On Combining Sequence Alignment and Feature-Quantization for Sub-Image Searching

Tomas Homola, Masaryk University, Czech Republic
Vlastislav Dohnal, Masaryk University, Czech Republic
Pavel Zezula, Masaryk University, Czech Republic

ABSTRACT
The availability of various photo archives and photo sharing systems made similarity searching much more important because the photos are not usually conveniently tagged. So the photos (images) need to be searched by their content. Moreover, it is important not only to compare images with a query holistically but also to locate images that contain the query as their part. The query can be a picture of a person, building, or an abstract object and the task is to retrieve images of the query object but from a different perspective or images capturing a global scene containing the query object. This retrieval is called the sub-image searching. In this paper, the authors propose an algorithm, called SASISA, for retrieving database images by their similarity to and containment of a query. The novelty of it lies in application of a sequence alignment algorithm, which is commonly used in text retrieval. This forms an orthogonal solution to currently used approaches based on inverted files. They improve efficiency of SASISA by applying vector-quantization of local image feature descriptors. The proposed algorithm and its optimization are evaluated on a real-life data set containing photographs where images of logos are searched. It is compared to a state-of-the-art method (Joly & Buisson, 2009) and the improvement of 16% in mean average precision (mAP) is obtained.

Keywords: Image Matching, Local Image Features, Performance Evaluation, Sequence Alignment, Sub-Image Retrieval

INTRODUCTION
The complexity of search in current business intelligence systems, academic research, or even home audio-visual databases grows up rapidly and users of such systems require searching their data by content. For example, the user sees a cathedral while watching a movie and by taking a snapshot, his or her private collection of holiday photos can be searched for images containing that cathedral. Thus, it is not sufficient to store data and search it by exact match but rather

DOI: 10.4018/jmdem.2012070102

Copyright © 2012, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.
by means of similarity. Similarity searching is especially requested in multimedia databases, but also in digital right management systems, computer aided diagnosis and natural sciences.

In this paper, we focus on the issue of similarity searching in images with a particular interest of searching for parts of images—sub-image searching, where a query contains a picture of an object and the task is to retrieve database images that contain it as their part. In general, state-of-the-art sub-image searching methods build on three main pillars: (i) local features as the basis of image characteristic; (ii) similarity of the local features; and (iii) verification of spatial (geometrical) distribution of the local features. All these tasks are computationally intensive, since the number of local features obtained from a photograph is usually very high, ranging up to thousands. This leads to databases managing billions of data items. Moreover, individual local features are expressed as high-dimensional vectors, which further complicate algorithms from the performance point of view. Here, we do not tackle the issue of extracting local features from images. We rather focus on sub-image searching algorithms and we propose an algorithm that applies a sequence alignment method to solve the issue of the latter two of three pillars, namely identification of similar local features and verification of their spatial correspondence. This algorithm combines both of them into one procedure. This means finding a group of features between two images that are mutually similar and have an appropriate geometrical arrangement simultaneously.

The remaining parts of this paper are structured as follows. The necessary background is given in the next subsection. In Section Sub-image Searching Algorithm, the proposed algorithm is presented including a discussion on its properties and its improvements. In Section Experiments, the paper continues with experimental evaluation and comparison with a state-of-the-art solution. The related work is summarized in Section Related Work. The paper concludes with a Conclusion Section.

**NECESSARY BACKGROUND**

Multimedia data objects are usually characterized as sets of local features, which are spatially distributed in a \( d \)-dimensional space (\( d=1 \) in the case of sound, \( d=2 \) for the images, etc.). Local features \( F \) are likely to be viewed as tuples \((pos_1, ..., pos_d, desc)\), where \( pos \) indicates the spatial position in the \( i \)-th coordinate and is denoted as \( F[|pos|_i] \), while the whole position vector is denoted as \( F[|pos|] \). The actual local descriptor, referenced as \( F[|desc|] \), describes the area surrounding the point \( F[|pos|] \). In addition, the descriptor may contain other attributes internally, e.g., expressing its importance or orientation. In SIFT (Lowe, 2004), the scale attribute is used for such.

**SUB-IMAGE SEARCHING ALGORITHM**

We propose a novel algorithm for sub-image searching based on sequence alignment, called SASISA (Sequence-Alignment-based Sub-Image Searching Algorithm). Its idea is based on taking a set of local features as a sequence and applying a local/global sequence alignment algorithm to reveal whether a query is contained in an image or not. In particular, the sequences are strings of local features (instead of letters of a regular alphabet). A sequence is obtained by projecting positions of local features to a line and ordering them linearly. Formally, the projection is defined as 

\[ P : \mathbb{R}^d \rightarrow \mathbb{R} \]

The sequence \( S = (F_{i_1}, ..., F_{i_n}) \) for a set of local features \( I = \{F_0, ..., F_m\} \) is defined as a permutation \((i_0, ..., i_m)\), where \( i_j \in \{0, ..., m\} \) and
Music Control in an Interactive Conducting System Using Kinect
Yi-Shin Chen, Leng-Wee Toh and Yi-Lan Liu (2013). International Journal of Multimedia Data Engineering and Management (pp. 35-57).
www.igi-global.com/article/music-control-in-an-interactive-conducting-system-using-kinect/103010?camid=4v1a