

# Chapter VIII

## A Machine Learning–Based Model for Content–Based Image Retrieval

**Hakim Hacid**

*University of Lyon 2, France*

**Abdelkader Djamel Zighed**

*University of Lyon 2, France*

### ABSTRACT

*A multimedia index makes it possible to group data according to similarity criteria. Traditional index structures are based on trees and use the  $k$ -Nearest Neighbors ( $k$ -NN) approach to retrieve databases. Due to some disadvantages of such an approach, the use of neighborhood graphs was proposed. This approach is interesting, but it has some disadvantages, mainly in its complexity. This chapter presents a step in a long process of analyzing, structuring, and retrieving multimedia databases. Indeed, we propose an effective method for locally updating neighborhood graphs, which constitute our multimedia index. Then, we exploit this structure in order to make the retrieval process easy and effective, using queries in an image form in one hand. In another hand, we use the indexing structure to annotate images in order to describe their semantics. The proposed approach is based on an intelligent manner for locating points in a multidimensional space. Promising results are obtained after experimentations on various databases. Future issues of the proposed approach are very relevant in this domain.*

### INTRODUCTION

Data interrogation is a fundamental problem in various scientific communities. The database and statistics communities (with their various fields such as data mining) are certainly the most implied. Each

community considers the interrogation from a different point of view. The database community, for example, deals with great volumes of data by organizing them in an adequate structure in order to be able to answer queries in the most effective way. This is done by using index structures. The statistics community deals only with data samples in order to produce predictive models that are able to draw conclusions on phenomena; these conclusions then are generalized to the whole items. This is achieved using various structures such as decision trees (Mitchell, 2003), Kohonen maps (Kohonen, 2001), neighborhood graphs (Toussaint, 1991), and so forth.

Dealing with multimedia databases means dealing with content-based retrieval. There are two fundamental problems associated with content-based retrieval systems: (a) how to specify a query and (b) how to access the intended data efficiently for a given query. The main objective is to capture the semantics of the considered data. For traditional database systems, the semantics of content-based access are finding data items that match exactly the specified keywords in queries. For multimedia database systems, both query specification and data access become much harder (Chiueh, 1994).

To give the computer the ability to mimic the human being in scene analysis needs to explicit the process by which it moves up from the low level to the highest one. Multimedia processing tools give many ways to transform an image/video into a vector. For instance, MPEG-7 protocol associates a set of quantitative attributes to each image/video. The computation of these features is integrated and automated fully in many software platforms. In return, the labels basically are given by the user, because they are issued from the human language. The relevance of the image/video retrieval process depends on the vector of characteristics. Nevertheless, if we assume that the characteristics are relevant in the representation space, that it is supposed to be  $R^p$ , the images that are neighbors should have very similar meanings.

In order to perform an interrogation in a multimedia database, it must be structured in an adequate way. For that, an index is used. Indexing a multimedia database consists of finding a way to structure the data so that the neighbors of each multimedia document can be located easily according to one or more similarity criteria. The index structures used in databases are generally in a tree form and aim to create clusters, which are represented by the sheets of the tree and contain rather similar documents. However, in addition to the fact that a traditional index cannot support data with dimensions higher than 16, dealing with multimedia databases needs more operations such as classification and annotation. This is why the use of models issued from the automatic learning community can be very helpful.

The rest of this chapter is organized as follows. The next section introduces the point location and the database indexing problems. Section 3 presents the motivation and the contributions of our work. Section 4 describes the neighborhood graphs that are the foundation of this contribution. Our contributions are addressed in Section 5. The indexing method and the optimization of the neighborhood graphs are discussed in Section 5.1. Semi-automatic annotation is discussed in Section 5.2. Section 6 gives some experiments that were performed in order to evaluate and validate our approach. We conclude and give some future issues in Section 7.

## **POINT LOCATION AND DATABASES INDEXING PROBLEMS**

### **Point Location Problem**

Neighborhood search is a significant problem in several fields. It is handled in data mining (Fayyad, Piatetsky-Shapiro, & Smith, 1996), classification (Cover, & Hart, 1967), machine learning (Cost, &

18 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:  
[www.igi-global.com/chapter/machine-learning-based-model-content/4156](http://www.igi-global.com/chapter/machine-learning-based-model-content/4156)

## Related Content

---

### Collaborative Exploration Based on Simultaneous Localization and Mapping

Domenec Puig (2013). *Robotic Vision: Technologies for Machine Learning and Vision Applications* (pp. 303-332).

[www.irma-international.org/chapter/collaborative-exploration-based-simultaneous-localization/73197](http://www.irma-international.org/chapter/collaborative-exploration-based-simultaneous-localization/73197)

### Building a Multiple Object Tracking System with Occlusion Handling in Surveillance Videos

Raed Almomani and Ming Dong (2013). *Robotic Vision: Technologies for Machine Learning and Vision Applications* (pp. 98-111).

[www.irma-international.org/chapter/building-multiple-object-tracking-system/73186](http://www.irma-international.org/chapter/building-multiple-object-tracking-system/73186)

### Audiovisual Facial Action Unit Recognition Using Feature Level Fusion

Zibo Meng, Shizhong Han, Min Chen and Yan Tong (2018). *Computer Vision: Concepts, Methodologies, Tools, and Applications* (pp. 636-655).

[www.irma-international.org/chapter/audiovisual-facial-action-unit-recognition-using-feature-level-fusion/196972](http://www.irma-international.org/chapter/audiovisual-facial-action-unit-recognition-using-feature-level-fusion/196972)

### Image Processing for Localization and Parameterization of the Glandular Ducts of Colon in Inflammatory Bowel Diseases

Stanislaw Osowski, Michal Kruk, Robert Koktysz and Jaroslaw Kurek (2013). *Image Processing: Concepts, Methodologies, Tools, and Applications* (pp. 688-708).

[www.irma-international.org/chapter/image-processing-localization-parameterization-glandular/77570](http://www.irma-international.org/chapter/image-processing-localization-parameterization-glandular/77570)

### Key Frame Extraction from Video: Framework and Advances

Sergii Mashtalir and Olena Mikhnova (2014). *International Journal of Computer Vision and Image Processing* (pp. 68-79).

[www.irma-international.org/article/key-frame-extraction-from-video/115840](http://www.irma-international.org/article/key-frame-extraction-from-video/115840)