

SIFT highlight descriptor is invariant to uniform scaling, introduction, and in part invariant to relative bending and brightening changes. Filter key purposes of items are initially separated from an arrangement of reference images[6] and stored in a database.
A question is perceived in another picture by independently contrasting every element from the new picture to this database and discovering competitor coordinating components in view of Euclidean separation of their element vectors. From the full arrangement of matches, subsets of key focuses that concur on the question and its area, scale, and introduction in the new picture are distinguished to sift through great matches. The assurance of steady groups is performed quickly by utilizing a productive hash table usage of the summed up Hough change.

Every bunch includes that concur on a question and its posture is then subject to advance model confirmation and along these lines exceptions are disposed of. Finally the likelihood that a specific arrangement of components demonstrates the nearness of a question is registered, given the precision of fit and number of plausible false matches.

We start by distinguishing purposes of intrigue in the SIFT structure. The picture is convolved with Gaussian channels at various scales, and afterward the distinction of progressive Gaussian-obscured pictures is taken. Key focuses are then taken as maxima/minima of the Difference of Gaussians (DoG) that happen at different scales[17].

2.5 Comparison of various techniques

Developments of straight forward things can be enhanced by associated volume-sparing sinusoidal based expand and squash impacts for totally misleadingly created characters, through, including style is not too unmistakable. [7][8]

The thinking is clear and straightforwardness enough to make development getting data all the additionally captivating. ELM beats the moderate preparing speed and over-fitting issues.

ELM depends on experimental hazard minimization hypothesis and its learning procedure needs just a solitary emphasis. The calculation maintains a strategic distance from numerous cycles and nearby minimization. It has been utilized as a part of different fields and applications due to better speculation capacity, vigor, and controllability and quick learning rate. [9]

Spatio worldly descriptor matches two arrangements of descriptors by processing the separations between the conceivable correspondences, and after that applying a worldwide improvement calculation to authorize spatial consistency. Choice trees are simple to comprehend and translate. Individuals can comprehend choice tree models after a brief clarification.

It requires little information planning and is able to handle both numerical and clear cut information. Different strategies are normally represented considerable authority in breaking down datasets that have stand out sort of factor. SIFT descriptor highlights show the most elevated coordinating exactness for a relative change of 50 degrees.

After this change restrict, comes about begin to wind up questionable. Peculiarity of descriptors is measured by summing the eigen estimations of the descriptors, acquired by the principal segments examination of the descriptors standardized by their fluctuation. This compares to the measure of change caught by various descriptors, consequently, to their peculiarity.

III. METHODOLOGY

Our algorithm for local descriptors (termed PCA-SIFT) accepts the same input as the standard SIFT descriptor: the sub-pixel location, scale, and dominant orientations of the key point. We extract a 41×41 patch at the given scale, centred over the key point, and rotated to align its dominant orientation to a canonical direction. PCA-SIFT can be summarized in the following steps: pre-compute an Eigen space to express the gradient images of local patches; given a patch, compute its local image gradient; project the gradient image vector using the Eigen spaceto derive a compact feature vector. This feature vector is significantly smaller than the standard SIFT feature vector, and can be used with the same matching algorithms. The Euclidean distance between two feature vectors is used to determine whether the two vectors correspond to the same key point in different images. Principal Component Analysis (PCA) is a standard technique for dimensionality reduction and has been applied to a broad class of computer vision problems, including feature selection, object recognition and face recognition. While PCA suffers from a number of shortcomings, such as its implicit assumption of Gaussian distributions and its restriction to orthogonal linear combinations, it remains popular due to its simplicity. Our contribution lies in rigorously demonstrating that PCA is well-suited to representing key point patches (once they have been transformed into a canonical scale, position and orientation), and that this representation significantly improves SIFT's matching performance[10].

IV. EXPERIMENTAL WORK.

In this paper we have used the PCA-SIFT (Principal Components Analysis connected to SIFT descriptors) algorithm to identify the facial feelings. GLOH and SIFT highlights give the most noteworthy qualities. Filter based descriptors outflank other contemporary nearby descriptors on both finished and organized scenes, with the distinction in execution bigger on the finished scene. The basic blocks of the facial emotion detection is shown the figure below. [11]

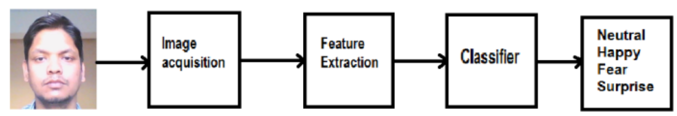


Figure 1.Architecture of Facial emotion recognition

MATLAB software was used to implement the PCA-SIFT algorithm. The facial features were extracted and compared with the available database. Location of the facial features was determined by computing the centroid of the face.

V. SIMULATION RESULTS

While training the classifier we do it by providing five instances of each motion.

The face data was taken for 1 individual who represents different emotions such as Neutral, Happy, Fear and Surprise. Five to six clips were extracted from the videos, each lasting between 3 to 6 seconds.

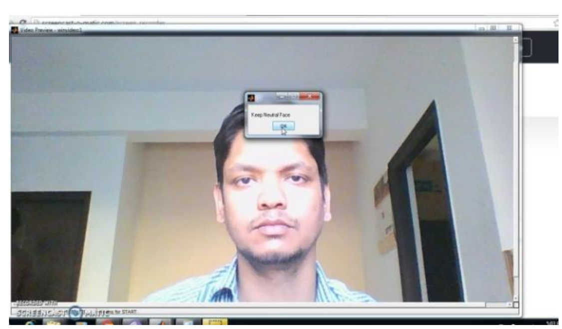
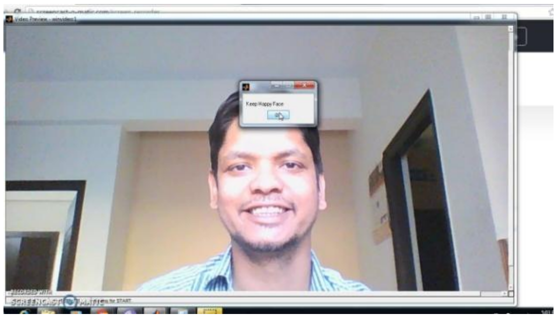


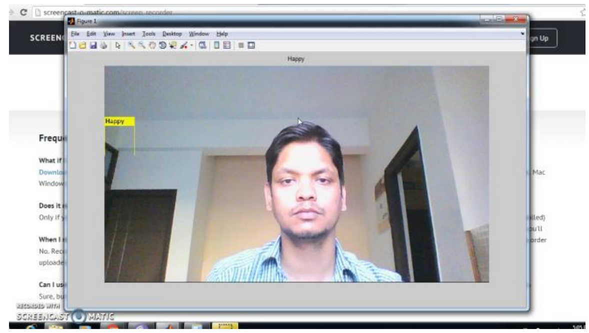
Fig.2.a Training for Neutral Face


2.b Training for Happy Face

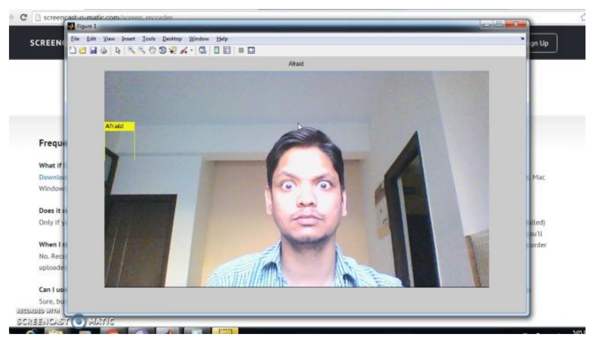

2.c Training for Fear Face

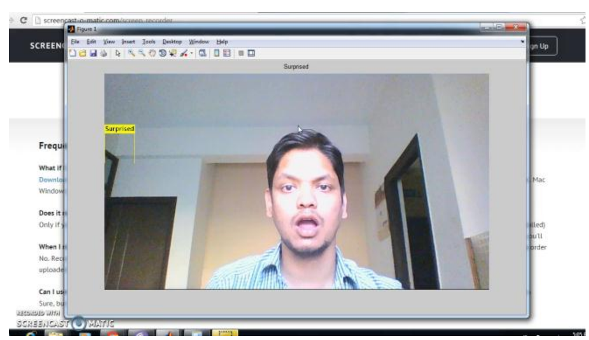

2.d Training for Surprised Face

The database consists of neutral, surprise, happy, fear facial expressions. The clips extracted from the video are compared with the database. The first image in database is neutral facial expression and the clips extracted from the video are compared with neutral face. An accuracy of 78.4% for happy and surprise faces, 78.2% for anger and fear were obtained.
While testing, the emotions were checked online and the video frames are tagged with the emotion.


3.a Neutral Face Detected


3.b Happy Face Detected


3.c Fear Face Detected


3.d Surprise Face Detected

VI. RESULTS AND DISCUSSION

We have performed probes on both acted and unconstrained feeling databases to assess the adequacy of the proposed highlight extraction and arrangement plots under various situations by using the PCA SIFT algorithm. The dimensions of the data was reduced, effective face recognition and better matching results were achieved.

Extreme learning machine algorithm characterizes intensity and dynamics of facial emotions while the head is in stable state with a recognition rate of 63%[13]. Spatio temporal descriptor algorithm achieved a recognition rate of 72% [14]and produced lowest errors[14] while the Binary decision tree produced 60.5% [15].

By using the PCA SIFT algorithm we were able to achieve a recognition rate of 78.43%. It can be observed that our proposed algorithm was able to achieve significant increase in the recognition rate.

Author	Algorithm	Purpose	Accuracy
Mr.R.Sathish Kumar[13]	Extreme Learning Machine	Characterizes intensity and dynamics of facial emotions while head is in stable state.	63%
Mengyi Liu[14]	Spatio Temporal Descriptor	Used to produce lowest errors.	72%
CC Lee[15]	Binary Decision Tree	Used to mitigate the error propagation.	60.5%
G. Zhao[12]	PCA-SIFT	Used to achieve significant increase in recognition rate. It is used for representation of key point patches and improves SIFT's matching performanc.	78.43%

VII. CONCLUSION

This paper gives an overview of the facial feeling identifying algorithms. ESL consolidates remaking properties of meagre representation and discriminative force of a nonlinear ELM for powerful arrangement and a novel OF-based spatio-transient descriptor for posture invariant facial feeling discovery[12].

In this paper an advanced version of SIFT i.e, PCA-SIFT algorithm is used and is demonstrated. Even though the SIFT algorithm is invariant for many images, but it fails if the image size is larger or not under the scope which is defined. Hence in this paper we use the PCA-SIFT algorithm for the extraction of facial emotions. The recognition rate for various facial emotions is calculated, which on an average is 68.25%. The accuracy of recognition rate is improved from 60.5% to 78.43% from Binary Decision Tree method to PCA-SIFT method.

VIII. REFERENCES

[1] W. Gu, C. Xiang, Y. V. Venkatesh, D. Huang, and H. Lin, "*Facial Expression Recognition using Radial Encoding of Local Gabor Features and Classifier Synthesis*," Pattern Recog., vol.45,no.1,pp.80–91,2012.

[2] T. Wehrle, S. Kaiser, S. Schmidt, and K. R. Scherer, "*Studying the Dynamics of Emotional Expression Using Synthesized Facial Muscle Movements*," J. Pers. Soc. Psychol., vol. 78, no. 1, pp. 105–119, 2000.

[3] Guang-Bin Huang, Qin-Yu Zhu, and Chee-KheongSiew "*Extreme Learning Machine: A New Learning Scheme of Feed forward Neural Network Extreme Learning Machine: Theory and Applications*", Neurocomputing, vol. 70, pp. 489-501, 2006

[4] Alexander Klaser, Marcin Marszałek and Cordelia Schmid "*A Spatio-Temporal Descriptor Based on 3D-Gradients*" INRIA Grenoble, LEAR, LJK,pp.1-10.

[5] D. H. Kim, S. U. Jung, and M. J. Chung, "*Extension of Cascaded Simple Feature based Face Detection to Facial Expression Recognition*," Pattern Recog. Letters, vol. 29, no. 11, pp. 1621–1631, 2008.

[6] Smith, T., Jones, M.: *'The title of the paper'*, IET Syst. Biol., 2007, **1**, (2), pp. 1–7

[7] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and M. Yi, "*Robust Face Recognition via Sparse Representation,*" IEEE Trans. PAMI,vol. 31, no. 2, pp. 210–227, 2009.

[8]K. Huang and S. Aviyente, "*Sparse Representation for Signal Classification*," in Adv. NIPS, 2006.Van der Geer J, Hanraads JAJ, Lupton RA. The art of writing a scientific article. J SciCommun 2000;163:51-9,pp.472-478.

[9] P. S. Aleksic and A. K. Katsaggelos, "*Automatic Facial Expression Recognition using Facial Animation Parameters and Multistream HMMs*," IEEE Trans. Inf. Forensics Security, vol. 1, no. 1, pp. 3–11, 2006.

[10]Y Ke, R Sukthankar - … , 2004. CVPR 2004. Proceedings of the IEEE Computer Society Conference 2004 - ieeexplore.ieee.org.pp,1312-1324.

[11] Y. Zhang and Q. Ji, "*Active and Dynamic Information Fusion for Facial Expression Understanding from Image Sequences*," IEEE Trans. PAMI, vol. 27, no. 5, pp. 699–714, 2005.

[12] G. Zhao and M. Pietikainen, "*Dynamic Texture Recognition Using Local Binary Patterns with an Application to Facial Expressions*," IEEE Trans. PAMI, vol. 29, no. 6, pp. 915–928, 2007.

[13] Mr.R.Sathish Kumar 1, G.Mohanraj M.Srivathsan" *Recognition of Facial Emotions Structures Using Extreme Learning Machine Algorithm*" (IRJET)Volume:3,pp.1187-1190, Issue: 04 Apr 2016.

[14] Mengyi Liu , Shiguang Shan "*Learning Expressionlets on Spatio-Temporal Manifold for Dynamic Facial Expression Recognition*" Proceedings of the CVPR Confernce Volume:2, pp.4321-4328, Issue: 02 Oct 2014 - ieeexplore.ieee.org

[15] CC Lee, E Mower, C Busso, S Lee, S Narayanan "*Emotion recognition using a hierarchical binary decision tree approach*" - Speech Communication, 2011 – Elsevier, pp.1162-1171.

[16] Stefano Berretti, Boulbaba Ben Amor, Mohamed Daoudi, Alberto Del Bimbo."*3D facial expression recognition using SIFT descriptors of automatically detected keypoints*". Visual Computer, Springer Verlag, 2011, 27 (11), pp.1021-1036

[17] Mohammad Mohsen Ahmadinejad, Elizabeth Sherly,"*An advanced scale invariant feature transform algorithm for face recognition* ", IJCSE, Vol. 7 No. 3, pp.82-90.

Authors Profile

Ms. M Jyothirmai pursed Bachelor of Technology from JNTU Hyderabad in 2000 and Master of Technology from JNTU Kakinada in year 2009. She is currently pursuing Ph.D. and currently working as Assistant Professor in Departmentm of Electronics and Communication Engineering, St. Peter"s Engineering College since 2011. She is a life member of the ISTE since 2016. She has published 2 research papers in International conferences. Her main research work focuses on Image Processing, Compression Algorithms and Cloud Computing She has 12 years of teaching experience and 2 years of Research Experience.

Ms Y V Sree Chandana is pursuing Bachelor of Technology from St. Peter"s Engineering College, Hyderabad, Her areas of interest are Image Processing, Signal Processing and VLSI Desinging.

Mr C Vishnu Vardhan is pursuing Bachelor of Technology from St. Peter"s Engineering College, Hyderabad, His areas of interest are Image Processing, Artificial Intelligence and embedded systems.

