

Cairo University

Egyptian Informatics Journal

[www.elsevier.com/locate/eij](http://www.elsevier.com/locate/eij) [www.sciencedirect.com](http://www.sciencedirect.com/science/journal/11108665)



FULL-LENGTH ARTICLE

Distance selection based on relevance feedback in the context of CBIR using the SFS meta-heuristic with one round



# Mawloud Mosbah [\*](#_bookmark0), Bachir Boucheham

*University 20 Aouˆt 1955 of Skikda, Algeria*

Received 20 August 2015; revised 24 June 2016; accepted 1 September 2016

Available online 19 October 2016

Abstract In this paper, we address the selection in the context of Content Based-Image Retrieval (CBIR). Instead of addressing *features’ selection* issue, we deal here with *distance selection* as a novel paradigm poorly addressed within CBIR field. Whereas distance concept is a very precise and sharp mathematical tool, we extend the study to weak distances: Similarity, quasi-distance, and diver- gence. Therefore, as many as eighteen (18) such measures as considered: distances: {Euclidian,

.. .}, similarities{Ruzika, .. .}, quasi-distances: {Neyman-*X*2, .. .} and divergences: {Jeffrey, .. .}. We specifically propose a hybrid system based on the Sequential Forward Selector (SFS) meta- heuristic with one round and relevance feedback. The experiments conducted on the *Wang database* (*Corel-1K*) using color moments as a signature show that our system yields promising results in terms of effectiveness.

KEYWORDS

CBIR;

Distance-selection; Relevance feedback; SFS meta-heuristic; Precision and Recall

© 2016 Production and hosting by Elsevier B.V. on behalf of Faculty of Computers and Information,

Cairo University. This is an open access article under the CC BY-NC-ND license [(http://creativecommons.](http://creativecommons/)

org/licenses/by-nc-nd/4.0/).

1. Introduction

As any information retrieval system, a Content Based-Image Retrieval (CBIR) system aims at satisfying the user need through extracting, from the image database, a subset of

\* Corresponding author.

E-mail addresses: [mos\_nasa@hotmail.fr](mailto:mos_nasa@hotmail.fr), [m.mosbah@univ-skikda.dz](mailto:m.mosbah@univ-skikda.dz) (M. Mosbah).

Peer review under responsibility of Faculty of Computers and Information, Cairo University.

**Production and hosting by Elsevier**

images deemed as similar to the submitted query, let alone rel- evant to the user expectations. For doing so, a CBIR system utilizes some low-level features such as color, e.g. [[1]](#_bookmark18), texture,

e.g. [[2,3]](#_bookmark19) and shape, e.g. [[4]](#_bookmark20). A comparative study of some CBIR works is reported in [[5]](#_bookmark22). Unfortunately, users are still usually unsatisfied with results answered by actual CBIR sys- tems, owing to the semantic gap problem. Indeed, there is a gap between the relevance notion from the user viewpoint and the automatic relevance of the system. For improving results given by a CBIR system, one must, therefore then, reduce the gap between the two previous cited kinds of rele- vance. The relevance from the user perspective is related to what he/she has in his/her mind about his/her needs, whereas relevance from the system viewpoint is related to the query.

<http://dx.doi.org/10.1016/j.eij.2016.09.001>

1110-8665 © 2016 Production and hosting by Elsevier B.V. on behalf of Faculty of Computers and Information, Cairo University. This is an open access article under the CC BY-NC-ND license [(http://creativecommons.org/licenses/by-nc-nd/4.0/).](http://creativecommons.org/licenses/by-nc-nd/4.0/))

Among suggested solutions to the semantic gap, some authors used multiple query techniques, e.g. [[6]](#_bookmark22).

A basic key affecting the system relevance, and in conse- quence its accuracy, is the matching measure to work with. Review of literature shows that there are many matching mea- sures ranging from distances and similarities to quasi-distances and divergences. To the best of our knowledge, few works have addressed the matching measure as a point of interest in the context of CBIR, e.g. [[7–10]](#_bookmark22). The question to ask then is what matching measure should one use when building a CBIR sys- tem. Similarly, what matching measure should be used with respect to a specific query? This question leads consequently to the legitimate issue of *matching measure selection*.

A natural answer to the question of matching measure selection is the learning process. Review of literature shows that there are two manners for implementing the matching measure selection in the CBIR field: utilizing selection problem tools or learning through relevance feedback.

The proposed work falls within both aforementioned areas of CBIR field: selection paradigm and relevance feedback. These notions are explained in the following subsections.

*1.1. Selection paradigm*

One-Round algorithm of being very efficient. In the following point, we explain briefly the SFS algorithm.

*1.1.1. SFS algorithm*

In this work, we use the SFS algorithm rather than other meta- heuristic algorithms such as Genetics Algorithm and Cuckoo Search Algorithm (CSA) owing to its simplicity. Other meta- heuristics than the SFS algorithm are of course of great inter- est as subject of study. Indeed, review of literature reveals that there exist many meta-heuristic algorithms applied in a variety of fields, e.g. [[17–20]](#_bookmark28). However, choosing the best meta- heuristic algorithm for selecting the adequate matching mea- sure goes beyond the scope of this paper.

Because we do not want to combine matching measures, we believe that one round is enough to answer the question: ‘‘which matching measure is the best?”. The pseudo code of the SFS algorithm is presented in [Fig. 1](#_bookmark1). In this pseudo code, the fitness value has trade-off with the sum of the ranks of images labeled as relevant by the user. This fitness is given by the following equation:

X*n*

*fitness* = 1/*n* \*

*rank*(*i*) (1)

*i*=1

To the best of our knowledge, the selection paradigm in the CBIR field has been so far restricted to features selection aspect only, e.g. [[11]](#_bookmark22). Indeed, many authors have asked the fol- lowing question: ‘‘which features are most suitable for a speci- fic query?”. Features selection methods search for the most relevant feature subset, belonging to the original feature space, according to a user defined criterion [[12]](#_bookmark22). Features selection algorithms aim at choosing a reduced number of features that preserve the most relevant information of the dataset. Features selection is usually applied as a preprocessing step in data min- ing tasks by removing irrelevant or redundant features leading to more efficient and accurate classification, clustering and similarity searching processes [[12]](#_bookmark22). There are three broad classes of features selection: filter methods, e.g. [[13]](#_bookmark22), wrapper methods, e.g. [[14]](#_bookmark23), and hybrid methods. Filter methods use general characteristics of the data independently of the classi- fier for the evaluation process. The evaluation process is classifier-dependent in wrapper methods. Finally, hybrid mod- els use both filtering and wrapping methods for improving the performance of the selection process.

The problem with the selection tools is that the learning stage is expensive in terms of computing time. Therefore, it is done offline. In addition to that, tools, utilized in the learn- ing stage, require evaluation according to a fitness measure. This poses other questions about the processed dataset devoted to learning. The retrieval problem, in the case of sys- tems based on feature selection, can, therefore, be viewed as a classification problem. Evidently, in this case, the learning stage is crucial.

In this paper, we address *the matching measure selection paradigm* rather than *the feature selection paradigm*. More specifically, we aim to select, for each query, one matching measure that would be the best for a given query from the per- spective of effectiveness. For doing so, we utilize the *Sequential Forward Selector* (SFS) algorithm [[15,16]](#_bookmark24) with one round. This choice has been motivated by the characteristic of the SFS-

where *n* is the number of images labeled as relevant by the user.

The SFS algorithm with one round uses the pre-cited pseudo code from Step 1 to Step 4.

*1.2. Relevance feedback*

The relevance feedback concept, coming from documentary information retrieval [[21,22]](#_bookmark25), has received, in last few years, a lot of attention in the CBIR field, e.g. [[23]](#_bookmark25). This scheme con- sists of receiving additional information from the user after visualizing the initial results. This additional information is simply the judgment of some visualized results by the user as relevant or non-relevant to his/her requirement. According to this judgment, the system proceeds to adjust its processing behavior for improving performances. The relevance feedback mechanism then is an additional tool for reducing the angle between the user relevance and system relevance by giving a

SFS Algorithm

Step 1: as initialization the algorithm starts with the following weighting (0, 0,..0) (no selected matching measure).

Step 2: each weight will be set 1 separately to generate many configurations.

Step 3: to evaluate each configuration based on the fitness. Step 4: selecting the best configuration.

Step 5: comparing the actual selected configuration with the selected configuration of the previous iteration, if there is no improvement so go to the Step 8.

Step 6: set the other weights 0 except the weight of the selected matching measures is still 1.

Step 7: go to the Step 2. Step 8: END.

Figure 1 Pseudo code of the SFS algorithm.

more clear vision on the user expectations and adjusting the inside system behavior in hope to bridge the semantic gap. Review of literature shows that there exist a lot of approaches for exploiting the feedback. The first approach consists of shifting query, in a way based on the new generated query, the images deemed as relevant by the user will be better ranked while the images judged as non-relevant will be ranked on the bottom. Query Point Movement [[24]](#_bookmark25), Standard Rocchio’s For- mula [[25]](#_bookmark25), and Adaptive Shifting Query [[26]](#_bookmark25) are three alterna- tives of this approach. Feature weighting [[27]](#_bookmark25), which is close to the feature selection, and the optimization of the parameters of the similarity metrics [[28,29]](#_bookmark26), utilizing of K-nearest Neighbors (KNN) classification algorithm [[30]](#_bookmark27) are other techniques for exploiting feedback. A comparative study of these approaches is given in [[31]](#_bookmark29). The approach adopted in this work is close to ‘‘Optimization of the parameters of the similarity metric”. This approach consists of optimizing the parameters in the case of many similarities or distances. The parameters that make the rate of images labeled as relevant by the user better than the rate of images labeled as non-relevant are the best configura- tion to look for.

To note that methods based on relevance feedback suffer generally from the scarcity of images being judged by the user. Generally, it is not possible to build a good model based on few deemed images which requires to enlarge the subset of judged images.

The difference between the two pre-cited approaches: *fea- ture selection* and *learning using feedback* is that *the feature selection* is a broad approach which explores all the possible cases and proceeds to designate for each class of images the

best configuration. Therefore feature selection is expensive in terms of resources especially in terms of processing time but this does not a matter owing to the fact that learning stage is done offline. For *the features weighting* based on the relevance feedback information, the task is done in deep way. In other words, the system looks for the configuration which ranks, on the top, the images deemed as relevant by the user. The learning then will be stopped when this condition is satisfied. In this paper, we introduce a novel approach which combi- nes the two approaches: *selection* and *relevance feedback* but we focus on the matching measure rather than features. To the best of our knowledge, the distance selection has not been addressed as a point of interest in the context of CBIR. Our proposed approach takes advantage then of both approaches: the effectiveness of selection paradigm and the efficiency of rel-

evance feedback.

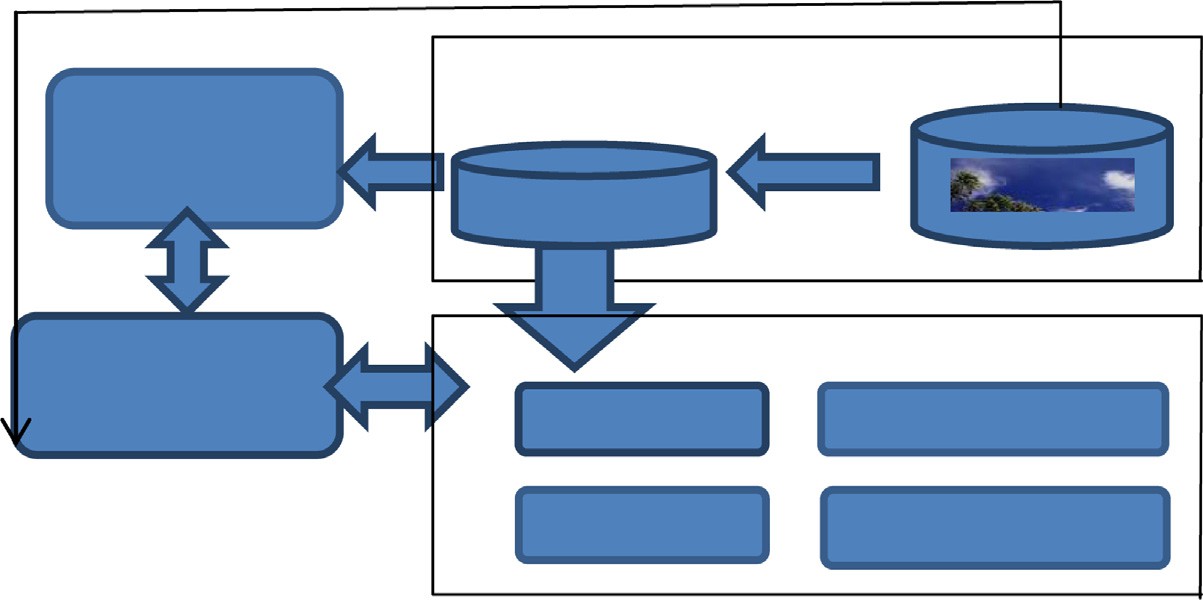
The rest of this paper is arranged as follows: Section [2](#_bookmark2) pre- sents the proposed approach. In Section [3](#_bookmark9), we discuss the con- sidered materials and settings. Section [4](#_bookmark21) shows experiments conducted and the obtained results. We conclude the paper with conclusion and some perspectives.

1. Our proposed approach

The execution scenario of our approach is as follows: after receiving the submitted query, the system answers by a set of images as initial results applying one matching measure. After that, the user has to designate some relevant images from those answered by the system. The first images not judged by the user are considered as non-relevant. The system will then



User



Matching Process

Indexing Process

Index

Interaction Module

SFS Selection Process

Image Results

Distance

Quasi-distance

Similarity

Divergence

Image Query

Figure 2 General architecture of the proposed approach.



Figure 3 Some images representing the 10 classes of the Wang database.

**120**

**100**

**80**

**Recall**

**60**

**40**

**20**

**0**

**Precision**

**120**

**100**

**80**

**Recall**

**60**

**40**

**20**

**0**

**Precision**

**Ruz Rob Mot**

**Cos**

**Eu In Sor Kulc Soe**

**Cheb Man Squar Maha**

**Can**

Figure 4 Average Precision/Recall over considered similarities (Ruz: Ruzicka, Rob: Roberts, Mot: Motyka, Cos: Cosine).

**350**

**340**

**330**

Figure 6 Average Precision/Recall over considered distances (Eu: Euclidean, In: Intersection, Sor: Sorensen, Kulc: Kulczunsky, Soe: Soergel, Cheb: Chebyshev, Man: Manhattan, Squar: Squared, Maha: Mahalanobis, Can: Canberra).

## Euc



**320**

**310**

**300**

**290**

## Can Maha

**Squar**

**400**

## 200

**0**

**Int**

## qua Kulc



**Utility Value**

**280**

**Utility Value**

**270**

## Manh

## Cheb

**Soer**

**260**

**Ruz Rob Mot Cos**

Figure 7 Comparison between the effectiveness of distances based on the utility concept.

Figure 5 Comparison between the effectiveness of similarities based on the utility concept.

* The first color moment

X*N*

launch the execution of the SFS algorithm with one round which designates the best matching measure. Selected match- ing measure will be then applied on the entire asked database and results will be visualized again to the user (see [Fig. 2](#_bookmark3)).

1. Settings and experiments

*m* = 1/*N fij* (2)

*j*=1

where *N* is the number of pixels in the image. *fij* is the value

of the pixel of *i*th row and *j*th column.

* + The second color moment

vuﬃﬃﬃﬃﬃﬃﬃﬃﬃXﬃﬃ*N*ﬃﬃﬃﬃﬃﬃﬃﬃﬃﬃﬃﬃﬃﬃﬃﬃﬃﬃﬃﬃﬃﬃﬃ

The results obtained in this paper are conducted on the hetero-

*v* = t1/*N*

*j*=1

(*fij*

— *m*)2

(3)

geneous *Wang database* (*Corel-1K*) [[32]](#_bookmark30). This dataset is com-

posed of 1000 images of 10 semantic classes and is widely used in the CBIR field. A sample of this base is presented in

* + The third color moment

vu3 ﬃﬃﬃﬃﬃﬃﬃﬃﬃXﬃﬃ*N*ﬃﬃﬃﬃﬃﬃﬃﬃﬃﬃﬃﬃﬃﬃﬃﬃﬃﬃﬃﬃﬃﬃ

[Fig. 3](#_bookmark4) where one image for each semantic class is shown. We utilize the three first low color moments [[33]](#_bookmark31) as a signature, 4

*s* = t1/*N*

*j*=1

(*fij* — *m*)

(4)

similarities: *Ruzicka*, *Roberts*, *Motyka* and *Cosine*, 10 dis-

tances: *Euclidean distance*, *Intersection distance*, *Sorensen dis- tance*, *Kulczunsky distance*, *Soergel distance*, *Chebyshev distance*, *Manhattan distance*, *Squared distance*, *Mahalanobis distance*, and *Canberra*. 3 quasi-distances are used here: *X2 dis- tance*, *Neyman-X2 distance* and *Separation distance* [[34]](#_bookmark32). For the divergence, we use the *Jeffrey divergence*. All these settings

* + Ruzicka similarity min{*xi*, *yi* } max{*xi*, *yi* }
  + Roberts similarity

P

P(*xi* + *y* )

min{*xi*,*yi* }

*i* max{*xi*,*yi* }

*i*

*i*

(5)

are given in the following formulas:

P(*x* + *y* ) (6)

**120**

**100**

* + Motyka similarity min{*x* , *y* }

P(*x* + *y* ) (7)

P *i i*

*i*

*i*

**80**

**Recall**

**60**

**40**

* + - Cosine similarity

→ →

*T* · *S*

cosine(*T*, *S*)=

qﬃﬃﬃﬃﬃﬃﬃﬃﬃﬃ

→ →



*T S*

(8)

**20** ● Euclidean distance

**X2**

**Ney\_X2 Sep**

**0** qXﬃﬃﬃﬃﬃﬃﬃﬃ(ﬃﬃ*a*ﬃﬃﬃﬃﬃ—ﬃﬃﬃﬃ*b*ﬃﬃﬃﬃ)ﬃﬃ2ﬃ

**0 20 40 60 80 100**

*i*

*i*

(9)

**Precision**

Figure 8 Average Precision/Recall over considered quasi-dis- tances (*X*2: Neyman-*X*2, Sep: Separation).

* Manhattan distance

X |*ai* — *bi*| (10)

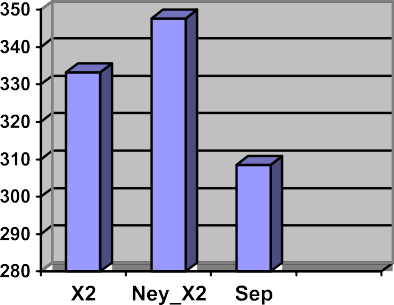
* Intersection distance

min{*a* , *b* }

P *i i*

1 — min{P *a* , P *b* } (11)

*i* *i*

* Sorensen distance

|*ai* — *bi*|

P(*ai* + *bi*)



(12)

Figure 9 Comparison between the effectiveness of quasi-dis- tances based on the utility concept.

* Kulczunsky distance

|*ai* — *bi*|

PP

min{*ai*, *bi*}

* Soergel distance

|*ai* — *bi*|

P(*ai* + *bi*)

(13)

(14)

Table 1 Correspondence between the query class and the adequate matching Measure.

Class

Corresponding matching formula

Africa

Ruzicka

Bus Monument

Separation Intersection

Horse

Separation

Dinosaur Elephant Flower Mountain Beach Food

Kulczunsky Ruzicka Jeffrey Cosine

Soergel Ruzicka Separation

Chebyshev Roberts

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 2 Precision/Recall values for each query class after applying SFS.  Recall (%) Precision (%) | | | | | | | | | | | |
|  | African | Bus | Horse | Monument | Dinosaur | Elephant | Flower | Mountain | Beach | Food | Average |
| 10 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 20 | 100 | 100 | 10.52 | 10.52 | 100 | 100 | 100 | 100 | 100 | 66.66 | 78.77 |
| 30 | 60 | 100 | 13.63 | 15 | 100 | 100 | 100 | 60 | 16.66 | 60 | 62.53 |
| 40 | 66.66 | 44.44 | 14.28 | 19.04 | 100 | 100 | 100 | 66.66 | 18.18 | 66.66 | 59.59 |
| 50 | 55.55 | 20.83 | 17.24 | 22.72 | 100 | 100 | 100 | 29.41 | 18.51 | 71.42 | 53.56 |
| 60 | 33.33 | 22.22 | 8.82 | 10.71 | 100 | 85.71 | 100 | 31.57 | 16.66 | 60 | 46.90 |
| 70 | 25.92 | 20 | 8.43 | 12.06 | 100 | 77.77 | 87.5 | 28 | 18.91 | 43.75 | 42.23 |
| 80 | 27.58 | 19.04 | 9.41 | 12.69 | 100 | 72.72 | 80 | 28.57 | 15.09 | 25.80 | 39.10 |
| 90 | 20.45 | 20.93 | 9.67 | 11.11 | 100 | 31.03 | 75 | 27.27 | 13.63 | 25.71 | 33.48 |
| 100 | 16.94 | 10.75 | 10.41 | 10.52 | 100 | 33.33 | 47.61 | 18.86 | 14.70 | 21.27 | 28.44 |
|  |  |  |  |  |  |  |  |  |  |  |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Recall (%) | Precision (%) |  | | | |
|  | Ruzicka | Canberra *X*2 Jeffrey SFS | | | |
| 10 | 100 | 100 | 100 | 100 | 100 |
| 20 | 72.20 | 68.09 | 70.83 | 82.48 | 78.77 |
| 30 | 58.89 | 60.47% | 61.64 | 59.44 | 62.53 |
| 40 | 55.35 | 55.13 | 55.34 | 45.83 | 59.59 |
| 50 | 52.91 | 51.12 | 45.34 | 41.68 | 53.56 |
| 60 | 45.36 | 46.20 | 33.87 | 40.37 | 46.90 |
| 70 | 42.92 | 41.58 | 32.67 | 35.49 | 42.23 |
| 80 | 33.83 | 31.78 | 29.51 | 32.29 | 39.10 |
| 90 | 28.29 | 29.97 | 27.33 | 30.36 | 33.48 |
| 100 | 27.28 | 27.31 | 23.54 | 24.84 | 28.44 |

**120**

Table 3 Average Precision/Recall values of the best formulas against the SFS.

**100**

**80**

**Recall**

**60**

**40**

**20**

**0**



**0 10 20 30 40 50 60 70 80 90 100**

**People Bus Horse Building Dinosaur Elephant Flower Mountain Beach Food**

**Average**

**Precision**

Figure 10 Precision/Recall curves for each asked class of the Wang database after applying the SFS with relevance feedback.

**370**



**360**



**350**

**340**

**330**

**320**

**Utility Value**

**Ruz Can Ney Jeff SFS**

Figure 12 The best measures against the SFS with relevance feedback.





Figure 11 SFS against the best measures in terms of average Precision/Recall (Jef: Jeffrey divergence).

* + Chebyshev distance

max{|*xi* — *yi* |} (15)

* Squared distance
  + Canberra distance

|*ai* — *bi*|

X

|*ai*|+ |*bi*|

* + *X*2 quasi-distance

X (*p* (*x*)— *p* (*x*))2

*x*

*p*2(*x*)

1 2

(18)

(19)

*D* = q(ﬃﬃ*f*ﬃﬃﬃﬃﬃ—ﬃﬃﬃﬃﬃ*f*ﬃﬃﬃ)ﬃﬃ*T*ﬃﬃ*A*ﬃﬃﬃﬃ(ﬃﬃ*f*ﬃﬃﬃﬃ—ﬃﬃﬃﬃﬃ*f*ﬃﬃﬃ)ﬃﬃ

where *A* = [*a* ] and *a*

= 1 — *dij*

*ij*

*ij*

max(*dij* )

*Q* 1 2 1 2

(16)

* + - Neyman-*X*2 quasi-distance

X (*p* (*x*)— *p* (*x*))2

1

2

(20)

*p* (*x*)

* + - Mahalanobis distance *x* 1

*Dm* = q(ﬃﬃ*f*ﬃﬃﬃﬃﬃ—ﬃﬃﬃﬃﬃ*f*ﬃﬃﬃ)ﬃﬃ*T*ﬃﬃ*C*ﬃﬃﬃﬃ—ﬃﬃ1ﬃﬃ(ﬃﬃ*f*ﬃﬃﬃﬃ—ﬃﬃﬃﬃﬃ*f*ﬃﬃﬃ)ﬃﬃ

1

2

1

2

(17)

* + - Separation quasi-distance

max

1 — *p*1(*x*)

(21)

where *C* is the co-variance matrix.

*x p*2(*x*)



Figure 13 Results returned by the system after applying the SFS for a query belonging to the flower class (the formula chosen by the SFS is the Jeffrey quasi-distance).

* Jeffrey divergence

X(*p* (*x*)— *p* (*x*)) ln *p*1(*x*)

1 2

(22)

where *P* is the *precision value* and s is a constant belongs to the range [0 1].

[Fig. 4](#_bookmark5) shows the average precision/recall values of the 4

*x p*2(*x*)

The effectiveness of our system is evaluated using the

Precision and Recall metrics [[35]](#_bookmark33). These metrics are given as follows:

Precision = NRIR/TNIR (23)

Recall = NRIR/TNRI (24)

where NRIR is number of relevant images retrieved. TNIR is

total number of images retrieved and TNRI is total number of relevant images in the asked database.

Even with precision/recall values, it is difficult to compare between the effectiveness of matching measures. Therefore Precision/Recall values will be changed to only one value using the Utility concept inspired from [[36]](#_bookmark34) as depicted by the follow- ing equation:

X1

*v* =

*P* \* (1 — *s*) (25)

*s*=0

considered similarities.

[Fig. 5](#_bookmark8) clearly shows the outperformance of the *Ruzika sim- ilarity* and to some extent that of the Roberts over the other similarities.

[Fig. 6](#_bookmark6) depicts the average precision/recall values of the 10 considered distances.

According to [Fig. 7](#_bookmark7), *Canberra* is the best distance in terms of precision.

[Fig. 8](#_bookmark10) illustrates the precision/recall values of the 3 consid- ered quasi-distances.

As described in [Fig. 9](#_bookmark11), Neyman-*X*2 is the best in terms of performance.

[Table 1](#_bookmark12) shows the correspondence between the query class and the adequate matching measure found by the *SFS algo- rithm* of one round with relevance feedback. For the dinosaur class, there are 4 measures selected. The 4 measures yield the same high performance. [Table 2](#_bookmark13) shows Precision/Recall values after applying *SFS* with relevance feedback for each query

class of 10 classes of the *Wang database*.



Figure 14 Results returned by the system without applying the SFS for a query belonging to the flower class (the formula utilized here is intersection distance).

[Table 3](#_bookmark14) labels average Precision/Recall values of the best measures against applying of the *SFS algorithm* with *relevance feedback* (see [Figs. 11 and 12](#_bookmark16)).

According to [Table 3](#_bookmark14) and [Fig. 12](#_bookmark15), *SFS algorithm* of one round with relevance feedback improves the performance in terms of precision (see [Figs. 13 and 14](#_bookmark17)).

1. Conclusion

This study was focused on the paradigm of matching measure selection within the CBIR systems. The study has considered as many as 18 *matching measures*, including *similarities*, *dis- tances*, *quasi-distances* and *divergences*. The selection process was based on the *SFS algorithm* with one round and *relevance feedback* in order to determine the best *matching measure* for a specific query. As such, we introduced a novel approach to the *distance-selection* paradigm in the context of CBIR, rather than the classical and well-known *features selection* paradigm. The obtained results show that our approach yields promising results in terms of precision, recall and utility value.

As a perspective, the results achieved will be compared to other *relevance feedback* techniques, especially in terms of dis- tance combination, such as the ‘‘optimization of the parame- ters of similarities” method. Moreover, we plan to address the selection of both features and matching measures utilizing other meta-heuristic algorithms and why not addressing the selection, in the context of CBIR, of different selection algorithms.

References

1. [Qiu GP. Embedded color image coding for content based](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0005) [retrieval. J Vis Commun Image Represent 2004;15(4):507–52](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0005).
2. [Haralick R. Statistical and structural approaches to texture. In:](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0010) [Proceedings of IEEE, vol. 67. p. 786–804](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0010).
3. [Seetharaman K, Sathiamoorthy S. Color image retrieval using](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0015) [statistical model and radial basis function neural network. Egypt](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0015) [Inf J 2014;15:59–68](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0015).
4. [Ko BC, Byun H. FRIP: a region-based image retrieval tool using](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0020) [automatic image segmentation and stepwise Boolean and match-](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0020) [ing. IEEE Trans Multimedia 2005;7(1):105–13](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0020).
5. [El-gayar MM, Soliman H, Meky N. A comparative study of](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0025) [image low level feature extraction algorithms. Egypt Inf J](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0025) [2013;14:175–81](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0025).
6. [Jin Xiangyu, French James C. Improving image retrieval effec-](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0030) [tiveness via multiple queries. In: Proceeding of MMDB’03, New](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0030) [Orleans, Louisiana, USA](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0030).
7. [Nielsen Frank. Steering self-learning distance algorithms. Com-](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0035) [mun ACM 2009;52(11)](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0035).
8. [Gomez Faustino J. Sustaining diversity using behavioral infor-](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0040) [mation distance. In: GECCO’09, Montre´ al Que´ bec, Canada](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0040).
9. [Cha Sung-Hyuk. Comprehensive survey on distance/similarity](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0045) [measures between probability density functions. Int J Math](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0045) [Models Methods Appl Sci 2007;1(4)](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0045).
10. [Zhang Zhenjie, Ooi Beng Chin, Parthasarathy Srinivasan, Tung](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0050) [Anthony KH. Similarity search on bergman divergence: towards](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0050) [nonmetric indexing. In: VLDB ‘09, August 2428](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0050).
11. [Jiang Wei, Er Guihua, Dai Qionghai, Jinwei Gu. Similarity-based](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0055) [online feature selection in content-based image retrieval. IEEE](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0055) [Trans Image Process 2006;15(3):702–12](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0055).
12. [da Silva SF et al. Improving the ranking quality of medical image](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0060) [retrieval using a genetic feature selection method. Decis Support](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0060) [Syst 2011](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0060).
13. [Dy JG, Brodley CE, Kak A, Broderick LS, Aisen AM. Unsuper-](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0065) [vised feature selection applied to content-based retrieval of lung](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0065) [images. IEEE Trans Pattern Anal Mach Intell 2003;25(3):373–8](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0065).
14. [Lu J, Zhao T, Zhang Y. Feature selection based-on genetic](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0070) [algorithm for image annotation. Knowl-Based Syst](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0070) [2008;21:887–91](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0070).
15. [Whitney AW. A direct method of non-parametric measurement](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0075) [selection. IEEE Trans Comput. 1971;C-20(9):1100–3](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0075).
16. [Molina LC, Belanche L, Nebot A](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0080)` [. Feature selection algorithms: a](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0080)

[survey and experimental evaluation. In: Proceedings of the IEEE](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0080) [international conference on data mining, ICDM 2003. IEEE;](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0080) [2002. p. 306–13](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0080).

1. [Kumar M, Rawat TK. Optimal design of FIR fractional order](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0085) [differentiator using cuckoo search algorithm. Expert Syst Appl](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0085) [2014](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0085).
2. [Kumar M, Rawat TK. Optimal fractional delay-IIR filter design](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0090) [using cuckoo search algorithm. ISA Trans 2015](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0090).
3. [Aggarwal Apoorva, Kumar Manjeet, Rawat Tarun Kumar,](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0095) [Upadhyay Dharmendra Kumar. Optimal design of 2D FIR filters](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0095) [with quadrantally symmetric properties using fractional derivative](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0095) [constraints. Circ Syst Signal Process 2016](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0095).
4. [Kumar Manjeet, Rawat Tarun Kumar. Fractional order digital](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0100) [differentiator design based on power function and least-squares.](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0100) [Int J Electron 2016](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0100).
5. [Salton G. Automatic text processing. Reading, Mass: Addison-](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0105) [Wesley; 1989](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0105).
6. [Salton G, McGill MJ. Introduction to modern information](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0110) [retrieval. New York: McGraw-Hill; 1998](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0110).
7. [Zhou Xiang Sean, Huang Thomas S. Relevance feedback in](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0115) [content based image retrieval: some recent advances. J Inform Sci](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0115) [2012;148:129–37](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0115).
8. [Porkaew Kriengkrai, Chakrabarti Kaushik, Mehrotra Sharad.](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0120) [Query refinement for multimedia similarity retrieval in MARS. In:](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0120) [Multimedia’99 Proceedings of the seventh ACM international](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0120) [conference on multimedia (Part 1). p. 235–8](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0120).
9. [Rui Y, Huang TS, Mehrotra S. Content based image retrieval](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0125) [with relevance feedback. In: MARS, Proc IEEE intern conf on](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0125) [image processing, Santa Barbara, CA. p. 815–8](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0125).
10. [Giacinto Giorgio, Roli Fabio, Fumera Giorgio. Adaptive query](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0130) [shifting for content based image retrieval. In: Perner P, editor.](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0130) [MLDM 2001, LNAI 2123. p. 337–46](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0130).
11. [Kherfi ML, Ziou D. Relevance feedback for CBIR: a new](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0135) [approach based on probabilistic feature weighting with positive](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0135) [and negative examples. IEEE Trans Image Process](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0135) [2006;15:1017–30](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0135).
12. [Ishikawa Y, Subramanys R, Faloutsos C. MindReader: querying](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0140) [databases through multiple examples. In: Proceedings of the 24th](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0140) [VLDB conference. p. 433–8](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0140).
13. [Rui Y, Huang TS. Relevance feedback techniques in image](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0145) [retrieval. In: Lew MS, editor. Principles of visual information](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0145) [retrieval. London: Springer; 2001. p. 219–58](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0145).
14. [Duda RO, Hart PE. Pattern classification and scene analysis. New](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0150) [York: Wiley; 1973](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0150).
15. [Mosbah Mawloud, Boucheham Bachir. Relevance feedback](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0155) [within CBIR systems. Int J Comput Inf Sci Eng 2014;8(4):19–23](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0155).
16. <http://Wang.ist.psu.edu/docs/related.shtml>.
17. Stricker M, Orengo M. Similarity of color images: storage and retrieval for image and video database III; 1995.
18. [Marie Deza Michel, Elena Deza. Encyclopedia of distance.](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0170) [Dordrecht Heidelberg, London, New York: Springer; 2009](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0170).
19. [Babu GP, Mehre BM, Kankanhalli MS. Color indexing for](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0175) [efficient image retrieval. Multimedia Tools Appl 1995;1:327–48](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0175).
20. [Fishburn P. Non-linear preference and utility theory. Johns](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0180) [Hopkins University Press; 1998](http://refhub.elsevier.com/S1110-8665(16)30035-4/h0180).