Exploiting Contextual Information for Image Re-Ranking and Rank Aggregation in Image Retrieval Tasks

Daniel Carlos Guimarães Pedronette[1], **Ricardo da S. Torres**[1]

[1]Institute of Computing (IC), University of Campinas (UNICAMP)
Campinas, Brazil

{dcarlos,rtorres}@ic.unicamp.br

Abstract. *In Content-based Image Retrieval (CBIR) systems, accurately ranking collection images is of great relevance, since users are interested in the returned images placed at the first positions. However, CBIR systems often consider only pairwise image analysis, i.e., they compute the similarity measures by taking into account 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. This PhD work proposed five novel re-ranking and rank aggregation algorithms aiming at exploiting contextual information for improving the effectiveness of CBIR systems. We also proposed approaches for combining re-ranking and rank aggregation methods and for executing them efficiently on GPUs.*

Resumo. *Em Sistemas de Recuperação de Images baseados no Conteúdo, a ordenação eficaz de imagens de uma coleção é de grande relevância, já que usuários estão interessados nas imagens retornadas nas primeiras posições. Entretanto, geralmente, estes sistemas analisam apenas pares de imagens para a geração das listas de resultados, ignorando importantes informações codificadas nos relacionamentos entre as imagens. Em contrapartida, a percepção dos usuários usualmente considera a especificação da consulta e os resultados em um dado contexto. Esse trabalho de doutorado propôs cinco novos algoritmos de re-ranking e rank aggregation com o objetivo de explorar informação contextual para melhorar a eficácia de sistemas de CBIR. Foram propostas também abordagens para combinar os métodos de re-ranking e rank aggregation e para executar eficientemente os métodos propostos em GPUs.*

1. Introduction

Given the huge growth of image collections and multimedia resources, traditional retrieval systems based on image metadata can be unfeasible in many situations, specially for large datasets, since they require large amount of human intervention. A promising approach relies on supporting image searches by taking into account image content information, using the well-known Content-Based Image Retrieval (CBIR) systems. Basically, a CBIR system aims at retrieving the most similar images in a collection by taking into account image visual properties (*e.g.*, shape, color, and texture).

Several efforts have been proposed for improving the effectiveness of CBIR approaches, considering more accurate features or similarity measures. However, in general, those initiatives 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 relationships among images. On the other hand, the user perception usually considers the query specification and the query responses in a given *context*. Recently, the relationships among images encoded in ranked lists and distances have been used for extracting *contextual information* [Jegou et al. 2010].

In this PhD work, we proposed five novel *re-ranking* methods aiming at improving the effectiveness of image retrieval tasks replacing pairwise similarities by more global affinities. For each re-ranking method, we proposed a *rank aggregation* approach that uses the *re-ranking* method to combine different CBIR descriptors. In summary, we put efforts on post-processing the distance/similarity scores, by taking into account the *contextual information* available in relationships among images in a given collection. The overall goal of proposed methods is to mimic the human behavior on judging the similarity among objects. An example of the iterative behavior of proposed re-ranking algorithms can be observed in the results showed in Figure 1. The figure illustrates the evolution of rankings (and their effectiveness) along iterations. Each row presents 20 results for a query image (first column with green border) for each iteration. The re-ranking algorithm moves the non-similar images down in the ranked lists, and as a result of this process, the quality of the ranked lists is improved.

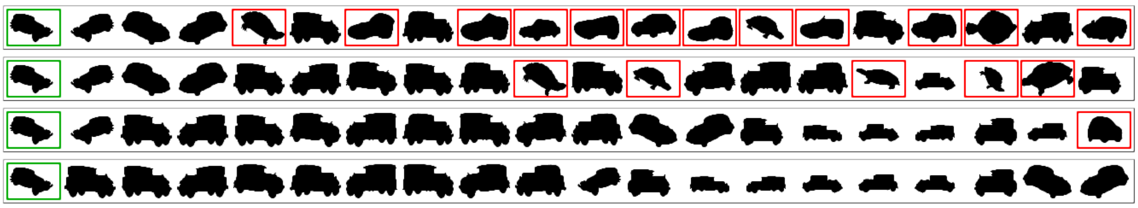

Figure 1. Evolution of rankings along iterations using a re-ranking algorithm.

The proposed methods require no user intervention, training or labeled data, operating on an absolutely *unsupervised* way. The use of different strategies by the proposed methods contributes for analysing different, and complementary aspects encoded in relationships among images. Therefore, we also proposed approaches for combining the outputs of *re-ranking* and *rank aggregation* methods, aiming at further improving the effectiveness of CBIR results. In addition, we also proposed an efficient parallel approach for computing re-ranking methods on GPUs.

2. Main Contributions

2.1. Distance Optimization Algorithm

The *Distance Optimization Algotihm - DOA* ([Pedronette and da S. Torres 2010c, Pedronette and da S. Torres 2010a, Pedronette and da S. Torres 2011a]) is a re-ranking method that uses an iterative clustering approach for performing image re-ranking in CBIR tasks. 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. The main idea of the algorithm consists in clustering images and then using the created clusters for updating distances and performing image re-ranking. These steps are repeated in an iterative manner until a convergence criterion is reached.

2.2. Pairwise Recommendation

We proposed a novel re-ranking method conceptually based on recommender systems [Pedronette and da S. Torres 2012c]. Recommender systems attempt to reduce information overload by selecting automatically items that match the personal preferences of each user. Although inspired by the concept of recommendation, our *Pairwise Recommendation* approach does not require any user intervention. Recommendations are made based on *image profiles*. These image profiles are created by taking into account the relationships among images encoded in ranked lists.

2.3. Context Spaces

We proposed the creation of *contextual spaces* [Pedronette et al. 2013c, Pedronette and da S. Torres 2011c] for encoding contextual information. A *contextual space* is a bidimensional representation of the image collection, considering two reference images. In the proposed approach, the *contextual spaces* are constructed considering the most similar images to a given query image. Later, new distance scores are computed by taking into account the distances among these neighbours to other collection images. The image collection is re-ranked based on the new distances and this process is repeated along iterations.

2.4. RL-Sim

We proposed the *RL-Sim Re-Ranking* algorithm [Pedronette and da S. Torres 2013, Pedronette and da S. Torres 2011d], a new re-ranking method that considers the similarity among ranked lists (***R****anked****L****ists-****Sim****ilarities*) for characterizing contextual information in CBIR systems. The modeling of contextual information considering only the similarity between ranked lists represents an advantage of this algorithm. In this way, the re-ranking method does not depend on distances or similarity scores 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.

2.5. Contextual Re-Ranking

The *Contextual Re-Ranking* algorithm [Pedronette and da S. Torres 2011b, Pedronette and da S. Torres 2012b, Pedronette and da S. Torres 2010b] is a novel approach for retrieving contextual information, by creating a gray scale image representation of distance matrices computed by CBIR descriptors (referenced as *context image*). The *context images* represent a great source of information about the entire image collection. A single *context image* contains information about all distances among images and their spatial relationship defined by the ranked lists of the reference images. The contextual information encoded in the context images are later processed using image processing techniques. This last step represents the main novelty of this work.

2.6. Efficient Image Re-Ranking Computation on GPUs

The usefulness of CBIR systems depends on both the *effectiveness* and the *efficiency* of the retrieval process. While the effectiveness is related to the quality of retrieved images, the efficiency is related to the time spent to obtain the results. High effectiveness and efficiency are indispensable for the construction of useful real-world retrieval systems. We addressed the image re-ranking performance challenges by designing and implementing a re-ranking algorithm that takes advantage of the massive amount of parallelism offered by GPUs. We proposed a parallel GPU-based solution [Pedronette et al. 2012a] which can speed up the computation of *Contextual Re-Ranking* algorithm. We used the OpenCL standard, a new industry standard for task-parallel and data-parallel heterogeneous computing. Experimental results demonstrated that very significant speedups (up to $10.4\times$ in the core of the algorithm) can be achieved by our parallel solution.

2.7. Combining Re-Ranking and Rank Aggregation

Although a lot of efforts have been employed to develop new re-ranking and rank aggregation methods, there are few initiatives aiming at *combining* the existing methods. Besides that, in the same way that different CBIR descriptors produce different and complementary rankings, results of *re-ranking* and *rank aggregation* methods can also be combined to obtain more effective results. We proposed three novel approaches [Pedronette and da S. Torres 2012a] aiming at improving the effectiveness of CBIR systems by combining: (*i*) re-ranking algorithms; (*ii*) rank aggregation algorithms; (*iii*) and both re-ranking and rank aggregation algorithms.

3. Results and Conclusions

In this PhD work, we proposed five novel re-ranking and rank aggregation methods, using different approaches for exploiting contextual information in image retrieval tasks. A large experimental evaluation [Pedronette 2012] was conducted, considering different visual (*e.g.*, shape, color, and texture) and multimodal (visual and textual) features and comparisons with state-of-the-art methods. The obtained results demonstrated the effectiveness of our methods, achieving effectiveness gains up to 29% for some CBIR descriptors. We have also showed that the methods can be combined and efficiently executed using parallel computing.

This PhD work was developed in three years (March/2009 - April/2012) and resulted in several publications in selective journals and conferences of the pattern recognition, computer vision, image processing, and multimedia retrieval areas. The proposed contributions were published in 13 scientific papers (5 international journal papers, 7 international conference papers, and one national conference paper).

Published Journal Papers:

- **Pattern Recoginition** (*Qualis A1*) [Pedronette and da S. Torres 2013];
- **Information Sciences** (*Qualis A1*) [Pedronette and da S. Torres 2012c];
- **Journal of Visual Languages & Computing** (*Qualis B1*) [Pedronette and da S. Torres 2011a];
- **Multimedia Tools and Applications** (*Qualis B2*) [Pedronette et al. 2013c];
- **International Journal of Multimedia Information Retrieval** [Pedronette and da S. Torres 2012b];

Published Conference Papers:

- **ICIP 2011** (*Qualis A1*) [Pedronette and da S. Torres 2011b]
- **ICMR 2011** (*Qualis A2*) [Pedronette and da S. Torres 2011c]
- **SIBGRAPI 2010** (*Qualis B1*) [Pedronette and da S. Torres 2010a]

- **CAIP 2011** (*Qualis B2*) [Pedronette and da S. Torres 2011d]
- **CIARP 2012** (*Qualis B2*) [Pedronette and da S. Torres 2012a]
- **CIARP 2010** (*Qualis B2*) [Pedronette and da S. Torres 2010b]
- **ISPA 2012** (*Qualis B3*) [Pedronette et al. 2012a]
- **VISAPP 2010** (*Qualis B3*) [Pedronette and da S. Torres 2010c]

Other related contributions were obtained by extending the proposed approaches after the thesis defense. The rank aggregation methods have already been applied with good results in different scenarios such as multimodal image retrieval [Kozievitch et al. 2011] and multimodal geocoding [Li et al. 2012a, Li et al. 2012b, Li et al. 2013]. Advances have also been obtained in terms of scalability [Pedronette et al. 2013a] and efficient execution of re-ranking methods [Pedronette et al. 2012b, Pedronette et al. 2013b] by using parallel computing.

Figure 2 summarizes the overall organization of the PhD work, considering the main concepts and their relationships: the main subjects related to the research (in shades of red), the main proposed contributions (in shades of green), the background concepts used to address the contributions (in blue), the associated publications (in shades of yellow), and possible future work (in gray).

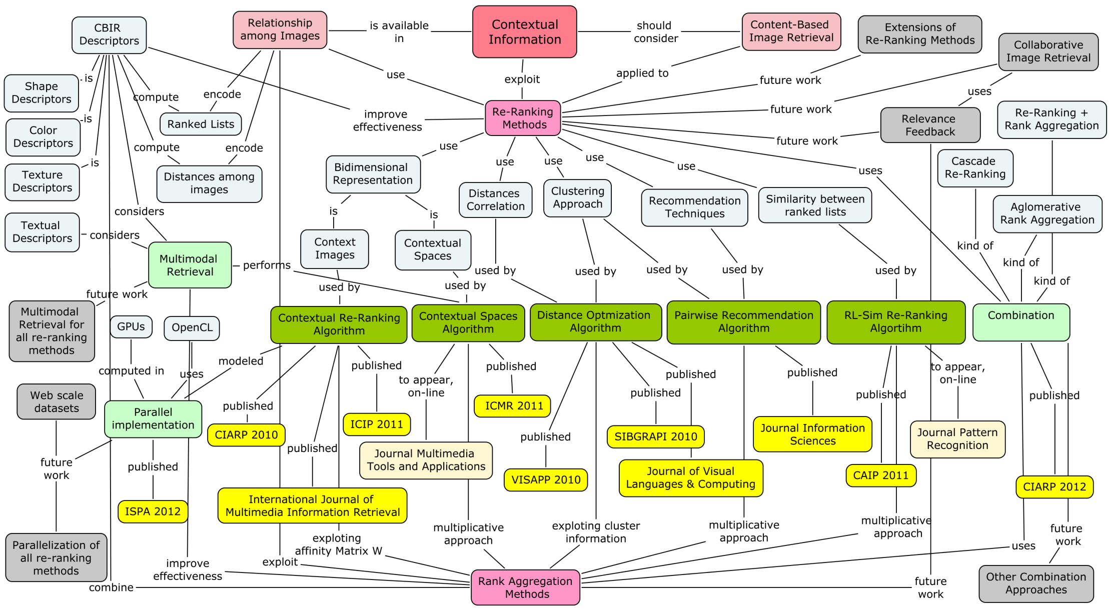

Figure 2. Main concepts, contributions, publications, and possible extensions of the PhD work.

Acknowledgment

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

References

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

Kozievitch, N., Pedronette, D. C. G., da S. Torres, R., Santanchè, A., Fox, E., and Calumby, R. T. (2011). An infrastructure for searching and harvesting complex image objects. *I3 Journal*, 11(2):39–68.

Li, L. T., Almeida, J., Pedronette, D. C. G., Penatti, O. A. B., and da Silva Torres, R. (2012a). A multimodal approach for video geocoding. In *MediaEval'2012*.

Li, L. T., Pedronette, D. C. G., Almeida, J., Penatti, O. A. B., Calumby, R. T., and da Silva Torres, R. (2012b). Multimedia multimodal geocoding. In *SIGSPATIAL/GIS'2012*, pages 474–477.

Li, L. T., Pedronette, D. C. G., Almeida, J., Penatti, O. A. B., Calumby, R. T., and da Silva Torres, R. (2013). Accurate video multimodal geocoding. *Multimedia Tools and Applications*. Major Revision.

Pedronette, D. C. G. (2012). *Exploiting Contextual Information for Image Re-Ranking and Rank Aggregation in Image Retrieval Tasks*. PhD thesis, Campinas, SP, Brazil. Advisor: Ricardo da S. Torres.

Pedronette, D. C. G., Almeida, J., and da S. Torres, R. (2013a). A scalable re-ranking method for content-based image retrieval. *Information Sciences*. Major Revision.

Pedronette, D. C. G. and da S. Torres, R. (2010a). Distances correlation for re-ranking in content-based image retrieval. In *Conf. on Graphics, Patterns and Images (SIBGRAPI'2010)*, volume 1, pages 1 – 8.

Pedronette, D. C. G. and da S. Torres, R. (2010b). Exploiting contextual information for image re-ranking. In *Iberoamerican Congress on Pattern Recognition (CIARP'2010)*, pages 541–548.

Pedronette, D. C. G. and da S. Torres, R. (2010c). 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, pages 197 – 202.

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

Pedronette, D. C. G. and da S. Torres, R. (2011b). Exploiting contextual information for rank aggregation. In *International Conference on Image Processing (ICIP'2011)*, pages 97–100.

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

Pedronette, D. C. G. and da S. Torres, R. (2011d). Image re-ranking and rank aggregation based on similarity of ranked lists. In *C. Analysis of Images and Patterns (CAIP'2011)*, volume 6854, pages 369–376.

Pedronette, D. C. G. and da S. Torres, R. (2012a). Combining re-ranking and rank aggregation methods. In *Iberoamerican Congress on Pattern Recognition (CIARP'2012)*, pages 170–178.

Pedronette, D. C. G. and da S. Torres, R. (2012b). Exploiting contextual information for image re-ranking and rank aggregation. *International Journal of Multimedia Information Retrieval*, 1(2):115–128.

Pedronette, D. C. G. and da S. Torres, R. (2012c). Exploiting pairwise recommendation and clustering strategies for image re-ranking. *Information Sciences*, 207(1):19–34.

Pedronette, D. C. G. and da S. Torres, R. (2013). Image re-ranking and rank aggregation based on similarity of ranked lists. *Pattern Recognition*. To appear, On-Line.

Pedronette, D. C. G., da S. Torres, R., Borin, E., and Breternitz, M. (2012a). Efficient image re-ranking computation on GPUs. In *Int. Symposium Parallel Distributed Processing (ISPA'2012)*, pages 95–102.

Pedronette, D. C. G., da S. Torres, R., Borin, E., and Breternitz, M. (2012b). Efficient image re-ranking computation using GPUs. In *Escola Regional de Alto Desempenho - São Paulo (ERAD'2012)*.

Pedronette, D. C. G., da S. Torres, R., Borin, E., and Breternitz, M. (2013b). Efficient image re-ranking and rank aggregation on GPUs. *Parallel Computing*. Submitted.

Pedronette, D. C. G., da S. Torres, R., and Tripodi, R. C. (2013c). Using contextual spaces for image re-ranking and rank aggregation. *Multimedia Tools and Applications*. To appear, On-Line.