



Universidade de São Paulo

Biblioteca Digital da Produção Intelectual - BDPI

Departamento de Ciências de Computação - ICMC/SCC

Comunicações em Eventos - ICMC/SCC

2014-10

Have you met VikS? A novel framework for visual diversity search analysis

Brazilian Symposium on Databases, 29th; Demos and Applications Session, 2014, Curitiba.
<http://www.producao.usp.br/handle/BDPI/48639>

Downloaded from: Biblioteca Digital da Produção Intelectual - BDPI, Universidade de São Paulo

paper:4

Have You Met VikS? A Novel Framework for Visual Diversity Search Analysis*

Lucio F. D. Santos¹, Rafael L. Dias², Mônica R. P. Ferreira¹,
Marcela X. Ribeiro², Agma J. M. Traina¹, Caetano Traina Jr.¹

¹Database and Images Group (GBDI)
Institute of Mathematics and Computer Science (ICMC)
University of São Paulo (USP)
São Carlos – SP – Brazil

{luciodb,monika,agma,caetano}@icmc.usp.br

²Department of Computing
Federal University of São Carlos (UFSCar)
São Carlos – SP – Brazil

{rafael.dias,marcela}@dc.ufscar.br

Abstract. Searching images based on their pictorial content, or content-based image retrieval (CBIR), instead of using traditional tags and labeling attached to them has attracted considerable attention. However, this retrieval by content may often retrieve images too similar among themselves. Considering a diversity factor has been a way of improving the quality of results retrieved by user queries. There are still questions about how this factor is used in the searches. In this paper, we present VikS¹, a CBIR system that answers queries based on the similarity and diversity paradigms and supports visual data mining techniques, becoming the user an active agent in the query process and enhancing the understanding of the impact of the diversity in the k -nearest neighbor queries. This framework provides implementations of a wide suite of algorithms to compute and compare diverse results. Users can tune diversification parameters, combine similarity with diversity and see how diverse results are in a projecting space which highlight the distance distribution of the elements.

1. Introduction

Most Content-Based Image Retrieval (CBIR) systems query image sets employing the similarity paradigm [Santos et al. 2013]. They retrieve images that are similar to the given query image based on images' pictorial content instead of using traditional data (text or numerical attributes annotated to the image). Typically, each image is represented by a feature vector, which is a numeric representation summarizing the image visual characteristics, such as color, shape or texture. To establish a similarity degree among images, the feature vectors are compared by a distance function. Based on a similarity criteria (either range or k -nearest neighbor), the system retrieves images that are closer/similar to the given query image. A range query (Rq) retrieves all images closer to or at the same

*This research was supported by FAPESP, CNPq, Capes, RESCUER, and SticAMSUD.

¹<https://db.tt/Px8yzUng>

distance from the query image, while the k -nearest neighbor ($kNNq$) retrieves the k ones most similar to the query image.

However, in some scenarios, querying a dataset employing $kNNq$ and/or Rq criteria may often return images that present too much similarity among themselves, specially if the data has clusters. The problem of having elements too similar in result sets induces the user to think that the dataset does not have his/her answer, often requiring query reformulations, which can lead the user to give up the search. A solution is to provide a diversity degree between images in the results, aiming at approximating them to the user's expectation. The intuition behind diversity is generating a result set that includes elements not only similar to the query element, but also diverse among themselves.

Diversity has been explored in various ways in different areas [Drosou and Pitoura 2013]. Most approaches use metadata related to the elements besides the distances among them, which restrict their usage to datasets that have this information in advance [Vieira et al. 2011]. Others have pursued diversity without using extra information, called the distance-based approach. Here, similarity and diversity measurements must compete with each other, exploiting a user-defined diversity preference parameter. The result of traditional similarity algorithms is re-ranked, inducing diversity among elements based on a bi-criteria objective function, which results in an NP-hard problem [Santos et al. 2013]. To overcome their high computational cost, randomized or greedy algorithms are used to provide approximate results in feasible time.

In the literature, there are some prototypes aimed at optimizing and evaluating the distance-based approaches, such as DivDB [Vieira et al. 2011] and POIKILO [Drosou and Pitoura 2013]. In DivDB, the user can manipulate the *trade-off* between diversity and similarity and other query parameters. However, even though the user can participate more actively in the query process, the comprehension of how the diversity factor is being applied in the whole dataset is not explored [Dias et al. 2013]. In a similar way, POIKILO provides tools to tune several query parameters and diversity models. Geographical dataset produces some visual information about the diversification process, but there is no feedback regarding the relationship among elements in the image dataset. Despite their efforts, these systems do not provide information to help the user to understand the diversified answer.

Using visualization techniques together with user feedback is a way to bring the user as an active agent in the diversity query process. Visual Data Mining (VDM) techniques [Keim 2002] allow creating a process for knowledge extraction that combines computational efficiency with human judgment. Different visualization and interaction strategies were developed, turning easier the analysis and exploration of image datasets and related information by the user. An example of a VDM technique is the multidimensional projection LAMP (Local Affine Multidimensional Projection) [Joia et al. 2011] that employs the local mapping, using an initial sample subset to guide the position of other elements in the final projection. Thus, the user can manipulate the elements and bend the projection according to his/her knowledge about the correlation between the control points previously projected, driving interactively to the final projection.

In this paper we present the Visual Interactive k Diverse Neighbor System (VikS), a CBIR system that answers queries based on the similarity and diversity paradigms and

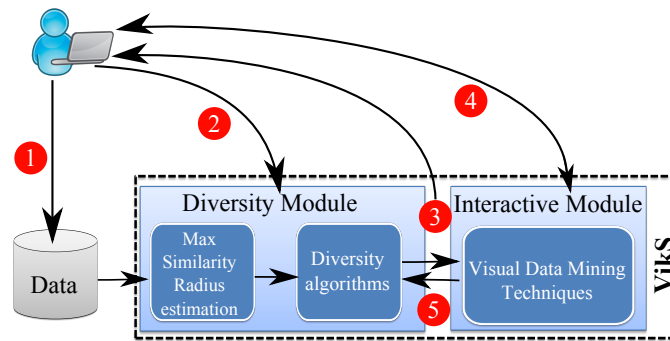


Figure 1. The VikS architecture.

supports VDM techniques to visualize and interact with query results, semantically improving them.

2. The VikS Framework

The *Visual Interactive k Diverse Neighbor System* – VikS – is a framework that helps users to understand and interact with the answers provided by similarity with diversity search algorithms. The framework takes advantage of VDM techniques to provide tools for the user to explore interactively the dataset, engaging more actively in the process and steering the results according to his/her needs. Therefore, the user becomes an active agent in the analysis process, providing feedback about the relationship of the elements in the search space.

Figure 1 presents an overview of VikS architecture. Starting VikS, the user should provide the dataset to be analyzed and the query image (Step 1). Due to the high computational cost to answer diversity search, the user must define the ‘Max Similarity Radius Estimation’ parameter in the ‘Diversity Module’ (Step 2) to restrict the search space for the most similar elements to the query image, since too much dissimilar elements may induce users to think that the query was misinterpreted. Therefore, the diverse candidate elements should not be farther than a distance threshold $MaxSim$ to the query image. However, the definition of this threshold for the user is not easy, because it can vary according to the distance distribution of the dataset. To surpass such problem, VikS takes advantage of an optimization technique that applies the dataset fractal dimension analysis to estimate the $MaxSim$ just using the number of elements expected in the candidate set, being more friendly for the user. Another parameter that should be provided in the ‘Diversity Module’ is the ‘Diversity Algorithms’. In this parameter, the user chooses which algorithm will be used for search. Viks has the following literature algorithms available: Greedy with Marginal Contribution (GMC) [Vieira et al. 2011], Greedy Randomized with Neighborhood Expansion (GNE) [Vieira et al. 2011], Maximal Marginal Relevance (MMR) [Carbonell and Goldstein 1998], SWAP [Vieira et al. 2011], Max-Sum Dispersion (MSD) [Vieira et al. 2011], MOTLEY [Vieira et al. 2011] and a parallel version of the brute-force algorithm (OPD) [Dias et al. 2013]. Those algorithms differ in the result set strategy construction which impacts in the running time and the result quality.

The plain result (only the images) is presented to the user using the n -dimensional features space (Step 3), that is, the search was employed using the original space that

has the low-level features (basic characteristics such as color and shape) extracted by an automatic image processing algorithm, which may not entirely represent the semantic expected by the user cognition about the images. At this point, the user interacts with VikS executing queries with different adjustment and algorithms, as already happens in the other literature systems.

In addition, the user can activate the ‘Interactive Module’ (Step 4) and manipulate the search space, since VikS uses the LAMP VDM technique to map n -dimensional data in a 2-dimensional space, preserving most of the relationships between distances of elements from the original space. Thus, projections of the query image and initial control points are presented to the user, which allow interactive edition, such as shifting elements in the space and changing the position of the query image, following prior knowledge of the user about the dataset. To assist the activity of bending and edition of the initial projection space, the user can request to visualize the image identifier or the preview (thumbnail) of a projected instance. Besides that, each class of images (this information is retrieved from metadata) is identified by different colors, enhancing the user understanding about the relationship among the elements and their classes.

The edited projection space (Step 5) is sent back to the ‘Diversity Module’, which is used to generate a new diverse result set using the information provided by the user in the Step 4. As a result, VikS returns the plain images and the projection space with information about the elements selected in each space, n -dimensional (extracted features) and the edited one. Aimed at highlighting the advantages of using the user’s knowledge in the query process, VikS provides information about which images were considered as candidate subsets to the diversity algorithm, i.e., depicts the ‘Max Similarity Radius’ estimated to prune the search space, as well as the proximity (similarity) in which they are positioned in projected space.

To the best of our knowledge, VikS is the first framework allowing users to compare different diversifying algorithms and participate in the query process bending and steering the results through visual data mining techniques, providing a better understanding of the impact of the diversity factor in the $kNNq$ and how those algorithms differ from each other.

3. Demonstration of the ViKs Framework

Aimed at providing portability, the ViKs framework was coded in Java EE. Its main interface is presented in Figure 2. Users are allowed to submit queries, see results diversified, tune the diversification parameters and interact with the search space steering the result to its needs. To allow a download-and-run experience, two real images datasets are pre-loaded in the system: MRIBalan [Balan et al. 2005] and Faces [Vieira et al. 2011]. The former is composed of 704 images of magnetic resonance imaging obtained from the Clinical Hospital at Ribeirão Preto of University of São Paulo (USP). The image feature vectors were obtained by the method proposed in [Balan et al. 2005]. The latter is a collection of 300 gray images of 30 people’s faces with slight variations of the head positions. The image feature vectors were extracted using gray-level histogram. For both datasets, the feature vectors are compared using the Euclidean distance function (L_2) evaluated over the full feature vectors extracted.

Upon starting the framework, a panel is presented, where the user should config-

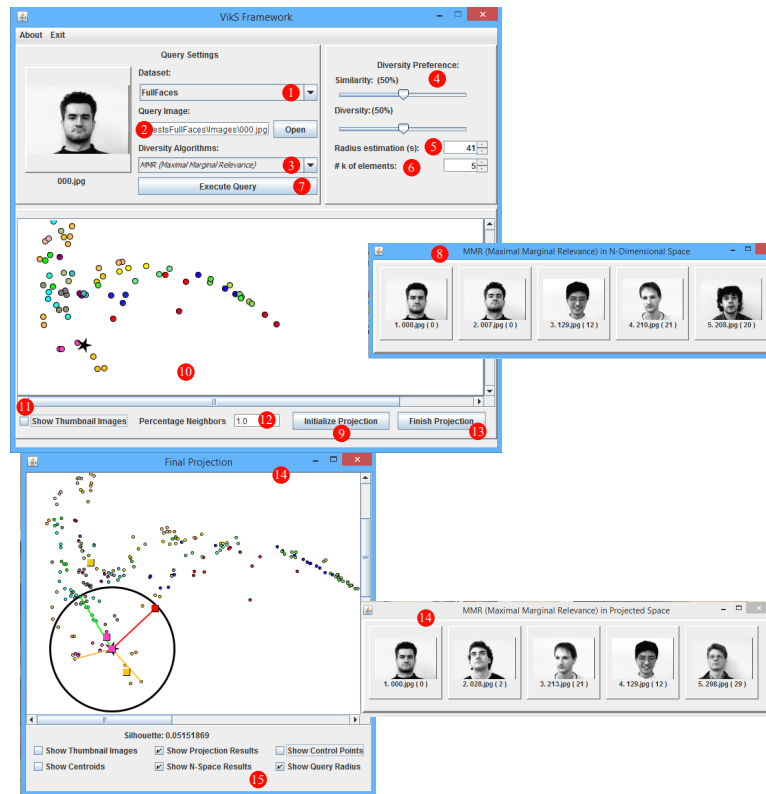


Figure 2. The VikS framework interface.

ure the query settings, selecting the dataset (Step 1), the query image (Step 2), and the diversity algorithm to be executed (Step 3). Then, the user defines which is his/her preference for the results, starting from pure similarity to full diversity answers (Step 4). The number $MaxSim(s)$ of elements to be estimated by the max similarity radius parameter (Step 5) and the number of k elements to be returned (Step 6) should also be specified. The diversity algorithm, the diversity preference and the $MaxSim$ are the diversification parameters of the ‘Diversity Module’ (Figure 1). At this point, the user can submit the query to the system (Step 7) using only the n -dimensional space without providing any more feedback. The diversified result is showed in the result window (Step 8).

Another option is to activate the ‘Interactive Module’ clicking on the ‘Initialize Projection’ button (Step 9), which provides feedback about the similarity of the elements in the search space and the position of the query image (Step 10). The colors indicate the class of each image, and the user can see the preview (thumbnail images) marking the check box (Step 11) as well as defining the number of neighbors to control the projection (Step 12), which has 1.0 as the default value.

The user provides feedback about the search space handling/bending the information presented in the initial projection (Step 10). It is also possible to redo the query using the new distribution of the images (Step 13). Finally, VikS returns the result in a projected space (Step 14). At this point, the user have the option to tune the visualization by zooming-in or out of the presented projection. It is also possible to highlight the factors involved in this process (Step 15) showing the results in the n -dimensional space, the projection results, the max similarity radius, the clusters centroids, the thumbnail images

and the element color indicating the different classes of the dataset (this information must be defined in advance).

4. Conclusions

Querying by image content shows up as a complement to text-based query. In certain situations, there is a greater interest in analyzing the image by its visual features than just using the textual information related to them. However, the manipulation of information concerning images' visual content can create some drawbacks to the user analysis when the returned images for a query are too similar to the query image and among themselves. One way to reduce the problem is by adding a diversity factor to the query, using besides the similarity between the query image and other images in the dataset, some dissimilarity in the resulting subset. However, specifying dissimilarity is not an easy task, since the impact of diversification process in the dataset is almost always unclear to the user.

In this paper, we presented the VikS framework, which uses the user's knowledge to assist in the query diversification process. Additionally, visualization and interaction techniques were used in order to engage the user in a more active role in the query process, enabling a better understanding of the results and relationships between images and categories. To the best of our knowledge, VikS is the first framework to integrate similarity with diversity search to visual data mining technique, enhancing the understanding of the impact of the diversity in the k -nearest neighbor queries.

References

- Balan, A., Traina, A. J. M., Traina, A., and Azevedo-Marques, P. (2005). Fractal analysis of image textures for indexing and retrieval by content. In *Proc. of CBMS*, pages 581–586.
- Carbonell, J. G. and Goldstein, J. (1998). The use of MMR, diversity-based reranking for reordering documents and producing summaries. In *Proc. ACM SIGIR*, pages 335–336, Melbourne, Australia. ACM.
- Dias, R. L., Bueno, R., and Ribeiro, M. X. (2013). Reducing the complexity of k -nearest diverse neighbor queries in medical image datasets through fractal analysis. In *Proc. CBMS.*, pages 101–106.
- Drosou, M. and Pitoura, E. (2013). Poikilo: A tool for evaluating the results of diversification models and algorithms. *Proc. VLDB Endow.*, 6(12):1246–1249.
- Joia, P., Coimbra, D., Cuminato, J. A., Paulovich, F. V., and Nonato, L. G. (2011). Local affine multidimensional projection. *IEEE Trans. Visual. Comput. Graphics*, 17(12):2563–2571.
- Keim, D. A. (2002). Information visualization and visual data mining. *IEEE Trans. Visual. Comput. Graphics*, 8(1):1–8.
- Santos, L. F. D., Oliveira, W. D., Ferreira, M. R. P., Cordeiro, R. L. F., Traina, A. J. M., and Jr., C. T. (2013). Evaluating the diversification of similarity query results. *Journal of Information and Data Management*, 4(3):188–203.
- Vieira, M. R., Razente, H. L., Barioni, M. C. N., Hadjieleftheriou, M., Srivastava, D., Jr., C. T., and Tsotras, V. J. (2011). Divdb: A system for diversifying query results. *PVLDB*, 4(12):1395–1398.