Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Engineering 41 (2012) 827 – 832

Engineering Procedia

www.elsevier.com/locate/procedia

International Symposium on Robotics and Intelligent Sensors 2012 (IRIS 2012)

Static Hand Gesture Recognition Using Local Gabor Filter

Shikha Gupta[a*], Jafreezal Jaafar[b], Wan Fatimah Wan Ahmad[a,b,]

Department of computer and Information Science, Universiti Teknologi Petronas
Tronoh, Perak, Malaysia

Abstract

With the increasing demand of Human-Computer interaction it is expected that in future human computer interfaces will provide more convenient, natural and comfortable communication same like human-human interaction. Usage of static gestures in our daily life to convey certain meaning leads to recognition of static hand gestures as an important aspect in HCI. Recognition systems involve various processes such as feature extraction, features reduction and classification. Gabor filter is being used for feature extraction since long time in image processing due to its remarkable mathematical and biological properties. Gabor filters have high dimensionality so in the proposed method instead of using 40 Gabor filters only 15 local Gabor filters have been used. The main objective behind using 15 Gabor filter responses is to reduce the complexity with better accuracy. In the proposed method after using local Gabor filters the features are being reduced by PCA to overcome small sample size problem. So that can use LDA which is most popular linear projection technique for feature extraction and feature reduction. This combination of PCA and LDA has been used in many applications such as face recognition for security purpose. But here in this paper this is being used with Gabor filter for hand gesture recognition. Till now in hand gesture recognition the main focus is on Gabor filter with PCA but here in this paper will use PCA and LDA for hand gesture recognition. The use of local Gabor filter helps in reducing the redundant data as instead of using 40 filters here in this paper 15 filters are being used which is a subset of the global filters parameters. Classification of the gestures as per their classes will be done with the help of one against one multiclass SVM.

Keywords: Static hand gesture; Local Gabor-filter bank; PCA; LDA; One against one SVM

1. Introduction

To facilitate efficient human computer interaction many special devices are being used as an interface between human and computer. But still, gestures are powerful means for the communication among the human. Many devices have been developed so that computer vision system would be able to understand gestures. The use of such devices became very familiar but still it bounds the speed and naturalness by which users can communicate with the computers. It became more serious after the evolution of different technologies for example Virtual reality [1]. Many efforts have been carried out for the detection and recognition of faces, palm, emotional expression and hand gestures. Recognition systems plays very important role in many applications such as telemedicine, biometrics and advanced interfaces for Human-Computer Interaction.

* Corresponding author. Tel.: +601 65445759.
E-mail address: shikha.cs88@gmail.com

1877-7058 © 2012 Published by Elsevier Ltd.
doi:10.1016/j.proeng.2012.07.250

Particularly, Gesture is nothing but a form of communicative conversation which can be used to impart information among people. Gestures can differ from easy way to more complex way of using hand for verbalize feelings such as pointing an object to more complex one. But to get the meaning of gestures for being used in Human Computer Interaction it's a big challenge. It requires some means by which gesture recognition process can become easy and efficient for understanding the intended gestures. Gesture recognition process requires features extraction based on which classifier classifies gesture with respect to their respective classes accurately. The proposed method in this paper is a step towards developing a system with less complexity and high accuracy.

2. Related Work

Gesture recognition which is new form of interaction between human and computer was first proposed by Myron W. Krueger in the middle of 70s. With the passage of time it is becoming very important research area. Many methods were introduced for hand gesture recognition. Gabor filter followed by PCA for the American Sign Language [8]. In this system they used global Gabor filter (5X8) for the representation of various ASL alphabets. PCA to reduce the features because of the very large dimensions of Gabor filters. For the classification Fuzzy-C-Mean clustering was applied. The accuracy rate was 93.23% but the combination of PCA and Fuzzy-C-Mean is very complicated and consumes more time. For being a good recognition system along with high accuracy complexity parameter contributes equally.

Haar-like features and AdaBoost learning algorithm [9] was applied for gesture recognition with the accuracy rate of 90%. This method can achieve the real time performance. But this algorithm was used only for four hand gestures with the angle difference of 15 degree. A novel Gabor-SVM method was used for hand gesture recognition with the accuracy rate of 95.2% [10]. This method performed better than other methods such as the Euclidean and Cosine measures. But due to the use 40 Gabor filters the complexity was very high in terms of both the time and space. Same novel method used again for gesture recognition with complex background and this time instead of global Gabor filter with (5x8) here (3X8) Gabor filters were used. The accuracy rate for this system was 96.1%. But the complexity was very high as 24Gabor filters response were used for one gesture (24x160x100) [11].

In the proposed method instead of 40 Gabor filter 15 local Gabor filters have been used to reduce the dimensions which lead to better performance and also reduce the complexity. The Gabor filters are not orthogonal wavelets so selection of optimal parameters is very important to reduce the complexity due to the redundant data. Also LDA which is one of the most important linear projection techniques for projection may encounter small sample size problem so by using Local filters followed by PCA can helps to overcome this problem also can make recognition process scale, rotation and light invariant.

3. Gabor filter

Wavelet transform can extract spatial and frequency information from a given signal. As compare other wavelet transforms, the Gabor wavelet has impressive properties in terms of mathematics and biology. Among various wavelet Gabor also provides optimal resolution in time as well as in frequency domain. Gabor wavelet models are very well known in simulations of the receptive profile of visual cortical cells [12]. The simple cells of the visual cortex of mammalian brains can be best model as self- similar 2D Gabor wavelets. Gabor filters can derive multi- orientation information from a hand gesture images at different scales. The derived information it is local in nature. Basic approach while constructing a Gabor filter for hand gesture or in any other application such as face recognition or emotion detection is to construct a filter bank with different scales and orientations. Hence Gabor filters are used for features detection at various angles and scales [13-17].The main Principle of Gabor filters is that they can capture visual properties, such as spatial locality, orientation selectivity, and spatial frequency characteristics. Due to these characteristics many applications choose Gabor filter for feature representation. In 2D Gabor filter is represented as multiplication of a 2D Gaussian and a complex sinusoidal function sometimes also called as a complex exponential function. The general expression is:The Gabor filter is represented as [18]:

$$g(x,y,\lambda,\theta,\psi,\sigma,\gamma)=\exp\left(-\frac{x'^2+\gamma^2y'^2}{2\sigma^2}\right)\cos\left(2\pi\frac{x'}{\lambda}+\psi\right)$$

$$x' = x\cos\theta + y\sin\theta \qquad y' = -x\sin\theta + y\cos\theta$$

λ shows the wavelength of the filter. The standard deviation σ of the Gaussian factor determines the size of the receptive field. Spatial frequency bandwidth can be calculated as σ/ λ, which determines the numbers of visible parallel stripes zone. In the proposed method this ratio has been fixed to σ/ λ=0.50. Ψ- Phase offset to get the symmetry of the kernel in terms of origin. γ- Aspect ratio this gives the ellipticity of the receptive field. Here, in this proposed method, local Gabor filter bank with 3 different scale and 5 different orientations σ = {1, 2, 3} and θ= $\{0^0, 36^0, 72^0, 108^0, 144^0$ is being used. After the calculation of Gabor function will be displayed on the output window.

Basically Gabor filter bank with 5 different scales and 8 different orientations is being used where input image is convolved with these 40 Gabor filters (5 scales and 8 orientations). But here the image is being convolved with 15 filters to reduce the complexity. As per the above discussion the higher the value of σ makes the image blur so here only three scale values are being selected to reduce the complexity and also to get better results [18, 19]. The responses on the orientation [0, pi] are complex conjugate of the responses on orientation [pi, 2pi] so by selecting orientation [0, pi] the computation is being reduced to half. The figure 4 shows the hand gestures in the taken dataset. The figure 5 shows the kernel response and figure 6 shows the magnitude of the applied Gabor filter.

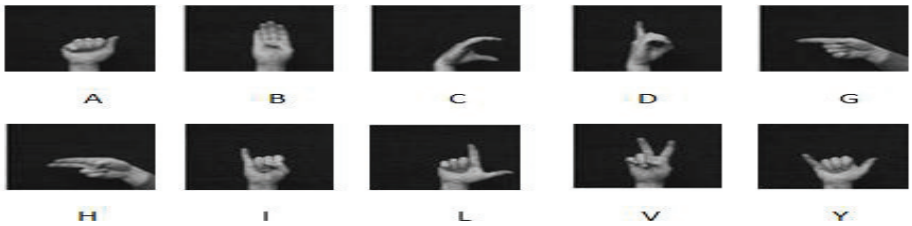

Figure 4: Sample hand posture images from Jochen Triesch dataset

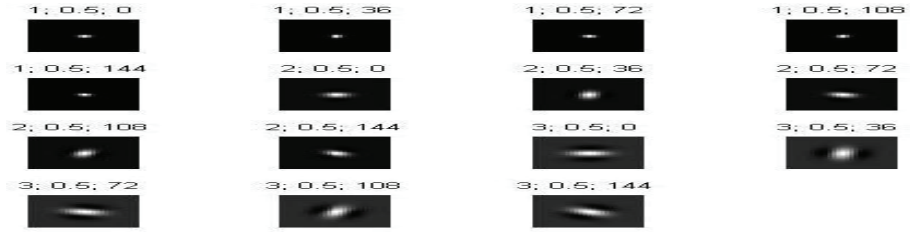

Figure 5: Kernel response with different scales and orientations

Figure 6: Gabor filter response of a typical hand gesture

As in case of gesture recognition while using Gabor filter for the feature extraction its output vector has high dimensionality. If the input image is of 64x64 pixels after convolution with Gabor filter bank of 3 scales and 5 orientation s the dimensionality will be 15x 64x64=3840.

The 2D Gabor filtered images is converted into a pattern vector this process is repeated for all other 15 responses. In the last after getting all the pattern vectors for 15 responses concatenate all of them either in rows or columns.

4. PCA and LDA

When dealing with any application either face recognition, emotion detection or gesture recognition usually have a high dimensionality. In order to handle such data its dimensionality needs to be reduced. Dimensionality reduction is nothing just the transformation of high- dimensional data into efficient, meaningful representation of reduced dimensionality. Basically, the reduce dimensions should represents the minimum number of features needed to account the observed property of the data. Dimension reduction techniques facilitate classification and compression of high-dimension data. PCA and LDA these are very powerful techniques for dimension reduction and feature extraction. Many applications use these techniques for recognition purpose especially in face recognition LDA plays a very important roll but its performance get effected if there are less number of observation as compare to the dimensionality of the given samples [25]. To overcome this here instead of using global Gabor filter, local filter with 15 banks has been used. To reduce the dimensionality of the Gabor filtered image PCA is being used so that LDA performance will not get degraded due to small sample size problem.

As many recognition systems use PCA (Principal components analysis) [8, 20]. PCA is an effective method which can reduce the dimensionality of the data and can effectively extract the required information of the image. It provides data which has no redundancy as Gabor filter wavelet are not orthogonal wavelets. In case of image processing the complexity of grouping the images can be reduced. As among other dimension reduction methods PCA is the fastest algorithm [21] so it will reduce the time complexity. PCA outperform when Dataset is small so in this paper by using local Gabor filter bank it enhance the computation efficiency of the PCA. But in case of PCA its does not emphasis on the class just features are builded on the basis of the difference but in case of LDA when data is map from one dimensional space to another it maximized scatter between two classes but minimized scatter within classes.

5. Classification

Classification it is the last stage in the recognition system. The choice of classification algorithm it is highly depending on the recognition system. There are many classifiers working for hand gesture classification. Neural network is one of the much known classifier. It functioning is based on the human brain. NN has three layers: First layer represents neurons which give hand gesture samples to the neural network. Next, the hidden layer which is responsible for error reduction to achieve the desired output. Finally, the last layer the output layer having neurons equal to the output classes. Back-propagation NN is very common in image recognition [7]. NN are learning based classifier instead of programming, flexible to changing environment. But consume time to train a model also face problem with generalization.

Decision tree-based classifier works by breaking a complex decision into simple decisions to get the final solution which resembled to the desired solution. Internal nodes in decision tree represent different attribute values forming the training data in order to decimate the various classes most clearly. The branches connecting the nodes denote the possible values these attributes can have and the terminal nodes show the final value that is the intended class. Tree-based classification is conceptually very simple and computation is also very effective. Have simple form to store the data. But in case of multiclass classification the output has a very large and complex tree. Also it does not provide robustness to the recognition system and the intended classes can overlapped. SVM it is a supervised learning method. Basically, SVM is a linear classifier that maximizes the distance between the decision lines. The main advantage of SVM is that it can use kernels for non-linear data transformation. The main principle behind using SVM is to divide the given data into two distinct categories and then to get hyper-plane to separate the given classes as shown in figure 7(a) and 7(b). Even if it is OAO or OAA still need to divide the classes into binary form. Among the nearest neighbor, NN SVM is best method for classification [22].

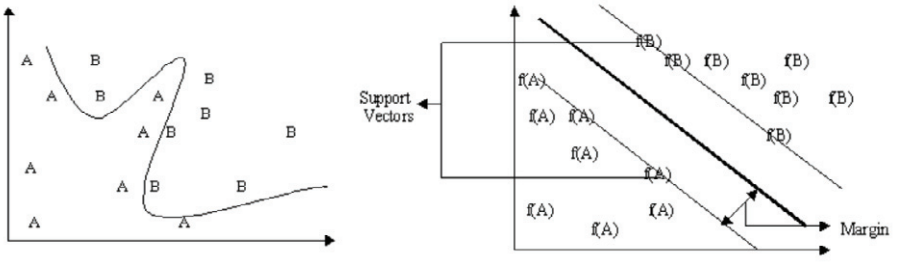

Figure 7(a): Original data; (b): SVM classification

Among both the strategies OAO(one against one) and OAA(one against all), if very few classes are there then OAA gives better accuracy but if the number of classes increases then OAO is better than OAA as it speed up the decision making process [23]. Also when number of samples is very high like in case sign language recognition systems then OAO is a better strategy in terms of training time due to this it is more suitable to practical use [24].

6. Expected results and discussion

In many applications a combination of PCA followed by LDA is being used for feature extraction and feature reduction and for classification either SVM or K-Nearest Neighbor has been used with very good accuracy [25]. In this paper instead of using directly raw image a method is proposed where local Gabor filter with 3 different scales and 5 different orientations is being used. Then PCA and LDA will be used for feature reduction and extraction respectively. By using Gabor filter the proposed work can be invariant to the different and orientation. Also a combination of Gabor filter and PCA can over lighting problem which is very common. Again the combination of PCA and LDA can help for better performance since LDA performance get degraded due to small sample size problem. For classification as per the above discussion Multiclass SVM will be used. SVM has high accuracy rate in non-linear environment. Among OAA and OAO, One against One strategy will be taken in use.

Many Methods are working to recognized hand gestures by using Gabor filter and achieved fairly good results, but have very high computation. The aim of this method is to achieve good results but at the same time a tradeoff must be considered between time and accuracy. So here instead of using Global Gabor filter local Gabor filter is being used with a combination of PCA and LDA to provide better accuracy also making system invariant to scale and rotation problem. Along with high accuracy and less complexity focus is done on robustness too.

7. Conclusion

This paper addresses about the local Gabor filter with a combination of PCA and LDA followed by SVM for classification. From the above discussion it is expected that a combination of Gabor filter along with PCA and LDA will achieve high accuracy with less complexity. On the same side there are many classifiers are there each of them have some advantage and disadvantage but to get right combination with underlying application is very important. From the discussion it is shown that SVM can provide better classification with multiple classes with less complexity and high accuracy.

Acknowledgements
This research was fully supported by a grant from UTP and also thankful to the Support staff of institute for their additional.

References

[1]J. J. Stephan and S. a. Khudayer, Gesture recognition for Human Computer Interaction, International Journal of Advancements in computing Technology, vol. 2, 4 November 2010.

[2]Pavlovic, et al., Visual interpretation of hand gestures for human-computer interaction: A review, IEEE Trans. on Pattern Recognition and Machine Intelligence, vol. 19, p. 19, JULY 1997.

[3]S. K. Kang, et al., Color based hand and finger detection technology for user interaction, presented at the International conference on convergence and hybrid information technology 2008.

[4]S. Padam and K.Prabin.Bora, A Study on Static Hand Gesture Recognition using Moments, presented at the International Conference on Signal Processing and Communications (SPCOM), 2010

[5]K. Abe, et al., Virtual 3D interface system via hand motion recognition from two cameras, IEEE Transactions on Systems, Man and Cybernetics, Part A, vol. 32, p. 5, JULY 2002.

[6]T. Heap and D. Hogg, Wormholes in shape space: tracking through discontinuous changes in shape, presented at the Proc. Sixth International Conference on Computer Vision, ICCV98.

[7]H. Hasan and S. Abdul-Kareem, Static hand gesture recognition using neural networks, 2012.

[8]M. A. amin and H. Yan, Sign Language Finger Alphabet Recognition from Gabor-PCA Representation of hand gestures, presented at the Proceeding of the sixth International Conference on Machine Learning and Cybernetics, Hong Kong, 2007.

[9]Chen, et al., Hand gesture recognition using Haar-like features and a stochastic context-free grammar, IEEE Transactions on Instrumentation and Measurement vol. 57, p. 9, 2008.

[10]D.-Y. Huang, et al., Vision-based Hand Gesture Recognition Using PCA+ Gabor filters and SVM, presented at the Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2009.

[11]D.-Y. Huang, et al., Gabor filter-based hand-pose angle estimation for hand gesture recognition under varying illumination, Expert Systems With Applications, vol. 38, p. 12, 2011.

[12]K. Y. C. Chung, Facial Expression Recognition by Using Class Mean Gabor Responses with Kernel Principal Component Analysis, 2010.

[13]S. Bashyal and G. K. Venayagamoorthy, Recognition of facial expressions using Gabor wavelets and learning vector quantization, Engineering Applications of Artificial Intelligence, vol. 21, p. 10, 2008.

[14]C. Liu., Gabor-Based Kernel PCA with Fractional Power Polynomial Models for Face Recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, p. 10, MAY 2004.

[15]L. Wang, et al., 2D Gabor face representation method for face recognition with ensemble and multichannel model, Image and Vision Computing, vol. 26, p. 9, 2008.

[16]W. Xu, et al., A Scale and Rotation Invariant Interest Points Detector Based on Gabor Filters, signal processing image processing and pattern recognition Communications in Computer and Information Science, , vol. 61, p. 8, 2009.

[17]Z. Huang, et al., Study of Sign Language Recognition Based on Gabor Wavelet Transforms, International Conference on Computer Design and applications (ICCDA 2010).

[18]Y. B. Jemaa and S. Khanfir, Automatic local Gabor features extraction for face recognition, (IJCSIS) International Journal of Computer Science and Information Security, vol. 3, 2009.

[19]A.-A. Bhuiyan and C. H. Liu, On Face Recognition using Gabor Filters, World Academy of Science, Engineering and Technology, vol. 28, 2007.

[20]R.PATIL and P. D. S. M.B., Dimensionality Reduction of Satellite images using Principal Component analysis, International Journal of Communication Engineering Applications-IJCEA, vol. 02, July- Oct 2011.

[21]K. Burgers, et al., A Comparative Analysis of Dimension Reduction Algorithms on Hyperspectral Data, 2009.

[22]A. Savaris and A. vonWangenheim, "Comparative evaluation of static gesture recognition techniques based on nearest neighbour, neural networks and support vector machines, Journal Brazilian computer society, vol. 16, p. 16, 2010.

[23]J. Milgram, et al., Speeding Up the Decision Making of Support Vector Classifiers, presented at the Proceedings of the 9th Int'l Workshop on Frontiers in Handwriting Recognition (IWFHR-9 2004).

[24]J. Milgram, et al., "One against one" or "One against all": which one is better for handwriting recognition with SVMs.

[25]A.Khan and H. Farooq, "Principal Component Analysis-Linear Discriminant Analysis Feature Extractor for Pattern Recognition," IJCSI International Journal of Computer Science, vol. 8, 2011.