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

International Journal of Innovative Research in Computer and Communication Engineering
Vol. 1, Issue 5, July 2013

"A Novel Approach For Automatic Image Annotation Using Color Saliency"

Shalini K.Kharkate[1], Prof.Nitin J.Janwe[2]

Dept. Of Computer Technology, Rajiv Gandhi College of Engg, Research and Technology,Chandrapur,M.S., India[1]

Professor, Dept. Of Computer Technology, Rajiv Gandhi College of Engg, Research and Technology, Chandrapur, M.S., India[2]

ABSTRACT: Image annotation task consists to assign a set of semantic tags or labels to a novel image based on some models learned from certain training data. Automatically assigning keywords to images is of great interest as it allows one to index, retrieve, and understand large collections of image data. Many techniques have been proposed for image annotation in the last decade that gives reasonable performance on standard datasets. This paper proposed algorithm to annotate image by comparing test image feature vector with feature matrix of training data sets and similar and dissimilar image pairs. Ranking technique is used for transferring the key word from similar image pair to the test image by counting local frequency of keywords. Performance is evaluated using precision and recalle proposed method.

Keywords**:** Automatic image annotation, feature vector, feature matrix, COREL 5K, precision and recall.

I. INTRODUCTION

Automatic image annotation is a challenging problem in the field of image retrieval. It can be used to facilitate semantic search in large image databases. However, retrieval performance of the existing annotation schemes is far from the users' expectation. For years Automated image annotation has been an active and challenging research topic in computer vision and pattern recognition several techniques have been propose[10,11,12,13,14,15,16,17,18,19] Automated image annotation is essential to make huge unlabeled digital photos index able by existing text based indexing and search solutions. In general, an image annotation task consists to assign a set of semantic tags or labels to a novel image based on some models learned from certain training data.

A large number of image search engines mainly employ the surrounding texts around the images and the image names to index the images. However, this limits the capability of the search engines in retrieving the semantically related images using a given query. On the other hand, although the current state-of-the-art in content-based image retrieval is progressing, it has not yet succeeded in bridging the semantic gap between human concepts, e.g., keyword-based queries, and low-level visual features that are extracted from the images .Hence, it has become an urgent need for developing novel and effective paradigms that go beyond these conventional approaches or retrieval models.

In the image retrieval problem, given an input image, the algorithm needs to discover similar and relevant images.Images are annotated to simply access to them by using metadata that being added to images in order to allow more effective searches. If the images are described by textual information, then text search technique can be used to perform images searches [20]. Many researchers have proposed various techniques in attempting to bridge the well known semantic gap. In [7,21] segmental approaches images are segmented into region and relation is find out between image region and word. Segmentation process is fragile and erroneous which make annotation process unreliable. Holistic approach [27] is to estimate the probabilities of images queries that then will be ranked according to their probabilities. No segmentation in holistic approach makes fast feature extraction but no direct correspondence between image region and word. Many of them realize another problem which is dependency on the training dataset to learn the models [21]. Image annotation surveys have been reviewed by many researchers according to the demanding the needs for annotating images. The graph model based image annotation methods' time complexity and space complexity are always high, and it is difficult to apply it directly in real world image annotation Jiayu [22] has classified image annotation approaches into statistical approaches, vector-space related approaches and classification approaches. Probabilistitic approaches have computational overhead. Classification model performance is superior to probabilistic. However classification approach cannot be extended to unsupervised learning which is inherently supervised. Each model has its own advantages and disadvantages.

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

International Journal of Innovative Research in Computer and Communication Engineering
Vol. 1, Issue 5, July 2013

Our method includes training and testing procedures training part selects low-level features (e.g., bins in the feature histogram) using saliency detection technique.
These priors improve the model's robustness to noise. Testing a part automatically annotates input images by transferring keywords from similar images.

Sun ,sky, water, tree
Figure 1: . An annotation example from Core5k.

II. RELATED WORK

Recent techniques for AIA based image retrieval generally divided into two types of approaches, the probabilistic modeling methods and the classification methods. the probabilistic modeling methods and the classification methods. The probabilistic modeling methods aim to develop a relevance model to represent the correlation or joint probability distribution between images and keywords [2]. [1] propose to treat image annotation as a process of machine translation. They introduced a Translation Model (TM) based on statistics. They used the method to translate a visual vocabulary into keywords. The other typical method is the Latent Dirichlet Allocation model [3]. However, in the two models above, the probability distributions may not reflect the actual distributions. The process of the parameter estimation is also complex and expensive. [4] proposed Cross Media Relevance Model (CMRM) where the vision information of each image was denoted as blob set which is to manifest the semantic information of image. However, blob set in CMRM was erected based on discrete region clustering which produced a loss of vision features so that the annotation results were too perfect. In order to compensate for this problem, a Continuous-space Relevance Model (CRM) was proposed in [5]. Furthermore, in [6]Multiple-Bernoulli Relevance Model was proposed to improve CMRM and CRM. These methods employ a non-parametric method to estimate a Gaussian distribution. Compared with other discrete models, these methods can evidently improve annotation accuracy. Tianxia Gong, Shimiao Li, Chew Lim Tan[7]roposed a framework of using language models to represent the word-to-word relation utilizing probabilistic models.

On the other hand, the discriminative model trains a separate classifier from visual features for each tag. These classifiers are used to predict particular tags for test image samples [8], [9]. Similarly, we can also train a regression model (regression coefficients) to predict tags for test images, taking features as predictors (input variables) and tags as responses (output labels). In image annotation and retrieval, SVM is a widely used machine learning method. SVM can generate a hyperplane to separate two data sets of features and provide good generalization.

A. GLOBAL FEATURE EXTRACTION

The global feature representation techniques have been extensively studied in image processing and content-based image retrieval. In contrast to the local feature-based approaches the global feature is very efficient in computation and storage due to its compact representation. A wide variety of global feature extraction techniques have been proposed in the past decade. In this paper, we extract features using residual spectral method.

Human eye is perceptually more sensitive to certain colors and intensities and objects with such features are considered more salient. Salient regions are most important point in image which attracts greater attention by visual system than other part of the image. These regions has distinctive features when compared with others in image. Eg. a polar bear is salient on dark rocks, but almost invisible in snow.

Recently, several saliency approaches came up that are based on computational and mathematical ideas and usually less biologically motivated. These approaches range from the computation of entropy[23] over determining features that best discriminate between a target and a null hypothesis to learning the optimal feature combination with machine learning techniques.

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

International Journal of Innovative Research in Computer and Communication Engineering
Vol. 1, Issue 5, July 2013

In[24] proposed saliency detection method spectral residual. Spectral residual is the difference between original log spectrum and its mean-filtered version. The saliency map is obtained by applying inverse Fourier transform to the spectral residual. We compute the color histogram of saliency regions for the color space RGB..

III.PROPOSED FRAMWORK

A training set S consisting of N images with n feature vectors. n feature vectors forms a feature matrix and A pair of similar and dissimilar images(L). The main purpose of this paper is to investigate the feature selection properties in the image annotation task. This image pair setting helps us to create a feature matrix that contains the same groups of features. Thus, we can directly do feature analysis on this matrix within the same framework.

Calculate the weight assign to each feature vector by using feature matrix and L, which is final step of training stage. Weight vector is used to find relevancy of keyword to the image. Sufficient training is necessary to have correct annotations to the image in testing stage.

Feature vector of input image is compared with feature matrix of the training images. Based on the weights calculated, most similar images are find out from L. Keywords from L are getting assigned to test image which are annotations.

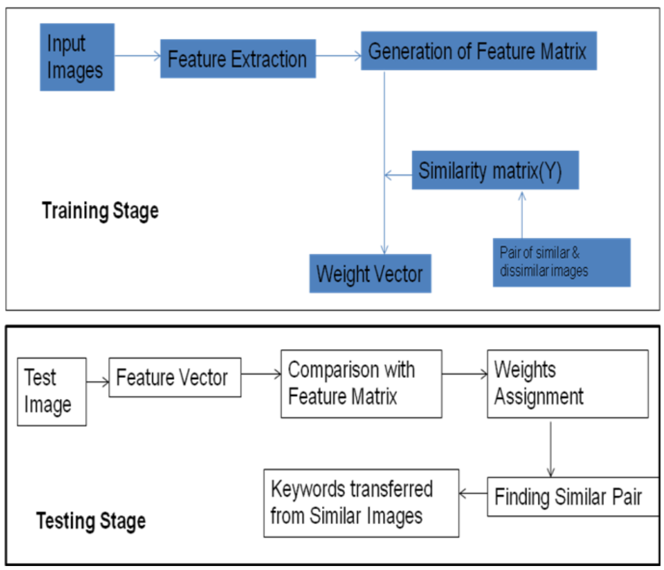

Figure 2: Proposed Architecture.

A. WEIGHT VECTOR CALCULATION

Here weight of each feature vector is calculated by using feature matrix obtained using SR method and set of similar and dissimilar image pair. In this setting we consider any pair of images that share enough keywords to be positive training samples and any pair with no keywords in common to be negative example. In this work we obtained training samples from the designated training set of the Corel5K dataset. Image paires that had at least four common keywords were treated as positive sample for training and those with no common keywords were used as negative samples[25].

Weighted least square is an efficient method that makes good use of small data sets.It also shares the ability to provide different types of easily interpretable statistical intervals for estimation , prediction

The most popular loss function to calculate w in this regression problem is the least square estimate, which is also named as the minimizer of the residual sum of squared errors and is given as

$$w = \arg\min_{w \in \mathbb{R}p} \|Xw - Y\|^2_2$$

.

This weight is used for testing stage in image annotation task.

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

International Journal of Innovative Research in Computer and Communication Engineering
Vol. 1, Issue 5, July 2013

Input : The Weight vector w
Feature vector of test images t
Feature matrix of training images F
Pair of Similar and Dissimilar images
Keywords

Algorithm:

for all rows in F
Compute Similarity array as
$s_i = (t_i - f_i) * w_i$
End for
Compare with Pair of Similar and Dissimilar images
Compute the local frequency of keywords and rank it.
Take predefine number of values from s
Transfer keyword according to their local frequency

Output:
Annotated images with defined keywords

Figure 3: The proposed Algorithm

IV. EXPERIMENTAL RESULTS

A. COREL DRAW DATASET

Photo Collection are used containing 5000 images each image is annotated with 3 to 4 keywords. There are 374 distinct words in the dataset. Whole datset is taken for training. To test image first feature vector of test image calculated. It is then compare with feature matrix to find the similarity vector. It is then applied with weight vector to find 5 most similar image pairs, from that image pairs keywords are transferred to testing image by using ranking.

B PERFORMANCE EVALUATION.

In most of the literature the performance of annotation system is calculated by using prescision and recall. Precision and Recall values in annotation system are evaluated for each word and the mean of all words are consider as the performance of the system. Accordingly,

$$\text{Precision} = \frac{\text{no of correct annoted label}}{\text{total annoted label}}$$

$$\text{recall} = \frac{\text{no of correct annoted label}}{\text{total label in testset}}$$

For comparing system using only single value F-score is good choice.

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

International Journal of Innovative Research in Computer and Communication Engineering
Vol. 1, Issue 5, July 2013

$$\text{F-score} = \frac{2 \cdot \text{precision} \cdot \text{recall}}{(\text{precision} + \text{recall})}$$

10 images are taken for testing purpose. Average of its F score is measured by above mentioned formula.

Predicted: 'beach' *'palm'* 'people' 'tree' *'city'*
Human annotation
Tree, beach, people

.

Predicted: 'clouds' 'sky' 'sun' 'tree' *'mountain'*
Human annotation
'clouds' 'sky' 'sun' 'tree' 'mountain'

Predicted: field' 'foals' 'horses' 'mare' *'flower'*

Human annotation
Field, foals. horse ,mare

Predicted: 'jet' 'plane' 'smoke' 'sky' *'f-16'*

Human annotation
Sky ,jet, plane, smoke

Figure4. Predicted keywords using PROPOSED METHOD versus the human annotations for a sampling of images in the Corel5K dataset (using all 374 keywords).

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

International Journal of Innovative Research in Computer and Communication Engineering
Vol. 1, Issue 5, July 2013

Model	F-score
Co-occurrence	0.02
Translation Model	0.05
CMRM	0.09
Max. Entropy	0.10
CRM	0.17
CRM-Rectangles	0.20
MBRM	0.22
CSD-prop	0.24
HANOLISTIC	0.28
Proposed Method	0.44

Figure 5 . Comparison of Proposed method with other systems in literature

V. CONCLUSION AND FUTURE WORK

In this study we proposed a framework and algorithm for automatic image annotation problem. We took a holistic approach and saliency detection technique .We compared the obtained results with other studies in the literature.

Compared with other existing methods, it shows higher performance in the image annotation task .In future we add texture feature for better result along with color and compare with only color feature.

REFERENCES

[1] Duygulu, P., Barnard, K., Freitas, J., Forsyth, D.A., Object recognition as machine translation: learning a lexicon for a fixed image vocabulary. In:Proceedings of 7th European Conference on Computer Vision (ECCV'02), Copenhagen, Denmark, pp. 97–112.

[2] Zhu, S., Liu, Y.." Semi-supervised learning model based efficient Image annotation. " 2009,IEEE Signal Process. Lett. 16 (11), 989–992.

[3]. Blei, D.M., Jordan, M.I "Modeling annotated data." In: Proceedings of 26th Annual International ACM SIGIR Conference e on Research and Development
in Information Retrieval (SIGIR'03), Toronto, Canada, pp. 127–134

[4] Jeon J, Lavrenko V, Manmatha R.," Automatic image annotation and retrieval using cross-media relevance models". In: Proc. of Int. ACM SIGIR Conf. on
Research and Development in Information Retrieval (ACM SIGIR'03), Toronto, Canada, Jul. 2003: 119-126.

[5] Lavrenko V, Manmatha R, Jeon J." A model for learning the semantics of pictures". In: Proc. Of Advances in Neural Information Processing Systems
(NIPS'03), 2003.

[6] Feng S, Manmatha R , Lavrenko V. "Multiple bernoulli relevance models for image and video annotation." In: Proc. of IEEE Int. Conf. on Computer Vision and Pattern Recognition (CVPR'04), Washington DC, USA, Jun. 2004: 1002-1009.

[7] Tianxia Gong, Shimiao Li, Chew Lim Tan,"A Semantic Similarity Language Model to Improve Automatic image annotation" 22nd International Conference on Tools with Artificial Intelligence, 2010

[8] D. Grangier and S. Bengio, "A discriminative kernel-based approachto rank images from text queries," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 8, pp. 1371–1384, Aug. 2008.

[9] Y. Chen, J. Wang, and D. Geman, "Image categorization by learningand reasoning with regions," *J. Mach. Learn. Res.*, vol. 5, pp. 913–939, Dec. 2004.

[10] F Tianxia Gong, Shimiao Li, Chew Lim Tan,,"Semantic Similarity Language Model to Improve Automatic image annotation" 22nd International Conference on Tools with Artificial Intelligence, 2010

[11] Rami albatal, Philippe mulhem, Yues chiaramella, "A new ROI grouping schema for automatic image annotation"

[12] J Jing Liua,, Mingjing Lib, Qingshan Liua, Hanqing Lua, Songde Ma "Image annotation via graph learning", Pattern Recognition 42 (2009) 218 – 228

[13] Yunhee Shin, Youngrae Kim, Eun Yi Kim," Automatic textile image annotation by predicting emotional conceptsfrom visual features". Image and Vision Computing 28 (2010)

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

International Journal of Innovative Research in Computer and Communication Engineering
Vol. 1, Issue 5, July 2013

[14] Y. Alp Aslandogan and Clement T. Yu, Senior Member, IEEE, Techniques and Systems for Image and Video Retrieval VOL. 11, NO. 1, JANUARY/FEBRUARY 1999
[15] Lei Ye, Philip Ogunbona and Jianqiang Wang" Image Content Annotation Based on Visual Features" Proceedings of the Eighth IEEE International Symposium on Multimedia (ISM'06).
[16] D Dongjian He & Yu Zheng, Shirui Pan, Jinglei Tang,"Ensemble Of Multiple Descriptors For Automatic Image Annotation" -2010 3rd International Congress on Image and Signal Processing
[17] Y Hamid ansari, Mansour Jamzad "Large-Scale Image Annotationusing Prototype-based Models"7th International Symposium on Image and Signal Processing and Analysis (ISPA 2011) September 4-6, 2011, Dubrovnik, Croatia.
[18] Golnaz Abdollahian, Murat Birinci †, Fernando Diaz-de-Maria ‡, Moncef Gabbouj †, Edward J. Delp "A region-dependent image matching methodfor image and video annotation" 2011 IEEE
[19] Yinjie Lei Wilson Wong Mohammed Bennamoun Wei Liu "Integrating Visual Classifier Ensemble with TermExtraction for Automatic Image Annotation",2011
[20] Hua Wang, Heng Huang and Chris Ding, "Image Annotation Using Bi-Relational Graph of Images and Semantic Labels" pages 126–139, 2011
[21] Ran Li, YaFei Zhang, Zining Lu, Jianjiang Lu, Yulong Tian-"Technique of Image Retrieval based on Multi-label Image Annotation" ,2010 Second International Conference on MultiMedia and Information Technology
[22] T. Jiayu, "Automatic Image Annotation and Object Detection," PhD thesis, University of Southampton, United Kingdom, May 2008.
[23]T. Kadir, A. Zisserman, and M. Brady. "An affine invariant salient region detector." In Proc. of ECCV, 2004
[24]Xiaodi Hou and Liqing Zhang "Saliency Detection: A Spectral Residual Approach"
[25]Ameesh Makadia1, Vladimir Pavlovic2, and Sanjiv Kumar "A New Baseline for Image Annotation"
[26] F. Tsai and C. Hung, "Automatically Annotating Images with Keywords: A Review of Image Annotation Systems," Recent Patents on Computer Science,vol1,pp55-68,Jan.,2008