Categorization based Relevance Feedback Search Engine for Earth Observation Images Repositories

Mihai Costache
GET, Telecom Paris
46 rue Barrault
75013 Paris, France
Email: costache@enst.fr

Henri Maitre
GET, Telecom Paris
46 rue Barrault
75013 Paris, France
Email: henri.maitre@enst.fr

Mihai Datcu
German Aerospace Center (DLR) and GET, Telecom Paris
D-82234 Oberfaffenhaffen
Germany
Email: mihai.datcu@dlr.de

***Abstract*— Presently Earth Observation (EO) satellites acquire huge volumes of high resolution images very much over-passing the capacity of the users to access the information content of the acquired data. Thus, in addition to the existing methods for EO data and information extraction, new methods and tools are needed to explore and help to discover the information hidden in large EO image repositories. This article presents a categorisation based Relevance Feedback (RF) search engine for EO images repositories The developed method is presented as well results obtained for a SPOT5 satellite image database.**

I. Introduction

Presently Earth Observation (EO) satellites acquire huge volumes of high resolution images, very much over-passing the capacity of the users to access the information content of the acquired data. In addition to the existing methods for EO, data and information extraction, are needed new methods and tools to explore and help to discover the information hidden in large EO image repositories.

In this article we present a concept of search engine using RF based on Support Vector Machine (SVM) classifier. The convergence speed and precision of the queries are enhanced due to the knowledge stored by the assigned categories.

The proposed search engine supports users to search images of interest in a large repository, by visual ranking of automated suggested images, grouped in classes of relevance.

The system diagram is presented in Fig. 1. It is composed of four blocks: a SPOT5 satellite image database, a primitive feature extraction block, a classification component and a Graphical User Interface (GUI) which allows the human - machine communication.

The link between the visual similarity of images and their information content is a complicated matter. Solutions investigated and proposed are the use of Quadrature Mirror Filters (QMF) coefficients to derive image descriptors for further similarity computation.

The SVM methods have been used for classification and learning in Content Based Image Retrieval (CBIR) systems with impressive results. These methods are kernel-based learning machines and they have the capability of performing high accuracy classification for different type of data starting only from small number of examples for each class.

Searching by category generation (relevant/irrelevant) is equivalent with performing a two classes classification. Thus SVM methods are used in the process of RF in the part of selecting the next images which are shown to the user in order to proceed to the next step.

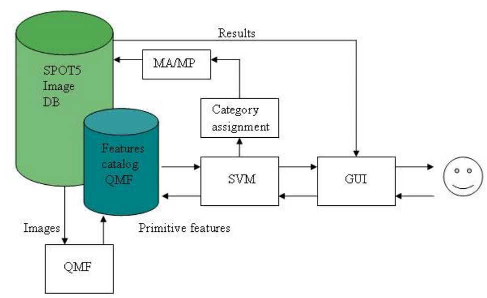

Fig. 1. System diagram. The system is composed of SPOT5 image database, feature extraction block, classifier and communication between user and machine. Based on the user and classifier decisions, categories are generated.

Based on the user's previous decisions the system can employ SVM-one class or SVM-two classes.

One class SVM is limited in capturing the relevance of the user feedback. SVM-one class performs well in the first steps as it provides as results images which are comprised within the sphere constructed around the relevant examples. Once the first irrelevant examples are given the precision decrease as it takes additional iterations to learn the system the irrelevant category.

A modification of the first iteration step is done in order to use the SVM binary classifier described in subsection III-B. Using two classes SVM classifier along with category assignment provides the tool to better delimitate the relevant images within the database as this ensures diversity of the images presented to the user.

Experimental results on a data base of 600 SPOT5 images comprising six classes manually obtained with comments and explanation are presented.

II. Primitive features and database images

As primitive features for our study the QMF sub-band decomposition was selected. QMF filters which are a pair of non separable filters are used to decompose the signal in frequency domain into low-pass and high-pass frequency bands. They allow a perfect reconstruction of the original signal.

In the bi-dimensional satellite image case the filters are obtained by convoluting the filters in one dimension.

The decomposition is applied recursively which implies octave band split of the original spectrum. Taking into consideration that visual human receptors are spaced at octave distances [2] the utilisation of QMF bank filters is obvious for modelating the human visual system.

The standard deviation and of each output and the mean value of the QMF coefficients are computed and stored as features. For our purpose we use a two level decomposition thus giving us a collection of eight features.

The test database was obtained by manually cropping SPOT5 scenes (of resolution 5m/pixel) into small images of dimension 64x64 pixels. A total set of 600 small images was obtained, divided in 6 classes as fallows: clouds, sea, desert, city, forest and fields (100 images/class). The 64x64 gray level images were not normalised before feature extraction. However in order to employ SVM the feature vectors were individually normalised with mean 0 and variance 1.

III. Relevance Feedback

Of high importance in retrieval systems is the problem of RF, which is a form of query-free retrieval and which allows that documents, in our case satellite images, to be retrieved according to a similarity measure to a given document.

The goal of the system is that it retrieves documents based on the user's judgement. Based on his interest the user labels the documents as relevant or irelevant.

A RF algorithm starts with a query for document (i.e. satellite image), in the action of labelling which is performed by the user, and a retrieval method which, based on the user's judgement, tries to return the most pertinent documents from the data base.

The RF algorithms employed in CBIR were developed from classical algorithms implemented for textual document retrieval systems. They can be split in two categories as fallows:

1) query point movement algorithms and
2) re-weighting algorithms

While the query point movement methods [1], [10] try to move the query vector far from non-relevant documents and towards relevant documents, the re-weighting schemes [5], [8] try to enhance the importance of the features which help in retrieving relevant documents and in the same time to minimise the importance of those which are not efficient for this task.

Recently SVM was studied and used to enhance CBIR performances.

A. Relevance Feedback based on SVM

In this section a general presentation of the active learning problem in relevance feedback which employs SVM is done based on the work presented in [3], [4].

SVMs techniques were previously applied for text classification, object recognition and recently in CBIR with good results. The powerfull tool which SVM represents for classification problem with an active learning provides a good solution for RF.

The RF based SVM implementations can be seen as a loop process with two steps. The first is the learning step where using a small number of examples the systems learns how to separates images which belong to different classes. For the case of RF there are only two classes: the relevant images class and irrelevant images class. In the second step the classification of the remained documents is made in order to determine the next images which user has to label as relevant/irrelevant.

B. SVM binary classification

In the case of RF in CBIR systems the user is interested in finding images with similar content as his target. This means that they can be considered as a class of relevant images while all the other images will be considered as the class of irrelevant ones. This comes to the problem of binary classification. The relevant class is labelled by $\{+1\}$ while the irrelevant one by $\{-1\}$. A short description of the binary classification using SVM technique is done bellow.

Having a data base of images in which each image is described by a vector x_i in M-dimensional Vector Space Model and each image is labelled by y_i, where $y_i \in \{-1, +1\}$ indicates the the class of the image. Because usually in the M-dimensional space the problem of separability is not linear a transformation, $\Phi(x_i)$, into another space where a linear discrimination between examples can be made is done. The classification is done with the help of the kernel function defined as

$$K(x_i, x_j) = < \Phi(x_i), \Phi(x_j) > \quad (1)$$

without having to know the expression of function Φ.

The problem is to maximise the margin area between the support vectors of the two classes. After this optimisation phase the separation surface is determined as the middle plane which is passing through the margin area. The decision function for an image described by primitive feature vector x_i is

$$g(x_i) = sign(f(x_i)) \quad (2)$$

where

$$f(x_i) = \sum_{l=1}^{n} \alpha_l y_l K(x_l, x_i) + b \quad (3)$$

where $f(x_i)$ represents the distance from the point x_i to the separation plane.

If the sign is positive then x_i belongs to the class labelled by $\{+1\}$ if not it belongs to the class labelled by $\{-1\}$. The parameters w, b and α_i are determined in the training step while the next step is the classification step. The number n in Eq. 3 represents the number of support vectors in the model.

Concerning the kernel type for our implementation we are using a Gaussian kernel ($K(x_i, x_j) = \exp(-\gamma||x_i - x_j||^2)$) with the scale factor γ = 0.5.

C. Relevance Feedback based on SVM protocol

Based on RF evaluation algorithms presented in literature [3], [6], [9], [11] we developed our version which is presented in the following subsection. The process starts with a query phase in which the user is selecting among a randomly machine generated image sequence one which best describes his interest. The selected image is then used by the system for retrieval purposes in the following way: based on a measure of similarity (e.g. Euclidean distance) the system performs a ranking of the images and returns the top N images. The similarity measure can generate only relevant images for the user. This would make necessary the use of SVM-one class which is desired to be avoided as after some iterations the performances suffer from accumulating too much relevant information as it will be shown later in this article.

To overcome this, a change is performed in the first retrieval step: the system performs a ranking of the images and returns the top N-m images and the bottom m.

The number N of images shown at each iteration to the user by means of the implemented GUI is equal to 8 and the number m is equal to 2.

The returned images are then labelled by the user as relevant/irrelevant and are used together with the previous annotated collection to train a SVM classifier. SVM determines based on the leaning collection a separation surface between available examples. The systems performs again a ranking but this time, based on the absolute value of the distance function *f*, in ascending order. The top N images with the smallest absolute value of distance function *f* are selected to be shown to the user at the next step. They are the most ambiguous (MA) images as they have a very small decision function value. In this way one step after another the separation surface between the relevant and irrelevant images is better traced.

While the RF process is based on annotating MA images returned by the system, the RF performances are obtained by evaluating the images with high positive distance function *f*.

They are the most pertinent (MP) images belonging to the class of interest to the user as they are situated far from the ambiguous area on the positive side of the separation surface.

An exemplification is shown in the Fig.2

D. Evaluation measure

In classical Information Retrieval Systems the performances evaluation are done by employing different measures such as Precision - Recall curves or Mean precision plots [7].

Precision-Recall curves are considered as a good tool to evaluate the properties of a retrieval system.

Let us denote the retrieved images by A and the relevant ones by B. The precision P is defined as the fraction of the retrieved images which are relevant and the recall R as the fraction of relevant images which have been retrieved:

$$P = \frac{\mid A \bigcap B \mid}{\mid A \mid} \quad and \quad R = \frac{\mid A \bigcap B \mid}{\mid B \mid} \tag{4}$$

A precision of 1 represents that all retrieved images are relevant to the user, while a recall value of 1 indicates that all the relevant images have been retrieved. An ideal system should retrieve only relevant documents no matter the recall value. It is desirable that Precision to be equal to unit for all recall values.

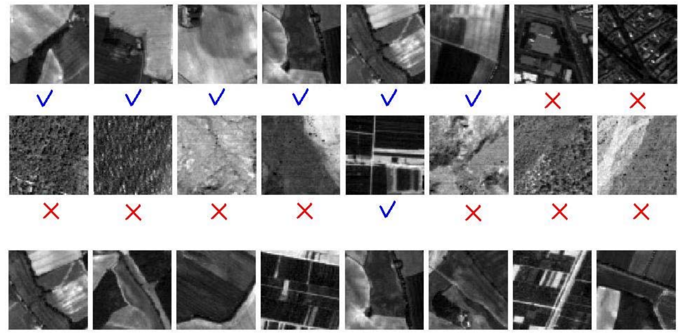

Fig. 2. RF process example. A query (top left image) is made by the user and in successive RF steps, from top to bottom, the systems returns to the user images based on the described protocol and the user is labelling them according to his target interest as relevant/irrelevant. The relevant images are images which represent fields. All other images are labelled as irrelevant. The final line represents the most pertinent (MP) images within the database after two RF steps.

Another measure used in evaluating a RF retrieval system is the mean precision. It is computed as the mean value of all precision values obtained after each iteration using the following formula:

$$p = \frac{1}{N} \sum_{i=1}^{N} P(i). \tag{5}$$

where the mean precision represents the capability of the RF system to retrieve top-ranked images which best describe the user's interest.

E. Evaluation protocol

In order to evaluate from statistical point of view the performances of the implemented system the human decision is simulated taking into account that the database has been built with human supervision and thus the class for each image is known. Taking this into consideration, each image in the database is considered as an initial query and 9 RF steps are simulated by automatic annotations of the retrieved images. The obtained results after each iteration are saved and Precision - Recall curves are computed. In order to obtain an overall information after each RF step mean values for Precision are computed over the 600 input images.

F. Results

Taking into account the Evaluation Protocol presented in the previous section Precision - Recall and mean precision curves are obtained.

1) Precision - Recall curve: Figure 3 presents a classical Precision - Recall curve for the first case after 1, 5 and 9 RF steps. The retrieval performances are affected after 9 steps as too much relevant information was accumulated. The same curves in the modified RF version are plotted in Fig. 4. It can be seen the increase in the value of precision with the number of iterations. After 9 RF steps the system returns quite a few of irrelevant images (high precision) while in the same time is well identifying the relevant images in the database (high recall). In general there is a trade-off between the values of Precision and Recall meaning that while the Recall is improved the Precision is altered. The point on the curve can be considered as the one which is closer to the ideal case when both the Precision and the Recall are equal to 1.

Taking this into consideration it can be seen that the above defined trade-off point is converging to the ideal one while the RF process advances.

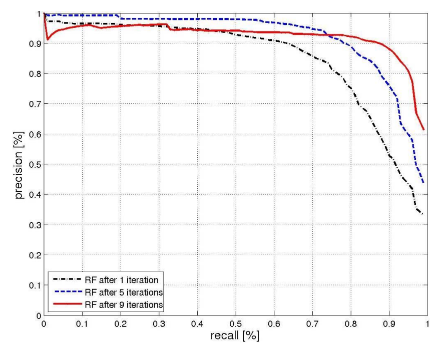

Fig. 3. Precision - Recall after 1, 5 and 9 RF steps in the first case. The retrieval properties are affected after 9 RF steps due to too much relevant information.

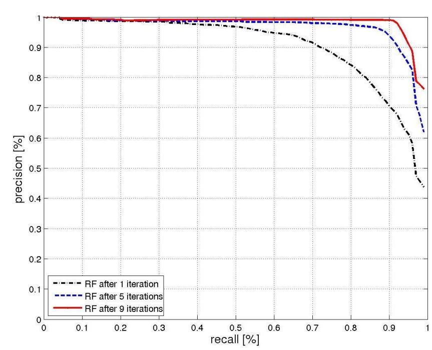

Fig. 4. Precision - Recall after 1, 5 and 9 RF steps in the improved version. The retrieval properties improve with the number of iterations. The initial change ensures diversity in the learning process.

2) Mean precision: In Fig. 5 the mean precision for each iteration step is plotted. Each new RF step brings new information into the systems and in this way the separation plane is better traced which determines the increase in the mean precision. For the first case it can be seen a decrease in the mean precision value but after 2 RF steps the system improves its performances. In the changed case the mean precision is increasing with the number of iterations and overall it can be seen that the performances are better than in the classical case.

IV. Conclusion

The present study describes a concept of a search engine which allows image categorisation based RF for EO images

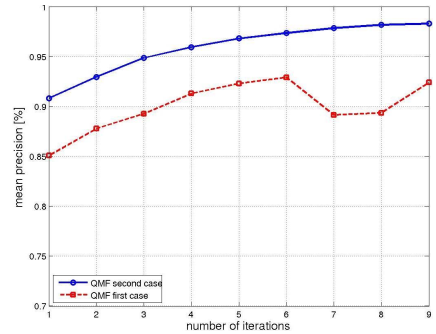

Fig. 5. Mean precision for QMF for the two cases

repositories. The developed method based on SVM is presented as well results obtained for a SPOT5 satellite image database. As well an improved version of the first RF iteration step was introduced in order to allow the SMV classifier to correctly learn the generated categories and to avoid the accumulation of too much relevant information which as it was shown determines a decrease in retrieval performances.

Acknowledgment

The work was performed within The CNES/DLR/ENST Competence Centre on Information Extraction and Image Understanding for Earth Observation

SPOT5 images have been provided by "Centre National d'Etudes Spatiales" (CNES).

References

[1] I. J. Cox and M. L. Miller and T. P. Minka and T. V. Papathomas and P. N. Yianilos, The Bayesian Image Retrieval System, PicHunter. Theory, Implementation and Psychophysical Experiments, IEEE Transactions on Image Processing, 2000, 20, 1, 20-37, January

[2] J. G. Daugman, Entropy reduction and decorrelation in visual coding by oriented neural receptive fields, I.E.E.E. Transactions on Biomedical Engineering, 36, 1, 1989

[3] M. Ferecatu and M. Crucianu and N. Boujemaa, Reducing the redundancy in the Selection of samples for SVM-based Relevance Feedback, INRIA, No. 5258, 5 May, 2004

[4] P. Hong and Q. Tian and T. S. Huang, Incorporate Support Vector Machines to Content-Based Image Retrieval Feedback, IEEE International Conference on Image Processing, Vancouver, Canada, 3, 750-753, Septembre, 2000

[5] Y. Ishikawa and R. Subramanya and C. Faloutsos, MindReader: Query Databases Through Multiple Examples, Proc. of the 24th VLDB Conference, New York, August 24-27, 1998

[6] T. Onoda and H. Murata and S. Yamada, Relevance feedback document retrieval using support vector machines, Active Mining: Second International Workshop, AM 2003, Maebashi, Japan, 2003

[7] G. Salton and M. J. McGill, Introduction to Modern Information Retrieval, New-York: Mcgraw-Hill Book Company, 1982

[8] S. Sclaroff and L. Taycher and M. L. Cascia, ImageRover: a content-based image browser for the World Wide Web, Proc. of IEEE Workshop on Content-based Access of Image and Video Libraries, 2-9, 1997

[9] S. Thong and Chang E., Support Vector Machine Active Learning for Image Retrieval, Proc. ACM Multimedia, 2001, 5, 107-118

[10] N. Vasconcelos and A. Lippman, A Bayesian Framework for Content-Based Indexing and Retrieval, Data Compression Conference, 580, 1998

[11] J. Yang and Q. Li and Y. Zhuang, Image Retrieval and Relevance Feedback using Peer Indexing, Proc. of IEEE Int'l Conf. on Multimedia and Expo, Lausane, Switzerland, 2002