

Blur Classification Using Segmentation Based Fractal Texture Analysis

Shamik Tiwari

University of Petroleum and Energy Studies, India

Article Info

Article history:

Received Mei 8, 2018

Revised Sep 7, 2018

Accepted Dec 2, 2018

Keyword:

Blur Classification

Hand Gesture Recognition

Neural Network

ABSTRACT

The objective of vision based gesture recognition is to design a system, which can understand the human actions and convey the acquired information with the help of captured images. An image restoration approach is extremely required whenever image gets blur during acquisition process since blurred images can severely degrade the performance of such systems. Image restoration recovers a true image from a degraded version. It is referred as blind restoration if blur information is unidentified. Blur identification is essential before application of any blind restoration algorithm. This paper presents a blur identification approach which categories a degradation available in hand gesture image into one of the sharp, motion, defocus and combined blurred categories. Segmentation based fractal texture analysis extraction algorithm is utilized for featuring the neural network based classification system. The simulation results demonstrate the preciseness of proposed method than other methods.

Copyright © 2018 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

Shamik Tiwari,

University of Petroleum and Energy Studies,

Bidholi, Via Prem Nagar, Dehradun, Uttarakhand 248007, India.

Email: shamik.tiwari@ddn.upes.ac.in

1. INTRODUCTION

Gesture recognition is an area where meaningful physical movement of the fingers, hands, arms, face or body are used to convey information for human computer interaction [1]. These massive gestures may be identified either as the static gestures or as dynamic gestures. Static gestures require less computational complexity rather than dynamic gestures, which are complex but suitable for real time environments. Dynamic gesture recognition requires a decent interpretation of the body or its part movement to have effectively expressive directives.

Hand Gestures compromise a common and natural communication modality for Human Computer Interaction (HCI) [2]. Proficient human computer interfaces (HCIs) need to be established to permit systems to visually recognize hand gestures in real time. However, vision-based hand gesture recognition is challenging problem due to complexness of hand gestures and other image degradation factors during image acquisition.

1.1 Use of Mobile Camera

These days, a wide range of general-purpose hand-held devices, such as mobile phones, come with an optical imaging system [3]. Enabling these devices with the capability to recognize gestures is a cost-effective alternative to conventional static camera with the system. A gesture image captured by camera phone can be transmitted to the system for decoding purpose. With such a mechanism, end-user will not only benefitted from communication provided by the phone, but the camera phone also provides the full mobility. The hardware is readily available to billions of people. Combined with new services, this can revolutionize HCI environment.

1.2 Challenges with Mobile Camera

The availability of imaging phones provides people a mobile platform for decoding of gesture rather than the use of the conventional camera, which has lack of mobility. Unfortunately, sometimes the deprived quality of the images taken by digital cameras makes it difficult to correctly decode gestures. Identifying gesture from images taken by general-purpose handheld devices is particularly challenging due to limitations of the integrated imaging system and the processing capabilities of the device. These devices often have lower quality lens systems and lower resolution imaging circuitry compared to dedicated digital cameras [4], [5].

Image blurring is often a major factor influencing the performance of a gesture recognition system. It is very difficult to recognize right gesture from a blurred image as shown in Figure 1 Image blurring is usually inevitable in a camera-based imaging system, especially in the case of the camera that does not have auto focus or macro mode [6], [7]. Although some high-end camera phones, which integrate a high resolution and auto focus/macro mode are available, the low-end camera phone user segment is huge.

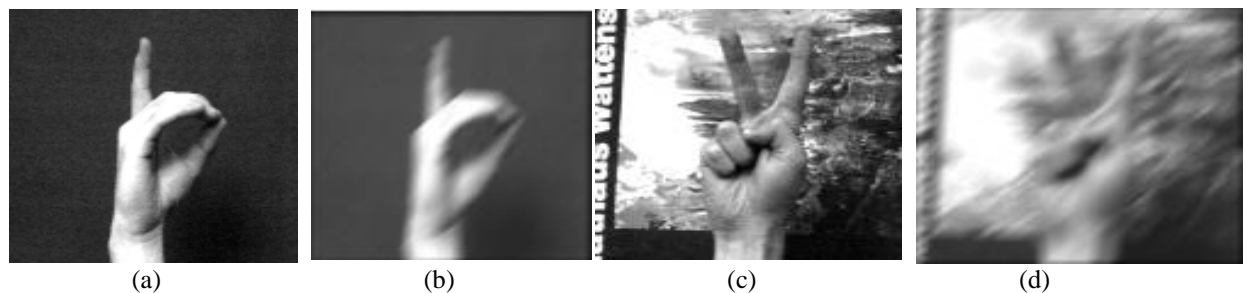


Figure 1. (a) A sharp hand gesture image with simple dark background, (b) Blurred version of image (a), (c) A sharp hand gesture image with complex background and (d) Blurred version of image (c)

Generally, motion and defocus blur degrades the performance of mobile camera based HCI systems. Image restoration is highly desirable since blurring diminishes the high frequency components of image, which are highly desirable for gesture recognition. Image restoration methods can be categorized into blind and non-blind restoration. For blind restoration of images, it is necessary to guess blur parameters for deblurring. However, before use of any blur parameters estimation method blur identification is very required (i.e. to know if image is blurred than what type of blur is present).

1.3 Blur Detection and Classification

While most previous work focuses on blur detection i.e. categorization of image as sharp or blurred image and blur classification separately, not as much research has been done on combined framework for blur identification and blur classification, which is more applied.

Tong et al. [8] proposed a scheme to decide whether an image is blurred or not and if blurred then up to what extent it is blurred. They used Harr wavelet transform to determine different types of edges and presence of blur is decided by this assessment. Yang et al. [9] addressed the motion blur detection method using support vector machine and natural image statistics. In photographic camera, the optical lenses may be set in a way to clearly distinct two areas in the image: the blurry one and the non blurry one. Such a method in which an automatic segmentation coupled to specific descriptors first allow to describe any region of the image and then uses a supervised learning process that decides for each unknown region as "blurry" or "sharp" as suggested by Runga et al. [10].

Crete et al. [11] used intentional blurring pixel difference (IBD) algorithm. This method does not require the use of edge detection. It is motivated with the observation that the intentional blurring of a sharp image provides significant gray scale variations. Instead, intentional blurring of an existing blurred image gives small gray scale variations. Chi [12] proposed an unsupervised method to isolate focused main subject regions from defocused background. This procedure first calculates the blurring level using the bivariate kurtosis of all Discrete Cosine Transform (DCT) blocks of a photographic image with low field depth. Then these blocks are clustered to blurry regions and sharp regions.

Chong and Tanaka [13] proposed a scheme that simultaneously detects and identifies blur. This method is based on the analysis of extrema values in an image. Liu et al. [14] offered a framework for partial blur detection and classification i.e. whether some portion of the image is blurred as well as blur types. They considered maximum saturation of color, gradient histogram span and spectrum details as blur features. Aizenberg et al. [15] presented a work, which identifies blur type, estimates blur parameters and perform image

restoration using neural network. They considered four kinds of blur namely rectangular, motion, defocus, and Gaussian blurs as a pattern classification problem.

Bolan et al. [16] proposed an image blurred region detection and classification method, which can automatically detect blurred image regions and blur type using singular value feature. This scheme has also analyzed the alpha channel information and classifies the blur type into defocus blur and motion blur categories. Yan and Shao [17] made an attempt to find a general feature extractor for common blur kernels with various parameters, which is closer to realistic application scenarios and applied deep belief neural networks for discriminative learning. In this method Fourier spectrum of blurred images is passed to the neural network as input. However, utilizing spectrum as an input is not a good idea due to large size of feature set. Blur patterns in frequency spectrum are utilized by Tiwari et al. [18], [19] for blur classification into one of the three blur categories namely motion, defocus and combined blur for barcode images. They have used statistical and wavelet features for classification.

2. IMAGE DEGRADATION MODELS

Blur is an artefact that occurs in images due to irregularities in image acquisition. The blurring process can be described as some kind of weighted averaging of pixel values in a certain neighborhood. Blurs are treated as low pass filters, which smoothens out the abrupt changes in the gray level of an image. There are different analytical models used to represent the shift invariant degradation model. The description of blurs models considered in this work is given below.

2.1 Motion Blur

When the scene to be recorded translates relative to the camera at a constant velocity (v_{relative}) under an angle of θ degree with the horizontal axis during the exposure interval $[0, t_{\text{exposure}}]$, the distortion is one dimensional. Defining the length of motion as $L = v_{\text{relative}} \times t_{\text{exposure}}$, the point spread function (PSF) for uniform linear motion blur is described as [13]

$$h_m(x, y) = \begin{cases} \frac{1}{L} & \text{if } 0 \leq |x| \leq L \cos\theta, y = L \sin\theta \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

The frequency response of PSF is called Optical Transfer Function (OTF). The frequency response of h_m , is a SINC function given by

$$H(u, v) = \text{sinc}(\pi L(u \cos \theta + v \sin \theta)) \quad (2)$$

When the blur angle is zero degree, it is called as horizontal motion blur.

2.2 Defocus Blur

Images that are captured due to a deprived convergence of light from an object on the sensor plane results in what is known as out-of-focus or defocus blur. The consequence of a defocus blur on an image is the scattering of the pixel intensities around its neighbors in a circle. The centre of such a circle is called the Centre of Confusion (COC). The out of focus blur caused by a system with circular aperture is modeled as a uniform disk of radius R as [14]

$$h_d(x, y) = \begin{cases} \frac{1}{\pi R^2} & \text{if } \sqrt{x^2 + y^2} \leq R \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

The frequency response of Equation (2.6) is given by

$$H(u, v) = \left[\frac{J_1(R\sqrt{u^2+v^2})}{R\sqrt{u^2+v^2}} \right] \quad (4)$$

where J_1 is the Bessel function of first kind and R is radius of uniform disk.

2.3 Combined Blur

In general, image degradation is very complex in nature, but in many situations, it can be modeled as a linear shift-invariant process. The combined blur (also referred as joint or simultaneous blur) considered in this thesis is assumed to be motion blur and defocus blur mixed together, and the two blur effects are supposed to be independent linear shift-invariant processes. Therefore, the combined blur is treated as if it has two steps:

first the original image is degraded by motion blur, and then it is further degraded by defocus blur [14]. In the case where both out-of-focus blur and motion blur are present in the same image, the blur model is

$$g(x, y) = f(x, y) * h_d(x, y) * h_m(x, y) + \eta(x, y) \quad (5)$$

Since convolution is associative, so PSF of combined blur can be obtained as convolution of two blur functions as

$$h(x, y) = h_d(x, y) * h_m(x, y) \quad (6)$$

where $h_d(x, y)$, $h_m(x, y)$ are point spread functions for motion and defocus blur respectively and $*$ is the convolution operator. The PSF estimation for blur is corresponding to estimation of the three parameters angle (θ), length (L) and radius (R).

2.4 Noise Model

The image degradation model in (1) includes an additive noise term that represents the effects of various types of noise introduced at the time of image capturing. This additive term sufficiently models the noise in Charged Coupled Devices (CCD) cameras. There are several sources of noise inherent when using CCD cameras. Only one of these noise sources known as Gaussian noise is modelled for the research in this thesis which can be occurred at the time of image capturing or transmission. This noise is modeled randomly by the Gaussian distribution as [14].

$$\eta(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \quad (7)$$

where σ is the standard deviation of the Gaussian distribution.

The variance of noise added to the blurred images is defined using the Blurred Signal-to-Noise Ratio (BSNR) metric in this work. This metric for a blurred image with variance σ_b^2 is given by

$$BSNR = 10 \log_{10} \left(\frac{\sigma_b^2}{\sigma_n^2} \right) \text{ dB} \quad (8)$$

where σ_n^2 is the variance of noise.

3. SEGMENTATION BASED FRACTAL TEXTURE ANALYSIS (SFTA)

Fractal dimension is extensively applied texture measure. Fractal mentions to the self-similarity at multiple scales of some object pattern. For example, a bounded set S is supposed to be self-similar when S is the union of m distinct copies of itself, where each copy is scaled down by a proportion of p . The fractal dimension (f_d) defined by

$$f_d = \frac{\log m}{\log p}$$

From the fractal dimension, one can estimate the irregularity or roughness of the surface of an object. The high value of fractal dimension denotes the more coarse texture.

The Segmentation Based Fractal Texture Analysis (SFTA) algorithm is a proficient texture feature extraction method; proposed by Costa et al. [20]. This algorithm has two main steps. In the first step, a gray level image is decomposed into a set of binary images using the Two-Threshold Binary Decomposition (TTBD) algorithm. This algorithm decomposes a gray scale image into $2 * n$ binary images, where n is the desired number of thresholds. Figure 2 shows the binary decomposed images for input image with $n = 4$. in the second step, the area, mean gray level and fractal dimension are calculated using these decomposed binary images. SFTA successfully used by researchers for multiple applications such as Content-Based Image Retrieval (CBIR) and image classification in the past. The performance of SFTA based features extraction method is showed better than other widely employed feature extraction methods such as Haralick and Gabor filter banks. SFTA achieved higher accuracy for CBIR and image classification. Moreover, SFTA is quicker than Gabor feature extractor and Haralick feature extraction method.

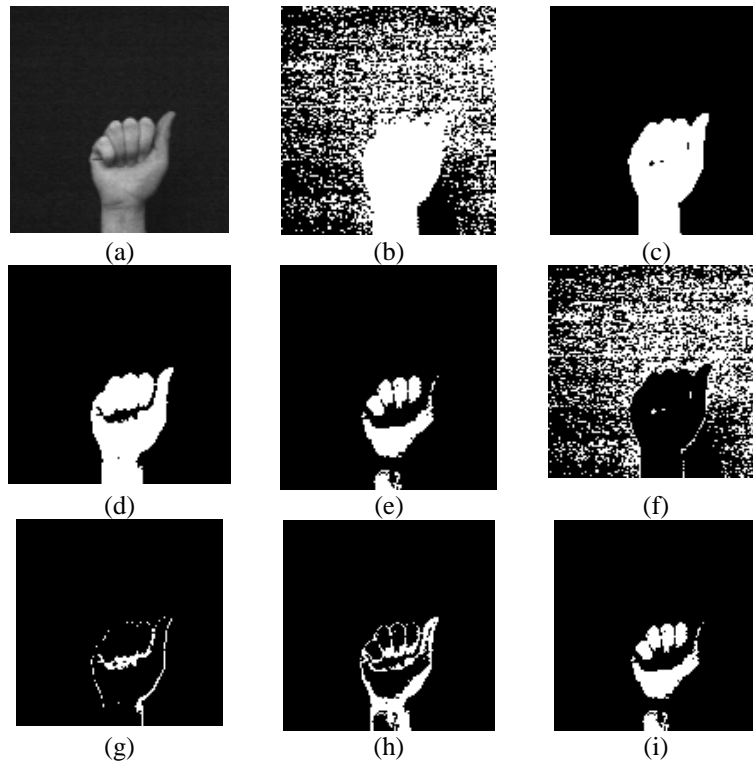


Figure 2. (a) Input image, (b-i) binary decomposed images from input image using TTBD algorithm

Algorithm: Segmentation Based Fractal Texture Analysis (SFTA)

Input: Gray level image (I) and number of thresholds (n).

Output: Feature Vector (F) consisting of three features for each thresholded image namely fractal Dimension, mean gray level and region area.

Step1: $Th := MOTsu(I, n)$

/ Apply multi-level Otsu algorithm to compute the initial set of thresholds*/*

Step2: $Th^C := \{(th_i, th_{i+1}) : th_i, th_{i+1} \in Th \text{ and } i \in [1 \dots |Th| - 1]\}$

/ Add all pairs of contiguous thresholds in Th to Th^C */*

Step3: $Th^P := \{(th_i, l_{max}) : th_i \in Th \text{ and } i \in [1 \dots |Th|]\}$

/ Add pairs of threshold in Th and maximum gray level (l_{max}) in input image to Th^P */*

Step4: $j := 1$.

for $i := 1 : n$

$I_C := TTS(I, t)$, where $t \in Th^C$

/ Apply Two-Threshold Binary Decomposition using each pair of thresholds from Th^C */*

$I_P := TTS(I, t)$, where $t \in Th^P$

/ Apply Two-Threshold Binary Decomposition using each pair of thresholds from Th^P */*

$\partial^C(x, y) := FB(I_C)$.

$\partial^P(x, y) := FB(I_P)$.

/ Find Border (FB) of each thresholded image*/*

$F[j] := BC(\partial^C)$.

$F[j + 1] := MGL(\partial^C)$.

$F[j + 2] := PC(\partial^C)$.

$F[j + 3] := BC(\partial^P)$.

$F[j + 4] := MGL(\partial^P)$.

$F[j + 5] := PC(\partial^P)$.

$j := j + 6$.

/ calculate features fractal dimension using Box Counting (BC), Mean Gray Level (MGL) and regional area using Pixel Count (PC) from border of each thresholded Image*/*

end for

Step5: return F .

4. ARTIFICIAL NEURAL NETWORK

The feed forward artificial neural network with back propagation learning algorithm is preferred as pattern classification tool over other classifiers in this work. Conventional statistical classifiers such as discriminant analysis are designed on the Bayesian decision theory. To make a classification decision, these classifiers requires a prior probability model in order to calculate the posterior probability. One main restraint of the statistical models is that they perform well only when the prior assumptions are satisfied. The efficiency of these models depends largely on the different assumptions or conditions under which the models are developed. Before application of such models, users must aware with both data properties and model capabilities. To overcome these limitations, ANN has utilized as widely acceptable tool for classification. It is a preferred tool due following properties [21]:

- ANN is data determined self-adaptive method in that it can adjust itself to the data without any explicit condition of functional or distributional form for the underlying model.
- ANN is universal functional approximator in that it can approximate any function.
- It is based nonlinear models, which makes it flexible in modeling of real world complex relationships.

5. PROPOSED METHOD

Blurring reduces significant features of image such as boundaries, shape, regions, objects etc., which creates problem for image analysis. It is difficult to identify blur type in spatial domain since all the blurs have same effect in spatial domain. However, motion and defocus blurs have different appearance in frequency space. So, the blur categorization can be easily done using these specific appearance in frequency domain. If the blurred image is transformed in frequency domain, it can be seen from frequency response of motion blurred image that the dominant parallel lines appear which are orthogonal to the motion blur orientation with near zero values [22]. In defocused blur one can see appearance of some circular zero crossing patterns [23] and in case of coexistence of both blurs combined effect of both blurs becomes visible. The power spectrum of sharp image has not such specific patterns. Figure 3 shows the patterns of different blurs on the power spectrum of image. These patterns in frequency domain are considered as textured image itself for blur classification

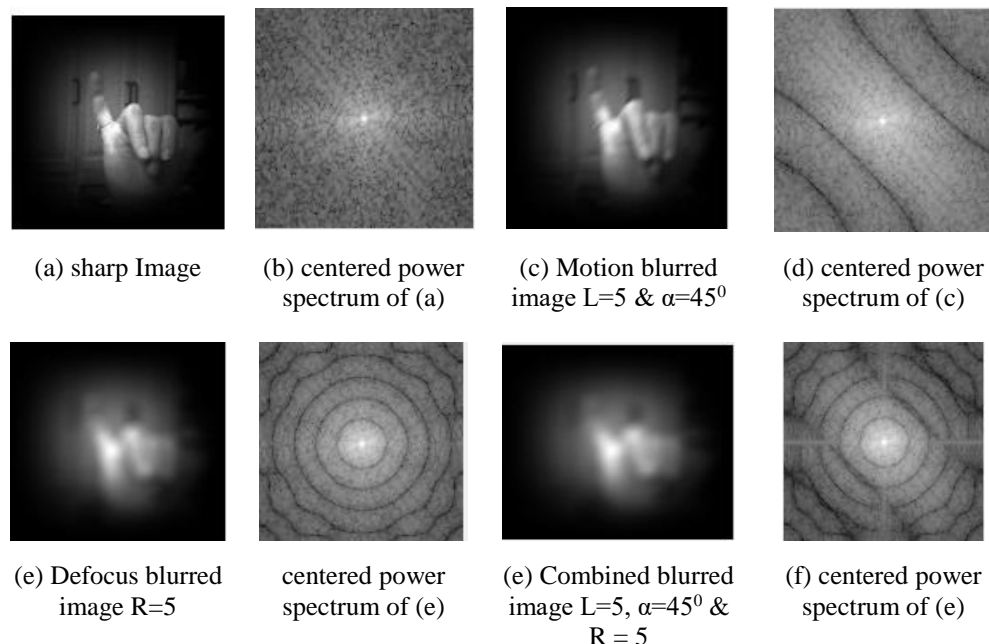


Figure 3. Power spectrum of sharp and blurred hand gesture images

The steps of the algorithm to classify blur are detailed in Figure 4 These are five major steps pre-processing of images, find logarithmic power spectrum, two-threshold binary decomposition, feature extraction, training of neural network classifier system and result analysis.

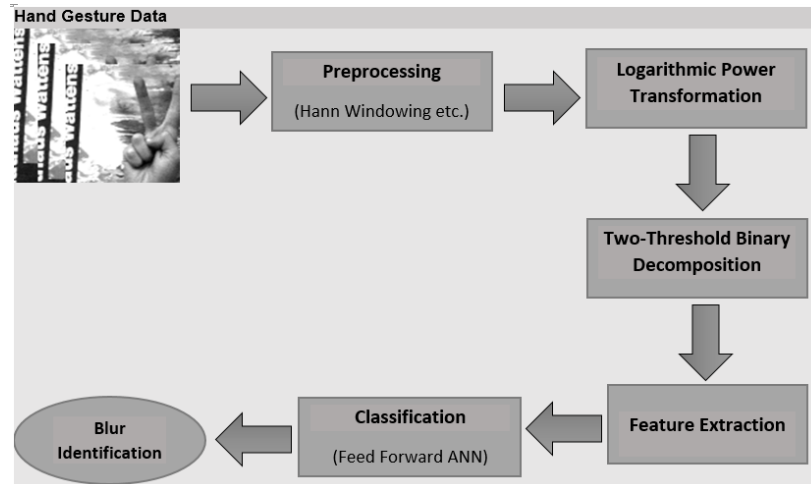


Figure 4. Blur Identification Framework

When achieving a Fourier transform, periodic boundary conditions are forced and sharp discontinuities among the edges of an image cause a cross patterned artifact along the reciprocal space axes. These are visible as vertical and horizontal lines in the power spectrum of image. Because these lines may distract from or even superpose the stripes caused by the blur, they have to be removed by applying a windowing function before frequency transformation. The Hanning window gives a fine trade-off between forming a smooth transition towards the image borders and maintaining enough image information in power spectrum. A 2-D Hann window of size $N \times M$ defined by the product of two 1-D Hann windows as [24]

$$w[n, m] = \frac{1}{4} \left(1 - \cos\left[2\pi \frac{n}{N-1}\right] \right) \cdot \left(1 - \cos\left[2\pi \frac{m}{M-1}\right] \right) \quad (9)$$

for $0 \leq n < N$ and $0 \leq m < M$.

After this step, the windowed image can be transferred into the frequency domain by performing a Fast Fourier Transform (FFT). The power spectrum is calculated to facilitate the identification of particular features of the Fourier spectrum. However, as the coefficients of the Fourier spectrum decrease rapidly from its centre to the borders, it can be hard to identify local differences. Taking the logarithm of the power spectrum helps to balance this fast drop off. In order to obtain a centered version of the spectrum, its quadrants have to be swapped diagonally. Figure 5 shows the effects of Hann windowing.

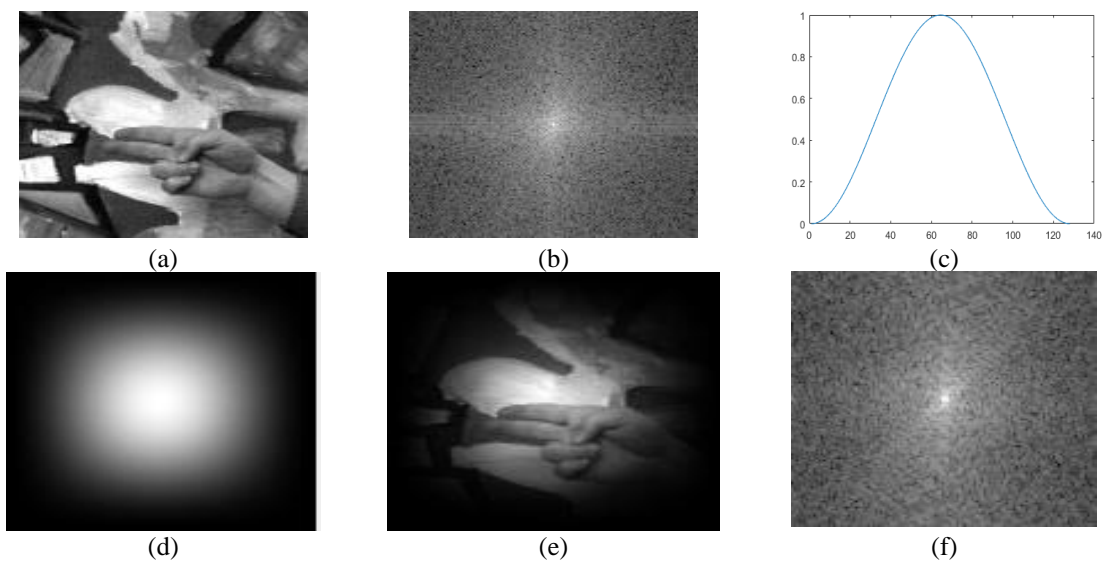


Figure 5. (a) a hand gesture image, (b) Centered power spectrum of (a), (c) Hann Window, (d) Hann windowed image, (e) Centered power spectrum of (d)

6. Simulation

The proposed method is evaluated using Triesch gesture database [25]. This database consists of 720 images taken by 24 different persons in three different backgrounds comprise of 10 hand gestures. All the images are synthetically blurred to create the database. This complete work is implemented using neural network and image processing toolboxes of Matlab 6.5.

SFTA algorithm needs to set the number of thresholds parameter n_t that is used in the input image decomposition. An analysis of classification performance is carried out for different number of thresholds. Considering best accuracy results, the n_t is set to 4 and 5 respectively for blurred and blurred & noisy hand gesture database. So, total $6 * 4 = 24$ and $6 * 5 = 30$ features are extracted for each experiment.

The classification is achieved using a feed forward neural network containing back propagation as network training algorithm, where the training dataset is designed by the extracted features of the blur patterns in frequency spectrum. The whole training and testing features set is normalized into the range of [0, 1], whereas the output class is assigned to zero for the lowest probability and one for the highest probability. In this work, the transfer functions of hidden layers are sigmoid functions. The numbers of neurons used in input layer are equal to the extracted features from image dataset which are 24 and 30 respectively for each experiment. To select the optimum number of hidden layers, the neural network is trained and tested with different number of hidden layers.

The final architecture is selected with single hidden layer of ten neurons, which gives best performance. The out layer consists of 4 neurons corresponding to four class's viz. no-blur, motion blur, defocus blur and combined blur.

The term epoch in back-propagation learning algorithm is used to specify one weight update or training iteration. For each epoch, the learning algorithm makes a dissimilar model with a different set of weights. If an ANN is trained with 100 epochs, the learning algorithm investigates through 100 dissimilar models. So, ANN learning is kind a searching of best model among a large number of models. However, running too many epochs may over train the network and result in overfitting. Early stopping has been used to avoid overfitting of the neural network. In early stopping method, it is possible to test the neural network after successive epoch of training on a validation data set to ensure that over-fitting is avoided.

In order to apply early stopping, the available data is divided into three subsets namely training set, validation set and the testing set in a ratio of 50:20:30 respectively. The validation error is observed after each epoch. When the network starts to over fit, the error on the validation set usually starts to increase. When the validation error grows for a specified number of epochs, the training is terminated, and the trained model at the epoch with minimum validation error is used.

Although the ideal initial values for weights (i.e., those that will maximize the efficiency and momentum with which a neural network trains) cannot be determined theoretically [26], it is common practice to initialize these values randomly. Even if, such a random initialization may lead to a different learning time or result but such a random distribution can minimize the possibilities of the network becoming stuck in local minima [27]. In this work weight are repeatedly randomly initialized and model is preferred with best generalization ability.

To assess the blur classification model, confusion matrix and four statistical metrics as explained in Section 6.1 are utilized. Finally, overall classification results are used to evaluate the blur classification model. A receiver operating characteristic (ROC) curve is also plotted to validate the results. It plots the true positive rate (fraction of true positives over all positives) versus the false positive rate (fraction of false positives over all negatives). A large area under this curve confirms that a classifier has high true positive rate and low false positive rate, which is desired.

6.1 Performance Evaluation

To assess the blur classification model, confusion matrix is formed and five statistical metrics precision, recall, F-score, error rate and accuracy are used as specified below [19].

6.1.1 Precision (Positive Predictive Value): Precision measures the fraction of samples that truly turns out to be positive in the group the classifier has categorized as a positive class. The higher the precision is, the lower the number of false positive errors committed by the classifier.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \times 100 \quad (10)$$

6.1.2 Sensitivity (Recall): It measures the actual members of the class, which are accurately recognized as such. It is also referred as true positive rate (TPR).

$$\text{Recall} = \frac{TP}{TP+FN} \quad (11)$$

6.1.3 F-Score: F-Score, or F1, measurement is a standard metric, which combines precision and recall into one value. It is the weighted average of precision and recall. A high F-score means that almost all images are classified accurately, along with minimal misclassified images.

$$F - \text{Score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (12)$$

6.1.4 Accuracy & Error Rate: Accuracy is the percentage of overall correctness of the model and is calculated as the sum of correct classifications divided by the total number of classifications.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FN+FP} \quad (13)$$

The error rate is given by 1 - Accuracy.

Where:

- True Positive (TP) and True Negative (TN) if the system correctly predicts the label.
- False Positive (FP) if the system labels a negative point as positive.
- False Negative (FN) if the system labels a positive point as negative.

6.2 Experiment 1: Performance Analysis with Blurred Hand Gesture Images

In the first experiment, all the 720 images of the hand gesture database are synthetically blurred with the three different classes of blur (motion, defocus and combined blur) using varying blur parameters to make the database of 2160 blurred images (i.e., 720 images with each class of blur) for each type of barcode images. So, total 2880 images (i.e. 2180 blurred and 720 original sharp) are used to carry out the experiment. Out of 2880 images, 1440, 576 and 864 images are used as training, validation and testing datasets respectively to design the neural network model. The training process is continued till the best validation performance achieved. The training stopped when the validation error increased for 6 iterations. The best validation performance is 0.0065, at epoch 75 as shown in Figure 6.

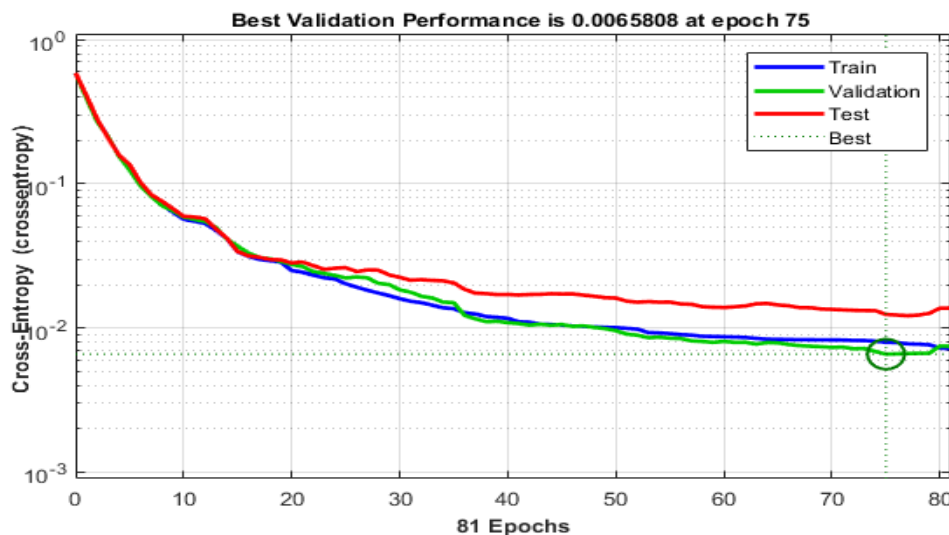


Figure 6. Neural network performance for training with blurred hand gesture images

The trained network model is used subsequently to categorize the blur type from degraded images. The overall blur classification results are presented as confusion matrix in Table 1. Using this confusion matrix statistical quantities precision, recall, F-score, error rate and accuracy are calculated. The high values of these metrics in Table 2 show effectiveness of this classification model. The overall classification accuracy of the model is 99.2 %. The ROC curves plotted in Figure 7 also confirms the high performance.

Table 1. Confusion matrix for blur identification results of blurred hand gesture images

	No-Blur	Motion Blur	Defocus Blur	Simultaneous Blur
No-Blur	718	0	0	0
Motion Blur	0	706	4	8
Defocus Blur	0	3	715	0
Simultaneous Blur	0	9	0	709

Table 2. Blur Identification results for blurred hand gesture images

	Precision	Recall	F-Score	Error Rate	Accuracy
No-Blur	1	1	1	0	1
Motion Blur	0.983	0.983	0.983	0.017	0.983
Defocus Blur	0.994	0.995	0.995	0.006	0.994
Simultaneous Blur	0.989	0.987	0.998	0.011	0.989

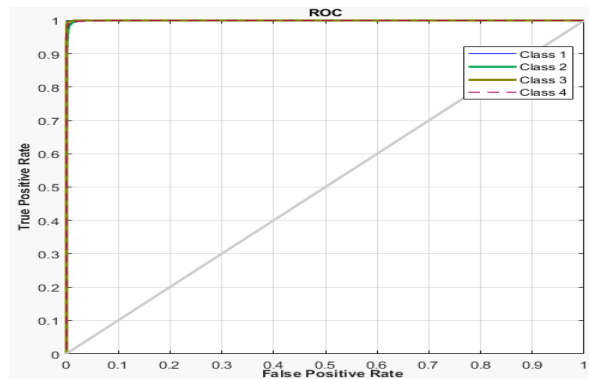


Figure 7. ROC curves for each class of blur identification where classes 1, 2, 3 and 4 representing no-blur, motion, defocus and combined blur respectively

6.3 Experiment 2: Performance Analysis with Blurred and Noisy Hand Gesture Images

To test the efficiency of the proposed scheme in presence of noise, another neural model is created using all the hand gesture images. All the images are degraded by different kinds of blur similar to *Experiment 6.1*. Subsequently, zero mean Gaussian noise of 40 dB Blurred Signal-to-noise ratio (BSNR) is added to the blurred and sharp images to create a database of blurred and noisy samples. The training process is continued till the best validation performance achieved. The training stopped when the validation error increased for 6 iterations. The best validation performance is 0.024, at epoch 33 as shown in Figure 8.

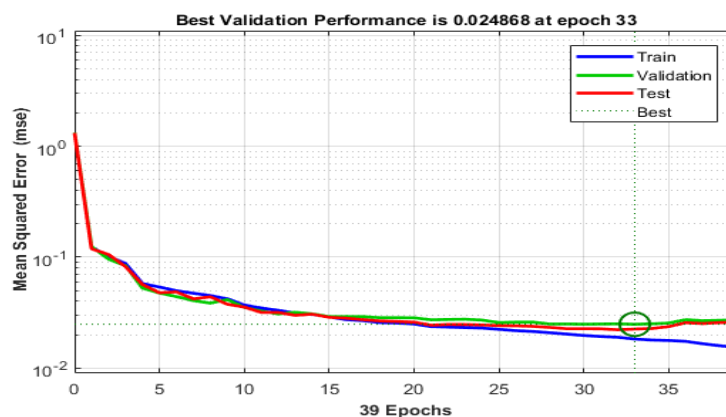


Figure 8. Neural network performance for training with blurred and noisy hand gesture images

The trained network model is used afterwards to classify the blur type from degraded images. Table 3 shows the confusion matrix of the proposed blur classification system. Table 4 shows the performance metrics of the proposed model. The high values of these metrics show efficiency of this model. The overall classification accuracy is 95.6 %. The ROC curves plotted in Figure 9 also approves the high performance.

Table 3. Confusion matrix for blur identification results of blurred and noisy hand gesture images

	No-Blur	Motion Blur	Defocus Blur	Simultaneous Blur
No-Blur	718	0	0	0
Motion Blur	0	657	61	0
Defocus Blur	0	65	653	0
Simultaneous Blur	0	0	0	718

Table 4. Blur Identification results for blurred and noisy hand gesture images

	Precision	Recall	F-Score	Error Rate	Accuracy
No-Blur	1	1	1	0	1
Motion Blur	0.91	0.915	0.913	0.09	0.91
Defocus Blur	0.915	0.91	0.912	0.085	0.915
Simultaneous Blur	1	1	1	0	1

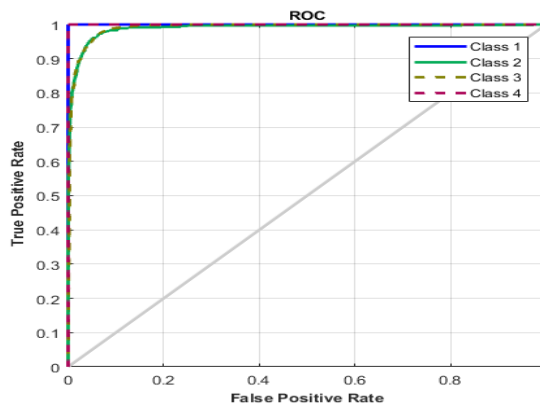


Figure 9. ROC curves for each class of blur identification in presence of noise where classes 1, 2, 3 and 4 representing no-blur, motion, defocus and combined blur respectively

6.4 Comparative Analysis

The results of the proposed methods are compared with two earlier offered methods based on statistical features [18] and Curvelet transform based features [19]. Table 5 and Table 6 present the comparative results. From the results, it evident that proposed method is superior than other methods.

Table 5. Comparative results for blurred hand gesture images

	Proposed Method using SFTA features	Statistical Features	Curvelet Transform Based Features
No-Blur	1	0.98	0.98
Motion Blur	0.983	0.96	0.968
Defocus Blur	0.994	0.92	0.94
Simultaneous Blur	0.989	0.97	0.98
Overall Accuracy	0.992	0.957	0.967

Table 6. Comparative results for blurred and noisy hand gesture images

	Proposed Method using SFTA features	Statistical Features	Curvelet Transform Based Features
No-Blur	1	0.93	0.92
Motion Blur	0.91	0.90	0.95
Defocus Blur	0.915	0.89	0.90
Simultaneous Blur	1	0.91	0.91
Overall Accuracy	0.956	0.907	0.920

7. Conclusion

A blur identification approach is presented in this paper, which categorize a degraded hand gesture image into one of the four categories viz. sharp, motion, defocus and coexistence of both blurs. Such a framework is very required before blind restoration of degraded hand gesture images. This blur identification framework gives impressive results with the accuracies of 99.2 % and 95.6 % with nonexistence and existence of noise respectively. The results are also compared with other methods which shows the superiority of the proposed method. The limitation of these methods is with the level of noise. The blurred images with a high degree of noise will not be suitable since disappearance of blur patterns in frequency spectrum.

REFERENCES

- [1] Escalera S., Athitsos V., and Guyon I., "Challenges in Multi-modal Gesture Recognition. In *Gesture Recognition*," Springer, Cham, pp. 1-60, 2017.
- [2] Singha J., and Laskar R. H., "Hand gesture recognition using two-level speed normalization, feature selection and classifier fusion," *Multimedia Systems*, vol. 23(4), pp. 499-514, 2017.
- [3] Nicholas C. G., Marti L. M., van der Merwe R., and Kassebaum J., "U.S. Patent No. 9,679,414," *U.S. Patent and Trademark Office*, Washington, DC, 2017.
- [4] Eisaku O., Hiroshi H., Lim A. H., "Barcode readers using the camera device in mobile phones," in *Proc. of Internet. Conf. on Cyber worlds*, pp. 260–265, 2004.
- [5] Thielemann J. T., Schumann-Olsen H., Schulerud H., and Kirkhus T., "Handheld PC with camera used for reading information dense barcodes," in *Proc. IEEE Int. Conf. on Computer Vision and Pattern Recognition, Demonstration Program*, Washington, DC, pp. 102-112, 2004.
- [6] Joseph E., and Pavlidis T., "Bar code waveform recognition using peak locations. Pattern Analysis and Machine Intelligence," *IEEE Transactions on*, vol. 16(6), pp. 630-640, 1994.
- [7] Selim E., "Blind deconvolution of barcode signals," *Inverse Prob*, vol. 20(1), pp. 121– 135, 2004.
- [8] Tong H., Li M., Zhang H., and Zhang C., "Blur detection for digital images using wavelet transform," in *proceedings of IEEE international conference on Multimedia and Expo*, vol. 1, pp. 17-20, 2004.
- [9] Yang Q., Yi. X., and Yang X., "No-reference image blur assessment based on gradient profile sharpness," *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, pp. 1-4, 2013.
- [10] Rugna J. D., and Konik H., "Blur identification in image processing," *International Joint Conference on Neural Networks (IJCNN '06.)*, pp. 2536-2541, 2006.
- [11] Crete F., Dolmiere T., Ladret P., and Nicolas M., "The blur effect: Perception and estimation with a new no-reference perceptual blur metric," in *SPIE Human Vision & Electronic Imaging*, vol. 6492, pp. 1-11, 2007.
- [12] Chi Z., "An unsupervised approach to determination of main subject regions in images with low depth of field," *IEEE 10th Workshop on Multimedia Signal Processing*, pp. 650-653, 2008.
- [13] Chong R.M., and Tanaka T., "Image extrema analysis and blur detection with identification," *IEEE International Conference on Signal Image Technology and Internet Based Systems (SITIS '08)*, pp. 320-326, 2008.
- [14] Liu R., Li. Z., and Jia J., "Image partial blur detection and classification," in *Proc. CVPR*, pp. 23–28, 2008.
- [15] Aizenberg I., Paliy D. V., Zurada J. M., and Astola J. T., "Blur identification by multilayer neural network based on multivalued neurons," *IEEE Transactions on Neural Networks*, vol. 19(5), pp.883-898, 2008.
- [16] Bolan S., Lu S., and Tan C., "Blurred image region detection and classification," in *Proc. ACM Multimedia*, pp.1397–1400, 2011.
- [17] Yan R., and Shao L., "Image Blur classification and parameter identification using two-stage deep belief networks," *British Machine Vision Conference (BMVC)*, Bristol. UK, pp. 1-11, 2013.
- [18] Tiwari S., Shukla V. P., Biradar S., and Singh A., "Texture features based blur classification in barcode images," *International Journal of Information Engineering and Electronic Business*, vol 5(5), pp 34, 2013.
- [19] Tiwari S., "A Pattern Classification Based approach for Blur Classification," *Indonesian Journal of Electrical Engineering and Informatics (IJEI)*, vol 5(2), 2017.
- [20] Costa A. F., Humpire-Mamani G., and Traina A. J. M., "An efficient algorithm for fractal analysis of textures. in Graphics, Patterns and Images (SIBGRAPI), 2012 25th SIBGRAPI Conference on," *IEEE*, pp. 39-46, August 2012.
- [21] Tripathy R. K., Acharya A., Choudhary S. K., and Sahoo S. K., "Artificial neural network based body posture classification from EMG signal analysis," *Indonesian Journal of Electrical Engineering and Informatics (IJEI)*, vol 1(2), pp 59-63, 2013.
- [22] Dobes M., Machala L., and Frst T., "Blurred image restoration: A fast method of finding the motion length and angle," *Digital Signal Processing*, vol. 20(6), pp. 1677–1686, 2010.
- [23] Sakano M., Suetake N., and Uchino E., "A robust point spread function estimation for out-of-focus blurred and noisy images based on a distribution of gradient vectors on the polar plane," *Journal of Optical Society of Japan*, vol. 14(5), pp. 297-303, 2007.
- [24] Tiwari S., Shukla V. P., Biradar S. R., and Singh A. K., "Blur parameters identification for simultaneous defocus and motion blur," *CSI transactions on ICT*, vol 2(1), pp 11-22, 2014.
- [25] Triesch J., and Von Der Malsburg C., "Robust classification of hand postures against complex backgrounds. In Automatic Face and Gesture Recognition," *Proceedings of the Second International Conference on IEEE*, pp. 170-175, 1996.
- [26] Thimm G., and Fiesler E., "High-order and multilayer perceptron initialization," *IEEE Transactions on Neural Networks*, vol 8(2), pp. 349-359, 1997.
- [27] Azriyenni A., Mustafa M. W., Sukma D. Y., and Dame M. E., "Backpropagation Neural Network Modeling for Fault Location in Transmission Line 150 kV," *Indonesian Journal of Electrical Engineering and Informatics (IJEI)*, vol 2(1), pp 1-12, 2014.