Electronic Letters on Computer Vision and Image Analysis 13(2):45-46, 2014

Towards an interactive index structuring system for content-based image retrieval in large image databases

LAI Hien Phuong

L3I, Université de La Rochelle, Avenue M. Crépeau, 17042 La Rochelle cedex 1, France
Advisors: Muriel VISANI, Alain BOUCHER, Jean-Marc OGIER
Date and location of PhD thesis defense: 2 October 2013, University of La Rochelle

Received 9th Feb 2014; accepted 27th May 2014

Keywords: Semi-supervised clustering, interactive learning, image indexing.

1 Abstract

In recent years, the expansion of acquisition, storage and transmission techniques and the success of tablet computers facilitate the development of many large image databases as well as the interactions with the users. This thesis [1] deals with the problem of Content-Based Image Retrieval (CBIR) on these huge masses of data. Among three CBIR phases (feature extraction, feature space structuring and retrieval), we are particularly interested in the structuring phase (normally called indexing phase), which plays a very important role in finding information in large databases. This phase aims at organizing the visual feature descriptors of all images into an efficient data structure in order to facilitate, accelerate and improve further retrieval. Instead of traditional structuring methods, clustering methods which organize image descriptors into groups of similar objects (*clusters*), without any constraint on the cluster size, are studied. The aim is to obtain an indexed structure more adapted to the retrieval of high dimensional and unbalanced data. Clustering can be done without prior knowledge (unsupervised clustering) or with a limited amount of prior knowledge (semi-supervised clustering).

Due to the "semantic gap" between high-level semantic concepts expressed by the user via the query and the low-level features automatically extracted from the images, the clustering results and therefore the retrieval results are generally different from the wishes of the user. In this thesis, we proposed to involve the user in the clustering phase so that he/she can interact with the system so as to improve the clustering results as well as the performance of the further retrieval. The idea is as follows. Firstly, an initial clustering is used to organize images into clusters. Then, the user visualizes the clustering result and provides feedback for guiding the re-clustering phase. The system then re-organizes the dataset by using not only the similarity between objects, but also the user feedback in order to reduce the semantic gap. The interactive loop can be iterated until the clustering result satisfies the user. In the case of large database indexing, we assume that the user has no prior knowledge about the image database. Therefore, an unsupervised clustering method is suitable to be used for the initial clustering, when no supervised information is available yet. Then, after receiving the user feedback in each interactive iteration, a semi-supervised clustering can be used for the re-clustering process.

Based on a deep study of different unsupervised clustering methods [3] as well as semi-supervised clustering approaches [2], we propose in this thesis a new interactive semi-supervised clustering model [2] involving the

Correspondence to: <hien_phuong.lai@univ-lr.fr>

Recommended for acceptance by <Alicia Fornés and Volkmar Frinken>
ELCVIA ISSN:1577-5097
Published by Computer Vision Center / Universitat Autònoma de Barcelona, Barcelona, Spain

user in the clustering phase. From the formal analysis of different unsupervised clustering methods [3], we chose to experiment some methods which appear to be the most suitable to be used in an incremental context involving the user. The hierarchical BIRCH unsupervised clustering (Zhang *et al.*, 1996) which gives the best performance from these experiments [3] is chosen to be used as the initial clustering in our model. Then, an interactive loop in which the user provides the feedback and the system re-organizes the database using the new semi-supervised clustering method proposed in this thesis is iterated until the clustering result satisfies the user. As the user has no prior knowledge about the image database, it is difficult for him/her to label the clusters or the images in the clusters using classes. Therefore, we provide an interactive interface allowing the user to easily visualize the clustering result and give feedback. Based on the majority of the images displayed for each cluster, the user can specify relevant or non-relevant images for each cluster. The user can also drag and drop images between clusters in order to change the cluster assignment of some images. Then, supervised information is deduced from the feedback in order to be used for the re-clustering phase. According to our study of different semi-supervised clustering methods, supervised information may consist of class labels for some objects or pairwise constraints (must-link or cannot-link) between objects. The experimental analysis of different semi-supervised clustering methods in the interactive context [2] shows a high performance of the HMRF-kmeans (Basu *et al.*, 2004) which uses pairwise constraints. Inspired from the HMRF-kmeans method, we proposed a new semi-supervised clustering method [2] for the re-clustering phase. Instead of using pairwise constraints between images, our method uses pairwise constraints between the leaf entries (CF entries) of the BIRCH tree as supervised information for guiding the re-clustering of the CF entries. As each CF entry groups a set of similar images, pairwise constraints between images can be replaced by a smaller number of pairwise constraints between CF entries, without reducing the quality of supervised information. And therefore, the processing time could be reduced without decreasing the performance. In our model, after receiving user feedback in each interactive iteration, pairwise constraints can be deduced based not only on the user feedback but also on the neighbourhood information. Neighbourhood information groups images according to the willingness of the user to classify them in the same clusters (via user feedback of all interactive iterations). This kind of information helps to maximize the supervised information gained from a same number of user clicks.

In order to avoid the subjective dependence of the results on the human user, a software agent simulating the behaviour of the human user for providing feedback is used for the experimental analysis of our system using different image databases of increasing sizes (Wang, PascalVoc2006, Caltech101, Corel30k). Moreover, different strategies for deducing pairwise constraints from user feedback and neighbourhood information were investigated. Among these strategies, the strategy which keeps only the most "difficult" constraints (must-link between the most distant objects and cannot-links between the closest objects) was shown to give the best trade-off between the performance and the processing time. Furthermore, the experimental results show that the involvement of the user helps to improve the results and that our semi-supervised clustering outperforms the HMRF-kmeans, in both performance and processing time. Note that our clustering structure can be used not only for facilitating the further retrieval, but also for helping the navigation in large image databases. Moreover, in this thesis, we propose a 2D interface for visualizing the group structure of high dimensional image databases.

References

[1] H.P. Lai (2013). *Towards an interactive index structuring system for content-based image retrieval in large image databases*. PhD Thesis. Université de La Rochelle. (http://tel.archives-ouvertes.fr/docs/00/93/48/42/PDF/thesis.pdf)

[2] H.P. Lai, M. Visani, A. Boucher, J.M. Ogier, "A new interactive semi-supervised clustering model for large image database indexing", *Pattern Recognition Letters*, 37:94-106, 2014.

[3] H.P. Lai, M. Visani, A. Boucher, J.M. Ogier, "An experimental comparison of clustering methods for content-based indexing of large image databases", *Pattern Analysis and Applications*, 15(4):345-366, 2012.