An evolutionary approach for optimizing content-based image retrieval using a support vector machine

T. Kanimozhi*, K. Latha

Department of Computer Science and Engineering, University College of Engineering, BIT Campus, Tiruchirappalli 620024, Tamilnadu, India

*Corresponding author, e-mail: csrkani@gmail.com

Received 31 Aug 2014
Accepted 20 Jul 2016

ABSTRACT: One of the vital challenges in the field of image retrieval is the semantic gap between visual features and high-level semantic concepts. Various kinds of relevance feedback approaches have been developed to deal with this semantic gap. Of these, the support vector machine is important. The support vector machine relies upon its parameters and the number of feedback samples. This paper uses a Gaussian firefly algorithm along with support vector machine to raise the relevance feedback performance. The proposed approach determines the parameters of the support vector machine and increases the number of relevant feedback samples, thereby optimizing the retrieval process. The performance of the proposed approach is compared with other existing retrieval methods.

KEYWORDS: image mining, relevance feedback, parameter selection, Gaussian firefly algorithm

INTRODUCTION

Recently, multimedia information retrieval has become the most examined area as there is tremendous increase in the technology of digital world for storing multimedia data among which content-based image retrieval (CBIR) is one of the significant topic, where many proposals have been made in the past 10 years. Even though vast studies have been conducted, image retrieval remains one of the key challenging issues. This event is broad because of two major causes. One being the gap between the information signified by computers and the object of the world also called sensor gap, and the other is the semantic gap which is the gap between the low-level visual features and high-level human observation and understanding.

Relevance feedback (RF), as an unconventional and a more viable technique to bridge the semantic gap and an intensive research, have been also carried out in recent years. In the example of image retrieval, relevance feedback initially solicits the user’s relevance opinion on the retrieved images returned by image retrieval schemes. Side by side, it refines retrieval results by breaking through the query targets from the presented relevance information. More recently, Broilo and de Natale investigated the possibility of embedding relevance feedback with particle swarm optimization (PSO) into the image retrieval process in which the problem of image stagnation in local optima is avoided and the improved exploration of image space is provided. Lai and Chen employed the genetic algorithm as an interactive process in the relevance feedback of content based image retrieval system and reduced the gap between the image retrieval results and the user’s preferences.

In late years, support vector machine (SVM) has been tremendously investigated in machine learning as its performance in real-world applications is comparatively higher in pattern classification. However, most of the SVM-based RF approaches do not consider the fundamental difference between relevant and irrelevant feedback images, that is, unlike the relevant feedback images which have a similar notion for all the images, the irrelevant feedback images have various notions in a different manner in each image.

As reported, conventional SVM based relevance feedback learning treats relevant and irrelevant feedback images equally. Directly utilizing the SVM learning as an RF method degrades the efficiency of CBIR systems due to the facts such as “dissimilar semantic model reside in different subspaces and every image can reside in various different subspaces” and it is the purpose of relevance feedback to point out “which one” and due to disproportion in labelling the feedback images, where the number of relevant feedback images are fewer than the irrelevant feedback images. In addition, despite the fact
that SVM generally furnish reasonable accuracies, 
the training data have been frequently not linearly 
separable, SVM introduces the concept of “kernel 
influenced feature space” which maps the data into 
a higher-dimensional space where the data is sep-
arable into two classes. Normally, such a mapping 
space would create problems computationally. This 
leads to long iteration problems and stagnation in 
local optimal solutions. These drawbacks play a 
vital role in lessening the performance of SVM based 
RF for CBIR. Nevertheless, the results of SVM 
can be made more robust by tuning the internal 
parameters, such that they reduce the error, that is, 
the misclassification effect.

Consequently, Arevalillo-Herráez et al. worked 
on a different approach to relevance feedback CBIR. 
Their technique integrates an interactive GA with 
an extended nearest neighbour approach to miti-
gate the existing gap between the high-level semantic 
contents of images and the information given 
by their low level descriptors. Renukadevi and 
Thangaraj enhanced the performance of SVM ra-
dial basis function on classifying computed tomog-
raphy images by optimizing SVM parameters using 
PSO. Imran et al. introduced a PSO-based SVM 
approach to improve the performance of content 
based image retrieval by boosting the number of 
positive images compared to user selected relevant 
images.

Motivated by these approaches, the process of 
SVM learning is embedded into a strong, well-
structured, meta-heuristic algorithm, called a Gaussi-
n firefly algorithm to increase the performing 
efficiency of content based image retrieval process. 
During the relevance feedback iterations, the SVM 
finds the hyperplane that separates relevant and 
irrelevant images and it is adapted to optimum level 
along the iteration by the firefly algorithm. Hence 
the degrading performance of the SVM classifier is 
enhanced by the assistance of firefly algorithm (FA) 
and optimizes the relevance feedback process. In 
the proposed approach, the firefly optimizer not 
merely functions as an efficient optimization engine, 
but searches the search space effectively and avoids 
converging to local minima.

THE EVOLUTION OF FIREFLY ALGORITHM

Recently, FA was introduced by Yang based on 
the flashing (social) behaviour of fireflies. In social 
insect colonies, each individual seems to possess its 
own agenda and yet the group as a whole seems 
to be extremely coordinated. A swarm is a group 
of multi-agent systems such as fireflies, in which 
simple agents coordinate their activities to solve the complex problem of the allocation of communica-
tion to multiple forage sites in dynamic environments. The flashing light helps fireflies (agents) in 
attracting their potential prey, protecting themselves 
from their enemy, including their successful repro-
duction, for finding mates, and to share their food.

Agents are considered as fireflies that discharge 
light relative to the value of luciferin which is carried 
in the abdomen of firefly. The idealized firefly algo-


The query selection and distance calculation

The first and foremost step operation is to express the images in terms of features. The visual signature of the pth image is composed of: \( M_{ch} \) colour histograms \( y_{p}^{ch} \), \( M_{cm} \) colour moments \( y_{p}^{cm} \), \( M_{edh} \) edge histograms \( y_{p}^{edh} \), and \( M_{wt} \) texture wavelet feature values \( y_{p}^{wt} \). The feature vectors \( V_{p} = y_{p}^{ch} + y_{p}^{cm} + y_{p}^{edh} + y_{p}^{wt} \) of dimension \( D = M_{ch} + M_{cm} + M_{edh} + M_{wt} \) provide the overall depiction of the image. The feature vectors of database images are computed off-line and those of the query image are computed on-line. From there, each one is absolutely represented as a feature vector in \( D \)-dimensional space. The system then illustrates the most \( M_{FB} \) nearest images from the entire database after choosing and mapping of query image in the feature space based on the following distance equation:

\[
\text{Dist}(V_q, V_s) = \text{WMSE}(v_q^{ch}, v_s^{ch}) + \text{WMSE}(v_q^{cm}, v_s^{cm}) \\
+ \text{WMSE}(v_q^{edh}, v_s^{edh}) + \text{WMSE}(v_q^{wt}, v_s^{wt})
\]

where \( V_q \) is the query feature vectors and \( V_s \) is the database feature vectors; \( s = 1, \ldots, M_{DB} \) is the total number of database images. WMSE is the weighted Euclidean distance computed among a pair of feature vectors

\[
\text{WMSE}(v_q; v_s) = \frac{1}{N} \sum_{r=1}^{N} (v_{qr} - v_{sr})^2 w_r^k
\]

where \( w_r^k \) is weight vectors related to the features, \( k \) is the iteration number, and \( N \) is equal to \( M_{ch}, M_{cm}, M_{edh}, \) or \( M_{wt} \). All the features are uniformly significant at the first iteration \((k = 1)\), i.e., \( w_r^1 = 1 \), \( r = 1, \ldots, N \), after calculating \( \text{Dist}(V_q, V_s) \) and \( V_s \), \( s = 1, \ldots, M_{FB} \), the system ranks the entire database based on the distance from the query and sorts the results. The first feedback is then requested by presenting the ranked list to the user.

Relevance feedback learning via SVM

The SVM works on the fact that minimum the Vapnik-Chervonenkis dimensions minimum is the structural risk. The basic performance of SVM is as follows. Initially the input vectors are mapped into higher dimensional feature space, then a hyperplane is constructed which can distinguish the two categories within this feature space. A mapping function involves comparatively low-dimensional vectors within the input sample space and dot product within the feature space. The mapping of input vectors into feature space is carried out by a kernel function.

During the relevance feedback process, the user labels each retrieved image \( y_{f} \), as relevant \((y_{f} = 1)\) and irrelevant \((y_{f} = 0)\) according to the input query image. The two image subsets namely, relevant and irrelevant are now created and they are updated in all the iterations, conserving the record of the earlier retrieval process. To perform a classification fairly well with the created relevant and irrelevant image subsets, SVM as a binary classifier is significant which confines the query concept of separating the relevant images from the irrelevant images using a hyperplane in the non-linearly transformed feature space.
space. We exploit the set of collective retrievals $R = (V_f, y_f)$, $f = 1, \ldots, l$, where $y_f \in \{0, 1\}$, as training data in the following standard SVM with radial basis function in order to compute SVM.

$$\min_{\alpha_f} \frac{1}{2} \sum_{f=1}^{l} \sum_{g=1}^{l} \alpha_f \alpha_g y_f y_g K_s(V_f, V_g) - \sum_{f=1}^{l} \alpha_f$$

subject to $\sum_{f=1}^{l} \alpha_f y_f = 0, \quad 0 \leq \alpha_f \leq C$ (3)

where $\alpha_f$ are known as Lagrange multipliers and the parameter $C$ controls the separation error. The term $K_s(V_f, V_g)$ denotes radial basis function kernel or nonlinear similarity measure between two feature vectors which is to optimize the hyper-plan with $\sigma$ being standard deviation

$$K_s(V_f, V_g) = \exp \left( -\frac{\|V_f - V_g\|^2}{2\sigma^2} \right). \quad (4)$$

Then the score $F(V)$ for each database image $V$ is calculated using the following SVM decision function and the images are sorted based on this score.

$$F(V_s) = \sum_{f=1}^{R} \alpha_f y_f K_s(V_s, V_f) + b \quad (5)$$

for $V_s, s = 1, \ldots, M_{DB}$, where $b$ is the biases and $R$ is the number of support vectors. The $N$ highest score images are then fed to a firefly algorithm for stochastic optimization. Previous studies show that selecting the parameters of support vector machine $C$ and $\sigma$ plays an important role in determining the accuracy of the image retrieval process, so in this paper Gaussian firefly algorithm is employed to optimize the parameter of SVM and thereby to improve the retrieval accuracy of the image retrieval process.

### Optimization using a stochastic firefly algorithm

In the proposed system, each firefly or agent is a parameter set of the SVM learning process and can be represented as $V_f = (C, \sigma)$. A set of agents or fireflies is described as a population or swarm and the image retrieval process is formulated as a process of optimization. Firefly algorithm is therefore considered for the retrieval problem. For this purpose, the swarm of agents $A_n; n = 1, \ldots, N$ or a swarm space, i.e., lies is defined as points inside the solution space, i.e., as $D$-dimensional feature vectors in the solution space. In this application, the decision variables of firefly algorithm are the feature vectors. At present, each firefly is associated with $N$ highest score images in order to initialize the swarm. Then the light intensity and the attractiveness of each firefly are calculated to initialize the stochastic optimization. Our hypothesis here is that the tuning of SVM internal parameters can improve the content based image retrieval results.

In an optimization procedure, specifying the role that has to be maximized or minimized which is also referred as fitness is a significant point. This fitness value must specify the effectiveness of the position reached by fireflies. When irrelevant and relevant images are taken into account, the weight cost function $\phi(A_n)$ expresses the fitness associated with the solution space found by the swarm of firefly $A_n$:

$$\phi(A_n) = \frac{1}{M_{rel}} \sum_{r=1}^{M_{rel}} \text{Dist}(A_{n; \text{rel}}^k; V_r^k) + \frac{1}{M_{irr}} \sum_{i=1}^{M_{irr}} \text{Dist}(A_{n; \text{irr}}^k; V_i^k) \quad (6)$$

where $k$ is the iteration number, $V_r^k, r = 1, \ldots, M_{rel}^k$ and $V_i^k, i = 1, \ldots, M_{irr}^k$ are the images in the relevant and irrelevant image subsets, respectively. The computation of Dist$(A_{n; \text{rel}}^k; V_r^k)$ is the same as that defined in the previous step. It is to be noted that the cost function produces a lower fitness value, when the firefly is close to relevant images and far from the irrelevant images. Thus lower the fitness value, the better the positions achieved by the firefly towards the relevant images. Grounded along the value of fitness, it is viable to re-order the firefly to obtain new ranking. It is worth mentioning that in most examples the number of relevant images collected across iterations is smaller than the number of irrelevant images; this view makes the fitness associated with the solution space depends only on relevant images when the average firefly’s distance grows from the irrelevant images. Having defined all the elements of the optimization process, we still need to identify how to make the swarms evolve in time. To do that we have to define some attributes of the fireflies. There are two stages of firefly algorithm which are identified as follows:

(i) Attraction towards new firefly.

The attractiveness function $\beta(r)$ of the firefly is described by

$$\beta(r_n) = \beta_0 e^{-r_n^2} \quad (7)$$
where $\beta_0$ is the attractiveness with distance $r = 0$ and $\gamma$ stands for the light absorption coefficient. A firefly attractiveness $\beta$ depends on the light intensity of adjacent fireflies. The brightness $I$ of a firefly is equivalent to its fitness value $f(x)$.

\[ I_n = f(x_n) = \phi(A_n). \]  

(8)

Thus an agent with high/low light intensity will attract another firefly high/low light intensity. Each firefly has its unique attractiveness which means how strong it attracts other members of firefly. Yet the attractiveness is relative, it will differ with the distance between two fireflies $n$ and $s$ located at positions $x_n$ and $x_s$, respectively, and it is given as

\[ r_{ns} = \|x_n - x_s\| = \text{Dist}(A_n; V_s) \]  

(9)

for $n = 1, \ldots, N$ and $s = 1, \ldots, M_{DB}$.

(ii) Movement of fireflies.

The movement of firefly $n$ located at $x_n$ is attracted to another more attractive firefly $s$ located at $x_s$, i.e., $n$th firefly changes its position if there exists any $x_s$ such that $I_s < I_n$, which is determined by

\[ x_n(g + 1) = x_n(g) + \beta_0 e^{-\gamma r_{ns}}(x_s - x_n) + \alpha \varepsilon_n, \]  

(10)

where the first expression signifies the current location of a firefly, the second expression is the attractiveness, and the third expression represents the randomization with a vector of random variable $\varepsilon_n$ being drawn from the Gaussian distribution with $\mu = 0$ and $\sigma = 1.0$. Then from this Gaussian distribution, a random number is drawn such that it is tied into each firefly probability $p$. The behaviour of firefly is then introduced by

\[ x_n(t + 1) = x_n(t) + \alpha(1 - p) \text{Rand}(0, 1) \]  

(15)

where $\text{Rand}(0, 1)$ is a random number between 0 and 1. The new solutions are evaluated and then updated its light intensity. After evaluation, if the new position $x_n(t + 1)$ of a firefly shows a better cost then it will move towards that new position.

If the dimension is high, the value of $m$ is low which means the algorithm converges more accurate. This strategy changes the step length of time and thus it makes the firefly to search the solution space globally in the initial stage and in the end of iterative process, the firefly exploits the solution space locally to obtain better solutions. Hence the movements of firefly are stabilized.

As standard normal distribution is employed, the value of $\mu$ is set to 0 and $\sigma$ is set to 1. Then from this Gaussian distribution, a random number is drawn such that it is tied into each firefly probability $p$. The movement of firefly is then introduced by

\[ x_n(g + 1) = x_n(g) + \beta_0 e^{-\gamma r_{ns}}(x_s - x_n) + \alpha \varepsilon_n, \]  

(10)

where the first expression signifies the current location of a firefly, the second expression is the attractiveness, and the third expression represents the randomization with a vector of random variable $\varepsilon_n$ being drawn from the Gaussian distribution with $\mu = 0$ and $\sigma = 1.0$. Then from this Gaussian distribution, a random number is drawn such that it is tied into each firefly probability $p$. The behaviour of firefly is then introduced by

\[ x_n(t + 1) = x_n(t) + \alpha(1 - p) \text{Rand}(0, 1) \]  

(15)

where $\text{Rand}(0, 1)$ is a random number between 0 and 1. The new solutions are evaluated and then updated its light intensity. After evaluation, if the new position $x_n(t + 1)$ of a firefly shows a better cost then it will move towards that new position.

It makes sense to observe the possibility of examining the fireflies as query points which will travel around the $D$-dimensional search space that constitutes image features with its own light intensity. After the two phases of the firefly algorithm at the first iteration, an updating process is carried out at every further iteration. The significance of the objective function, attractiveness and movement of firefly towards other firefly are again determined if new relevant images are labelled by the user. As a consequence, after subsequent user feedback, the fireflies travel towards new locale in the search space where other new relevant images may be found.

It is to be noted that the fireflies move in a continuous way inside the search space or feature
space while the database images \( (V_i; s = 1, \ldots, M_{DB}) \) are distinct and fixed set of points. Thus further operation is needed to complete a single iteration. The swarm agents ranked according to (6) is placed at "correct position" in the solution space that is to be related to their nearest images in the database according to (1). Hence a new set of images is obtained which is then shown to the user. If the swarm of fireflies points to the irrelevant image which is already classified or more than one swarm of fireflies points to the same image, these images are discarded and the next closest or nearest images are conceived until a different \( M_{FB} \) set of images are gathered from the user. After the user feedback, the above said process of firefly swarm updating and SVM learning is iterated. The process ends, when the user verifies one of the following conditions: (i) the result of search satisfies user; (ii) the relevant number of images targeted is attained, or (iii) when it reaches the predefined number of iterations. All the relevant images are shown to the user, after the process terminates. Thus based on swarm intelligence, the FA find the global optima of the objective function by investigating the foraging behaviour of fireflies.

**EXPERIMENTAL RESULTS**

**Experimental setup**

In our experiments, we used Corel photo galleries covering a wide range of semantic categories of natural scenes with artificial objects. The dataset is partitioned into 25 categories, including butterfly, buildings, hills, flowers, earth, sky, trees, boats, birds, statue, horses, and elephants, etc., and each category are represented by 100 images, for a total of 2500 images. In our experiments, 600 images are randomly taken out of the whole dataset as test queries.

As usual, the feature computation process is on-line for query image and off-line for database image. The feature components of database images are stored in a database for run time access. The visual signature of each image is composed of four different feature vectors. The first one is 32-bin colour histogram calculated in the HSV colour space, the second one is 9-bin colour moments extracted from HSV colour space, the third one is 8-bin edge direction histogram \(^{16}\) is obtained from the edge map of an image and the forth one is 18-wavelet texture energy values \(^{17}\). Hence the size of feature vector was set to 67.

**Results**

For each image database in our experiments, we simulated the presence of users by using each image as initial query point. For a query image, 10 iterations of user-and-system interaction were carried out. At each iteration, the system examined the top 18 images and images from the same (different) category of the query image were considered equally relevant (irrelevant) images.

To evaluate the effectiveness of the proposed approach, we use precision and recall as performance measurements which are defined as

\[
\text{Precision} = \frac{N_{p(q)}}{N_{r(q)}}, \quad \text{Recall} = \frac{N_{p(q)}}{N_{T(q)}},
\]

where, \( N_{p(q)} \) represents the number of retrieved relevant images, \( N_{r(q)} \) denotes the total number of retrieved images for a given query \( q \) and \( N_{T(q)} \) indicates the total number of relevant images available in the database for a given query \( q \). Since the FA, PSO stochastic nature, all precision and recall values are calculated by averaging five consecutive runs for every query image.

**Performance measures**

Three relevance feedback methods: PSO and SVM in CBIR\(^ {18} \), GA and SVM in CBIR\(^ {19} \), and SVM in CBIR\(^ {20} \) are compared with the proposed method. The firefly algorithm is compared with PSO and genetic algorithm for content based image retrieval along with SVM since these 2 methods are in the same category of population based optimization techniques. We chose the third one as SVM in content based image retrieval, as it is based on machine learning techniques. Altogether the three comparable methods make use of relevant and irrelevant images.

**Parameter tuning**

In our experiments, the values for the firefly algorithm parameters are the same as suggested by Yang\(^ {12} \). The swarm of fireflies is randomly initialized in the solution space. The parameter values used in this paper are: number of fireflies = 18; light absorption coefficient \( \gamma = 1 \); attractiveness \( \beta_0 = 1 \); number of generations = 110. After the fireflies are randomly distributed in the solution space, the parameter value of \( \beta_0 \) strongly determines the firefly positions which are neighbourhood to the brightest firefly. This is tantamount to the co-operative local search strategy. The value of light absorption coefficient \( \gamma = 1 \) determines the value of light intensity
Average precision versus recall curve, the better its performance is) with the number of iterations from 3–10. These curves demonstrate that the performance of the proposed method is improved (the higher the precision versus recall curve, the better its performance is) with the number of iterations. Figs. 2, 3, and 4 show the experimental results of precision versus recall curve for 5, 7, and 10 iterations of the relevance feedback process. As can be observed from the figure, the proposed method consistently outperforms all other baselines.

**CONCLUSIONS**

The SVM classifier, being a statistical method of pairwise classification, enables one to make use of the statistical properties of a data, to classify a dataset. The efficiency of this classifier can be further improved by overcoming a few defects. This approach lead to the technique of integrating SVM learning with a Gaussian firefly algorithm as a unique optimization model for the image retrieval problem. The use of firefly algorithm finds optimum parameters for SVM accordingly achieves a number of relevant feedback images and thus optimizes the relevance feedback process. This introduced diversity so as to enlarge the exploration and to converge showing that importance of classification step in relevance feedback based learning process. This entails the significance of incorporating SVM with firefly algorithm and improves the model in terms of precision and recall. Fig. 1 shows the experimental results of precision versus recall curve by varying the number of iterations from 3–10. These curves demonstrate that the performance of the proposed method is improved (the higher the precision versus recall curve, the better its performance is) with the number of iterations. Figs. 2, 3, and 4 show the experimental results of precision versus recall curve for 5, 7, and 10 iterations of the relevance feedback process. As can be observed from the figure, the proposed method consistently outperforms all other baselines.

**Comparison methods**

The performance of this experiment is evaluated using average precision versus recall curve, and aims to verify whether the proposed method have a significant difference from all the other baselines or comparing methods. Table 1 shows the average performance comparison of FA and SVM in CBIR with FA in CBIR for different categories of Corel database. From the table, it is noticed that precision and recall values of FA and SVM in CBIR system outperforms FA in CBIR system showing that importance of classification step in relevance feedback based learning process. This entails the significance of incorporating SVM with firefly algorithm and improves the model in terms of precision and recall. Fig. 1 shows the experimental results of precision versus recall curve by varying the number of iterations from 3–10. These curves demonstrate that the performance of the proposed method is improved (the higher the precision versus recall curve, the better its performance is) with the number of iterations. Figs. 2, 3, and 4 show the experimental results of precision versus recall curve for 5, 7, and 10 iterations of the relevance feedback process. As can be observed from the figure, the proposed method consistently outperforms all other baselines.

**Table 1** Average performance comparison for FA+SVM in CBIR with FA in CBIR for categories of Corel database.

<table>
<thead>
<tr>
<th>Categories</th>
<th>FA in CBIR</th>
<th>FA+SVM in CBIR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Precision</td>
<td>Recall</td>
</tr>
<tr>
<td>Butterfly</td>
<td>84.12</td>
<td>60.14</td>
</tr>
<tr>
<td>Buildings</td>
<td>80.23</td>
<td>56.11</td>
</tr>
<tr>
<td>Hills</td>
<td>83.21</td>
<td>60.52</td>
</tr>
<tr>
<td>Flowers</td>
<td>77.42</td>
<td>54.13</td>
</tr>
<tr>
<td>Earth</td>
<td>84.31</td>
<td>61.12</td>
</tr>
<tr>
<td>Sky</td>
<td>83.24</td>
<td>58.25</td>
</tr>
<tr>
<td>Trees</td>
<td>83.21</td>
<td>55.13</td>
</tr>
<tr>
<td>Boat</td>
<td>83.25</td>
<td>62.14</td>
</tr>
<tr>
<td>Bird</td>
<td>83.23</td>
<td>58.01</td>
</tr>
<tr>
<td>Statue</td>
<td>83.15</td>
<td>58.14</td>
</tr>
</tbody>
</table>

as the distance increases from the communicated firefly which results in the complete random search.
towards global optima rather than local optima, when optimizing the performance of SVM classifier which improves the retrieval efficiency. Experiments were carried out on a large-size database of images, and the experiment results reveal that it has a better retrieval efficiency compared to existing methods.

REFERENCES