International Journal of
Emerging Research in Management &Technology
ISSN: 2278-9359 (Volume-4, Issue-5)

Research Article **May 2015**

HIK Based Image Classification Using Saliency Based Multiscale Representation

Anju Davis *, Suhara S
Dept of CSE & University of Calicut,
KMCT College of Engineering,
Kerala, India

***Abstract*—**

HIK based image classification method is proposed which efficiently classifies the images with less background. HIK kernel is used with saliency driven multiscale fusion. The generated scale space in general preserves or even enhances semantically important structures such as edges, lines, or flow like structures in the foreground, and smoothes clutter in the background. The image is classified using multiscale information fusion based on the original image, the image at the final scale at which the diffusion process converges, and the image at a midscale. The proposed method emphasizes the foreground features, which are important for image classification. The background image regions, whether considered as contexts of the foreground or noise to the foreground, can be globally handled by fusing information from different scales.HIK kernel is used for classification, to handle the cases of incorrect detection of saliency.HIK based visual code book generation algorithm is used for visual vocabulary creation, in order to classify the images with less background effectively. Experiments were conducted in 17 Oxford Flowers, 102 Oxford Flowers and Caltech 101 datasets. The proposed method has an accuracy of 81.3%.

***Keywords*— multiscale fusion,visual vocabulary, classification,saliency, HIK.**

I. INTRODUCTION

Image classification[1] is a very active research topic which has stimulated researches in many important areas of computer vision, including feature extraction and feature fusion, the generation of visual vocabulary [5], the quantization of visual patches to produce visual words , pooling methods , and classifiers. Classification includes a broad range of decision theoretic approaches to the identification of images .

In image classification, it is an important but difficult task to deal with the background information. The background is oftentreated as noise; nevertheless, in some cases the background provides a context, which may increase the performance of image classification. Zhang et al. [33] experimentally analyzed the influence of the background on image classification. They demonstrated that although the background may have correlations with the foreground objects, using both the background and foreground features for learning and recognition yields less accurate results than using the foreground features alone. Overall, the background information was not relevant to image classification. Heitz and Koller [10] showed that spatial context information may help to detect objects. Shotton et al. [21] proposed an algorithm for recognizing and segmenting objects in images, using appearance, shape, and context information. They assumed that the background is useful for classification and there are correlations between foreground and background in their test data. Galleguillos et al. [5] proposed an algorithm that uses spatial context information in image classification. The input image was first segmented into regions and each region was labeled by a classifier. Then, spatial contexts were used to correct some of the labels based on object co-occurrence. The results show that combining co-occurrence and spatial contexts improves the classification performance. From the previous work, we conclude that image classification is faced with the partial matching problem [8], [14]: some features obtained from images in the same class differ significantly from one image to another because of background clutter and occlusion of the foreground objects by other objects. The influence of background on image classification varies. Only semantically important contexts, such as object co-occurrence, or particular object spatial relations are helpful for image classification. Backgrounds which contain only clutter provide no information to support image classification. It is interesting to filter out background clutter and simultaneously use the background context to increase the performance of image classification.

Context aware image classification is a very difficult task in computer vison.The background of the image can be a context for the image or it can be just the clutter. An idealistic approach has to be developed where the background is considered when it is the context for the object in the scene. The background has to be avoided when it is just the clutter. Most of the image classification methods fail to detect the objects with less background since they do not give any importance to the background. In the proposed method, when the background is only clutter it will be removed in the multi scale space generation.

Goal of the project is to build an image classification system, which classifies the images with less background accurately. In multiscale saliency driven diffusion,the background information is preserved in the original image and the image at the midscale and the background information is smoothened in the larger scales. In the visual vocabulary creation the less important image features will be filtered out especially when the euclidean distance is used to create the visual vocabulary. In k-means clustering algorithm for visual vocabulary creation,euclidean distance is used to find the distance between the image features. Sometimes the same feature will be assigned to more than one cluster centres,this can lead to losing the information. In the proposed system HIK kernel based distance is used for creating the visual dictionary. In order to achieve a better classification performance HIK kernel based SVM classifier is being used.

II. PROPOSED METHOD

Image classification using multiscale fusion uses information fusion to improve the performance of image classification. Saliency map is used to differentiate between the foreground and the background. The multiscale space is generated by applying the saliency map as a mask on the l2 norm of the gradients. The multiscale space images are given to bag of words model for classification.The muliscale information fusion method classifes the images with less background inaccurately. Histogram Intersection Kernel[3] based multiscale fusion method is proposed to overcome the drawbacks of the saliency driven multiscale fusion. HIK based mutiscale fusion generates a scale space using Perona-Malik nonlinear anisotropic diffusion.

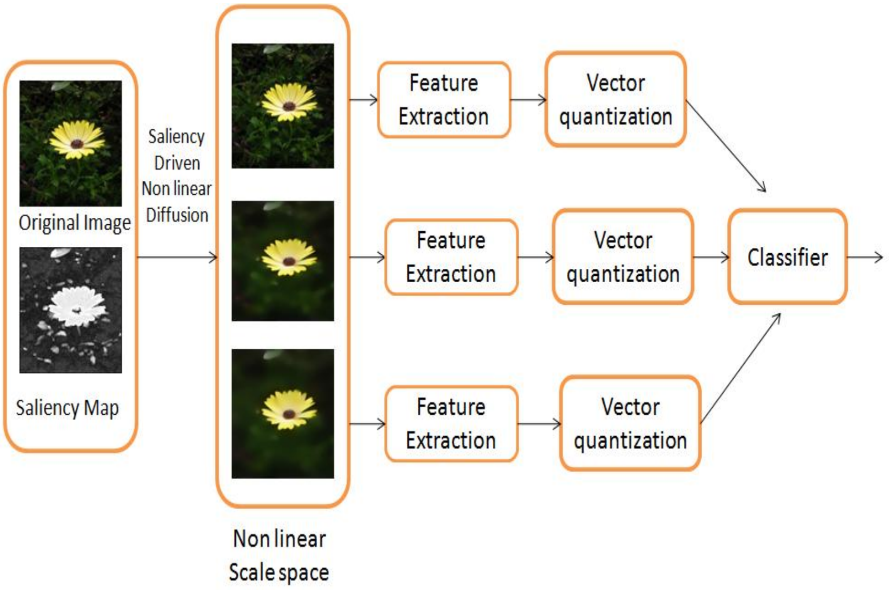

Fig 2.1 The framework of HIK based Multiscale Fusion.

A. Saliency Map Generation

An average of the saliency maps generated using Cheng's[13] method and Goferman's[2]method is used for saliency driven nonlinear diffusion. Cheng's method[13] is based on the Histogram Contrast. An image histogram is created by color quantization(median cut algorithm). In the image histogram the colour difference between the pixels are computed. The sum of the colour difference is taken as the saliency value and each of the pixels are later on replaced by this saliency value. Goferman's method[14] is based on context aware saliency where the image is divided into patches of 7×7.The colour difference between the patches are taken as the saliency value. The color distance between the pixels are calculated in the CIE L*a*b color space instead of the RGB color space.

B. Saliency Driven Nonlinear Diffusion

Nonlinear anisotropic diffusion[15] is used for generating the scale space for classification. Saliency map is used as the mask for performing the diffusion. The saliency mask Is is applied on the norm of the gradient,such that Is works as a mask that indicate the region of interest.
The diffusion equation is defined by

$$\mathrm{u(x,y,t)} = \mathrm{f(x,y)} \text{ if } \mathrm{t{=}0}$$

$$\partial_t \mathrm{u} = \mathrm{div}(\mathrm{g}(\nabla \mathrm{u}, I_s)\nabla \mathrm{u}) \text{ if } \mathrm{t{>}0}$$

Where the diffusivity is defined by

$$g(\nabla u, I_s) = \{1 - exp(-\frac{c}{(\frac{I_s\|\nabla u\|}{\lambda})^m})$$

Here λ is the contrast parameter and m controls the speed of diffusivity. In the experiments c is taken as 1 and m is 100. The value of λ depends on the image. The λ value is determined by doing an edge detection on both the original image and the saliency map. The λ value is the difference between the number of the salient edges and the nonsalient edges.

C. Multiscale representation and classification

Harris laplace sampling and dense sampling are used for generating the local patches,where each of the patch corresponds to the point of interest. Features are extracted from the patches using SIFT[15] and four colour descriptors,opponentSIFT, rgSIFT, CSIFT and RGBSIFT. For each of the descriptors the training images are clustered using k-means to generate a visual vocabulary of 4000 words. Soft coding is used to generate feature vector of images. SVM classifier is learned to classify bike,car and ower images. A training set was created by including 25 images from each of the category for each of the datset. Validation set was created by including 20 images from each of the category for each of the data set. 20 images from each of the category is used during testing time. The HIK kernel used to find the histogram distance. HIK kernel has a time complexity of O(n log m). For multi scale space the distances obtained at multiple scales are combined using weighted averaging[29].

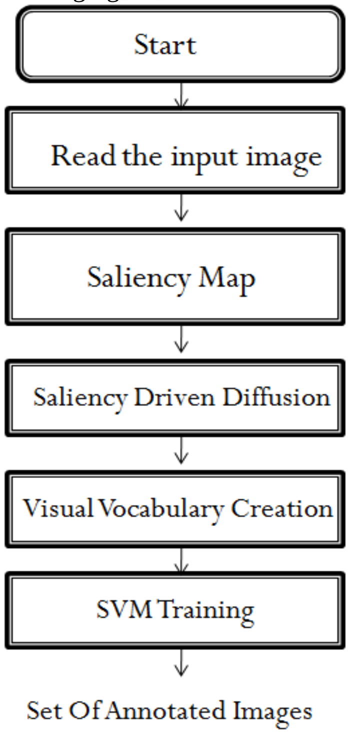

Fig 2.2 Flow Chart For SVM Training

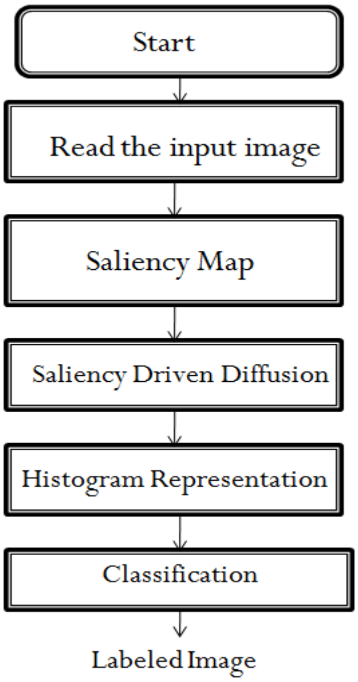

Fig 2.3 Flow chart for Image Identification

The proposed HIK based multiscale fusion has several advantages.

- The nonlinear diffusion based multiscale space image representation can preserve or enhance semantically important image structures at large scales.
- The saliency driven nonlinear diffusioncan divide the foreground from the background at large scales, with only a little loss of the foreground information.

- Saliency driven nonlinear diffusion can deal with the background information no matter whether it is a context or noise, and the same method can be applied to video as well.
- Proposed method can partly handle incorrect detection of saliency, by including the original image at scale 0 in the set of scaled images used for classification.
- HIK based multiscale representation can be easily combined with any existing image classification algorithms.

III. RESULTS

The multiscale information fusion is implemented using java.The proposed HIK based multiscale fusion is tested in three public datasets the 17 Oxford Flowers dataset, the 102 Oxford Flowers dataset and Caltech101 dataset. The 17 Oxford Flowers dataset [19] contains images from 17 flower categories with 80 images per category. Fig. 12 shows some example images in the dataset. For each flower category, 40 images were used for training, 20 for validation, and 20 for testing. Two types of samplings the Harris-Laplace sampling and the dense sampling. Five types of descriptor – the SIFT and four types of the color-SIFT descriptors were used for experiments. In the 17 Oxford Flowers dataset the proposed method has a recognition accuracy of 92%. The 102 Oxford Flowers dataset [18] contains 8189 images from 102 flower categories with 40-250 images per category.For each category, 10 images were used for training, 10 for validation, and the rest for testing. For each image, two sampling methods, the Harris-Laplace point sampling and dense sampling, were used to generate local patches where each patch corresponds to a point of interest. Then, each image patch was further represented by the SIFT and the four color-SIFT descriptors : OpponentSIFT, rgSIFT, C-SIFT, and RGB-SIFT. An SVM classifier was trained using the training images. The parameters were estimated on the validation set and further used on the test set.In this dataset the proposed method has a recognition accuracy of 82%. The Caltech 101 dataset, consists of images from 101 object categories.From this datasets experiments were conducted in car and bike dataset.The same experimental setup used in the above datasets was used in Caltech 101.The proposed method has recognition accuracy of 85%.

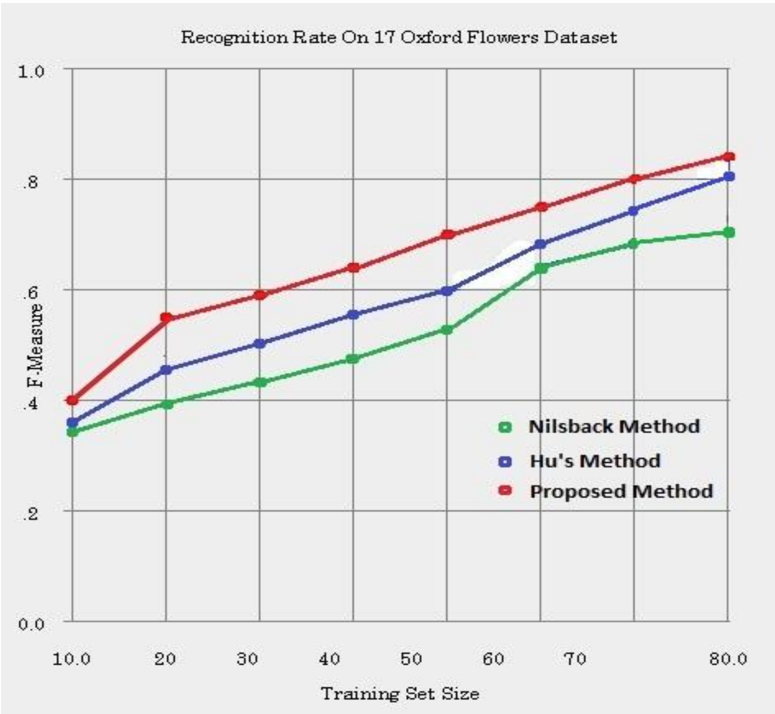

Fig 3.1 F-measure Analysis

IV. CONCLUSION

HIk based image classification based on multiscale fusion is proposed in this paper to effectively classify the images with less background. The saliency driven multi-scale nonlinear diffusion filtering, by modifying the mathematical equations for nonlinear diffusion filtering, and determining the diffusion parameters using the saliency detection results is proposed in this work. The new method is applied to image classification and it outperforms the available methods for classification. The proposed anisotropic nonlinear diffusion method for scale space generation leads to edge preserving smoothing and the automatic segmentation of the image. The saliency driven nonlinear multi-scale space preserves and even enhances important image local structures, such as lines and edges, at large scales. The proposed method uses HIK distance for code book generation in order to effectively classify the images with less background. Multi-scale information has been fused using a weighted function of the distances between images at different scales. The saliency driven multi-scale representation can include information about the background in order to improve image classification. Experiments have been conducted on widely used datasets, namely the Caltech101 dataset and the Oxford 17 owers[17] dataset. The results have demonstrated that HIK based saliency driven multi-scale information fusion improves the accuracy of image classification.

V. FUTURE SCOPE

Video classification can be carried out using the proposed method.HIK based multiscale fusion can be used for satellite image classification.

REFERENCES

[1] W. Hu, R. Hu, H. Ling, N. Xie, and S. Maybank, *Image Classi_cation Using Multiscale Information Fusion Based on Saliency Driven Nonlinear Diffusion Filtering*, IEEE Trans.on Image Processing, vol. 23, no. 4,Apr. 2014,pp. 1513-1526.

[2] J. Yang, K. Yu, and T. Huang, *Supervised translation invariant sparse coding* ,in Proc.IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2010, pp. 3517-3524.

[3] F. Li, G. Lebanon, and C. Sminchisescu, *Chebyshev approximations to the histogram 2 Kerne*l,in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Apr. 2012, pp. 2424-2431.

[4] S. Maji, A. C. Berg, and J. Malik, *Classification using Intersection Kernel Support Vector Machines is Efficient*, in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp. 40-48 ,Jun.2008.

[5] J. James M. Rehg, *Beyond the Euclidean distance: Creating e_ective visual codebooks using the histogram intersection kernel*, in Proc IEEE Con. Vis. Computer Vision.,2009, pp. 630-637.

[6] S. Gao, I. W. Tsang, and L. Chia, *Sparse representation with kernels*, IEEE Trans. Image Process., vol. 22, no. 2,pp. 423-434, February 2013..

[7] S. Lazebnik, C. Schmid, and J. Ponce, *Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories*, in Proc. IEEE Conf. Comput. Vis. Pattern Recognit.,pp. 2169-2178, vol. 2. Jun. 2006.

[8] M. Varma and D. Ray, *Learning the discriminative power invariance trade-off*, in Proc.IEEE Int. Conf. Comput. Vis., pp. 18,Oct. 2007.

[9] M.-E. Nilsback and A. Zisserman, *A visual vocabulary for flower classification*, in Proc.IEEE Conf. Comput. Vis. Pattern Recognit.,vol. 2. pp. 1447-1454. Jun. 2006.

[10] J. Zhang, M. Marszalek, S. Lazebnik, and C. Schmid, *Local features and kernels for classification of texture and object categories: A comprehensive study*, Int. J. Comput. Vis., vol.73, no. 2, pp. 213-238, Jun. 2007.

[11] C. Galleguillos, A. Rabinovich, and S. Belongie, *Object categorization using co-occurrence,location and appearance*, in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp. 18,Jun.2008.

[12] G. Heitz and D. Koller, *Learning spatial context: Using stuff to find things*, in Proc. Eur.Conf. Comput. Vis., 2008, pp. 30-43.

[13] M. M. Cheng, G. X. Zhang, N. J. Mitra, X. L. Huang, and S. M. Hu, *Global contrast based salient region detection*, in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Dec. 2011, pp.409-416.

[14] S. Goferman, L. Zelnik-Manor, and A. Tal, *Context-aware saliency detection*,in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2010, pp. 2376-2383.

[15] J.Weickert, *A review of nonlinear diffusion filtering*, in Scale- Space Theory Compute Vision (Lecture Notes in Computer Science). New York, NY, USA: Springer-Verlag, 1997, pp. 128.

[16] K. E. A. van de Sande, T. Gevers, and C. G. M. Snoek, *Evaluation of color descriptors for object and scene recognition* ,IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 9, pp.1582-1596, Sep. 2010.

[17] M. E. Nilsback and A. Zisserman, *Automated flower classification over a large number of classes*, in Proc. Indian Conf. Comput. Vis., Graph. Image Process., Feb. 2008, pp. 722-729.

[18] J. Harel, C. Koch, and P. Perona, *Graph-based visual saliency*, in Proc. Annu. Conf. Neural Inf. Process. Syst., 2007, pp. 545-552.

[19] L. Itti, C. Koch, and E. Niebur, *A model of saliency-based visual attention for rapid scene analysis*, IEEE Trans. Pattern Anal. Mach. Intell., vol. 20, no. 11, pp. 1254-1259, Nov. 1998.

[20] Liu, Z. Yuan, J. Sun, J. Wang, N. Zheng, X. Tang, et al., *Learning to detect a salient object*, IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, no. 2, pp. 35-36, Feb. 2011.

[21] R. Valenti, N. Sebe, and T. Gevers, *Image saliency by isocentric curvedness and color*, in Proc. IEEE Int. Conf. Comput. Vis., Sep. 2010, pp. 2185-2192.

[22] D. Walther and C. Koch, *Modeling attention to salient proto-objects*, Neural Netw., vol. 19,no. 9, pp. 1395-1407, 2006.

[23] N. Xie, H. Ling, W. Hu, and X. Zhang, *Use bin-ratio information for category and scene classification*, in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2010, pp. 2313-2319.

[24] J. Wu and J.M. Rehg, *Beyond the Euclidean distance: Creating effective visual codebooks using the histogram intersection kernel*, in Proc. IEEE Conf. Comput. Vis.,Jun. 2009,pp.630-637

[25] M.T. Mahmood and T.-S. Choi, *Nonlinear approach for enhancement of image focus volume in shape from focus*, IEEE Trans. Image Process., vol. 21, no. 5, pp. 2866-2873, May 2012.

[26] M. R. Hajiaboli, M. O. Ahmad, and C. Wang, *An edge-adapting Laplacian kernel for non-linear diffusion filters*, IEEE Trans. Image Process., vol. 21, no. 4, pp. 1561-1572, Apr. 2012.

[27] P. Rodrigues and R. Bernardes, *3-D adaptive nonlinear complex- diffusion despeckling filter*, IEEE Trans. Med. Imaging, vol. 31, no. 12, pp. 2205-2212, Dec. 2012.

[28] B. Abdollahi, A. El-Baz, and A. A. Amini, *A multi-scale non-linear vessel enhancement technique*, in Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., May 2011, pp. 3925-3929.

[29] J. L. Wu, *Collaborative filtering on the netix prize dataset*, Ph.D. dissertation, Dept. Com-put Sci., Peking Univ.,

Beijing, China, 2010,pp. 82-85.

[30] C.-C. Chang and C.-J. Lin. (2001). *LIBSVM: A Library for Sup- port Vector Machines [Online]*. Available: http://www.csie.ntu.edu.tw/cjlin/libsvm/

[31] M. E. Nilsback and A. Zisserman, *Automated flower classification over a large number of classes*, in Proc. Indian Conf. Comput. Vis., Graph. Image Process., Feb. 2008, pp. 722-729.

ABOUT AUTHOR

Ms. Anju Davis completed her B.Tech in Computer Science & Engineering from University Of Calicut in the year 2007.She has 6 years of IT Industry experience with iGATE global Solutions Bangalore. She is currently pursuing M.Tech in Computer Science & Engineering from KMCT College Of Engineering, Calicut, Kerala. Her research interests include image processing.

