Available online at www.sciencedirect.com
ScienceDirect
Pattern Recognition 40 (2007) 2621–2632
PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY
www.elsevier.com/locate/pr

Applying logistic regression to relevance feedback in image retrieval systems

T. León[a], P. Zuccarello[b], G. Ayala[a,*], E. de Ves[b], J. Domingo[c]

[a]Department of Statistics and Operations Research, University of Valencia, Spain
[b]Department of Informatics, University of Valencia, Avda. Vicente Andrés Estellés, 1. 46100 Burjasot, Spain
[c]Institute of Robotics, University of Valencia, Spain

Received 14 July 2006; received in revised form 21 December 2006; accepted 5 February 2007

Abstract

This paper deals with the problem of image retrieval from large image databases. A particularly interesting problem is the retrieval of all images which are similar to one in the user's mind, taking into account his/her feedback which is expressed as positive or negative preferences for the images that the system progressively shows during the search. Here we present a novel algorithm for the incorporation of user preferences in an image retrieval system based exclusively on the visual content of the image, which is stored as a vector of low-level features. The algorithm considers the probability of an image belonging to the set of those sought by the user, and models the *logit* of this probability as the output of a generalized linear model whose inputs are the low-level image features. The image database is ranked by the output of the model and shown to the user, who selects a few positive and negative samples, repeating the process in an iterative way until he/she is satisfied. The problem of the small sample size with respect to the number of features is solved by adjusting several partial generalized linear models and combining their relevance probabilities by means of an ordered averaged weighted operator. Experiments were made with 40 users and they exhibited good performance in finding a target image (4 iterations on average) in a database of about 4700 images. The mean number of positive and negative examples is of 4 and 6 per iteration. A clustering of users into sets also shows consistent patterns of behavior.

Keywords: Visual information retrieval; Low-level image descriptors; Content-based image retrieval systems; Logistic regression

1. Introduction

The last few years have witnessed an increasing amount of pictorial information in different digital formats. Thus large image databases raise the need to retrieve relevant data efficiently. In this framework, content-based image retrieval (CBIR) systems are one of the most promising techniques for retrieving multimedia information [1–3]. CBIR systems are thought of as an improvement on traditional image retrieval systems based on textual information such as keywords. The new CBIR systems take advantage of valuable digital information held by the image itself. Visual features related to color, shape and texture are extracted in order to describe the image content [4]. The main drawback of textual image retrieval systems, that is, the annotator dependency, would be overcome in pure CBIR systems. Several papers have been published trying to integrate both approaches: textual and CBIR [5,6].

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 [7]. 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

* Corresponding author. Tel.: +34 963543977; fax: +34 963544735.
E-mail addresses: Teresa.Leon@uv.es (T. León), pezuc@alumni.uv.es (P. Zuccarello), Guillermo.Ayala@uv.es (G. Ayala), Esther.deVes@uv.es (E. de Ves), Juan.Domingo@uv.es (J. Domingo).

0031-3203/$30.00
doi:10.1016/j.patcog.2007.02.002

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.

The iterative algorithms which, in order to improve the set resulting from a query, require the user to enter his/her preferences in each iteration, are called relevance feedback algorithms [8]. These algorithms have been shown to provide a dramatic boost in retrieval system performance. Being part of this mainstream, this paper presents a new algorithm for relevance feedback in image databases based on logistic regression models.

A query can be seen as an expression of an information need to be satisfied. Any CBIR system aims 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 the image I and whose output is a function of $\pi(I)$. In logistic regression analysis, one of the key features to be established is the order of the model to be fitted. The order of logistic regression model, the number of image characteristics, and the number of relevant (positive/negative) images the user is prompted to select, are strongly related. 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, each consisting of semantically related characteristics. The outcome of this strategy is that n smaller regression models must be adjusted: each sub-model will produce a different relevance probability $\pi_k(I)$ $(k = 1, \ldots, n)$. 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. We tackled this problem by making use of the so-called OWA (*ordered weighted averaging*) operators which were introduced by Yager [9] and provide a consistent and versatile way of aggregating multiple inputs into one single output.

Section 2 describes related work addressing issues of feature relevance computation. Section 3 presents and explains our approach in detail. Next, Section 3.1 describes the low-level features extracted from the images and used to retrieve them. After that, in Section 4 we present experimental results which evaluate the performance of our technique using real-world data. Finally, in Section 5 we extract conclusions and point to further work.

2. Related work

Relevance feedback is a term used to describe the actions performed by a user to interactively improve the results of a query by reformulating it. An initial query formulated by a user may not fully capture his/her wishes. This is due to several reasons: the complexity of formulating the query, lack of familiarity with the data collection procedures, or inadequacy of the available features. Users then typically change the query manually and re-execute the search until they are satisfied. By using relevance feedback, the system learns a new query that better captures the user's need for information.

In recent years, several methods have been developed to guide the searching process in a retrieval system. All these techniques can be roughly classified into two different groups:

- *Query point movement*: The method of the query point movement approach is to construct a new query point that is supposed to be close to the relevant results and far from those which are non-relevant. The best-known approach for achieving query point movement is based on a formula initially developed by Rocchio in the context of textual information retrieval [10].
- *Reshaping distance functions*: the objective of this approach is to modify the distance function in such a way that it can improve the query results according to the user's criterion.

A procedure belonging to the query point movement group was proposed by Ciocca and Schettini [11], who introduce a very simple algorithm for computing a new query point $\mathbf{Q}$ that can better represent the images of interest to the user. The procedure takes the set of relevant images the user has selected and computes a new point based on the standard deviation of the features used, computed separately one by one. Obviously, this ignores the dependency between image features, which is particularly important when they are values sampled from continuous functions.

Another implementation of point movement strategy consists of using the Bayesian methodology. Cox uses an adaptive Bayesian scheme which incorporates user preferences by means of a model of the user [12]. This model, together with the prior, gives rise inductively to a probability distribution on the event space. Experiments show that retrieval performance can be improved considerably by using such relevance feedback approaches. Relevance feedback has been also considered as a Bayesian classification problem by Duan et al. [13].

Yet another approach was taken by Rui et al., who propose an interactive retrieval approach which takes into account the