Article

Necessary Morphological Patches Extraction for Automatic Micro-Expression Recognition

Yue Zhao *[ID] **and Jiancheng Xu**

School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710129, China; xujchg@nwpu.edu.cn
*   Correspondence: zy1990@mail.nwpu.edu.cn; Tel.: +86-158-9130-3906

check for updates

Featured Application: Micro-expression recognition is calculated to attract wide attention in the fields of psychology, mass media, and computer vision. Artificial identification of micro-expressions is not only time-consuming but also inefficient. This paper focuses on the realization of micro-expressions in a real-time computer system.

Abstract: Micro expressions are usually subtle and brief facial expressions that humans use to hide their true emotional states. In recent years, micro-expression recognition has attracted wide attention in the fields of psychology, mass media, and computer vision. The shortest micro expression lasts only 1/25 s. Furthermore, different from macro-expressions, micro-expressions have considerable low intensity and inadequate contraction of the facial muscles. Based on these characteristics, automatic micro-expression detection and recognition are great challenges in the field of computer vision. In this paper, we propose a novel automatic facial expression recognition framework based on necessary morphological patches (NMPs) to better detect and identify micro expressions. Micro expression is a subconscious facial muscle response. It is not controlled by the rational thought of the brain. Therefore, it calls on a few facial muscles and has local properties. NMPs are the facial regions that must be involved when a micro expression occurs. NMPs were screened based on weighting the facial active patches instead of the holistic utilization of the entire facial area. Firstly, we manually define the active facial patches according to the facial landmark coordinates and the facial action coding system (FACS). Secondly, we use a LBP-TOP descriptor to extract features in these patches and the Entropy-Weight method to select NMP. Finally, we obtain the weighted LBP-TOP features of these NMP. We test on two recent publicly available datasets: CASME II and SMIC database that provided sufficient samples. Compared with many recent state-of-the-art approaches, our method achieves more promising recognition results.

Keywords: micro-expression; local movement; necessary morphological patches (NMPs); local binary pattern; feature selection

1. Introduction

Micro-expression is a brief, involuntary, and external representation of a real emotion that can be exploited to determine the "real" behaviors and feelings of an individual [1]. Micro-expression was first discovered by psychologists Ekman and Friesen in 1969 [2]. Compared with ordinary facial expressions, micro-expressions have three significant characteristics: short duration, (generally lasts 1/25 to 1/3 s), low intensity, and (usually) local movement [3]. Based on these characteristics, micro-expressions are very difficult to detect by human beings [4]. Only highly trained individuals may be able to distinguish them, but even with proper training, the accuracy of their recognition falls below 50% in general [5]. However, it is very crucial to have the capability to detect and recognize

micro-expressions in many areas, such as psychological and clinical diagnosis, police interrogation, and national security [6]. Up till now, there have been plentiful research works for this area in the literature [7–13], in which computer-aided techniques have been established to automatically recognize micro-expressions.

According to the known research results [8–10], extracting features from the whole facial region can be found in most experiments, thus generating a lot of unnecessary redundant features so that the efficiency of recognition is greatly reduced. Psychologists have also found that micro-expressions tend to be partial movements that do not appear in the upper and lower portions of the facial region simultaneously [5] and are generally concentrated near the region of eyes, nose, and mouth. Ekman proposed that facial expression represents slight changes in several discrete facial motion units [5]. He also established Facial Action Coding System (FACS) [14] to describe the relationship between facial muscle changes and emotional states. This system illustrates that facial expressions are associated with subtle changes in several action units (AUs). Compared with ordinary expression, micro-expressions have extremely brief facial representations and invoke less muscle action units. For example, even without the involvement of eyebrows and mouth, one can express their inner state of surprise only by raising the upper eyelids. Later on, the theory of necessary morphological patches (NMPs) for micro-expression has established [15]. NMPs denote that some regions are indispensable in micro expression. For example, in all facial representations of disgust, the eyebrows and mouth do not need to be raised and open respectively, but the upper lip must be raised and appearing nasolabial sulcus on both sides of the nose. Figure 1 shows the NMPs of disgust, which are the indications of micro-expression to judge whether a person has a disgust emotional state. Almost all emotional information of micro-expressions is concentrated in these patches, therefore we need to separate these patches from the whole facial area.

| 79 | 99 | 109 | 133 | 156 (frames) |
| (onset) | | (apex) | | (offset) |
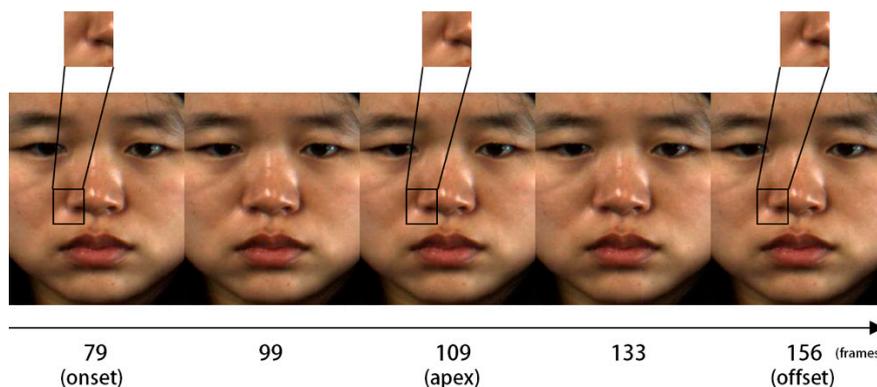
Figure 1. One necessary morphological patch (NMP) of disgust, micro-expression sequence.

In this paper, we aim to find some NMPs that play a key role in micro-expression recognition, and use these patches to train and learn how to identify micro-expression sequences. Automatic facial landmark detection is the first step in this work. This technique can detect the face in the video sequences, and adjust the position of the face in order to cut out the facial region. Face alignment is applied to extract facial active patches. In this study, in order to locate the facial active patches more accurately, we chose the algorithm of 68 landmarks [10]. After finding these patches in the face, we calibrated the active patches of the eyebrows, eyes, nose, and mouth based on the FACS criterion and landmarks technology. Later, we manually cut out 18 active patches, known as regions of interest (ROI) [13] in the whole face area. Moreover, we need to extract more effective features from these active patches [16–19]. Many existing works apply optical flow [12] and Local Binary Patterns from Three Orthogonal Panels (LBP-TOP) [8] algorithm to extract feature form dynamic micro-expression video sequences. The optical flow method reflects the close correlation between the frames by calculating the two adjacent frames. However, the changes between adjacent frames in micro-expressions are very weak and hence the algorithm does not reflect the changes of the facial

active patches. The LBP-TOP algorithm analyzes image texture from temporal and spatial pattern. Texture is a feature that shows the spatial distribution property of pixels, and can convey the necessary information of micro-expression. In addition, it can also display the local structural information of facial images. Compared to ordinary expressions, micro-expressions call for less muscle motion to convey emotions and to evaluate the current emotional state we can we identify only some NMPs.

To reduce the dimensionality and improve the recognition efficiency, we use the Entropy-Weight method to screen the NMPs, which are essential for micro-expression recognition from 18 active patches. The concept of entropy is first introduced into information theory by Shannon, which describes the size of the average information amount of events. The entropy weight has been widely used in engineering, social and economic fields, whose basic idea is to determine objective weights according to the variability of indexes. In general, the smaller information entropy of an index indicates the greater the variability of the index value. Therefore, this index affects more in terms of comprehensive evaluation, and its weight will be greater. In order to assess the contribution of 18 active patches to micro-expression recognition, we use Entropy-Weight method to evaluate its weights. Entropy-Weight method can not only filter out NMPs from active patches but weight these patches, thus increasing the discriminative ability of our algorithm. Finally, the multi-class SVM classifier is used to identify these NMPs, and the recognition rate is obtained from CASME II and SMIC databases.

The rest of this paper is organized as follows. The next section describes the related work on facial landmark detector and NMP selection, feature extraction, and weighting these necessary regions. Section 3 illustrates the particulars of the databases and discusses the experimental results in detail. Finally, Section 4 concludes the paper.

2. Methods

The subtle local movement of the facial muscles allows the facial expression to change and involve the relative positions of facial landmarks. The texture information of these regions also changes as the expression changes. In this paper, we aim at exploring different facial areas towards the recognition accuracy based on the subtle, local qualities of micro-expression. In other words, our goal is to identify the crucial facial areas corresponding to different emotions of the micro expression. The framework of the proposed algorithm is shown in Figure 2.
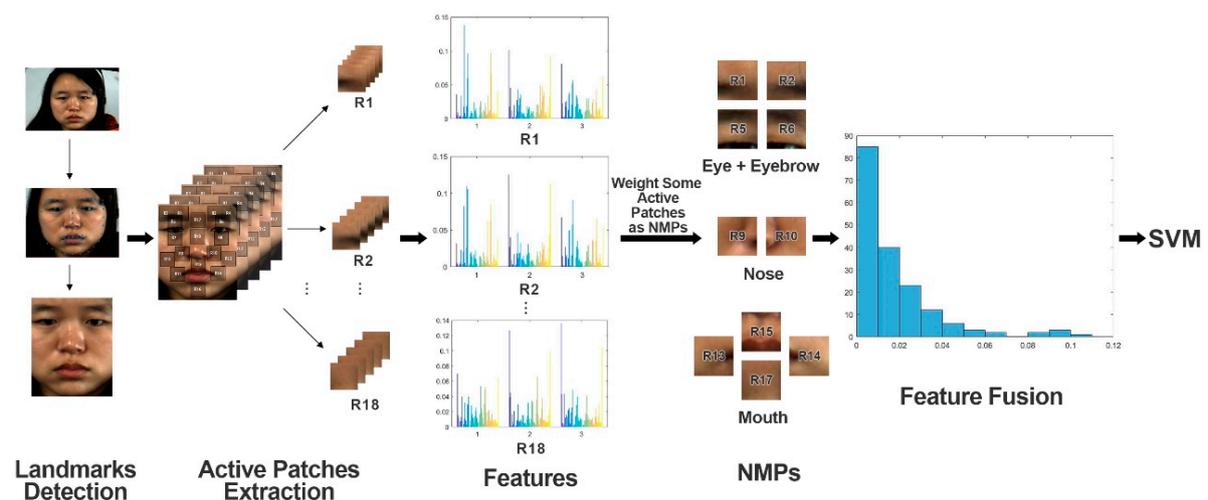

Figure 2. The proposed algorithm framework.

2.1. Facial Landmark Location

The goal of facial landmark detection is to accurately locate the key points of the face through the detection algorithm. The landmarks generally refer to the points around eyes, eyebrows, nose, mouth, and face contour. Studies have shown that the active facial areas are mainly

concentrated in the interlaced area of eyebrows and nasal bridge, as well as the corners of the eyes and mouth. In this paper, we firstly use face detection and landmarks detection to accurately locate the active facial patches. After that, we cut out the regions and extract the necessary features. Therefore, in order to get better location effect in active facial patches, 68 landmarks algorithm is used to calibrate micro-expression sequence.

To the best of our best knowledge, there are many machine learning methods to locate 68 landmarks, such as Active Appearance Model (AAM), Active Shape Model (ASM), and deep neural network algorithm, etc. Taking into account the real time and accuracy of position, we used ASM to localize the landmarks of micro-expression images [20]. The algorithm learns facial images, which are calibrated by using a training set, then the best matching points are searched on the test set and the landmarks of the face are located accordingly. We located facial landmarks in micro-expression images based on a previously published algorithm [21]. The 68 landmarks we drew on a facial image are shown in Figure 3. This method is applied in our algorithm and 68 landmarks are employed to align the active areas. These landmarks indicate the shape of eyebrows, eyes, nose, mouth, and the whole face, which are beneficial for researchers to cut the active patches.
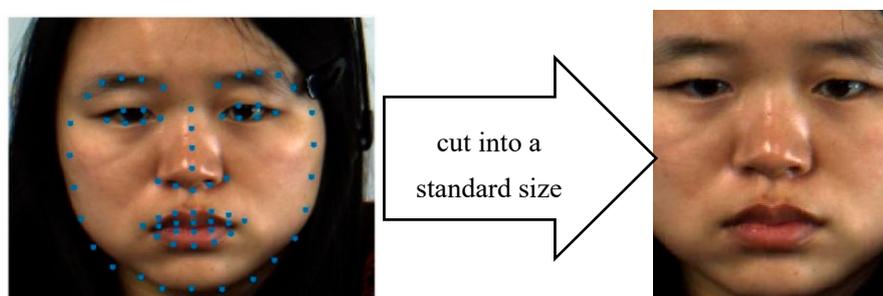

Figure 3. Landmarks on the face (68 total) and cut the image.
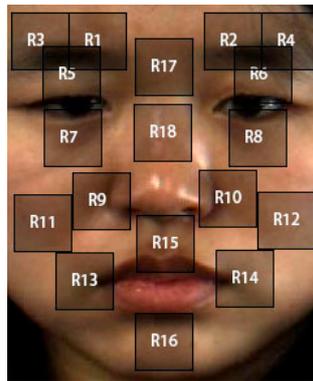
2.2. Extraction of Facial Active Patches

There are two main drawbacks in direct training of the classifier through the whole face: (1) The dimensions of the features are too large and the training time is relatively long; (2) Some regions on the face do not express emotion and contribute little to the representation of facial expressions. Hence, the features obtained from these regions are most likely to introduce noise.

The face must be partitioned appropriately for micro-expression recognition to be feasible [22]. The FACS criterion quantifies several muscle movements of the face and reveals 57 elementary components of the expression. These elementary components are known as the action units (AUs) and action descriptors (ADs). Similar to other facial expressions, a micro-expression is also a spatial combination of AUs. Each AU describes a local movement of micro-expression. Table 1 defines several relationships between AUs and facial movements [14].

Table 1. Relationships between AUs and facial movements.

ROIs	AUs	Facial Movements
R1, R2	AU1, AU4	inner eyebrows
R3, R4	AU2	outer eyebrows
R5, R6	AU5	upper eyelid
R7, R8	AU7	lower eyelid
R9, R10	AU6, AU10	side of the nose
R11, R12	AU6	cheeks
R13, R14	AU12, AU13	mouth corner
R15	AU15	upper lip
R16	AU17	chin
R17	AU9	eyebrow center
R18	AU9	nose root

Considering that micro-expressions only involve certain local muscle movements and AUs, extracting only a few active facial patches instead of the entire image of the face is an effective approach to recognition. The eyebrows and eyes, for example, are involved in nearly all basic emotions [23]. The morphological characteristics of the eyes and brow are important cues of different micro-expressions. The mouth is another key discriminant area for expression recognition. Here we manually choose a frontal neutral face image as a template, and divide the image into 18 ROIs, as shown in Figure 4.

Figure 4. Position of facial active patches.

The patches are separated according to the movements of micro-expressions. Each patch represents the active facial area of the micro-expression. We maintain the same size of each patch and extract the sequences of active patches for subsequent research.

2.3. Extraction Features

Micro expressions differ from ordinary ("macro") expression in regards to their low intensity, short duration, and local movements. It is unreasonable to use ordinary expression recognition methods to deal with micro-expression sequences. Here, we extend the classic LBP descriptor to a LBP-TOP to manage dynamic textures and events across spatial-temporal dimensions [24,25].

The LBP-TOP operator extends the LBP to three orthogonal planes, it was first proposed by Ojala et al. This operator reveals the local binary pattern of each image as well as the motion features of the spatial-temporal domain on the whole sequence. The LBP-TOP operator firstly divides the temporal and spatial domain into three orthogonal planes (XY, XT, and YT), then calculates the LBP values of the center pixels in each plane and eventually yields statistics of the expression information in three directions.

In practical applications, the spatial and temporal feature scales are different due to the unpredictable texture orientation, and the differences in the image resolution as well as frame rate. Here, we use an elliptical structure to define all neighboring points on the three orthogonal planes respective to the center point between frames, as shown in Figure 5.

The LBP code is extracted from the XY, XT, and YT planes and denoted as XY-LBP, XT-LBP, and YT-LBP. The statistics of three different planes were obtained for all pixels, and then concatenated into a single histogram [24]. In this paper, we extract LBP-TOP features and generate feature histograms for 18 active patches sequences. Only a few facial muscles are called on because of the micro expression. If 18 active patches are used to represent micro-expression sequences, the dimension of feature is too large, which makes the feature matching extremely complex and consumes too much system resources. Moreover, the movement range of the micro expression is much smaller than that of ordinary expression so that micro-expressions can be represented by some NMPs. For the next stage, we estimate some NMPs that are of significance for the micro expression from these 18 active patches.
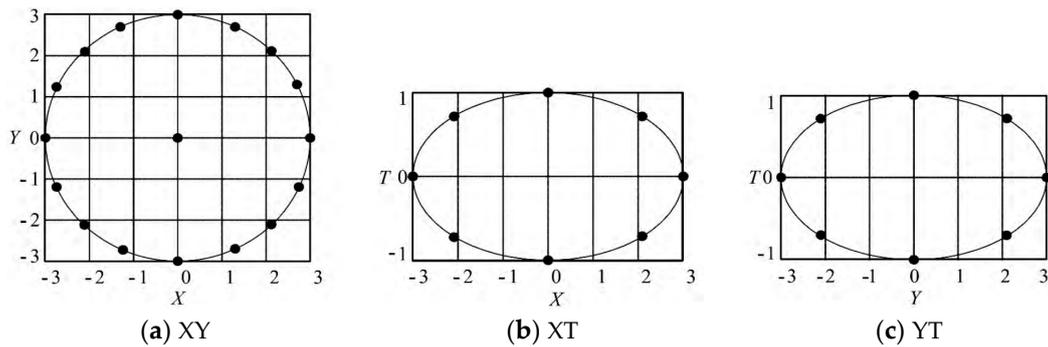
Figure 5. Different radius and numbers of neighboring points on three planes. (**a**) XY orthogonal planes; (**b**) XT orthogonal planes; (**c**) YT orthogonal planes.

2.4. Learning Crucial Facial Patches

Only a few NMPs play key roles in micro-expression recognition [5], because each active patch has a different importance for micro-expression recognition. For example, the eye area and the mouth area are highly distinguishable for people to express their emotion. Therefore, we should set different weights for each active patches so as to find out the NMPs and improve the subsequent recognition accuracy.

In this paper, the Entropy-Weight method is used to calculate the weights of each active patch and select the NMPs essential for micro-expression recognition. Information entropy can represent the information content of an image, and express the richness of the image texture. When an image is divided into many sub-patches, the local information entropy can partly reflect the quantity of information for each patch. We therefore can calculate the contribution of the texture feature of each patch based on local information entropy. The weight of the histogram of the patches is given by information entropy, which can absolutely embody the importance of each patch.

Information entropy of the local patch indicates the information contained in the pixel. The greater the amount of information, the more abundant texture information of the patch. Considering the strong discriminating ability of texture features to the expression details, our paper introduces the concept of entropy by using the entropy weight to express the NMPs weight.

The steps of determining the weight by the Entropy-Weight method are as follows:

Suppose that there are m objects to be evaluated and n evaluation indexes, the original data matrix of the image is as follows:

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix}_{m \times n}$$

(1) Step 1: Standardization of the original matrix, thus the normalized matrix is obtained.

where $R = \left(r_{ij} \right)_{m \times n}$ is the Establishing evaluation matrix.

$$r_{ij} = \frac{x_{ij} - min\{x_{ij}\}}{max\{x_{ij}\} - min\{x_{ij}\}}$$

where the formula is the standard value of the jth evaluation index on the ith evaluation object.

(2) Step 2: Calculating the proportion f_{ij} of index value of the i-th object the j-th index.

$$f_{ij} = \frac{r_{ij}}{\sum_{i=1}^{m} r_{ij}}$$

(3)　Step 3: Determining information entropy. In the case of *m* objects and *n* indexes, the information entropy of the *j*th index is defined as follows:

$$H_j = -k \sum_{i=1}^{m} f_{ij} \ln f_{ij}$$

where, $k = \frac{1}{\ln m}$.

(4)　Step 4: Defining entropy weight. The entropy weight of the *j*th index is defined as follows:

$$w_j = \frac{1 - H_j}{m - \sum_{j=1}^{n} H_j}$$

where, $0 \leq w_j \leq 1, \sum_{j=1}^{n} w_j = 1$.

2.5. Multi-Class Classification

In this study, we used Support Vector Machine (SVM) [25] as a classifier for micro-expression recognition. It projects feature vectors to a higher dimensional plane by nonlinear mapping and finds a linear hyperplane for classification. SVM is a linear two-class model that maximize the margin in the feature space maximum. Micro-expression recognition is a multi-classification problem, in this paper, so we used Leave One Sample Out Cross Validation (LOSOCV) and 10-fold cross validation to it. However, micro-expression recognition is a multi-classification problem, there are two common methods to solve this problem: one-versus-rest (OVR) and one-versus-one (OVO). In this paper, we use the OVO SVMs. The approach is to design a SVM between any two classes of samples, so we need to design $k(k - 1)/2$ SVMs. After that, when classifying an unknown sample, the sample will select the class with the largest number of votes. The advantage of this method is that it does not need to retrain all SVMs, but only needs to retrain and add classifiers related to the samples. In addition, we also need to use the kernel function to map the sample from the original space to a higher dimensional feature space, so that the sample is linearly separable in this feature space. The kernel functions include the linear kernel, polynomial kernel, and Radial Basis Function (RBF). In this paper, RBF kernel: $k(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right)$ is used as our classifier.

3. Experiments and Discussion

3.1. Datasets

Compared with the macro-expression, there are only a few micro-expression databases for investigation. In this section, we evaluate the proposed algorithm using CASME II [26] and SMIC [27] databases.

3.1.1. CASME II

The CASME II database was established in 2014 as an upgraded version of the CASME database [26]. The time resolution of the new database changed from 60 fps to 200 fps, and the spatial resolution increased to 280×340. The database was obtained under a strict laboratory environment and appropriate light conditions to a total of 247 micro-expression videos. The film clips either have a total duration with less than 0.5 s or an onset duration (i.e., time from onset to apex) with less than 0.25 s. The ground truth information includes the onset and offset frames, represented emotions, as well as the AUs being provided. It consists of five classes of emotions, namely, happiness (32 samples), disgust (60 samples), surprise (25 samples), repression (27 samples), and tense (102 samples).

3.1.2. SMIC

The Spontaneous Micro-Expression Database (SMIC) was designed by the Zhao's team of the machine vision research center of the University of Oulu, Finland [27]. The researchers asked subjects to watch movie clips, which induced disgust, fear, sadness, and surprise, while attempting to suppress their facial expressions. This experiment used a 100 fps camera. The resulting database includes 164 videos of 16 subjects. The micro-expressions have a maximum total duration of 0.5 s and the longest video sequence contains 50 frames. There are three main emotion categories: positive (happiness; 51 samples), negative (sad, fear, disgust; 70 samples), and surprise (43 samples).

3.1.3. Experiment Settings

Above all, we use ASM to locate 68 landmarks points in all the images of the micro-expression sequence, then cut the face area and the size of each frame is normalize to 164×196.

These databases are micro-expression sequences captured by high-speed camera for several different individuals. The frames in the database are different, which will degrade the recognition rate if different sequence samples are used to extract and classify the micro-expressions. With the reference to the literature, ref. [20] we use the time interpolation model (TIM) to normalize all the frames of the micro-expression sequences. Table 2 shows the relationship between the number of frames with the experimental time and accuracy. Based on the table below, the frames of all samples were normalized to 10.

Table 2. Relationship between time interpolation model (TIM) length with time and accuracy.

TIM Length	No TIM	8	10	20	30	40
time (s)	87.8	16.7	22.1	49.6	77.2	104.9
accuracy	54.84%	57.15%	59.05%	57.80%	57.40%	56.93%

We use a 68-point ASM to locate the facial key points and then to establish landmarks [21]. The whole face can be divided into active patches based on these landmarks. Eighteen facial active patches are generated based on the FACS rule and AUs. In addition, the number of frames of the active patches sequences is 10. And the size of all facial active patches keeps equal at approximately one-eighth of the width of the face, as shown in the Figure 6.
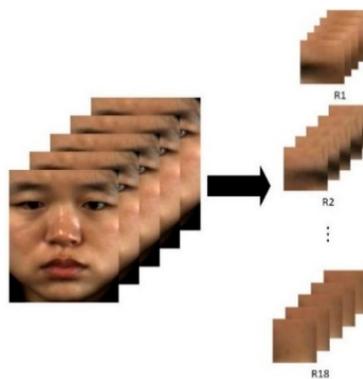
Figure 6. Active facial patches of micro-expression sequence.

In this paper, we use a LBP-TOP operator to extract the feature vector for micro-expressions recognition and the descriptor contains two important parameters: radius and neighbor points. For convenience, we wrote $LBP - TOP_{R_X, R_Y, R_T, P_{XY}, P_{XT}, P_{YT}}$ as R_X, R_Y, R_T; $P_{XY} = P_{XT} = P_{YT} = P$. The parameter comparison of the LBP-TOP algorithm is shown in Table 3.

Table 3. Recognition accuracies of different parameters (%).

R_X,R_Y,R_T, P	Accuracy	R_X,R_Y,R_T, P	Accuracy	R_X,R_Y,R_T, P	Accuracy
111, 4	46.20	111, 8	47.85	111, 16	47.02
331, 4	49.50	331, 8	50.33	331, 16	47.85
333, 4	45.37	333, 8	47.03	333, 16	47.68

In this paper, SVM is selected as the experimental classifier, and the choice of kernel function is very important for its performance. The experiments are conducted on the CASME II and SMIC databases by using weight LBP-TOP feature extraction methods. A SVM classifier with a kernel function was used to evaluate the proposed method. In this multi-subject level analysis, both LOSOCV and 10-fold cross validation are utilized to validate the effectiveness of the proposed method in all the experiments. In LOSOCV, the video sequence of one subject is treated as the testing data and the remaining frames as the training data. Such a process is repeated for *k* times, where *k* denotes the number of subjects in the database. Then the recognition results for all subjects are averaged to form the final recognition accuracy. In 10-fold cross validation, the data set is divided into ten parts, nine of which are taken as training data in turn, and the other one as test data. The correct rate is obtained from each test, and the average value of the correct rate is estimated as the accuracy of the algorithm. Finally, we do ten times 10-fold cross validation and get the mean value as the ultimate accuracy. In this paper, two common cross validation methods are used to evaluate the classification and recognition ability of SVM, as shown in Table 4.

Table 4. Recognition rate of different kernel functions on CASME II and SMIC database (%).

Kernel Function	CASME II		SMIC	
	LOSOCV	10-Fold	LOSOCV	10-Fold
SVM (polynomial)	64.35	61.90	60.92	59.04
SVM (linear)	62.89	60.02	58.57	58.39
SVM (RBF)	67.95	65.27	64.19	62.04

4. Results and Discussion

In this section, we conduct extensive experiments to evaluate the performance of the proposed micro-expression method on two widely-used micro-expression databases.

4.1. NMPs

The number of facial active patches also affects the performance and recognition rate. Figure 7 shows the relationship between the number of patches and the recognition rate. Even the use of a single crucial patch yields a recognition rate of 31.62%; the use of all 18 active patches produces a recognition rate of 56.93%. We found that the features of some unimportant patches do not play a significant role in identifying micro-expression. Instead of applying all 18 facial active patches, we can extract some crucial patches (NMPs) with discriminant ability for micro-expression recognition.

In the figure, N represents the number and location of NMPs. The micro expression only involves a few of facial active patches. The recognition rate reaches to the highest when the number of patches is 10. High values of N contain some extra patches that contribute less to subtle movement and micro expression recognition. Moreover, low values of N lost some important information.

The occurrence of micro-expressions is very weak, most of the movements focus on the corners of the eyes, eyebrows, nose, and mouth. The Entropy-Weight method as a feature selection algorithm can evaluate the importance of each feature on the classification problem. This paper compares the contribution of the Entropy-Weight method and other feature selection algorithms in the number of NMPs and recognition accuracy. The experimental results are shown in Table 5. Then, comparing with the data, each algorithm chooses the number of NMPs in the region of the eyes, eyebrows,

and mouth, which are basically the same while the number for the cheek and nose area is different. This is because the muscle movement of micro-expressions is mainly concentrated in the eye, eyebrow, and mouth regions, and the action units of the cheek and nose regions are very few. Micro-expressions are usually restrained facial movements, which are very subtle and easily overlooked. The Pearson coefficient is insensitive and misleading to these micro-expression areas because of the small correlation between the motions. The lasso model is very unstable, when the data changes slightly, it may lead to great changes in the model. The Entropy-Weight method has the advantages of high robustness and easy use, and the experimental results show that the NMPs selected by this method are basically in line with the most representative facial muscle motion patches proposed by psychologists when micro-expressions occur. In addition, the Entropy-Weight method can give weight to feature vectors, which can better represent the motion characteristics of micro-expressions in the classification process.
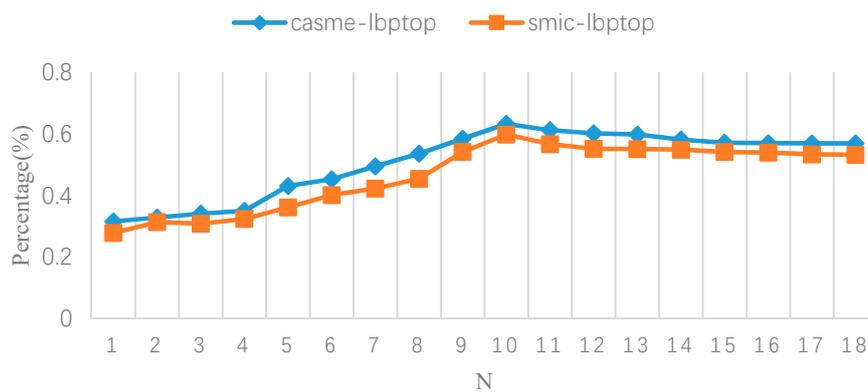

Figure 7. Recognition accuracy (%) of numbers of NMPs.

Table 5. Accuracy rate and NMPs numbers of different feature selection algorithms.

NMP Number	Pearson Coefficients	Mutual Information	Lasso	Entropy-Weight Method
eye + eyebrow	8	6	4	4
nose	0	3	0	2
cheek	0	0	10	0
mouse	5	8	5	4
Recognition rate	57.37%	55.09%	60.10%	67.95%

Each micro expression affects a few specific facial muscles. In other words, only part of the AUs are crucial for micro expression. In this paper, we use Entropy-Weight method to determine the location of NMPs. The optimum number of NMPs and location corresponding to different micro expressions are shown in Table 6.

Table 6. NMP of micro-expression in the CASME II database.

Emotion	Number of the Crucial Patches	Action Unit
Happiness	R9, R10, R13, R14	AU6 or AU12
Disgust	R17, R9, R10, R1, R2, R7, R8	AU9 or AU10 or AU4+7
Surprise	R1, R2, R3, R4, R5, R6	AU1+2 or AU2 or AU5
Repression	R15, R16	AU15 or AU17
Tense	R1, R2, R15, R16	AU4 or AU15 or AU17

The subtle muscle movements of micro expressions mainly concentrate in the patches of the eyes, the eyebrows, the alar sides, and the mouth according to the weight value derived. The proposed method chooses 10 patches (R1, R2, R5, R6, R9, R10, R13, R14, R15, R17) which get the highest weights of these regions as NMP.

4.2. Recognition Performance

Firstly, we use a LBP-TOP descriptor to extract features in the whole face area, 18 facial active patches, and 10 facial NMP, respectively. Following the original implementations [26,27], we employ $LBP-TOP_{888331}$ for CASME II and $LBP-TOP_{444113}$ for SMIC. We summarize the results in Table 7. It can be concluded from the experimental result that a higher recognition efficiency can be obtained by using some regions of NMPs, while a lot of redundant features are introduced by using the whole face area. Because of the local characteristic and low intensity, micro expressions involve a few muscle movements when people try to conceal and suppress their true emotions. Therefore, using 18 facial active patches to identify micro-expression can also increase the dimensions and complexities of features, and, meanwhile, reduce the accuracy.

Table 7. Recognition rate in different regions of micro expression images in the CASME II database (%).

Facial Region	LOSOCV	10-Fold
CASME II-whole face	50.72	48.15
CASME II-18 active patches	56.93	52.57
CASME II-10 NMPs	63.24	60.78
SMIC-whole face	45.04	43.98
SMIC-active patches	53.19	50.17
SMIC-10 NMP	59.80	56.01

Next, we calculate the information entropy to weight the active patches, and then obtain the 10 regions of NMPs with the greatest weight. The weights not only represent the importance of these patches for micro expressions, but also shows the information they could convey. Furthermore, the weighted values of these NMPs generate weighted LBP-TOP features for classifying and recognizing the micro-expressions. The results are shown in Tables 8 and 9.

Table 8. Comparison of dimension reduction results (CASME II).

Algorithm	Recognition Accuracy (%)	Between-Class Accuracy (%)				
		Disgust	Happiness	Repression	Surprise	Others
LBP-TOP	63.24	56	73	77	38	57
LBP-TOP + PCA	64.67	58	70	79	40	56
LBP-TOP + LDA	65.92	60	70	80	40	57
weighted LBP-TOP	67.95	60	72	83	42	60

Table 9. Comparison of dimension reduction results (SMIC).

| Algorithm	Recognition Accuracy (%)	Between-Class Accuracy (%)		
		Negative	Positive	Surprise
LBP-TOP	58.90	60	64	55
LBP-TOP + PCA	61.36	60	66	58
LBP-TOP + LDA	62.05	62	68	58
weight LBP-TOP	64.19	64	68	60

We compare the recognition rates between the dimensionality reduction algorithm and the original method. The weighted LBP-TOP achieves the highest accuracy rate. Compared with other traditional dimensionality reduction algorithms, this proposed algorithm has an obvious advantage that can directly explain the importance of NMPs for micro-expression recognition. The proposed method not only increases the significance of the features, but also improves the robustness of the algorithm. Although the traditional dimensionality reduction algorithm can reduce the high-dimension of the feature in a simple manner, it tends to lose some vital and useful information

for recognition. Because the micro expression is a subtle facial muscle movement, many useful features would be filtered out as redundant information when applying the traditional algorithm to identify them.

Table 10 shows the accuracy rate of different algorithms for micro expression recognition on the CASME II [28–31]. Our method achieves a higher recognition rate than the others. Compared with other algorithms, our proposed method is more effective in the extractions of different feature regions for different micro-expression sequences compared with other algorithms and it also reveals the discriminative feature vectors in each NMP more accurately. We get the recognition performances of all kinds of emotion in CASME II and SMIC databases as show in Figure 8.

Table 10. Micro-expression recognition rates (%) in CASME II database.

Method	Task	Recognition Rate
TICS	happiness, surprise, disgust, repression, others	61.76%
CUDA based DCNN	happiness, surprise, disgust, repression, others	64.90%
HIGO-TOP	happiness, surprise, disgust, repression, others	55.87%
HOG-TOP	happiness, surprise, disgust, repression, others	57.49%
The proposed method	happiness, surprise, disgust, repression, others	67.95%
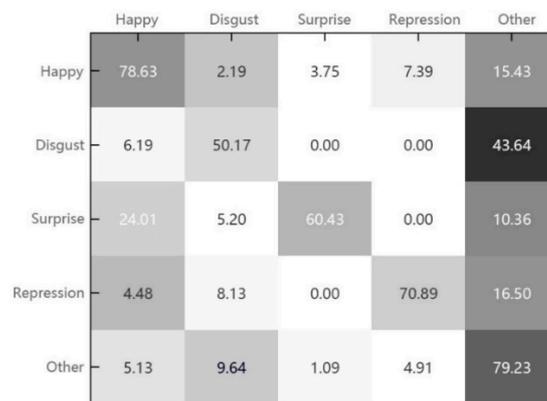
(**a**)

(**b**)

Figure 8. Confusion matrix for micro-expression recognition (**a**) Recognition rate of all emotion types on CASME II database (**b**) Recognition rate of all emotion types on SMIC database.

From the experiment results, the emotions of "surprise" and "disgust" acquire a lower recognition rate since they are very easy to cause confusion with other micro-expressions. This is due to using some single muscle action units or the combines of these units to convey them. Because of that, many NMP regions can express them, which shows that these two micro expressions lack uniqueness. Compared to

other emotional types, the stimulus of the two emotions are relatively simple, especially the emotion of "surprise", because many emotions begin with surprise. Happiness is easily recognized by a machine because it has unique muscle movements, particularly in the mouth area. Therefore, they are more likely to be captured by machines or humans when people are happy. Furthermore, for SMIC, the figure also illustrates the accuracy of three expressions: negative, surprise, and positive.

Table 11 shows the accuracy of different methods for micro-expression recognition [32–36]. Our algorithm outperforms the others on the SMIC database, which is benefit from the extraction of NMP on the whole face that effectively increases the micro-expression recognition rate.

Table 11. Micro-expression recognition rates (%) in SMIC database.

Method	Recognition Rate
CNN+SFS	53.60%
HIGO-TOP	57.93%
HGO-TOP	59.15%
FDM	54.88%
NMFL	62.33%
The proposed method	64.19%

5. Conclusions

Many research organizations pay close attention to the automatic recognition of micro expression in the field of computer vision. However, according to most literature, researchers have used the whole face region as a feature to test their algorithm. Psychologists have shown that micro expression is a kind of emotional category with local characteristics. Compared to facial macro-expressions, micro-expression is usually produced when people try to suppress their own emotions. Based on this situation, this is due to the inadequate contraction of facial muscles and involve less action units. This paper proposes a recognition method which exploits the local motion characteristics of micro-expressions.

In this paper, we studied the correlation between different facial regions. Based on the ground of the NMP of micro-expressions provided by a psychologist, we extract the active facial patches representing the features of facial deformation. After analyzing the active areas, we apply the Entropy-Weight method to identify some active patches as NMP. Thereby, we calculate the weighted LBP-TOP in these patches. These features reduce dimensions and improve recognition accuracy. Experiments on two public micro-expression databases demonstrates that our method achieves a remarkably high micro-expression recognition accuracy rate.

In addition, the proposed algorithm manually extracts 18 facial active patches, which may increase the complexity of the algorithm. Moreover, the artificial extraction of these patches is bound to have subjectivity, in future work we would be eager to construct an algorithm that automatically extracts the NMP regions.

References

1. Warren, G.; Schertler, E.; Bull, P. Detecting deception from emotional and unemotional cues. *J. Nonverbal Behav.* **2009**, *33*, 59–69. [CrossRef]
2. Yan, W.; Wu, Q.; Fu, X. How Fast are the Leaked Facial Expressions: The Duration of Micro-Expressions. *J. Nonverbal Behav.* **2013**, *37*, 217–230. [CrossRef]

3. Ekman, P.; Friesen, W. Detecting deception from emotional and unemotional cues. *Psychiatry* **1969**, *32*, 88–106. [CrossRef] [PubMed]

4. Shen, X.; Wu, Q.; Fu, X. Effects of the duration of expressions on the recognition of micro expressions. *J. Zhejiang Univ. Sci. B* **2012**, *3*, 221–230. [CrossRef] [PubMed]

5. Carroll, J.M.; Russell, J.A. Facial Expressions in Hollywood's Portrayal of Emotion. *J. Personal. Soc. Psychol.* **1997**, *72*, 164–176. [CrossRef]

6. Frank, M.; Herbasz, M.; Sinuk, K.; Keller, A. I see how you feel: Training laypeople and professionals to recognize fleeting emotions. In *International Communication Association*; Sheraton New York: New York, NY, USA, 2009.

7. Huang, X.; Wang, S.; Liu, X. Discriminative Spatiotemporal Local Binary Pattern with Revisited Integral Projection for Spontaneous Facial Micro-Expression Recognition. *IEEE Trans. Affect. Comput.* **2017**. [CrossRef]

8. Shreve, M.; Godavarthy, S.; Manohar, V. Towards macro-and micro-expression spotting in video using strain patterns. In Proceedings of the 2009 Workshop on Applications of Computer Vision (WACV), Snowbird, UT, USA, 7–8 December 2009; pp. 1–6.

9. Shreve, M.; Godavarthy, S.; Goldgof, D.; Sarkar, S. Macro-and micro-expression spotting in long videos using spatio-temporal strain. In Proceedings of the Face and Gesture AFGR, Santa Barbara, CA, USA, 21–25 March 2011; pp. 51–56.

10. Polikovsky, S.; Kameda, Y.; Ohta, Y. Facial micro-expressions recognition using high speed camera and 3d-gradient descriptor. In Proceedings of the 3rd International Conference on Imaging for Crime Detection and Prevention (ICDP 2009), London, UK, 3 December 2009.

11. Zhou, Z.; Zhao, G.; Pietikainen, M. Towards a practical lip reading system. In Proceedings of the Computer Vision and Pattern Recognition (CVPR), Colorado Springs, CO, USA, 20–25 June 2011; pp. 137–144.

12. Pfister, T.; Li, X.; Zhao, G.; Pietikainen, M. Recognizing spontaneous facial micro-expressions. In Proceedings of the 12th IEEE International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2011; pp. 1449–1456.

13. Patel, D.; Hong, X.; Zhao, G. Selective deep features for micro-expression. In Proceedings of the 2016 23rd International Conference on Pattern Recognition (ICPR), Cancun, Mexico, 4–8 December 2016; pp. 2259–2264.

14. Ekman, P.; Friesen, W. Facial Action Coding System (FACS): A Technique for the Measurement of Facial Actions. *Rivista Di Psichiatria* **1978**, *47*, 126–138.

15. Jiang, Z.Y. *Micro-Expression*; Changjiang Literature & Art Press: Wuhan, China, 2015; pp. 10–132.

16. Lajevardi, S.M.; Hussain, Z.M. Automatic facial expression recognition: Feature extraction and selection. *Signal Image Video Process* **2012**, *6*, 159–169. [CrossRef]

17. Happy, S.L.; Routray, A. Automatic facial expression recognition using features of salient facial patches. *IEEE Trans. Affect. Comput.* **2015**, *6*, 1–12. [CrossRef]

18. Liu, F.Y.; Cao, Y.; Li, Y. Facial expression recognition with PCA and LBP features extracting from active facial patches. In Proceedings of the IEEE International Conference on Real-time Computing & Robotics, Angkor Wat, Cambodia, 6–10 June 2016; pp. 368–373.

19. Koutlas, A.; Fotiadis, D. An automatic region based methodology for facial expression recognition. In Proceedings of the IEEE International Conference on Systems, Singapore, 12–15 October 2009; pp. 662–666.

20. Cootes, T.; Taylor, C.; Cooper, D.; Graham, J. Active shape models—Their training and application. *Comput. Vis. Image Underst.* **1995**, *61*, 38–59. [CrossRef]

21. Goshtasby, A. Image registration by local approximation methods. *Image Vis. Comput.* **1988**, *6*, 255–261. [CrossRef]

22. Shan, C.; Gong, S.; Mcowan, P.W. Facial expression recognition based on local binary patterns: A comprehensive study. *Image Vis. Comput.* **2009**, *27*, 803–816. [CrossRef]

23. Zhao, G.; Pietikainen, M. Dynamic texture recognition using local binary patterns with an application to facial expressions, Pattern Analysis and Machine Intelligence. *IEEE Trans. Pattern Anal. Mach. Intell.* **2007**, *29*, 915–928. [CrossRef] [PubMed]

24. Wang, Y.; See, J.; Phan, R.C.W. LBP with Six Intersection Points: Reducing Redundant Information in LBP-TOP for Micro-expression Recognition. In Proceedings of the Asian Conference on Computer Vision, Singapore, 1–5 November 2014; Volume 3, pp. 21–23.

25. Ojala, T.; Pietikainen, M.; Maenpaa, T. Multiresolution gray-scale and rotation invariant texture classification with Local Binary Pattern. *IEEE Trans. Pattern Anal. Mach. Intell.* **2001**, *24*, 971–987. [CrossRef]

26. Yan, W.; Li, X.; Wang, S.; Zhao, G.; Liu, Y.; Chen, Y. CASME II: An improved spontaneous micro-expression database and the baseline evaluation. *PLoS ONE* **2014**, *9*, e86041. [CrossRef] [PubMed]

27. Li, X.; Pfister, T.; Huang, X.; Zhao, G.; Pietikainen, M. A spontaneous micro-expression database: Inducement, collection and baseline. In Proceedings of the IEEE International Conference and Workshops on Automatic Face and Gesture Recognition, Shanghai, China, 22–26 April 2013.

28. Wang, S.; Yan, W.; Li, X.; Zhao, G.; Fu, X. Micro-expression recognition using dynamic textures on tensor independent color space. In Proceedings of the 2014 22nd International Conference on Pattern Recognition (ICPR), Stockholm, Sweden, 24–28 August 2014.

29. Mayya, V.; Pai, R.; Pai, M. Combining temporal interpolation and DCNN for faster recognition of micro-expressions in video sequences. In Proceedings of the 2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Jaipur, India, 21–24 September 2016; pp. 699–703.

30. He, X.; Cai, D.; Niyogi, P. Monogenic Riesz wavelet representation for micro-expression recognition. In Proceedings of the 2015 IEEE International Conference on Digital Signal Processing (DSP), Singapore, 21–24 July 2015; pp. 1237–1241.

31. Li, X.; Hong, X.; Moilanen, A.; Huang, X. Towards reading hidden emotions: A comparative study of spontaneous micro-expression spotting and recognition methods. *IEEE Trans. Affect. Comput.* **2017**. [CrossRef]

32. Zhang, S.; Feng, B.; Chen, Z.; Huang, X. Micro-expression recognition by aggregating local spatio-temporal pattern. In Proceedings of the International Conference on Multimedia Modeling (MMM), Reykjavik, Iceland, 4–6 January 2017; pp. 638–648.

33. Ngo, A.; Phan, R.; See, J. Spontaneous subtle expression recognition: Imbalanced databases and solutions. In Proceedings of the Asian Conference on Computer Vision (ACCV), Singapore, 1–5 November 2014; pp. 33–48.

34. Xu, F.; Zhang, J.; Wang, J. Micro-expression identification and categorization using a facial dynamics map. *IEEE Trans. Affec. Comput.* **2016**, *8*, 254–267. [CrossRef]

35. He, J.; Hu, J.; Lu, X.; Zheng, W. Mult-task mid-level feature learning for micro-expression recognition. *Pattern Recognit.* **2017**, *66*, 44–52. [CrossRef]

36. Kim, D.; Baddar, W.; Ro, Y. Micro-expression recognition with expression-state constrained spatio-temporal feature representation. In Proceedings of the ACM on Multimedia Conference, Amsterdam, The Netherlands, 15–19 October 2016; pp. 382–386.