International Journal of Scientific Research & Engineering Trends
Volume 7, Issue 2, March-April-2021, ISSN (Online): 2395-566X

Facial Expressions Recognition Based on LBP & SVM

M.Tech. Scholar Nagesh Patel, Asst. Prof. & Head Ashish Tiwari
Department of computer science
Vindhya Institute of Technology & Science,
Indore, (M.P)
nageshpatel878@gmail.com, ashishtiwari205@gmail.com

***Abstract*- Facial expression analysis is a compelling and demanding problem affecting important applications in various fields such as human-computer interaction and data-driven animation. The development of effective facial expressions from the original facial images is an important step in gaining facial expression recognition. The actual evaluation is based on the face representation of statistical local functions, local binary pattern (LBP). Several machine learning techniques have been thoroughly observed on various databases. Researchers usually use the effective and competent LBP feature of facial expression recognition. Cohn Kanade is the database for the current work and the programming language used is MATLAB. First, the face area is divided into small areas through which histograms are extracted, local binary pattern (LBP) and then connected as a single function vector. This feature vector outlines a well-organized representation of face and is helpful in determining the resemblance among images. These operators along with other proposed techniques were experimented considering different settings viz. with or without localization and registration errors, person dependent or independent, operator scales, number of grids (hence the size) and number of available LBP based codes. The FR system was configured in verification mode using Eigen face approach derived on these LBP based histogram feature vectors. The FER system was configured for multi-class facial expression classification mode using Support Vector Machine (SVM). Experiments on facial expression databases reveal that maximum recognition rate would be obtained for the scale in which width is larger than height of the operator. Both proposed operators are sensitive to registration errors. However, these operators could be applied in automatic FERS.**

***Keywords*- Facial expression, local binary pattern, Support Vector Machine, FER**

I. INTRODUCTION

Facial expression recognition (FER) has become a hot topic of research in human-computer interaction (HCI) and has attracted attention because of its enormous potential in multimedia applications. Digital entertainment, customer service, driver monitoring, etc If the computer recognizes emotions as human beings, HCI will become more friendly and natural, which can benefit from solving the FER problem.

FER aims to analyze and classify a given facial image into six of the six common Types of Emotions One of the common types of emotions is anger, disgust, fear, joy, sadness and surprise. In the last few years, many FER algorithms have been proposed in the literature, including the recognition of facial expressions from frontal and non-frontal facial images. Compared to positive FER, non-positive FER is more challenging and more suitable for practical scenarios. However, among the various proposed methods, only a small part of the algorithm solves this problem. For positive and non-positive FER problems, a common identification framework has been shown in most previous cases.

The work can be divided into two main steps, one is extraction of the function and the other classification structure. To extract face features, various image features were used in previous papers, such as local binary pattern (LBP) local phase quantification, directional gradient histogram, and scale-invariant function transformation (SIFT).

Among various facial features, SIFT shows encouraging performance due to its powerful features for image scaling, rotation, occlusion and light differences. Communication between two or more individuals can be in verbal or non-verbal language.

Non-verbal communication involves many different aspects as well as synonyms (physical and interpersonal space: distance and region), kinematics (body orientation: posture, postures and movements), appearance (physical attraction and clothing), touch (touch), secondary language (voice), intonation, speed, tone and volume of the speaker) and facial expressions (Sarter, 2010). All this non-verbal communication needs to be explained, they occupy more than 60% of the position in the communication process (Foley & Gentile, 2010). This nonverbal communication

© 2021 IJSRET

plays an important role in conveying emotions, attitudes and emotions.

Emotions in themselves arise from the physiological process of conscious and / or unconscious consciousness that stimulates any event or object related to a person's mental state, characteristics, and nature (College et al.,2014). Emotions play an important role in communication between individuals.

Personal feelings can affect relationships with others, such as the family, family and friends in the family, the work place, or other environments that connect with others. Emotions are expressed through the tone of voice, movements and body positions, where the most common are facial expressions.

II. LITERATURE REVIEW

Many prior imaging methods discard low-frequency components of the image to extract illumination variance for face recognition. However, this method can cause the processed image to be distorted and the effect is not good under normal lighting.

Although 3D facial imaging is becoming increasingly popular, many 3D facial imaging systems have obvious noise components. If meaningful recognition results are needed, post-processing is required to reduce these noise components. Biometric image recognition is the process of examining the closest matching area between images. Recognize the spatial pixels (pixels) in the image. Try to identify two different biological features, namely fingerprints and facial images. One of the biggest challenges facing the current face recognition technology (FR) is the difficulty of handling different positions and lighting.

In this article, we propose three new technologies, namely. Face recognition (FR) under changing lighting conditions and positions is extremely challenging. This article proposes a new method to improve the performance of the FR system. It uses a unique combination of active lighting equalization (AIE) and image grinding (IS). When the background, attitude and lighting change, the appearance of the human face will sharply Wide range. Changes to these conditions allow facial recognition (FR).

1. Discrete Emotion Perspective:
Each emotion has unique characteristics in terms of experience, physiology, and behavior (Panksepp, 2019). For this purpose, emotions are divided into several basic emotions, such as: fear, anger, joy, sadness, disgust and surprise. According to this view, basic emotions can be mixed, which will result in a large number of changes in a single emotional event. The mood setting depends on the theoretical background used. The basic facial expressions for Ekman 6 are one example (Ekman, 2019).

3. Dimensional Perspective:
Emotions are divided between three independent dimensions: joy-discomfort, rest-activation and relaxation-awareness. Recently, the two-dimensional approach has been adapted by assigning emotions to pleasant, uncomfortable, and rest-activated dimensions because these dimensions are sufficient to describe emotions.

Haifeng Zhang et.al (2019) Research shows that the facial expression recognition is strongly related to the person's identity. This paper presents an expression-identity fusion network to address the great inter-subject variations in facial expression recognition. The model is designed to jointly learn identity-related features and expression-related features via two branches with the same expression image input. A bilinear module is introduced to fuse two kinds of features and learn the relationship between them. Experimental results show that identity-related features can greatly boost the performance of facial expression recognition. Our method outperforms most of the state-of-the-art. On two popular facial expression databases (CK+ and Oulu-CASIA), our method achieves 96.02% and 85.21% recognition accuracy, respectively.

III. PROPOSED APPROCH

In this process, a facial descriptor is proposed that uses a local binary pattern algorithm (LBP) for the recognition of facial expression. Therefore, LBP is used to extract feature information from emotion-related features by using directional information and a ternary pattern to take fine edges in the face area while the face has a smooth area. By extracting a histogram-based face description method (this method breaks the face into small blocks and then displays the code evenly), this method performs better than other methods. Then, while sampling the information related to the expression at different scales, the grid used to construct the face descriptor is classified.

Reducing the number of dimensions by extracting distinctive features is based on the idea of maximizing total data distribution and minimizing differences within the class. It can be seen that the characteristic values of the six categories are highly merged, which can lead to a higher misclassification rate. Note that the actual number of functions can exceed three, but for the sake of intuition, the first three functions were selected for creation. Therefore, this work uses powerful features.

This is easy to explain, has good predictive power, and is computationally cheaper than other existing methods. As mentioned above, the existing HCRF uses the Gaussian distribution of diagonal covariance in the characteristic function, and cannot guarantee that its parameters converge to some specific value, in which the conditional probability is modeled as a mixture of functions of normal density. Due to this feature, the existing HCRF will lose a lot of information. This is one of the main disadvantages

© 2021 IJSRET

of the existing HCRF model. To resolve this limitation, we explicitly include the Gaussian distribution of full covariance in the observation level characteristic function.

For the classification module, a large number of methods have been adopted to perform an accurate expression classification. The author used the artificial neural network (ANN) to classify different facial expressions and obtained a recognition rate of 73%. However, the artificial neural network is a black box and does not have the full ability to clearly classify possible basic relationships. Also, artificial neural networks can take a long time to train and can fall into bad local minima.

Furthermore, the authors adopted a support vector machine (SVM) for their FER system. However, in support vector machines, the probability of observation is calculated using indirect techniques.

In other words, there is no direct estimate of probability. Additionally, SVM simply ignores the time dependency between video frames, making each frame statistically independent of the remaining frames.

To show the effectiveness of the SH-FER sub-assembly, three sub-experiments were carried out. For this reason, from the first experiment, the best case was selected according to the recognition rate. Three sub-experiments were then performed using the 10-fold verification rule. In the first case, ICA is used with HCRF instead of SWLDA. In the second case, the ICA is coupled to the LDA before the function is provided to HCRF.

We test and verify the accuracy of the crop / weed classification of the wavelet support and transformation vector machine (SVM) for the selective herbicide system in real time. Unlike the previous systems, the proposed algorithm involves a pre-processing step that helps eliminate the effects of lighting to ensure high precision in real life.

We tested a large number of wavelets and broke them into four levels to classify weed images into broadleaf and narrowleaf weeds. Reduce the function space by extracting only the most significant functions. Finally, SVM provides functions for classification. The method was tested on a database of 1,200 samples, which is much larger than the previously studied database (200-400 samples).

Pre-processing refers to "preparing" samples / images for insertion into algorithms for specific tasks: target tracking, recognition, feature extraction, etc. Data pre-processing is a data mining technique that involves converting raw data into an understandable format. . Data in the real world is generally incomplete, inconsistent, and / or lacks certain behaviours or trends, and can contain many errors. Data pre-processing is an effective method of solving such problems.

We can convert the data file (.xlsx) to a .mat file and align the data. The purpose of optimization is to achieve a "better" design with respect to a set of criteria or priority constraints. Choosing the correct function is very important to the ground model, which will seriously affect the results. The quality and quantity of the function have a great influence on whether the model is good. To improve the prediction model, different data pre-processing techniques are applied. Data integration is combining data from multiple sources into consistent storage. In embedded processing, redundant data problems always occur because the same real-world entity, attribute values from different sources has different names, or an attribute may be a derived attribute in another table. Therefore, researchers should use correlation analysis to carefully identify real-world entities from multiple data sources. Otherwise, careful integration of data from multiple sources can help reduce/avoid redundancies and in consistencies, and improve the speed and quality of mining [9].

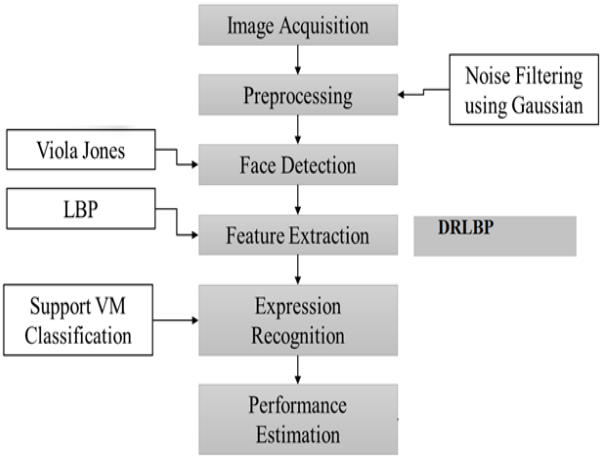


Fig 1. Proposed flow Diagram.

The proposed system has an architecture. It consists of different modules depending on the connection of the input image, and the facial database shows the method of storing the recognition results. The image obtained from the dataset library, the first module deals with "resize and normalize", and then pre-processes, the pre-processing module converts the color image to its grayscale representation and then performs detection of faces.

In the next step, the detected face is further removed from the image. This module can also detect the face of the object in the detected face area and transform and resize the face. The feature extraction module is used to create facial representations. Detects SIFT key points and creates a set of SIFT descriptors for facial image representation.

The following module is used to select the most representative face vector to create an accurate face model, then implement the algorithm in this module, and then perform face recognition. The recognized face is

© 2021 IJSRET

compared to the face model stored in the "face gallery", and the most similar model is selected as the recognized face model.

The final confidence measurement module is dedicated to identifying whether the recognition result is correct and then classifying it.

1. Binary Image:
This is the simplest type of image with two gray values (0 and 1 or black and white). Each pixel is represented by a bit. These types of images are useful in computer vision applications that only require information about images or contours. You can create an image from a grayscale image. The grayscale image uses 0 for pixels with grayscale values below the threshold and 1 for other pixels, but this method of creation is useless because most of the information is lost and the image results are more little ones.

2. Grayscale Images:
These images contain brightness information. The number of bits used to represent each pixel is related to the number of different brightness levels available. A typical image contains 8 bits per pixel, so there may be 256 shades of gray (Ng) or different intensity values, ranging from 0 to 255.

3. Colour Image:
Generally, the image is expressed as an RGB model (red, green, blue), each pixel in the model has 24 bits. In many applications, the brightness information and colour information are coupled and represented. The RGB information is passed to a mathematical function to separate the two information.

4. System Implementation:
- Image Aquistation
- Pre-processing
- Face Detection
- Feature Extraction
- Expression Recognition
- Performance Estimation

4.1 Image Aquistation: Use the imread command to read the image into the workspace. This example reads one of the sample images included with the toolbox, the image, and stores it in an array named I. imread infers from the file that the graphics file format is "tagged image file format (TIFF) " Imshow function You can also view images in the "Image Viewer" application.

The imtool function opens the Image Viewer application, which provides an integrated environment for displaying images and performing some common image processing tasks. The "Image Viewer" application provides all of imshow's image viewing functions, and also provides access to various other tools for navigating and exploring images, such as scroll bars, "pixel area" tools, "image information" tools image "and" Contrast Adjustment "Tool.

4.2 Pre Processing:
- Noise reduction is the process of removing noise from the signal.
- All recording devices, including analog and digital recording devices, have characteristics that make them susceptible to noise. The noise can be random or white without coherence, or it can be coherent noise introduced by the device's processing mechanism or algorithm.
- In electronic recording equipment, the main form of noise is hiss, which is caused by random electrons, which can deviate from their designated path under the severe influence of heat. These stray electrons will affect the output signal voltage, thereby producing detectable noise.

4.3 Face Detection:
- Face detection is a computer technology used in a variety of applications to recognize faces in digital images.
- Face detection also refers to the psychological process of humans locating and caring for faces in visual scenes.
- The Viola-Jones Object Detection Framework is the first object detection framework proposed by Paul Viola and Michael Jones in 2001 to provide competitive real-time object detection rates.
- Although he can be trained to detect various categories of objects, his main motivation is face detection.

4.4 Feature Extraction: Pattern recognition is a branch of machine learning that focuses on pattern recognition and data regularities, although in some cases it is considered almost synonymous with machine learning.
- Local Ternary Mode (LTP) is an extension of Local Binary Mode (LBP). Unlike LBP, it does not limit the pixel threshold to 0 and 1, but uses a threshold constant to limit the pixel threshold to three values.
- Use k as a threshold constant, use c as the value of the center pixel, and use the adjacent pixel p as a result of the threshold Classification:

4.4.1 Multiple Support Vector Machine: In machine learning, support vector machines (SVMs, also support vector networks) are supervised learning models with related learning algorithms that analyze data for classification and regression analysis. Given a set of training examples, each training example is marked as belonging to one or the other of the two categories, the SVM training algorithm will build a model that assigns the new examples to one category or another category, making it become a classifier non-probabilistic binary linear (although the method has problems like Platt's scale to use SVM in the probabilistic classification configuration). The SVM model represents the examples as points in space

© 2021 IJSRET

and maps them so that the examples in each category are separated by as wide a gap as possible. Then assign the new examples to the same space and predict which category they belong to based on the side of the hole they fall into.

In addition to performing a linear classification, SVM can also effectively perform a non-linear classification using so-called kernel techniques, implicitly mapping its input into a high-dimensional feature space.

If the data is not tagged, supervised learning cannot be done, and supervised learning requires a learning method that tries to naturally group the data into groups and then assign the new data to these formed groups. An improved grouping algorithm for support vector machines is called support vector grouping, which is generally used in industrial applications. It does not mark data or only marks some data as a pre-processing of the classification process.

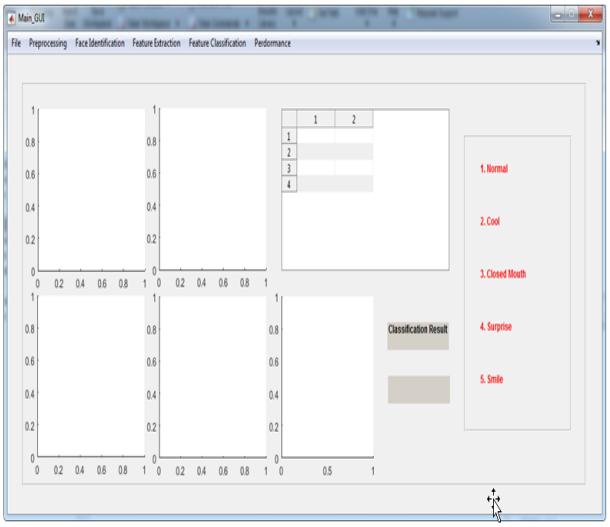


Fig 2. GUI Window.

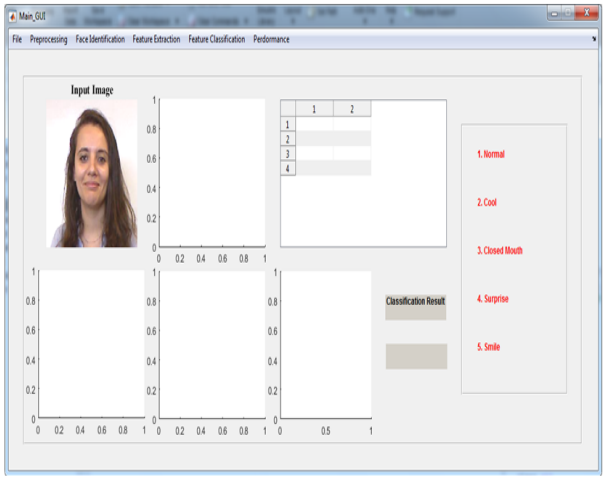


Fig 3. GUI Window Input Image.

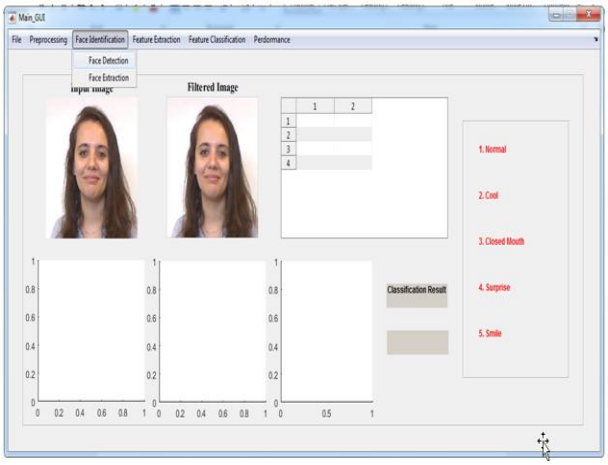


Fig 4. GUI Window Filter Image.

4.5 Expression Recognition: Large amounts of data are used to represent an image, so the analysis of an image needs a Lots of memory, so it takes longer. To reduce the amount of data, use a set of functions to represent the image. Feature extraction is the original type of pattern recognition, which is very important for pattern recognition. This step extracts some characteristics, such as colour, shape, and texture.

Features contain image-related information and will be used for image processing (for example, search, retrieval, storage). It is divided into two parts, Feature Construction and Feature Selection. The elements are divided into different categories according to the types of attributes they describe. Some important functions are as follows.

From a mathematical point of view, the color signal is an extension of the scalar signal to the vectoamr signal. Colours can be represented by average colors (three scalars) or color histograms (three functions). Color characteristics can be derived from the image histogram. Multidimensional histograms are used for color images. From the luminance histogram, the color characteristics are derived as RGB components.

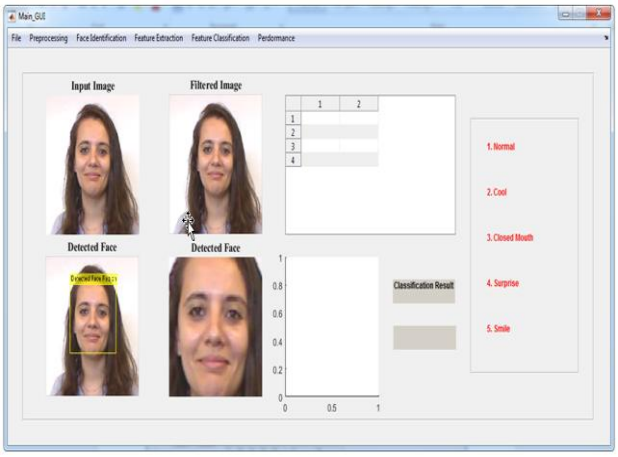


Fig 5. GUI Detected Image.

© 2021 IJSRET

This module also detects the face of object in the detected face region and transforms and resizes the face. The feature extraction module is used to create the face representation.

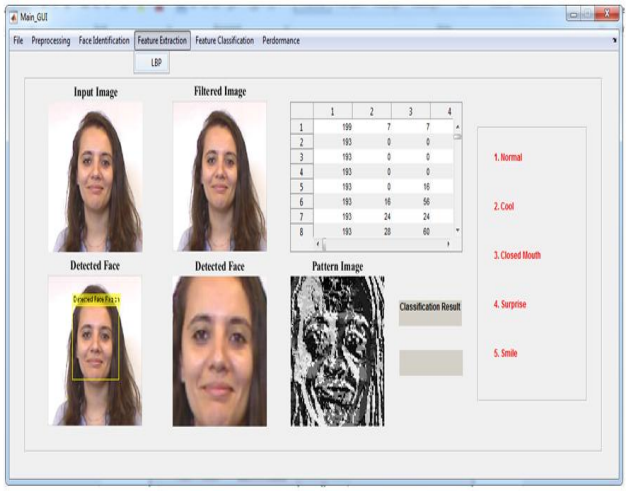


Fig 6. GUI Result Window.

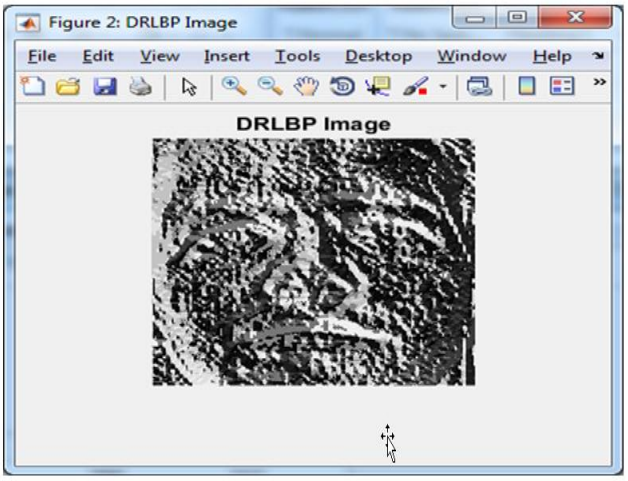


Fig 7. DRLBP Window.

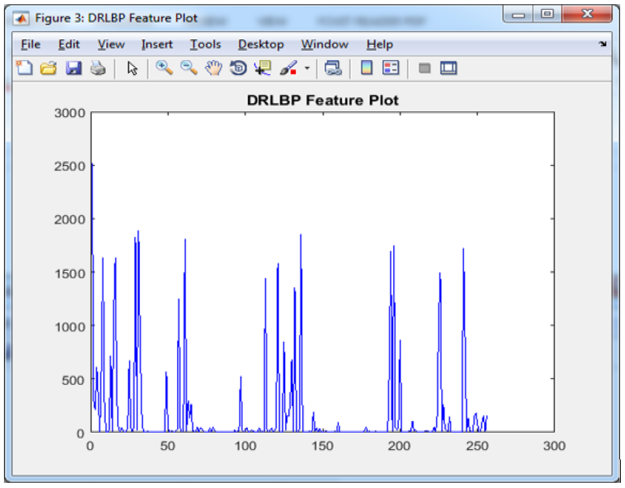
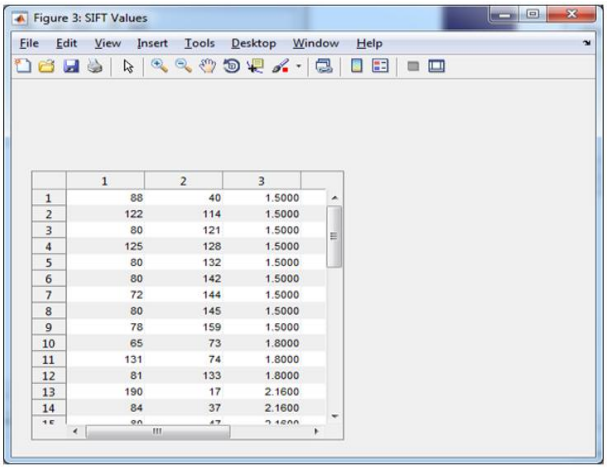


Fig 8. DRLBP Histogram.

Rotational invariance is obtained by calculating descriptors relative to a reference in a local neighbour-hood. References make calculations fast and keep calculations in Local Binary Mode (LBP) simple. The proposed method not only preserves the complete structural information extracted with LBP, but also captures complementary information using amplitude information, thereby gaining greater discriminatory power. For selecting function.

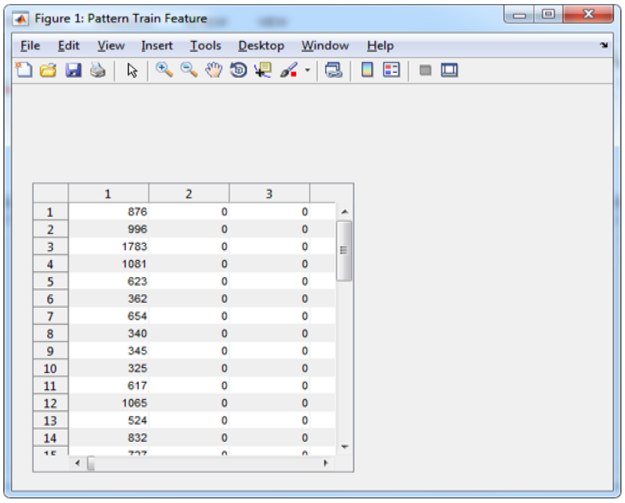


	1	2	3
1	88	40	1.5000
2	122	114	1.5000
3	80	121	1.5000
4	125	128	1.5000
5	80	132	1.5000
6	80	142	1.5000
7	72	144	1.5000
8	80	145	1.5000
9	78	159	1.5000
10	65	73	1.8000
11	131	74	1.8000
12	81	133	1.8000
13	190	17	2.1600
14	84	37	2.1600

Fig 9. SIFT Values Window.

4.5.1 Classification: The purpose of the classifier is to compare the characteristics of the test face image with the characteristics of the template and to make a final decision based on some similarity measures. For facial recognition, the most widely used classifier is the nearest neighbour classifier (NN).

An important issue in designing NN classifiers is how to measure similarity. The direct method is to measure the similarity of the two compared images. Another possible method is to calculate the distance between two image characteristics.



	1	2	3
1	876	0	0
2	996	0	0
3	1783	0	0
4	1081	0	0
5	623	0	0
6	362	0	0
7	654	0	0
8	340	0	0
9	345	0	0
10	325	0	0
11	617	0	0
12	1065	0	0
13	524	0	0
14	832	0	0

Fig 10. GUI feature value Window.

© 2021 IJSRET

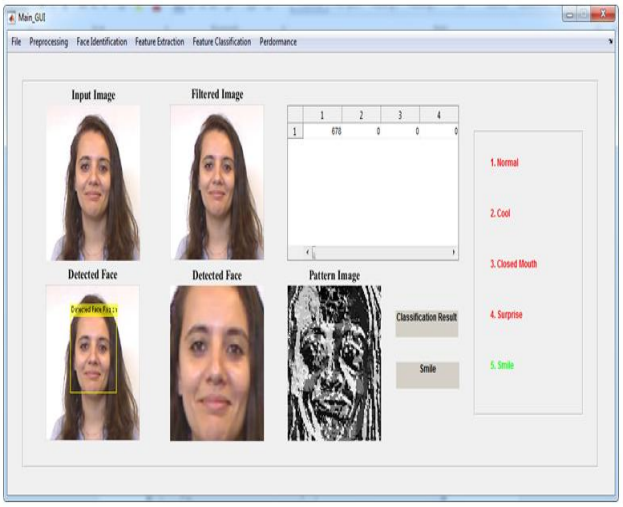


Fig 11. GUI smile image.

The last confidence measure module is dedicated to identifying whether the recognition result is correct or not followed by classifier.

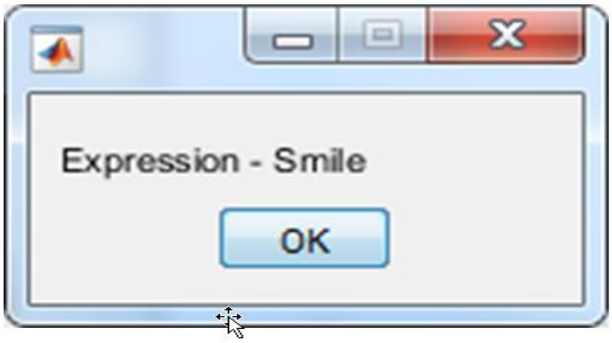


Fig 12. GUI Result Window.

This module is used to select the most representative face vector to create an accurate face model and implement algorithms in this module.

4.5.2 Key Point Matching: The next step is to apply these SIFT techniques to video frame sequences for object tracking. Filter characteristics pass through the continuous set of data characteristics and are retained by their key-point descriptors.

Each key point is assigned 4 parameters, which are 2D area (x direction and y coordinate), introduction and scale.

By independently considering each feature point found in the new feature table layout to each feature point on the object of interest, each item is followed in another feature edge grouping. Euclidean separation is expressed as a comparable estimate of the characteristics of the trait.
When the Euclidean distance between the two characteristics is greater than the edge specified in the past, the applicant can be protected. Therefore, the best match can be selected by parameter estimation, and on another route, the consistency of your region can be selected,

Table 1.Confusion Matrix and Automatic Facial Expression Recognition.

	Normal	Cool	Surprise	Closed Mouth	Smile
Normal	25	06	0	0	0
Cool	02	21	0	1	0
Surprise	20	10	25		
Closed Mouth	10	10	0	24	4
Smile	0	10	2	1	25

IV. CONCLUSION

Goal tracking and tracking is an important task in computer vision. In object detection and tracking, it contains two main processes, namely object detection and object tracking. The target detection in the video image obtained from a single camera with a static background means that the camera is attached by background subtraction.

In this article, we try to use a single camera fixed object and multiple objects to shoot different videos to understand that it can detect the object. A motion-based system for detecting and tracking a given moving object of interest can be created. The SIFT function is used to extract the first function of the object and it is detected that the frame matches the interesting object. As the SIFT algorithm is used in feature extraction, the tracker representation of the object remains unchanged.

We propose a model that primarily improves current work and finds that local facial features and expressions are interconnected. We divide the facial area according to the theory of cognitive neurology to identify and compare with the theory.

Due to occlusion and other complex conditions, this is beneficial for future recognition of facial expressions. We use LBP codes to describe images and consider the distance between image distribution points in binary space that cannot be measured directly at European distance.

Finally, the basic emotion classification model based on dimensional space theory and the circular emotion model is discussed and compared. In this section, we propose a proposed method for texture classification tasks. First, we introduce the RLBP operator obtained by cyclically shifting the weight of the LBP operator.

© 2021 IJSRET

In addition, the inherent structure of the pattern is used by combining the principle of a uniform pattern to generate a uniform RLBP (uRLBP).

The main direction is defined as the index of the circular neighbourhood with the largest difference. As the image is rotated, the nearby dominant direction is also rotated at the same angle. In the future, we can expand our work on detecting moving objects with a non-static background and having more cameras that can be used to monitor real-time applications.

REFERENCE

[1] Face Recognition Using DRLBP and SIFT Feature Extraction M. Sushama E.Rajinikanth 2018 International Conference on Communication and Signal Processing (ICCSP)
[2] Representation of Pose Invariant Face Images Using SIFT Descriptors Nthabiseng Mokoena ; Kishor Nair2018 International Conference on Advances in Big Data, Computing and Data Communication Systems (icABCD)
[3] PNN-SIFT: An enhanced face recognition and classification system in image processing W. Sylvia Lilly Jebarani ; T. Kamalaharidharini 2017 4th International Conference on Electronics and Communication Systems (ICECS)
[4] Robust face recognition and classification system based on SIFT and DCP techniques in image processing W. Sylvia Lilly Jebarani ; T. Kamalaharidharini 2017 International Conference on Intelligent Computing and Control (I2C2)
[5] Image matching with SIFT feature Rajkumar N. Satare ; S. R. Khot 2018 2nd International Conference on Inventive Systems and Control (ICISC)
[6] 6 B. Moghaddam, C. Nastar and A. Pentland, “A Bayesian similarity measure for deformable image matching,” Image and Vision Computing, vol. 19, no. 5, pp. 235-244, 2001.
[7] B. Shan, “A Novel Image Correlation Matching Approach,” JMM, vol. 5, no. 3, 2010.
[8] David G Lowe, “Distinctive Image Features from Scale-Invariant Keypoints,” International Journal of Computer Vision, vol.50, No. 2, 2004, pp.91-110.
[9] 9.E. Karami, M. Shehata, A. Smith, “Image Identification Using SIFT Algorithm: Performance Analysis Against Different Image
[10] Deformations,” in Proceedings of the 2015 Newfoundland Electrical and Computer Engineering Conference,St. john’s, Canada, November, 2015.
[11] 10 M. Güzel, 'A Hybrid Feature Extractor using Fast Hessian Detector and SIFT', Technologies, vol. 3, no. 2, pp. 103-110, 2015.
[12] Liang-Chi Chiu, Tian-Sheuan Chang, Jiun-Yen Chen and N. Chang, 'Fast SIFT Design for Real-Time Visual Feature Extraction', IEEE Trans. On Image Process., vol. 22, no. 8, pp. 3158-3167, 2013.
[13] A. Mehrabian, J. A. Russell, ”An approach to environmental psychology”, Cambridge, MA, US: The MIT Press, 1974.
[14] P. Ekman, W. V. Friesen, ”Facial action coding systema technique for the measurement of facial movement”, Palo AltO: Consulting Psychologists Press, pp. 271-302, 1978.
[15] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, ”The Extended Cohn-KanadeDataset (CK+) A complete dataset for action unit andemotion-specifiedexpression”, IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp. 94-101, 2010.
[16] M. Turk and A. Pentland, ”Eigenfaces for Recognition”, Cognitive Neuroscience, vol. 3, no. 1, pp. 71-86, 1991.
[17] P. N. Belhumeur, J. P. Hespanha and D. J. Kriegman, ”Eigenfaces vs. Fisherfaces”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 711-720, 1997.
[18] K. Etemad and R. Chellappa, ”Discriminant Analysis for Recognition of Human Face Images”, Optical Society of America A, vol. 14, no. 8, pp. 1724-1733, 1997.
[19] L. Wiskott, J. M. Fellous, N. Kuiger, and C. V. Malsburg, ”Face Recognition by Elastic Bunch Graph Matching”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 775-779, 1997.
[20] M. S. Bartlett, J. R. Movellan and S. Sejnowski, ”Face Recognition by Independent Component Analysis”, IEEE Transactions on Neural Networks, vol. 13, no. 6, pp. 1450-1464, 2002.
[21] J. Yang, D. Zhang, A. F. Frangi, and J. Yang, ”Two-dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no. 1, pp. 131-137, 2004.
[22] D. J. Kim, K. W. Chung and K. S. Hong, ”Person authentication using face, teeth, and voice modalities for mobile device security”, IEEE Transactions on Consumer Electronics, vol. 56, no. 4, pp. 2678-2685, 2010.

© 2021 IJSRET