International Journal of Engineering, Science and Mathematics
Vol. 6 Issue 8, December 2017,
ISSN: 2320-0294 Impact Factor: 6.765
Journal Homepage: http://www.ijesm.co.in
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

SPATIO-FREQUENCY FEATURE EXTRACTION METHODS FOR CONTENT BASED IMAGE RETRIEVAL

Gagan Sharma[1]
Research Scholar
Department of Computer Science
Asra College of Engineering and Technology,
Bhawanigarh, Punjab

Shevata Sehgal[2]
Assistant Professor:
Department of Computer Engineering,
Asra College of Engineering and Technology,
Bhawanigarh, Punjab

ABSTRACT: Texture plays a significant role in most of the Content Based Image Retrieval systems (CBIR) when compared to other low level visual descriptors. Since there are plenty of methods available for extracting the texture, among which Local Binary Pattern (LBP) is considered the state of the art method for effectively describing the texture. It is not only used for describing the texture apart from that it has been applied to many applications in image and signal processing such as texture classification , face and facial expression recognition ,object tracking and leaf image classification etc. As this considers spatial features only, a similar another method has been existed called local phase quantization which uses the frequency features of the texture by considering neighborhood of a pixel.The Local Binary Patterns (LBPs) are first proposed for representing the texture by encoding the pixel wise information. In this method, given a center pixel in the (3 x 3) window, LBP value is computed by comparing its neighborhoods grayscale value with center pixel. Then the neighboring pixels are assigned with a binary label, which can be either 0 or 1 depending on whether the center pixel has higher intensity value than the neighboring pixel. These binary values multiplied by specific weights and summed up. But they gives variant values when an image is rotated etc. hence an improved LBP has been used which gives same texture properties even on image rotation. After that GLCM metrices has been evaluated from both spatial and frequency descriptors. For retrieving the images by inputting a query image, For different types of similarity metrics has been used in which all four measures gives almost same accuracy rate in retrieving the images but Canberra distance gives highest accuracy on all types of images.

Keywords:*CBIR, Texture, LBP, GLCM, Similarity Measure.*

I. INTRODUCTION

Although the technology of CBIR has developed rapidly, it is still a very challenging problem nowadays. The essence of CBIR is to compare the resemblance of visual features extracted from the query image and the database images. Therefore, extracting appropriate and effective visual features to represent image content correctly is the principal issue in CBIR [1]. These features are expected to be invariant to geometrical transformation, viewpoint lighting condition and so on. Various visual features have been designed to transform image visual content into vector representation since two decades ago. These features can be broadly categorized into global features and local features. The commonly used global features are the holistic statistics of colour, texture and shape that symbolize an image by a single vector, termed feature descriptor. The local features, such as Scale Invariant Feature Transform (SIFT), Speeded Up Robust Features (SURF), Maximally Stable Extremal Regions (MSER), Histogram of Oriented Gradient (HOG) and Local Binary Patterns (LBP), also have been used to represent an image in the form of a collection of feature vectors (descriptors) [2]. However, all these features are low-level visual features without providing semantic information of the underlying images [558, 9], unable to achieve effective image representation and have become one of the main sources of challenges to an efficient CBIR system.

To deal with this difficulty, many machine learning methods have been applied to learning higher level visual features to obtain a semantic representation. One of the successful methods is the bag of visual words (BoVW) model [5]. The key idea of the BoVW is to learn a visual dictionary by clustering local feature descriptors of training images into some representative visual words and quantise each local feature descriptor to the most similar visual word in the visual dictionary such that an image is represented by a histogram of all visual words in the dictionary. To reduce the quantisation error of the BoVW, another similar approach is to seek image sparse representation. In this approach, sparse coding is performed on local features of an input image with trained visual dictionary based on the principle of reconstruction error minimisation and the sparsity constraint on coding coefficients for reconstruction [6]. Next, a high-dimensional sparse vector is generated to represent an image by applying some feature pooling strategy, e.g. sum pooling [5511] and max pooling [7]. The existing methods include Linear Spatial Pyramid Matching Using Sparse Coding (ScSPM) [7], local coordinate coding (LCC) [8], locality-constrained linear coding (LLC) [9] and so on, which provide state-of-the-art performances in image classification [10].

II. RELATED WORKS

The multimedia database Image Retrieval algorithms may be assigned to three types of algorithms:

• based on textual descriptions, most often keywords - Keywords Based Image Retrieval (KBIR) algorithms

• based on semantic information extracted from the image - Semantic Based Image Retrieval (SBIR) algorithms

• based on information which is present in the image - Content Based Image Retrieval (CBIR) algorithms

1) Keywords Based Image Retrieval (KBIR) algorithms

The Keyword Based Image Retrieval algorithms practice textual annotations with the aim of describing the whole image or their parts. Most frequently descriptions are made by humans and the precision of keywords is limited to the knowledge and perception of a person [12]. For objects which are well known it is easy to represent them by annotations and the results of retrieval are very satisfactory. For example a car object may be named very precisely by the brand, model name, version, production year and color.When the object is not well known or it is not calm to describe it precisely by keywords, the results may be imprecise. This is due to the fact that annotations are very subjective and different person may use different words as keywords for the same objects [13], [14]. For example a landscape with trees and water may be annotated by one person as a forest and a river, but by another one as trees and a lake. The third person contrary may use the name of place where the photo was taken. In this situation the results of a query may be imprecise and unsatisfactory for the user. Another drawback of the KBIR approach is that it is hard to automatically add keywords without human interaction.

2) Semantic Based Image Retrieval (SBIR) algorithms

The Semantic Based Image Retrieval algorithms are similar to Keyword Based Image Retrieval algorithms because they use also words to perform queries. However, contrary to them, they permit users to write queries as phrases that are more natural form for them. Such different query interface is used in order to overcome the so called 'semantic gap' which is a difference between what a human could describe and what is present in the image [14], [15]. After defining by an user, the phrases are mapped onto so called semantic features which are correlated with the content of the image [16]. The use of semantic based textual approach is more comfortable and easy for users but still if they do not have the full knowledge about searched images the results may be insufficient. There are also approaches which uses graphical queries which are then transformed into textual description. One of the most interesting research is [17] which uses a sketch as a query, then extracts textual annotations - semantic features and then finds 3D models of objects which are described by similar or the same set of features.

3) Content Based Image Retrieval (CBIR) algorithms

The Content Based Image Retrieval algorithms use information present in the image to perform queries [12].

In this area two types of algorithms could be distinguished:

1)Low level

2)High-level[11].

The first type of algorithms is based on extraction of features for the whole image. There may be statistical image features used, e.g. a normalized color histogram [18]. Another method may be aalteration moment and entropy [19], a spatial domain image representation [20] or a bag of words histogram [21].

There are also approaches which use different MPEG-7 descriptors, e.g. shape and texture descriptors [22]. Since the features describe the whole image, the low-level CBIR algorithms provide very satisfactory results when a query is performed in order to find similar images. However, when an user would like to obtain images with the same object but with different backgrounds, the low level algorithms are not efficient and the results may be insufficient. The high-level CBIR algorithms are more suitable for that situations.

- **Content-Based Image Retrieval Systems**

The greater part of these frameworks have a fundamentally the same architecture for perusing and documenting/ordering images involving tools for the extraction of visual features, for the storage and efficient retrieval of these features, for remotenessdimensions or similarity calculation and a type of graphical user interface (GUI). This general framework setup is shown in Fig. 1.

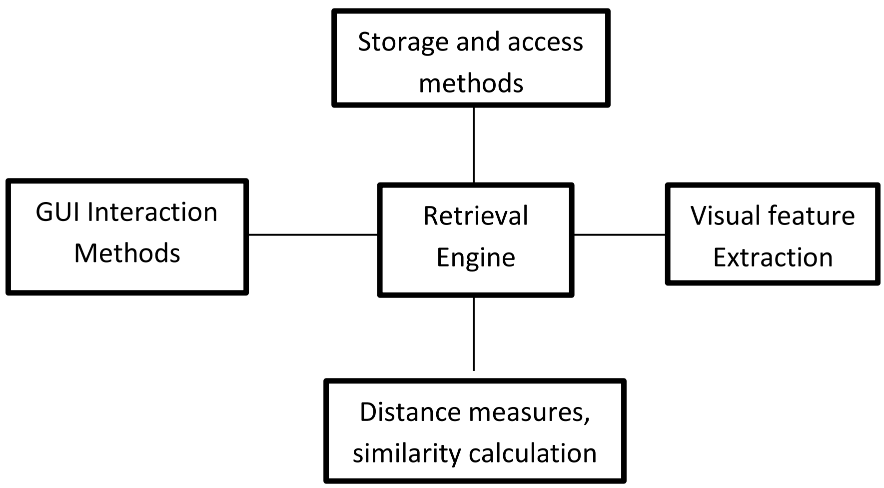

Figure: 1 The principal components of all content-based image retrieval systems.

III. LITERATURE SURVEY

WangmingXuet. al. (2017)presented a new non-negative sparse feature learning approach for CBIR. The key idea is to extract effective visual features to build holistic image representation. The proposed approach comprises three key steps: visual dictionary learning, non-negative sparse feature encoding and feature pooling. In their approach, a modified spectral clustering method is proposed to generate a non-negative visual dictionary.

StanisławDeniziaket. al. (2017) presents a new Content Based Image Retrieval database structure. Each object can be represented as a set of predefined shapes: a line segment, a polygon, a polyline, an arc, a polyarc and an arc-sided polygon. All shapes are connected into a graph, in order to store the mutual relations between them.

KaoutherZekriet. al. (2017) presented an appropriate descriptors selection method for the analysis of image texture in CBIR. Their work is based on the extraction of descriptors attributes GLCM, Log-Gabor and Tamura in order to model the visual content of textured images.

Muhammad HammadMemonet. al. (2016) proposed a technique based on image matching by region based similarity technique for content based image retrieval. The given method might be depends upon simple but effective metrics that have been explained to calculate similarities.

S. Sathiya Devi et. al. (2017) proposed a local binary pattern called Modified Dominant Directional Local Binary Pattern (MDDLBP) for texture feature extraction. It considers four dominant directions such as horizontal, vertical, diagonal and anti-diagonals to compute the binary pattern. The proposed method captures the structural information and the dimensionality is reduced.

IV. PROPOSED WORK

The basic steps in the algorithm are written as below

1. Extract the features of query image using proposed algorithm.
2. Encrypt the features using XOR operation.
3. Compute similarity index between query image feature vector and every database image feature vector.
4. Sort the similarity index.
5. Retrieve images as final results which correspond to shorter distance.
6. Evaluation measure.

The flowchart of the proposed algorithm is shown below:

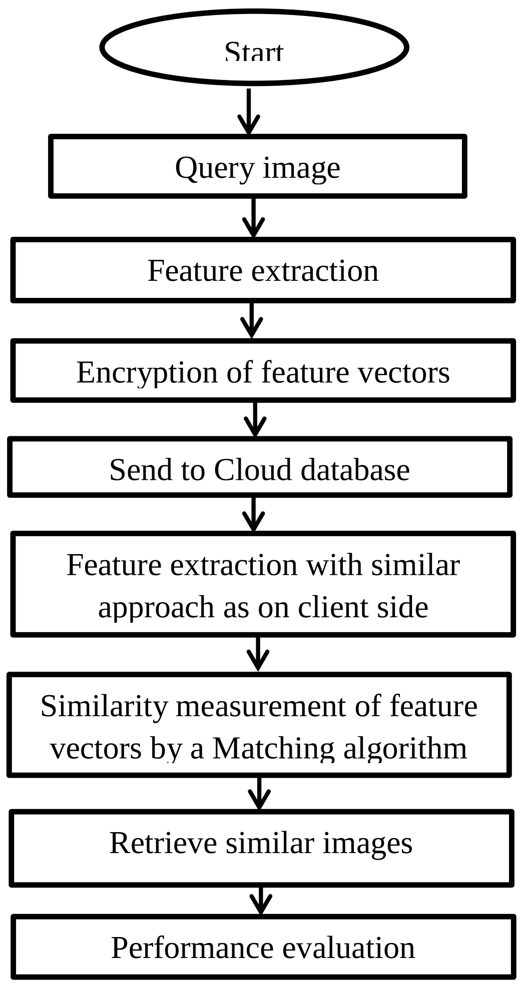

Figure 2: Flowchart for Proposed algorithm

International Journal of Engineering, Science and Mathematics
Vol. 6 Issue 8, December 2017,
ISSN: 2320-0294 Impact Factor: 6.765
Journal Homepage: http://www.ijesm.co.in
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

In feature extraction phase we have used Rotated LBP and LPQ feature matrices which are further used to calculate GLCM feature vector. The problem of variations to rotations in LBP arises due to the fixed arrangement of weights. As the weights are aligned in a circular manner, the effect of image rotations can be countered by rotating the weights by the same angle while computing the descriptor. Since the angle of the rotation cannot be known, we propose an adaptive arrangement of weights based on the locally computed reference direction. LPQ is another widespread histogram-based feature extractor from the family of local texture descriptors, which performs the assessment of phase in a local window at the pixel position [28]. Local phase analysis in frequency domain leads to a detailed illumination insensitive texture description of input images [29][30]. LPQ is also insensitive against image degradation, blur effect, which happens usually in real world applications, such as video surveillance, which is caused by out of focus of camera or object motion. After evalutating features different similarity measures has been used to retrieve the images. It has been found that canbarra distance similarity measures show much improved results than other similarity measures techniques.

IV. RESULT AND DISCUSSION

We have applied different similarity measures on many images. The results on an image are shown below:

Here it is the input query image taken to apply different similarity measure.

Figure 3: Input query image two

Canbarra distance similarity measure is applied on given input query image.

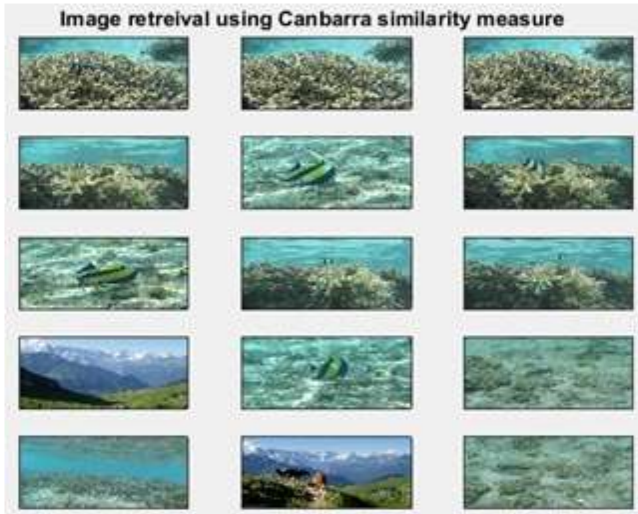

Figure 4: Retrieved images using Canbarra distance similarity measure

The table 1 shows the results of different similarity measures applied on above image. It has been found that Canberra distance similarity measures show much improved results than other similarity measures techniques.

Table 1: Recall and precision ratio for image two

Similarity measure used	Recall ratio	Precision ratio
Euclidian distance	1	0.9230
Chi squre distance	1	0.8461
Canbarra distance	1	0.9230
Manhattan distance	1	1

The bar chart shown below represent the results for given image which also shows that the results for canbera distance similarity measures are much improved than other similarity measures techniques.

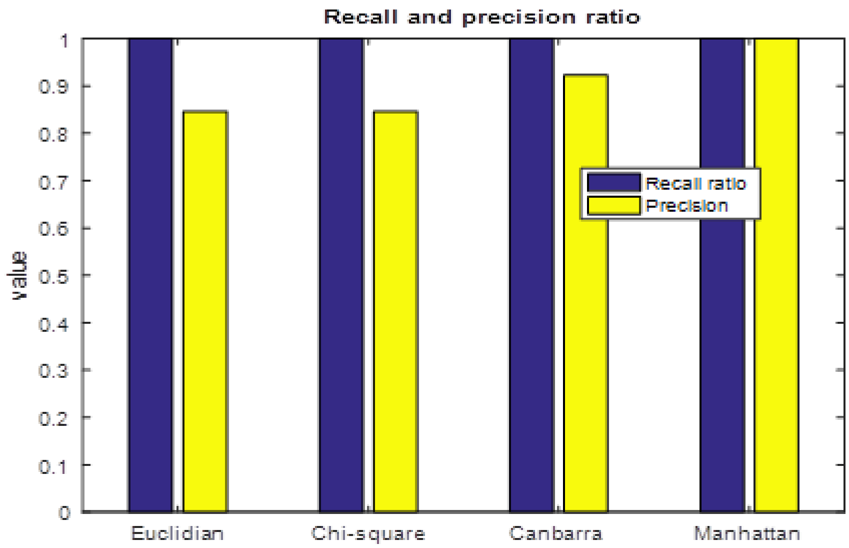

Figure 5: Recall and precision ratio

V. CONCLUSION

Proposed feature extraction method based on LBP is one of the most successful local feature extractors, which extracts texture features of the image by comparing each pixel with its neighbors in a small neighborhood. There is no training requirement, which makes it fast and easy to integrate into the new data sets. Furthermore, because of the application of histograms as the feature sets, it is robust against rotation and scaling. Also, the image-size dimension can be reduced to the number of histogram bins. Since the grey values of neighbors are compared, it is robust against monotonic variations in the images. Similarly the local phase quantization (LPQ) method is based on the blur invariance property of the Fourier phase spectrum. It uses the local phase information extracted using the 2-D DFT or, more precisely, a short-term Fourier transform (STFT) computed over a rectangular M-by-M neighborhood. Local phase analysis in frequency domain leads to a detailed illumination insensitive texture description of dataset images. After feature extraction,encryption has been applied to the feature set and is tested with the encrypted features of the dataset images over the cloud. Four different similarity measure metrics has been used in order to check the performance of the proposed algorithm. It has been found that all similarity measures gives efficient results in retrieving similar images but similarity metrics named as canbarra distances are more proficient than the Euclidian, manhattan and Chi-square distances.

References

[1]Gandhani, S., Singhal, N.: 'Content based image retrieval: survey and comparison of CBIR system based on combined features', Int. J. Signal Process. Image Process. Pattern Recognit., 2015, 8, (10), pp. 155–162

[2]Mikolajczyk, K., Schmid, C.: 'A performance evaluation of local descriptors', IEEE Trans. Pattern Anal. Mach. Intell., 2005, 27, (10), pp. 1615–1630

[3]Xu, G., Xu, L., Li, X., et al.: 'An image retrieval method based on visual dictionary and saliency region', Int. J. Signal Process. Image Process. Pattern Recognit., 2016, 9, (7), pp. 263–274

[4] Tsai, C.F.: 'Bag-of-words representation in image annotation: a review', ISRN Artif. Intell., 2012, 2012, pp. 1–19

[5]Sivic, J., Zisserman, A.: 'Video google: a text retrieval approach to object matching in videos'. Proc. the 9th Int. Conf. Computer Vision (ICCV), Nice, France, October 2003, pp. 1470–1477

[6] Lee, H., Battle, A., Raina, R., et al.: 'Efficient sparse coding algorithms'. Proc. the 21st Annual Conf. Advances in Neural Information Processing Systems (NIPS), Vancouver, BC, Canada, December 2007, pp. 801–808

[7] Yang, J., Yu, K., Gong, Y., et al.: 'Linear spatial pyramid matching using sparse coding for image classification'. Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), Miami, FL, USA, June 2009, pp. 1794– 1801

[8] Yu, K., Zhang, T., Gong, Y.: 'Nonlinear learning using local coordinate coding'. Proc. 23rd Annual Conf. Advances in Neural Information Processing Systems (NIPS), Vancouver, BC, Canada, December 2009, pp. 2223–2231

[9] Wang, J., Yang, J., Yu, K., et al.: 'Locality-constrained linear coding for image classification'. Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), San Francisco, CA, USA, June 2010, pp. 3360–3367

[10] Chatfield, K., Lempitsky, V., Vedaldi, A., et al.: 'The devil is in the details: an evaluation of recent feature encoding methods'. Proc. the 22th British Machine Vision Conf. (BMVC), Dundee, UK, 2011, pp. 1–12

[11] S. Deniziak and T. Michno, "Content based image retrieval using query by approximate shape," in 2016 Federated Conference on Computer Science and Information Systems (FedCSIS), Sept 2016, pp. 807–816.

[12] S. law Deniziak and T. Michno, "Query by shape for image retrieval from multimedia databases," Beyond Databases, Architectures and Structures, p. 377.

[13] C.-Y. Li and C.-T. Hsu, "Image retrieval with relevance feedback based on graph-theoretic region correspondence estimation," IEEE Transactions on Multimedia, vol. 10, no. 3, pp. 447–456, April 2008.

[14] H. H. Wang, D. Mohamad, and N. A. Ismail, "Approaches, challenges and future direction of image retrieval," CoRR, vol. abs/1006.4568, 2010.

[15] A. Singh, S. Shekhar, and A. Jalal, "Semantic based image retrieval using multi-agent model by searching and filtering replicated web images," in Information and Communication Technologies (WICT), 2012 World Congress on, Oct 2012. pp. 817–821.

[16] C.-Y. Li and C.-T. Hsu, "Image retrieval with relevance feedback based on graph-theoretic region correspondence estimation," Multimedia, IEEE Transactions on, vol. 10, no. 3, pp. 447–456, April 2008.

[17] B. Li, Y. Lu, and J. Shen, "A semantic tree-based approach for sketchbased 3d model retrieval," in 2016 23rd International Conference on Pattern Recognition (ICPR), Dec 2016.

[18] M. Mocofan, I. Ermalai, M. Bucos, M. Onita, and B. Dragulescu, "Supervised tree content based search algorithm for multimedia image databases," in 2011 6th IEEE International Symposium on Applied Computational Intelligence and Informatics (SACI), May 2011.

[19] H. P. Kriegel, P. Kroger, P. Kunath, and A. Pryakhin, "Effective similarity search in multimedia databases using multiple representations," in 2006 12th International Multi-Media Modelling Conference, 2006.

[20] T. K. Shih, "Distributed multimedia databases," T. K. Shih, Ed. Hershey, PA, USA: IGI Global, 2002, ch. Distributed Multimedia Databases, pp. 2–12. ISBN 1-930708-29-7.

[21] A. Sluzek, "Machine vision in food recognition: Attempts to enhance CBVIR tools," in Position Papers of the 2016 Federated Conference on Computer Science and Information Systems, FedCSIS 2016, Gda´nsk, Poland, September 11-14, 2016., M. Ganzha, L. A. Maciaszek, and M. Paprzycki, Eds., 2016.

[22] C. Lalos, A. Doulamis, K. Konstanteli, P. Dellias, and T. Varvarigou, "An innovative content-based indexing technique with linear response suitable for pervasive environments," in 2008 International Workshop on Content-Based Multimedia Indexing, June 2008.

[23]W. Xu, S. Wu, M. J. Er, C. Zheng and Y. Qiu, "New non-negative sparse feature learning approach for content-based image retrieval," in *IET Image Processing*, vol. 11, no. 9, pp. 724-733, 9 2017.

[24]S. Deniziak and T. Michno, "New content based image retrieval database structure using query by approximate shapes," *2017 Federated Conference on Computer Science and Information Systems (FedCSIS)*, Prague, Czech Republic, 2017, pp. 613-621.

[25]K. Zekri, A. G. Touzi and Z. Lachiri, "A comparative study of texture descriptor analysis for improving content based image retrieval," *2017 International Conference on Control, Automation and Diagnosis (ICCAD)*, Hammamet, Tunisia, 2017, pp. 247-253.

[26]M. H. Memon, J. P. Li, I. Memon, Q. A. Arain, S. Jamil and M. H. Memon, "Localization in images matching through region-based similarity technique for content-based image retrieval," *2016 13th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, Chengdu, 2016, pp. 269-272.

[27]S. S. Devi and R. Balasundaram, "Content based texture image retrieval based on modified dominant directional local binary pattern," *2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Coimbatore, 2017, pp. 1-6.

[28] Dhall, A., Asthana, A., Goecke, R., Gedeon, T.: 'Emotion recognition using PHOG and LPQ features'. Proc. IEEE Int. Conf. on Automatic Face and Gesture Recognition and Workshops (FG11), Santa Barbara, CA, 2011, pp. 878–883

[29] Zang, D.: 'Illumination invariant object tracking based on multiscale phase'. Proc. IEEE Int. Conf. on Image Processing (ICIP10), Hong Kong, 2010, pp. 357–360

[30] Li, Y., Shan, S., Zhang, H., Lao, S., Chen, X.: 'Fusing magnitude and phase features for robust face recognition'. Proc. Asian Conf. on Computer Vision (ACCV12), Daejeon, 2012, pp. 601–612