Assessment of the diploma thesis entitled

„Recognition algorithms for image viewers“

submitted by: Matúš Dekánek

This thesis attempts to solve a relatively ambitious goal of automatic tagging of image collections based on previously assigned tags. Machine learning techniques are used for this purpose, where a mapping between visual features detected in the images and the assigned tags is learned from a given collection of tagged images. Subsequently, given an untagged image, the algorithm decides which tags are appropriate for the image based on the previously learned information. The thesis describes experiments that led to the choice of specific technical solutions for the individual building blocks of this general framework (such as image feature extraction, suitable representation of the mapping between the extracted features and the tags, etc.). In addition, a simple program for demonstrating the developed technology is included.

I found the thesis' goal quite ambitious for several reasons, one of the most important being that many of the tags users assign to images are unrelated to the actual content of the image (for example "holidays 2011" may, visually, refer to just about anything). For this reason, I don't find it particularly surprising that the proposed algorithm, in spite of being technically sound, has troubles fulfilling its original purpose. I do not consider this to be the failure of the candidate – instead, I believe the inherent difficulty of the addressed problem is to blame. I would ask the candidate to elaborate on the feasibility of finding a successful solution to the problem during the defense.

Not being an expert in the topic in question, I cannot judge whether or not the chosen techniques were the most appropriate for the problem at hand. Nonetheless, the impression I had from reading the text is that the candidate did a good job at exploring the space of possible approaches. I'd like to ask the candidate to comment on other possible techniques that, in his opinion, could improve the efficacy of the overall solution.

The text of the thesis is well structured and delivers the information in an easy-to-understand manner. The text does contain quite a number of spelling and grammar errors but I do not consider them a significant flaw of the work. Another problem of the text is perhaps a lack of diagrams that would help in understanding the presented material. Improving the graphical execution of the pseudocode blocks would be highly appreciated, too.

In summary, I believe the submitted thesis reports on a solid piece of research and implementation work and recommend it for acceptance.

Praha, 23.4. 2012

Jaroslav Křivánek