Journal of Theoretical and Applied Information Technology
31st October 2012. Vol. 44 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

RESEARCH ON MULTI-LEVEL LOG-BASED RELEVANCE FEEDBACK SCHEME FOR IMAGE RETRIEVAL

[1]WEIFENG SUN, [2]JING LUO, [3]KAIXIAN HU, [4]CHUANG LIN
[1]School of Software of Dalian University of Technology 116620 Dalian Liaoning, China

ABSTRACT

Content-based Image Retrieval (CBIR) using relevance feedback technique is applied to improve the results of traditional techniques in image retrieval. Since the results returned by system cannot fully satisfy users and the iteration process of feedback can be very time-consuming and tedious, log-based relevance feedback is introduce to the system. In previous work, we have already introduced multi-level log-based relevance feedback scheme for image retrieval to accelerate the iteration process and to increase the hit rate. In this paper, we improve the novel algorithm and apply it in a demo image retrieval system which presents refined results based on multi-level log-based relevance feedback for Content-based Image Retrieval.

Keywords: *Image Retrieval, Multi-Level, Content-Based, Log-Based*

1. INTRODUCTION

Nowadays, with the development of digital era, governments, enterprises and personal users produce a huge amount of data, the majority of which is image. And the number keeps going up. Effective data management and image retrieval tool is necessary. Traditional text-based image retrieval face the difficulties that manually image annotation requires a vast amount of labor and that annotation asks for objectivity while people always work subjectively. To overcome the difficulties, Content-based image retrieval is introduced, which can automatically detect the low-level features, including color, shape, texture, and find similar target images (Ru Liyun, Peng X, Su Z, and Ma S, 2003). Considering the complexity of images, judging an image by its low-level features is semantically ambiguous (LI Xiangyang, Zhuang Y, and Pan Y, 2001). To make up the semantic gap between low-level features and high-level concepts, relevance feedback is applied to CBIR and has caused a wide research interests among scientists. User's feedback improves the results of CBIR greatly, and meanwhile the tedious iteration process and time-consuming property makes users impatient to wait for the refined results. However, the previous users' feedback log contains lots of effective user information about images, which can help to cut off the semantic gap between high-level concept and low-level features as well as shorten the iteration process. In this paper we use a novel algorithm developed in previous work to discuss log-based image relevance and apply it in an image retrieval system to test the results.

The rest of this paper is organized as follows. Related work is presented in Section 2. Section 3 introduces a novel algorithm which mines the multi-level log-based relevance of images. Section 4 presents performance in an image retrieval system. Section 5 is the conclusion.

2. RELATED WORK

In a traditional Contend-based Image Retrieval (CBIR) system, users will get some given samples after submitting the query. Those sample images can be both relevant and irrelevant to the desired target image users want. In CBIR, the system retrieves images based on the low-level features of the sample image given by users. The whole system works as follows (Huang Xianglin, Shen L S, 2002): first, user will give a sample image to the system, then, system extract low-level features from the sample image, and next the system will measure the similarity between the sample image and images in the image database. If the similarity reaches a certain threshold, the two images will be marked as relevant. After the whole matching process ends, all the relevant images will be ranked according to their similarity degree and then they will be presented to the user. If the user is satisfied with the result then the whole process ends whereas the user will give a feedback to the system, and then the system will run the matching process again until user finds his desired target. Though attracting a wide scale of interests and having many algorithms, the performance of CBIR is not satisfying. In order to acquire more similar images in a more efficient way, relevance feedback technique is introduced to

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

the system. After the same procedure mentioned above, users will be asked to label the relevant images and irrelevant images. Then, the relevance feedback learning machine (Meng W, Hua X, 2011) will update the search conditions and rematch all the images in the image database, refine the result after every feedback session until user finds desired target (Wu Hong, Lu H, Ma S D, 2005). Though it enhances the results, the iteration process and response time are too much to cause the users of impatience. To reduce the waiting time for users finding the desired target, log-based relevance feedback is introduced. Informative previous users' logs help construct the relevance relationship between the images in the database and rank them according to the log-based relevance degree computed by the system. Some typical approaches use support vector machine (Fu Yan, Wang Y W, Wang W Q, Gao W, 2003) or soft label support vector machine (S. C. H. Hoi, M. R. Lyu, and R. Jin, 2006) to solve the problem, however, there still are drawbacks in those approaches.

Different from schemes mentioned above, we use a multi-level algorithm to compute the log-based relevance degree in a CBIR system. When computing the log-based relevance, our multi-level log-based relevance feedback (MLLR) scheme is more efficient. It fully uses the connections between positive samples to find the target image for user.

3. BRIEF INTRODUCTION ON MULTI-LEVEL LOG-BASED RELEVANCE

Typical Log-based Relevance Feedback Scheme Using SVM : Typical log-based relevance feedback scheme uses user logs in two ways. First, the system counts the log-based relevance image to target image; second, those log-based images are used as the training samples for the learning machine. Learning machine' performance can be affected by the quality of samples. Among all the samples, those relevant sample are marked as positive samples while irrelevant ones are marked negative. After all the samples are sorted into positive and negative, they will be used to classify all the rest images in database. In most log-based relevance feedback scheme using SVM (Fu Yan, Wang Y W, Wang W Q, Gao W, 2003) or SLSVM (S. C. H. Hoi, M. R. Lyu, and R. Jin, 2006) , both positive and negative feedbacks are used, we can call it a two-class SVM learning in image retrieval (Y Chen, X S Zhou, T S Huang, 2001). After samples are engaged to the learning machine classifier, images in the dataset will be ranked by the support vector machine or soft label support vector machine. If two images are fed back as positive samples in one round of feedback session, these two images are marked as relevant "1" while if one image is fed back as positive and the other one as negative, those two images are marked as irrelevant "-1". And by adding up all the corresponding relevance degree in every round of feedback session, the degree of the log-based relevance can be computed. According to the former user logs database, the log-based relevance of image to the user's desired image is the difference between the image's log-based relevance to user's positive feedback and log-based relevance to user's negative feedback.

Having improved the result of the content-based image retrieval, we find some drawbacks in those schemes using SLSVM (S. C. H. Hoi, M. R. Lyu, and R. Jin, 2006). First, there may be some images which are relevant to the positive samples linked to the samples indirectly rather than directly. In this case, it may be due to the following reasons: first, when user is giving feedbacks, user may choose images only at the top of the choosing list and ignores images at a lower place; second, in the initialization of the system, user's log has not been fully initialized. Secondly, while all the positive samples are relevant to the target image, they surely belong to a certain sort of image category. Meanwhile negative samples can be negative in all means, that's to say that negative sample can be unorganized, e. g. A is positive to the desired target, B is negative to the desired target, C is negative to the negative sample B, so the system may judge C as positive while C can still be negative to A, which may cause some "bad" sample marked as positive too. And the third problem is that when giving feedback, users choose those "same" ones to be relevant, and those "similar" ones to be negative. It's very unfair, because, we have a great possibility to consider those "similar" ones as useful samples which means that they can be positive. That they are no better than "same" ones doesn't make them irrelevant. And the last problem is a typical one with content-based image retrieval system with relevance feedback. Usually, the number of training sample can be relatively small, and the dimension of the feature space can be quite high. The system can only choose some images out of the database for user to label. The image relevant to user's desired target is a very small part of the whole image database, in which most images remained to be unlabeled. Therefore, it is reasonable to assume that positive samples can cluster in a certain way while negative samples do not cluster because they can belong to any class. It is not possible to

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

estimate the distribution of negative images in the database based on relevant feedback (H Yu, J Hen, and K C Chang, 2004). Fig 1 shows what it looks like in a more direct way.

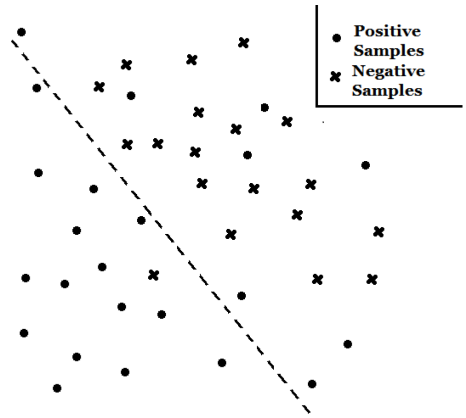


Fig 1. Dot Represents Positive Samples, And Cross Represents Negative Samples

Figure 1 shows that the system may have a high chance to miss some "good" samples if not knowing the distribution of negative samples.

To solve the problems, we use a novel algorithm named multi-level log based relevance (MLLR) feedback scheme (H C Zhang, W F Sun, S C Dong, 2010). We made an assumption that when a user gives a feedback containing two positive images, there is a virtual link built between those two images. As the rest images are done in the same manner, the virtual links between images construct a multi-level structure of log-based relevance. Assuming that A is a positive sample fed back by user, B is relevant to A, C is relevant to B, and D is fed back being relevant to C, then D will be relevant to A, though indirectly. In the case above, A is in the top level, B is in a lower level than A, C is in a lower level than B, and the same manner for D. In MLLR, upper level images' log-based relevance can be transmitted to the lower level images through those logical virtual links which makes the algorithm full uses both direct and indirect multi-level relevance. Therefore, the algorithm can full mine the log-based relevance. Besides, MLLR only use positive feedbacks which can both save users time from picking up irrelevant images and reduce the noise impact negative feedbacks bring. Also, manually collected negative samples could be biased and detrimental due to human's unintentional prejudice.

Multi-level Log-based Relevance Feedback Scheme： In multi-level log-based relevance feedback scheme, when user cannot find his desired target in the initial results which come from the content-based relevance feedback system, he will choose to start the relevance feedback procedure. The system will give user some images, then user can pick some of which as positive samples. After that the system will begin the iteration session. Each round of iteration process is regarded to be one feedback session. Through all the sessions, how the log information is stored is vital. MLLR assumes that there exists a virtual link between two images and has c_{ij} represents the number of virtual links between image i and image j. After one feedback session is finished, all the c_{ij} in the database will be updated. In order to manage the logs well, we build a virtual link vector (H C Zhang, W F Sun, S C Dong, 2010):

$$ILV = \{i, j, c_{ij}\} \qquad (1)$$

The vector is used to represent the information about the number of virtual links between any too images in the database. The first element represents image i, the second element represents image j and the third element c_{ij} represents the number of virtual links between image i and image j. After each feedback session, the CBIR system knows which images are marked positive by user. And if any of the positive images has a virtual link with images from former feedback session, the number of virtual links of the image (c_{ij}) should be updated. For n images, the MLLR system has to update C_n^2 ILV (there is C_n^2 ways to choose every pair of images from those n images) in the log database by adding a certain number to the corresponding c_{ij} .

MLLR measures the relevance degree by set a criterion value c. That the number of virtual links between image i and image j c_{ij} is larger than c means image i and image j are same whereas MLLR will uses c_{ij} / c as the relevance degree if is c_{ij} smaller. If c is manually set in the system, there will be a problem that we set it subjectively depend on the image content, i.e. if the pair of images are about a latest hot topic, the possibility to build a virtual link between them will be much higher than the situation that those two images are about an old topic. Besides, evaluating all the relevance degree

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

by the same c is not reasonable while manually setting c in every query is no possible, so we have to use a dynamic value of c. In MLLR, c is defined as following equation (2):

$$c = (\sum_{x=1}^{n-1} \sum_{y=x+1}^{n} c_{i_x i_y}) / C_n^2 \quad (2)$$

Since MLLR assumes all the feedback positive images equals, so c equals the average number of virtual links between any two images in the database. While c_{ij} is less than c, it happens that sometimes c_{ij} is too small to use the methods mentioned above. That's what we call as minority user noise. This phenomenon is due to disuse of the search system or special perception of images. So, here we try to find a suitable threshold value t. If c_{ij} is less than t, the relevance degree will be set to zero. t is defined as the equation t = a * c. a is a coefficient which can be manually set ranging from 0 to 1. To sum up, the relevance degree r_{ij} can be computed as the equation (3) below:

$$r_{ij} = \begin{cases} 1 & c_{ij} \geq c \\ c_{ij} / c & t < c_{ij} < c \\ 0 & c_{ij} \leq t \end{cases} \quad (3)$$

In MLLR, each image is given a label level based on the round of the feedback session. In the first round, user gives back the feedback images $i_1, i_2, ..., i_n$ as the initial value and all the images in the first round are given level label 0. Images in the second round of iteration process will be give a level label 1, and the rest can be done in the same manner. The reason we give each image only one level label is that each image can only belong to one level, that's one round of iteration process. For in every round of iteration process, only images which don't have a level label will be computed. Another reason is that log-based relevance of each image is only effected by the image that in the upper level.

MLLR has R(i) denotes the log-based relevance between the image to the desired target and Set $R(i_1) = R(i_2)... = R(i_n) = 1$ as the initial value since all the user's feedback images are treated the same.

In every iteration, the log-based relevance of all the unlabeled i is computed as the following equation (4).

$$R(i) = (\sum_{x=1}^{n} (r_{ii_x} * R(i_x))) / n \quad (4)$$

$r_{ii_x} * R(i_x)$ is the log-based relevance from i_x to i via virtual links, n is the number of the image. And the average R(i) of all the images to i will be used as the final log-based relevance.

Images with R(i)>0 will be labeled level 1 that means images in the first round of iteration. And we have $I_1 = \{i_{11}, i_{12}, ..., i_{1m}\}$ represents level 1 images.

In the second round of iteration, all the images which have not been labeled in the image database will be computed the relevance between user's desired image and themselves according to the equation (5).

$$R(i) = (\sum_{x=1}^{m} (r_{ii_{1x}} * R(i_{1x}))) / m \quad (5)$$

And again, images with R(i)>0 will be labeled level 2 that means images in the second round of iteration.

The iteration keeps going until:

$$R(i) = (\sum_{x=1}^{p} (r_{ii_{yx}} * R(i_{yx}))) / p = 0 \quad (6)$$

$I_y = \{i_{y1}, i_{y2}, ..., i_{yp}\}$ represents the images labeled level y. If for all the unlabeled images, R(i)=0, the whole multi-level log-based relevance feedback iteration process is finished.

When the iteration ends, all the images in the database has a log-based relevance value according to previous user logs.

4. EVALUATION

Simulation Of Collecting Log Data： Different from the method used in S. C. H. Hoi, M. R. Lyu, and R. Jin (2006), we take only positive feedback into account. The whole process starts with the users. They will choose a query at beginning, and le t CBIR system retrieval the results after clicking the search button. If users weren't able to get the desired image, they give feedback by choosing the

Journal of Theoretical and Applied Information Technology
31st October 2012. Vol. 44 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

relevant images. When users finish choosing and press search button again, the feedback will be given to the MLLR and new retrieval results will be presented. The whole process is shown in Fig 2.

Evaluate MLLR In An Image Retrieval System：In the experiment, we pick up 20 categories each of which contains 100 images. First our tester selects a query, and pick up some images as positive feedback. Meanwhile, taking noise into account, we inject 15% of the whole number of feedback images in to the log session. Then, after computing, system will retrieve images from image database and rank them according to relevance degree.

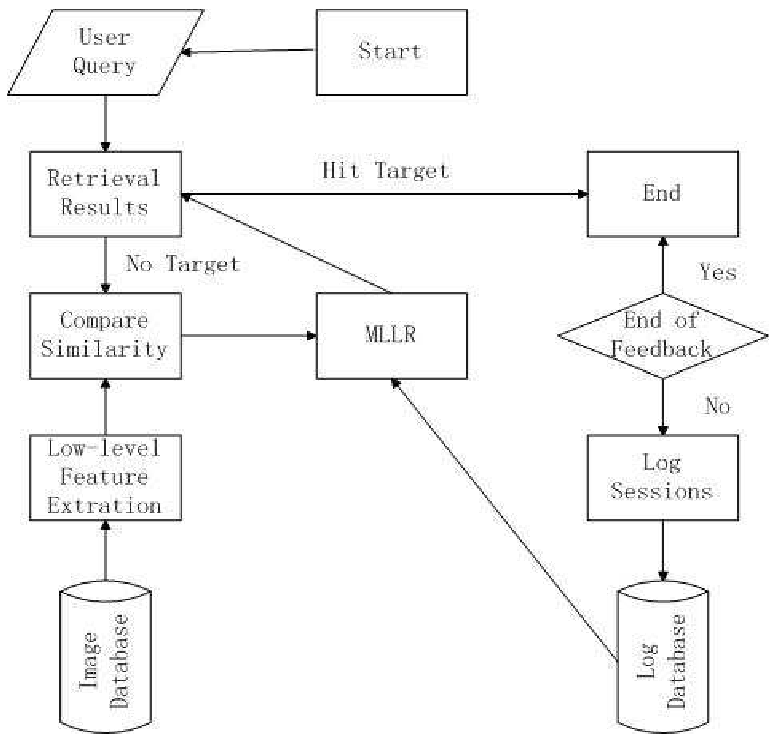


Fig 2. The Experiment Process.

In the experiment, we pick up 20 categories each of which contains 100 images. First our tester selects a query, and pick up some images as positive feedback. Meanwhile, taking noise into account, we inject 15% of the whole number of feedback images in to the log session. Then, after computing, system will retrieve images from image database and rank them according to relevance degree.

To evaluate our multi-level log-based relevance feedback (MLLR) system, we use the performance results in the log-based relevance feedback using SLSVM (LRF) (S. C. H. Hoi, M. R. Lyu, and R. Jin, 2006) and compare it to MLLR.

In the first experiment, we have q represents the query the user selects from the query database we mentioned above, PF represents user's positive feedback image samples, and NF represents user's negative feedback image samples. Num1(q) represents the number of images that in the same category as q and it is given positive relevance based on the PF. Num2(q) represents the number of images that in the different category from q and it is given negative relevance based on the NF.

The experiment compared schemes which are evaluated on 50 queries that are selected randomly from the data set. For each query, the number of labeled sample acquired from the real-world user feedback is 10. Two measure metrics are applied in first experiment to evaluate the quality of multi-level log-based relevance feedback scheme.

The first measurement is about the average ratio of Num1(q) of images which are in the same category as q, i.e. 100, over the 50 queries. In Table 1, we can find out that our multi-level log-based relevance feedback (MLLR) scheme retrieval more images in the same category as q than log-based relevance feedback (LRF) schemes using SVM or

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

SLSVM. This is because MLLR has virtual links between images and therefore can fully mine the direct and indirect relevant images while LRF only count those direct relevant images.

Table 1. Average Ratio Of Num1

	Average Num1/100
MRRL	0.40
LRF	0.34

Table 2. Average Num2

	Average Num2
MRRL	20.3
LRF	349.68

In the second measurement, the average Num2(q) is counted over 50 queries. In table 2, we can find out that MLLR gives less irrelevant images than LRF. This is because MLLR only uses positive feedback which can avoid negative effects caused by negative feedback as we mentioned above while LRF uses negative feedback. For detail, logically, a negative image to a negative feedback can be positive, so some positive images LRF choose can actually be strongly negative. This is why MLLR can perform better than LRF.

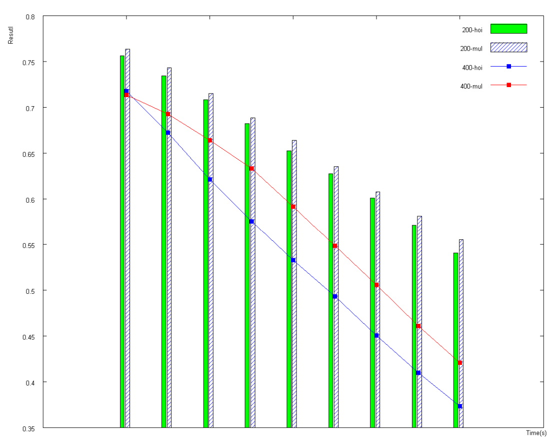


Fig. 3. The Green Bar Represents LRF In First Round While The While Bar Represents MLLR. The Blue Line Represents LRF And The Red Line Represents MLLR In Second Round.

In the second experiment, In first round, we collected 200 images as the scale of images; in second round, a scale of 400 images is used. It turned out that MLLR performed better than LRF. In Fig 3, we found out that MLLR has a higher precision ratio than LRF, and as the number of retrieved images goes up, both MLLR and LRF's precision ratio go down. Therefore, we can tell that the experiment matches real world situation, and MLLR does improve the results of log-based relevance feedback image retrieval.

In the third experiment, we apply MLLR in a test image retrieval system. First, we choose the query and search "white tiger" in the search engine without using MLLR. Then, we record the top18 images retrieved by the search engine and name them as group A. Second, we use the same query in the search engine using MLLR. Again, we record the top 18 images and name them as group B. After comparison in Fig 4, we can see that in group A, 10 out of 18 images are our desired target while in group B, 15 of the whole 18 images are our desired target. So, it proves that a search engine with MLLR performs better than one does not apply MLLR (Huanchen Zhang, H J Li, S C Dong, and W F Sun, 2011).

5. CONCLUSION AND FUTURE WORK

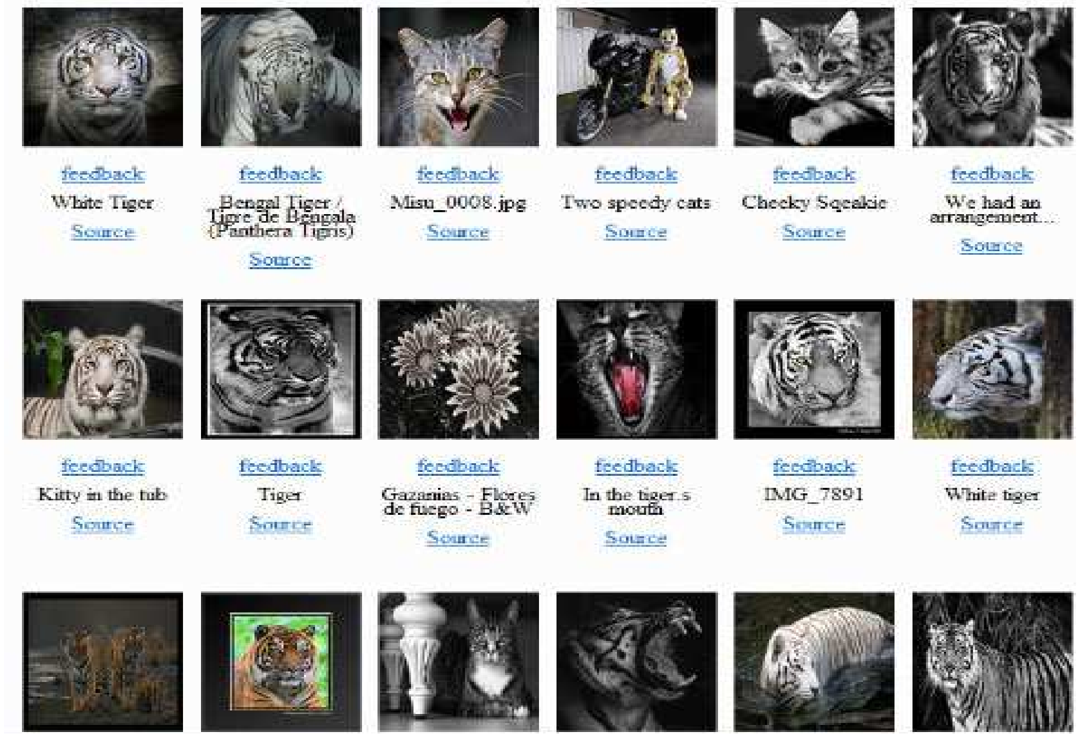
Multi-level Log-based Relevance Feedback scheme for image retrieval prove to have a better retrieval results and shorter iteration process in the experiment compared to Log-based Relevance Feedback scheme. It fully mines the potential images via using the valuable user log information, and therefore gives users a better user experience. However, limitation still exists in MLLR. First, the noise ratio cannot be decided automatically. Manually choosing noise ration cannot avoid subjective judgment. Second, the lower limit of c_{ij} does not have a standard method to compute which may affect the final result. We plan to investigate into problems mentioned above and enlarge the image database to have a further study about MLLR in the future.

ACKNOWLEDGEMENT

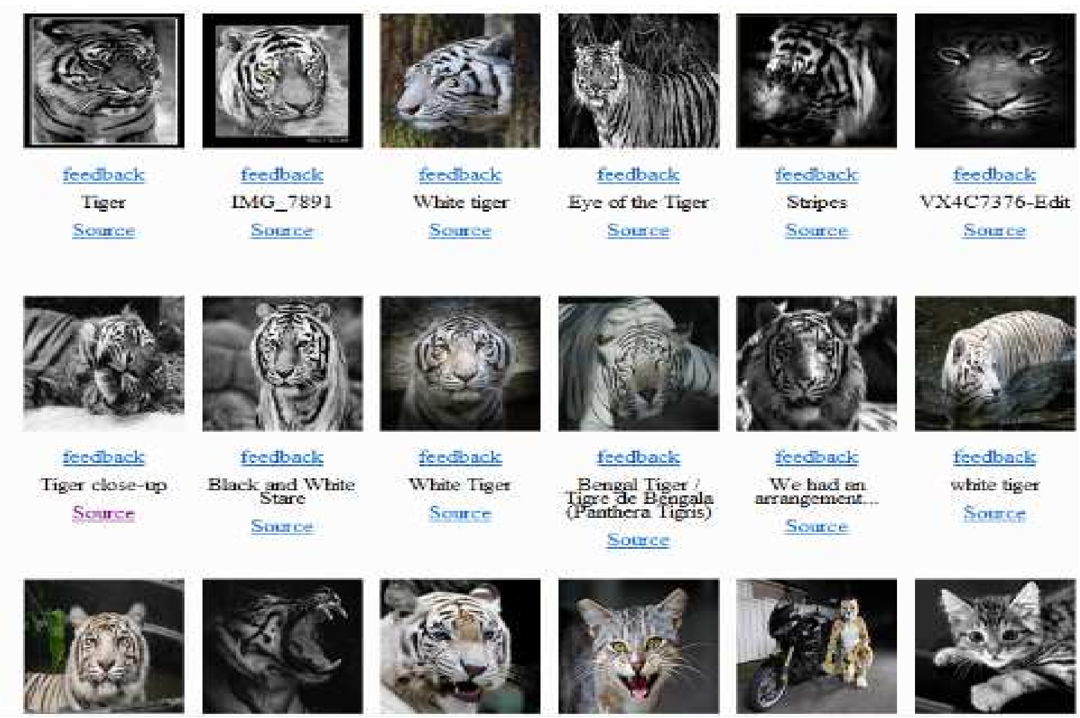
This work is supported in part by Natural Science Foundation of China under grant No. 61103233, 60903153, NSFC-JST under grant No.51021140004; Nature Science Foundation of China under grant No. 60673046, 90715037. Zhejiang Provincial Natural Science Foundation of China under Grant No. Y108685, the Fundamental Research Funds for Central Universities (DUT10ZD110)

31st October 2012. Vol. 44 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195



Group A



Group B

Fig 4. The Retrieval Results Of The Search Engine Without MLLR In Group A And Search Engine Using MLLR In Group B

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

REFERENCES

[1] Huang Xianglin, Shen Lansun, 2002. Research on content-based image retrieval techniques. Acta Electronica Sinica, 30(7): 1065-1071

[2] H. Yu, J. Hen, and K. C. Chang, 2004. PEBL : Web page classification without negative examples. Trans. on Knowledge and Data Engineering, 16:70–81

[3] Huanchen Zhang, Weifeng Sun, Shichao Dong, 2010. Multi-level log-based relevance feedback scheme for image retrieva. Computer Science

[4] Huanchen Zhang, Haojie Li, Shichao Dong, and Weifeng Sun, 2011. RoboGene: An Image Retrieval System with Multi-Level Log-Based Relevance Feedback Scheme. Computer Science

[5] Fu Yan, Wang Yaowe, Wang Weiqiang, Gao Wen, 2003. Content-based Natural Image Classification and retrieval using SVM. Chineses Journal of Computers, 26(10): 1261-1265

[6] LI Xiangyang, Zhuang Yueting, and Pan Yunhe, 2001. The technique and systems of content-based image retrieval. Journal of Computer Research & Development, 38(3): 344-353

[7] Meng, W., Xian Sheng, H, 2011. Active learning in multimedia annotation and retrieval: a survey. ACM Transactions on Intelligent Systems and Technology, ACM New York, NY, USA

[8] Ru Liyun, Peng xiao, Su Zhong, and Ma Shaoping, 2003. Feature performance evaluation in content-based image retrieval. Journal of Computer Research and Development, 40(11): 1566-1570

[9] S. C. H. Hoi, M. R. Lyu, and R. Jin, 2006. A unified log-based relevance feedback scheme for image retrieval. IEEE Trans. Knowl. Data Eng, 18(4): 509–524

[10] Wu Hong, Lu Hanqing, Ma Songde, 2005. A survey of relevance feedback techniques in content-based image retrieval. Chinese Journal of Computers, 28(12): 1969-1979

[11] Y. Chen, X. S. Zhou, T. S. Huang, 2001. One-class SVM for Learning in Image Retrieval. IEEE Int’l Conf. on Image Proc. (ICIP'2001), Thessaloniki, Greece