

# Research Visit Report — NII, Tokyo, Japan

Michal Uříčář  
Center for Machine Perception  
Czech Technical University in Prague

February 2015

## 1 Introduction

From the 6th of October, 2014 till the 6th of February, 2015 I was visiting the National Institute of Informatics (NII) in Tokyo, Japan. I was working as an internship student in the group of prof. Akihiro Sugimoto.

## 2 Research topic

### Multi-view Facial Landmark Detection

Facial landmark detection is a crucial step of the face recognition pipeline (e.g. identity recognition, gender detection or age estimation), since the correct face alignment has a substantial impact on the overall accuracy of the face recognition system. Apart from the face recognition pipeline, the facial landmark detection can be used also as a pre-processing step to some stand-alone applications like head-pose estimation, 3D face reconstruction, face tracking, facial expression analysis or Human-Computer Interaction. Despite the increasing interest in the facial landmark detection by the computer vision community, there is only a few detectors working on the wide yaw-range angle interval ( $-90^\circ, 90^\circ$ ), e.g. Zhu and Ramanan [1].

During the internship I proposed a novel learning scheme for the multi-view facial landmark detector within a deformable part based model framework. The main difference to [1] is in the objective function — while [1] is minimizing the number of missed faces and the landmark detection accuracy is not taking into account, the proposed method optimizes the landmark accuracy directly. Experimental evaluation showed that our proposed method performs better than [1] in terms of both the precision accuracy and speed. The results were summarized as a submission to the Biometrics in the Wild (BWILD) 2015 workshop which is held in connection with the Face and Gesture (FG) 2015 conference. The improvements of the landmark detection framework, which I did during the internship also allowed me

to participate in the 300 Faces in the Wild challenge for the special issue of the Image and Vision Computing (IMAVIS) Journal.

### 3 Procession of the research

- **October** — studied relevant papers on (multi-view) facial landmark detection, implemented learning of the multi-view facial landmark detector.
- **November** — implemented several speed-up techniques for the facial landmark detection framework, which made the multi-view detector variant learning feasible.
- **December** — worked on the learning and experiments for the BWILD 2015 paper.
- **January** — submitted a paper to BWILD 2015, worked on the 300-W IMAVIS competition implementation and submitted the binaries for the competition.
- **February** — started writing the paper for the 300-W IMAVIS competition.

### 4 Results

- The main research results achieved during the internship were formulated as a submission to BWILD 2015 (Biometrics in the wild workshop held in the conjunction with FG 2015) called Real-time Multi-view Facial Landmark Detector Learned by the Structured Output SVM. In this paper we present a multi-view facial landmark detector based on deformable part models which outperforms current state of the art.
- The progress on the implementation made during NII internship allowed also a submission to 300-W IMAVIS competition.
- Ideas on further research of the topic.
- Invaluable experience with working in a stimulating international environment.

### 5 Presentation

The research topic was presented at the Tokyo University of Agriculture and Technologies (TUAT) as an invited talk for a master course.

## 6 Feedback

NII is a stimulating and friendly international environment both of which make it a good place for research and study. I am very thankful for the opportunity of participating in the NII internship program. I made many professional and personal connections among fellow researchers/interns. Most of all, I would like to thank my supervisor professor Akihiro Sugimoto and Diego Thomas for their patience and kind and helpful advice during my internship.

## References

- [1] Xiangxin Zhu and Deva Ramanan, "Face detection, pose estimation, and landmark localization in the wild." *IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2012*. IEEE, 2012.