Preprint version of paper published in Pattern Recognition journal, 2013.
DOI: http://doi.org/10.1016/j.patcog.2013.01.004

Image Re-Ranking and Rank Aggregation based on Similarity of Ranked Lists

Daniel Carlos Guimarães Pedronette and Ricardo da S. Torres

RECOD Lab - Institute of Computing (IC)
University of Campinas (UNICAMP)
Campinas, Brazil
dcarlos@ic.unicamp.br, rtorres@ic.unicamp.br

Abstract

In Content-based Image Retrieval (CBIR) systems, ranking accurately collection images is of great relevance. Users are interested in the returned images placed at the first positions, which usually are the most relevant ones. Collection images are ranked in redincreasing order of their distance to the query pattern (*e.g.*, query image) defined by users. Therefore, the effectiveness of these systems is very dependent on the accuracy of the distance function adopted. In this paper, we present a novel context-based approach for redefining distances and later re-ranking images aiming to improve the effectiveness of CBIR systems. In our approach, distances among images are redefined based on the similarity of their ranked lists. Conducted experiments involving shape, color, and texture descriptors demonstrate the effectiveness of our method.

Keywords: content-based image retrieval, re-ranking, ranked lists

1. Introduction

The huge growth of image collections and multimedia resources available and accessible through various technologies is remarkable [22]. Technological improvements in image acquisition and the decreasing cost of storage devices have enabled the dissemination of large image collections. redTraditional image retrieval approaches based on keywords and textual metadata face serious challenges, since describing the image content with textual descriptions is intrinsically very difficult [8], mainly due to the huge growth of image collections. Many applications, especially those dealing with large general datasets face obstacles to obtain textual descriptors, since manual annotation is prohibitively expensive. It is laborious and time-consuming. This task has not been made easier by the diversification of image collections.

One of the most common approaches to overcome these limitations relies on the use of Content-Based Image Retrieval (CBIR) systems. The objective of CBIR systems is to return the most similar images given an image query, considering visual features, such as shape, color, and texture.

In this scenario, ranking accurately collection images is of great relevance. Collection images are ranked in increasing order of their distance to the query pattern (*e.g.*, query image) defined by users. Therefore, choosing a good distance measure is often critical to building an effective CBIR system. In general, CBIR systems consider only pairwise image analysis, that is, compute similarity measures considering only pairs of images, ignoring the rich information encoded in the relations among several images. On the other hand, the user perception usually considers the query specification and responses in a given *context*. Therefore, the distance between two images can be correctly described only if it is considered in the *context* of other images that are similar to them. This requires having a model to capture the essence of a similarity among images instead of viewing each image as a set of points or a feature vector [42].

red Context can be broadly defined as all information about the whole situation relevant to an application and its set of users. In the image retrieval scenario, ranked lists represent a relevant source of contextual information, since given a query image, users do not analyze only pairs of images, but the ranked list as a whole. It is expected, for example, that images ranked at the top positions of ranked lists are similar to each other. It is also expected that, if we take one of these images as a query image, the computed ranked list contains many images in common. In this scenario, we use the term context for denoting any analysis that, instead of considering only pairwise image comparisons, takes into account other information encoded in both ranked lists and distances among all images. More specifically, the notion of *context* can refer to updating image similarity measures by taking into account information encoded in the ranked lists defined by a CBIR system.

In the past few years, there has been considerable research on improving the distance measures in CBIR systems [42, 43, 13, 12, 38, 44, 3, 19, 26, 29, 22, 20, 24]. Promising results have been obtained considering several approaches and techniques. Several unsupervised approaches have been proposed aiming to improve the effectiveness of retrieval tasks, replacing pairwise similarities by more global affinities measures that also consider the relation among the database images. The objective of these methods is somehow mimic the human behavior on judging the similarity among objects by considering specific *contexts*.

In this paper, we present the *RL-Sim (red**R**anked **L**ists **Sim**ilarities) Re-Ranking Algorithm* , a new post-processing method that considers the similarity among ranked lists for characterizing contextual information in CBIR systems. The main motivation of our re-ranking algorithm relies on the conjecture that *contextual information encoded in the similarity between ranked lists can provide resources for improving the effectiveness of CBIR descriptors.* red In general, only two images are considered for distance computation and, if the distance measure adopted is not accurate, the two images will be wrongly placed in the ranked lists of each other. On the other hand, for obtaining a ranked list, many distances have to be computed. These incorrect scores are often mixed with correct values, specially in the beginning of the ranked lists. In this way, the contextual information provided by the ranked lists can be used for correcting the wrong scores. Beyond that, if two images are similar, their ranked lists should be similar as well [21]. It is somehow close to the the cluster hypothesis [27], which states that *"closely associated documents tend to be relevant to the same requests"*. redThe main contributions of this paper are:

(***i***) the modelling of contextual information considering only the similarity between ranked lists, independent of distance (or similarity) scores between images. Distance scores computed by different image descriptor usually are in different scales and requires normalization procedures. These variations can affect the effectiveness of re-ranking approaches. Since the proposed re-ranking method does not depend on distances or similarity scores, it can be easily used for different CBIR tasks and can be adapted for other information retrieval tasks (e.g., text or multimodal retrieval);

(***ii***) the proposed contextual distance measure does not depend on specific approaches for comparing the ranked lists. In this way, our re-ranking algorithm can use different similarity/distance measures among ranked lists, a well-established research area [7, 39, 41]. Therefore, the re-ranking algorithm can be easily extended for using and combining different approaches for comparing ranked lists. The proposition of a generic iterative approach based on ranked lists represents an important contribution, since other re-ranking methods can be proposed just by creating new metrics to compare ranked lists.

This paper differs from previous work [25] as it presents a deeper analysis of RL-Sim Re-Ranking Algorithm, discusses different distance measures used for computing the similarity of ranked lists and extends the experimental protocol.

A large experimental evaluation was conducted, considering three different datasets and twelve different image descriptors (shape, color, and texture descriptors). Other aspects of the proposed algorithm were also considered, such as the analysis of the efficiency of the method and the impact of parameters. Our experimental evaluation demonstrates that the proposed method can achieve significant improvements in various CBIR tasks. In addition, we also evaluated the proposed RL-Sim in comparison with several other state-of-the-art approaches considering a common shape dataset. Experimental results demonstrate that the proposed method yields better results in terms of effectiveness performance than various post-processing algorithms recently proposed in the literature.

The paper is organized as follows: Section 2 discusses related work; Section 3 discusses the definition of the image re-ranking problem; in Section 4, we present our approach for unsupervised distance learning based on the similarity of ranked lists. Section 5 discusses approaches for comparing ranked lists; Section 6 presents the RL-Sim Re-Ranking Algorithm; in Section 7, we describe how the proposed algorithm can be used in rank aggregation tasks; Section 8 presents the experimental evaluation and, finally, Section 9 discusses the conclusions and presents future work.

2. Related Work

Determining the appropriate distance measures plays a key role in many multimedia applications, including classification, clustering, and retrieval tasks. For example, choosing a good distance measure is often critical for building an effective content-based image retrieval (CBIR) system. In general, aiming at retrieving the most similar images to a given query image, CBIR systems compute a predefined distance measure between the query image and each collection image. Traditional distance measures, as Euclidean distance, are often adopted and consider the pairwise similarity between any two images. In many situations, these approaches fail to return satisfactory results, mainly due to the well-known semantic gap challenge [10].

In the past few years, there has been considerable research on improving the distance measures in CBIR systems [42, 43, 13, 12, 38, 44, 3, 19, 26, 29, 32, 46, 22, 20, 24, 23]. Promising results have been obtained considering several approaches and techniques. In this paper, we focus on unsupervised approaches. In *unsupervised learning* approaches, the "*learning*" method considers only the domain of object instances and no training labeled data are provided. Since labeling often is a laborious and time consuming task, whereas unlabeled data is far easier to obtain, unsupervised learning represents a very attractive solution in many situations.

In CBIR applications, the use of context may play an important role. In general, traditional CBIR systems perform only pairwise image analysis, that is, they compute similarity (or distance) measures considering only pairs of images, ignoring the rich information encoded in the relations of several images. However, in recent years, several CBIR approaches [42, 43, 13, 12, 38, 44, 3, 19, 26, 29, 22, 20, 24, 25] have been proposed aiming to improve the effectiveness of retrieval tasks replacing pairwise similarities by more global affinity measures that also consider the relation among the database objects [44]. Although using a very diverse nomenclature (re-ranking [19, 29, 22, 20, 24, 25], graph transduction [42, 3], diffusion process [43], affinity learning [44], contextual similarity/dissimilarity measures [38, 12, 26]), these *post-processing* methods have in common the fact that all approaches propose improving the effectiveness of image searches by exploiting the information about the relationships among collection images in an *unsupervised* way (with no training data). Another important common point consists in the use of an iterative strategy adopted to process contextual information [42, 43, 12, 22, 20, 24, 25].

red A *graph-based transductive learning algorithm* [42] was proposed for shape retrieval tasks. It learns a better metric through graph transduction by propagating the model through existing shapes, in a way similar to computing geodesics in dataset manifold. The method does not require learning the shape manifold explicitly and it does not require knowing class labels of existing shapes. The better metric is learned by collectively propagating the similarity measures to the query shape and between the existing shapes through graph transduction. Although inspired by label propagation algorithm [47], which is semi-supervised, the shape retrieval was treated as an unsupervised problem.

red The *locally constrained diffusion process* [43] considers that the distance between two shapes can be correctly described only if it is considered in the context of other shapes similar to them. The work observes that, since differences between shapes in the same class can be very large and differences between shapes in different classes can be very small, no pairwise shape comparison can describe shape dissimilarity correctly. The influence of other shapes is propagated as a diffusion process on a graph formed by a given set of shapes. The weights of graph edges are defined by applying a Gaussian to the shape distance. A reversible Markov chain based on the graph is constructed and used to propagate the influence of shapes. Another approach based on propagating the similarity information in a weighted graph is called *affinity learning* [44]. Instead of propagating the similarity information on the original graph, it uses a tensor product graph (TPG) obtained by the tensor product of the original graph with itself.

Graphs are also used in other approaches. redA modified mutual kNN graph [13] is proposed as the underlying representation used for shape retrieval. The structure of the shape manifold is estimated from the shape similarity scores among all the shapes within a database. redA shortest-path propagation algorithm [38] was also proposed for shape/object retrieval tasks. Given a query object and a target database object, it explicitly finds the shortest path between them in the distance manifold of the database objects. Then a new distance measure is learned based on the shortest path and it is used to replace the original distance measure.

Beside graph methods, *context* is a term frequently used for designating post-processing methods that consider relationships among images. In general interactive applications, the use of *context* can play an important role, which can be broadly defined as all information about the whole situation relevant to an application and its set of users [1]. In CBIR systems, it is related to the fact that, when humans have to judge the similarity between two images, they always do so in a given context, i.e. they do not consider only the two objects to be compared [26].

redA *contextual dissimilarity measure* [12] was introduced aiming to improve the accuracy of bag-of-features-based image search. The proposed measure takes into account the local distribution of the vectors and updates distances by modifying the neighborhood structure. The dissimilarity measure improves the symmetry of the k-neighborhood relationship by iteratively regularizing the average distance of each vector to its neighborhood. The method performs a global analysis of properties in small overlapping neighborhoods, resembling methods for non-linear dimensionality reduction, inspired by ISOMAP [34] and LLE [28].

redA family of contextual measures [26] was proposed considering the similarity between two distributions measured in the context of a third distribution. These contextual measures are then applied to the image retrieval problem. In such a case, the context is estimated from the neighbors of a query. Using different contexts, and especially contexts at multiple scales (i.e., broad and narrow contexts), provides different views on the same problem, while combining the different views can improve retrieval accuracy.

Contextual information has also been exploited for *re-ranking* methods. Re-ranking can be broadly defined as a process of refining the search results: the re-ranking methods take an initial ranking and aggregate some information for improving the effectiveness of the retrieval process. redA re-ranking approach based on contextual spaces [24] aims at exploiting the relationships among images to improve the effectiveness of CBIR tasks. Information encoded in both distances among images and ranked lists computed by CBIR systems are used for analyzing contextual information.

Clustering approaches are also closely related to re-ranking methods that exploit contextual information in CBIR domain. redA re-ranking framework for CBIR systems [29] based on contextual dissimilarity measures uses a clustering approach. The contexts are modeled using a clustering algorithm to group similar images given their ranked lists. redA re-ranking algorithm that uses post-retrieval clustering [19] was proposed for color retrieval tasks. In the first step, images are retrieved using visual features such as color histogram. Next, the retrieved images are analyzed using hierarchical agglomerative clustering methods and the returned ranked lists is adjusted according to the distance of a cluster to a query. redThe Distance Optimization Algorithm (DOA) [22] considers an iterative clustering approach based on distances correlation and on the similarity of ranked lists. The algorithm explores the fact that if two images are similar, their distances to other images and therefore their ranked lists should be similar as well.

This paper presents the *RL-Sim Re-Ranking Algorithm*, a new post-processing method that considers the similarity between ranked lists for encoding contextual information in CBIR systems. We believe that the modeling of contextual information considering only the similarity between ranked lists represents an advantage of our strategy. Since the re-ranking method does not depend on distances or similarity scores, it can be used for different CBIR tasks and can be easily adapted for other information retrieval tasks (*e.g.*, text or multimodal retrieval). Beyond that, the re-ranking method can use different similarity/distance measures among ranked lists, a well-established research area [7, 39, 41].

3. Image Retrieval Model

Let $\mathcal{C}$=$\{img_1, img_2, \ldots, img_N\}$ be an *image collection*.

Let $\mathcal{D}$ be an *image descriptor* which can be defined [35] as a tuple (ϵ, ρ), where:

- ϵ: $\hat{I} \rightarrow \mathbb{R}^n$ *is a function, which extracts a feature vector* $v_{\hat{I}}$ *from an image* $\hat{I}$.
- ρ: $\mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ *is a distance function that computes the distance between two images according to the distance between their corresponding feature vectors.*

In order to obtain the distance between two images img_i and img_j, it is necessary to compute the value of $\rho(\epsilon(img_i),\epsilon(img_j))$. For simplicity and readability purposes we use the notation $\rho(img_i,img_j)$ along the paper.

The distance $\rho(img_i,img_j)$ among all images $img_i,img_j \in \mathcal{C}$ can be computed to obtain an $N \times N$ distance matrix A, such that $A[i,j] = \rho(img_i,img_j)$.

Given an image query img_q, we can compute a ranked list R_q in response to the query, based on distance matrix A. The ranked list R_q=$\{img_1$, img_2, $\ldots$, $img_N\}$ can be defined as a permutation of the collection $\mathcal{C}$. A permutation σ_q is as a bijection from the collection $\mathcal{C}$ onto the set $[N] = \{1, 2, \ldots, N\}$ where N is the size $\mid \mathcal{C} \mid$ of collection $\mathcal{C}$. For a permutation σ_q, we interpret $\sigma_q(i)$ as the position (or rank) of image img_i in the ranked list R_q. Therefore, we can say that, if img_x is ranked before img_y, that is $\sigma_q(x) < \sigma_q(y)$, then $\rho(img_q,img_x) \leq \rho(img_q,img_y)$.

We also can take every image $img_i \in \mathcal{C}$ as an image query img_q, in order to obtain a set $\mathcal{R} = \{R_1, R_2, \ldots, R_N\}$ of ranked lists for each image of collection $\mathcal{C}$.

red Our goal is to propose a re-ranking algorithm (represented by function f_r) which uses a contextual distance measure based on ranked lists. The re-ranking algorithm takes as input the distance matrix A and the set of ranked lists $\mathcal{R}$ for computing a new and more effective distance matrix $\hat{A}$:

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

Given the new distance matrix $\hat{A}$, a new set $\hat{\mathcal{R}}$ can be obtained. $\hat{\mathcal{R}}$ contains the new ranking positions of all collection images, that is, the collection images are re-ranked. Note that the main aspect of f_r consists in exploiting all relationships encoded in both A and $\mathcal{R}$.

4. Contextual Distance Measure based on Ranked Lists

In this section, we define a contextual distance measure based on similarity/dissimilarity of ranked lists. The contextual distance measure represents the basis of our proposed re-ranking algorithm. According to the formalization presented in the previous section, a given image descriptor $\mathcal{D}$ can compute a distance $\rho(img_i,img_j)$ between two images img_i, $img_j \in \mathcal{C}$. Therefore, this distance value considers only the two images img_i, img_j.

In order to compute the ranked lists R_i, R_j for images img_i, img_j, distances from these images to all other collection images need to be computed. In this way, the ranked lists represent, by itself, a contextual description of images with regard to the whole dataset. The images at top positions of ranked lists often represent the most relevant images, in the sense that they usually represent the results in which users are interested. In this scenario, we conjecture that *given any two images and their respective ranked lists, a new and more effective distance measure between the two images can be computed by considering the images at top positions of their ranked lists.*

The proposed contextual distance measure is based on this conjecture. In the commom case, the top positions of ranked lists contain many images that are similar to the query image and some "*wrong*" (non-similar) images. Those images placed at top positions usually are similar to each other and, therefore, there are many images in common in their ranked lists. We can observe that this set of images (similar to the query image and similar to each other) appears in the ranked lists of all images that compose the set. The same behavior can not be observed when analysing the top positions of the ranked lists of non-similar images (the same set of images does not appear at top positions). In this scenario, a low contextual distance score is produced, since there are few images in common at top positions of ranked lists of non-similar images. The objective of the proposed re-ranking algorithm is to move the non-similar images down in the ranked lists, and as a result of this process, the effectiveness of ranked lists is improved. We should note that, in extreme situations, in which the CBIR descriptors completely confuse similar and non-similar images, there is no contextual information available for improving the ranked lists.

red A new contextual distance measure, defined in the following, is iteratively learned in a unsupervised setting which is able to incorporate the contextual information, improving retrieval results. Our approach does not require any user intervention, but can be combined with other techniques that take into account the user's preferences, such as relevance feedback approaches.

Let us consider the *neighborhood set* $\mathcal{N}(i)$ of an image img_i, which contains images similar to img_i according to a given distance, say ρ defined by the image descriptor. The set $\mathcal{N}(i)$ can be obtained, for example, by the well-known *k-Nearest Neighbor* approach, where the cardinality of the set is denoted by $\mid \mathcal{N}(i) \mid = k$. In Section 5.1, we formally define approaches for obtaining the neighborhood set $\mathcal{N}$.

In the following, we formally define the top positions of a ranked list as a *top k list*, according to [7]. We define a ranked list R_i as a permutation of collection $\mathcal{C}$, given by a bijection σ_i from the collection $\mathcal{C}$ onto the set $[N] = \{1, 2, \ldots, N\}$. Similarlly, a *top k list* τ_i is a bijection from a domain $\mathcal{N}(i)$ (the members of the top k list) to $[k] = \{1, 2, \ldots, k\}$. We say that img_j appears in the top k list τ_i if $img_j \in \mathcal{N}(i)$. We interpret $\tau_i(j)$ as the position (or rank) of image img_j in τ_i. For the well-known *k-Nearest Neighbor* approach, we can say that if img_1 is ranked before img_2 ($\tau_i(1) < \tau_i(2)$), then $\rho(img_i,img_1) \leq \rho(img_i,img_2)$.

Approaches for computing top k lists are formally defined in Section 5.1.

Assume that τ_i and τ_j are top k lists computed for images img_i, img_j respectively. Several similarity (or dissimilarity) measures for comparing τ_i and τ_j can be defined [7, 39, 41] (different distance measures are discussed in Section 5.2). Let $d(\tau_i, \tau_j, k)$ denote a given distance measure for comparing top k lists, we define a non-iterative contextual distance measure $\rho_c(img_i, img_j)$ based on comparison of the top k lists, as follows:

$$\rho_c(img_i, img_j) = d(\tau_i, \tau_j, k) \tag{2}$$

Based on the conjecture that the contextual distance measure ρ_c represents a more effective distance between images, we can recompute the distance among all images in a collection based on this measure. In this way, a new set of ranked lists (and their respective top k lists) can be obtained, such that the contextual distance can also be recomputed. Therefore, this process can be repeated in an iterative manner. Let $^{(t)}$ be a superscript that denotes the iteration. Let $\tau_i^{(t)}$ be the top k list for image img_i at iteration t, which is computed based on contextual distance $\rho_c^{(t)}$. 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 $img_i, img_j \in \mathcal{C}$. We can define an iterative contextual measure as follows:

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

Once the effectiveness of the contextual distance measure improves along iterations, the effectiveness of ranked lists also improve. Non-relevant images are moved out from the first positions of the ranked lists and, therefore, k can be increased for considering more images. In this way, a larger k can be considered for computation of top k lists along iterations, as follows:

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

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

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

Finally, a new distance matrix $\hat{A}$ can be computed based on $\hat{\rho}$, such that for all images $img_i, img_j \in \mathcal{C}$ we have $\hat{A}_{ij} = \hat{\rho}(img_i, img_j)$. Based on $\hat{A}$, a new set of ranked lists $\hat{\mathcal{R}}$ can be computed completing the re-ranking process.

5. Comparing Ranked Lists

The comparison between ranked lists is the basis of our proposed contextual measure. This section discusses and formalizes the process of comparison, which can be divided in two main steps: *(i)* how to retrieve a *neighborhood set* for a given image img_i, which is used to compose a top k list τ_i; *(ii)* how to compute a distance $d(\tau_i, \tau_j, k)$ between two top k lists τ_i and τ_j.

Sections 5.1 and 5.2 discuss respectively steps *(i)* and *(ii)*.

5.1. Neighborhood Set

This section presents and formally defines two different approaches for computing the top k lists for a given image: the well-known k-Nearest Neighbors (kNN) method and the Mutual k-Nearest Neighbors ($MkNN$).

5.1.1. k-Nearest Neighborhs

Let img_i be a collection image $img_i \in \mathcal{C}$ whose the k most similar images (neighborhood set) we want to select. Let $\mathcal{N}_{kNN}(i)$ be the neighborhood set obtained using the k-nearest neighbors method, which is defined as follows:

$$\mathcal{N}_{kNN}(i,k) = \{\mathcal{R} \subseteq \mathcal{C}, |\mathcal{R}| = k \wedge \forall x \in \mathcal{R}, y \in \mathcal{C} - \mathcal{R} : \rho(i,x) \leqslant \rho(i,y)\} \tag{6}$$

Based on the neighborhood set $\mathcal{N}_{kNN}(i)$ we want also define the top k list $\tau_{i_{kNN}}$ using the k-Nearest Neighbors. Let $\tau_{i_{kNN}}(j)$ be the position (or rank) of image img_j in $\tau_{i_{kNN}}$, we can say that if img_x is ranked before img_y, that is $\tau_{i_{kNN}}(x) < \tau_{i_{kNN}}(y)$, then $\rho(img_i,img_x) \leq \rho(img_i,img_y)$.

More formally, let's consider that there is no equal distances from img_i to images in a neighborhood set $\mathcal{N}_{kNN}(i,k)$ (or there is a pre-processing step for tie breaking distances), such that $\{\forall x, y \in \mathcal{N}_{kNN}(i,k) : \rho(i,x) = \rho(i,y)\} = \emptyset$. The top k list $\tau_{i_{kNN}}(j)$ is a permutation of $\mathcal{C}$, that can also be considered as a bijection $\tau_{i_{kNN}} : \mathcal{C} \to [1, \ldots, k]$ defined as follows:

$$\tau_{i_{kNN}}(j) = |\{j \in \mathcal{N}_{kNN}(i,k), \forall x \in \mathcal{N}_{kNN}(i,k) : \rho(i,x) < \rho(i,j)\}| + 1 \tag{7}$$

5.1.2. Mutual k-Nearest Neighborhs

Let $\tau_{i_{kNN}}(j)$ be the position (or rank) of image img_j in the top k list $\tau_{i_{kNN}}$. Let $\tau_{j_{kNN}}(i)$ be the position of img_i in the top k list $\tau_{j_{kNN}}$, it is very common in CBIR systems that $\tau_{i_{kNN}}(j) \neq \tau_{j_{kNN}}(i)$. However, when the difference between these positions is large, it may indicate an incorrect position of the image in one of the top k lists.

Based on this observation, we define a *Mutual k-Nearest Neighbors* method that considers reciprocal positions of images in their ranked lists. In fact, we select the k-Nearest Neighbors considering $c \times k$ neighbors (where c is a constant[1]). Given a neighborhood set $\mathcal{N}_{kNN}(i, c \times k)$, we select the k most similar of this set by taking into account both: *(i)* the position of images in ranked list of img_i as *(ii)* the position of img_i in the ranked list of these images. We

[1]We used $c = 2$ in our experiments.

formally define the neighborhood set based on the Mutual k-Nearest Neighbor as follows:

$$\mathcal{N}_{MkNN}(i,k) = \{\mathcal{R} \subseteq \mathcal{N}_{kNN}(i, c \times k), |\mathcal{R}| = k \wedge \forall x \in \mathcal{R}, y \in \mathcal{C} - \mathcal{R} : \tau_{i_{kNN}}(x) + \tau_{x_{kNN}}(i) \leqslant \tau_{i_{kNN}}(y) + \tau_{y_{kNN}}(i)\} \tag{8}$$

We also define the top k list $\tau_{i_{MkNN}}$ using the Mutual k-Nearest Neighbors:

$$\tau_{i_{MkNN}}(j) = |\{j \in \mathcal{N}_{MkNN}(i,k), \forall x \in \mathcal{N}_{MkNN}(i,k) : \tau_{i_{kNN}}(x) + \tau_{x_{kNN}}(i) \leqslant \tau_{i_{kNN}}(y) + \tau_{y_{kNN}}(i)\}| + 1 \tag{9}$$

5.2. Distance Measures between Top k Lists

Given the methods for obtaining neighborhood sets and top k lists, we now discuss how to compute a distance $d(\tau_i, \tau_j, k)$ between two retrieved top k lists τ_i, τ_j. Note that the distance measure adopted $d(\cdot, \cdot, k)$ does not depend on the method used for computing the neighborhood set $\mathcal{N}$ and top k lists τ. Therefore, different combinations can be used in our re-ranking algorithm.

5.2.1. Intersection Measure

An approach to define the distance between two top k lists τ_i and τ_i proposed in [7] is to capture the extent of overlap between τ_i and τ_i. This idea of overlap can be extended by considering not only the overlap at depth k, but also the cumulative overlap at increasing depths [39, 25]. For each $k_c \in \{1 \dots k\}$, it is computed the overlap at k_c, and then those overlaps are averaged to derive a similarity measure. The measure gives higher weights to the first positions of top k lists, which are considered many times. Equation 10 formally defines the intersection similarity measure ψ.

$$\psi(\tau_i, \tau_j, k) = \frac{\sum_{k_c=1}^{k} \mid \mathcal{N}(i, k_c) \cap \mathcal{N}(j, k_c) \mid}{k} \tag{10}$$

Note that if two ranked lists present the same images at the first positions, the size of the intersection set is greater, and the value of ψ is greater as well. Figure 1 illustrates the computation of ψ considering multiscale values of k.

Since we are interested in a distance measure between top k lists, we define d_ψ as follows:

$$d_\psi(\tau_i, \tau_j, k) = \frac{1}{1 + \psi(\tau_i, \tau_j, k)} \tag{11}$$

5.2.2. Kendal's Tau Measure

The Kendall's tau is a 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 [7]. The normalized Kendall's tau measure is defined as follows:

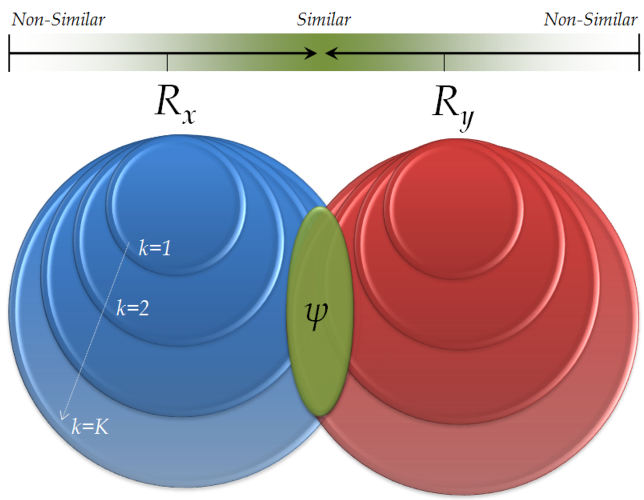

Figure 1: Computation of measure ψ: intersection of ranked lists with different sizes.

$$d_\tau(\tau_i, \tau_j, k) = \frac{\sum_{x,y \in \mathcal{N}(i,k) \cup \mathcal{N}(j,k)} \bar{K}_{x,y}(\tau_i, \tau_j)}{k \times (k-1)}, \tag{12}$$

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 R_i and R_j. This function can be defined as follows:

$$\bar{K}_{x,y}(\tau_i, \tau_j) = \begin{cases} 0 & \text{if } (\sigma_i(x) \leqslant \sigma_i(y) \wedge \sigma_j(x) \leqslant \sigma_j(y)) \;\vee\; (\sigma_i(x) \geqslant \sigma_i(y) \wedge \sigma_j(x) \geqslant \sigma_j(y)) \\ 1 & \text{otherwise} \end{cases}$$

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

Note that, although our goal is to compute the distance between the top k lists τ_i and τ_j, we considered the ranked lists positions σ_i and σ_j because we may have an image that is in only one of the top k lists (for example, $img_x \in \tau_i$ and $img_x \notin \tau_j$).

6. The RL-Sim Re-Ranking Algorithm

red In this section, we present an algorithmic view of the proposed re-ranking approach. The goal of our re-ranking algorithm is to exploit the initial set of ranked lists $\mathcal{R} = \{R_1, R_2, \ldots, R_N\}$ for computing a more effective distance matrix $\hat{A}$ and, therefore, a more effective set of ranked lists $\hat{\mathcal{R}}$. The *RL-Sim Re-Ranking Algorithm* is based on the presented contextual measure ρ_c, which takes into account the similarity between ranked lists on an iterative way.

An iterative approach is proposed. Let the superscript $^{(t)}$ denotes the current iteration, a new (and more effective) set of ranked lists $\mathcal{R}^{(t+1)}$ is computed by taking into account distances among top k 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 several iterations aiming to improve the effectiveness incrementally. After a number T of iterations a re-ranking is performed based on the final distance matrix $\hat{A}$. Based on matrix $\hat{A}$, a final set of ranked lists $\hat{\mathcal{R}}$ can be computed. Algorithm 1 outlines the proposed *RL-Sim Re-Ranking Algorithm*.

Algorithm 1 RL-Sim Re-Ranking Algorithm

Require: Original set of ranked lists $\mathcal{R}$ and parameters k_s, T, λ
Ensure: Processed set of ranked lists $\hat{\mathcal{R}}$

1: $t \leftarrow 0$
2: $\mathcal{R}^{(t)} \leftarrow \mathcal{R}$
3: $A^{(t)} \leftarrow A$
4: $k \leftarrow k_s$
5: **while** $t < T$ **do**
6:   **for all** $R_i \in \mathcal{R}^{(t)}$ **do**
7:     $counter \leftarrow 0$
8:     **for all** $img_j \in R_i$ **do**
9:       **if** $counter \leq \lambda$ **then**
10:         $A^{(t+1)}[i,j] \leftarrow d(\tau_i, \tau_j, k)$
11:       **else**
12:         $A^{(t+1)}[i,j] \leftarrow 1 + A^{(t)}[i,j]$
13:       **end if**
14:       $counter \leftarrow counter + 1$
15:     **end for**
16:   **end for**
17:   $\mathcal{R}^{(t+1)} \leftarrow perfomReRanking(A^{(t+1)})$
18:   $k \leftarrow k + 1$
19:   $t \leftarrow t + 1$
20: **end while**
21: $\hat{\mathcal{R}} \leftarrow \mathcal{R}^{(T)}$

Observe that the distances are redefined considering the function $d(\tau_i, \tau_j, k)$ for the first λ positions of the each ranked list, such that $\lambda \in \mathbb{N}$ and $0 \leq \lambda \leq N$. For images in the remaining positions of the ranked lists, the new distance is redefined (Line 12) based on the current distances. In these cases, the function $d(\tau_i, \tau_j, k)$ does not need to be computed, considering that relevant images should be at the begining of the ranked lists. In this way, the computational efforts decrease, making this step of the algorithm not dependent on the collection size N.

In Line 18, at each iteration t, we increment the number of k neighbors considered. The motivation behind this increment relies on the fact that the effectiveness of ranked lists increase along iterations. In this way, non-relevant images are moved out from the first positions of the ranked lists and k can be increased for considering more images. red The increment value of 1 was chosen because it represents the smallest possible increment, since the greater the increment value, the greater the risk of considering non-relevant images in the ranked list.

Note that the re-ranking algorithm does not depend on specific measures between top k lists. In this way, an important advantage of our re-ranking algorithm is the possibility of using different approaches for retrieving the neighboorhood set (we discussed the *kNN* and *Mutual kNN* methods) and different measures for comparing top k lists (we discussed the *intersection* and *Kendall's tau* measures). Therefore, the proposed RL-Sim Re-Ranking algorithm can be easily extended in order to consider different and even more complex approaches to compute the similarity between top k lists.

7. Rank Aggregation

Let $\mathcal{C}$ be an image collection and let $\mathcal{D}_s = \{D_1, D_2, \ldots, D_m\}$ be a set of CBIR descriptors. We can use the set of descriptors $\mathcal{D}$ for computing a set of distances matrices $\mathcal{A}_s = \{A_1, A_2, \ldots, A_m\}$. Our approach for combining descriptors works as follows: first, we combine the set $\mathcal{A}$ in a unique matrix A_c. For the matrices combination we use a multiplicative approach. Each position (i, j) of the matrix is computed as follows:

$$A_c[i,j] = (1 + A_1[i,j]) \times (1 + A_2[i,j]) \times \ldots (1 + A_m[i,j]) \tag{13}$$

Once we have a matrix A_c, we can compute a set of ranked lists $\mathcal{R}_c$ based on this matrix. Then, we can submit the matrix A_c and the set $\mathcal{R}_c$ for our original re-ranking algorithm.

8. Experimental Evaluation

This section presents the set of conducted experiments for demonstrating the effectiveness of our method. We analyzed and compared our method under several aspects. Section 8.1 presents an analysis of the re-ranking algorithm considering the impact of parameters. Section 8.2 presents a brief discussion about complexity and efficiency.

Section 8.3 discusses the experimental results for our re-ranking method. Section 8.3.1 presents results of the use of our method for several shape descriptors, considering the well-known MPEG-7 database [15]. Sections 8.3.2 and 8.3.3 aim to validate the hypothesis that our method can be applied to other image retrieval tasks. In addition to shape descriptors, we conduct experiments with color and texture descriptors.

Section 8.4 presents experimental results of our method on rank aggregation tasks. Finally, we also conduct experiments aiming to compare our results to state-of-the-art related post-processing and rank aggregation methods in Section 8.5.

All experiments were conducted considering all images in the collections as query images. Results presented in the paper (MAP and Recall@40 scores) represent an average result.

8.1. Experiment 1: Impact of Parameters

The execution of Algorithm 1 considers three parameters: *(i)* k_s - number of neighbors considered when algorithm starts; *(ii)* λ - number of images of each ranked list that are considered for redefining distances; and *(iii)* T - number of iterations along which the algorithm is executed.

To evaluate the influence of different parameter settings on the retrieval scores and for determining the best parameters values, we conducted a set of experiments considering the MPEG-7 [15] database. The MPEG-7 [15] database is a well-known shape database, composed by 1400 shapes divided in 70 classes. The size of images range from (50×48) to (526×408) pixels. For evaluation, the so-called bullseye score was considered, which 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. For distance computation, we used the ASC [17] shape descriptor.

Retrieval scores are computed considering the kNN method for the intersection measure. Parameter k_s varies in the interval [1,20] while parameter T varies in the interval [1,7]. Figure 2 illustrates the results of precision scores for different values of parameters k_s and T. We observed that best retrieval scores increased along iterations yielding the best precision score (94.69%) for $k_s = 15$ and $T = 3$. We used these values in all experiments involving the intersection measure (for *kNN* and *Mutual kNN*). Analogous experiments were conducted for the Kendall's tau measure and very similar values were obtained: $k_s = 15$ and $T = 2$. Those values were also used in all experiments involving the Kendall's tau measure.

We also analyzed the impact of parameter λ on precision. As discussed before, the objective of λ consists in decreasing computation efforts needed for the algorithm. It can be seen as a tradeoff between effectiveness and efficiency. In this way, we ranged λ in the interval [0,N] (considering the MPEG-7 collection). Results are illustrated in Figure 3. Note that the precision scores achieve the stability near to $\lambda = 700$ (value used in our experiments).

8.2. Experiment 2: Aspects of Efficiency

This paper has as its focus on the presentation of *RL-Sim Re-Ranking Algorithm* and on its effectiveness evaluation. The focus on effectiveness is justified by the fact that the execution of the algorithm is expected to be off-line, as in other post-processing methods [38]. This subsection aims to briefly discuss some aspects of efficiency and computational complexity.

The complexity of computation of comparison between top k lists, considering both intersection and Kendall's tau measures, is $O(k^2)$. The number of comparisons that should be processed is equal to ($N \times \lambda$). Since the parameters λ and k have fixed values independent on N, the asymptotic computional complexity of the main step of the algorithm (computing distance between top k lists) is $O(N)$. Note, however, that the parameter λ is given in the interval $0 \leqslant \lambda \leqslant N$ and, if defined with the maximum value $\lambda = N$, that makes this step of the algorithm $O(N^2)$.

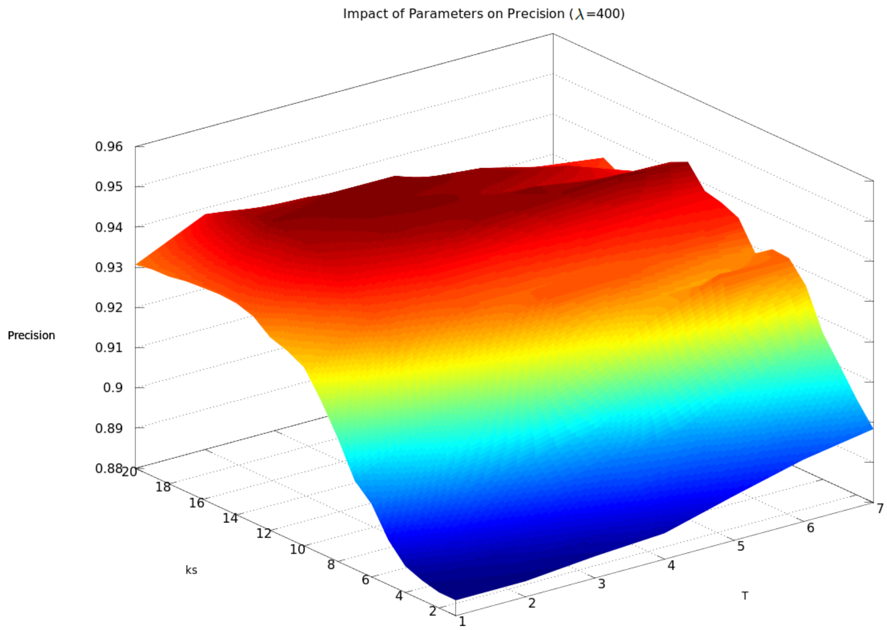

Figure 2: Impact of parameters: k_s and T.

Other steps of the algorithm have different complexities. The matrix A are recomputed ($O(N^2)$) at each iteration. The re-ranking step computes a sort operation ($O(NlogN)$) for all images ($O(N^2 logN)$). However, these steps admit optimizations: the matrix does not require to be totally recomputed and the ranked lists do not require to be totally sorted again at each iteration. The *RL-Sim Re-Ranking Algorithm* can also be massively parallelized, since there is no dependence between comparisons between ranked lists in a same iteration. Optimizations and parallelization issues will be investigated in future work. Note also that other post-processing methods use matrices multiplication approaches [42, 43] and graph algorithms [38], both with complexity of $O(N^3)$.

We evaluated the computation time of RL-Sim Re-Ranking algorithm for MPEG-7 dataset ($N = 1400$), using the parameters defined in Section 8.1 ($k_s = 15$, $T = 3$, and $\lambda = 700$), executing in a Linux Laptop Core i3 and using a C implementation. This execution took approximately 7.6 s.

8.3. Experiment 3: Re-Ranking Evaluation

In this section, we present a set of conducted experiments for demonstrating the effectiveness of our method. We analyzed our method in the task of re-ranking images considering shape, color, and texture descriptors.

8.3.1. Shape Descriptors

We evaluate the use of our method with five shape descriptors: Segment Saliences (SS) [36], Beam Angle Statistics (BAS) [2], Inner Distance Shape Con-

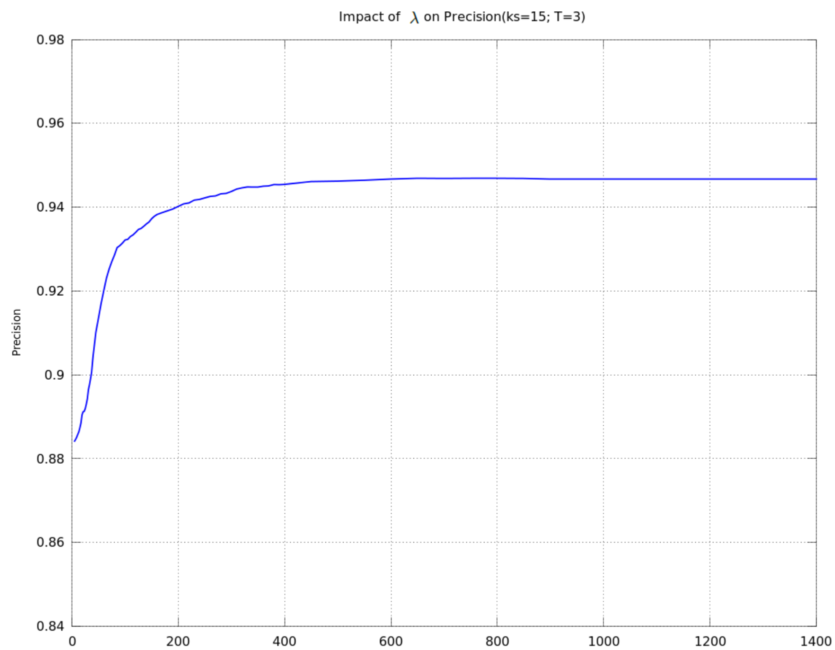

Figure 3: Impact of parameter λ on precision.

text (IDSC) [16], Contour Features Descriptor (CFD) [21], and Aspect Shape Context (ASC) [17]. We consider the MPEG-7 [15] database, described in Section 8.1.

Table 1 presents results (bullseye score - Recall@40) for shape descriptors using the *intersection measure* on MPEG-7 [15] database. Both *kNN* and *Mutual kNN* methods are considered in the experiments. We can observe significative gains from +6.62% to +30.90% in relation to the initial descriptor's results.

The iterative behavior of the *RL-Sim Re-Ranking* algorithm can be observed in results illustrated in Figure 4. The figure shows the evolution of rankings (and their precision) along iterations. The first row presents 20 results for a query image (first column with green border) according to the CFD [21] shape descriptor. We can observe that wrong results contain images from different classes, situation where the re-ranking algorithm can *correct* the rankings based on the contextual information. The remaining rows present the results for each iteration of *RL-Sim Re-Ranking* algorithm, considering the *Mutual kNN* and *intersection measure* approaches. We can observe the significant improvement in terms of precision, ranging from 45% (on the ranking computed by the CFD [21] descriptor) to 100% at the third iteration of the re-ranking algorithm.

Table 2 presents results for shape descriptors using the *Kendall's tau measure* on MPEG-7 database. Results of both *kNN* and *Mutual kNN* methods are presented. We can also observe significative gains ranging from +5.98% to +27.44%.

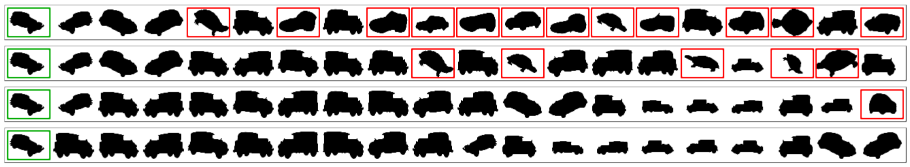

Figure 4: Evolution of rankings according to iterations on MPEG-7 [15] database. The first column (green border) contains the query image. The first row presents the results of CFD [21] shape descriptor (wrong results with red borders). The remaining rows present the results of *RL-Sim Re-Ranking* algorithm for each iteration.

Table 1: **RL-Sim Re-Ranking using Intersection Distance Measure** for shape descriptors on MPEG-7 dataset *(Recall@40)*.

Shape Descriptor	**Score**	**kNN + Intersection**	**Gain**	**Mutual kNN + Intersection**	**Gain**
SS [36]	43.99%	53.15%	+20.82%	57.58%	+30.90%
BAS [2]	75.20%	82.94%	+10.29%	85.87%	+14.19%
IDSC [16]	85.40%	92.18%	+7.94%	92.62%	+8.45%
CFD [21]	84.43%	94.13%	+11.49%	95.33%	+12.91%
ASC [17]	88.39%	94.69%	+7.13%	95.75%	+8.33%
AIR [9]	93.67%	99.90%	+6.65%	99.87%	+6.62%

Results for MAP (*Mean Average Precision*) score are presented in Table 3, considering the insersection measure; and Table 4, considering the Kendall's tau measure. Both tables present results considering the *kNN* and *Mutual kNN* approaches. We can observe positive gains for all shape descriptors in all combinations of approaches, ranging from +2.42% to +26.63%.

In addition to shape descriptors, we conducted experiments with color and texture descriptors, considering 12 image descriptors in 3 different datasets. Experiments with color and texture descriptors are described in next Sections.

8.3.2. Color Descriptors

We evaluate our method with three color descriptors: Border/Interior Pixel Classification (BIC) [30], Auto Color Correlograms (ACC) [11], and Global Color Histogram (GCH) [31]. The experiments were conducted on a database used in [40] and composed by images from 7 soccer teams, containing 40 images

Table 2: **RL-Sim Re-Ranking using Kendall's Tau Distance Measure** for shape descriptors on MPEG-7 dataset *(Recall@40)*.

Shape Descriptor	**Score**	**kNN + Kendall's Tau**	**Gain**	**Mutual kNN + Kendall's Tau**	**Gain**
SS [36]	43.99%	52.67%	+19.73%	56.06%	+27.44%
BAS [2]	75.20%	81.16%	+7.93%	83.44%	+10.96%
IDSC [16]	85.40%	91.12%	+6.70%	92.06%	+7.80%
CFD [21]	84.43%	93.12%	+10.29%	94.27%	+11.65%
ASC [17]	88.39%	93.68%	+5.98%	94.56%	+6.98%
AIR [9]	93.67%	99.94%	+6.69%	99.93%	+6.68%

Table 3: MAP scores for **RL-Sim Re-Ranking using Intersection Distance Measure** in different CBIR tasks.

Descriptor	**Type**	**Dataset**	**Score** (MAP)	**kNN + Intersection**	**Gain**	**M-kNN + Intersection**	**Gain**
SS [36]	Shape	MPEG-7	37.67%	43.06%	+14.31%	47.70%	+26.63%
BAS [2]	Shape	MPEG-7	71.52%	74.57%	+4.25%	78.16%	+9.28%
IDSC [16]	Shape	MPEG-7	81.70%	86.75%	+6.18%	87.67%	+7.31%
CFD [21]	Shape	MPEG-7	80.71%	88.97%	+10.23%	90.78%	+12.48%
ASC [17]	Shape	MPEG-7	85.28%	88.81%	+4.14%	90.88%	+6.57%
AIR [9]	Shape	MPEG-7	89.39%	93.54%	+4.64%	93.52%	+4.62%
GCH [31]	Color	Soccer	32.24%	33.66%	+4.40%	33.84%	+4.96%
ACC [11]	Color	Soccer	37.23%	43.54%	+16.95%	44.78%	+20.28%
BIC [30]	Color	Soccer	39.26%	43.45%	+10.67%	44.08%	+12.28%
LBP [18]	Texture	Brodatz	48.40%	47.77%	-1.30%	48.51%	+0.23%
CCOM [14]	Texture	Brodatz	57.57%	62.01%	+7.72%	63.48%	+10.27%
LAS [33]	Texture	Brodatz	75.15%	77.81%	+3.54%	78.11%	+3.94%

Table 4: MAP scores for **RL-Sim Re-Ranking using Kendall's Tau Distance Measure** in different CBIR tasks.

Descriptor	**Type**	**Dataset**	**Score** (MAP)	**kNN + Kendall's Tau**	**Gain**	**M-kNN + Kendall's Tau**	**Gain**
SS [36]	Shape	MPEG-7	37.67%	44.24%	+17.92%	46.74%	+24.08%
BAS [2]	Shape	MPEG-7	71.52%	73.25%	+2.42%	75.38%	+5.40%
IDSC [16]	Shape	MPEG-7	81.70%	85.93%	+8.18%	86.53%	+5.91%
CFD [21]	Shape	MPEG-7	80.71%	88.40%	+9.53%	89.50%	+9.55%
ASC [17]	Shape	MPEG-7	85.28%	88.10%	+3.31%	89.92%	+5.44%
AIR [9]	Shape	MPEG-7	89.39%	96.27%	+7.70%	95.72%	+7.08%
GCH [31]	Color	Soccer	32.24%	32.96%	+2.18%	33.76%	+4.71%
ACC [11]	Color	Soccer	37.23%	44.29%	+18.96%	46.02%	+23.61%
BIC [30]	Color	Soccer	39.26%	43.76%	+11.46%	45.58%	+16.35%
LBP [18]	Texture	Brodatz	48.40%	45.20%	-6.61%	45.78%	-5.41%
CCOM [14]	Texture	Brodatz	57.57%	60.30%	+4.74%	61.41%	+6.67%
LAS [33]	Texture	Brodatz	75.15%	75.62%	+0.63%	76.13%	+1.30%

per class. The size of images range from (198×148) to (537×672) pixels.

Table 3 presents the experimental results considering the insersection measure while Table 4 considers the Kendall's tau measure. Both tables present results considering the *kNN* and *Mutual kNN* approaches. We can observe a positive gain for all color descriptors for approaches ranging from +2.18% to +23.18% (considering MAP as score).

8.3.3. Texture Descriptors

The experiments consider three well-known texture descriptors: Local Binary Patterns (LBP) [18], Color Co-Occurrence Matrix (CCOM) [14], and Local Activity Spectrum (LAS) [33]. We used the Brodatz [5] dataset, a popular dataset for texture descriptors evaluation was considered. The Brodatz dataset is composed of 111 different textures of size (512×512) pixels. Each texture is divided into 16 blocks (128×128) pixels of non-overlapping sub images, such that 1776 images are considered.

Table 3 presents the experimental results considering the insersection measure while Table 4 considers the Kendall's tau measure. Both tables present

Table 5: MAP scores regarding the use of RL-Sim Algorithm for Rank Aggregation

Descriptor	**Type**	**Dataset**	**Neighbor Set**	**Measure**	**Score** (MAP)
CFD [21]	Shape	MPEG-7	-	-	80.71%
ASC [17]	Shape	MPEG-7	-	-	85.28%
CFD [21] + ASC [17]	Shape	MPEG-7	kNN	Intersection	98.75%
CFD [21] + ASC [17]	Shape	MPEG-7	M-kNN	Intersection	98.96%
CFD [21] + ASC [17]	Shape	MPEG-7	kNN	Kendall's Tau	98.57%
CFD [21] + ASC [17]	Shape	MPEG-7	M-kNN	Kendall's Tau	98.57%
ACC [11]	Color	Soccer	-	-	37.23%
BIC [30]	Color	Soccer	-	-	39.26%
BIC [30] + ACC [11]	Color	Soccer	kNN	Intersection	44.49%
BIC [30] + ACC [11]	Color	Soccer	M-kNN	Intersection	44.16%
BIC [30] + ACC [11]	Color	Soccer	kNN	Kendall's Tau	44.45%
BIC [30] + ACC [11]	Color	Soccer	M-kNN	Kendall's Tau	45.16%
CCOM [14]	Texture	Brodatz	-	-	57.57%
LAS [33]	Texture	Brodatz	-	-	75.15%
LAS [33] + CCOM [14]	Texture	Brodatz	kNN	Intersection	80.26%
LAS [33] + CCOM [14]	Texture	Brodatz	M-kNN	Intersection	83.39%
LAS [33] + CCOM [14]	Texture	Brodatz	kNN	Kendall's Tau	80.51%
LAS [33] + CCOM [14]	Texture	Brodatz	M-kNN	Kendall's Tau	81.68%

Table 6: Post-processing methods comparison on MPEG-7 database *(Recall@40)*.

Algorithm	**Shape Descriptor**	**Score**	**Gain**
Shape Descriptors			
Data Driven Generative Models (DDGM) [37]	-	80.03%	-
Contour Features Descritpor (CFD) [21]	-	84.43%	-
Inner Distance Shape Context (IDSC) [16]	-	85.40%	-
Shape Context (SC) [4]	-	86.80%	-
Aspect Shape Context (ASC) [17]	-	88.39%	-
Articulation-Invariant Representation (AIR) [9]	-	93.67%	-
Post-Processing Methods			
Graph Transduction (LP) [42]	IDSC [16]	91.00%	+6.56%
Distance Optimization Algorithm [22]	CFD [21]	92.56%	+9.63%
Locally Constrained Diffusion Process [43]	IDSC [16]	93.32%	+9.27%
Shortest Path Propagation [38]	IDSC [16]	93.35%	+9.31%
Mutual kNN Graph [13]	IDSC [16]	93.40%	+9.37%
Contextual Re-Ranking [20]	CFD [21]	94.55%	+11.99%
RL-Sim Re-Ranking [kNN+Intersection]	**ASC [17]**	**94.69%**	**+7.13%**
RL-Sim Re-Ranking [M-kNN+Intersection]	**CFD [21]**	**95.33%**	**+12.91%**
RL-Sim Re-Ranking [M-kNN+Intersection]	**ASC [17]**	**95.75%**	**+8.33%**
Locally Constrained Diffusion Process [43]	ASC [17]	95.96%	+8.56%
RL-Sim Re-Ranking [M-kNN+Kendall's tau]	**AIR [9]**	**99.93%**	**+6.68%**
RL-Sim Re-Ranking [kNN+Kendall's tau]	**AIR [9]**	**99.94%**	**+6.69%**
Tensor Product Graph [44]	AIR [9]	99.99%	+6.75%
Rank Aggregation Methods			
Borda [45]	CFD [21]+IDSC [16]	91.06%	-
Borda [45]	CFD [21]+ASC [17]	93.51%	-
Reciprocal Rank Fusion [6]	CFD [21]+IDSC [16]	94.98%	-
Reciprocal Rank Fusion [6]	CFD [21]+ASC [17]	96.25%	-
Co-Transduction [3]	IDSC [16]+DDGM [37]	97.31%	-
Co-Transduction [3]	SC [4]+DDGM [37]	97.45%	-
Co-Transduction [3]	SC [4]+IDSC [16]	97.72%	-
RL-Sim Re-Ranking [kNN+Intersection]	**CFD [21]+IDSC [16]**	**99.31%**	-
RL-Sim Re-Ranking [kNN+Intersection]	**CFD [21]+ASC [17]**	**99.44%**	-
RL-Sim Re-Ranking [M-kNN+Intersection]	**CFD [21]+IDSC [16]**	**99.49%**	-
RL-Sim Re-Ranking [M-kNN+Intersection]	**CFD [21]+ASC [17]**	**99.65%**	-

results considering the *kNN* and *Mutual kNN* approaches. We can observe that our re-ranking methods presents positive gains ranging from +0.63% to +10.27%, except for LBP [18], which presents loss of effectiveness in some cases. The LBP [18] descriptor on Brodatz dataset represents the extreme situations, discussed in Section 4, in which the CBIR descriptor confuses classes of images. In these situations, there is no enough contextual information available in the ranked lists to distinguish the classes, causing the loss of effectiveness. This situation is contrary to that illustrated in Figure 4, in which wrong results contain few images from different classes.

8.4. Experiment 4: Rank Aggregation Evaluation

We evaluate the use of our re-ranking method to combine different CBIR descriptors. We select two descriptors for each visual property. Table 5 presents results of MAP score of these descriptors. We can observe that significatve gains are obtained when compared with the results of descriptors in isolation.

8.5. Experiment 5: Comparison to Other Approaches

Finally, we also evaluate our method in comparison with other state-of-the-art post-processing methods. We use the MPEG-7 [15] dataset, with the called bullseye score (*Recall@40*), commonly used for post-processing methods evaluation and comparison. Table 6 presents results of our *RL-Sim Re-Rannking* algorithm (considering selected variations of shape descriptors, *kNN*, *Mutual kNN*, intersection and Kendall's tau measure approaches) in comparison with several other post-processing and rank aggregation methods recently proposed in the literature. The gains in relation to initial descriptors results are also presented.

Note that the results of our **RL-Sim Re-Ranking** method presents better effectiveness performace when compared to various methods. In re-ranking tasks, the best combination (*M-kNN* + Kendall's tau) results reached 99.94%. In rank aggregation tasks comparison, we consider as baselines the traditional Borda [45] method, the recently proposed Reciprocal Rank Fusion [6] method, and the recently proposed Co-Transduction [3] method, proposed for CBIR applications. We can observe that the best combination of *RL-Sim Re-Ranking* algorithm (*M-kNN* + Intersection measure) in rank aggregation tasks reached 99.65%.

9. Conclusions

In this work, we have presented a new re-ranking method that exploits contextual information for improving CBIR tasks. The main idea consists in analyzing similarity between ranked lists for redefing distance among images. We conducted a large set of experiments and experimental results demonstrated the applicability of our method to several image retrieval tasks based on shape, color, and texture descriptors.

Future work focuses on: (i) considering other different measures between top k lists; (ii) combining results obtained from different measures; (iii) optimizing the proposed re-ranking algorithm by considering parallel architectures; red(vi) combining our re-ranking algorithm with other supervised methods, such as relevance feedback approaches.

10. Acknowledgments

Authors thank AMD, CAPES, FAPESP, FAEPEX, and CNPq for financial support. Authors also thank DGA/UNICAMP for its support in this work.

References

[1] Abowd, G.D., Dey, A.K., Brown, P.J., Davies, N., Smith, M., Steggles, P.. Towards a better understanding of context and context-awareness. In: 1st international symposium on Handheld and Ubiquitous Computing. 1999. p. 304–307.

[2] Arica, N., Vural, F.T.Y.. BAS: a perceptual shape descriptor based on the beam angle statistics. Pattern Recognition Letters 2003;24(9-10):1627–1639.

[3] Bai, X., Wang, B., Wang, X., Liu, W., Tu, Z.. Co-transduction for shape retrieval. In: European Conference on Computer Vision (ECCV'2010). volume 3; 2010. p. 328–341.

[4] Belongie, S., Malik, J., Puzicha, J.. Shape matching and object recognition using shape contexts. IEEE Transactions on Pattern Analysis and Machine Intelligence 2002;24(4):509–522.

[5] Brodatz, P.. Textures: A Photographic Album for Artists and Designers. Dover, 1966.

[6] Cormack, G.V., Clarke, C.L.A., Buettcher, S.. Reciprocal rank fusion outperforms condorcet and individual rank learning methods. In: ACM SIGIR Conference on Research and Development in Information Retrieval. 2009. p. 758–759.

[7] Fagin, R., Kumar, R., Sivakumar, D.. Comparing top k lists. In: ACM-SIAM Symposium on Discrete algorithms (SODA'03). 2003. p. 28–36.

[8] Faria, F.F., Veloso, A., Almeida, H.M., Valle, E., da S. Torres, R., Goncalves, M.A., Jr., W.M.. Learning to rank for content-based image retrieval. In: Multimedia Information Retrieval (MIR'2010). 2010. p. 285–294.

[9] Gopalan, R., Turaga, P., Chellappa, R.. Articulation-invariant representation of non-planar shapes. In: 11th European Conference on Computer Vision (ECCV'2010). volume 3; 2010. p. 286–299.

[10] Hoi, S.C., Liu, W., Chang, S.F.. Semi-supervised distance metric learning for collaborative image retrieval and clustering. ACM Transactions on Multimedia Computing and Communication Applications 2010;6(3):18:1–18:26.

[11] Huang, J., Kumar, S.R., Mitra, M., Zhu, W.J., Zabih, R.. Image indexing using color correlograms. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR'97). 1997. p. 762–768.

[12] Jegou, H., Schmid, C., Harzallah, H., Verbeek, J.. Accurate image search using the contextual dissimilarity measure. IEEE Transactions on Pattern Analysis and Machine Intelligence 2010;32(1):2–11.

[13] Kontschieder, P., Donoser, M., Bischof, H.. Beyond pairwise shape similarity analysis. In: Asian Conference on Computer Vision. 2009. p. 655–666.

[14] Kovalev, V., Volmer, S.. Color co-occurence descriptors for querying-by-example. In: International Conference on Multimedia Modeling. 1998. p. 32.

[15] Latecki, L.J., Lakmper, R., Eckhardt, U.. Shape descriptors for non-rigid shapes with a single closed contour. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR'2000). 2000. p. 424–429.

[16] Ling, H., Jacobs, D.W.. Shape classification using the inner-distance. IEEE Transactions on Pattern Analysis and Machine Intelligence 2007;29(2):286–299.

[17] Ling, H., Yang, X., Latecki, L.J.. Balancing deformability and discriminability for shape matching. In: European Conference on Computer Vision (ECCV'2010). volume 3; 2010. p. 411–424.

[18] Ojala, T., Pietikäinen, M., Mäenpää, T.. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence 2002;24(7):971–987.

[19] Park, G., Baek, Y., Lee, H.K.. Re-ranking algorithm using post-retrieval clustering for content-based image retrieval. Information Processing and Management 2005;41(2):177–194.

[20] Pedronette, D.C.G., da S. Torres, R.. Exploiting contextual information for image re-ranking. In: Iberoamerican Congress on Pattern Recognition (CIARP'2010). 2010. p. 541–548.

[21] Pedronette, D.C.G., da S. Torres, R.. Shape retrieval using contour features and distance optmization. In: International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISAPP'2010). volume 1; 2010. p. 197 – 202.

[22] Pedronette, D.C.G., da S. Torres, R.. Exploiting clustering approaches for image re-ranking. Journal of Visual Languages and Computing 2011;22(6):453–466.

[23] Pedronette, D.C.G., da S. Torres, R.. Exploiting contextual information for rank aggregation. In: International Conference on Image Processing (ICIP'2011). 2011. p. 97–100.

[24] Pedronette, D.C.G., da S. Torres, R.. Exploiting contextual spaces for image re-ranking and rank aggregation. In: ACM International Conference on Multimedia Retrieval (ICMR'11). 2011. p. 13:1–13:8.

[25] Pedronette, D.C.G., da S. Torres, R.. Image re-ranking and rank aggregation based on similarity of ranked lists. In: Computer Analysis of Images and Patterns (CAIP'2011). volume 6854; 2011. p. 369–376.

[26] Perronnin, F., Liu, Y., Renders, J.M.. A family of contextual measures of similarity between distributions with application to image retrieval. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR'2009). 2009. p. 2358–2365.

[27] Rijsbergen, C.J.V.. Information Retrieval. London: Butterworth-Heinemann, 1979.

[28] Roweis, S.T., Saul, L.K.. Nonlinear dimensionality reduction by locally linear embedding. Science 2000;290(5500):2323–2326.

[29] Schwander, O., Nielsen, F.. Reranking with contextual dissimilarity measures from representational bregmanl k-means. In: International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISAPP'2010). volume 1; 2010. p. 118–122.

[30] Stehling, R.O., Nascimento, M.A., Falcão, A.X.. A compact and efficient image retrieval approach based on border/interior pixel classification. In: ACM Conference on Information and Knowledge Management (CIKM'2002). 2002. p. 102–109.

[31] Swain, M.J., Ballard, D.H.. Color indexing. International Journal on Computer Vision 1991;7(1):11–32.

[32] Tan, H.K., Ngo, C.W.. Fusing heterogeneous modalities for video and image re-ranking. In: Proceedings of the 1st ACM International Conference on Multimedia Retrieval. ICMR '11; 2011. .

[33] Tao, B., Dickinson, B.W.. Texture recognition and image retrieval using gradient indexing. Journal of Visual Comunication and Image Representation 2000;11(3):327–342.

[34] Tenenbaum, J.B., Silva, V.d., Langford, J.C.. A global geometric framework for nonlinear dimensionality reduction. Science 2000;290(5500):2319–2323.

[35] Torres, R.d.S., Falcão, A.X.. Content-Based Image Retrieval: Theory and Applications. Revista de Informática Teórica e Aplicada 2006;13(2):161–185.

[36] Torres, R.d.S., Falcão, A.X.. Contour Salience Descriptors for Effective Image Retrieval and Analysis. Image and Vision Computing 2007;25(1):3–13.

[37] Tu, Z., Yuille, A.L.. Shape matching and recognition - using generative models and informative features. In: European Conference on Computer Vision (ECCV'2004). 2004. p. 195–209.

[38] Wang, J., Li, Y., Bai, X., Zhang, Y., Wang, C., Tang, N.. Learning context-sensitive similarity by shortest path propagation. Pattern Recognition 2011;44(10-11):2367–2374.

[39] Webber, W., Moffat, A., Zobel, J.. A similarity measure for indefinite rankings. ACM Transactions on Information Systems 2010;28(4):20:1–20:38.

[40] Weijer, J.V.D., Schmid, C.. Coloring local feature extraction. In: European Conference on Computer Vision (ECCV'2006). volume Part II; 2006. p. 334–348.

[41] Wu, S., Crestani, F.. Methods for ranking information retrieval systems without relevance judgments. In: ACM Symposium on Applied Computing (SAC'03). 2003. p. 811–816.

[42] Yang, X., Bai, X., Latecki, L.J., Tu, Z.. Improving shape retrieval by learning graph transduction. In: European Conference on Computer Vision (ECCV'2008). volume 4; 2008. p. 788–801.

[43] Yang, X., Koknar-Tezel, S., Latecki, L.J.. Locally constrained diffusion process on locally densified distance spaces with applications to shape retrieval. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR'2009). 2009. p. 357–364.

[44] Yang, X., Latecki, L.J.. Affinity learning on a tensor product graph with applications to shape and image retrieval. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR'2011). 2011. p. 2369–2376.

[45] Young, H.P.. An axiomatization of borda's rule. Journal of Economic Theory 1974;9(1):43–52.

[46] Zhou, W., Tian, Q., Lu, Y., Yang, L., Li, H.. Latent visual context learning for web image applications. Pattern Recognition 2011;44(10-11):2263–2273.

[47] Zhu, X.. Semi-supervised learning with graphs. Ph.D. thesis; Pittsburgh, PA, USA; 2005. John Chair-Lafferty and Ronald Chair-Rosenfeld.