ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 2, February 2016

A Survey on Saliency Driven Nonlinear Diffusion Filtering for Classifying Image Using Multi-scale Information Fusion

Swapna R. Kharche[1], Prof.B.K.Chaudhari[2]

M.E. Student, Department of Computer Engineering, V B Kolte College of Engineering, Malkapur, M.S, India[1]

Assistant Professor, Department of Computer Engineering, V B Kolte College of Engineering, Malkapur, M.S., India[2]

ABSTRACT: Image classification is one of the most challenging problems in computer vision. It refers to the labeling of images into one of the predefined categories. In image classification, it is very difficult to deal with background information. A nonlinear diffusion which has been widely used in image de-noising, enhancement, etc, can preserve or even enhance the semantically important image structures, such as edges and lines. The image is classified using multi-scale 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. In this paper we will see how saliency driven nonlinear diffusion filtering and multi-scale information fusion help to improve classification performance.

KEYWORDS*:* saliency detection, nonlinear diffusion, information fusion, image classification

I. INTRODUCTION

Classification between the objects is easy task for humans but it has proved to be a complex problem for machines. The raise of high-capacity computers, the availability of high quality and low-priced video cameras, and the increasing need for automatic video analysis has generated an interest in object classification algorithms. A simple classification system consists of a camera fixed high above the interested zone, where images are captured and consequently processed. Classification system consists of database that contains predefined patterns that compares with detected object to classify into proper category. Image classification is an important and challenging task in various application domains, including biomedical imaging, biometry, video surveillance, vehicle navigation, industrial visual inspection, robot navigation, and remote sensing.

Humans routinely and effortlessly judge the importance of image regions, and focus attention on important parts. Saliency originates from visual uniqueness, unpredictability, rarity, or surprise, and is often attributed to variations in image attributes like color, gradient, edges, and boundaries. Extracted saliency maps are widely used in many computer vision applications including object of-interest image segmentation, object recognition, adaptive compression of images, content aware image resizing, and image retrieval.

In image classification, it is an important but difficult task to deal with the background information. The background is often treated as noise; nevertheless, in some cases the background provides a context, which may increase the performance of image classification. In order to deal effectively with the background information, a saliency driven nonlinear diffusion filtering is used to generate a multi-scale space, in which the information at a scale is complementary to the information at other scales [1].

As shown in Fig. 1, at large scales, the background is filtered out and the foreground which contains importance structure is preserved. At small scales, background and foreground regions are both preserved. The fusion of information from different scales helps to improve the image classification performance.

Fig.1. An example of saliency driven diffusion filtering. (a) The original image. (b) The image after diffusion. (c) The saliency map.

II. RELATED WORK

Heitz and Koller [2] showed that spatial context information may help to detect objects. Shotton et al. [3] 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. Zhang et al. [4] 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. Itti, Koch, and Niebur (1998) introduced a model for bottom-up selective attention based on serially scanning a saliency map, which is computed from local feature contrasts, for salient locations in the order of decreasing saliency. C.C. Chang and C.J. Lin, et al., [5] LIBSVM employs Classification methods for more number of images. However, this article does not intend to teach the practical use of LIBSVM, for instructions of using LIBSVM.

Galleguillos et al. [6] proposed an algorithm that uses spatial context information to classify image. 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 result shows that combining co-occurrence and spatial contexts improves the classification performance.

III. PREPROCESSING OF IMAGE

The aim of preprocessing is an improvement of image data that suppress unwanted distortions or enhance some image features important for further processing. Pre-processing could be a method to get rid of noises from the linear image. Fig.2. shows the system flow diagram. Preprocessing of input image gives filtered image by avoiding noise or distortion in image, then nonlinear diffusion which filter out background. The saliency detection carried out to find the interesting area in target image and feature calculation by SIFT. The multi-scale information fusion carried out to correctly identify image.

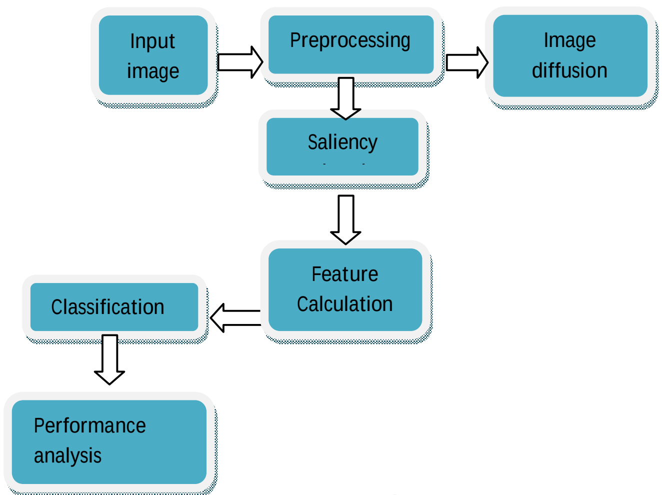

Fig.2 System Flow Diagram

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 2, February 2016

IV. SALIENCY DRIVEN NONLINEAR DIFFUSION

A. *Nonlinear Diffusion*

Let $u(x, y, t)$ be the grey value at position (x, y) and scale t in the multi-scale space. The image diffusion filtering is defined by the diffusion equation [7]:

$$\partial_t u = div\ (D \cdot \nabla u) = div \cdot (D \cdot \nabla u) \qquad (1)$$

where $\nabla$is gradient operator: $\nabla = (\partial/\partial x, \partial/\partial y)$, "*div*" is the divergence operator, and D is the diffusion tensor which is a positive definite symmetric matrix.

If the D in (1) is a function $g(\nabla u)$ of the gradient ∇u of the evolving image u itself, then Equation (1) defines a nonlinear diffusion filter [7], [8].

The nonlinear diffusion filtering is represented as:

$$\partial_t u = div\ (D \cdot \nabla u) = div\ (g\ (\nabla u)\ \nabla u). \qquad (2)$$

The nonlinear diffusion preserves and enhances image structures defined by large gradient values. If image structures with large gradients are all in the foreground, nonlinear diffusion filters out the background.

B. *Saliency Detection*

Saliency detection methods can be grouped as supervised and unsupervised. We first introduce the supervised methods and then the unsupervised methods.

Supervised methods [6], [9] estimate the saliency map using a classifier which is trained with samples for which saliency is well labeled. Marchesotti et al. [10] trained a classifier each target image using the images most similar to an annotated database for saliency detection. The basic assumption is that images sharing a globally similar visual appearance are likely to share similar saliencies. This supervised saliency detection needs a very large well-labeled database, which is not easy to obtain.

Unsupervised saliency detection [11], [12] usually starts with features extraction of image structures known to be salient for the human visual system (HVS). These structure features include the intensity

of salient regions, and the orientation, position and color of edges. Goferman's method[13] is based on context aware saliency detection where the image is divided into patches of 7×7.The color 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, so need conversion of RGB to L*a*b color space. He summarized the following three principles for saliency detection by the HVS.

• Local low level considerations, including factor such as color and contrast [14], [15], [16].
• Global consideration with frequently occurring features should be suppressed [16], [17], [18].
• The salient pixels should be grouped together, rather than scattered across the image.

Cheng et al. [19] proposed a histogram-based contrast method to measure saliency. Their algorithm separates a large object from its surroundings, and enables the assignment of similar saliency values to homogenous object regions, and highlights entire objects. An image histogram is created by color quantization (median cut algorithm). In the image histogram the color difference between the pixels are computed. The sum of the color difference is taken as the saliency value and each of the pixels are later on replaced by this saliency value.

C. *Saliency Driven Nonlinear Diffusion*

Nonlinear anisotropic diffusion [7] 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 [1]

$$U(x, y, t) = f\ (x, y) \text{ if } t = 0$$
$$\partial_t\ u = div\ (g\ (\nabla u, I_s \nabla u) \text{ if } t > 0. \qquad (3)$$

Saliency driven nonlinear diffusion preserved foreground regions and largely smoothed the background regions.

V. MULTI-SCALE IMAGE REPRESENTATION AND CLASSIFICATION

Images whose foregrounds are clearer than their backgrounds are more likely to be correctly classified at a large scale, and images whose backgrounds are clearer are more likely to be correctly classified at a small scale. So, different scales information can be combined to obtain more efficient results of image classification. At a large scale,

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 2, February 2016

the background is filtered out and the foreground is preserved. At a small scale, both the background and the foreground regions are preserved. When the background is relative to the foreground, at small scale the images from the same category are more similar than at a large scale. When the background is noise, the images are more similar at a large scale. Through this multi-scale representation, background information can be utilized.

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 color descriptors, opponentSIFT, rgSIFT, CSIFT and RGBSIFT.

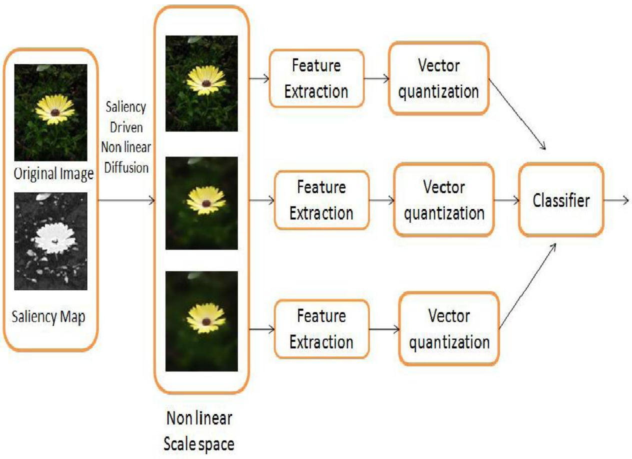

Fig. 3. The framework of the multi-scale representation for image classification

Image classification framework is shown in Fig. 3. Each image is represented by its multi-scale images. Then, for each scale t, scale invariant feature transform (SIFT) features, which are widely used to represent image regions, are extracted, and the bag-of-words model is used to generate a word frequency histogramh_t. The dissimilarity between images 1 and 2 at scale t is represented by the $\chi 2$ distance d (h_t^1,h_t^2) between histograms h_t^1 andh_t^2. The distances {d (h_t^1,h_t^2)}$t\epsilon T$ between images 1 and 2 obtained at different scales are combined to yield the final distance $d(h_1, h_2)$ between images 1 and 2 [1]:

$$d(h_1,h_2) = \frac{\sum_{t\in T} w_t d(h_t^1,h_t^2)}{\sum_{t\in T} w_t} \qquad (4)$$

where w_t is a weight for scale t, and T is a chosen set of scales. Weighted averaging, which is a general way for information fusion, is used to fuse information from different scales. By selecting appropriate weights, the distances between samples in the same class can be reduced and the distances between samples in different classes can be enlarged. Then, more accurate classification results can be obtained. This weighted averaging has been widely used in many applications such as for product recommendation and it was shown that the weighted averaging improves the prediction accuracy.

The saliency driven nonlinear multi-scale representation has several advantages. First, the nonlinear diffusion-based multi-scale space can preserve or enhance semantically important image structures at large scales. Second, this saliency driven multi-scale representation can deal with the background information no matter whether it is a context or noise, and then can be adapted to backgrounds which change over time. Finally, this saliency driven multi-scale representation can be easily combined with any existing image classification algorithms (e.g. bag-of-words).

VI. CONCLUSION

The saliency driven nonlinear multi-scale space preserves and even enhances important image local structures, such as lines and edges, at large scales. The diffusion process of nonlinear diffusion filters has capabilities of noise removal and edge preservation of images. The image classification results are improved by inclusion of background information in saliency driven multi-scale representation.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 2, February 2016

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

[3] J. Shotton, J. Winn, C. Rother, and A. Criminisi, "Textonboost: Joint appearance, shape and context modeling for multi-class object recognition and segmentation," in *Proc. Eur. Conf. Comput. Vis*., 2006, pp. 1–15.

[4] 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.

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

[6] Galleguillos, A. Rabinovich, and S. Belongie, "Object categorization using co-occurrence, location and appearance," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–8.

[7] 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

[8] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 7, pp. 629–639, Jul. 1990.

[9] P. Gehler and S. Nowozin, "On feature combination for multiclass object classification," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Oct. 2009, pp. 221–228.

[10] L. Marchesotti, C. Cifarelli, and G. Csurka, "A framework for visual saliency detection with applications to image thumb nailing," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Oct. 2009, pp. 2232–2239.

[11] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency-tuned salient region detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 1597–1604.

[12] X.-H. Shen and Y. Wu, "A unified approach to salient object detection via low rank matrix recovery," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jan. 2012, pp. 853–860.

[13] 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.

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

[15] 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.

[16] T. 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. 35–367, Feb. 2011.

[17] X. Hou and L. Zhang, "Saliency detection: A spectral residual approach," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–8.

[18] 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.

[19] 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.

.