Review Statement for Roman Juránek's Doctoral Thesis

"Acceleration of Object Detection using Classifiers"

May 15, 2012

As the reviewer nominated by Brno University of Technology (BUT), I respectfully make the following statement concerning the doctoral thesis of Mr. Roman Juránek, submitted for the fulfillment of the requirements of the PhD degree in Computer Science and Engineering. I consider the following details of the thesis: the position in the research field, originality, and contributions, including also the candidate's publications.

This thesis investigates how to accelerate object detection in digital image processing and analysis. Object detection is one of the most important tasks in understanding details in digital images. The most important object should be detected for further scene understanding. This task consists of two critical matters: accuracy of detection and computation time of detection, usually containing a trade-off between them. Due to many real-world practical applications solutions must be robust, and thus in practice, in many cases real-time solutions. In this thesis, the focus is on research of computation time, considering selected feature extraction and classification methods. One of the main research topics of the thesis is how to robustly divide the classifiers between hardware and software implementations minimizing a proposed cost function. The work is well motivated and the considered literature represents the current state-of-the-art.

As a conclusion, the topic is appropriate to the particular area of dissertation and it is up-to-date from the viewpoint of the present level of knowledge.

The thesis consists of eight chapters. In Chapter 1, the research field is shortly introduced, the focus of the thesis is defined, and the objectives and contributions are given. Also the structure of the thesis is summarized. The exact research questions have not been declared, but however, "the introduction of a technique for the composition of different implementations of object detection in order to enhance its performance with respect to a user defined cost function" has been stated as the main contribution of the thesis. The WaldBoost algorithm was selected as the classifier and Local Binary Patterns (LBP) and Local Rank Functions (LRF) were selected as the image features. Object detection, focusing on the AdaBoost type of machine learning algorithms is described in Chapter 2 and the motivation to select WaldBoost, one of the current hot topic methods, is given, considering implementation issues which justify the research of the thesis. Feature extraction methods are introduced in Chapter 3, and the selection of LBP and LRF for the given object detection problem is explained. These choices are well grounded since LBP is one of the most popular feature extraction methods at moment and LRF has been recently proposed by the candidate's research group, related to LBP. The selected variants of LRF Local Rank Difference (LRD) and Local Rank Pattern (LRP) are described. Chapter 4 focuses on different levels of acceleration of object detection, mainly taking advantage of

algorithmic accelerations and computational architectures, including SIMD, GPU, and FPGA.

Chapter 5 presents methods that use properties of SIMD architectures for the acceleration of an evaluation of a feature response. Implementations of LBP, LRD, and LRP are explained in details. Besides being responsible for the implementations, the candidate has partially contributed to the method development as a member of the research group. Chapter 6 considers classification cost and its minimization, especially the minimization of classification cost of the WaldBoost classifiers through a combination of different run-time implementations of object detection. This is the main contribution of the thesis. The properties of the classifier (here in practice a learning process) and the classification engine (methods and platforms) affect the computation time significantly. The equations for the cost of the computations have been defined and the total cost has been minimized so that the division point between the hardware implementation (FPGA) and the software implementation can be detected. Chapter 7 consists of experiments and discussion on results. Experiments were performed with the publicly available well-known CMU dataset for face detection containing 130 images. 12 classifiers were used for each three feature type. The figures of the results could have been presented better adding more clear information and interpretation. However, the results show that the varying training properties of the classifiers and using different platforms the proposed cost minimization model works in real-world experiments. Besides the theoretical work, the candidate has done comprehensive work in practical implementations. Conclusions including discussion on future work are given in Section 8.

Based on the considered matters, the work is original and contains a sufficient contribution to the area.

The candidate has published four journal articles and eight conference papers. Two of the journals contain an impact factor and most of the conferences are well-known international conferences. The candidate has written many joint publications with other researchers, and thus, the candidate has proven to be able to co-operate efficiently with other scientists. According to Harzing's Publish or Perish, the candidate's h-index is 4 and there are 42 citations to the candidate's research work. I encourage the candidate to publish more in high-level international journals.

As a conclusion, the doctoral thesis has been published at an appropriate level and the candidate has published actively.

Besides the clear merits of the thesis, there are also shortcomings, especially in the presentation. The research work itself is well motivated, the plan of the storyline of the thesis is good, and the candidate has followed in his writing the storyline very well. However, there are quite many fundamental problems in the presentation in the thesis, mainly presentation mistakes, but also mistyping and misspelling. The candidate should notice that this kind of careless may undermine the credibility of the thesis. Especially presenting equations is not according to the scientific practice. First, the candidate refers in the middle of a sentence to an equation number (before introducing the equation itself), then later he explains the details of the equation, and only then afterwards he shows the equation, not connected to the previous sentence, but totally as an independent sentence in the middle of text. This style is exactly in the wrong

way which makes these parts of the thesis difficult to read. Moreover, some figures are not mentioned in the text at all, and those figures which are mentioned in the text are presented randomly before or after a paragraph where they are mentioned. This also makes the thesis difficult to read. The references have been written in the list of references using totally random style. For example, CVPR as a publication forum has been written in five different ways. There are also other shortcomings in the presentation. Besides improving the presentation, scientific improvements and challenges could have been discussed more, including comments on how general the proposed solutions and the obtained results are.

The candidate has shown a good understanding of the key issues in the research field. The thesis clearly contains contribution to knowledge in the field of computer science and engineering. There are several references to related work and the problems are considered properly.

Based on the considerations presented in this review statement I conclude that the doctoral thesis meets the requirements of the proceedings leading to the PhD title conferment.

Professor Heikki Kälviäinen
Lappeenranta University of Technology (LUT)
Finland