International Journal of Latest Engineering and Management Research (IJLEMR)
ISSN: 2455-4847
www.ijlemr.com || Volume 02 - Issue 09 || September 2017 || PP. 41-49

Heuristic Model for Web Image Annotation: A Statistical Approach

Khasim Syed[1], Dr. S V N Srinivasu[2]
[1]*Research Scholar, Department of CSE, Rayalaseema University, Kurnool, (AP) – India.*
[2]*Professor & Principal Department of CSE, IITS, Markapur, (AP) – India.*

Abstract: Web image annotation plays a critical role in modern keyword-based image retrieval systems. It is an essential tool for handling web scale images for retrieval, index and other management tasks. The goal of image annotation is to find suitable annotation words to represent the visual content of an untagged or noisily tagged image. Differing from the established lexicon or training data based keyword correlation assessment; we proposed a heuristic model for web image annotation. First, it selects the majority relevant features from all the data points by using a heuristic model. Then, it can uncover the shared subspace of original features, which is beneficial for multi-label learning. To further improve web image annotation performance, we uses statistical model to estimate the semantics of the web image from the corresponding associated text. We integrate the proposed methods into our web image annotation framework and conduct experiments on a genuine Web image data set. The statistical results demonstrating its feasibility and can improve the annotation performance significantly of being applied to web image annotation.
Keywords: Image Retrieval, Heuristic model, tagged web image, Statistical Analysis.

I. Introduction

Due to the explosive development of digital technologies, yet rising visual data are created and stored. Nowadays, visual data are as general as textual data. There is a burning need of effective and efficient tool to discover visual information on demand. A large amount of research has been carried out on image retrieval (IR) in the last two decades. In general, IR research efforts can be divided into three types of approaches. The first approach is the traditional text based annotation. In this approach, images are annotated manually by humans and images are then retrieved in the same way as text documents. However, it is impractical to annotate a huge amount of images manually. Furthermore, human annotations are usually too subjective and ambiguous. The second type of approach focuses on content based image retrieval (CBIR), where images are automatically indexed and retrieved with low level content features like color, shape and texture. However, recent research has shown that there is a significant gap between the low level content features and semantic concepts used by humans to interpret images. In addition, it is not practical for common users to use a CBIR system because users are necessary to offer query images. The third method of image retrieval is the automatic image annotation (AIA) so that images can be retrieved in the similar way as text documents. The key idea of AIA methods is to automatically study semantic concept models from large number of image samples, and use the concept models to label new images. Once images are annotated with semantic labels, images can be retrieved by keywords, which is similar to text document retrieval. The key characteristic of AIA is that it offers keyword searching based on image content it is mainly used for visual information management and can be applied in a variety of domains such as entertainment, commerce, education, biomedicine, military, web image classification and search, etc.

In particular, image annotation can aid in image retrieval since annotated keywords greatly narrow the semantic gap between low-level features and high-level semantics. Automatic image annotation is a challenging task due to various imaging conditions, complex and hard-to-describe objects, a highly textured background, and occlusions. In general, most approaches use learning-based techniques to train manually categorized images and test the uncategorized images based on the training results. Because the selected training images are usually very limited and cannot represent all the aspects of real life, automatic annotation may not achieve high accuracy using the current computer vision and image processing technologies.

Now a day, image annotation has become one of the central parts of research topics in computer vision and multimedia areas due to its potential impact on both image understanding and semantic based image retrieval. In general, it indicate the task of learning statistical models from a training set of pre-annotated images in order to produce annotations for unseen images using visual feature extracting technology. The target of image annotation is to discover appropriate annotation words to show the visual content of an untagged or noisily tagged image. In other words, the correlation between images and annotation words is a vital problem in sight of technical solutions. However, the results of the state-of-the-art image annotation techniques are often unsatisfactory. Therefore, it is require refining the imprecise annotations obtained by existing annotation techniques. In this paper, we proposed a heuristic model web image annotation.

The remainder of the paper is organized as follows. Section 2 briefly reviews the related work and defines the problem statement. Section 3 describes the generic framework of our proposed annotation system. Section 4 illustrates the proposed framework. Section 5 concludes with a brief discussion of our approach.

II. Problem Statement

Image annotation can be viewed as a classification task. It aims to correlate concept labels with specific images by classifying images to different classes. The ultimate goal is that the predicted labels via annotation algorithms can precisely reflect the real semantic contents of images. Content based image retrieval (CBIR) systems retrieve images linked to the query image (QI) from enormous databases. The feature sets extracted by the present CBIR systems are limited. This limits the systems effectiveness. The correlation between keywords has been exploited to improve Automatic Image Annotation (AIA). It can automatically annotate images with semantic labels to significantly facilitate image retrieval and organization. Classical web image annotation techniques often calculate approximately specific label relevance to image, and ignore the relevance of the assigned label set as a whole.

Chaoran Cui et al. proposed modern keyword-based image retrieval systems. For this process, the nearest-neighbor–based scheme works in two steps: first, it discovers the most similar neighbors of a novel image from the set of labeled images; then, it propagates the keywords associated with the neighbors to the new image. Wang et al. proposed a search-based annotation system – AnnoSearch. This system requires an initial keyword as a seed to speed up the search by leverage text-based search technologies. However, the initial keyword might not always be available in real environment. In the case there is no initial keyword available for the query image, the system will encounter a serious efficiency problem. Furthermore, the system tends to be biased by the quality of initial keywords. If the initial keywords are not accurate, the annotation performance will degenerate. Duygulu et al. proposed a translation model to label images at region level under the assumption that each blob in a visual vocabulary can be interpreted by certain word in a dictionary. The latent Dirichlet allocation model and the hierarchical aspect model were investigated in.

Wu et al. focused on learning an optimal distance metric by exploring implicit side information associated with web images, such as surrounding text and existing tags. Guillaumin et al. proposed the TagProp model, which weighted different base distances by maximizing the likelihood of the annotations of training images. To boost the recall of rare keywords, they also introduced a word-specific modulation of the weighted neighbor prediction Jeon et al. proposed cross-media relevance model to predict the probability of generating a word given the blobs in an image. In the scenario that each word is treated as a distinct class, image annotation can be viewed as multi-class classification problem. Yang et al. use multiple-instance learning to identify particular keywords from image data using labeled bags of examples. The basic intuition is to learn the most representative image region for a given keyword.

Nonetheless, the web image resources are innumerable so it is infeasible to annotate all of them manually. Hence, automatic image annotation becomes a necessary tool for handling web scale images for retrieval, index and other management processes. Thanks to the continuous effort made by researchers, we have witnessed enormous advance in automatic annotation for web images. However, the performance of automatic image annotation has yet to be satisfactory, thus requiring extra research work in this domain. Inspired by the current advanced methods of feature selection and shared feature subspace uncovering, we propose a novel framework to extract the most discriminating features to boost the image annotation performance.

III. Generic Framework

Automatically assigning keywords to images is of great interest as it allows one to index, retrieve, and understand huge collections of image data. Several methods have been proposed for image annotation in the last decade that provides reasonable performance on standard datasets. Image annotation can be showed as a classification task. It aims to correlate concept labels with specific images by classifying images to dissimilar classes. The final goal is that the predicted labels via annotation algorithms can accurately reflect the authentic semantic contents of images. Nonetheless, the web image sources are countless so it is infeasible to annotate all of them manually. Hence, image annotation becomes an essential tool for handling web scale images for retrieval, index and other management tasks.

The Overview of Web Image Annotation Framework

For a given training set L_{train}, each labeled image $S \in L_{train}$ is demoted by $S = \{U, V, T\}$,

Where

U is a binary annotation keyword vector indicating whether a keyword is the annotation of S

V is a set of region based visual features of J; and

$T = \{T_1, T_2, \ldots, T_n\}$ is a set of the types of associated texts.

Given a new image I, the annotation keywords set after i−1 iterative is denoted as AN_{i-1} ($i = 1, \ldots, k$, AN_0 = null, where k is the size of the annotation keywords set). In the i^{th} iterative, the probability of keyword w to be annotated for I is:

$$P(u|AN_{i-1}) = \frac{P(u|I)P(u|AN_{i-1})}{P(u)} = \frac{P(u|I_v,\ I_T)P(u|AN_{i-1})}{P(u)} \quad \ldots\ldots (1)$$

Where I_V and I_T is the visual and textual feature of image I respectively. Assuming that P (w) is uniformly distributed, and I_V and I_T are independent, we have:

$$w_i^* = \mathrm{argmax}_w\ P(u|I)P(u|AN_{i-1}) \quad \ldots\ldots (2)$$

$$= \mathrm{argmax}_w P\left(u\middle|I_{v,}\ I_T\right)P(u|AN_{i-1})$$

Then

$$AN_i = AN_{i-1} \cup w_i^* \quad \ldots\ldots (3)$$

Note that the maximum likelihood estimation for P (w | AN_i) is:

$$P_M\ (u|AN_i) = \frac{\#\{S/u,AN_i\ \in J\}}{\#\{S/AN_i \in J\}} \quad \ldots\ldots (4)$$

Where #{S |u, $AN_i \in J$} denotes the number of images in which keyword w and keywords subset AN_i appear together. For a limited training set, when $|AN_i|$ is large, the co-occurrence of w and AN_i is rare, which means there will be many zero values in the probability estimation. However, a zero probability event in the limited training set does not mean it never happen in the future, thus smoothing is necessary.

In the text information retrieval, smoothing is usually performed by making use of a huge background collection to allocate a non-zero probability to not happened event in current model. For instance, we can choose a larger training image set for smoothing. However, it is hard to obtain sufficient training images for the Web-scale image annotation task. Therefore, we propose to explore the Web Social Multimedia to infer semantic correlation, rather than maintaining a large scale image database by ourselves, or using the limited training image set to generate the keywords correlation graph. It is expected that Web-scale semantic space learning is more flexible to deal with the scalability problem of Web images annotation. Denoting the keywords correlation graph as Sim, then we can smooth the maximum likelihood estimation for P (u |AN_i) by the keyword semantic similarity graph as follows:

$$P(u|AN_i) = (1-\gamma)\, P_M\ (u|AN_i) + \gamma \sum_{v\in V} \frac{\mathrm{Sim}\ (u,v)}{\mathrm{Degree}\ _+\ (v)} P(v|AN_i), \quad \ldots\ldots (5)$$

Where γ is the smoothing factor. V is the vertex set of the graph Sim. Degree +(v) is the outside degree of vertex v in Sim, that is:

$$\mathrm{Degree}_+\ (v) = \sum_{u\in V} \mathrm{Sim}(u,v) \quad \ldots\ldots (6)$$

Dissimilar keywords have special importance for smoothing, here $Degree_+(v)$ captures the importance of keyword v, that is, if keyword v only associates to some keywords, then v is more important for smoothing than those associating to more keywords. Eqn.5 represents that the more similar between keyword v and u, more significant of keyword v for smoothing the probability of u. The visual generation probability P (u |I_V) is processed as the expectation over the images in the training set, that is:

$$P\ (u|I_i)\ \ \alpha\ \ P\ (u,\ I_v\ ) = \sum_{i=1}^{|T|} P(u, I_v|ZS_i)\ \ P(\ S_i\ ) = \sum_{i=1}^{|T|} P_v\ (I|S_i)\ \ P\ (u|S_i)\ \ P\ J_i \quad \ldots\ldots (7)$$

Where P_V (I |S_i) is the probability of I being generated from S_i based on their visual features. P (u| S_i) represents the probability of word u generated from S_i, which can be estimated by maximum likelihood estimation. And we suppose P (S) is uniformly distributed.

Based on the supposition that the regions of image are independent each other, P_V (I| S_i) equals to the product of the regional generation probabilities. The regional generation probability P_V (f_s |S_i) can be estimated by non-parameter kernel-based density estimation.

IV. Heuristic Model

Heuristic model refers to experience-based techniques for problem solving, learning, and discovery. Where an exhaustive search is impractical, heuristic methods are used to fast up the process of finding a satisfactory solution. This method includes using a rule of thumb, an educated imagine, an intuitive judgment, or common sense. In additional precise terms, heuristics are strategies using readily accessible, though loosely applicable, information to control problem solving in human beings and machines. In computer science, mathematical optimization, artificial intelligence and a heuristic is a technique designed for solving a problem more rapidly when classic methods are too slow, or for finding an approximate solution when classic methods fail to discover any exact ones, but they do not guarantee that the best will be found, therefore they may be considered as approximately and not precise algorithms. These algorithms usually discover a solution close to the most excellent one and they discover it quick and easily. Sometimes these algorithms can be accurate, that is they actually discover the best solution, but the algorithm is still called heuristic until this best solution is proven to be the best.

Approximate problems separated into two categories Heuristics algorithm and Meta-Heuristics algorithm. Heuristics algorithm includes local search, branch & bound, dynamic programming, cutting plane, and Branch algorithms and so on. Whereas meta-heuristics algorithms give best solution or near-optimal solution instead of accurate one, they are again separated into two categories: (1) P pupation-Based and (2) Trajectory-Based. These algorithms include genetic algorithm, tabu search, ant colony optimization, simulated annealing ,GRASP, hybrid search and so on, which are suitable for big size problem to discover best solution but they necessary massively parallel processing on very large instances. Some algorithms are solution-based (LSM) such as simulated annealing, tabu search, evolutionary algorithm, scatter search etc. And few of them are population–based algorithms such as evolutionary algorithms, ant colony optimization and so on. The following figure 1 shows Taxonomy of Approximate Problems:

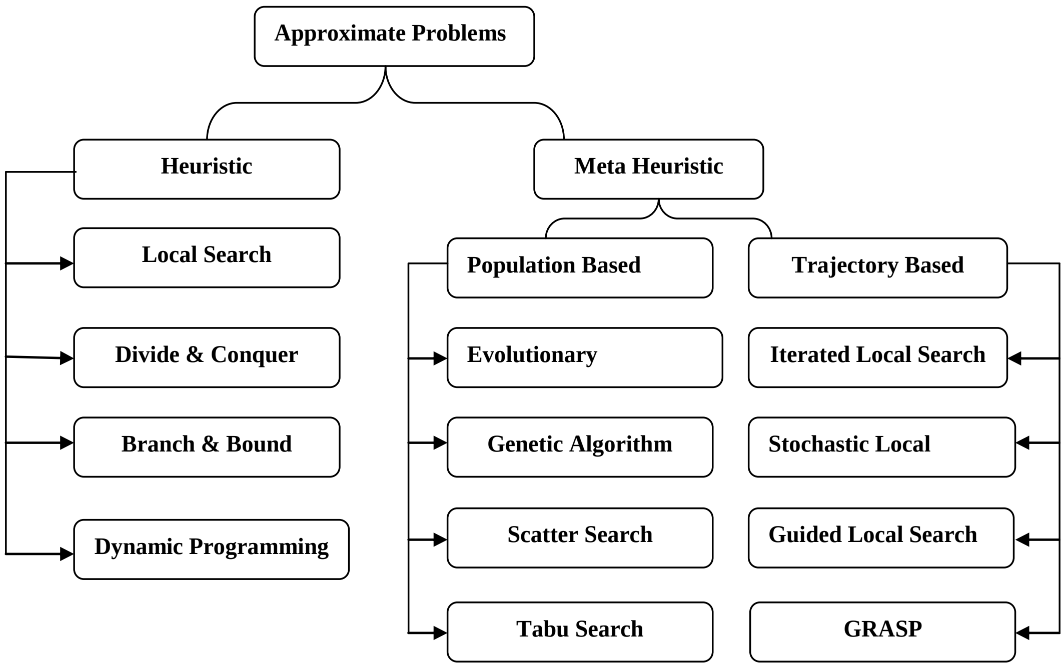


Figure 1: Taxonomy of Approximate Problems

Heuristic features for every image are stored in the database. 'Bark', 'Water', 'Cloud', 'Sand', 'Brick', 'Grass', 'Sky' etc. are some of the mostly used heuristic features. For color images color based recognition is used. Color correlogram is used as low level color feature. Calculating correlogram for so many colors are space and time consuming. It has been seen local correlation is more important than correlation with large distances. Hence to decrease the time and space complexity correlogram is calculated is for 8 distances and 8 colors. But to distinguish about similar color images like water and sky texture based recognition is used. 'car' and

'Pedestrian', are some of the heuristic features used and can be efficiently recognized by shape based recognition method. Heuristic Search for web image annotation stated as follow:

Algorithm: Heuristic Search for Web Image annotation

Input: Graph G, integer n (colors)
Output: h (B^*) the best conflict number ever found
Parameters:

i. B and B^*: present coloring and finest coloring discover so far
ii. J: iteration counter
iii. H and H': Heuristic table and dynamic tabu tenure
iv. α: table |U | × m of conflict number variations induced by chaque move

Begin

1. J = 0 (iteration counter)
2. H = 0 (Heuristic table contains no Heuristic move)
3. B = any initial n-coloring
4. B^* = B (save the finest coloring discover so far)
5. Initialize α (for every possible move, calculate the induced conflict number variation)
6. While (h (B) > 0 and time/iteration limit not reached)
   i. Pick the finest acceptable move < j, a_0 (j) > from the neighborhood (if added moves lead to a finest conflict number, a random choice is taken using the analysis function)
   ii. H [j, a (j)] = J + H' (set the color of j Heuristic)
   iii. a (j) = a_0(j) (perform the move)
   iv. Update α
   v. If (h (B) < h(B^*)) then B^* = B (Enhanced coloring found)
   vi. J = J + 1;
7. Return f (B*)

End

The following figure 2 shows step- by-step graphical representation of our proposed method:

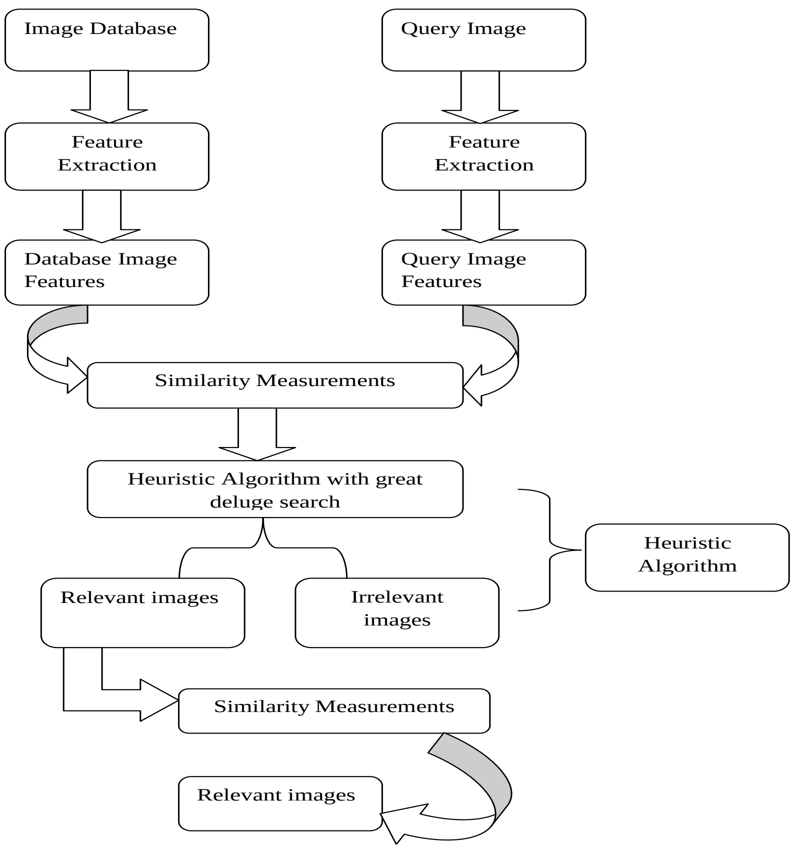


Figure 2: Block diagram of proposed method

In this method annotation of image is done by estimating the combined probability of an image with a set of terms. Duygulu et al make use of the statistical machine translation model and applied the Expectation Maximization algorithm to study a maximum likelihood relationship of terms to image regions using a bilingual corpus.. The pre-processed COREL data-set made available by Duygulu et al has become a widely used and popular benchmark of annotation systems in the literature. The major core of the algorithm is that it utilizes the probability table to estimated correspondences and using it to refine the estimate of the probability table. It annotates the image by separate the segments into blobs and finding the connection of words and blobs by selecting the words with highest probability. Feng et al proposed a probabilistic generative model which is based on Bernoulli compute to produce words and kernel density estimate to generate image features. It simultaneously learns the joint probabilities of associating words with image features using a training set of images with keywords and then generates multiple probabilistic annotations for each image.

This approach uses multiple grid segmentation and feature extraction is done using color and texture characteristics of image. Each training image has many annotations .This approach focuses on the presence or absence of words in the annotation rather than its importance. It does not rely on clustering and models continuous features. Mori et al. proposed a Co-occurrence Model which is based on the co-occurrence of words with image regions created using a regular grid. The annotation process began by partitioning images into rectangular tiles of the same size. Then, for each tile, a feature descriptor which was fusion of color and texture is calculated. All the descriptors were then clustered into a number of groups which is represented by the centroid. Each tile inherited the whole set of labels from the original image. Then, the estimation of the probability of a label related to a cluster by the co-occurrence of the label and the image tiles within the cluster is done. Wang et al proposed progressive model to approximate the joint probability of words in for a given an image, the word with highest probability is first annotated. Then, the successive words are annotated by incorporating the information of previously annotated words. In this model joint probability of words is calculated on basis of greedy algorithm.

V. Statistical Analysis

Web image annotation is an effective way for managing and retrieving abundant images on the internet. It is widely acknowledged that image annotation is an open and very difficult problem in computer vision. In this work, the implementation of heuristic model for web image annotation is presented. To deal with the problem of web image annotation, we have presented a new web image annotation approach by integrating visual feature and textual information in this paper. First, it selects the most relevant features from all the data points by using a heuristic model. Then, it can uncover the shared subspace of original features, which is beneficial for multi-label learning. Specifically, we use the Social Media Web site: Flickr as Web scale image semantic space to determine the annotation keyword correlation graph to smooth the annotation probability estimation. To further improve web image annotation performance, we uses statistical model to estimate the semantics of the web image from the corresponding associated text. The main goal of proposed is to furnish a good support to conceptualize images and resolve the problems of traditional image annotation. We integrate the proposed approaches into our web image annotation framework and conduct experiments on a real Web image data set.

The experiments in this study employed the Corel dataset which contains 10,908 different images with each image in the size of 256*384 or 384*256. As such, the outcomes were reported utilizing the ten semantic sets with every comprising of 100 images. These datasets are in the groups of Food, Buses, Elephants, Mountains, Beach, Buildings, Flowers, Africa, Horses and Dinosaurs. These groups were used in reporting the results owing to the fact that the majority of the outstanding researches. The performance of image annotation can be evaluated using precision / recall parameters. Precision of image annotation is the fraction of retrieved documents that are relevant to the query:

$$\text{Precision} = \frac{|\{\text{releavant document}\} \cap \{\text{retrived documents}\}|}{|\{\text{retrived documents}\}|} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots..(8)$$

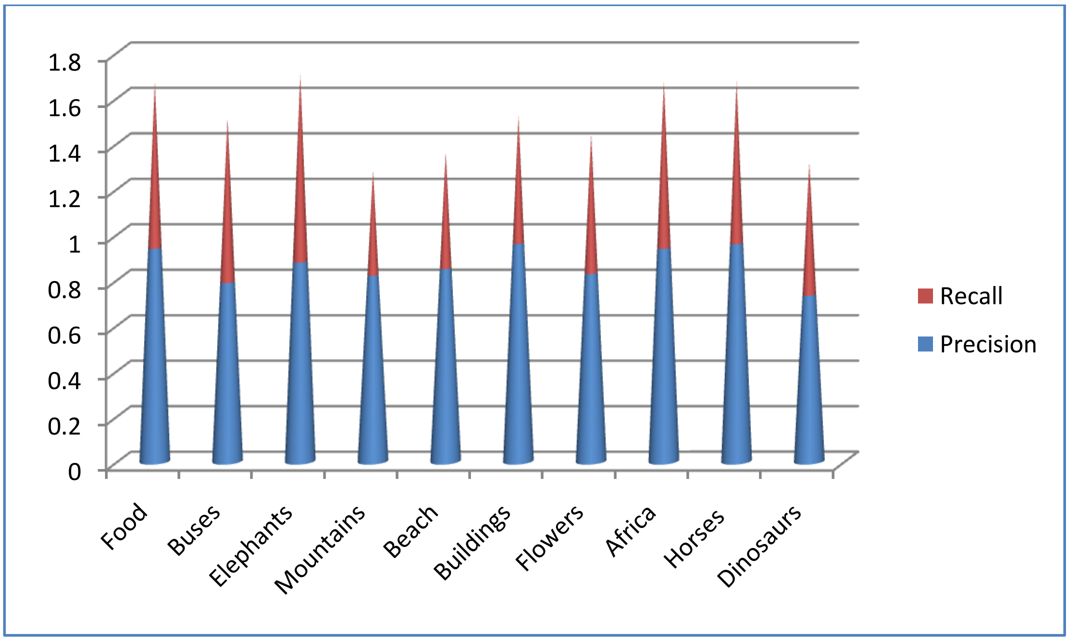
Precision takes all retrieved documents into account, but it can also be evaluated at a given cut-off rank, considering only the topmost results returned by the system. Precision is used with recall, the percent of all relevant documents that is returned by the search. The two measures are sometimes used together in the F1 Score (or f-measure) to provide a single measurement for a system. Recall of image annotation is the fraction of the relevant documents that are successfully retrieved.

$$\text{Recall} = \frac{|\{\text{releavant document}\} \cap \{\text{retrived documents}\}|}{|\{\text{releavant document}\}|} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots..(9)$$

For a text search on a set of documents, precision is the number of correct results divided by the number of all returned results and Recall is the number of correct results divided by the number of results that should have been returned. Statistically, when more similar images are retrieved, the precision and recall will be better. The reported results by means of the extracted features combined with proposed method show very promising improvements with respect to the efficiency and accuracy of the overall web image annotation process. The precision / recall parameters are evaluated on corel image dataset as shown in the following table 1:

Table 1: Evaluation of precision / recall parameters on Corel image dataset

Attribute group	**Precision**	**Recall**
Food	0.94	0.73
Buses	0.79	0.72
Elephants	0.88	0.821
Mountains	0.82	0.46
Beach	0.85	0.51
Buildings	0.96	0.56
Flowers	0.829	0.61
Africa	0.94	0.73
Horses	0.96	0.71
Dinosaurs	0.734	0.58



In specific, the average precision and recall rates obtained were 0.8703 and 0.6431 respectively. The statistical results demonstrating its feasibility and can improve the annotation performance significantly of being applied to web image annotation.

VI. Conclusion

The amount of web images has been explosively growing due to the improvement of network and storage technology. These images make up a huge amount of current multimedia data and are intimately related to our daily life. To efficiently browse, retrieve and arrange the web images, various numerous approaches have been proposed. Most established annotation methods use image features that are often noisy and redundant. Hence, feature selection can be abused for a additional precise and compact demonstration of the images, thus improving the annotation performance. Differing from the traditional lexicon or training data based keyword correlation estimation; we propose a new feature selection technique called heuristic model and apply it to web image annotation. Specifically, we use the Social Media Web site: Flickr as Web scale image semantic space to determine the annotation keyword correlation graph to smooth the annotation probability estimation. To further improve Web image annotation performance, we uses probability model to estimate the semantics of the Web image from the corresponding associated text. We integrate the proposed methods into our Web image annotation framework and conduct experiments on a genuine Web image data set. The statistical results representing its feasibility and can improve the annotation performance significantly of being applied to web image annotation.

References

[1]. Mutasem K. Alsmadi,” Query-sensitive similarity measure for content-based image retrieval using meta-heuristic algorithm”, Journal of King Saud University – Computer and Information Sciences, http: //dx.doi.org/10.1016/j.jksuci. 2017. 05.002. 2017.

[2]. Kai Zhou and Feng Tian,” heuristic Label Set relevance Learning for Image annotation”, IJCSNS International Journal of Computer Science and Network security, Vol. 14 NO.6, June 2014.

[3]. Chaoran Cui, Jun Ma, Tao Lian, and Zhumin Chen,” Improving Image Annotation via Ranking-Oriented Neighbor Search and Learning-Based Keyword Propagation”, Journal of the Association For Information Science And Technology, 66(1):82–98, 2015.

[4]. Hongtao Xu, Xiangdong Zhou, Lan Lin, Yu Xiang, and Baile Shi,” Automatic Web Image Annotation via Web-Scale Image Semantic Space Learning”, APWeb/WAIM 2009, LNCS 5446, pp. 211–222, 2009. Springer-Verlag Berlin Heidelberg 2009.

[5]. Ashraf, R., Bashir, K., Irtaza, A., Mahmood, M.T., 2015. Content based image retrieval using embedded neural networks with bandletized regions. Entropy 17, 3552– 3580.

[6]. Ashraf, R., Bashir, K., Mahmood, T., 2016. Content-based image retrieval by exploring bandletized regions through support vector machines. J. Inform. Sci. Eng. 32, 245–269.

[7]. Avci, M., Topaloglu, S., 2017. A multi-start iterated local search algorithm for the generalized quadratic multiple knapsack problem. Computer Operations Res. 83, 54–65.

[8]. Nikoo, H., Talebi, H., Mirzaei, A., 2011. A supervised method for determining displacement of gray level co-occurrence matrix. In: Machine Vision and Image Processing (MVIP), 2011 7th Iranian, pp. 1–5.

[9]. Syam, B., Rao, Y., 2013. An effective similarity measure via genetic algorithm for content based image retrieval with extensive features. Int. Arab J. Inform. Technol. (IAJIT) 10.

[10]. Selvarajah, S., Kodithuwakku, S., 2011. Combined feature descriptor for content based image retrieval. In: Industrial and Information Systems (ICIIS), 2011 6th IEEE International Conference on, 2011, pp. 164–168.

[11]. Bertelli, L., Yu, T., Vu, D., & Gokturk, B. (2011). Kernelized structural svm learning for supervised object segmentation. In Proceedings of the 2011, IEEE Conference on Computer Vision and Pattern Recognition (pp. 2153–2160). Los Alamitos, CA, USA: IEEE Computer Society.

[12]. Cao, Z., Qin, T., Liu, T.-Y., Tsai, M.-F., & Li, H. (2007). Learning to rank: From pair wise approach to listwise approach. In Proceedings of the Twenty-Fourth International Conference on Machine Learning (pp. 129–136). New York, NY, USA: ACM.

[13]. arneiro, G., Chan, A., Moreno, P., & Vasconcelos, N. (2007). Supervised learning of semantic classes for image annotation and retrieval. IEEE Transactions on Pattern Analysis and Machine Intelligence, 29, 394–410.

[14]. Li, X., Snoek, C., & Worring, M. (2009). Learning social tag relevance by neighbor voting. IEEE Transactions on Multimedia, 11(7), 1310–1322.

[15]. Liu, J., Li,M., Liu, Q., Lu, H., & Ma, S. (2009). Image annotation via graph learning. Pattern Recognition, 42, 218–228.

[16]. Niu, S., Guo, J., Lan, Y., & Cheng, X. (2012). Top-k learning to rank: Labeling, ranking and evaluation. In Proceedings of the Thirty-Fifth International ACM SIGIR Conference on Research and Development in Information Retrieval (pp. 751–760). New York, NY, USA: ACM.

[17]. Siddiquie, B., Feris, R., & Davis, L. (2011). Image ranking and retrieval based on multi-attribute queries. In Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition (pp. 801–808). Los Alamitos, CA, USA: IEEE Computer Society.

[18]. Tang, J., Hong, R., Yan, S., Chua, T.-S., Qi, G.-J., & Jain, R. (2011). Image annotation by knn-sparse graph-based label propagation over noisily tagged web images. ACM Transactions on Intelligent Systems and Technology, 2(2), 14:1–14:15.

[19]. Yakhnenko, O., & Honavar, V. (2008). Annotating images and image objects using a hierarchical dirichlet process model. In Proceedings of the Ninth International Workshop on Multimedia Data Mining (pp. 1–7). New York, NY, USA: ACM.

[20]. Torralba, A., Fergus, R., & Freeman, W. (2008). 80 million tiny images: A large data set for nonparametric object and scene recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 30(11), 1958–1970.

[21]. Tang, J., Yan, S., Hong, R., Qi, G.-J., and Chua, T.-S. 2009. Inferring semantic concepts from community-contributed images and noisy tags. In *Proceedings of the seventeen ACM international conference on Multimedia*. 223– 232.

[22]. Tseng, V., Su, J.,Wang, B., Lin, Y.:Web image annotation by fusing visual features and textual information. In: SAC, pp. 1056–1060 (2007).

[23]. Zhou, X., Wang, M., Zhang, Q., Zhang, J., Shi, B.: Automatic image annotation by an iterative approach: incorporating keyword correlations and region matching. In: CIVR, pp. 25–32 (2007).

[24]. Jin, Y., Khan, L., Wang, L., Awad, M.: Image annotations by combining multiple evidence & wordnet. In: ACM Multimedia, pp. 706–715 (2005).

[25]. Feng, H., Shi, R., Chua, T.: A bootstrapping framework for annotating and retrieving www images. In: ACM Multimedia, pp. 960–967 (2004).

[26]. Tseng, V.S., Su, J.-H., Wang, B.-W., Lin, Y.-M., 2007. Web image annotation by fusing visual features and textual information. In: Proceedings of the 2007 ACM symposium on Applied computing, pp. 1056–1060.

[27]. Tseng, V.S., Su, J.-H., Huang, J.-H., Chen, C.-J., 2008. Integrated mining of visual features, speech features, and frequent patterns for semantic video annotation. IEEE Trans. Multimedia 10, 260–267.

[28]. Yang, X.-S. 2009. Firefly algorithms for multimodal optimization. In: O. Watanabe, T. Zeugmann (eds.), Stochastic Algorithms: Foundations and Applications: 5th International Symposium, SAGA 2009, Sapporo, Japan, October 26–28, 2009. Proceedings, Berlin, Heidelberg, Springer, Berlin Heidelberg, pp. 169–178.

[29]. Jyothi, G., Sushma, C., Veeresh, D.S.S., 2015. Luminance based conversion of gray scale image to RGB image. Int. J. Comput. Sci. Information Technol. Res. 3, 279– 283.

[30]. Kanimozhi, T., Latha, K., 2013. A Meta-heuristic optimization approach for content based image retrieval using relevance feedback method. In: Proceedings of the World Congress on Engineering 2013 London, U.K.