
The topic of the doctoral thesis of Mgr. Juraj Mosko focuses on a very important and challenging topic, that lies in the interdisciplinary field of multimedia retrieval and human-computer interaction, namely multimedia content exploration. These days, where multimedia data is used ubiquitously in many different areas (e.g., news, entertainment, surveillance, life-logging, health and medical care, product manual/instructions, etc.), it is increasingly challenging to retrieve desired information from multimedia collections that grow on a daily basis. One of the problems - and this is very nicely and clearly motivated by Mgr. Juraj Mosko in this thesis - is the insufficient support of traditional multimedia retrieval tools for content inspection that goes beyond the query-by-example approach. In practice we often have no good example of an image or video we want to find in a database we typically don't know yet, but still we want to explore its content and navigate through its structure, for example in a similarity arrangement that is presented on a mobile device with intuitive interaction. Instead of a traditional multimedia retrieval system with inconvenient querying features, a retrieval engine that acts like a ‘black box’ and finally presents a one-dimensional presentation of ranked retrieval results with low support for interactivity, it is often much more successful to put the user into the center and build an exploration system around him. In order to find specific information, or simply to investigate image and video content, users often focus on subtle details of some content aspects, which a static retrieval engine simply cannot do. In order to make this work, users need a flexible/powerful but still intuitive/easy-to-use and (probably most important) responsive system with flexible features to browse the content structure and explore the multimedia data at high speed (i.e., with high efficiency).

In his doctoral thesis, Mgr. Juraj Mosko targets exactly that challenging field and proposes an architecture for a multimedia exploration system that is designed for exploring large-scale multimedia data in real-time. For that purpose, he designed a real-time exploration system with several novel components (such as a query analyzer, a scheduler, and timely-limited similarity queries, as well as a new interface for multimedia exploration that build on these components). He proposes an interesting multi-layer model for multimedia exploration, implements a prototype system and finally evaluates his approach in a large user study.
Mgr. Juraj Mosko starts with a good introduction to the topic (Chapter 1), where he explains the shortcomings of the query-by-example approach that is typically used with traditional multimedia retrieval systems. He describes the idea of multimedia exploration as a powerful alternative, and introduces four requirements for it: (1) a simple and intuitive user interface, (2) an initial set of multimedia items to start with the exploration, (3) continuous interaction with the system to constantly refine retrieved items, and (4) high responsiveness that allows a user to efficiently perform similarity queries.

In Chapter 2, which covers more than 80 references (!), Mr. Mosko gives an excellent overview of existing work in the field of multimedia exploration and retrieval. After a summary of the shortcomings of the query-by-example approach, he discusses many exploration systems that were proposed in the literature and are based on visualization and mapping approaches. Moreover, he surveys user interfaces that were developed for image and video exploration, with several examples/illustration. He also provides an interesting discussion about multimedia indexing and describes metric indexing, non-metric indexing, and approximation search. The chapter is closed by an overview of real-time queries, that are the basis for a highly responsive exploration system.

The novel ideas for a new multimedia exploration system that supports these real-time queries are described in Chapter 3, where Mr. Mosko introduces a sound architecture for such a system. It consists of several layers (presentation layer, logic layer, data layer, and communication layer), and contains interesting components that were not considered yet in the community for the sake of multimedia exploration. The logic layer is particularly interesting because it contains several components that can improve the responsiveness of an exploration system: a query analyzer, a scheduler, and a dispatcher. It closely communicates with the data layer, where similarity queries are performed. Another interesting part is the user interface where zooming and panning actions are used to navigate through a similarity arrangement of images. The most interesting and novel part of the proposed architecture, however, is the invention of instant similarity queries for the purpose of multimedia exploration. Mr. Mosko identifies monotonically approximative results as an important requirement to allow early termination of similarity queries (the basis for real-time operation). For that purpose, he discusses the metric access methods (1) sequential scan, (2) pivot tables, and (3) the M-tree/PM-tree and conducts a detailed evaluation of the effect of early termination of similarity queries. The evaluation is performed on two large databases with MPEG-7 descriptors (scalable color and color structure) and different similarity distance functions. The results show that the PM-tree works best for early termination (in terms of error) with the Euclidian distance, while pivot tables work better when the SQFD distance is used.

Chapter 4 provides an interesting discussion about exploration structures and related requirements for interactive use, i.e., required navigation actions (zoom-in, zoom-out, pan) for hierarchical exploration, as typically used when exploring large multimedia data sets. This is followed by a description of the demo application for image exploration, which is available online, as well as a discussion of how to enable indexing such multi-layer exploration structures.

In Chapter 5, Mgr. Juraj Mosko finally presents results from an impressively large user study with 94 participants from different countries, who performed a total number of 1661 search tasks with the goal 'to find as most images as possible of a specific topic in a large data set' (almost 22,000 images). The results are very interesting and show that already with a few exploration actions (e.g., five in average) the users were able to find the first item of the topic (which then allowed to continue more easily by using similarity search). While of course not the complete set of items from a topic could be found by the users, the results show that about 20% of them could be retrieved in average, which is a good result for such a retrieval task (MAP of traditional content-based retrieval systems is typically lower). Also, the results show a monotonic (i.e., linear) increase of retrieved objects with the progress of exploration steps.
The doctoral thesis of Mr. Mosko is closed with his conclusions of the whole work, which is presented in Chapter 6.

**Summary**

Mgr. Juraj Mosko shows with his doctoral thesis that he is able to systematically investigate a series of scientific research questions in a challenging field of computer science. He obviously has detailed knowledge about the literature, is able to critically analyze these works and understand their current shortcomings, and apply that knowledge to make new contributions in the field. Further, he shows that he is able to work methodically correct, both in the technical parts and in user experiments.

The presented work (in particular in Chapter 3, Chapter 4, and Chapter 5) makes several novel contributions, that are of clear interest to the multimedia research community. It provides the basis for additional investigations and potentially also practical applications in the field of multimedia exploration. Parts of his work are already published in scientific papers, and it can be expected that several others will follow.

His doctoral thesis is well structured and a real joy to read. It is written in excellent English, and shows that the author is able to compactly but precisely describe detailed technical knowledge in great detail. Moreover, it gives a clear evidence that Mr. Mosko is capable of performing substantial scientific tasks on his own, in a methodically correct way.

Therefore, I suggest the best score, without restriction.

Assoc.-Prof. DI Dr. Klaus Schöffmann