Improving Retrieval Quality Using Pseudo Relevance Feedback in Content-Based Image Retrieval

Dinesha Chathurani Nanayakkara Wasam Uluwitige, Timothy Chappell, Shlomo Geva, Vinod Chandran
School of Electrical Engineering and Computer Science
Queensland University of Technology
Brisbane, Australia
(d.nanayakkara@hdr.qut.edu.au, t.chappell@connect.qut.edu.au, s.geva@qut.edu.au, v.chandran@qut.edu.au)

ABSTRACT

The increased availability of image capturing devices has enabled collections of digital images to rapidly expand in both size and diversity. This has created a constantly growing need for efficient and effective image browsing, searching, and retrieval tools. Pseudo-relevance feedback (PRF) has proven to be an effective mechanism for improving retrieval accuracy. An original, simple yet effective rank-based PRF mechanism (RB-PRF) that takes into account the initial rank order of each image to improve retrieval accuracy is proposed. This RB-PRF mechanism innovates by making use of binary image signatures to improve retrieval precision by promoting images similar to highly ranked images and demoting images similar to lower ranked images. Empirical evaluations based on standard benchmarks, namely Wang, Oliva & Torralba, and Corel datasets demonstrate the effectiveness of the proposed RB-PRF mechanism in image retrieval.

General Terms

Performance, Measure.

Keywords

CBIR, Content based image retrieval, Image signature, Pseudo relevance feedback, Scaling factor.

1. INTRODUCTION

The development of the Internet created an ever-growing need for efficient and effective image browsing, searching, and retrieval tools. Retrieving relevant images accurately to satisfy an information need from a large, diversified collection using visual queries is a challenging task in information retrieval. Despite many years of research in content-based image retrieval (CBIR), an effective general solution - in terms of speed, precision and scalability still eludes researchers.

Relevance feedback (RF) [1] is an approach that seeks to improve the precision of search results through the incorporation of user feedback in response to an initial results list. When interacting with a RF interface, the user informs the search engine as to whether the results that have already been inspected are useful or not. With this feedback the system can attempt to improve upon the initial query results, usually by reformulating the original query or by re-ranking the initial result list on the basis of available feedback, at the cost of requiring additional input from the user.

Even in cases where explicit user feedback is not available, some aspects of the relevance feedback approach can be used by making the often reasonable assumption that the first pass results returned by the search engine will include a sufficiently large number of images that match the user's information need. The system then makes use of that set of top ranked initial results as if it were actual positive feedback provided by the user. This technique is often referred to as pseudo-relevance feedback (PRF). The assumption is that despite being noisy feedback, it will still help to promote relevant documents in the initial retrieval list.

Motivated by the concept of PRF, the RB-PRF approach is proposed to improve retrieval performance by applying PRF on binary image signatures. While PRF is not a new concept, two original contributions are made - the first is the original use of document signatures directly in feedback processing - as opposed to traditional approaches, which return to the original images (or documents). The second is the incorporation of the rank order of the initial results in utilising feedback to re-rank the results - rather than assume equal importance as traditionally done.

Image representation is derived through sub-image decomposition, and a full image signature is generated by the combination of sub-image features. Sub image signatures are generated from low-level features of each sub image. Image signatures are fixed length binary strings derived through a form of locality sensitive hashing [2]. The derivation of an image signature from image features is described in more detail in section 2.

A retrieved image list of length K, (much smaller than the collection size) which is small enough to be practical, is re-ranked through RB-PRF by using the top N image signatures as relevant image signature examples and the bottom N as non-relevant examples, where $N << K$. K is selected to be large enough that it can be assumed that the bottom N results are unlikely to be relevant, yet small enough to ensure that the bottom N results are still similar to the query signature, albeit irrelevant. The entire feedback-based re-ranking is performed in signature space. There is no need to return to the original image representation, as is

SIGIR '16, July 17-21, 2016, Pisa, Italy.
© 2016 ACM. ISBN 978-1-4503-4069-4/16/07…$15.00
DOI: http://dx.doi.org/10.1145/2911451.2914747

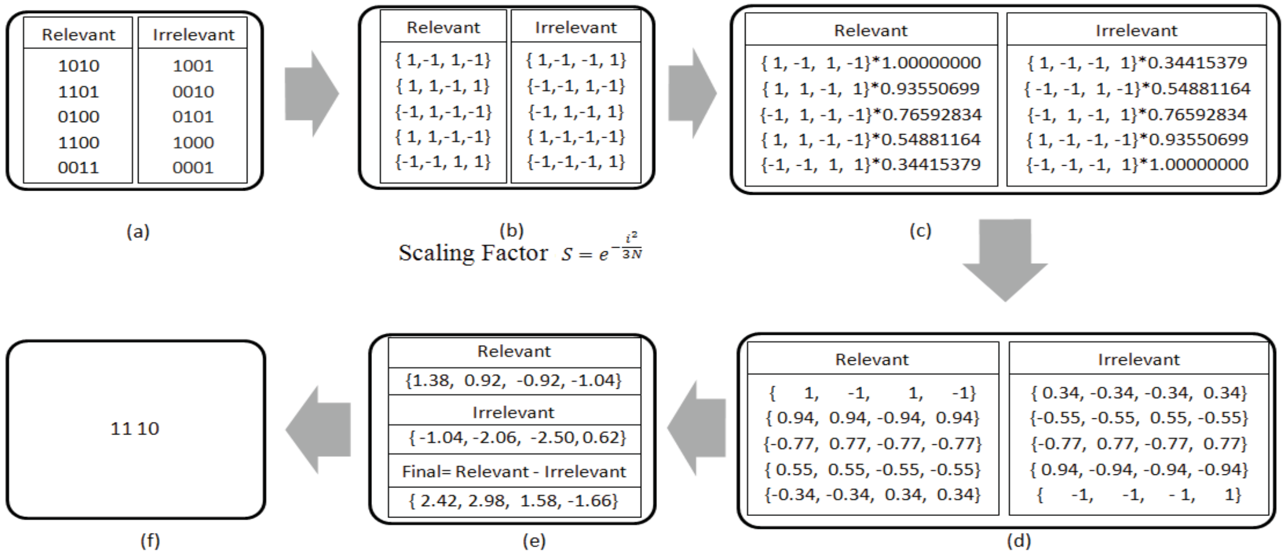

Figure. 1. Toy example to show the process of PRF using a toy dataset with signature size four bits and sample size (N) five (which are considered as relevant).

done with conventional PRF. Here we re-rank only the retrieved list of first *K* binary image vectors against the generated FB signature. Results presented here show that RB-PRF achieves effective and efficient relevance feedback in CBIR and a considerable improvement in retrieval performance over earlier approaches.

The rest of the paper is organized as follows: Section 2 provides an overview of the proposed PRF method. Section 3 provides a detailed explanation of the experiment and the results obtained relative to published baselines. Conclusions drawn from these findings are included in section 4.

2. PROPOSED RANK BASED PSEUDO RELEVANCE FEEDBACK TECHNIQUE

Our approach divides each image into a 3x3 non-overlapping grid and extracts visual features from each sub-image using global feature descriptors such as colour, texture, and shape. Each feature set is clustered using the k-means algorithm; with similar feature vectors placed together and independent visual vocabularies per feature generated. Then each sub-image is represented by those visual words from these vocabularies through codebook lookup of each raw image feature and finally the full image feature set is constructed. This representation is finally translated into a binary image signature using random indexing for efficient retrieval. This process is performed by producing text files to be parsed by the TopSig signature based search system (www.topsig.org).

Initially the search produces a list of images with signatures ranked from the most similar to least similar. The top-ranked *K* signatures are then used in RB-PRF process. The RB-PRF process considers the binary vectors of the top *N* (size of the PRF list) and the bottom *N* form the list of *K* binary vectors as relevant and irrelevant respectively, to generate FB signature. Each selected signature is represented by values of {-1, 1} with 1-bits being interpreted as 1 and 0-bits interpreted as -1. The binary signatures are transformed into real valued vectors as a weighted linear combination of the feedback signatures, weighted by rank (with negative feedback ranked in reverse order.) The feedback vectors are then added together independently and two vectors are generated - one from the pseudo-relevant signatures and one from the pseudo-irrelevant signatures. The vector generated from irrelevant signatures is then subtracted from the vector generated by relevant signatures. Finally the vector is 'squashed' back to a binary representation {1,0} simply taking the sign bit. The resulting binary signature is the feedback (FB) signature. The signatures in the result list that are to be re-ranked are then sorted according to the Hamming distance from the new FB signature. Note that the entire re-ranking process takes place in signature space without ever going back to the original image features. Furthermore, the initial list of signatures that is being re-ranked is already in memory following the initial search so the process is computationally efficient.

Figure 1 gives detailed description of the RB-PRF process for an example problem with 4-bit signatures and *N*=5 (Figure 1a). These are converted to real valued vectors as shown in Figure 1b, and multiplied by the scaling factor *S* (with $N = 5$ and *i* ranging from 0 to 4 for this example). The scaling formula used here ensures a smooth decay of feedback with increased rank position and was chosen after experiments with cross-validation. Scaling performs significantly better than without scaling. Moreover, it is not very sensitive and it is easy to tune. Figure 1c shows the process of scaling and Figure 1d shows the results. Vectors are added and subtracted as shown in Figure 1e. Finally a binary image signature is generated as shown in Figure 1f. This is then used to re-rank the initial search results.

3. EVALUATION

The RB-PRF approach is evaluated on different datasets using different evaluation measures to confirm the effectiveness in terms of precision.

3.1 CBIR System

In the CBIR system, global features such as colour, texture, and shape are used to represent images. Colour is often the most cognitively informative feature and it is relatively robust and simple to represent. Texture is also an important feature and it is very helpful in describing real world images. Shape features are vital for describing the shape of an object. Three colour descriptors are used in the CBIR system. The color histogram [3] is generated with 64 bins. First order, second order, and third order moments are calculated. Colour coherence vectors [4] are used with 16 buckets. For texture, Gabor wavelets [5] are used with five scales and eight orientations. GIST features are used with eight orientations and three scales as they describe the spatial envelope of the image. The Edge histogram descriptors are using five filters which represent shape features as well as texture. Seven invariant moments are used to further describe the shapes. YCbCr, HSV and RGB color spaces are used in these descriptors. Image representation and searching are described in section 2.

3.2 Datasets

The Wang dataset [6] of 1000 images is used for evaluation of the system and for comparisons with earlier work. It is a subset of manually selected images from the standard Corel image database and consists of 10 classes with 100 images each. For further validation the Oliva and Torralba dataset [7] is also used. It includes 2688 images classified into eight categories. As these datasets are well classified, it was possible to quantitatively evaluate and compare the performance.

In order to compare our PRF method with existing baselines that use PRF we conducted another set of experiments on subset of images from the Corel Photo Gallery [8] for which we have results for three earlier systems [8,9,10]. That data set is larger and it consists of about 12K images which are classified into 83 semantic concepts with about 100 images in each. These images are JPEG files with a resolution of 120 x 80 or 80 x 120. Even though many images containing the same semantic content are distributed across different Corel categories we use the same set as it provides consistent comparison with the PRF systems from the literature.

3.3 Evaluation Measures

Precision is used to evaluate the proposed RB-PRF approach. Precision is defined as the fraction of retrieved images in a result list that are relevant to a given query. Specifically the average precision (AP) at *n*, where *n* is the length of the result list is used. Images are considered to be correct matches if they are in the same class as the query image.

3.4 Approaches and Settings

AP@*n* is calculated for both datasets. AP@20, AP@50 and AP@100 are calculated for each class to compare with existing systems. RB-PRF is compared with the baseline signature based system (CBIR-ISIG) [11], and with existing systems from the literature. It may be noted that we do use proxies for both positive FB and negative FB in our approach as described in section 2.

3.5 Results

Tables 1 and 2 show how the baseline signature-based CBIR (CBIR-ISIG) system performs. Table 2 shows the results of CBIR-ISIG compared with other baseline systems for AP@20 and it shows that the CBIR-ISIG system generates slightly better results [11] using the signature-based approach. The differences are not statistically significant at 95% significance level but are significant at the 90% significance level. CBIR-ISIG with RB-PRF is statistically significantly better at 95% significance level. In addition Table 2 shows the results of CBIR-ISIG system compared with other systems for AP@100. They are not statistically significantly better but averages are higher. Table 3 shows the results of comparing our CBIR-ISIG system with another baseline [12] on the Oliva and Torralba dataset. In tables

Table 1. AP@20 of Wang with the performance in the literature

Category	2005 [13]	2007 [14]	2011 [15]	2011 [16]	2013 [17]	2015 [11]	+VE PRF	+&- PRF
1	0.23	0.48	0.57	**0.90**	0.7	0.75	0.82	*0.83*
2	0.23	0.34	0.58	0.38	0.28	0.68	0.74	***0.76***
3	0.23	0.36	0.43	**0.72**	0.56	0.53	0.56	*0.60*
4	0.23	0.61	**0.93**	0.49	0.84	0.86	0.90	***0.93***
5	0.23	0.95	0.98	**1.00**	0.81	**1.00**	**1.00**	***1.00***
6	0.23	0.48	0.58	0.39	0.58	0.86	**0.89**	0.87
7	0.23	0.61	0.83	0.56	0.55	0.98	0.99	***1.00***
8	0.23	0.74	0.68	0.87	0.87	0.98	1.00	***1.00***
9	0.23	0.42	0.46	0.45	0.48	0.77	**0.77**	0.74
10	0.23	0.50	0.53	0.87	0.66	0.86	0.88	***0.91***
AP	0.23	0.55	0.66	0.66	0.63	0.83	0.86	0.87

Table 2. AP@100 of Wang with the performance in the literature

Category	2000 [18]	2002 [19]	2008 [20]	2009 [21]	2011 [22]	2015 [11]	+VE PRF	+&- PRF
1	0.48	0.47	0.48	0.45	0.49	0.52	**0.62**	0.60
2	0.33	0.33	0.34	0.35	0.40	0.46	0.55	***0.56***
3	0.33	0.33	0.33	0.35	0.39	0.38	0.39	***0.40***
4	0.36	0.60	0.52	0.60	0.58	0.63	0.69	***0.74***
5	0.98	0.95	0.95	0.95	0.96	0.95	0.98	***0.99***
6	0.40	0.25	0.40	0.6	0.50	0.57	0.67	***0.68***
7	0.40	0.63	0.60	0.65	0.75	0.84	**0.95**	***0.95***
8	0.72	0.63	0.70	0.70	0.80	0.76	0.87	***0.89***
9	0.34	0.25	0.36	0.40	0.40	0.47	**0.51**	0.48
10	0.34	0.49	0.46	0.40	0.51	0.60	**0.67**	***0.67***
AP	0.47	0.49	0.51	0.55	0.55	0.62	0.69	***0.70***

Table 3. AP@50 of Oliva & Torralba

Category	2007 [12]	2015 [11]	+VE PRF	+&-PRF
1	0.84	0.63	0.72	***0.79***
2	0.50	0.50	0.48	***0.50***
3	0.76	0.85	0.96	***0.97***
4	0.80	0.73	**0.75**	0.72
5	0.62	0.75	0.81	***0.81***
6	0.44	0.94	0.95	***0.96***
7	0.38	0.67	0.76	***0.80***
8		0.73	**0.75**	0.73
AP	0.62	0.73	0.77	0.79

Table 4. AP@20 of Corel with the performance in the literature

Evaluation Criteria	No FB	+&- PRF	Simulated RF
MTSVM [8] (2006)	0.28		0.37
KBMCM [9] (2007)	0.28		0.44
BDEE [10] (2010)	0.28		0.37
CBIR-ISIG	0.29	0.31	0.44

1-3 bold values signify the highest AP for each class among the compared systems. The underlined values signify that CBIR-ISIG system [11], our underlying signature retrieval system, outperforms the compared system without any FB. Furthermore italic values signify cases where RB-PRF is higher when using both positive and negative FB than using only positive FB. We observe that the underlying CBIR-ISIG system outperforms the baseline systems on the Wang, and Oliva and Torralba datasets and the proposed additional RB-PRF mechanism works well having further considerable performance improvement.

Table 4 shows the results of CBIR-ISIG system with PRF compared with other systems for AP@20, using a sample size of

500 random images and 15-fold cross-validation. Here images are considered as positive FB (relevant) if and only if they are in the same class as the query image. All others are considered as negative FB (non-relevant).

Finally, on the time efficiency of retrieval; the use of signature-based methods, which are designed to make efficient retrieval of high-dimensionality data feasible, allows our approach to work quickly and achieve retrieval times in the milliseconds through the use of scalable signature search approaches [23]. The PRF process can also be performed efficiently due to the fact that the signatures for the retrieved images remain resident in memory after the initial search and performing the requisite re-ranking only requires the calculation of a small number of Hamming distances and a sort.

4. CONCLUSIONS

In this paper a RB-PRF approach is described for the application of pseudo relevance feedback in CBIR. The original contribution is in the first application of PRF to signatures and in taking into account the initial rank order of results to weight the feedback from each image. This approach has balance of retrieval speed and quality. The proposed RB-PRF approach performs well and outperforms several systems with previously published results on the same data sets. Our experiments demonstrate that the RB-PRF approach is effective and efficient on signature-based representation for image retrieval.

5. REFERENCES

[1] D. Harman," Relevance feedback revisited," *in Proc. of the Fifteenth Annual International ACM Special Interest Group on Information Retrieval Conference on Research and Development in Information Retrieval*, pp. 110,1992.

[2] S. Geva, and C. M. De Vries, “TOPSIG: Topology Preserving Document Signatures.” Proceedings of the 20th ACM international conference on information and knowledge management. CIKM, 2011.

[3] G. Qiu, "Indexing chromatic and achromatic patterns for content-based image retrieval," *Journal of pattern recognition,* pp. 1675-1686, 2002.

[4] G. Pass, R. Zabih, and J. Miller, “Comparing images using color coherence vectors.” *Proceedings of the fourth ACM international conference on Multimedia*. ACM, 1997.

[5] M. H. Rahmana, M. R. Pickering, and M.R. Frater, "Scale and Rotation Invariant Gabor Features for Texture Retrieval," in *Proc. DICTA International Conference on Digital Image Computing Techniques and Applications*, pp. 602-607, 2011.

[6] J. Li, and J. W. Wang, "Automatic linguistic indexing of pictures by a statistical modeling approach," *IEEE transaction on Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, pp. 1075-1087, 2003.

[7] A. Oliva, and A. Torralba, "Modeling the shape of the scene: A Holistic representation of the spatial envelope," *International Journal of Computer Vision*, vol. 42, no. 3, pp. 145-175, 2001.

[8] J. Li, N. Allinson, D. Tao, and X. Li, "Multitraining Support Vector Machine for Image Retrieval," *IEEE Transaction on Image Processing*, vol. 15, no. 11, pp. 3597-3601, 2006.

[9] D. Tao, X. Li, and J. M. Stephen, "Negative Samples Analysis in Relevance Feedback," *IEEE Transaction on Knowledge and Data Engineering*, vol. 19, no. 4, pp. 568-580, 2007.

[10] W. Bian, and D. Tao, "Biased Discriminant Euclidean Embedding for Content-Based Image Retrieval," *IEEE Transaction on Image Processing*, vol. 19, no. 2, pp. 545-553, 2010.

[11] N. W. U. D. Chathurani, S. Geva, and V. Chandran, "Conversion of an Image to a Document Using Grid-based Decomposition For Efficient Content-Based Image Retrieval." *International Journal of Information Science and Intelligent Systems*, vol.4, no 3, 2015.

[12] G. Demir, and A. Selim, "Scene Classification Using Bag-of Region Representations," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-8, 2007.

[13] V. Takala, T. Ahoen, and M. Pietikainen, "Block-based methods for image retrieval using local binary patterns," *in Proc. of the 14th Scandinavian Conference on Image Analysis*, pp. 882-891, 2005.

[14] P. S. Hiremath, and J. Pujari, "Content based image retrieval based on color, texture and shape features," *in Proc. 15th International Conference on Advance Computing and Communications*, pp. 780-784, 2007.

[15] X. Yuan, J. Yu, Z. Qin, and T. Wan, "A SIFT-LBP image retrieval model based on bag-of-features," *in Proc. IEEE ICIP 18th International Conference on Image Processing*, pp. 1061-1064, 2011.

[16] M. H. Saad, H. I. Saleh, H. Konbor, and M. Ashour, "Image retrieval based on integration between YCbCr color histogram and shape feature," *in Proc. ICENCO 7th International Computer Engineering Conference*, pp. 97-102, 2011.

[17] N. S. Mansoori, M. Nejati, P. Razzaghi, and S. Samavi, "Bag of visual words approach for image retrieval using color information," *in Proc. ICEE 21st Iranian Conference on Electrical Engineering*, pp. 1-6, 2013.

[18] J. Li, J. Z. Wang, and G. Wiederhold, "IRM: Integrated Region Matching for Image Retrieval," *in Proc. ACM Multimedia 00'*, pp. 147-156, 2000.

[19] Y. Chen, and J. Wang, "A region-based fuzzy feature matching approach to content-based image retrieval," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1252-1267, 2002.

[20] P. S. Hiremath, and J. Pujari, "Content Based Image Retrieval Using Color Boosted Salient Points and Shape Features of an Image," i*n Proc. International Journal of Image Processing*, pp. 10- 17, 2008.

[21] M. Banerjee, M. Kundu, and P. Maji, "Content-based image retrieval using visually significant point features," *Journal of Fuzzy Sets and Systems,* vol. 160, no. 23, 2009.

[22] M. Chowdhury, S. Das, and M. Kundu, "Interactive content based image retrieval using Ripplet transform and fuzzy relevance feedback," *in Proc. 1st Indo-Japan Conference*, pp. 243-251, 2012.

[23] T. Chappell, S. Geva, and G. Zuccon, "Approximate Nearest-Neighbour Search with Inverted Signature Slice Lists," in: Pro. of the 37th European Conference on IR Research, 2015.