Review Statement for Michal Hradiš's Doctoral Thesis

"Sharing local information for faster scanning-window object detection"

August 21, 2014

As the reviewer nominated by Brno University of Technology (BUT), I respectfully make the following statement concerning the doctoral thesis of Mr. Michal Hradiš's, 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 robust objection detection for image processing and analysis applications. The goal is to improve existing scanning-window based methods exploiting information shared among neighboring image windows. The focus is on AdaBoost-like classifiers, especially on WaldBoost where computationally expensive training of weak classifiers learns to detect objects of interest. In this approach it is important to understand the nature and effect of selected features in a behavior of a classifier to be trained. The task to be solved is to detect one type of an object, for example, face, where the variability of object representations should be considered highly adaptively and computationally efficiently. This makes the problem to be one of the challenging bottle-neck problems in Computer Vision. Thus, the topic of the thesis is well motivated. The candidate has proposed with his co-workers two algorithms, based on neighborhood suppression and early non-maxima suppression, and compared the proposed algorithms to the corresponding state-of-the-art methods using the known public datasets. The proposed methods have experimentally shown to be computationally efficient with significantly improved speed-precision trade-off.

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 nine chapters. In Chapter 1, the research field is shortly introduced and the focus of the thesis is defined. Instead of defining explicit research questions or objectives, the main contributions have been summarized as follows: (i) the object detection algorithm using neighborhood suppression, (ii) the object detection algorithm using early non-maxima suppression, and (iii) the evaluation of features with WaldBoost. Besides the structure of the thesis, the contributions among the co-workers related to this thesis are discussed. This discussion should have been more detailed since the study has been very collaborative with many joint publications with other researchers, including also many publications where the candidate is not the first author. Moreover, the contributions and the relevant publications could have been connected already here to emphasize the value of the established research although these references appear later in the thesis.

In Chapter 2 detection with boosted classifiers is overviewed. The existing boosted window-based detectors are considered and the AdaBoost learning approach is

algorithmically presented as the basis for the research of this thesis, extending it later to WaldBoost in Chapter 3 where sequential decision making is discussed, including the sequential probability ratio test. The close connection to the previous work in the research area is shown in the both chapters. Chapters 2 and 3 could have been presented as one joint chapter.

Chapter 4 presents an interesting comparison of existing state-of-the-art features used in appearance-based detectors giving comprehensive experimental results with known public datasets. The selected features are as follows: Haar-like features, Local Binary Pattern (LBP), Histograms of Oriented Gradients (HOG), Local Rank Differences (LRD), and Local Rank Patterns (LRP). The features could have been introduced more, especially telling which kinds of variants have been used and giving references in Table 4.1. EHOG is mentioned in Table 4.1, but explained later only. Moreover, LRD and LRP should have been introduced more, especially since they belong to research done by the candidate and his co-workers. Tables 4.1 and 4.4 should have been connected better in terms of the selected features and detectors. Are "Our LBP", "Our LRP", and "Our Haar" clear to a reader? Moreover, the sensitivity of selected parameter values could have been discussed more. It is also important to define the performance measurements of the results clearly as equations, in this chapter and also in the experiments in Chapters 6 and 7. In general, this chapter is useful for other researchers, containing also an overview of important public datasets.

Information sharing in scanning-window detection is discussed in Chapter 5 in a quite compact way. This chapter gives general background for Chapters 6 and 7 were the main contributions of the thesis are presented as the algorithms for neighborhood suppression and early non-maxima suppression. The main goal is to accelerate computation while processing local features and making boosted classification decisions. The algorithm to extend existing appearance-based detectors with an ability to suppress image positions in the neighborhood of the position being currently classified is proposed in Chapter 6. This neighborhood suppression method based on WaldBoost tries to reduce computation time so that it is guaranteed that accuracy is not decreased too much. The algorithm was tested in frontal face detection and in eye detection with the MIT+CMU and GroupPhoto datasets, and was compared to WaldBoost without neighborhood suppression. In Chapter 7 the algorithm based on early non-maxima suppression is suggested, built also on WaldBoost and besides the sequential probability ratio test. The algorithm was tested in face localization using two datasets formed from known scientific datasets, flicker, and Internet search. The content of these datasets could have been presented more clearly. Other new approaches related to the proposed algorithms are discussed, but there are no experimental comparisons. However, the both proposed methods are very promising, being twice faster than WaldBoost at the same detection rates, showing significant contributions to the research field.

The proposed methods and the results of the experiments are discussed in Chapter 8 and conclusions are given in Chapter 9. The objectives are stated to be fulfilled as the obtained contributions. Discussion about the future work is quite limited, mainly stating the approaches could be also applied to different problems and data.

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

The candidate has published 3 journal articles (two with an impact factor) and 14 conference papers, and 16 other scientific publications. Most of 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 8 and there are 179 citations to the candidate's research work. The candidate has published much more actively than the average doctoral student and his publications have attracted by other researchers, generating an exceptionally high h index for a doctoral student.

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

Besides the clear merits of the thesis, there are also presentation shortcomings, including minor mistyping and misspelling. Although the thesis is well structured there are numerous very short paragraphs and unfortunately many of them are including one sentence only. The use of the third level sections would have made the thesis easier to read since now there are only the corresponding unnumbered sections. The list of abbreviations and symbols is missing. This list is more important than to present all figures and tables as lists. The list could, for example, clarify meanings like EHOG (and EHOGS) which was introduced in the text later than EHOG was first time mentioned. It is very unusual to use "I" or "our" like "I propose", "I explore", or "Our LBP" in the scientific text although naturally research work done in a doctoral dissertation is very personal matter. The use of the emphasized font (also called italics) is too extensive. The expressions "some" and "etc." should be avoided. Also the future tense "will" should be avoided when it is meant to say "would" or to use the present tense. For pleasant readability, a figure should be presented after it has been mentioned in the text, not before. Moreover, too long figure and table captions should be avoided. For example, Figure 6.5 contains nine and Table 7.1 eight lines of text. A table caption should be before a table, not after a table. The presentation in the bibliography is not always uniform. Despite these presentation problems, in general, the candidate has written a well structured thesis which is comfortable to read.

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 many references to related work, research problems are considered properly, and there are several scientific papers published based on the results of the thesis.

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