

Unsupervised Distance Learning by Rank Correlation Measures for Image Retrieval

César Yugo Okada¹, Daniel Carlos Guimarães Pedronette¹, Ricardo da S. Torres²

¹Dept. of Statistic, Applied Math. and Computing, Universidade Estadual Paulista (UNESP), Rio Claro, Brazil

²RECOD Lab, Institute of Computing, University of Campinas (UNICAMP), Campinas, Brazil
okada@rc.unesp.br, daniel@rc.unesp.br, rtorres@ic.unicamp.br

ABSTRACT

Ranking accurately collection images is the main objective of Content-based Image Retrieval (CBIR) systems. In fact, the set of images ranked at the first positions generally defines the effectiveness of provided search services, i.e., they are used for assessing automatically the quality of search systems as this set usually contains the collection images that are of interest. Recently, the use of ranking information (e.g., rank correlation) has been used in different research initiatives with the objective of improving the effectiveness of image retrieval tasks. This paper presents a broad rank correlation analysis for unsupervised distance learning on image retrieval tasks. Various well-known rank correlation measures are considered and two new measures are proposed. Several experiments were conducted considering various image datasets involving shape, color, and texture descriptors. Experimental results demonstrate that ranking information can be exploited for distance learning tasks successfully. Evaluated approaches yield better results in terms of effectiveness than various state-of-the-art algorithms.

Categories and Subject Descriptors

H.3.3 [Information Search and Retrieval]: Search process

General Terms

Experimentation, Performance

Keywords

content-based image retrieval; rank correlation; measures

1. INTRODUCTION

The increasing availability of large image collections has demanded the development of novel search and indexing methods. In this scenario, Content-Based Image Retrieval (CBIR) systems have emerged as promising approaches, as they are able of considering the image visual content.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

ICMR '15, June 23–26, 2015, Shanghai, China.

Copyright © 2015 ACM

DOI: <http://dx.doi.org/10.1145/2671188.2749335>.

Typically, the implementation of a content-based image retrieval system relies on the definition of three important concepts in the retrieval pipeline [7]: (i) image representation, (ii) image similarity measure, and (iii) result refinement by post-processing methods. The first two steps were exploited for decades and several visual features and distance measures have been proposed for image retrieval tasks (based on shape, color, and texture properties).

Recently, post-processing methods have attracted a lot of attention, mainly due to the significant effectiveness gains obtained [27, 32, 35]. Different approaches have been employed, as diffusion process [35], graphs [32], and clustering [26]. More recently, ranking approaches have been proposed, presenting important advantages, such as simplicity, low-computational efforts, and total independence of distance measures [27]. Ranking scores and ranking correlation measures have been used as the basis of various of these methods.

This paper discusses methods for improving the effectiveness of image retrieval results based on rank correlation measures. Two novel rank correlation measures are proposed for unsupervised distance learning procedures on image retrieval tasks. The paper also presents an extended version of the recently proposed RL-Sim (*Ranked Lists Similarities*) algorithm [27] that exploits the proposed measures. In summary, the contributions of this paper are threefold: (i) presentation of the RL-Sim* Algorithm, which uses information from both rank correlation measures and top-*k* lists overlap for improving the effectiveness of distance measures; (ii) the evaluation of six different rank correlation measures for the proposed algorithm; and (iii) the proposal of two novel rank correlation measures for unsupervised distance learning.

A large experimental evaluation was conducted, considering different datasets and image descriptors. Experiments were conducted on four image datasets considering 16 different visual descriptors (shape, color, texture, and local descriptors). Other aspects of the proposed approach were also considered, such as the analysis of the impact of parameters and the relationship among the considered rank correlation measures.

The experimental evaluation demonstrates that the effectiveness performance of the proposed RL-Sim* Algorithm is superior to the original algorithm. Experimental results also show that the proposed rank correlation measures achieved significant effectiveness gains in several image retrieval tasks. The proposed RL-Sim* Algorithm is also evaluated in comparison with several state-of-the-art approaches considering a shape dataset commonly used as benchmarking. The pro-

posed unsupervised learning approach yields better results in terms of effectiveness performance than various methods recently proposed in the literature.

The paper is organized as follows: Section 2 discusses related work and Section 3 presents the problem formulation. Section 4 briefly describes the original RL-Sim Algorithm, while Section 5 presents the proposed RL-Sim* Algorithm. In Section 6, we describe the evaluated rank correlation measures and present two novel measures. Section 6 discusses our approach for measure combination. Section 7 presents and discusses conducted experiments and, finally, Section 8 discusses conclusions and presents future work.

2. RELATED WORK

Rank correlation measures have been used in information retrieval applications with the aim of computing similarities or distances between rankings produced by different systems. Generally, these measures are based on simple ranking functions, and provide information about the similarity among different systems used to compute the rankings. Many measures used to compare rankings are based on statistical associations of objects in different ranked lists, such as Kendall τ and Spearman's Footrule [10].

Although various measures have been used along decades, the research field remains active and, even recently, novel measures continue being proposed. The developing of new measures is mainly based on statistical or probabilistic models. The Rank Biased Overlap [33] (RBO) measure, for example, was recently proposed based on a model which compares the overlap of two ranked lists at different increasing depths. The RBO [33] measure consists in a measure adapted to information retrieval, assigning to the top- k objects high weights. In a different venue, some studies are proposing new measures that use classical methods as part of the process, as SKT [14] measure that uses Kendall τ .

More recently, rank correlation measures have been used [7,27] to compare ranked results, aiming at increasing the effectiveness of CBIR systems. Although several approaches were employed in post-processing methods (e.g., graph transduction [34], diffusion process [35], affinity learning [36]), the use of ranking information presents significant advantages and have attracted a lot of research interest.

A ranking consistency model is proposed in [7], using a result refinement procedure that re-ranks images based on an improved image similarity measure. Basically, the rank-consistency information indicates that query images are likely to contain similar content. Rank correlation analysis is also employed for metric learning in content-based medical image retrieval applications [14].

The RL-Sim Algorithm [27] is an unsupervised distance learning method recently proposed that considers the similarity between ranked lists for computing a more effective distance measure. Its main motivation relies on the conjecture that ranked lists encode contextual information in CBIR systems, providing not only a pairwise relation as traditional distance measures, but a more global contextual measure. In addition, the method presents important features such as scalability, and low-computational complexity [24].

This paper presents the RL-Sim* Algorithm, a new post-processing method, which extends the original algorithm by exploiting more contextual information. The proposed method is based on information regarding the overlap be-

tween ranked lists combined with rank correlation measures. In addition, two novel rank correlation measures are proposed aiming at characterizing distance learning procedures on image retrieval tasks.

3. PROBLEM DEFINITION

A formal definition of the image retrieval model considered is presented in this section. Let $\mathcal{C}=\{img_1, img_2, \dots, img_n\}$ be an image collection. Let $n = |\mathcal{C}|$ be the size of the collection \mathcal{C} . Let \mathcal{D} be an image descriptor, which can be defined [8] as a tuple (ϵ, ρ) , where $\epsilon: \hat{I} \rightarrow \mathbb{R}^n$ is a function, which extracts a feature vector v_f from an image \hat{I} ; and $\rho: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ is a distance function that computes the distance between two images according to the distance between their corresponding feature vectors, i.e., the distance between two images img_i and img_j is given by the value of $\rho(\epsilon(img_i), \epsilon(img_j))$. The notation $\rho(i, j)$ is used along the paper for readability purposes.

The distance $\rho(i, j)$ among all images $img_i, img_j \in \mathcal{C}$ can be computed to obtain a squared $n \times n$ distance matrix A , such that $A_{ij} = \rho(i, j)$. Also based on the distance function ρ , a ranked list τ_q can be computed in response to a query image img_q . The ranked list $\tau_q=(img_1, img_2, \dots, img_n)$ can be defined as a permutation of the collection \mathcal{C} . A permutation τ_q is a bijection from the set \mathcal{C} onto the set $[N] = \{1, 2, \dots, n\}$. For a permutation τ_q , we interpret $\tau_q(i)$ as the position (or rank) of image img_i in the ranked list τ_q . We can say that, if img_i is ranked before img_j in the ranked list of img_q , that is, $\tau_q(i) < \tau_q(j)$, then $\rho(q, i) \leq \rho(q, j)$. We also can take every image $img_i \in \mathcal{C}$ as a query image img_q , in order to obtain a set $\mathcal{R} = \{\tau_1, \tau_2, \dots, \tau_n\}$ of ranked lists for each image of the collection \mathcal{C} .

The objective of the unsupervised learning algorithm consists in redefining the initial distance ρ by computing a more effective distance function. The general objective is to improve the effectiveness of distances among images by using the contextual information encoded in the ranked lists defined by the set \mathcal{R} . More formally, we can define the algorithm as a function f_r :

$$\hat{A} = f_r(\mathcal{R}, A). \quad (1)$$

A new distance matrix \hat{A} can be computed by the function f_r , which takes as input the set of ranked lists \mathcal{R} .

4. RL-SIM ALGORITHM

The RL-Sim Algorithm [27] is a recently proposed unsupervised distance learning method that improves the effectiveness of image retrieval tasks though an iterative re-ranking scheme. The RL-Sim Algorithm [27] exploits contextual information encoded in the similarity between ranked lists with the objective of improving the effectiveness of CBIR descriptors. Ranked lists represent a relevant source of information, since ranked lists establish a relationship among a set of images contained in ranked lists, instead of only between pairs of images.

In general, if two images are similar, their ranked lists should be similar as well. Therefore, the main objective of the algorithm is to improve the effectiveness of distance measures by computing the similarity between the ranked lists. The modeling of contextual information considering only the similarity between ranked lists represents an advantage of this approach strategy. Instead of using the dis-

tance information, the algorithm requires only the ranking information.

Given an initial set of ranked lists, an iterative approach is used. Let the superscript (t) denote the current iteration, a new and more effective set of ranked lists $\mathcal{R}^{(t+1)}$ is computed by taking into account distances among ranked lists. Next, $\mathcal{R}^{(t+1)}$ is used for the next execution of our re-ranking algorithm and so on. These steps are repeated along iterations aiming to improve the effectiveness incrementally. After a number T of iterations a definitive re-ranking is performed. In the following, this procedure is detailed.

Ranking Contextual Distance Measure

In this section, we briefly describe the RL-Sim Algorithm [27] by using a ranking contextual distance measure based on similarity/dissimilarity of ranked lists. The ranking contextual distance measure is based on the conjecture that top-ranked images are similar to each other and their ranked lists contain many images in common [27]. In this scenario, one straightforward strategy for computing the similarity between images relies on the use of rank correlation measures.

A ranking contextual distance measure is iteratively learned in a unsupervised setting, by incorporating the contextual information provided by rank correlation measures. Let us consider the neighborhood set $\mathcal{N}(i, k)$ of an image img_i , which contains the k most similar images to img_i , according to a given distance (say ρ defined by the image descriptor). The set $\mathcal{N}(i, k)$ can be obtained by the well-known k -Nearest Neighbor approach, where the cardinality of the set is denoted by $|\mathcal{N}(i, k)| = k$.

Let $d(\tau_i, \tau_j, k)$ denote a rank correlation measure between ranked lists τ_i and τ_j , considering their top- k positions given by the sets $\mathcal{N}(i)$ and $\mathcal{N}(j)$ and defined in the interval $[0, 1]$. A non-iterative contextual distance measure $\rho_c(img_i, img_j)$ based on the comparison of ranked lists τ_i, τ_j can be defined as follows:

$$\rho_c(img_i, img_j) = d(\tau_i, \tau_j, k) \quad (2)$$

Based on the conjecture that the contextual distance measure ρ_c represents a more effective distance between images [27], the distance among all images in a collection can be recomputed based on this measure. Therefore, a new set of ranked lists can be obtained, such that the contextual distance can also be recomputed and the process can be repeated in an iterative way. Let (t) denote the current iteration and let $\tau_i^{(t)}$ denote the ranked list at iteration t . Let $\rho_c^{(0)}$ be the contextual distance at first iteration, which is equal to the distance defined by the image descriptor, such that $\rho_c^{(0)}(img_i, img_j) = \rho(img_i, img_j)$ for all images $img_i, img_j \in \mathcal{C}$. The iterative contextual measure is defined as:

$$\rho_c^{(t+1)}(img_i, img_j) = d(\tau_i^{(t)}, \tau_j^{(t)}, k) \quad (3)$$

It is expected that the effectiveness of the distance measure improves along iterations, so non-relevant images are moved out from the first positions of the ranked lists. In this way, the size of the neighborhood k can be increased for considering more images along iterations. Therefore the contextual measure can be redefined as:

$$\rho_c^{(t+1)}(img_i, img_j) = d(\tau_i^{(t)}, \tau_j^{(t)}, k + t) \quad (4)$$

After a given number of T iterations, a new distance $\hat{\rho}$ is computed based on contextual distance measure ρ_c :

$$\hat{\rho}(img_i, img_j) = \rho_c^{(T)}(img_i, img_j) \quad (5)$$

Finally, using the distance $\hat{\rho}$ a new distance matrix can be computed such $\hat{A}_{ij} = \hat{\rho}(img_i, img_j)$. Based on \hat{A} , a new set of ranked lists $\hat{\mathcal{R}}$ can be also computed.

5. RL-SIM* ALGORITHM

The RL-Sim Algorithm [27] computes a new distance between images img_i, img_j by analyzing the similarity between their respectively ranked lists τ_i, τ_j considering the top- k positions defined by the sets $\mathcal{N}(i, k)$ and $\mathcal{N}(j, k)$.

The effectiveness gain, however, is obtained by redefining the distances among the query image and images at initial positions of ranked lists. It occurs since is very unlikely to found similar images at the end of ranked lists. Therefore, the distances are redefined considering the rank correlation measure $d(\tau_i, \tau_j, k)$ for the first positions of the each ranked list, such that $L \in \mathbb{N}$ and $k \leq L \ll N$. For images in the remaining positions of the ranked lists, the new distance is redefined based on the current distances (or rank positions) and the function $d(\tau_i, \tau_j, k)$ does not need to be computed. As a result, this step of the algorithm depends only on a constant L , and not on the collection size N .

Although this approach allows decreasing the demanded computational costs, it still presents a limitation. Since the rank correlation measures are computed considering the top- k positions defined in terms of the neighborhood set $\mathcal{N}(i, k)$, their accuracy tends to be low when there is no overlap between the ranked lists being compared at top positions. For measures based on intersection analysis, it is still more critical, producing the same distance values for all pairs of images without overlap at top- k positions. In these situations, the effectiveness of distance can be worsened.

Based on this observation, we propose the extended *RL-Sim* Algorithm* which aims at computing a different distance when there is no overlap between the ranked lists being compared. In this way, considering a query image img_i we propose to divide the ranked list τ_i in three segments. Each segment defines a subset of the ranked list, which is processed differently:

(i) First Segment (top- L positions, with overlap): this segment contains an image img_j at top- L positions if the neighborhood sets of img_i, img_j present other images in common. Formally, if $(\tau_i(j) < L) \wedge (\mathcal{N}(i, k) \cap \mathcal{N}(j, k) \neq \emptyset)$. For these cases, the new distance between img_i and img_j is computed by the rank correlation measure (defined in the $[0, 1]$ interval).

(ii) Second Segment (top- L positions, no overlap): if the img_j appears at top- L positions of τ_i , but there is no overlap between top- k positions ($\mathcal{N}(i, k) \cap \mathcal{N}(j, k) = \emptyset$), the rank correlation measures have no enough information for improving the distance measure. In this situations, the current distance is only incremented by 1, such $A_{ij}^{(t+1)} = A_{ij}^{(t)} + 1$. Notice that images with no overlap always will be ranked after that with overlap.

(iii) Remaining Images: the remaining images, i.e., images after top- L positions have their distances incremented by 2, such that $A_{ij}^{(t+1)} = A_{ij}^{(t)} + 2$. This procedure keeps these images at the end of ranked lists, ensuring that they are not mixed with images at top- L positions.

Notice that the simple addition of constants (1 and 2)

are enough to differentiate images at different positions of ranked lists. Algorithm 1 outlines the proposed RL-Sim* Algorithm. We can observe that the conditional structure at Line 8 defines the first (Line 9) and second segments (Line 12). The remaining images are processed at Lines 17-19.

Algorithm 1 RL-Sim* Algorithm

Require: Distance matrix A , Set of ranked lists \mathcal{R} and parameters k, T, L

Ensure: Processed set of ranked lists $\hat{\mathcal{R}}$

```

1:  $t \leftarrow 0$ 
2:  $k \leftarrow k_s$ 
3:  $\mathcal{R}^{(t)} \leftarrow \mathcal{R}$ 
4:  $A^{(t)} \leftarrow A$ 
5: while  $t < T$  do
6:   for all  $\tau_i \in \mathcal{R}^{(t)}$  do
7:     for all  $\{img_j \in \mathcal{C} \mid \tau_i(j) \leq L\}$  do
8:       if  $\mathcal{N}(i, k) \cap \mathcal{N}(j, k) \neq \emptyset$  then
9:         //First Segment - With Overlap
10:         $A^{(t+1)}[i, j] \leftarrow d(\tau_i, \tau_j, k)$ 
11:       else
12:         //Second Segment - No Overlap
13:         $A_{ij}^{(t+1)} \leftarrow A_{ij}^{(t)} + 1$ 
14:       end if
15:     end for
16:   //Remaining Images - After top- $L$  positions
17:   for all  $\{img_j \in \mathcal{C} \mid \tau_i(j) > L\}$  do
18:      $A_{ij}^{(t+1)} \leftarrow A_{ij}^{(t)} + 2$ 
19:   end for
20: end for
21:  $\mathcal{R}^{(t+1)} \leftarrow \text{sortRankedLists}(A^{(t+1)})$ 
22:  $k \leftarrow k + 1$ 
23:  $t \leftarrow t + 1$ 
24: end while
25:  $\hat{\mathcal{R}} \leftarrow \mathcal{R}^{(T)}$ 

```

6. RANK CORRELATION MEASURES

The RL-Sim* (Algorithm 1) is not dependent on the rank correlation measure used, so we can define $d(\tau_i, \tau_j, k)$ using various approaches. However, only two measures were evaluated for the original RL-Sim Algorithm [27].

In fact, the effectiveness gains are directly associated with the measure used. Therefore, for obtaining high effectiveness gains it is required the definition of appropriate measures. In this paper, we perform a comparative study of eight correlation measures: six measures of the literature (Intersection, Kendall τ , Spearman, Goodman, Jaccard, and RBO); and two novel measures, named Jaccard _{l} and Kendall τ_w . These measures are described in the following.

6.1 Intersection Measure

An approach to define the distance between two top- k lists τ_i and τ_j proposed in [10] is to capture the extent of overlap between τ_i and τ_j . This idea of overlap can be extended by considering not only the overlap at depth k , but also the cumulative overlap at increasing depth. For each depth $d \in \{1 \dots k\}$, it is computed the overlap at d , and then those overlaps are averaged to derive a similarity measure. The measure assigns higher weights to the first positions of top k lists, which are considered many times. Equation 6 formally defines the intersection similarity measure ψ .

$$\psi(\tau_i, \tau_j, k) = \frac{\sum_{d=1}^k |\mathcal{N}(i, d) \cap \mathcal{N}(j, d)|}{k} \quad (6)$$

Note that if two ranked lists present the same images at the first positions, the size of the intersection set is larger, and therefore the value of ψ is higher as well. Since we are interested in a distance measure, we define d_I as follows:

$$d_I(\tau_i, \tau_j, k) = \frac{1}{1 + \psi(\tau_i, \tau_j, k)} \quad (7)$$

6.1.1 Kendall τ

The Kendall's τ is a traditional distance measure between permutations, used to measure rank correlation. Its value turns out to be equal to the number of exchanges needed in a bubble sort to convert one permutation to the other [10]. The normalized Kendall's tau measure is defined as follows:

$$d_\tau(\tau_i, \tau_j, k) = \frac{\sum_{x, y \in \mathcal{N}(i, k) \cup \mathcal{N}(j, k)} \bar{K}_{x, y}(\tau_i, \tau_j)}{k \times (k - 1)}, \quad (8)$$

where $\bar{K}_{x, y}(\tau_i, \tau_j)$ is a function that determines if images img_x and img_y are in the same order in compared ranked lists τ_i and τ_j . This function can be defined as follows:

$$\bar{K}_{x, y}(\tau_i, \tau_j) = \begin{cases} 0 & \text{if } (\tau_i(x) \leq \tau_i(y) \wedge \tau_j(x) \leq \tau_j(y)), \\ 0 & \text{if } (\tau_i(x) \geq \tau_i(y) \wedge \tau_j(x) \geq \tau_j(y)), \\ 1 & \text{otherwise.} \end{cases}$$

The maximum value of defined Kendall's tau measure is given by $k \times (k - 1)$, which occurs when $\mathcal{N}(i, k) \cap \mathcal{N}(j, k) = \emptyset$ and τ_i is the reverse of τ_j .

6.1.2 Spearman

The Spearman's metric, commonly denoted by the letter ρ , can be seen as the L1 distance between two permutations [10]. The metric can also be defined as a non-parametric measure, which evaluates the relationship between two variables. Equation 9 formally defines the measure.

$$d_\rho(\tau_i, \tau_j, k) = \frac{\sum_{x, y \in \mathcal{N}(i, k) \cup \mathcal{N}(j, k)} |\tau_i(j) - \tau_j(i)|}{2 \times k \times n}, \quad (9)$$

6.1.3 Goodman

The Goodman and Kruskal's measure [11], commonly denoted by letter γ , is a measure of rank correlation based on pairwise analysis. The measure considers the relationship between the number of discordant (N_d) and concordant pairs (N_s), in a given set. The number of discordant pairs N_d can be defined based on the function $\bar{K}_{x, y}$ (Equation 6.1.1), as follows:

$$N_d(\tau_i, \tau_j, k) = \sum_{x, y \in \mathcal{N}(i, k) \cup \mathcal{N}(j, k)} \bar{K}_{x, y}(\tau_i, \tau_j) \quad (10)$$

We can also define the number of concordant pairs N_s as the inverse of N_d , such $N_s(\tau_i, \tau_j, k) = |\mathcal{N}(i, k) \cup \mathcal{N}(j, k)| - N_d$. Finally, the distance d_γ is formally defined by Equation 11.

$$d_\gamma(\tau_i, \tau_j, k) = \frac{N_s(\tau_i, \tau_j, k) - N_d(\tau_i, \tau_j, k)}{N_s(\tau_i, \tau_j, k) + N_d(\tau_i, \tau_j, k)}. \quad (11)$$

6.1.4 Jaccard

The Jaccard coefficient is a well-known distance between sets. Given two non-empty sets, it measures the probability that an element of at least one of two sets is an element of both, and thus is a reasonable measure of similarity or overlap between the two [18]. The Jaccard Coefficient is defined as follows:

$$J(\tau_i, \tau_j, k) = \frac{|\mathcal{N}(i, k) \cap \mathcal{N}(j, k)|}{|\mathcal{N}(i, k) \cup \mathcal{N}(j, k)|}, \quad (12)$$

As the coefficient $J(\tau_i, \tau_j, k)$ computes a similarity score, the distance measure is defined as:

$$d_J(\tau_i, \tau_j, k) = \frac{1}{1 + J(\tau_i, \tau_j, k)}. \quad (13)$$

6.1.5 Rank Overlap Biased (RBO)

The Rank-Biased Overlap (RBO) is a recently proposed measure [33] that compares the overlap of the two rankings at incrementally increasing depths. The measure takes a parameter that specifies the probability of considering the overlap at the next level. The weight of the overlap measured at each depth is computed based on these probabilities. The RBO measure is defined by Equation 14:

$$RBO(\tau_i, \tau_j, k, p) = (1 - p) \sum_{d=1}^k p^{d-1} \times \frac{|\mathcal{N}(i, k) \cap \mathcal{N}(j, k)|}{d}, \quad (14)$$

where p is a constant. The distance measure can be computed as the inverse of RBO function, as follows.

$$d_R(\tau_i, \tau_j, k) = \frac{1}{1 + RBO(\tau_i, \tau_j, k)}. \quad (15)$$

6.1.6 Kendall τ_w

The original Kendall τ measure is a non-weighted measure. Pairs of images in a discordant order but with a small difference between their positions have the same weight of pairs very distant in the ranking. In this way, we propose a Kendall τ_w measure, which aims at giving higher weights to distant pairs. The measure is defined as follows:

$$d_{\tau_w}(\tau_i, \tau_j, k) = \frac{\sum_{x, y \in \mathcal{N}(i, k) \cup \mathcal{N}(j, k)} \bar{W}_{x, y}(\tau_i, \tau_j)}{n^2 \times k^2 \times (k - 1)}. \quad (16)$$

The measure is very similar to the original Kendall measure, except by the function $\bar{W}_{x, y}$ that computes the weight of each pair. The function $\bar{W}_{x, y}$ is defined as follows:

$$\bar{W}_{x, y}(\tau_i, \tau_j) = \begin{cases} 0 & \text{if } (\tau_i(x) \leq \tau_i(y) \wedge \tau_j(x) \leq \tau_j(y)), \\ 0 & \text{if } (\tau_i(x) \geq \tau_i(y) \wedge \tau_j(x) \geq \tau_j(y)), \\ f \times (k - \min(\tau_i(y), \tau_j(x), \tau_i(x), \tau_j(y))) & \text{otherwise,} \end{cases}$$

where f is a factor that aims at penalizing discordant pairs which are distant in the ranked lists. The pairs are considered distant when the difference between their positions are greater than k . Therefore, the value $f = 2$ is assigned when $(|\tau_i(x) - \tau_i(y)| + |\tau_j(x) - \tau_j(y)|) > (2 \times k)$, and $f = 1$, otherwise.

6.1.7 Jaccard_i

The traditional Jaccard coefficient establishes a relationship between the size of the union and the intersection sets, at a certain depth in the ranked list defined by k . However, in information retrieval tasks this approach ignores the information provided by top positions, smaller than k . In this way, we proposed to compute an accumulated Jaccard score considering different depths, as Intersection and RBO measures. Equation 17 formally defines the proposed approach:

$$J_I(\tau_i, \tau_j, k) = \frac{\sum_{d=1}^k J(\tau_i, \tau_j, d)}{k} \quad (17)$$

As the accumulated score gives a similarity score, the respective distance is defined as follows:

$$d_{J_I}(\tau_i, \tau_j, k) = \frac{1}{1 + J_I(\tau_i, \tau_j, k)}. \quad (18)$$

7. EXPERIMENTAL EVALUATION

A large experimental evaluation was conducted aiming at assessing the effectiveness of the presented approach. Section 7.1 describes the image descriptors and datasets considered. Section 7.2 discusses the impact of parameters. Sections 7.3 and 7.4 present the experimental results for the original and the proposed algorithm considering all discussed rank correlation measures. Various shape, color, and texture descriptors were considered. Section 7.5 presents a comparison with other state-of-the-art approaches and Section 7.6 discusses the correlation among different measures.

7.1 Datasets and Image Descriptors

The experimental evaluation was conducted considering four different datasets with diverse characteristics and size ranging from 280 to 10,200 images. Various local and global descriptors were used, considering shape, color, and texture properties. Table 1 summarizes information about datasets and descriptors. All images of each dataset are considered as query images for effectiveness evaluation purposes. The Mean Average Precision (MAP) was used as effectiveness measure for most of datasets. For the N-S [22] dataset, the N-S retrieval score is used and for the MPEG-7 [17] dataset, the Recall@40 is considered in addition to MAP.

7.2 Analysis of Parameters

The computation of RL-Sim* Algorithm considers three parameters: (i) k : the size of the neighborhood set; (ii) T : the number of iterations; and (iii) L : the position at which the ranked lists are considered in the distance learning procedure. To evaluate the impact of different parameter settings on the effectiveness of the method and for determining the best parameters values, we conducted various experiments considering the MPEG-7 [17] dataset.

The first experiment aims at analyzing the impact of parameters k and T on the effectiveness of the proposed method. We set $L = 700$ and varied the values of parameter k and T in the intervals $[5, 20]$ and $[1, 5]$, respectively. We considered the RBO [33] measure and the CFD [25] descriptor. The MAP was used as effectiveness measure. Figure 1 illustrates the obtained results. A large red region can be observed around $k = 15$, indicating high effectiveness scores and robustness of the method regardless the parameter value variations. We used $L = 700$ and $k = 15$ for all experiments. The only exception is the N-S [22] dataset, which has a very small number of relevant images per class. In this case, we used $k = 5$ and $L = 200$.

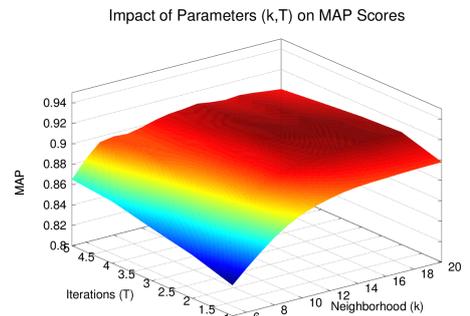


Figure 1: Impact of parameters k, T on MAP scores - CFD [25] descriptor and RBO [33] measure.

The number of iterations is more dependent on the rank correlation measure than the size of neighborhood. There-

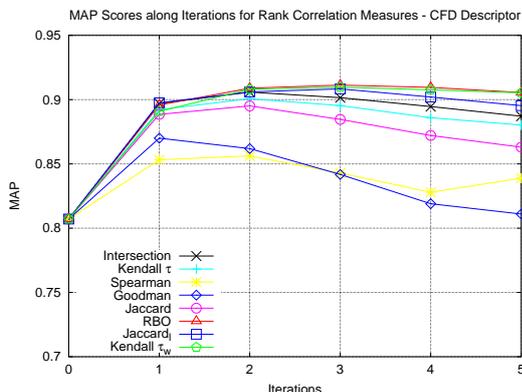
Table 1: Datasets and images descriptors used in the experimental evaluation.

Dataset	Size	Type	General Description	Descriptors	Effectiv. Measure
Soccer [31]	280	Color Scenes	Dataset composed of images from 7 soccer teams, containing 40 images per class	Border/Interior Pixel Classification (BIC) [28], Auto Color Correlograms (ACC) [13], and Global Color Histogram (GCH) [29]	MAP (%)
MPEG-7 [17]	1,400	Shape	A well-known dataset composed of 1400 shapes divided in 70 classes. Commonly used for evaluation of unsupervised distance learning approaches.	Segment Saliences (SS) [9], Beam Angle Statistics (BAS) [1], Inner Distance Shape Context (IDSC) [19], Contour Features Descriptor (CFD) [25], Aspect Shape Context (ASC) [20], and Articulation-Invariant Representation (AIR) [12]	MAP (%), Recall@40
Brodatz [4]	1,776	Texture	A popular dataset for texture descriptors evaluation composed of 111 different textures divided into 16 blocks	Local Binary Patterns (LBP) [23], Color Co-Occurrence Matrix (CCOM) [16], Local Activity Spectrum (LAS) [30]	MAP (%)
N-S [22]	10,200	Objects/Scenes	Composed of 2,550 objects or scenes. Each object/scene is captured 4 times from different viewpoints, distances and illumination conditions	ACC [13], BIC [28], Color and Edge Directivity Descriptor (CEED) [5], Fuzzy Color and Texture Histogram (FCTH) [6], Joint Composite Descriptor (JCD) [37], Scale-Invariant Feature Transform (SIFT) [21]	N-S score

Table 2: Effectiveness of the RL-Sim [27] Algorithm for various rank correlation measures. MAP scores (%) for shape, color, and texture descriptors considering different datasets. Best score for each descriptor in bold.

Descriptor	Type	Initial MAP	Intersection	Kendall τ	Spearman	Goodman	Jaccard	RBO	Jaccard $_l$	Kendall τ_w
SS [9]	Shape	37.67	43.06	44.24	44.53	40.95	45.18	43.34	46.11	46.60
BAS [1]	Shape	71.52	74.57	73.25	75.78	61.66	73.79	76.39	76.58	76.01
IDSC [19]	Shape	81.70	86.75	86.93	84.26	69.79	86.09	87.76	87.25	86.67
CFD [25]	Shape	80.71	88.97	88.40	85.32	72.43	88.72	90.64	90.51	89.13
ASC [20]	Shape	85.28	88.81	88.10	87.11	70.26	88.46	89.70	89.75	89.39
AIR [12]	Shape	89.39	93.54	96.27	97.86	84.84	96.68	94.84	96.43	96.44
GCH [29]	Color	32.24	33.66	32.96	34.66	32.16	32.65	33.13	33.23	33.60
ACC [13]	Color	37.23	43.54	44.29	45.23	39.49	44.81	41.83	43.33	43.59
BIC [28]	Color	39.26	43.45	43.76	46.28	39.31	43.19	42.58	43.86	42.23
LBP [23]	Texture	48.40	47.77	45.20	47.07	44.80	45.55	50.22	49.28	48.53
CCOM [16]	Texture	57.57	62.01	60.30	58.52	56.68	60.72	63.62	63.49	61.67
LAS [30]	Texture	75.15	77.81	75.62	75.13	69.29	76.04	79.31	79.00	78.30
Average	-	61.34	65.33	64.86	65.14	56.80	60.61	66.11	66.57	66.01

fore, we conducted an experiment for evaluating the impact of T considering different measures in order to defining the best values for this parameter. Figure 2 shows the results. The CFD [25] descriptor and the MPEG-7 [17] dataset were used, but other experiments were conducted considering other descriptors and datasets with similar results. The value of T was defined as 3, 2, 1, 1, 2, 3, 2, and 2 for the measures Intersection, Kendall τ , Spearman, Goodman, Jaccard, and RBO, Jaccard $_l$, and Kendall τ_w , respectively. We used $p = 0.9$ as suggested by [33] for RBO measure.


Figure 2: Evolution of MAP scores along Iterations for various Rank Correlation Measures.

7.3 Rank Correlation Measures for RL-Sim Algorithm

This section presents the evaluation of the two rank correlation measures proposed and the other six measures discussed in Section 6 for the original RL-Sim [27] Algorithm. Table 2 presents the results considering MAP scores for twelve descriptors and three datasets.

We can note that, considering shape descriptors, there is not a clear best measure, although RBO, Jaccard $_l$, and Kendall τ_w measures have reached higher scores. The Spearman measure obtained the best MAP scores for color descriptors, while RBO measure obtained the best results considering texture descriptors. In a more general analysis, considering the average of all descriptors, Jaccard $_l$, RBO, and Kendall τ_w obtained the best scores.

7.4 Rank Correlation Measures for RL-Sim* Algorithm

We conducted a set experiments aiming at evaluating the proposed RL-Sim* Algorithm, considering different rank correlation measures and comparing the results with the original RL-Sim [27] Algorithm.

Table 3 presents the experimental results considering MAP scores for shape, color, and texture descriptors. We can observe a significant increase in the best value observed for each descriptor in most cases, when compared with Table 2. The Kendall τ_w and RBO measures obtained the best scores for most descriptors. The Jaccard $_l$ and Kendall τ_w proposed measures reached the best average scores. We can also observe an increase of the average values for all measures evaluated in comparison with Table 2, demonstrating the effectiveness of proposed of RL-Sim* Algorithm.

We also considered the Bull's Eye Score, commonly used in the literature for the MPEG-7 [17] dataset. This score counts all matching objects within the 40 most similar candidates. Since each class consists of 20 objects, the retrieved score is normalized with the highest possible number of hits and is equivalent to Recall@40. Table 4 presents the obtained results.

Table 5 presents the experimental results for the N-S [22]

Table 3: Effectiveness of the RL-Sim* Algorithm for various rank correlation measures. MAP scores (%) for shape, color, and texture descriptors considering different datasets. Best score for each descriptor in bold.

Descriptor	Type	Initial MAP	Intersection	Kendall τ	Spearman	Goodman	Jaccard	RBO	Jaccard $_l$	Kendall τ_w
SS [9]	Shape	37.67	44.10	46.54	44.70	44.11	45.49	44.23	46.60	47.87
BAS [1]	Shape	71.52	76.05	76.12	75.94	72.61	74.87	77.34	77.51	77.75
IDSC [19]	Shape	81.70	87.38	87.59	85.03	83.50	87.03	88.26	88.08	88.02
CFD [25]	Shape	80.71	90.15	90.07	86.57	86.19	89.51	91.13	90.91	90.81
ASC [20]	Shape	85.28	89.96	90.14	88.03	85.75	89.54	90.57	90.77	90.84
AIR [12]	Shape	89.39	96.17	95.94	97.86	96.08	97.72	96.08	96.78	97.23
GCH [29]	Color	32.24	33.99	33.93	34.29	33.38	33.43	33.99	34.04	34.39
ACC [13]	Color	37.23	45.19	45.94	44.91	42.77	45.63	44.03	44.60	45.75
BIC [28]	Color	39.26	45.42	45.10	45.40	43.05	44.56	45.10	45.47	45.50
LBP [23]	Texture	48.40	48.83	48.94	47.32	49.06	46.53	51.00	50.10	49.92
CCOM [16]	Texture	57.57	62.89	62.44	59.02	61.19	61.37	64.23	64.06	63.53
LAS [30]	Texture	75.15	78.58	78.49	75.74	75.76	76.76	79.80	79.57	79.34
Average	-	61.34	66.56	66.77	65.4	59.93	66.04	67.15	67.37	67.59

Table 4: RL-Sim* Algorithm: Rank Correlation Measures on the MPEG-7 [17] dataset - Recall@40 (%).

Descriptor	Original	Intersection	Kendall τ	Spearman	Goodman	Jaccard	RBO	Jaccard $_l$	Kendall τ_w
SS [9]	43.99	52.92	53.96	51.97	51.5	53.76	51.61	53.92	54.92
BAS [1]	75.20	82.94	82.25	80.91	81.93	81.32	83.45	83.23	82.86
IDSC [19]	85.40	92.20	91.73	88.68	90.79	91.18	92.73	92.15	91.70
CFD [25]	84.43	94.16	93.77	89.64	92.32	93.20	94.67	94.43	94.24
ASC [20]	88.39	94.66	94.06	91.35	92.85	93.74	94.67	94.39	94.11
AIR [12]	93.67	99.90	98.27	100	99.79	99.95	99.87	99.91	99.92
Average	78.51	88.13	85.67	83.76	84.87	85.52	86.17	86.34	86.29

Table 5: Rank Correlation Measures for the RL-Sim* Algorithm on the N-S Dataset [22].

Descriptor	Initial Score	Intersection	Kendall τ	Spearman	Goodman	Jaccard	RBO	Jaccard $_l$	Kendall τ_w
ACC [13]	3.36	3.54	3.51	3.45	1.06	3.47	3.54	3.55	3.52
BIC [28]	3.04	3.20	3.17	3.02	1.05	3.13	3.19	3.21	3.19
CEED [5]	2.61	2.75	2.71	2.56	1.04	2.68	2.75	2.76	2.74
FCTH [6]	2.73	2.84	2.79	2.63	1.04	2.77	2.84	2.85	2.81
JCD [37]	2.79	2.92	2.88	2.72	2.87	2.85	2.92	2.93	2.9
SIFT [21]	2.54	2.81	2.82	2.86	1.03	2.77	2.79	2.80	2.80
Average	2.84	3.01	2.98	2.87	1.35	2.94	3.01	3.02	2.99

dataset.¹ Each image is used as query and the N-S retrieval score [22] between 1 and 4 is computed. The score corresponds to the number of relevant images among the first four image returned (the highest achievable score is 4). The propose Jaccard $_l$ measure obtained the best scores for this dataset.

A summary of the best MAP results is presented in Table 6, considering the twelve descriptors and three datasets. The table also presents the relative effectiveness gains obtained for each descriptor. We can observe very significant gains ranging from +5.37% to +27.08%.

Table 6: Summary of best MAP Scores (%).

Descriptor	Initial MAP	Rank Measure	RL-Sim* MAP	Relative Gain (%)
SS [9]	37.67	Kendall τ_w	47.87	+27.08
BAS [1]	71.52	Kendall τ_w	77.75	+8.71
IDSC [19]	81.70	RBO	88.26	+8.03
CFD [25]	80.71	RBO	91.13	+12.91
ASC [20]	85.28	Kendall τ_w	90.84	+6.52
AIR [12]	89.39	Spearman	97.86	+9.48
GCH [29]	32.24	Kendall τ_w	34.39	+6.67
ACC [13]	37.23	Kendall τ	45.94	+23.40
BIC [28]	39.26	Kendall τ_w	45.50	+15.89
LBP [23]	48.40	RBO	51.00	+5.37
CCOM [16]	57.57	RBO	64.23	+11.57
LAS [30]	75.15	RBO	79.80	+6.19

7.5 Comparison with Other Approaches

We also evaluated the proposed approach in comparison with state-of-the-art post-processing methods. We considered the MPEG-7 [17] dataset and the Bull’s Eye Score, commonly used for post-processing methods evaluation and

comparison. Table 7 presents the comparison, with two selected best results of our approach. We can observe that high scores were obtained by the proposed approach, comparable or superior to various recently proposed methods.

7.6 Correlation Among Measures

Finally, we also conducted an experiment aiming at evaluating the correlation among measures. The hypothesis is that non-correlated measures encode diverse information, what justifies the evaluation of different measures and enables the investigation of approaches for combining them. The Pearson Correlation Coefficient was computed between each pair of measures, based on distances given by each measure for top images at ranked lists. We considered top 200 images, $k = 15$ and $T = 1$. Table 8 presents the results, with hot colors indicating low correlation scores. As we can observe, various measures present low correlation. That opens a novel research venue related to the investigation of approaches for combining the information provided by ranked lists obtained considering the use of different correlation measures.

8. CONCLUSIONS

In this paper, we have presented an unsupervised distance learning approach based on rank correlation measures. We propose the RL-Sim* Algorithm, which considers the rank correlation measures and the overlap between the neighborhood sets aiming at computing a more effective distance measure. Six traditional measures were evaluated and two novel rank correlation measures were proposed specially for the unsupervised distance learning problem on image retrieval tasks.

¹For the N-S [22] dataset we used $T = 1$ for all measures.

Table 7: Post-processing methods comparison on the MPEG-7 [17] dataset - Bull’s Eye Score (Recall@40).

Shape Descriptors		
CFD [25]	-	84.43%
IDSC [19]	-	85.40%
SC [3]	-	86.80%
ASC [20]	-	88.39%
AIR [12]	-	93.67%
Post-Processing Methods		
Algorithm	Descriptor(s)	Score
Graph Transduction [34]	IDSC	91.00%
Locally C. Diffusion Process [35]	IDSC	93.32%
Shortest Path Propagation [32]	IDSC	93.35%
Locally C. Diffusion Process [35]	ASC	95.96%
Pairwise Recommendation [26]	CFD	96.15%
Tensor Product Graph [36]	ASC	96.47%
Self-Smoothing Operator [15]	SC+IDSC	97.64%
Co-Transduction [2]	SC+IDSC	97.72%
Pairwise Recommendation [26]	CFD+IDSC	99.52%
RL-Sim* with Jaccard	AIR	99.95%
Tensor Product Graph [36]	AIR	99.99%
RL-Sim* with Spearman	AIR	100%

Table 8: Correlation Among Measures.

Rank Measures	(I)	(K)	(S)	(G)	(J)	(R)	(J _i)	(K _w)
(I) Intersection	1	0.23	0.31	0.74	0.77	0.65	0.65	0.31
(K) Kendall τ	0.23	1	0.75	0.22	0.55	0.61	0.63	0.98
(S) Spearman	0.31	0.75	1	0.19	0.52	0.47	0.65	0.76
(G) Goodman	0.74	0.22	0.19	1	0.82	0.86	0.86	0.28
(J) Jaccard	0.77	0.55	0.52	0.82	1	0.86	0.91	0.55
(R) RBO	0.65	0.61	0.47	0.86	0.86	1	0.99	0.65
(J _i) Jaccard _i	0.65	0.63	0.49	0.86	0.91	0.99	1	0.66
(K _w) Kendall τ_w	0.31	0.98	0.76	0.28	0.55	0.65	0.66	1

We conducted a large set of experiments for assessing the effectiveness of the proposed approach, considering different descriptors, datasets and rank correlation measures. Experimental results demonstrated that high effectiveness results were obtained by the proposed approach when compared with recent related work [27] and other state-of-the-art methods. Future work focuses on the investigation of: (i) novel rank correlation measures; and (ii) approaches for combining different rank correlation measures.

9. ACKNOWLEDGMENTS

The authors are grateful to PROPe/UNESP, São Paulo Research Foundation - FAPESP (grants 2013/08645-0 and 2013/50169-1), CNPq (grants 306580/2012-8 and 484254/2012-0), CAPES, AMD, and Microsoft Research.

10. REFERENCES

- [1] N. Arica and F. T. Y. Vural. BAS: a perceptual shape descriptor based on the beam angle statistics. *Pattern Recognition Letters*, 24(9-10):1627–1639, 2003.
- [2] X. Bai, B. Wang, X. Wang, W. Liu, and Z. Tu. Co-transduction for shape retrieval. In *ECCV*, volume 3, pages 328–341, 2010.
- [3] S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. *PAMI*, 24(4):509–522, 2002.
- [4] P. Brodatz. *Textures: A Photographic Album for Artists and Designers*. Dover, 1966.
- [5] S. A. Chatzichristofis and Y. S. Boutalis. Cedd: color and edge directivity descriptor: a compact descriptor for image indexing and retrieval. In *ICVS’08*.
- [6] S. A. Chatzichristofis and Y. S. Boutalis. Fctth: Fuzzy color and texture histogram - a low level feature for accurate image retrieval. In *WIAMIS ’08*, pages 191–196, 2008.
- [7] Y. Chen, X. Li, A. Dick, and R. Hill. Ranking consistency for image matching and object retrieval. *Pattern Recognition*, 47(3):1349 – 1360, 2014.
- [8] R. da S. Torres and A. X. Falcão. Content-Based Image Retrieval: Theory and Applications. *Revista de Informática Teórica e Aplicada*, 13(2):161–185, 2006.

- [9] R. da S. Torres and A. X. Falcão. Contour Saliency Descriptors for Effective Image Retrieval and Analysis. *Image and Vision Computing*, 25(1):3–13, 2007.
- [10] R. Fagin, R. Kumar, and D. Sivakumar. Comparing top k lists. In *SODA’03*, pages 28–36, 2003.
- [11] L. A. Goodman and W. H. Kruskal. Measures of association for cross classifications. *Journal of the American Statistical Association*, 49(268):pp. 732–764, 1954.
- [12] R. Gopalan, P. Turaga, and R. Chellappa. Articulation-invariant representation of non-planar shapes. In *ECCV’2010*, volume 3, pages 286–299, 2010.
- [13] J. Huang, S. R. Kumar, M. Mitra, W.-J. Zhu, and R. Zabih. Image indexing using color correlograms. In *CVPR*, pages 762–768, 1997.
- [14] W. Huang, K. Chan, H. Li, J. Lim, J. Liu, and T. Wong. Content-based medical image retrieval with metric learning via rank correlation. In *MLMI*, pages 18–25, 2010.
- [15] J. Jiang, B. Wang, and Z. Tu. Unsupervised metric learning by self-smoothing operator. In *ICCV*, pages 794–801, 2011.
- [16] V. Kovalev and S. Volmer. Color co-occurrence descriptors for querying-by-example.
- [17] L. J. Latecki, R. Lakmper, and U. Eckhardt. Shape descriptors for non-rigid shapes with a single closed contour. In *CVPR*, pages 424–429, 2000.
- [18] M. Levandowsky and D. Winter. Distance between sets. *Nature*, 243:34 – 35, 1971.
- [19] H. Ling and D. W. Jacobs. Shape classification using the inner-distance. *PAMI*, 29(2):286–299, 2007.
- [20] H. Ling, X. Yang, and L. J. Latecki. Balancing deformability and discriminability for shape matching. In *ECCV*, volume 3, pages 411–424, 2010.
- [21] D. Lowe. Object recognition from local scale-invariant features. In *ICCV*, pages 1150–1157, 1999.
- [22] D. Nistér and H. Stewénius. Scalable recognition with a vocabulary tree. In *CVPR*, volume 2, pages 2161–2168, 2006.
- [23] T. Ojala, M. Pietikäinen, and T. Mäenpää. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *PAMI*, 24(7):971–987, 2002.
- [24] D. C. G. Pedronette, J. Almeida, and R. da S. Torres. A scalable re-ranking method for content-based image retrieval. *Information Sciences*, 265(0):104 – 104, 2014.
- [25] D. C. G. Pedronette and R. da S. Torres. Shape retrieval using contour features and distance optimization. In *VISAPP*, volume 1, pages 197 – 202, 2010.
- [26] D. C. G. Pedronette and R. da S. Torres. Exploiting pairwise recommendation and clustering strategies for image re-ranking. *Information Sciences*, 207:19–34, 2012.
- [27] D. C. G. Pedronette and R. da S. Torres. Image re-ranking and rank aggregation based on similarity of ranked lists. *Pattern Recognition*, 46(8):2350–2360, 2013.
- [28] R. O. Stehling, M. A. Nascimento, and A. X. Falcão. A compact and efficient image retrieval approach based on border/interior pixel classification. In *CIKM*, pages 102–109, 2002.
- [29] M. J. Swain and D. H. Ballard. Color indexing. *International Journal on Computer Vision*, 7(1):11–32, 1991.
- [30] B. Tao and B. W. Dickinson. Texture recognition and image retrieval using gradient indexing. *JVCIR*, 11(3):327–342, 2000.
- [31] J. van de Weijer and C. Schmid. Coloring local feature extraction. In *ECCV*, pages 334–348, 2006.
- [32] J. Wang, Y. Li, X. Bai, Y. Zhang, C. Wang, and N. Tang. Learning context-sensitive similarity by shortest path propagation. *Pattern Recognition*, 44(10-11):2367–2374, 2011.
- [33] W. Webber, A. Moffat, and J. Zobel. A similarity measure for indefinite rankings. *ACM Transactions on Information Systems*, 28(4):20:1–20:38, 2010.
- [34] X. Yang, X. Bai, L. J. Latecki, and Z. Tu. Improving shape retrieval by learning graph transduction. In *ECCV*, volume 4, pages 788–801, 2008.
- [35] X. Yang, S. Koknar-Tezel, and L. J. Latecki. Locally constrained diffusion process on locally densified distance spaces with applications to shape retrieval. In *CVPR*, pages 357–364, 2009.
- [36] X. Yang, L. Prasad, and L. Latecki. Affinity learning with diffusion on tensor product graph. *PAMI*, 35(1):28–38, 2013.
- [37] K. Zagoris, S. Chatzichristofis, N. Papamarkos, and Y. Boutalis. Automatic image annotation and retrieval using the joint composite descriptor. In *14th PCI*, pages 143–147, 2010.