A NOVEL RELEVANCE FEEDBACK PROCEDURE BASED ON LOGISTIC REGRESSION AND OWA OPERATOR FOR CONTENT-BASED IMAGE RETRIEVAL SYSTEM

P. Zuccarello, E. de Ves
Departamento de Informtica, University de Valencia, Avda. Vicente Andrés Estellés, 1. 46100-Burjasot, Valencia, Spain

T. Leon, G. Ayala
Department of Statistics and Operations Research, University of Valencia, Valencia, Spain

J. Domingo
Institut of Robotics, University of Valencia, Valencia, Spain

Keywords: Visual information retrieval,relevance feedback,logistic regresion.

Abstract: This paper presents a new algorithm for content based retrieval systems in large databases. The objective of these systems is to find the images which are as similar as possible to a user query from those contained in the global image database without using textual annotations attached to the images. The procedure proposed here to address this problem is based on logistic regression model: the algorithm considers the probability of an image to belong to the set of those desired by the user. In this work a relevance proabaility $\pi(I)$ is a quantity wich reflects the estimate of the relevance of the image I with respect to the user's preferences. The problem of the small sample size with respect to the number of features is solved by adjusting several partial linear models and combining its relevance probabilitis by means of an ordered averaged weighted operator. Experimental results are shown to evaluate the method on a large image database in term of the average number of iterations needed to find a target image.

1 INTRODUCTION

The increasing amount of information available in todays world raises the need to retrieve relevant data efficiently. Unlike text-based retrieval, where key words are successfully used to index documents, content-based image retrieval poses up-front the fundamental questions of how to extract useful image features and how to use them for intuitive retrieval (Smeulders et al., 2000). The main drawback of textual image retrieval systems, that is, the annotator dependency, would be overcome in pure CBIR systems.

Image features are a key aspect of any CBIR system. A general classification can be made: low level features (color, texture and shape) and high level features (usually obtained by combining low level features in a reasonably predefined model). High level features have a strong dependency on the application domain, therefore they are not usually suitable for general purpose systems. This is the reason why one of the most important and developed research activities in this field has been the extraction of good low level image descriptors. Obviously, there is an important gap between these features and human perception (a semantic gap). For this reason, different methods (mostly iterative procedures) have been proposed to deal with the semantic gap (Rui et al., 1998). In most cases the idea underlying these methods is to integrate the information provided by the user into the decision process. This way, the user is in charge of guiding the search by indicating his/her preferences, desires and requirements to the system. The basic idea is rather simple: the system displays a set of images (resulting from a previous search); the user selects the images that are relevant (desired images) and rejects those which are not (images to avoid) according to his/her particular criterion; the system then learns from these training examples to achieve an improved performance in the next run. The process goes on iteratively until the user is satisfied. This kind of procedures are called relevance feedback algorithms (Zhou and Huang, 2003), (de Ves et al., 2006).

A query can be seen as an expression of an information need to be satisfied. Any CBIR system aims

Zuccarello P., de Ves E., Leon T., Ayala G. and Domingo J. (2007).
A NOVEL RELEVANCE FEEDBACK PROCEDURE BASED ON LOGISTIC REGRESSION AND OWA OPERATOR FOR CONTENT-BASED IMAGE RETRIEVAL SYSTEM.
In *Proceedings of the Second International Conference on Computer Vision Theory and Applications - IU/MTSV*, pages 167-172

at finding images relevant to a query and thus to the information need expressed by the query. The relationship between any image in the database and a particular query can be expressed by a relevance value. This relevance value relies on the user-perceived satisfaction of his/her information need. The relevance value can be interpreted as a mathematical probability (a relevance probability). The notion of relevance probability is not unique because different interpretations have been given by different authors. In this paper a relevance probability $\pi(I)$ is a quantity which reflects the estimation of the relevance of the image I with respect to the user's information needs. Initially, every image in the database is equally likely, but as more information on the user's preferences becomes available, the probability measure concentrates on a subset of the database. The iterative relevance feedback scheme proposed in the present paper is based on logistic regression analysis for ranking a set of images in decreasing order of their evaluated relevance probabilities.

Logistic regression is based on the construction of a linear model whose inputs, in our case, will be the image characteristics extracted from a certain image I and whose output is a function of the relevance probability of the image in the query $\pi(I)$. In logistic regression analysis, one of the key features to be established is the order of the model to be adjusted. The order of the model must be in accordance with the reasonable amount of feedback images requested from the user. For example, it is not reasonable for the user to select 40 images in each iteration; a feedback of 5/10 images would be acceptable. This requirement leads us to group the image features into n smaller subsets. The outcome of this strategy is that n smaller regression models must be adjusted: each sub-model will produce a different relevance probability $\pi_k(I)$ $(k = 1 \dots n)$. We then face to the question of how to combine the $\pi_k(I)$ in order to rank the database according to the user's preferences. OWA (*ordered weighted averaging*) operators which were introduced by Yager in 1988 (Yager, 1988) provides a consistent and versatile way of aggregating multiple inputs into one single output.

Section 2 explains the logistic regression approach to the problem. Next, in section 3 the aggregation operators used in our work are introduced. Section 4 describes the low level features extracted from the images and used to retrieve them. An crucial part of this work, the proposed algorithm, is described in detail in section 5. After that, in section 6 we present experimental results which evaluate the performance of our technique using real-world data. Finally, in section 7 we extract conclusions and point to further work.

2 LOGISTIC REGRESSION MODEL

At each iteration, a sample is evaluated by the user selecting two sets of images: the examples or positive images and the counter-examples or negative images. Let us consider the (random) variable Y giving the user evaluation where $Y = 1$ means that the image is positively evaluated and $Y = 0$ means a negative evaluation.

Each image in the database has been previously described by using low level features in such a way that the j-th image has the k-dimensional feature vector x_j associated. Our data will consist of (x_j, y_j), with $j = 1, \dots, k$ where x_j is the feature vector and y_j the user evaluation (1= positive and 0= negative). The image feature vector x is known for any image and we intend to predict the associated value of Y. The natural framework for this problem is the generalized linear model. In this paper, we have used a logistic regression where $P(Y = 1 \mid x)$ i.e. the probability that $Y = 1$ (the user evaluates the image positively) given the feature vector x, is related with the systematic part of the model (a linear combination of the feature vector) by means of the logit function. Generalized linear models (GLMs) extend ordinary regression models to encompass non-normal response distributions and modeling functions of the mean. Most statistical software has the facility to fit GLMs. Logistic regression is the most important model for categorical response data. Logistic regression models are also called *logit* models. They have been successfully used in many different areas including business applications and genetics. For a binary response variable Y and p explanatory variables $X_1, \dots, X_p$, the model for $\pi(x) = P(Y = 1 \mid x)$ at values $x = (x_1, \dots, x_p)$ of predictors is

$$logit[\pi(x)] = \alpha + \beta_1 x_1 + \dots + \beta_p x_p \qquad (1)$$

where $logit[\pi(x)] = \ln \frac{\pi(x)}{1-\pi(x)}$. The model can also be stated directly specifying $\pi(x)$ as

$$\pi(x) = \frac{\exp(\alpha + \beta_1 x_1 + \dots + \beta_p x_p)}{1 + \exp(\alpha + \beta_1 x_1 + \dots + \beta_p x_p)}. \qquad (2)$$

The parameter β_i refers to the effect of x_i on the log odds that $Y = 1$, controlling the other x_j. The model parameters are obtained by maximizing the *likelihood equations*.

In the first steps of the procedure, we have a major difficulty when having to adjust a global regression model in which we take the whole set of variables into account, because the number of images (the number of positive plus negative images chosen by the user) is typically smaller than the number of characteristics. In this case, the regression model adjusted has as many parameters as the number of datum and many relevant variables could be not considered. On the other hand it is not realistic to ask the user to make a great number of positive and negative selections from the very beginning; therefore we think that the difficulty cannot be avoided in this way. In order to solve this problem, our proposal is to adjust different smaller regression models: each model considers only a subset of variables consisting of semantically related characteristics of the image. Consequently, each sub-model will associate a different relevance probability to a given image x, and we face the question of how to combine them in order to rank the database according to the user's preferences. We can see this question as an information fusion problem.

3 AGGREGATING THE RELEVANCE PROBABILITIES

Let us denote as $\pi_1(x), \pi_2(x), \ldots, \pi_n(x)$ the different relevance probabilities associated with a given image x. Each one of them has been obtained separately by using different regression models and we need to associate a final probability $\pi(x)$ by aggregating the information provided by each $\pi_j(x)$, $(j = 1 \ldots n)$. Mathematical aggregation operators transform a finite number of inputs into a single output and play an important role in image retrieval. In (Stejic et al., 2005)the authors compare the effect of 67 operators applied to the problem of computing the overall image similarity, given a collection of individual feature similarities. Their results show how important for retrieval performance the choice of the aggregation operator is. We have not used any of the 67 operators reviewed. Instead, we decided to use the so-called ordered weighted averaged (OWA) operators (Yager, 1988) since then they have been successfully applied in different areas such as decision making, expert systems, neural networks, fuzzy systems and control, etc. An OWA operator of dimension n is a mapping $f : \Re^n \to \Re$ with an associated weighting vector $W = (w_1, \ldots, w_n)$ such that $\sum_{j=1}^n w_j = 1$ and where $f(a_1, \ldots, a_n) = \sum_{j=1}^n w_j b_j$ where b_j is the j-th largest element of the collection of aggregated objects $a_1, \ldots, a_n$. The particular cases shown in table 1 can better illustrate the idea underlying OWA operators.

Notice that no weight is associated with any particular input; instead, the relative magnitude of the input decides which weight corresponds to each input. In our application, the inputs are relevance probabilities and this property is very interesting because we do not know, a priori, which set of visual descriptors will provide us with the *best* information.

Table 1: Illustrating examples of OWA aggregation values.

W	$f(a_1, \ldots, a_n)$
$(1, 0, \ldots, 0)$	$\max_i a_i$
$(0, 0, \ldots, 1)$	$\min_i a_i$
$(\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n})$	$\frac{1}{n} \sum_{j=i}^n a_i$.

As OWA operators are bounded by the max and min operators, Yager introduced a measure called *orness* to characterize the degree to which the aggregation is like an *or* (max) operation:

$$\text{orness}(W) = \frac{1}{n-1} \sum_{i=1}^{n} (n-i) w_i. \tag{3}$$

This author also introduced the concept of *dispersion* or *entropy* associated with a weighting vector:

$$Disp(W) = \sum_{i=1}^{n} w_i \ln w_i. \tag{4}$$

$Disp(W)$ tries to reflect how much of the information in the arguments is used during an aggregation based on W.

Clearly, the vector of weights W can be pre-fixed, but a number of approaches have also been suggested for determining it according to different criteria. One of the first methods developed was proposed by O'Hagan (O'Hagan, 1988). It provides us with the vector of weights for a given level of orness (optimism) which maximizes their entropy:

$$W = argmax \sum_{i=1}^{n} w_i \ln w_i$$

$$\text{subject to} \begin{cases} \alpha = \frac{1}{n-1} \sum_{i=1}^n (n-i) w_i, \\ \sum_{i=1}^n w_i = 1, w_i \in [0, 1]. \end{cases}$$

This problem is not computationally easy to solve. Fuller and Majlender (Fuller and Majlender, 2003) have obtained the analytical expression of the maximum entropy weights.

Figure 1 shows the aggregation of weights for $n = 10$ obtained with the above-mentioned method for orness value $\alpha \in [0.3, 0.7]$. In this work, the aggregation weights have been computed by using this method.

4 VISUAL FEATURES

This section deals with the low level features the system uses for predicting human judgment of image

is precisely to capture that notion of similarity that each user has, which can also change between different queries. Consequently, the valid criterion of similarity appears to be the user's opinion. This would have introduced an external variable into the experiment that would have masked the main goal: an objective evaluation of the system as such. That is why we have chosen to use an approach in which a given image has to be found. The search is considered successful if the image is ranked within the first 16. This number is arbitrary but we have checked that 16 images shown side by side is a reasonable number to localize a particular one at a first sight.

Once the criterion for termination has been adopted, the experiment will be designed by showing several images to the user; a choice of 6 images (the same for all users) was selected from a database of about 4700. These images are classified as belonging to different themes such as flowers, horses, paintings, skies, textures, ceramic tiles, buildings, clouds, trees, etc. even though the category is not used at all during the search. The 6 target images are in our experience, representative of different themes and levels of difficulty. They are displayed in figure 2.

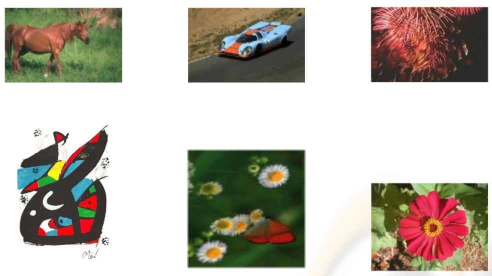

Figure 2: Target images used in experiments.

For each target image the search proceeds iteratively. In each iteration the user has to select some relevant images (similar to the target according to his/her judgment) and others significantly different from the target. The number of images of each type is left to the user, although two conditions must be fulfilled: at least one relevant and one irrelevant images must be selected and the total number of selections has to be greater than 4. The algorithm proceeds as explained in previous sections and the images are ranked. If the target appears in the first 16, it is considered to have been found; otherwise the user can move backwards or forwards to see more images in rank order and a new iteration of choosing/search/showing begins.

To ensure that the experiments are not biased, the query tasks were performed by a group of 40 users who had not been involved in the design and development of the system and had no knowledge of the content of the database or of the retrieval features and methods used (untrained users).

Table 2: Average, maximun, minimun iteration number to find a target image.

Image	It. Av.	max	min
Car	5.17(2.95)	12	1
Flower	4.17 (3.20)	17	1
Butterfly	4.71 (3.70)	19	1
firework	2.14 (1.81)	9	1
Miro	3.67 (1.55)	8	2
Glass	3.42 (1.52)	6	1
All	3.88(1.07)	19	1

Table 2 shows the average and standard deviation of the number of iterations needed to find images by these untrained users. The last row shows the average for all images and users. The experiments exhibit good performance in finding a target image (3.88 iterations in average) in the used database.

7 CONCLUSION

This paper addresses the problem of image retrieval by means of an algorithm based on logistic regression. The main advantage of the method is the facility of incorporating the feedback of the user. Its main drawback is the lack of sufficient information (too small sample) to fit the model, since the number of inputs (image features) is usually high. This has been addressed by means of partial models that get the output from each subset of the inputs. The problem of combining the information of the different models, which is a data fusion problem, is solved by using an ordered weighted averaging (OWA) operator.

Concerning the experimental results, the average number of iterations shown in 2 exhibits good performance of the procedure. Some further experimentation and results analysis is currently being carried out by our research group, where users are grouped and classified with regard to there interaction of the iterative process of image selection.

REFERENCES

de Ves, E., Domingo, J., Ayala, G., and Zuccarello, P. (2006). A novel bayesian framework for relevance feedback in image content-based retrieval systems. *Pattern Recognition*, 39:1622–1632.

Fuller, R. and Majlender, P. (2003). On obtaining minimal variability owa operator weights. *Fuzzy Sets and Systems*, 136:203–215.

O'Hagan, M. (1988). Aggregating template or rule antecedents in real-time expert systems with fuzzy set logic. In *Proc. of 22nd Annu. IEEE Asilomar Conf. on Signals*, pages 681–689, Pacific Grove, CA.

Rui, Y., Huang, S., Ortega, M., and Mehrotra, S. (1998). Relevance feeback: a power tool for interactive content-based image retrieval. *IEEE Transaction on circuits and video technology*, 8(5).

Smeulders, A., Santini, S., Gupta, A., and Jain, R. (2000). Content-based image retrieval at the end of the early years. *IEEE transactions on Pattern Analysis and Machine Intellingence*, 22(12):1349–1379.

Stejic, Z., Takama, Y., and Hirota, K. (2005). Mathematical aggregation operators in image retrieval: effect on retrieval performance and role in relevance feedback. *Signal processing*, 85:1297–324.

Yager, R. (1988). On ordered weighted averaging aggregation operators in multi-criteria decision making. *IEEE Trans. Systems Man Cybernet*, 18:183–190.

Zhou, X. and Huang, T. (2003). Relevance feedback for image retrieval: a comprehensive review. *Multimedia systems*, 8(6):536–544.