Projection-based Image Retrieval using Class-Specific Metrics

Paulo Joia, Erick Gomez Nieto, Glenda Botelho, João Batista Neto, Afonso Paiva, Luis Gustavo Nonato
ICMC/SME
Universidade de São Paulo
São Carlos, SP, Brazil
Email: {pjoia, egomezn, glenda, jbatista, apneto, gnonato}@icmc.usp.br

***Abstract*—Content-based image classification/retrieval based on image descriptors has become an essential component in most database systems. However, most existing systems do not provide mechanisms that enable interactive multi-objective queries, hampering the user experience. In this paper we present a novel methodology capable of accomplishing multi-objective searches while still being interactive. Our approach relies on a combination of class-specific metrics and multidimensional projection to devise an effective and interactive image retrieval system. Besides allowing visual exploration of image data sets, the provided results and comparisons show that the proposed approach outperforms existing methods, turning out to be a very attractive alternative for managing image data sets.**

***Keywords*-Multidimensional projection, Content-Based Image Retrieval.**

I. Introduction

Content-based image retrieval (CBIR) is a basic tool in any computational system aimed at supporting the cataloging and querying of images from large databases such as art collections, photograph archives, and medical diagnosis. Many different mechanisms have been proposed to accomplish the CBIR process, which vary as to the metric used to compare images as well as the way the search results are presented (see [6] for a comprehensive survey). Among the different approaches, multidimensional projection-based CBIR [8] has emerged as one of the most promising methodology, since it allows for simultaneously querying multiple images while visualizing the results in a two-dimensional manner that enables user interaction.

Despite the advances and good results, existing multidimensional projection (MP) methods have not yet exploited all its potential towards incorporating well-known pattern recognition tools into the CBIR process. For example, MP techniques are very flexible in terms of the metric used to measure the similarity between instances of data, thus, the metric could be adjusted according to the class of image one is searching for. In particular, the Least Square Projection (LSP) method [22], which has turned out to be a quite efficient MP method for CBIR [8], does not make use of any mechanism to tune similarity metrics so as to improve search results while querying multi-class image databases.

In this work we build on the flexibility provided by LSP and propose a new MP method, called *Class-Specific Multidimensional Projection* (CSMP), which preserves the good properties of LSP while still being more accurate. In contrast to other MP methods, CSMP uses a family of metrics rather than a single one to accomplish the projection. More specifically, CSMP combines pattern recognition tools such as feature extraction and feature selection to assign a distinct metric to each query image. Therefore, when the user performs a search with multiple query images, a particular metric is tailored from each query image, enabling a more accurate mechanism to compare images according to the class the images belong to. As we show in Section IV (results and comparison), the proposed class-specific metric construction increases accuracy considerably, outperforming other MP methods, including the original LSP method.

Contributions We can summarize the main contributions of this work as:

- The design of class-specific metrics to measure the similarity between images (Section III).
- A new multidimensional projection method (CSMP) that relies on the class-specific metrics to accomplish content-based image queries.
- A system capable of performing multiple queries in databases with images from many different classes.

To the best of our knowledge, this is the first time a MP method makes use of a set of metrics rather than a single one to accomplish the mapping to the visual space.

II. Related work

As mentioned above, the technique presented in this paper makes use of multidimensional projection to perform content-based image retrieval. Therefore, we provide a brief overview of both fields in order to better contextualize our approach.

A. Content-Based Image Retrieval

Content-based Image Retrieval (CBIR) is a branch of Computer Science that encloses techniques and methods aimed at organizing large digital image repositories by means of visual content. As such, any technique ranging from a similarity function to a robust image annotation system are strictly related to CBIR [6].

Started in the 90's [27], research on CBIR has evolved to more specific topics such as relevance feedback [34], face

recognition [33] and medical images [20]. Another comprehensive review has been written by *Datta et al.* [6].

In general, CBIR techniques seek to tackle two intrinsic problems: a) How can an image be mathematically described? and b) how to best compute the similarity between two images given their mathematical representation. The answer to these two questions are addressed in the following CBIR-related topics: feature extraction, feature selection, definition of distance functions and metrics, indexing of features and even user interfaces.

The framework presented in this paper addresses the topics of feature selection, definition of the most appropriate metrics and user interface issues, introducing the use of projections and user intervention to provide accurate image retrieval.

B. Multidimensional Projection

Multidimensional projection methods map instances from a high dimensional space into the visual space ($\mathbb{R}^2$ in our context) so as to preserve distances as much as possible. MP methods can be grouped into two main categories, namely, global and local methods.

Global Methods map data instances into the visual space using a single transformation. Techniques based on spectral decomposition, also known as *classical scaling*, are typical examples of global methods. Classical scaling embeds instances from eigenvectors of a double-centered transformation applied to a dissimilarity matrix (symmetric matrix containing the dissimilarity between each pair of data instances) [31]. Although many authors have proposed alternatives towards getting around the high computational cost and the lack of flexibility as to user interaction [2], [3], [9], [25], [29], [12], classical scaling is still computationally costly and cumbersome.

The global method proposed by Kruskal [13] uses nonlinear-optimization to map data instances to the visual space. Since finding the minimum of the energy functional (commonly called stress function) is costly, Pekalska et al. [24] proposed a technique that first embeds a subset of samples in the visual space using Kruskal optimization scheme and then maps the remaining instances using a linear mapping. Although Pekalska's approach may be tuned to allow user interaction, it requires that a minimum number of sample is projected in the first step, which hampers the user experience. The recent linear mapping PLMP [23] uses an approach similar to Pekalska's method, but PLMP makes use of faster mechanisms in both stages of the projection so as to enable the projection of out-of-core and streaming data.

Least Squares Projection [22] (LSP) is a two-step global technique that uses a non-linear scheme to position a subset of samples (control points) into the visual space. Assuming that each instance of data can be represented as an average of its neighbors, LSP builds Laplacian-like systems and uses the previously projected samples to constraint the system. Control points can be manipulated by user in order to facilitate grouping visualization and exploration.

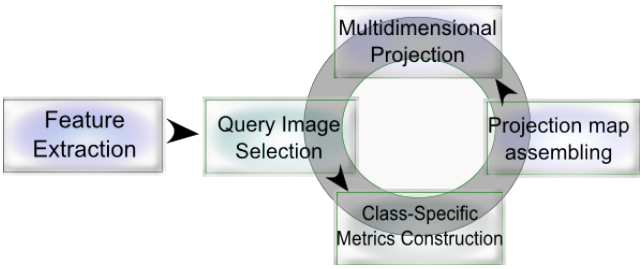


Fig. 1. Pipeline of our methodology.

Local Methods use, basically, two components to perform the multidimensional projection, namely the neighborhood information of each instance and the location of a subset of samples in the visual space. The mapping of an instance x is carried out by considering only the neighbors of x, which characterizes the local nature of the process. The well-known approach proposed by Chalmers [5] (and its hybrid variants [11], [19]) first maps a subset of the samples using a force-based scheme and then exploits neighborhood structure of those samples to embed the remaining data in the visual space.

The recent PLP method [21] uses a force-based scheme to place a subset of samples in the visual space. The remaining data instances are projected through a family of Laplacian-like operators, each one built using local information. PLP provides great flexibility to user interaction, however, the continuous manipulation of projected data demands costly structural updates, impacting robustness and interactivity.

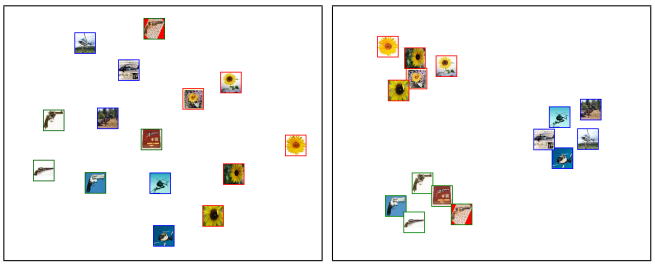
The technique proposed in this work can be seen as a mid term between global and local methods, since it uses a single transformation to project the instances onto the visual space while changing the metric so as to handle the data in a more localized manner. Therefore, our approach aims at holding the robustness of global methods while allowing for the flexible user interaction typically found in local methods.

III. CSMP Image Retrieval

This section presents the technicalities underling the proposed projection-based CBIR technique. Before going into the details, though, we provide an overview of the pipeline our method is built on.

A. Pipeline Overview

The proposed technique comprises five main steps (Fig. 1), namely, *feature extraction*, *selection of query images*, *class-specific metrics* construction, *projection map* assembling, and *multidimensional projection* itself. Having extracted the features, the next step, selection of query images, is carried out by the user and consists of selecting a set of images for which similar instances should be searched for (recall that our methodology allows multiple searches to be performed simultaneously). The user explicitly indicates the images belonging to distinct classes. From the class information the system automatically carries out a feature selection procedure so as to compute the subset of features that better represents



(a) (b)

Fig. 2. Multiple query images: (a) Before user manipulation; (b) After user intervention, similar instances are grouped together.

each class. The resulting subset of features is used to define the class-specific metrics which is used to assemble the matrix responsible to perform the multidimensional projection in the last step of the pipeline. Each step is fully detailed below.

B. Image Feature Extraction

Given a set of images $\mathscr{I}$, the very first step of our methodology corresponds to embed $\mathscr{I}$ in the feature space. In more mathematical terms, we have to specify a transformation $\Lambda : \mathscr{I} \to \mathbb{R}^k$, where k is the number of features used to represent each image.

The transformation Λ is defined from feature extraction mechanism widely employed by the computer vision and pattern recognition communities. More specifically, we start by applying a sequence of feature extractors in each image $\alpha \in \mathscr{I}$, concatenating the results in an array. The sequence of extractors we employ convey mainly texture (wavelet [15], Gabor [1], Tamura [7] and first order moment [30]) and color information (color moments [18]).

The extractors above results in an array with 220 entries, most of them not relevant to discriminate the images. In order to clean up the feature array, we apply the well-known SSFS subset evaluator [10] tool which filters out irrelevant entries, producing a much smaller set of features, with only 15% of the original amount of features, which defines the dimension of the feature space.

This feature extraction step is, indeed, run in a pre-processing step, since it has to be done only once.

C. Class-Specific Metrics

Suppose that $\mathscr{I}$ has already been embedded in the feature space $\mathbb{R}^k$. The CBIR process requires that a set of query images is specified in order to drive the search process. In our context, the set of query images, denoted by Q, are provided by the user. The images in Q can be either picked out from the database $\mathscr{I}$ or provided as a new instance. In the latter case, the query images has to go through the feature extraction procedure described above.

One strength of our methodology is that visual resources are available to the user during the process. Therefore, after selecting the query images the user can specify the class they belong to by interactively grouping similar query images in the visual space, as illustrated in Fig. 2. The system interprets each group of query images as a class to be searched. In other words, the user is implicitly labeling the query images according to the groups formed.

Accurate methods are known to select the features that best characterize each instance in a classified (labeled) data set. Therefore, the labeled query images can be submitted to a feature selection procedure in order to identify the subset of features that best represents each group of query images. We are using the Logistic Model Tree [17] as feature selector, since this method produced better results in most of the tests we carried out.

Let Q_i be a subset of the query images that shares the same label (grouped together by the user) and $I_{Q_i} = \{i_1, \ldots, i_r\}$ be the indexes of the features that best represent the images in Q_i (determined by the feature selector). Given an image $\alpha \in \mathscr{I}$, $\beta \in Q_i$, and $\gamma \in Q_j$, $i \neq j$ it is reasonable to expect that

$$d_{Q_i}(\alpha,\beta) \leq d_{Q_j}(\alpha,\gamma)$$

if α belongs to the same class as β, where the *class-specific* metric d_{Q_i} is defined as

$$d_{Q_i}(\alpha,\beta) = \sum_{j \in I_{Q_i}} (\alpha^j - \beta^j)^2 \qquad (1)$$

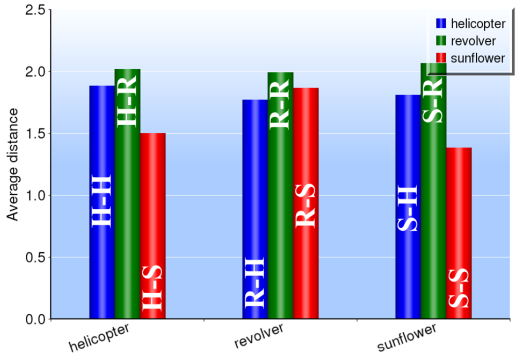
where α^j accounts for the j^{th} coordinate (feature) of α (resp. β). The metric (1) is indeed a pseudo-symmetric, since $d_{Q_i}(x,y) = 0$ does not imply that $x = y$, moreover, the triangle inequality may not be satisfied as well.

The rationale behind the class-specific distances defined in (1) is that if α is an image similar to $\beta \in Q_i$ then I_{Q_i} should be the best features to represent α, thus $d_{Q_i}(\alpha,\beta)$ should be small. However, one should expect a larger dissimilarity if a class-specific metric is used to measure the distance between α and a non-similar query image γ. Therefore, using class-specific metrics one avoids to compare features that do not represent the images properly, improving accuracy while increasing confidence in the value of the measured distance.

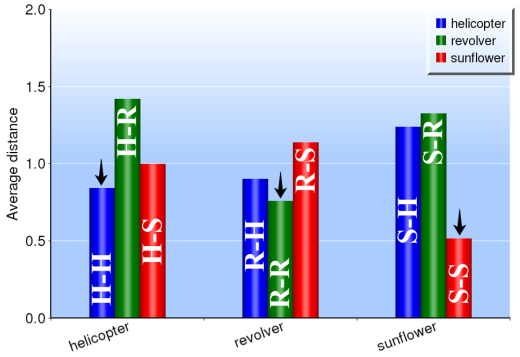
The bar graphs in Fig. 3 support our claim. Fig. 3(a) shows the average Euclidean distances from instances labeled as "Helicopter", "Revolver", and "Sunflower". Notice that, on average, the distance between similar instances such as helicopter to helicopter (left most blue bar, H-H) and revolver to revolver (middle green bar, R-R) are not much smaller than the distance between helicopter and revolver (left most green bar, H-R), for instance. Therefore, it would be difficult to distinguish elements from these classes. In contrast, the class-specific distance measure makes helicopters closer to their counterpart, the same for revolvers and sunflowers, as depicted in Fig. 3(b). It can be seen that the average distance computed from elements belonging to the same class is normally smaller, i.e. for one helicopter and other helicopters (left most blue bar, H-H), one revolver and other revolvers (middle green bar, R-R), one sunflower and other sunflowers (right most red bar, S-S).

D. The Class-Specific Multidimensional Projection

The proposed Class-Specific Multidimensional Projection method builds on the LSP technique [22] to perform visual



(a) Average Euclidean distance



(b) Average class-specific distance

Fig. 3. Dataset containing 3 distinct classes of images. (a) Distances between classes measured using Euclidean metric. (b) Distances measured using class-specific metric. Notice that average distance between elements belonging to the same class (H-H, R-R, S-S) are smaller for class-specific distance (indicated by arrows).

CBIR. However, in contrast to LSP, our approach makes use of the class-specific metrics defined in (1) so as to increase accuracy. Moreover, we constraint the linear system responsible for mapping instances to the visual space using a penalty method rather than the least square approach used by LSP. The penalty method has several advantages if compared to the least square method, as detailed below.

The CSMP relies on the assumption that each instance α of a data set $\mathscr{I}$ can be written as a linear combination of its nearest neighbors in the visual space. In more mathematical terms, let $N_\alpha = \{\alpha_1, \ldots, \alpha_s\}$ be the set of s nearest neighbors of α, and denote by (α_i^x, α_i^y) the coordinates of each element $\alpha_i \in N_\alpha$ when mapped to the visual space$\mathbb{R}^2$. From the linear combination hypothesis, one can compute the two-dimensional coordinates of α as:

$$(\alpha^x, \alpha^y) = \sum_{\alpha_i \in N_\alpha} c_{i\alpha}(\alpha_i^x, \alpha_i^y) \tag{2}$$

where $c_{i\alpha} > 0$.

Each image in $\mathscr{I}$ gives rise to a vectorial equation as the one given in (2), which can be assembled into two homogeneous linear systems:

$$L\mathbf{x} = 0; \quad L\mathbf{y} = 0 \tag{3}$$

where $\mathbf{x}$ and $\mathbf{y}$ account for the coordinates of the mapped elements and L is the matrix derived from equation (2).

The sets N_α define a *Nearest Neighbors Graph (NNG)* of $\mathscr{I}$, that is, a graph connecting each element in $\mathscr{I}$ to its nearest neighbors. It can be shown that the rank of L is $n-q$, where n is the number of elements in $\mathscr{I}$ and q is the number of connected components making up the NNG [28]. Thereby, in order to ensure a single non-trivial solution for the linear systems (3), the NNG should have only one connected component, which can be ensured by adding new edges linking disconnected components of the NNG.

The coefficients $c_{i\alpha}$ are defined as follows:

$$c_{i\alpha} = \begin{cases} d_{Q_i}(\alpha, \alpha_i) & \text{if either } \alpha \text{ or } \alpha_i \text{ is a query image} \\ d(\alpha, \alpha_i) & \text{if } \alpha \text{ and } \alpha_i \text{ are not query images} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

where d is the Euclidean distance and d_{Q_i} is the class-specific metric defined in (1). In order to ensure symmetry for L, we are assuming the convention $d_{Q_i}(\alpha, \alpha_i) = 0$, if α and α_i are query images from different classes.

The lack of geometric information in (3) may lead to solutions that are difficult to interpret and analyze. Geometrical information can be incorporated by using the coordinates of the user selected query images as constraints for (3).

We are using the penalty method [32] to constraint (3), which can be stated as follows: let Q be the query images and $\mathbf{b}^x$ (resp. $\mathbf{b}^y$) be the vector with zero in all entries except in the entries b_i corresponding to a query image α_i, where the value $b_i = \alpha_i^x$ is settled. The penalty method transforms the problem $L\mathbf{x} = 0$ into

$$(L+P)f = Pb \tag{5}$$

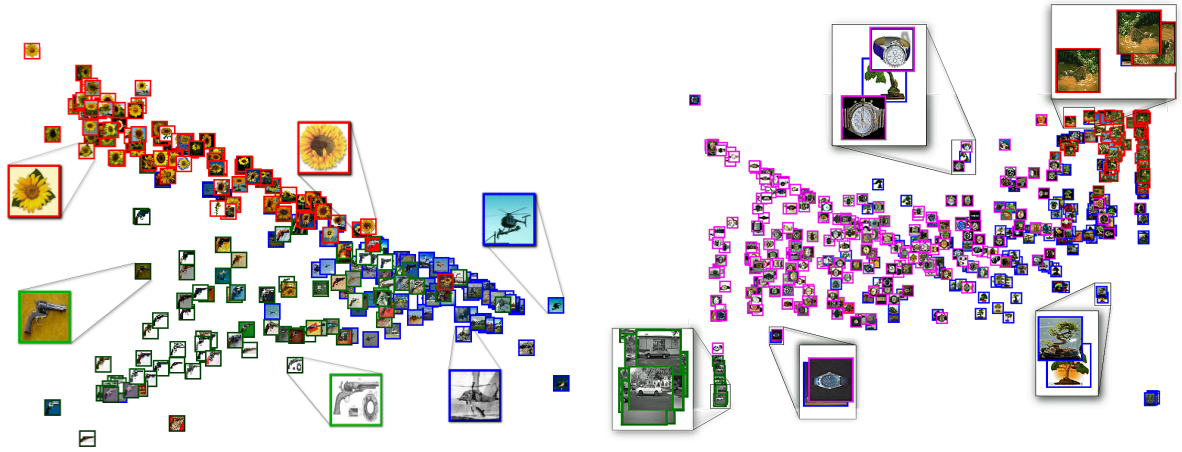
where P is the diagonal penalty matrix with non-zero diagonal elements $p_{ii} = p$ only in the positions corresponding to the query images and p a large value (10^8 in our implementation).

The penalty method holds several good properties. For instance, it preserves the symmetry and positive semi-definiteness of the system, thus allowing for Cholesky factorization. Moreover, adding a large positive value in some diagonal entries improves the conditioning number of the matrix, thus lessen numerical instabilities.

IV. Experimental Results and Comparisons

To evaluate the proposed approach, two experiments are provided. The first one aims to show how the projections produced by the CSMP compare with those computed by PLMP and LSP counterparts. Both qualitative and quantitative results are given. The second experiment describes the CSMP in a CBIR context.

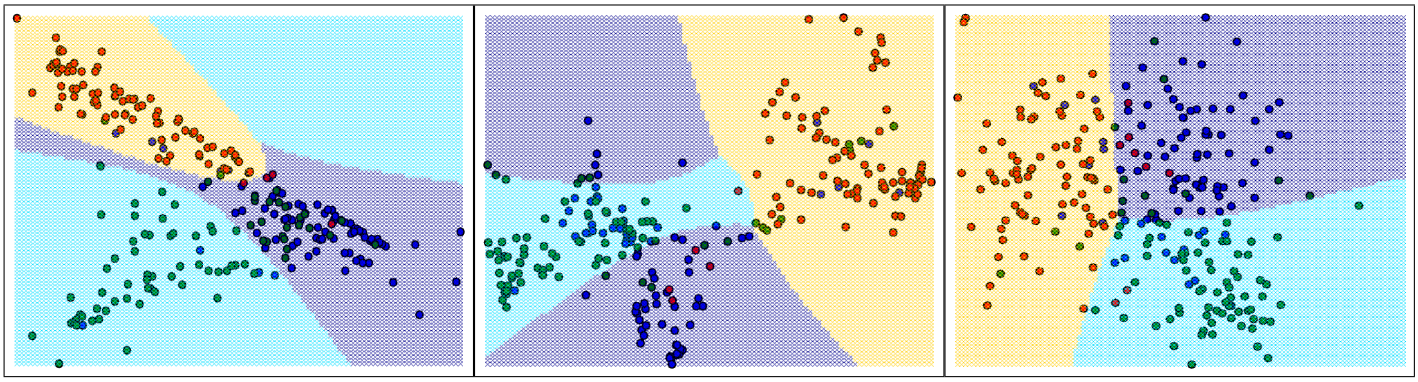
Two data sets, herein named dataset1 and dataset2, both created with 256x256 images from the Caltech 101 [16] image database, were used in the experiments. Dataset1 includes 255 samples from three distinct classes: Helicopter(88),



(a) 3-Class dataset: distinct color frames depict distinct classes.

(b) 4-Class dataset: distinct color frames depict distinct classes.

Fig. 4. Projection of the thumbnails obtained by the CSMP projections for the 3-class and 4-class data sets shown in Figs. 5(a) and 5(d), respectively.



(a) CSMP - Error 15.90%

(b) LSP - Error 23.32%

(c) PLMP - Error 17.16%

(d) CSMP - Error 12.63%

(e) LSP - Error 14.57%

(f) PLMP - Error 25.17%

Fig. 5. Comparing quality of projection. **Top row (3-class dataset)**: (a) CSMP (proposed method); (b) LSP and (c) PLMP. **Bottom row (4-class dataset)**: (d) CSMP (proposed method); (e) LSP and (f) PLMP.

Revolver(82) and Sunflower(85). Dataset2 is larger and includes 690 samples from four different classes: Bonsai(128), Car(123), Leopards(200) and Wrist Watches (239). Fig. 4 shows a projection of both data sets using the proposed CSMP.

We recall that to build the projection matrix, CSMP seeks for features that best represents each class in the dataset. These attributes change the metrics used for computing the dissimilarity among instances in the dataset. In general, traditional feature selection techniques rely on global information of all classes of the dataset to compute the most relevant features. Therefore, only one set of features is made available for each dataset. However we use a feature selection technique capable

of performing this task in class-specific mode, the Logistic Model Tree [17]. This technique produces the best features for each class, which is an essential characteristic for the success of our technique.

A. *Experiment 1 - Assessing the Quality of CSMP Projections*

In this experiment we compare the CSMP projections with LSP and PLMP. The top and bottom rows of Fig. 5 illustrate, respectively, the projections for dataset1 and dataset2. Projection on the left, center and right columns shows the projections computed with CSMP, SLP and PLMP, respectively.

To estimate the separability among classes, the visual space is divided into regions, according to a discriminant function (Discriminant Function Analysis). In this experiment we opted for a quadratic function so as to capture the largest possible number of elements of a given class within a region. This function determines which elements of a dataset may be considered as belonging to the same class by means of some statistical computation [14], [26]. Regions and projections can then be visualized as shown in Fig. 5. It can be seen that the CSMP produced the lowest rates for both dataset in comparison with LSP and PLMP projections. Moreover, the data becomes much better grouped when CSMP is used, an important aspect if data is not labeled, since cluster have to be identified visually.

The error rates are best described in Figs. 6(a) and 6(b), respectively, where the confusion matrices for the CSMP applied to the 3-Class and 4-Class data sets are depicted. Notice that CSMP is quite accurate for both data sets, producing an error smaller than the state-of-art methods.

Projections such as those of Fig. 5 can also be quantitatively evaluated by silhouette coefficient [4]. This coefficient measures both the cohesion and the separation between grouped instances. The cohesion a_x of x is calculated as the average of the distances between x and all other instances belonging to the same group as x. The separation b_x is the minimum distance between x and all other instances belonging to other groups. The silhouette of a projection is given by $Silh = \frac{1}{n}\sum_{x\in\mathcal{X}} \frac{(b_x - a_x)}{max(a_x, b_x)}$ where n is the number of instances. It is a sensitive coefficient in the interval $[-1, 1]$. The higher the values of silhouette for a given projection, the better the cohesion and the separation, that is, instances belonging to the same class are closer to each other, and yet, distinct classes are farther apart. Hence, higher values of silhouette indicate better projections. The bar graphs in Figs. 7(a) and 7(b) show the silhouette values computed from projections in Figs. 5(a) and 5(d), respectively. The CSMP silhouettes are higher than those for LSP and PLMP, in both data sets.

Classes	helicopter	revolver	sunflower
helicopter	77	4	7
revolver	23	57	2
sunflower	4	0	81

Error: 15.90%

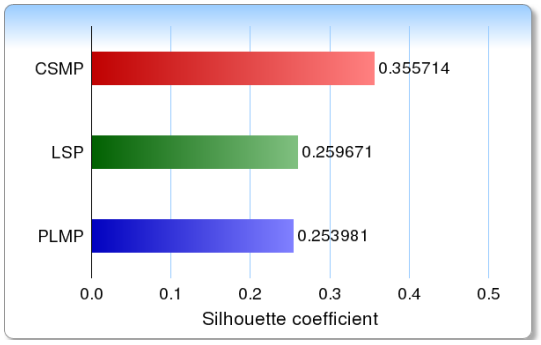
(a) CSMP Confusion Matrix 3-Class dataset

Classes	bonsai	car	leopards	watch
bonsai	93	0	8	27
car	0	123	0	0
leopards	7	0	193	0
watch	47	0	0	192

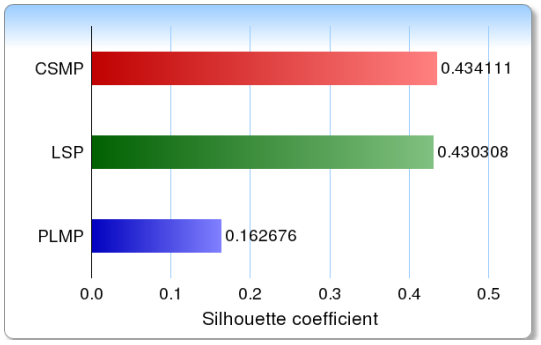
Error: 12.63%

(b) CSMP Confusion Matrix 4-Class dataset

Fig. 6. Classification by CSMP for 3-Class and 4-Class data sets.



(a) Silhouette for the 3-class dataset



(b) Silhouette for the 4-class dataset

Fig. 7. Silhouette of the projections shown in Fig. 5 for CSMP (red bar), LSP (green) and PLMP (blue). CSMP projection (proposed method) is better than the other two projections, for both data sets.

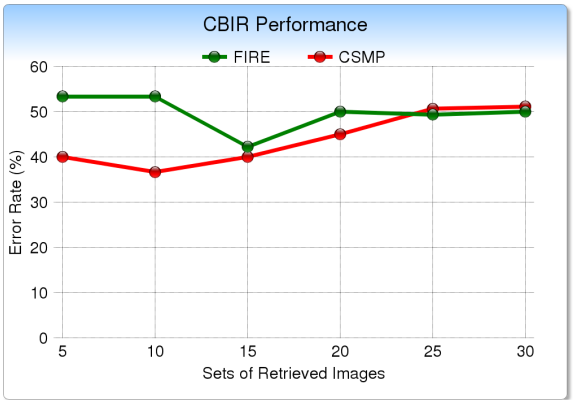
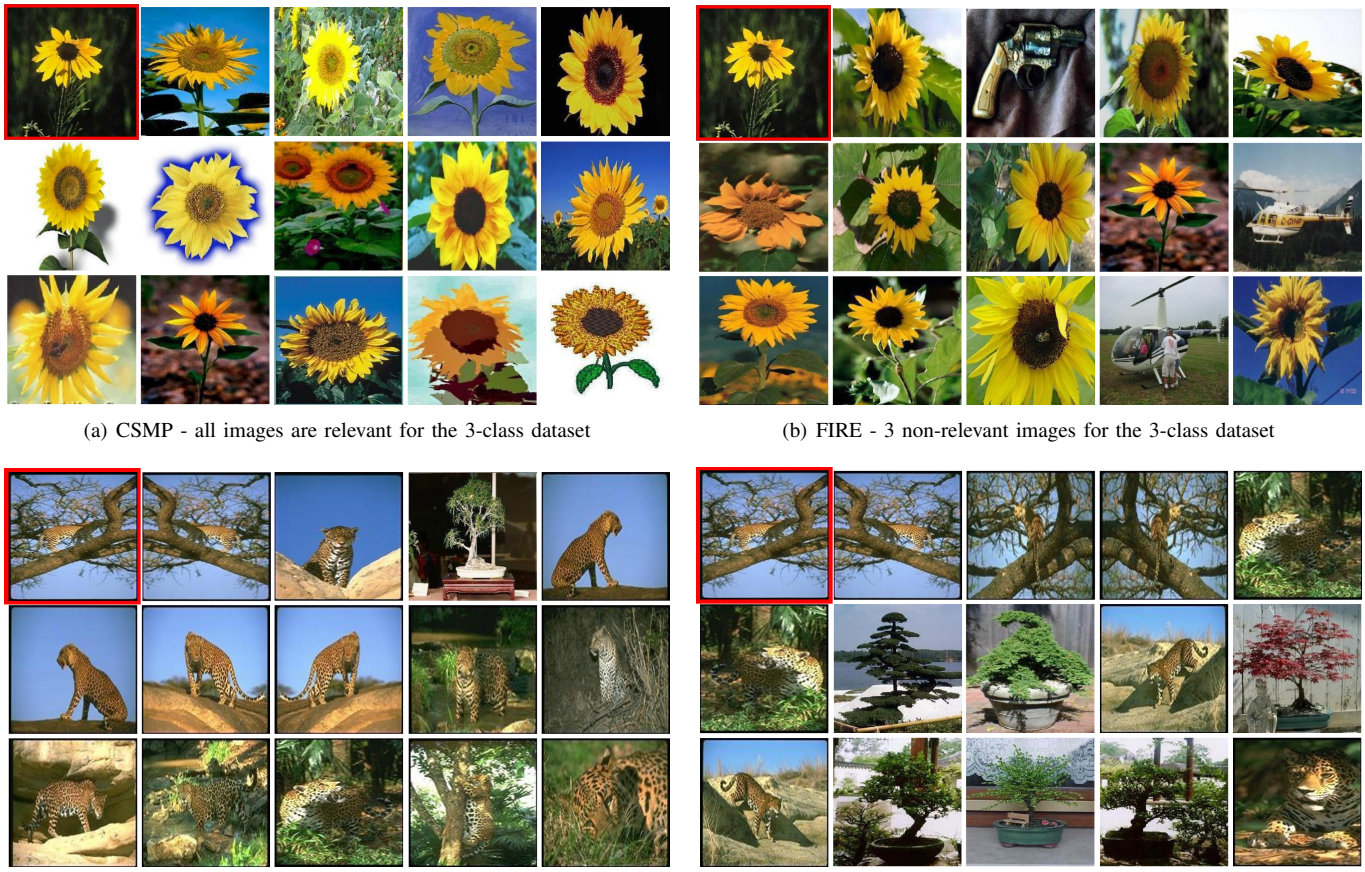


Fig. 8. CSMP in CBIR context for the 3-Class dataset.



(a) CSMP - all images are relevant for the 3-class dataset

(b) FIRE - 3 non-relevant images for the 3-class dataset

(c) CSMP - 1 non-relevant image for the 4-class dataset

(d) FIRE - 6 non-relevant images for the 4-class dataset

Fig. 9. Image retrieval for CSMP and FIRE with top 15 images. In red frame, the image used as query. **Top row (3-class dataset)**: all images in CSMP are relevant, whereas FIRE retrieved three non-relevant images. **Bottom row (4-class dataset)**: one image in CSMP is non-relevant, whereas in FIRE six images are non-relevant.

B. Experiment 2 - CSMP in a CBIR Context

In order to further attest the effectiveness of our methodology we compare CSMP against the public FIRE CBIR system [7] for dataset1. Fig. 8 shows the resulting error of the CBIR queries for both methods. The errors rates in the y-axis were computed as follows: for each set of retrieved images $(5, 10, \cdots, 30)$, we performed three queries, for three distinct images picked out randomly. We then computed the mean value for the three queries based on the relevant and non-relevant images retrieved. It can be seen that our method produces lower error rates up to 20 retrieved images. After that, both systems exhibit similar high error rates, which is expected due to a higher number of retrieved images from the database.

A visual comparison between CSMP and FIRE CBIR methods is presented in Figs. 9(a) to 9(d). The first thumbnail, in a red frame, is the query image used for both methods. The rank list of images is displayed in descending order by similarity. Notice that the CSMP outperforms the FIRE CBIR considerably for *dataset1* and *dataset2*. For the 3-class dataset shown in Fig. 9(a) all images displayed are relevant and belong to the class Sunflower, whereas three non-relevant images were retrieved by FIRE CBIR (Fig. 9(b)). For the 4-class dataset, CSMP retrieved one non-relevant image (Fig. 9(c)), whereas FIRE CBIR failed in six ocasions (Fig. 9(d)).

V. Discussion and Limitation

The comparisons presented in Section IV clearly show the effectiveness of the CSMP technique, surpassing, in requisites such as accuracy and flexibility, state-of-art methods. The superior performance of CSMP is a consequence of new class-specific distance measure it relies on, which ensures a more reliable distance definition among similar instances. Simplicity is another strength of CSMP, which essentially requires a linear solver library to be implemented.

The capability of performing multiple queries is another important property of CSMP that many applications can benefit from. This property allied to the possibility of interactively changing the position of control points in the visual space render CSMP a very attractive CBIR method.

In our experiments we notice that more "spread" layouts are produced when penalty factor p (set equal to 10^8 in all results presented in this paper) is decreased. The optimal value of p that produces the best trade-off between accuracy and pleasant layouts is an issue that deserves further investigation.

Choosing the ideal number of neighbors of each instance is another aspect we have to investigate more deeply. An alternative to the k-nearest neighbors scheme employed in our implementation would be to define a radius of influence to each point. However, finding the appropriate radius is not an easy task either.

VI. Conclusion and Future Work

In this work we proposed a new methodology for visual content-based image retrieval that relies on a class-specific distance measure to perform the search. The new metric we proposed turns out to be quite efficient in discriminating images that belong to the same class, rendering the proposed projection method superior to existing techniques. In fact, the evaluation we provided shows that our approach outperforms existing techniques in terms of accuracy and robustness while still being able to accomplishing multiple queries simultaneously. We are currently investigating how to adapt the proposed methodology to scenarios other than image retrieval, for instance music and video dataset.

Acknowledgment

We would like to thank the anonymous reviewers for their constructive comments. This research has been funded by FAPESP-Brazil and CAPES-Brazil.

References

[1] S. Arivazhagan, L. Ganesan, and S. Padam Priyal. Texture classification using gabor wavelets based rotation invariant features. *Pattern Recogn. Lett.*, 27:1976–1982, December 2006.

[2] Mikhail Belkin and Partha Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation*, 15(6):1373–1396, 2003.

[3] Ulrik Brandes and Christian Pich. Eigensolver methods for progressive multidimensional scaling of large data. In Michael Kaufmann and Dorothea Wagner, editors, *Graph Drawing*, volume 4372 of *Lecture Notes in Computer Science*, pages 42–53. Springer Berlin / Heidelberg, 2007. 10.1007/978-3-540-70904-6_6.

[4] Ricardo J. Campello, Eduardo R. Hruschka, and Vinícius S. Alves. On the efficiency of evolutionary fuzzy clustering. *Journal of Heuristics*, 15:43–75, February 2009.

[5] M. Chalmers. A linear iteration time layout algorithm for visualising high-dimensional data. In *Visualization '96. Proceedings.*, pages 127 –131, 27 1996-nov. 1 1996.

[6] Ritendra Datta, Dhiraj Joshi, Jia Li, and James Z. Wang. Image retrieval: Ideas, influences, and trends of the new age. *ACM Comput. Surv.*, 40:5:1–5:60, May 2008.

[7] Thomas Deselaers, Daniel Keysers, and Hermann Ney. Features for image retrieval: An experimental comparison. *Information Retrieval*, 11:77–107, 03/2008 2008.

[8] Danilo Eler, Marcel Nakazaki, Fernando Paulovich, Davi Santos, Gabriel Andery, Maria Oliveira, Joao Batista Neto, and Rosane Minghim. Visual analysis of image collections. *The Visual Computer*, 25:923–937, 2009. 10.1007/s00371-009-0368-7.

[9] Christos Faloutsos and King-Ip Lin. Fastmap: a fast algorithm for indexing, data-mining and visualization of traditional and multimedia datasets. *SIGMOD Rec.*, 24:163–174, May 1995.

[10] Martin Gutlein, Eibe Frank, Mark Hall, and Andreas Karwath. Large-scale attribute selection using wrappers. In *Proceedings of the IEEE Symposium on Computational Intelligence and Data Mining, CIDM 2009, part of the IEEE Symposium Series on Computational Intelligence 2009, Nashville, TN, USA,March 30, 2009 - April 2, 2009*, pages 332–339. IEEE, 2009.

[11] F. Jourdan and G. Melangon. Multiscale hybrid mds. In *Information Visualisation, 2004. IV 2004. Proceedings. Eighth International Conference on*, pages 388 – 393, july 2004.

[12] Yehuda Koren, L. Carmel, and D. Harel. Ace: a fast multiscale eigenvectors computation for drawing huge graphs. In *Information Visualization, 2002. INFOVIS 2002. IEEE Symposium on*, pages 137 – 144, 2002.

[13] J. Kruskal. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29:1–27, 1964. 10.1007/BF02289565.

[14] W. J. Krzanowski. *Principles of multivariate analysis: a user's perspective*. Oxford University Press, Inc., New York, NY, USA, 1988.

[15] D. Ashok Kumar and J. Esther. Article: Comparative study on cbir based by color histogram, gabor and wavelet transform. *International Journal of Computer Applications*, 17(3):37–44, March 2011. Published by Foundation of Computer Science.

[16] R. Fergus L. Fei-Fei and P. Perona. Learning generative visual models from few training examples: an incremental bayesian approach tested on 101 object categories. *IEEE. CVPR 2004, Workshop on Generative-Model Based Vision*, 2004.

[17] Niels Landwehr, Mark Hall, and Eibe Frank. Logistic model trees. *Mach. Learn.*, 59:161–205, May 2005.

[18] Priti Maheshwary and Namita Srivastav. Retrieving similar image using color moment feature detector and k-means clustering of remote sensing images. In *Proceedings of the 2008 International Conference on Computer and Electrical Engineering*, pages 821–824, 2008.

[19] A. Morrison, G. Ross, and M. Chalmers. A hybrid layout algorithm for sub-quadratic multidimensional scaling. In *Information Visualization, 2002. INFOVIS 2002. IEEE Symposium on*, pages 152 – 158, 2002.

[20] H. Muller, N. Michoux, D. Bandon, and A. Geissbuhler. A review of content-based image retrieval systems in medical applications - clinical benefits and future directions. *International Journal of Medical Informatics*, 73(1):1–23, 2004.

[21] F. V. Paulovich, D. M. Eler, J. Poco, C. P. Botha, R. Minghim, and L. G. Nonato. Piecewise laplacian-based projection for interactive data exploration and organization. *Computer Graphics Forum (EuroVis'2011)*, (to appear), 2011.

[22] F.V. Paulovich, L.G. Nonato, R. Minghim, and H. Levkowitz. Least square projection: A fast high-precision multidimensional projection technique and its application to document mapping. *Visualization and Computer Graphics, IEEE Transactions on*, 14(3):564 –575, may-june 2008.

[23] F.V. Paulovich, C.T. Silva, and L.G. Nonato. Two-phase mapping for projecting massive data sets. *Visualization and Computer Graphics, IEEE Transactions on*, 16(6):1281 –1290, nov.-dec. 2010.

[24] E. Pekalska, D. de Ridder, R. P. W. Duin, and M. A. Kraaijweld. A new method of generalizing sammon mapping with application to algorithm speed-up. In M. Boasson, J. A. Kaandorp, J.F.M. Tonino, and M. G. Vosselman, editors, *Annual Conf. Advanced School for Comput. Imag.*, pages 221–228, 1999.

[25] Sam T. Roweis and Lawrence K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500):2323–2326, 2000.

[26] G.A.F. Seber. *Multivariate Observations*. Wiley, New York, 1984.

[27] A. Smeulders, M. Worring, A. Gupta, and R. Jain. Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12):1349–1380, 2000.

[28] O. Sorkine, D. Cohen-Or, Y. Lipman, M. Alexa, C. Rössl, and H.-P.Seidel. Laplacian surface editing. In *SGP '04: Proceedings of the 2004 Eurographics/ACM SIGGRAPH symposium on Geometry processing*, pages 175–184, New York, NY, USA, 2004. ACM.

[29] Joshua B. Tenenbaum, Vin de Silva, and John C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323, 2000.

[30] Sergios Theodoridis and Konstantinos Koutroumbas. *Pattern recognition*. Academic Press, 2009.

[31] Warren Torgerson. Multidimensional scaling of similarity. *Psychometrika*, 30:379–393, 1965.

[32] K. Xu, H. Zhang, D. Cohen-Or, and Y. Xiong. Dynamic harmonic fields for surface processing. *Comput. Graph.*, 33(3):391–398, 2009.

[33] W. Zhao, R. Chellapa, P. Phillips, and A. Rosenfeld. Face recognition: A literature survey. *ACM Computing Surveys*, 8(4):399–458, 2000.

[34] X. Zhou and T. Huang. Relevance feedback in image retrieval: A comprehensive review. *Multimidia Systems*, 40:262–282, 2003.