9 August 2017

UNIVERSITY OF SURREY

Faculty of Engineering
and Physical Sciences (FEPS)

Centre for Vision,
Speech and Signal Processing
University of Surrey
Guildford, Surrey, GU2 7XH, UK

Professor Josef Kittler
MA PhD ScD(Camb) FREng

T: +44 (0)1483 689294

j.kittler@surrey.ac.uk
surrey.ac.uk/cvssp

CTU in Prague, Faculty of Electrical Engineering
Office for Science and Research
Ing. Marie Duchoslavová
Technicka 2
166 27 Prague 6
Czech Republic

Report on the PhD thesis of Ing Michal Uricar, entitled
Multi-view facial landmark detection

The thesis is concerned with the problem of automatic landmarking of unconstrained face images. The aim of the thesis is twofold. In the first instance the work is directed to the development of pose invariant facial landmarking method that is robust to various degradations encountered in the wild environment. The second aim is to enhance existing machine learning tools to accelerate the facial landmarking system training process. The ultimate goal is to design an automatic facial landmarking system, which is accurate and operates at speeds allowing video rate processing.

Facial landmark detection is a highly topical subject of research in computer vision and machine learning, as an effective solution to the problem would offer improved performance of face recognition systems and ultimately provide the enabling technology for a broad range of applications in the security and digital economy sectors, and for perceptual interfaces, computer graphics, face image retrieval, to mention just a few. The facial landmarking problem is very challenging, as pose and expression changes impact dramatically on the face appearance. The complexity of the landmarking problem is further aggravated by illumination variations and often by occlusion. All these factors render the annotation problem highly nonlinear.

The problem is introduced in Chapter 1, which also presents standard face databases used by the computer vision community, as well as in the thesis for facial landmarking studies and for landmarking tools evaluation. The review of the related literature presented in Chapter 2 provides a comprehensive background to the research topic. It covers generative model as well as regression approaches to the landmarking problem. The reviewed material is well structured and each group of papers is critically evaluated and their relevance to the task in hand established. The evaluation of the state of the art, and of the recent advances made in the literature motivate the work carried and help to

sharpen the objectives for the thesis. The detailed treatment of the material is a clear hallmark of the candidate's professionalism and the manifestation of his profound understanding of the subject, as well as clear evidence of the breadth of his knowledge in this field of research. The review motivates the choice of Deformable Part Based Model (DPM) as the core of the adopted facial landmarking framework. The landmark detector training problem is formulated as Structured Output learning to be solved by Structured Output Support Vector Machine.

The key contributions of the thesis are presented in Chapters 3 and 4. Chapter 3 gives an introduction to DPM and its multi-view extension. The complexity of optimization of the multi-view detector is mitigated in the thesis by an innovative proposal to detect facial landmarks by a two-stage process, which breaks the learning problem into two simpler sub-problems. The proposed detector exhibits other novel features, such as the adopted representation of the input image by features computed on a pyramid of image patches and the coarse to fine search strategy for landmarks position.

Chapter 4 is devoted to learning the parameters of the proposed multi-view facial landmark detector. The chapter makes interesting contributions to the machine learning methodology, first of all, by developing a principled loss function for the landmarking problem. Other contributions relate to the SO-SVM solver proposed for optimising the defined cost function. In particular, the candidate introduces a novel damping regularisation term to control the dynamics of the parameter vector evolving during iteration to avoid oscillation and to accelerate convergence. The extension of the cutting plane approximation of the empirical risk function to a multiple cutting plane model is also commendable.

Chapter 5 presents the results of an extensive and comprehensive evaluation conducted on standard benchmarking datasets using common protocols. The results validate the proposed methodology. The evaluation is structured so as to demonstrate the effect of the various innovative measures proposed in the thesis. The detector matches or outperforms state of the art algorithms in terms of facial landmarking accuracy, as well as speed of processing. It also discusses failure modes. The experimental results demonstrate the merit of the proposed techniques convincingly.

The thesis is well organised. The material of the thesis, including the underpinning theory, is developed in a self contained and coherent manner. The discussion of each topic is lucid and detailed. The thesis is very well written, with the text supported by helpful graphics. Although there is a separate conclusion chapter (Chapter 6), it would have been beneficial to include a brief conclusion section at the end of each chapter.

In summary, the submitted work is impressive in terms of scope, depth and readability and presentation. The mathematical developments are rigorous, and the methodological improvements extensively validated on diverse benchmarking data sets. The objectives defined at the outset have been achieved. In my opinion the work is independent and it makes a significant contribution to the scientific knowledge. The work has resulted in a number of high quality publications. The thesis is clearly of PhD standard and merits to be passed, and to proceed to the formal defence stage.