International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:3, No:6, 2009

Text Retrieval Relevance Feedback Techniques for Bag of Words Model in CBIR

Nhu Van NGUYEN, Jean-Marc OGIER, Salvatore TABBONE and Alain BOUCHER

***Abstract*—**The state-of-the-art Bag of Words model in Content-Based Image Retrieval has been used for years but the relevance feedback strategies for this model are not fully investigated. Inspired from text retrieval, the Bag of Words model has the ability to use the wealth of knowledge and practices available in text retrieval. We study and experiment the relevance feedback model in text retrieval for adapting it to image retrieval. The experiments show that the techniques from text retrieval give good results for image retrieval and that further improvements is possible.

***Keywords*—** Relevance feedback, bag of words model, probabilistic model, vector space model, image retrieval.

I. Introduction

ALL Content-Based Image Retrieval (CBIR) systems have a limited performance. This is mostly due to two factors: the impossibility to fully express all the user intent into a simply query for retrieval and the difference between the user interpretation and the computer description for an image, which is also called the semantic gap. Some information is lost because we use low level features to representing the image content seen through a high-level interpretation by the user. During mostly the last 10 years, relevance feedback (RF) techniques have been applied in CBIR to cope these problems [4], [7], [18]. The RF techniques are taken mostly from the traditional text retrieval domain in which they have shown very significant improvements in performance. And now, RF is becoming an essential component for a CBIR system.

The retrieval model for text retrieval is different from image retrieval. Images are represented by low-level features like color, texture, region-based, shape-based descriptors or salient point descriptor and more. Meanwhile, text documents are represented by the term weight features. Under the assumption that high-level concepts can be captured by low-level features, the RF techniques try to establish the link between these two levels of perception. Doing so, we have to modify the RF techniques from text retrieval for adapting them for image retrieval. Consequently, the result does not have the same performance as in text retrieval. For example, sometimes we have to convert image features into text features [14], [20] for using RF techniques. Sometimes we create a new RF technique for image retrieval only (for example re-weighting the similarity function).

Recently, the state-of-the-art model for CBIR became the bag of words (BoW) model which is also adopted from text retrieval. The image retrieval process is the same as for text retrieval, image features extracted from the images being considered as "words" [21]-[24]. Therefore, we believe that RF for CBIR has the potential for improving image retrieval performance by using the wealth of knowledge and practices available in text retrieval.

Our contributions in this paper are the study and the application of text retrieval models and theirs RF techniques combined with the BoW model for image retrieval. In particular, 2 models are further examined. The first model is the vector space model proposed by Salton et al. [1]. Three text retrieval ranking functions with the Rocchio's RF are applied for CBIR. This is a classical model with a powerful RF technique in text retrieval. All research of BoW model in CBIR use the vector space model for retrieval process. The second model involves the probabilistic model developed in the 1970s and 1980s by Robertson and Jones [25] which is widely used in text retrieval. This model is the state-of-the-art model in text retrieval which is designed for RF of BoW model in text retrieval. We demonstrate that both models are well-appropriate for CBIR.

The results show that text retrieval models can be easily applied in image retrieval. These RF techniques improve well the retrieval performance.

II. Related work

Relevance Feedback (RF) technique is an interactive strategy effective to improve the accuracy of information retrieval systems, in particular here CBIR systems [10], [17]. RF has a short-term memory. It is adapting the retrieval process for a specific user and a specific query. Short-term memory means that the user is searching by submitting a query, then sees some results and interacts in order to modify them by asking the system to change the weights of parameters

Nhu Van NGUYEN is with L3i – University of La Rochelle, 17042 La Rochelle cedex 1, France ; Institut de la Francophonie pour l'Informatique, MSI, UMI 209, Hanoi, Vietnam ; IRD, UMI 209, UMMISCO, IRD France Nord, Bondy, F-93143, France ; (phone: 06.26.15.112.58; fax: 05.46.45.82.42; e-mail: nnvan@ifi.edu.vn).

Jean-Marc OGIER is with L3i – University of La Rochelle, 17042 La Rochelle cedex 1, France (phone: 05.46.45.82.15; fax: 05.46.45.82.42; e-mail: jean-marc.ogier@univ-lr.fr).

Alain BOUCHER is with Institut de la Francophonie pour l'Informatique, MSI, UMI 209, Hanoi, Vietnam ; IRD, UMI 209, UMMISCO, IRD France Nord, Bondy, F-93143, France ; (e-mail: alain.boucher@auf.org).

Salvatore TABBONE is with QGAR – LORIA – University of Nancy 2, Campus scientifique - BP 239 - 54506 Vandoeuvre-les-Nancy, (e-mail: tabbone@loria.fr).

or to modify the query itself for adapting the result to the real user's intents. During that interaction time, the system can remember the results. But once it is finished, then the system cleans its memory and the next user starts from scratch.

The BoW model in Natural Language Processing is a popular method for representing documents. Each document is represented by a bag of words (which means into a vector where the order between elements is not considered). All words from the document database form a dictionary. A text is retrieved by computing the similarity between its vector of word frequencies and other documents. Researchers in computer vision are using the same idea for image representation. Images are treated as documents, and features extracted from images are considered as "words".

The BoW model is usually achieved by 3 steps: feature detection & description, codeword dictionary formation, image representation. The feature detection extracts local regions which are the candidates for "words". The simplest method for feature detection is regular grid [5], [8]. The image is segmented by some horizontal and vertical lines. A very popular method is interest point detector [5], [21]. In addition, random sampling and segmentation methods for feature detection are used in [6] and [3]. The most popular method for feature description is using the SIFT descriptor [2]. The SIFT descriptor describes each region as a 128-dimensional vector. The second step is to convert the region description vectors into a codeword dictionary. Each codeword is representing several similar regions by using a clustering method, as for example K-means clustering. The last step is to represent the image as a histogram of the codewords.

The BoW model has been researched for years. However the RF techniques for this model have not been fully investigated. To our knowledge, only few works address this issue. Rui et al. [14] used the query point movement for a look-like bag of word model, but the image feature is not the real "words". They converted the global image features such as color and texture into term-weight vector in text retrieval. Recently, [24] uses the discriminative RF for the BoW model. In their framework, the obtained RF scheme gives the most discriminative regions or keywords instead of just the important keywords for a particular class of images. They build a pseudo image from the top N discriminative keywords for the next iteration. In [19], authors present the RF technique for the bag of words model as a Multiple Instance Learning problem: finding the positive/negative words by using the MILL toolkit. In the 2 following sections we examine the 2 main models for RF in text retrieval: Vector space model and probabilistic model. Then we apply them to the BoW model for CBIR.

III. Relevance Feedback Techniques for Bag of Words

In this section, the two most popular models for RF in text retrieval are presented. In the *vector space model*, we introduce 3 ranking functions with a RF technique. In the *probabilistic model*, two ranking functions with a RF technique are presented.

A. Vector space model

The vector space model is proposed by Salton et al. [1]. Each document in a database D, as well as the query Q (Q is also a document or a set of documents), is represented as a vector of term weights:

$$\vec{d}_i = (w_{i,1}, w_{i,2} \dots w_{i,t}) \tag{1}$$

To compute the similarity between the query Q and the document D, we use the inner product between the corresponding vectors:

$$S(d_i, q) = \sum_j w_{i,j} \times w_{q,j} \tag{2}$$

Different methods for estimation $w_{i,j}$ and $w_{q,j}$ create different ranking functions for the vector space model. In this paper we investigate 2 methods: the classical TF-IDF weights and the state-of-the-art pivoted normalization weights.

The classic vector space model proposed in [1] computes weight of a term in a document as a product of the term frequency (TF) and the inverse document frequency (IDF). We get the ranking function for this method:

$$S(d_i, q) = \frac{\sum_j D_i \times D_q}{|D_i||D_q|} \tag{3}$$

with $D_k = \mathrm{tf}_k.\mathrm{idf}_k$ and $|D_i||D_q|$ the denominator needed for normalization.

Recently, this ranking function has been used for image retrieval [21]-[24].

The second ranking function is proposed in [11], [15]. Another method for document length normalization is used in the pivoted normalization weighting based document score which is:

$$S(d_j, q) = \sum_{t=1}^{n} \frac{1 + \log(1 + \log(tf_{t,D}))}{(1-s) + s\frac{dl}{avdl}} tf_{t,Q} \log\frac{N+1}{n_i} \tag{4}$$

with *dl* being the document length, *avdl* being the average document length, N being the number of documents in the collection, n_i being the number of documents that contain the term, and *s* being a constant, usually 0.20.

In [12], Singhal et al. are proposing a ranking function based on the pivoted normalization, called F2-EXP:

$$S(d_j, q) = \sum_{t=1}^{n} \frac{tf_{t,D}}{tf_{t,D} + s + s\frac{dl}{avdl}} tf_{t,Q} \left(\frac{N+1}{n_i}\right)^k \tag{5}$$

with $s = 0.5$ and $k=0.35$.

In text retrieval, the original RF process was designed for vector queries. The objective of RF is to reformulate query vector so that it is closer to the space containing the relevant documents. The optimal query maximizes similarity to relevant

documents, and minimizes similarity to irrelevant ones [13]:

$$\vec{q}_{i+1} = \alpha\vec{q}_i + \frac{\beta}{|D_r|}\sum_{\vec{d}\in D_r}\vec{d} - \frac{\gamma}{|D_n|}\sum_{\vec{d}\in D_n}\vec{d} \quad (6)$$

where D_r is the relevance set, D_n is the non-relevance set, α, β, and γ give relative weight of q, D_r, and D_n. In our experiment, the set of parameters $\alpha = \beta = \gamma = 1$ is used for the 3 ranking functions above (see (3), (4) and (5)).

The Rocchio's RF technique was used in CBIR since 1997 by Rui et al. [14]. They have converted the global image features, such as color and texture, into term-weight vectors for text retrieval to use the Rocchio's RF. The Rocchio's technique has not been yet investigated for BoW model. Doing so, we can avoid the previous conversion, because the BoW model is already using the term weight vector for representing images. As we will see that the visual words do not carry semantic like the real words in text retrieval, we do not have the same results when applying text retrieval techniques for image retrieval [26]. Moreover, in the BoW model for CBIR, images are normally small documents with less than 1000 words, and even 100 words for a simple image. For adapting to image retrieval, a manual tuning of parameters for the pivoted normalization weight is used. The value 0.20 for s in text retrieval is not appropriate for CBIR. Following our experiments, the best value obtained is 0.05.

B. Probabilistic model

In the probabilistic retrieval model, we estimate the probability of a document d_j to be relevant for a specific query q [25], [16], denoted as $P(R \mid q, d_j)$:

$$P(R \mid q, d_j) = \frac{P(d_j relevant - to - q)}{P((d_j non - relevant - to - q))} \quad (7)$$

We can represent the set of terms occurring in a document d_j as a binary vector $x = (x_1, x_2 ... x_n)$ with $x_i = 1$, if term i is present in d_j and $x_i = 0$ otherwise. Then the documents are ranked in decreasing order according to the following expression:

$$\log\frac{P(x \mid R)}{P(x \mid \overline{R})} \quad (8)$$

where R is the relevant set (positive examples) and $\overline{R}$ is the non relevant set (negative examples). $P(x \mid R)$ and $P(x \mid \overline{R})$ are the probabilities that a relevant or a non relevant item has a vector representation x. From (1), a retrieval status value can be derived [16]:

$$S(d_j, q) = \sum_{i=1}^{n}\log\frac{P(x_i \mid R)(1 - P(x_i \mid \overline{R}))}{(1 - P(x_i \mid R))P(x_i \mid \overline{R})} \quad (9)$$

Different methods for estimating $P(x \mid R)$ and $P(x \mid \overline{R})$ create different ranking functions. Although the probabilistic retrieval model is not used for BoW model in CBIR yet, it is the state-of-the-art model for text retrieval. We investigate 2 ranking functions for this model.

The Okapi weighting based document score by Sparck Jones et al. [25]:

$$S(d_j, q) = \sum_{t\in Q} W_t \frac{(k_1+1)tf_{t,D}}{k_1((1-b) + \frac{b.dl}{avdl}) + tf_{t,D}} \frac{(k_3+1)tf_{t,Q}}{k_3 + tf_{t,Q}} \quad (10)$$

with k_1 (between 1.0-2.0), b (usually 0.75) and k_3 (between 7-1000) are constants. The parameters set (k_1=1.2, b=0.75, k_3=1000) is used in the experiment. Wt is the Robertson/Sparck Jones weight of term t in the query:

$$W_t = \log\frac{(r_t + 0.5)(N - R - n_t + r_t + 0.5)}{(n_t - r_t + 0.5)(R - r_t + 0.5)} \quad (11)$$

where N is the collection size, R is the relevant set size, n_t is the number of documents in N, containing term t, r_t is the number of documents in R, containing term t. This weight is used for the RF. The initial search uses the reduced formula of the the Robertson/Sparck Jones weight (with R=0, r_t=0):

$$W_t = \log\frac{N - n_t + 0.5}{n_t + 0.5} \quad (12)$$

The second function is proposed by Singhal in [11]. The modified Okapi gives a better performance than the original Okapi as they claimed. To avoid the negative value in the logarithm factor they proposed to use the IDF factor in the pivoted normalization weigh, the Wt weight of (10) is then:

$$W_t = \frac{1 + \log(1 + \log(tf_{t,D}))}{(1-s) + s\frac{dl}{avdl}} \quad (13)$$

The probabilistic RF is based on the Robertson/Sparck Jones weight Wt. The documents are ranked basing on the relevant set R (see 11), The Wt weight of (13) does not have the relevant set R so we cannot use the modified Okapi with the probabilistic RF. The Rocchio's RF is used instead.

IV. Experiments

A. Experiment protocol

We present a total of 5 ranking functions: TF-IDF, pivoted normalization, F2-EXP, Okapi and modified Okapi. In text retrieval, the Okapi function uses the probabilistic RF and the 4 other functions use the Rocchio's RF. We propose to use the Rocchio's RF for all 5 ranking functions. This is reasonable as all ranking functions use the weight of query terms.

The Wang image database is used in our experiments[1]. It is a subset of the Corel database. This database contains 1000 images divided into 10 classes. The size of each image is 384 × 256 pixels.

For the experiments, a pseudo RF technique is used, simulating automatically the normal human interaction for RF.

[1] http://wang.ist.psu.edu

Using the existing labels for each image in the database (ground truth), the system chooses the nearest ones belonging to the query class for relevant examples (positive examples) and the nearest ones not belonging to the query class for non-relevant examples (negative examples). Following [4] one of the characteristics for RF is the small sample issue. The number of training examples is small, typically smaller than 20 per round of interaction. We believe that 20 examples per interaction is still a big number. In a real situation, it is difficult to imagine users clicking on even more than 10 examples. In our experiments, the assumption is made that only a maximum of 10 images can be selected by the user. We propose 5 strategies for the user relevance selection:

1. 5 relevant examples, 5 non-relevant examples.
2. 3 relevant examples, 3 non-relevant examples
3. 1 relevant example, 1 non relevant example
4. 5 relevant examples
5. 5 non relevance examples

To evaluate the performance of the ranking functions the precision/recall curve is used. The average precision is used for evaluation/comparison between the different RF strategies.

B. Comparison of the ranking functions

The retrieval performances of the 5 ranking functions are shown in Fig.1. The best ranking functions for image retrieval is the TF-IDF, the modified Okapi coming close in second position.

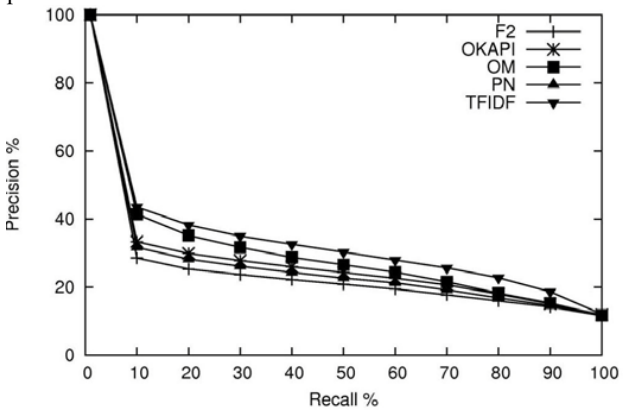

Fig. 1 Ranking function comparison for the first retrieval iteration for all classes in the Wang database (no relevance feedback). TF-IDF obtains the best results followed closely by the modified Okapi (OM).

We go further in details to analyze the retrieval performance of each ranking function by looking into each class of the database. The Wang database has 10 classes: ethnic group, beach, monument, car, dinosaur, elephant, flower, horse, mountain, food. In [9], the authors have already shown a big disparity in precision among the different classes of this database, some being very easy while some being too difficult. Our results show that TF-IDF works better than the other functions for the "easy" classes like 3, 4, 6 but it becomes much worse for the other classes. The modified Okapi (OM) has the closest performance with the TF-IDF. It works better than TF-IDF for the classes 0, 7, 8. But these good functions, TF-IDF and Okapi, suddenly result as much more worse than other functions for the difficult classes, like the example shown for class #2 in Fig. 2. This means that they work well in easy situations, but their performance decreases much faster than the other ranking functions. Depending on the type of images studied and the foreseen application, they are not necessarily the best choice.

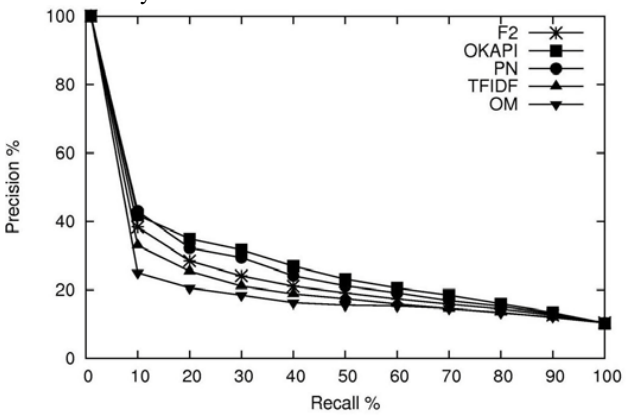

Fig. 2 The Recall/Precision for the class #2 of the Wang database. This class gets usually among some of the lowest precision result in image retrieval tests.

Following our experiments, other ranking functions can be used instead of the TF-IDF in different cases, especially when dealing with difficult databases. Among these ranking functions, the modified Okapi (OM) and the TF-IDF work best generally. Future work is needed to investigate these functions with different types of image databases.

C. Comparison of relevance selection strategies

Five relevance selection strategies are used with a maximum number of selected examples as 10. As one can expect, the experiments show that the strategy with 10 examples obtains the highest performance (see Fig 3).

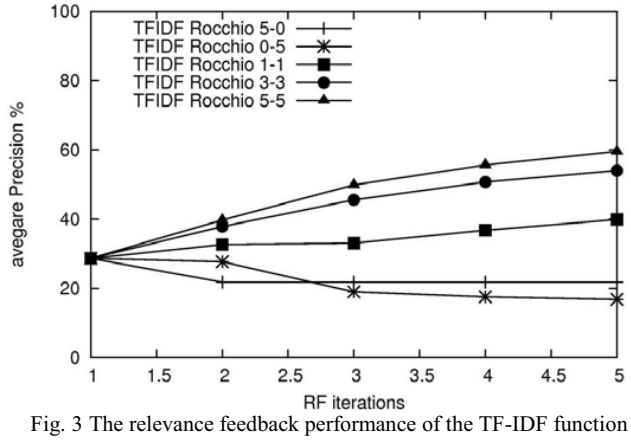

Fig. 3 The relevance feedback performance of the TF-IDF function with the Rocchio RF. It shows that more examples give better result.

The worst strategies are with 5 relevant/non-relevant examples for the Rocchio RF technique. The fact that the Rocchio RF can be used effectively depends on the presence

of both the relevant (or positive) examples and the non relevant (negative) examples. Also, using only non-relevant examples give worse result after the feedback than when using relevant examples. The probabilistic RF gives bad performance when using the strategy with the lowest number of examples (1 relevant or 1 non-relevant).

We compare in Fig. 4 the performance of RF techniques with different ranking functions. The Rocchio RF with the TF-IDF function shows the best result: +30.857% after 4 iterations, the Rocchio RF with the modified Okapi come second: +24.7608% after 4 iterations. The Okapi with probabilistic RF give the worst result: +14.542%.

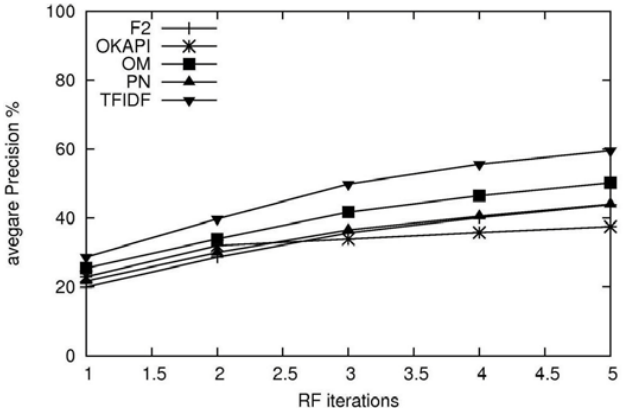

Fig. 4 Comparison on the relevance feedback performance of different strategies. The 2 best strategies are the TF-IDF function with the Rocchio RF and the Okapi function with the probabilistic RF.

Retrieval performance improvement after 4 iterations of RF for the 10 classes taken separately is shown in Fig. 5. The strategy of using 10 examples is used with 5 ranking functions.

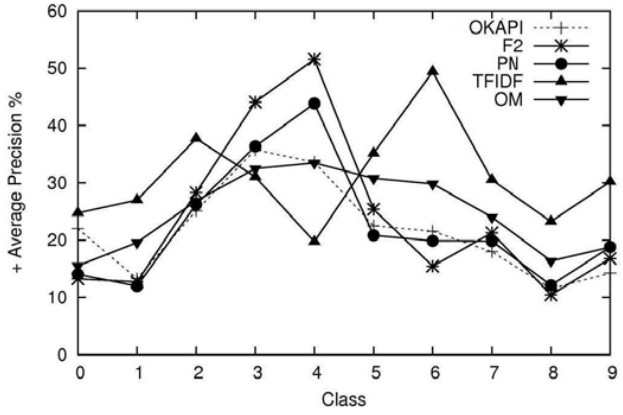

Fig. 5 Retrieval performance improvement after 4 iterations of RF for the 10 classes. The TF-IDF function with the Rocchio RF gives best result for most classes but not all.

The TF-IDF function with the Rocchio RF gives the best result for most classes, but not all. The F2-EXP and the pivoted normalization functions have a better average precision for class #4. This is reasonable because the TF-IDF has given a very good result in the 1st query for class #4 so for the next iterations of RF, the performance improves negligibly. But for the class #3, where the TF-IDF also give a best result in the 1st query among those ranking functions, the F2-EXP, the Okapi and the modified Okapi give better relevance feedback improvements. This shows that the TF-IDF with Rocchio RF give best performance in general, but there are cases where other methods give better results.

V. Conclusion

In this paper, we have investigated 2 text retrieval relevance feedback models with 5 ranking functions which were applied for content-based image retrieval. We show that techniques in text retrieval can be successfully used for CBIR. The experiments show that TF-IDF ranking function with the Rocchio relevance feedback technique gives good results, but not in all cases. In more difficult situations (difficult classes/databases with low retrieval performance) other methods are worth being investigated. As a future work, we will examine more in details some difficult situations where each function/technique can give the best results.

Acknowledgment

This project is supported in part by the ICT-Asia IDEA project from the French Ministry of Foreign Affairs (MAE), the DRI INRIA and DRI CNRS.

References

[1] G. Salton, A. Wong, and C. S. Yang, "A Vector Space Model for Automatic Indexing," *Communications of the ACM*, vol. 18, 1975, nr. 11, pp. 613–620.

[2] D. G. Lowe. Distinctive image features from scale-invariant key points. *International Journal of Computer Vision*, 60(2), 2004, pp. 91–110.

[3] K. Barnard, P. Duygulu, N. de Freitas, F. Forsyth, D. Blei, and M. Jordan. "Matching words and pictures," *Journal of Machine Learning Research 3*, 2003, pp. 1107–1135

[4] Y. Rui, T. S. Huang, M. Ortega, and S. Mehrotra. "Relevance feedback: A power tool in interactive content-based image retrieval," *IEEE Transactions on Circuits and Systems for Video Technology*, 8(5), 1998, pp. 644–655.

[5] L. Fei-Fei and P. Perona, "A Bayesian Hierarchical Model for Learning Natural Scene Categories," *Proc. of IEEE Computer Vision and Pattern Recognition*, 2005, pp. 524-531

[6] M. Vidal-Naquet and S. Ullman, "Object recognition with informative features and linear classification," *Proc. of IEEE International Conference on Computer Vision*, 2003, pp. 281-288

[7] J. Urban, J. Jose, and K. van Rijsbergen, "An adaptive approach towards content based image retrieval," *In Proceedings of the International Workshop Content-Based Multimedia Indexing*, 2003, Rennes, France.

[8] J. Vogel and B. Schiele, "On Performance Characterization and Optimization for Image Retrieval," *Proc. of European Conference on Computer Vision*, 2002, pp. 51-55

[9] A. Boucher, H. Dang, L. Le, "Classification vs recherche d'information : vers une caractérisation des bases d'images," 12èmes Rencontres de la Société Francophone de Classification (SFC), 2005, Montréal, Canada.

[10] M. Crucianu, M. Ferecatu and N. Boujemaa, "Relevance feedback for image retrieval: a short review, " *In State of the Art in Audiovisual Content-Based Retrieval*, *Information Universal Access and Interaction including Data models and Languages* (DELOS2 Report, 2004.

[11] A. Singhal. "Modern information retrieval: A brief overview," *Bulletin of the IEEE Computer Society Technical Committeeon Data Engineering*, 24(4), 2001, pp. 35-42

[12] A. Singhal, J. Choi, D. Hindle, D. Lewis, and F. Pereira. "AT&T at TREC-7," *In Proc. of the Seventh Text REtrieval Conference (TREC-7)*, NIST Special Publication, 1999, pp. 239–252.
[13] J.J. Rocchio, "Relevance feedback in information retrieval". In The SMART Retrieval System," *Experiments in Automatic Document Processing*, Prentice Hall Inc, 1971, pp. 313–323, Englewood Cliffs, NJ.
[14] Y. Rui, T. Huang and S. Mehrotra, "Content-based image retrieval with relevance eedback in Mars," *In Proceedings of the IEEE International Conference on Image Processing*, 1997, Washington, DC.
[15] H. Fang, C. Zhai, "An exploration of axiomatic approaches to information retrieval," *In Proceedings of the 28th ACM Conference on Research and Development in Information Retrieval*, 2005, pp. 480-487.
[16] N. Fuhr, "Probabilistic models in information retrieval," *The Computer Journal*, 35(3), 1992, pp 243-255.
[17] D. Heesch, S. Ruger, "Interaction models and relevance feedback in image retrieval," *Semantic-Based VisualInformation Retrieval*, 1st ed, IRM Press, pp. 160-186.
[18] Y. Rui and T. Huang, "Optimizing learning in image retrieval," *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2000, Hilton Head Island, SC.
[19] L. Le, A. Boucher, M. Thonnat, F. Bremond, "A framework for surveillance video indexing and retrieval," *Content-Based Multimedia Indexing, CBMI 2008*, 2008 pp. 338 – 345.
[20] H. Müller, W. Müller, D. Squire, S. Marchand-Maillet and T. Pun, "Strategies forpositive and negative relevance feedback in image retrieval," *In Proceedings of theInternational Conference on Pattern Recognition*, 2000, Barcelona, Spain.
[21] J. Sivic and A. Zisserman, "Efficient Visual Search for Objects in Videos," *In Proceedings of the IEEE, Special Issue on Advances in Multimedia Information Retrieval*, 96(4), 2008, pp. 548-566.
[22] J. Yang, C-W. Ngo, A. Hauptmann and Y-G. Jiang, "Evaluating Bag-of-Visual-Words Representations in Scene Classification," *ACM Multimedia Information Retrieval Workshop at ACM Multimedia 2007*, 2007, Augsburg, Germany.
[23] P. Tirilly, V. Claveau and P. Gros, "Language modeling for bag-of-visual words image categorization," *In CIVR '08: Proceedings of the2008 international conference on Content-based image and video retrieval*, 2008, pp. 249-258, New York, NY, USA.
[24] S. Karthik, C-V. Jawahar, "Discriminative Relevance Feedback With Virtual Textual Representation for Efficient Image Retrieval," *Visual Information Engineering VIE 2006. IET International Conference*, 2006, pp. 309 – 31.
[25] K. Spärck Jones, S. Walker, and S. E. Robertson, "A Probabilistic Model of Information Retrieval: Development and Comparative Experiments," *Information Processing and Management*, 36(6), 2000, pp.779-840.
[26] J. Yang, Y. G. Jiang, A. G. Hauptmann, C. W. Ngo, "Evaluating Bag-of-Visual-Words Representations in Scene Classification," *ACM SIGMM Int'l Workshop on Multimedia Information Retrieval (MIR'07)*, 2007, Augsburg, Germany,