ISSN 2320-9984

Volume 1, Issue 2, July 2013 ISSN: 2320-9984 (Online)

International Journal of Modern Engineering & Management Research

Website: www.ijmemr.org

Social Image Retrieval using Tag Relevance with Neighbourhood Voting

Sandeep Kumar Gupta
M.Tech Scholar
Patel Institute of Engineering and Science,
Bhopal, (M.P,) [INDIA]
Email: 28.sandeepgupta@gmail.com

Prof. Hitesh Gupta
Assistant Professor,
Patel Institute of Engineering and Science,
Bhopal, (M.P,) [INDIA]
Email: hitesh034@gmail.com

Abstract—*Within society image retrieval is important for exploiting the increasing amounts of nonprofessional-tagged multimedia. An algorithm is proposed that is scalable and reliably learns tag relieving by aggregating votes from visually similar neighbors. It is treated as tag frequency; learned tag relevance is flawlessly embedded into current tag-based social image retrieval hypotheses. In order to acquire image retrieval for tag based images, labeled images and unlabeled images we can use this proposed algorithm. On the basis of image ranking and tag ranking the proposed algorithm tag the images and retrieve them and according to the need either tag based retrievals, retrieval of labeled images or retrieval of unlabeled images. Preliminary experiments on one thousand images exhibit the prospective of the proposed algorithm. The results imply that the proposed algorithm is capable for real-world applications.*

Keywords- *Tag relevance, user contributed tag, social image tagging, and neighbour voting.*

1. INTRODUCTION

Image processing is any form of signal processing for which we give image as combine input, like a photograph or video frame. The output of image processing may be either an image or a set of characteristics related to the image. Usually Image Processing system includes treating images as two dimensional signals while applying previously set signal processing methods to them.

Chronological records show the use of images date back to paintings on walls of cave by early man. In the pre-Roman times images were seen mostly in the form of building plans and maps [1].

A. Image Tagging

One can correlate high-level meanings to images or image regions through image tagging, also known as captioning or annotations. Tagging enhances the content of images and efficiently retrieve desired images in response to text queries in image retrieval search engines. A social image tagging is a federal online service which enables users to add, annotate, edit, and share bookmarks of web documents. Tagging is a significant feature of social bookmarking systems, enabling users to organize their bookmarks in flexible ways and develop shared vocabularies known as folksonomies. A folksonomy is a arrangement of classification derived from the training and method of collaboratively creating and supervising tags to annotate and categorize content; this exercise is also known as collaborative tagging, social classification, social indexing, and social tagging.

B. Image Retrieval

Image retrieval is one of the demanding applications that develop along with the advancement of digital imaging technologies. An image retrieval system is a computer system for browsing, searching and retrieving images from a large database of digital images. Manual image annotation is time-consuming, laborious and costly; to address this, there has been a large amount of research done on automatic image annotation. in addition, the increase in social web applications and the semantic web have encouraged the development of numerous web-based image annotation tools. There are a variety of image retrieval techniques. These techniques can be categorized according to text, content, multimodal fusion, or semantic concepts. We discriminate these techniques by the category of features that are used to signify the images as well as the approaches that are used to retrieve similar images.

The text-based image retrieval techniques use keywords, the Content based image retrieval techniques use low-level image features, the multimodal fusion techniques use a combination of various image representative features, and the semantic-based techniques use concepts. [4].

The queries for images can be based on text descriptions or image content. Text-based description queries are posed to a text-based image retrieval system, whereas the content-based image queries are posed to a CBIR system. Text-based queries can be formulated in free-text or according to a query formation. Free-text queries are normally formulated for retrieving images using the full-text information retrieval approach. [4]

2. RELATED WORK

An image retrieval system is used for searching, browsing, and retrieving images from a large database of digital images. The techniques which were used prior for image and video retrieval consist of following steps:

- Organizing an image collection
- Classification and indexing schemes
- Current indexing practice

Improving Image Tagging

Image tagging can be improved by tagging the images on the basis of their features and tags should be relevant to the image and with the help of which image can be retrieved from pool of the databases. In text based image retrieval, images are retrieved on the behalf of the tags and tags are being given on the basis of their features and properties or characteristics of the image. Retrieval of image can also be done by multiple features together and is efficient for both unlabeled and labeled images. For labeled images tags are predicted on the behalf of the features.

The general idea of the algorithm is to progressively improve tagging accuracy by taking into account both the tags automatically predicted by an existing model and the tags provided by a user as implicit relevance feedback. In disparity to the model-based approaches, the model-free approaches attempt to predict relevant tags for an image by utilizing images on the Internet [6],[17]. These approaches imagine there exist a large well-labeled database such that one can find a visual characteristics, colour, texture etc and for unlabeled images tagging is done when we load a query image and get it neighbour images and tags are then predicted on the basis of the retrieved images and the features being exhibited by same tagged image. Depending on whether a target image is labeled, we can categorize existing methods into two main scenarios, explicitly improving image tagging for labeled images and automated image tagging for unlabeled images. In the first scenario, given an image labeled with some tags, one tries to improve image tagging by removing noisy tags [11], recommending new tags relevant to existing ones [12], or reducing tag ambiguity [5]. In [10] for instance, the authors assume that the majority of existing tags are relevant with respect to the image. They then measure the relevance of a tag by computing word similarity between the tag and other tags.

Improving Image Retrieval

Image retrieval can be improved on the basis of the content as well as the features, characteristics, colour etc of the image.

Given insufficient image tagging results, one might expect to improve image retrieval directly. Quite a few methods follow this research line, either by re-ranking search results in light of visual consistency. Re-ranking methods assume that the majority of search results are relevant with respect to the query and relevant examples tend to have similar visual patterns such as colour and texture. To find the dominant visual patterns, density estimation methods are often used, typically in the form of clustering [8] and random walk [9].

3. TRAINING TAG RELEVANCE BY NEIGHBOUR VOTING

In favour of achievement of image retrieval, a tag relevance measurement is seemed such that images appropriate with respect to a tag are ranked at the forefront of images irrelevant with respect to the tag. In the same time, to fulfill image tagging, the measurement should rank tags related with respect to an image ahead of tags inappropriate with respect to the image. On or after our past discussions we know that if different persons label visually analogous images using the same tags, these tags are most likely to reflect objective aspects of the visual content. This suggests that the significance of a tag given an image might be inferred from how visual neighbours of that image are tagged: the more regular the tag occurs in the neighbour set, the more relevant it might be, to the query image. Hence, a good tag relevance measurement should take into account the allocation of a tag in the neighbour set and in the entire collection, at the same time. Encouraged by the informal analysis above, we propose a neighbour voting algorithm for learning tag relevance. Though the proposed algorithm is simple, it is important to gain insight into the rationale for the algorithm. The following two subsections explain it. Firstly in Section 4.2 I have defined two conditions to describe the goal of tag relevance learning. After which, I have provided a formal analysis of user tagging and content-based nearest neighbour searches. After that we observe how our algorithm is naturally derived from the analysis.

A. Aim of Tag Relevance Training

A number of notations for the ease of explanation have been described.

A collection of user-tagged images is denoted as Ψ and a vocabulary of tags used in as W.

For an image $I_1 \in \emptyset$ and a tag w $\in$ W, let $r*(w, I)$:{W, $\emptyset$ } $\to$ R be a tag relevance measurement. It is called $r^*(w, I)$ an ideal measurement for image and tag ranking if it satisfies the following two conditions:

Condition 1: Image ranking. Given two images $I_1, I_2 \in \emptyset$ and tag w $\in$ W, if w is relevant to I_1, but irrelevant to I_2, then

$$r^*(w, I_1) > r^*(w, I_2) \quad (1)$$

Condition 2: Tag ranking. Given two tags $w_1, w_2 \in W$ and image $I \in \emptyset$, if I is relevant to w_1 but irrelevant to w_2, then

$$r^*(w_1, I) > r^*(w_2, I) \quad (2)$$

The goal is to find a tag relevance measurement satisfying the above two conditions.

B. Tag Relevance from Visual Neighbours

As mentioned, given an image I labeled with a tag w, the occurrence frequency of w in visual neighbours of I to some extent reflects the relevance of w with respect to I. Note that the neighbours can be decomposed into two parts according to their relevance to w, i.e., images relevant and irrelevant to w. If we know how relevant and irrelevant images are labeled with w and how they are distributed in the neighbour set, we can estimate the tag's distribution in the neighbours.

To formalize the above notions, we first define a few notations as listed in Table I. We now study how images relevant and irrelevant to a tag are labeled with that tag. In a large user -tagged image database, it is plausible that for a specific tag w, the number of images irrelevant to the tag is significantly larger than the number of relevant images,. Moreover, one might expect that user tagging is better than tagging at random such that relevant images are more likely to be labeled, meaning |Lw ∩ Rw| > |Lw ∩ Rcw|.

C. A Neighbour Voting Algorithm

That learning tag relevance ultimately comes down to computing (n_w[Nf (I, k)]- Prior (w, k)), i.e., the count of tag w in the k nearest neighbours of image I minus the prior frequency of w. Consider that each neighbour votes on w if it is labeled with w itself, n_w[Nf (I, k)] is then the count of neighbour votes on w. Thus, we introduced a neighbour voting algorithm: given a user-tagged image, a content-based k-nn search is performed to find its visual neighbours, and then for each neighbour image, its tags are used to vote on tags of the given image. The prior frequency of tag w is approximated as

$$\text{Prior}\,(w, k) \approx \frac{|L_w|}{\emptyset}\quad\frac{|L_w|}{\emptyset} \tag{3}$$

where k is the number of visual neighbours, | L_w | the number of images labeled with w, and | $\emptyset$ | the size of the entire collection. The minimum value of tagRelevance is set to 1. In other words, if the learned tag relevance value of a user-contributed tag is less than its original frequency in an image, the tag relevance learning result for that image is rejected. In addition, it is observed that the voting result might be biased by individual users who have a number of visually similar images. To make the voting decision more objective (which is the goal), we have introduced a unique-user constraint on the neighbour set. That is, each user has at most one image in the neighbour set per voting round. This unique-user constraint effectively reduces the voting bias. Finally I summarize the procedure for learning tag relevance by neighbour voting in Algorithm.

Proposed Algorithm:

Input: A user tagged image I.

Output: (tagRelevance(w, I, k), that is the tag relevance value of each tag w in I. Find the k-nearest visual neighbours of I from the collection with the unique user constraint that is a user has at most one image in the neighbour set.

for tag w in tags of I do

tagRelevance(*w*, *I*, *k*)= 0

end for

for image J in the neighbour set of I **do**

for tag *w* in (*tags _ of _ J* ∩ *tags _* ∩ *of _ I*) **do**

tagRelevance(*w*, *I*, *k*)= *tagRelevance*(*w*, *I*, *k*) + 1

end for

end for

tagRelevance (*w*, *I*, *k*)= *tagRelevance* (*w*, *I*, *k*) - *Prior* (*w*, k)

tagRelevance (*w*, *I*, *k*)= max (*tagRelevance* (*w*, *I*, *k*).1)

D. Flow Chart of Proposed Method

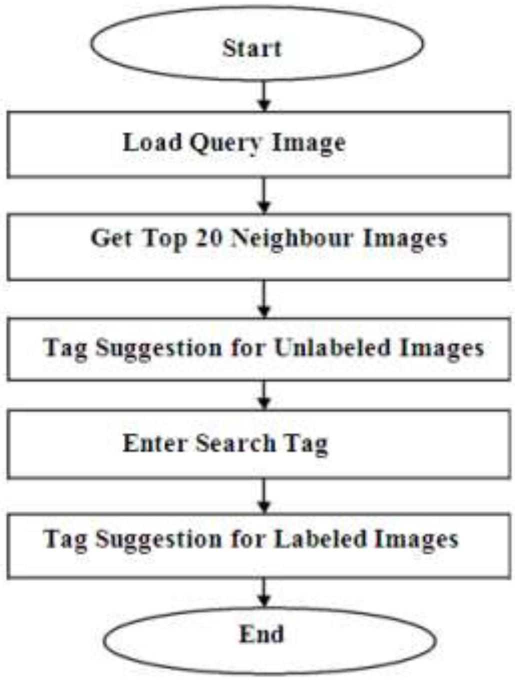

Figure 1. Shows Flow Chart of the Proposed Algorithm

E. Description of Proposed Algorithm

Image tagging can be done in two ways:

- Tag suggestion or prediction
- Tag based search button

First of all an input query image is loaded then we can search its neighbour images after that we get the top 20 neighbour images of the loaded query image retrieved on the basis of image ranking. Then in order to find tags for labeled images we enter user tag and then related to it we get the tag suggestion for labeled images. If images are unlabeled by clicking on the search button we can get the tag suggestion for unlabeled images also according to top rank priority. When a query image is loaded and we try to find its neighbour images then on the basis of image ranking the top 20 images are retrieved which are the most matching neighbour images of that query image.

If we want to search tag suggestion for labeled images then after entering the user tag on the basis of which searching is done, after that tags matching with that entered tags are retrieved, the top 5 tags which matches with the entered user tag are retrieved on the basis of priority matching with the called tag.

If we want to search images according to the search tag we have entered we can find it with the help of tag based search button, once we enter a tag it will retrieved all the top 15 images matching with that tag i.e. all the images which matches with the called tag will be retrieved on the basis of tag ranking algorithm. The images which match with the called tag and which are the top 15 images according to tag ranking algorithm and are retrieved on the basis of priority list matches with the searched tag. At the end we get all the matching neighbour images of the loaded query image as well as with tag suggestion for labeled as well as unlabeled images.

4. EXPERIMENTS AND RESULT

Experiment

The tag relevance learning algorithm is evaluated in both an image ranking scenario and a tag ranking scenario. For image ranking, three tag-based image retrieval methods with and without tag relevance learning are compared. For tag ranking, the potential of the algorithm is demonstrated in helping user tagging in two settings, namely, tag suggestion for labeled images and tag suggestion for unlabeled images. Specifically, the following three experiments have been designed.

Tag-Based Image Retrieval

A general tag-based retrieval framework widely used in existing systems is employed. A well founded ranking function for text retrieval [40] as a baseline is adopted. Given a query q containing keywords{*w*1........*wn*}, the relevance score of an image I is computed as

$$Score(q, I) = \sum_{w \in q} \frac{qtf(w) idf(w) \left(tf(w) \cdot (k_1 + 1)\right)}{tf(w) + k_1 \cdot \left(1 - b + b\frac{l_I}{l_{avg}}\right)} \quad (4)$$

where qtf(w) if the frequency of tag w in q, the frequency of w in the tags of I, lI the total number of tags of I, and lavrg the average value of lI over the entire collection.

The function idf(w) is calculated as,

$$\frac{\log(N - |L_w| + 0.5)}{|L_w| + 0.5}$$

where N is the number of images in the collection and | Lw | is the number of images labeled with w. By using learned tag relevance value as updated tag frequency in the ranking function namely substituting tagRelevance (w, I, k) for tf(w) in eq. (5.1), we investigated how the algorithm improves upon the baseline. The performance of the baseline method and our method has been studied, given various combinations of parameters. In total, there are three parameters to be optimized. One is k, the number of neighbours for learning tag relevance. k is chosen from{10; 20; 30; 40; 50; 100; 120; 150; 200}.The other two are b and k1 in ranking function. The parameter b ($0 \leq b \leq 1$) controls the normalization effect of document length. The document length is the number of tags in a labeled image. Let b range from 0 to 1 with interval 0.1. The variable k1 is a positive parameter for regularizing the impact of tag frequency. Since k1 does not affect ranking for single-word queries, k1 is set to 2, a generally common choice in text retrieval [40]. Considering that the ranking function originally aims for text retrieval and thus might not be optimal for tag-based image retrieval, further comparison with a recent achievement in web image retrieval by Jing and Baluja [9]. As depicted in [9], there are two parameters to optimize: a dump factor d (d > 0.8) controlling the restart probability of random walk and m the number of top ranked results in an initial list to calculate the prior probability. Various parameter combinations are tried, i.e.

d $\in$ {0.85;0 $\in$ .90;0.95} and m {5;10;15;20;30;50;100}.

Tag suggestion for labelled images

Given an image labeled with some tags, it is aimed for automated methods that accurately suggest new tags relevant to the image. We investigate how our algorithm improves upon a recent method by Sigurbjornsson and Van Zwol [12] by introducing visual content information into the tag suggestion process. Similar to [12], first x is computed, candidate tags having the highest co-occurrence with the initial tags. For each candidate tag, then compute its relevance score with respect to the image as,

$$score(c, I) = score(c, w_I).\lambda \frac{\square}{\lambda + (rank_c - 1)} \quad (5)$$

where c is the candidate tag, I the image, and w_i the set of initial tags. The function *Score* (*c*, w_I) computes a relevance score between the candidate tag and the initial tags. Vote+ is adopted, the best method in [12], as an implementation of the score function. The input $rank_c$ is the position of tag c in the candidate tag list ranked by tag relevance in descending order. The variable λ is a po λ sitive parameter for regularizing the effect of tag relevance learning.

Tag suggestion for unlabeled images

Compare with two model-free approaches: a tag frequency (*tf*) approach and an approach by Wang et al. [40] which re-weights the frequency of a tag by its inverse document frequency (*tf idf*). For our algorithm, since no user-defined tags are available, all tags in the vocabulary are considered as candidates. tagRelevance for each candidate tag is estimated with respect to the unlabeled image, and then rank the tags in descending order by tagRelevance. Care is taken to make the comparison fair. First, since the baselines do not consider user information, the unique-user constraint is removed from our algorithm. Second, for all methods the numbers of the visual neighbours are fixed to 100 as suggested in [41]. Finally, for each method, the top 5 tags are selected as a final suggestion for each test image.

In all the three experiments, baseline is used to represent he baseline methods, and tagRelevance for our method.

Data Collections

Social image twenty data group is chosen as a test case of user tagging. Images have been downloaded from social image twenty data group by randomly generating photo ids as query seeds. By removing images having no tags and those failed to extract visual features. 2000 labeled images were obtained in total. It is to be noted that the image retrieval experiment studies how well images are ranked, while the two tag suggestion experiments focus on how well tags are ranked. Different targets result in two different evaluation sets, one for image retrieval and the other for tag suggestion.

Table 1 **Experiment**. Each query has 1000 manually labeled examples.

User Tagging Accuracy = $\frac{\text{NUMBER OF RELEVANT IMAGES}}{1000}$

GROUND TRUTH STATISTICS FOR OUR IMAGE RETRIEVAL		
3.5 million user-tagged images		
Query	**Tag Frequency**	**User Tagging Accuracy**
car	37,614	0.548
cityscape	11,063	0.657
classroom	7,763	0.388
dog	52,981	0.764
flower	71,699	0.829
harbour	8,420	0.503
horse	27,008	0.736
kitchen	11,464	0.389
lion	8,509	0.326
mountain	36,844	0.502
rhino	4,929	0.346
sheep	3,603	0.525
street	40,772	0.426
tiger	8,214	0.224
airplane	15,231	0.447
beach	64,348	0.331
boat	25,385	0.424
bridge	25,197	0.762
bus	14,296	0.641
butterfly	8,476	0.701

Assessment set for image retrieval

A ground truth set has been created as follows. 20 diverse visual concepts have been selected as queries listed in Table II. As defined earlier, a query concept is considered and an image relevant if the concept is clearly visible in the image and we shall relate the concept to the visual content easily and consistently with common knowledge. Therefore, toys, cartoons, painting, and statues of the concept are treated as irrelevant. For each query, we randomly select 100 examples from images labeled with the query in our 2000 social image collection, and re label them according to our labeling criterion. For each query, we score its 100 test images with the two baseline methods and the proposed algorithm, respectively. The images are then ranked in light of their relevance scores. If two images have the same score, they are ranked according to photo ids in descending order so that latest uploaded images are ranked ahead.

Assessment set for tag suggestion

For evaluation of the performance of tag suggestion for labeled and unlabeled images, a ground truth set is adopted from [12], which is created by manually assessing the relevance of tags with respect to images. The set consists of 2000 images collection. Note that these tags might be predicted by tag suggestion methods. In that case, the tags are considered irrelevant. The number of tags per image in the evaluation set varies from 1 to 5.

C. Evaluation Criteria

For image retrieval, images relevant with respect to user queries should be ranked as high as possible. Meanwhile, ranking quality of the whole list is important not only for user browsing, but also for applications using search results as a starting point. For tag suggestion, tags relevant with respect to user images should be ranked as high as possible. Also, the candidate tag list should be short such that users pick out relevant tags easily and efficiently. Thus, the following two standard criteria are adopted to measure the different aspects of the performance. Given a ranked list of L instances where an instance is an image for image retrieval and a tag for tag suggestion, we measure the following:

Precision:

The proportion of relevant instances in the top n retrieved results, where $n \leq l$. The percentage of no. of relevant images out of retrieved images is known as precision.

Average precision (AP):

AP measures ranking quality of the whole list. Since it is an approximation of the area under the precision-recall curve [38], AP is commonly considered as a good combination of precision and recall, For evaluation of the overall performance, we use mean average precision abbreviated as MAP, a common measurement in information retrieval. MAP is the mean value of the AP over all queries in the image retrieval experiment and all test images in the tag suggestion experiments.

Ranking Function

A ranking function used by search engines to rank matching documents according to their relevance to a given search query. One of the most prominent instantiations of the function is as follows.

Given a query *Q*, containing keywords *q*1...........*qn*, the ranking function score of a document *D* is:

$$Score(D,Q) = \sum_{i=1}^{n} IDF(q_i) \cdot \frac{f(q_i, D) \cdot (k_1 + 1)}{f(q_i, D) + k_1 \cdot \left(1 - b + b \cdot \frac{|D|}{avgdl}\right)} \quad (6)$$

Where f(qi, D) is qi’s term frequency in the document D, |D| is the length of the document D in words, and avgdl is the average document length in the text collection from which documents are drawn. k1 and b are free parameters, usually chosen, in absence of an advanced optimization as k1 ∈ [1.2,2. ∈ 0] and b= 0.75. IDF(qi) is the IDF (inverse document frequency) weight of the query term qi It is usually computed as:

$$IDF(q_i) = \log \frac{N - n(q_i) + 0.5}{n(q_i) + 0.5} \quad (7)$$

where *N* is the total number of documents in the collection, and *n* -

(*qi*) is the number of documents containing *qi* .

E. Image Tagging

Image tagging can be done in two ways:-

Tag suggestion or prediction
Tag based search

1. Tag Suggestion

Tag suggestion is further divided into two parts. First is tag suggestion for labeled images and second is tag suggestion for unlabeled images. First of all we take an example of airplane for showing tag suggestion for labeled as well as unlabeled images and secondly tag search images for the entered tag.

Figure 2. Shows a query image of an airplane and its top 20 neighbour images.

In this image the top 20 neighbour images of the query images are retrieved on the basis of image ranking. The top 20 neighbour images are selected on the basis of image ranking algorithm in which a priority list is prepared in order to get the best matching neighbour images.

Figure 3. Shows manually entered tag

Once we enter a tag manually, after that we ask for tag suggestion for labeled images and get the relevant tags for the image. Here we enter tag airplane and get the tag suggestion for images labeled with a tag airplane. We get top 5 relevant tag to the query image airplane for labeled images more tag suggestions are airplane, airport, air show, plane, Boeing.

Figure 4. Shows tag suggestion for labeled images relevant to manually entered Tag

In this image the top 5 tags suggested for labeled images on the basis of tag entered manually are retrieved.

Table 2. Shows tag relevance values at different precision for labeled images

Evaluation criteria	Tag relevance
Precision at 5	0.4
Precision at 10	0.2
Precision at 15	0.266
Precision at 20	0.3

Table 3. Shows new tag suggestion for labeled images

Tag Suggestion For Labeled Images		
User Labeled		New Suggested Tags
Images	Tag	TagRelevance
	Airplane Moun- tains	Airplane Airport Air show Plane Boeing
	Beach Tree	Beach Ocean Sand Sea Vacation
	Bridge Light	Bridge Night Water River Top
	Car Road	Car Auto Car show 2006 Street
	Flower Leaves	Flower Butterfly Macro Flowers Nature

Table shows the relevant tags for the labeled images. These tags are retrieved on the basis of tag and image ranking. The top 5 tags according to priority list are retrieved which matches which the entered user tag.

Figure 5. Shows the tag suggestion for unlabeled images.

In this figure the tags for the unlabeled images are suggested which are suggested on the basis of image and tag ranking and the top 5 tags which matches with the retrieved images and are ranked according to priority list are retrieved.

2. Tag suggestions for unlabeled images
For each image we choose the top five ranked tags

Table 4. Shows tag relevance values at different precision for unlabeled images.

Evaluation criteria	Tag relevance
Precision at 5	0.2
Precision at 10	0.1
Precision at 15	0.133
Precision at 20	0.1

Table 5. Shows new tag suggestion for unlabeled image

Tag suggestion For Unlabeled Images	
Visual Search	Suggested Tags
Image	Tag Relevance
	xd0 Lion Zoo Rhino Sheep
	Zoo Rhino Animal Tiger Rhinoceros
	Horse Animal Rhino Zoo Tiger
	Zoo Tiger Animal Lion Park

REFERENCES :

[1] M.E. Osadebey, "*Integrated Content-Based* Image *retrieval Using Texture, Shape And Spatial Information,* "Master Thesis Report in Media Signal Processing, pp. 3-5 February 2006.

[2] Rafael C. Gonzalez and Richard E. Woods, "*Digital Image Processin,*" Pearson Edu 2^{nd} edition, 2005.

[3] Anil K. Jain, "*Fundamental Of Digital Image Processing,*" Pearson Education 1^{st} edition, 1989.

[4] Dayang Nurfatimah Awg Iskandar, "*Image Retrieval using Automatic Region Tagging*," Doctor of Philosophy Thesis report in RMIT University, March 2008.

[5] K. Weinberger, M. Slaney, and R. van Zwol, "*Resolving tag ambiguity*," in Proc. ACM Multimedia, 2008, pp. 111–119.

[6] X.-J. Wang, L. Zhang, X. Li, and W.-Y. Ma, "*Annotating images by mining image search results*," IEEE Trans. PAMI, vol. 30, no. 11, pp. 1919–1932, 2008.

[7] D. Grangier and S. Bengio, "*A discriminative kernel-based approach to rank images from text queries*," IEEE Trans. PAMI, vol. 30, no. 8, pp. 1371– 1384, 2008.

[8] G. Park, Y. Baek, and H.-K. Lee, "*Majority based ranking approach in web image retrieval*," in Proc. CIVR, 2003, pp. 499–504.

[9] Y. Jing and S. Baluja, "*Visual Rank: Applying page rank to large-scale image search*," IEEE Trans. PAMI, vol. 30, no. 11, pp. 1877–1890, 2008.

[10] G. Zhu, S. Yan, and Y. Ma, "*Image tag refinement towards low rank, content-tag prior and error sparsity*," ACM Multimedia, pages 461–470, 2010.

[11] Y. Jin, L. Khan, L. Wang, and M. Awad, "*Image annotations by*

combining multiple evidence & Wordnet,” in Proc. ACM Multimedia, 2005, pp. 706– 715.

[12] B. Sigurbj ¨ornsson and R. van Zwol, “*Flickr tag recommendation based on collective knowledge*,” in Proc. WWW, 2008, pp. 327–336.

[13] K. Barnard, P. Duygulu, D. Forsyth, N. de Freitas, D. M. Blei, and M. I. Jordan, “*Matching words and pictures*,” Jour. Machine Learning Research, vol. 3, no. 6, pp. 1107–1135, 2003.

[14] J. Li and J. Z. Wang, “*Real-time computerized annotation of pictures*,” IEEE Trans. PAMI, vol. 30, no. 6, pp. 985–1002, 2008.

[15] P. Quelhas, F. Monay, J.-M. Odobez, D. Gatica-Perez, and T. Tuytelaars, “*A thousand words in a scene*,” IEEE Trans. PAMI, vol. 29, no. 9, pp. 1575 – 1589, 2007.

[16] R. Datta, D. Joshi, J. Li, and J. Z. Wang, “*Tagging over time: realworld image annotation by lightweight meta -learning*,” in Proc. ACM Multimedia, 2007, pp. 393–402.

[17] C. Wang, F. Jing, L. Zhang, and H.-J. Zhang, “*Scalable search-based image annotation*,” Multimedia Systems, vol. 14, no. 4, pp. 205–220, 2008.