

Chapter 3

Mid-Level Image Descriptors

Jean Martinet

University of Lille, France

Ismail Elsayad

University of Lille, France

ABSTRACT

We present in this chapter a classification of image descriptors, from the low level to the high level, introducing the notion of intermediate level. This level denotes a representation level lying between low-level features – such as color histograms, texture or shape descriptors, and high-level features – semantic concepts. In a chain of process point of view, mid-level descriptors represent an intermediate step or stage between low and high level, dedicated to specific tasks such as annotation, object detection/recognition, or similarity matching. After introducing a definition for the three different levels, we review a number of approaches making use of such intermediate levels. We namely focus on different approaches making an analogy with text processing, by adapting and applying standard text processing techniques to image indexing.

INTRODUCTION

In typical Content-Based Information Retrieval (CBIR) systems, it is always important to select an appropriate representation for documents (Baeza-Yates & Ribeiro-Neto 1999). Indeed, the quality of retrieval results depends on the quality the internal representation of the content. Classical models of information retrieval usually consider that a document is described by a set of descrip-

tors. In text retrieval for instance, the descriptors take the form of representative index terms, that are keywords extracted from the collection.

When considering visual documents, the problem of the *semantic gap* arises. The notion of semantic gap has been defined a decade ago by Smeulders et al. (2000) as the lack of coincidence between the information that one can extract from the visual data and the interpretation that the same data have for a user in a given situation. In the case of images, because of the distance between the raw signal (i.e. the pixel matrix) and its interpretation,

DOI: 10.4018/978-1-61350-126-9.ch003

it is difficult to automatically extract an accurate semantic content representation of their content. Traditional automatic annotation techniques extract information from *low-level descriptors* (or *low level features*) to infer information about corresponding *high-level descriptors*. Note that although the two notations (*descriptor* and *feature*) are similar and can be assimilated, we preferably use the more general notation *descriptor* instead of *feature* in this chapter because of its dedication to indexing.

Because the step from the low level to the high level is not straightforward, many techniques make use of several chains of processes, in order to extract and refine the information incrementally. We call *mid-level descriptors* the result of such intermediate processes that help narrowing the semantic gap. The main target of techniques using such mid-level descriptors could be, for instance, to improve the results quality of a patch classifier (e.g. based on Support Vector Machines) by defining and using mid-level descriptors as an input, as compared using with only low-level descriptors.

The objective of this chapter is to introduce the emerging concept of *mid-level descriptors*, by identifying existing approaches making use of such descriptors, which illustrates and supports the proposed classification of descriptors. We review some of the widely used approaches in visual document indexing using such descriptors.

The remainder of the chapter is organized as follows. We introduce in the following section some definitions of the different levels of description of images. Then we review a number of approaches in image indexing, from the proposed *mid-level description* perspective. We namely explore an analogy between text and image, which is widely used in image indexing. We also introduce other *mid-level* descriptors like image epitomes and spatial data mining. We conclude this chapter by giving the main trends for the future of this research domain.

BACKGROUND

The notion of intermediate level (or mid-level) descriptor is not new. For instance, Luo & Savakis (2001) proposed the use of a Bayesian network for integrating knowledge from low-level to mid-level features for indoor/outdoor classification of images. The network integrates low-level features (color and texture) and so-called mid-level features (external knowledge about sky and grass) using a single classification engine. Mylonas et al. (2007) have also used some mid-level descriptors. Their work aims at improving both image segmentation and labeling of materials and simple objects at the same time, with obvious benefits for problems in the area of image understanding. The novelty of the proposed idea lies on blending well-established segmentation techniques with mid-level features. ARG (Attributed Relational Graphs) were used as mid-level descriptors in their approach, so that images can be described as structured sets of individual objects, thus allowing a straightforward mapping to a graph structure. We provide below process-oriented definitions of low-, high-, and mid-level descriptors. Figure 1 shows a schematic description of the situation each level of description.

Definition 1: Low-Level Descriptor

A low-level descriptor is a continuous or discrete numeric or symbolic measurement that is computed directly from the signal (e.g. image pixels), locally in a (part of a) document. Low-level descriptors (LLD) include usual color histograms, texture and shape descriptors. They qualify measurements operated directly from the signal, in a straightforward manner, involving neither external knowledge/learning process, nor global statistical analysis of other documents. For instance, the local binary pattern (LBP) operator and the Scale-Invariant Feature Transform (SIFT) feature (Lowe, 2004) are considered low-level descriptors.

13 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/mid-level-image-descriptors/59951

Related Content

Dependency Graph Based Detection of Semantically Equivalent Questions in Online Forums

Parmeet Kaur and Shrutika Gulati (2019). *International Journal of Information Retrieval Research* (pp. 50-64).

www.irma-international.org/article/dependency-graph-based-detection-of-semantically-equivalent-questions-in-online-forums/217483

Digital Image Authentication: A Review

Yue Li and Chia-Hung Wei (2013). *Modern Library Technologies for Data Storage, Retrieval, and Use* (pp. 102-128).

www.irma-international.org/chapter/digital-image-authentication/73769

A Novel Incremental Framework for Building Classifier Using Constraint Class Association Rules

B. Subbulakshmi, C. Deisy and S. Parthasarathy (2023). *International Journal of Information Retrieval Research* (pp. 1-21).

www.irma-international.org/article/a-novel-incremental-framework-for-building-classifier-using-constraint-class-association-rules/316125

Ensemble of Support Vector Machine and Ontological Structures to Generate Abstractive Text Summarization

Amita Arora (2022). *International Journal of Information Retrieval Research* (pp. 1-24).

www.irma-international.org/article/ensemble-of-support-vector-machine-and-ontological-structures-to-generate-abstractive-text-summarization/300294

A Data Mining Framework for Forest Fire Mapping

Ahmed Toujani and Hammadi Achour (2018). *Information Retrieval and Management: Concepts, Methodologies, Tools, and Applications* (pp. 771-794).

www.irma-international.org/chapter/a-data-mining-framework-for-forest-fire-mapping/198575