Vol. 7(24), Apr. 2017, PP. 3390-3396

Contents list available at IJMEC

International Journal of Mechatronics, Electrical and Computer Technology (IJMEC)

Journal Homepage: www.aeuso.org

PISSN: 2411-6173 - EISSN: 2305-0543

Automatic Image Annotation Based on Dense Weighted Regional Graph

Masoumeh boorjandi[1] and zahra Rahmani[2]

[1]*Faculty of Computer, Aliabad Branch, Islamic Azad University, Aliabad katoul, Iran*

[2]*Faculty of Computer, Ramsar Branch, Islamic Azad University, Ramsar, Iran*

***Corresponding Author's E-mail:** m.boorjandi@yahoo.com*

Abstract

Automatic image annotation refers to create text labels in accordance with images' context automatically. Although, numerous studies have been conducted in this area for the past decade, existence of multiple labels and semantic gap between these labels and visual low-level features reduced its performance accuracy. In this paper, we suggested an annotation method, based on dense weighted regional graph.In this method, clustering areas was done by forming a dense regional graph of area classification based on strong fuzzy feature vector in images with great precision, as by weighting edges in the graph, less important areas are removed over time and thus semantic gap between low-level features of image and human interpretation of high-level concepts reduces much more. To evaluate the proposed method, COREL database, with 5,000 samples have been used. The results of the images in this database, show acceptable performance of the proposed method in compare to other methods.

***Keywords:** automatic annotation, dense weighted regional graph, segmentation, feature vector*

1. Introduction

Due to the growing use of digital technologies, image data generated and stored every day in large numbers, and using this data as text data has become commonplace. Hence the need to search for video data according to different demands increased. One of the traditional methods for image retrieval, is content based image retrieval [1-3]. But these systems are not able to understand the meaning. Also in this system, user must express your wishes with the visual properties of image expression, which in turn is difficult for users [1]. This is a formidable challenge in content-based image retrieval, called semantic gap. Semantic gap, the gap between low-level visual content of the image and human interpretation of it, is a high-level concept. [2]Methods including automatic image annotation that have been proposed in recent decades to reduce semantic gap.[4] computer in automatic image annotation is used to describe words that are suitable for producing images. In this case, to recover the image of a set of annotated images, using text request is also possible. Using text of the question is far easier than using a sample image or characteristics of the image. [1] Great deal of research has been done in the field of image annotation that can be grouped into three models: probabilistic models, model-based on categories and models based on the nearest neighborhood. [5]Most probabilistic models [9-6] joint probability estimate on the image content and keywords. Model-based categories [11-10], the image annotation to be an issue with the supervisor behave category. Models based on the nearest neighborhood, is one of the oldest, simplest and yet most efficient models in the category, the model is k-nearest neighbor. This model is growing, especially in the absence of training samples, efficiently. One of the methods in this area includes paper [12] cited. But in all of the presented methods in the field of automatic annotation there are two challenging problems: First, annotation techniques available generally are a feature of regional or global brand used to describe alone. But national and regional features focused on different aspects of an image

Article History:
IJMEC DOI: 649123/10.225129
Received Date: Sep. 11, 2016
Accepted Date: Jan. 15, 2017
Available Online: Mar. 03, 2017

complement each other, so combining them together to describe the images will be beneficial. Second, in all delivery methods based on characteristics of the area, only the direct distribution areas as areas used for image annotation via image segmentation based on region or the nature of objects are obtained, and the relationship between areas does not be paid attention. If given the link between areas, each of which represents a word or concept, we can help improve final margin words.in order to fix the problem on the basis of a weighted graph, in this article we offer an area dense. So that the proposed method uses the theory of Rough and Fuzzy feature vector for regions resulting from segmentation to effectively classified images and graphs do make up a dense area of both national and regional characteristics for use together to annotate and significantly enhanced accuracy. It also uses weighting the edges between vertices in the graph area proposed for communication between areas of images detected by the system. This article is organized as in Section 2, the image automatic annotation method based on weighted graph describes dense in an area, in section 3, the proposed algorithm simulation results and comparison with other methods in this area are provided. In the fourth overall conclusion of the article offered.

2. Image annotation based on dense weighted regional graph

In this section of the paper, the proposed method for annotating images is explained. The construction of an area of dense graph is as follow:

1: First collect a free enlarge dataset of annotated images, and classify values into different categories according to their annotation keywords.

2: Pictures related to any particular class are divided with accurate and effective method to pieces Rough separate charges.

3: For Category parts resulting from segmentation of images related to any particular group, first we used the method of k-means, so that same parts are classified in the same group. But because of some shortcomings in this method and more accurate grouping, again we will classify low-density lightweight piece set of groups of pictures (which are included large amounts of each class of images) by considering fuzzy feature vector associated with them, and compare it with the original image feature vector corresponding to the category of other classes, categories reclassification of the charges.

4: At this point we create dense an area graph so that we put fences and high density near the center of the image in a category and get the fuzzy feature vector associated with the label of their respective classes, images are annotated. For groups with low density and outlying parts by determining the fuzzy feature vector and determine its similarity to other video groups from other classes, in Group of parts that are most similar to them categorized and tagged with annotations are related to their respective floors.

5: Then each vertex of the graph, which represents dense cluster of image segments with high similarity which are connected to each other by weighted edge with other categories that include other pieces from the collection of images related to each specific circuit. We are considering the joint probability of more accurate values for weight gain so that by taking pictures of each class, the weaker groups object to be removed from any particular class.

After creating an area of dense graph of images related to the training data, in order to annotate new image, first we will segmentation and then due to dense parts closer to the center of the image and fuzzy characterization of their right to obtain the original image, and in the following, based on Weighted area of dense graph, Image annotation associated by taking in account similarity of other image parts which are connected with more weighted vertices of main group of images. In the graph area of dense, lightweight piece groups are annotated with names of collected images. Thus, the number of categories in training data should be large enough to cover the meanings of these groups and pieces of the image.

2.1. Image Segmentation in every particular class

Maximum image components can be extracted through various ways such as image segmentation [13], dense samples [14], and recognize a specified area [15] and etc. But most of these methods are very expensive in terms of time and calculation. Among the proposed methods, Rough set theory has the ability to cluster profitably data that comes from image analysis [16] as efficiently identify the edge that is one of the effective methods of segmentation, convergence time So we proposed this theory we use for segmentation of images. Rough set more detailed description about the image segmentation in the paper [17] can see. After extracting the piece of the picture, we produce dense piece band, we take action the k-means clustering pieces by the charges. However, in clustering k-means, the number of clusters manually set-up, and to produce pieces that are freely distributed is such as parts noise or background unstable and cluttered, so the results so not ideal. So after the initial clustering for producing dense, we have tried to use the feature vector effective, low-density parts with high density split into smaller pieces and fit them with similar criteria in groups with high similarity of our Categories. In this way text communication between groups are clearly to be identified, so that the results show high accuracy of the proposed method in the categories of different parts of images.

2.2. Produce dense areas and annotate

Most of the annotations provided in part due to lack of identification with high density of low-resolution images. Since the identification of areas with low density, similar to other image areas are densely populated in the group, we have tried to use the feature vector effective, low-density parts with high density split into smaller pieces and fit them with similar standards in groups with high similarity of our markets. For this purpose, we used a mask with size 60 × 60 on low-density areas. We use to identify groups of two characteristic color and edge video compression phase for more accurate the obtained fuzzy logic. More details on how to determine vector features can be found in Articles [18] and [19] pictures. Since fuzzy feature vector includes three properties in each area of the image are color, location and edge color, therefore, the proposed fuzzy similarity measure using data from the three phase characteristics for more accurate similarity or dissimilarity to appoint different images.

The following formula shows the proposed fuzzy similarity measure:

$$sim(q,t) = (w_c \times (\frac{\sum_{j=1}^{10} \min(H_c^{q_i}(j), H_c^{t_i}(j))}{\min(\sum_{j=1}^{10} H_c^{q_i}(j), \sum_{j=1}^{10} H_c^{t_i}(j))} \times \sum_{j=1}^{10} (1 - (\sqrt{\frac{(\overline{x_{q_j}} - \overline{x_{t_j}})^2 + (\overline{y_{q_j}} - \overline{y_{t_j}})^2}{2}})))) +$$

$$(\sum_{i=1}^{4} w_{ei} \times \frac{\sum_{j=1}^{4} \min(H_e^{q_i}(j), H_e^{t_i}(j))}{\min(\sum_{j=1}^{4} H_e^{q_i}(j), \sum_{j=1}^{4} H_e^{t_i}(j))})$$

$H_c^{q_i}(j)$: J-th column of fuzzy color histogram, i-th class query image

$\overline{x_{q_j}}, \overline{y_{q_j}}$: Average Location j-th column of the query image color histogram

W_c: the importance of class background

$H_e^{q_i}(j)$: J-th bar of the histogram of the i-th edge to edge located

W_{ei}: the importance of the i-th to edge

As the numerical value for each component fuzzy weight in the literature [13] and [14] were determined as the following:

Very large = 1, large = 0.8, medium = 0.55, small = 0.3, very small = 0.1

This method to a large extent is resistant to the problems that the majority of edge detection methods are faced, such as sensitivity to noise and with thick lines. So low-density areas are divided into several regions with higher density and tagged with annotations are the most similar. In this way, for different images in the database area of dense graph is created.

2.3. Annotation

After you remove the piece of background object slice group dense pieces, we attempted to annotate them. To do this, we first collect band name as a label. Since a large number of large-scale collections are on the floor, we assume that most of groups extracted can be annotated with the name of the group. We do annotation groups based on two criteria:

(1) Visually similar groups must be annotated with similar tags.

(2) Groups that are distinct and belong to a particular class, are likely annotate with this class. This idea is illustrated in Figure 1.

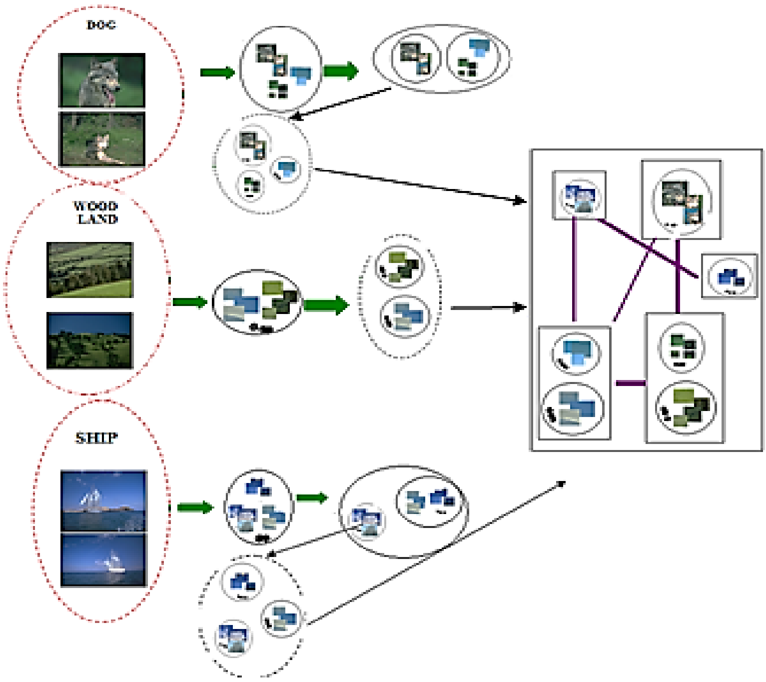

Fig. 1. piece for the dog, after compression zones

For example, in Figure 1 piece for the dog, after compression zones were discovered containing "dog", "woods", "sky" and so on, that piece band dogs due to high density and proximity to the center of the image the dog is the label class that is annotated. The following groups are fences "sky" and "woodlands" by comparing their characteristics with the vector of the main themes in classes other picture, as a new group piece which most closely resembles groups may be categorized with them And then the edge between the two vertices in the graph occurred that indicate the presence of similar images is common in these group. For example in Figure 1 edge between vertices of dogs and woodlands parts thicker than the edge between the vertices of the parts is dog heaven and this occurred due to high joint between the two parts of the head. To determine the edge weight between vertices, following formula is used:

$$W=k/n \quad (1)$$

W is Weight between two vertices in the graph, an area corresponding to each of the respective condensing And K, is the number of subscribers took place labels on two related helm annotation of images that class and n, is the number of images of respective class. So common they both took the helm that is more likely that more weight be given indicator is thicker than the edges.

image						
Correct lable	Cat, tiger, grass	Flower, leaf, grass, cloud, sky	Train, bus, grass, tree, sky	Eagle, bird, sky, cloud, mountain	Horse, tree, grass, sky, fance	Man, sea, beach, sky
Proposed model	Tiger, grass, tree	Flower, grass, leaf, tree, sea	Train, sky, grass	Eagle, sea, sky, tree	Horse, bush, fance	Man, sea, cloud, beach, sky

Fig. 2. The words predicted by the proposed method for multi-image annotation database of Corel

Then for more precise evaluation of effectiveness of the proposed model, the average precision, recall and F1 parameters have been calculated per word.

Precision $(v_i) = N_C / N_S$ **(3)**

Recall $(v_i) = N_C / N_R$ **(4)**

$F_1 = 2 \times \text{Precision} \times \text{Recall} / \text{precision} + \text{Recall}$ **(5)**

Accuracy, is the ratio of NC, the ratio of the number of images in the test phase to the NS, the number of images than in the test phase. Calling the ratio of NC, the ratio of the number of images in the test phase NR, number of images in the database for each v i is the word. Annotations to assess the quality, accuracy and calling for each and every word database (v i) are calculated. Table 1, shows mean precision, recall and F1 for the proposed methodology and the appropriate annotation recently presented three methods (IAGA -2014 [20], Feature fusion and semantic similarity-2014 [5], MLRank -2013 [21]). More areas have higher density than other areas in each image by weighting the edges of the graph. Less important areas are removed so that it causes a system closer to annotate people, in our proposed method is compared to other methods.

Table 1. Comparison of precision, recall and F 1 of the proposed method and other methods

F1	recalling	accuracy	Model name
0.34	0.31	0.34	IAGA-2014[12]
0.31	0.32	0.3	Feature fusion and semantic similarity-2014[2]
0.29	0.28	0.3	MLRank-2013[24]
0.36	0.33	0.39	Proposed method

Figure 3 percent identify 12 words sample to show the proposed method and model IAGA. The results of this graph shows that, using accurate segmentation according to the theory of Rough and strong feature vectors that lead to the formation of dense graph and highlight an area densely populated areas in each image, is the more precise identification of concepts such as stairs, flowers and lawns aircraft per cent lower compared to IAGA that have been identified. In fact, one of the main problems in annotation methods including IAGA non-designated areas in each image is important, we provided an area graph density to accurately classify areas according to each class video By weight of feature vectors appropriate and accurate to solve them.

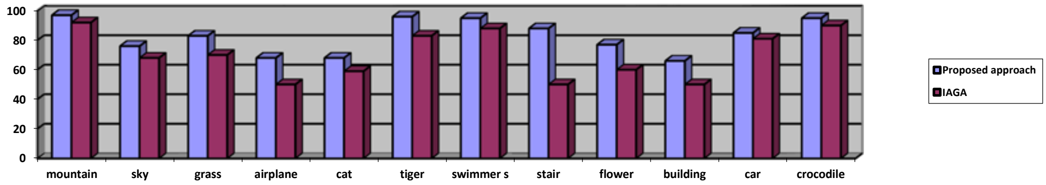

Fig 3. Percentage identify 12 words sample to the suggested model and model IAGA

Figure 4, the results retrieved for a query on the database shows Corel. For any query shows five images with the highest similarity. According to the results in the figure below, we conclude that the graph based image retrieval dense area for suggestions, will guide us to recover the database with accuracy and efficiency. Therefore In the proposed method, we were able to accurately and more efficient than the methods suggested in the annotation to achieve by Categorizin more accurate images with different areas of dense graph and also remove less important areas in each class by weighting the edges.

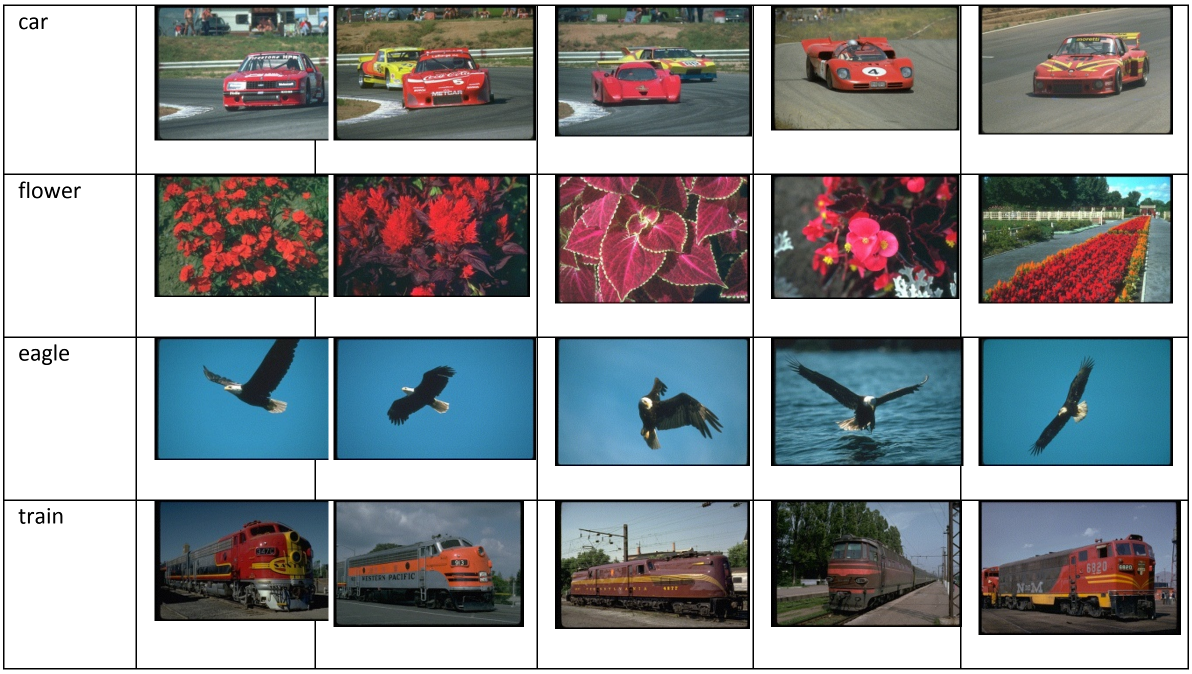

Fig 4. Results semantic image retrieval database corel. Each row of five top result query semantic meaning in accordance with the left-most column shows.

Conclusion

In this paper, a graph-based method for automatic image annotation densely an area is provided. In most of the methods presented in the context annotations there are two basic challenging problems including Lack of integration of national and regional characteristics for each of the images and lack of attention to the relationship between different areas in Pictures. In this article we formed an area graph, the relationship between different areas in the image are considered and weighted the edges in the graph, and done compression of areas so that Prominent areas on each floor image in low-density areas considered less important and have to be removed. Also by Using strong fuzzy feature vectors based on color and edge features for the considered areas we have done Aggregation of national and regional action features lightweight and have improved Annotations practice considerably. At the end we have implemented the proposed approach on Corel database. The results provided on the database show acceptable performance of the proposed method compared with other methods in this field.

References

[1] T. Pavlidis "Limitations of Content-based Image Retrieval", invited plenary talk at the 19 th International Conference on Pattern Recognition, Tampa, Florida, Dec. 8-11, 2008.

[2] R. Datta, D. Joshi, J. Li, J.Z. Wang, Image Retrieval: Ideas, Influences, and Trends of The New Age, ACM Comput. Surv, 40 (65), 2008.

[3] A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, R. Jain, Content-based image retrieval at the end of the early years, IEEE Trans. Pattern Anal. Mach.Intell, 22 (12) ,2000,p. 1349-1380.

[4] Jiayu Tang ,.Automatic Image Annotation and Object Detection, A thesis for the degree of Doctor of philosophy , University Of Southampton, 2008.

[5] X. Zhang and C. Liu, “Image annotation based on feature fusion and semantic similarity,” Neurocomputing , 2014.

[6] P. Duygulu, K. Barnard, J. F. de Freitas, and D. A. Forsyth, “Object recognition as machine translation: Learning a lexicon for a fixed image vocabulary,” in Computer Vision—ECCV 2002, Springer, pp. 97–112, 2002.

[7] J. Jeon, V. Lavrenko, and R. Manmatha, “Automatic image annotation and retrieval using cross-media relevance models,” in Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval, pp. 119–126, 2003.

[8] V. Lavrenko, R. Manmatha, and J. Jeon, “A model for learning the semantics of pictures,” in Advances in neural information processing systems, 2003.

[9] S. L. Feng, R. Manmatha, and V. Lavrenko, “Multiple bernoulli relevance models for image and video annotation,” in Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on, vol. 2, pp. II–1002, 2004.

[10] J. Li, J. Wang, Automatic linguistic indexing of pictures by a statistical modeling approach, PAMI 25 (19) ,2003,p. 1075-1088.

[11] E. Chang, K. Goh, G. Sychay, G. Wu, CBSA: CBSA: content-based soft annotation for multimodal image retrieval using Bayes point machines, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2003,pp. 26--38.

[12] J. Verbeek, M. Guillaumin, T. Mensink, and C. Schmid, “Image annotation with tagprop on the MIRFLICKR set,” in Proceedings of the international conference on Multimedia information retrieval, pp. 537–546, 2010.

[13] P. Felzenszwalb, D. Huttenlocher, Efficient graph-based image segmentation,IJCV 59 (2) (2004).

[14] S. Lazebnik, C. Schmid, J. Ponce, Beyond bags of features: spatial pyramidmatching for recognizing natural scene categories, CVPR (2006).

[15] J. Matas, O. Chum, M. Urba, T. Pajdla, Robust wide baseline stereo frommaximally stable extremal regions, BMVC (2002).

[16] B. Frey, D. Dueck, Clustering by passing messages between data points, Science315 (February) (2007) 972–976.

[17] T. Chua, J. Tang, R. Hong, H. Li, Z. Luo, Y. Zheng, NUS-WIDE: a real-world web image database from National University of Singapore, CIVR (2009).

[18] S. Bahrami, M. Saniee Abadeh, "Automatic image annotation using an evolutionary algorithm (IAGA)." In Telecommunications (IST), 2014 7th International Symposium on, pp. 320-325. IEEE, 2014.

[19] Z. Li, J. Liu, C. Xu, and H. Lu, “MLRank: Multi-correlation Learning to Rank for image annotation,” Pattern Recognit., vol. 46, no. 10, pp. 2700–2710, 2013.

[20] S. Bahrami, M. Saniee Abadeh, "Automatic image annotation using an evolutionary algorithm (IAGA)." In Telecommunications (IST), 2014 7th International Symposium on, pp. 320-325. IEEE, 2014. 23. Z. Li, J. Liu, C. Xu, and H. Lu, “MLRank: Multi-correlation Learning to Rank for image annotation,” Pattern Recognit., vol. 46, no. 10, pp. 2700–2710, 2013.