



A brief review of dissertation

Leipzig University of Telecommunications (HfTL)
Institute of Communications Engineering
Gustav-Freytag-Str. 43, 04277 Leipzig, Germany

Prof. Dr.-Ing. Oliver Jokisch
Chair of Signal and System Theory

Phone: +49 341 3062-224

Email: jokisch@hft-leipzig.de

Head of Scientific Advisory Council at the
Saxon Center of University Didactics (HDS)
Vice President of German ESSV Association
Member of University Council and Liaison
Professor of Cooperative Study Programs

University of West Bohemia
Doctoral Committee of the Faculty of Applied Sciences
Univerzitní, 301 00 Plzeň 3, Czech Republic

Dissertation “Heterogeneous Face Recognition from Facial Sketches”

Doctoral candidate: **Ivan Gruber**
Date, review location: 29th November 2019, Leipzig/Germany

I have reviewed the submitted dissertation “Heterogeneous Face Recognition from Facial Sketches” by Ivan Gruber, and I want to summarize my review as follows.

The dissertation thesis describes the development steps of a system for heterogeneous face recognition (FR) from facial sketches. At the beginning, available (advanced) state-of-the-art FR methods are analyzed. Afterwards existing methods for enabling a cross-modal bridge in heterogeneous FR tasks are surveyed, which can be utilized to bridge the differences in two different visual modalities. The proposed thesis work is re-combining these methods, and it tries to avoid some of the analyzed method shortcomings in a novel heterogeneous FR system.

Beyond the described own developments and experiments, the dissertation thesis manifests a good overview and extensive analysis of literature and practical work in the FR field.

The overall task of the combination and adaptation of the introduced FR methods and their application on photo-sketch recognition is quite challenging, in particular due to the limited, specialized datasets in this field, including specific peculiarities (e.g. the regional bias of sketch image databases).

The classification experiments show, that the proposed solutions of the candidate, including his proprietary X-Bridge concept, can outperform some existing methods.

On the other hand, the dissertation thesis has some shortcomings: The method investigation (in particular regarding neural network concepts) is too detailed, considering the broad literature in this field, which reduces the focus of the reader on the core parts of the thesis. Furthermore, thesis and synopsis are dealing with some method names and acronyms before introducing or defining them. Some motivations are based on speculations of the candidate (e.g. the photo-realistic performance of GANs and the concluded appropriateness for synthesis-based cross-modal bridges). Selected claims are superficial (e.g. “Biometrics cannot be stolen.”), which does not influence the useful results of the thesis apart from that. The introduced, novel X-Bridge method seems to fail for real-world samples or other sketch styles, which might be an indication for overfitting (to the specific datasets and processing chain in the thesis).

In general, the thesis describes interesting research, contains some novelty, and the clear writing style, including the content structure, the relevant tables and diagrams and attachments can be lauded.

With regard to the Study and Examination Regulations of the University of West Bohemia, I want to conclude in a structured form as follows:

- a) Evaluation of the significance of the doctoral thesis for the given field: **High;**
- b) Statement on the approach to solving problems, methods used, fulfilment of the given objectives: **Scientific assertions are fully justified;**
- c) Results of the thesis and specification of the student's original contribution to the given area of knowledge: **Relevant contribution incl. some novelty;**
- d) Statement on the systematic approach, clarity, appropriateness of form and language: **Content and presentation mostly satisfactory, good form and language;**
- e) Statement on the student's publications: **Average number with good quality,**
- f) Unambiguous statement about the thesis status: **Recommended for defense.**

Hereby, I summarize as follows: The state of the art in (heterogeneous) face recognition systems is sufficiently described, and the bibliography is up-to-date. All results are discussed comprehensively, and the conclusions are widely correct and relevant. This thesis is directly related to seven international and national (peer-reviewed) publications. In addition, the candidate contributed to bi-national projects and to further publications in the area of image recognition.

According to my brief review, the submitted dissertation fulfils the quality requirements for being presented in a public thesis defense.

Please, contact me for further information if necessary.

Yours sincerely

Prof. Dr.-Ing. Oliver Jokisch

Review of Dissertation Thesis
“Heterogeneous Face Recognition from Facial Sketches”

submitted by Ing. Ivan Gruber

at the Faculty of Applied Sciences, University of West Bohemia

The thesis proposes the method for facial photo-sketch recognition based on the combination of generative and discriminative neural networks. The addressed problem is relevant and practically significant, and the proposed solution is based on the advanced state-of-the-art methods. The task of their adaptation for the problem of facial photo-sketch recognition poses a significant challenge. One of the reasons for this is that only small photo-sketch datasets are currently publicly available.

The primary goal of the dissertation thesis is the development of a heterogeneous face recognition system for facial photos and sketches, which outperforms the existing approaches. This goal is decomposed into the following objectives:

1. *Analysis of the existing face recognition methods.* The analysis presented in Chapters 2-8 covers all components of a heterogeneous face recognition system.
2. *Development of a cross-modal bridge.* Chapter 9 describes a novel GAN-based cross-modal bridge named “X-bridge”.
3. *Development of a face recognition system based on the proposed cross-modal bridge.* Chapter 9 presents the proposed pipeline, which is composed of the cross-modal bridge (X-Bridge) and the feature extractor (based on DenseNet architecture).
4. *Experimental assessment of the developed system and its comparison with the existing solutions.* In Chapter 9, the Facial Features Preservation Score (FPPS) metric is proposed. In Chapter 10, FPPS is used to compare different cross-modal bridges.

Thus, it can be stated that the primary goal of the thesis is successfully achieved. A strong point of the Mr. Gruber’s work is that he has considered and implemented all components of the heterogeneous face recognition pipeline and performed quantitative study of the proposed approach. The obtained results reliably prove the superiority of X-Bridge approach over existing cross-modal bridges. On the other hand, it would be interesting to compare the performance of the proposed pipeline with other approaches, e.g. recognition without image translation.

Mostly, the text is written in good and clear English. The description of the approach and experimental results is sufficiently illustrated by appropriate figures. Several insignificant flaws can be found in the text, e.g. the figure captions in the text begin with Russian word “Рисунок” (instead of “Figure”), which is only expected in the Russian version of the synopsis.

One of the weak points of the dissertation thesis is its structure. The major part of the text (Chapters 2-8) represents the survey of the existing methods, and only two chapters (Chapters 9-10) describe the results obtained by Mr. Gruber. In fact, these two chapters share the same description with the synopsis. One may expect a lesser survey part and more detailed description of the author’s novel results.

Ivan Gruber has 10 publications related to his dissertation thesis, three of which are indexed in Scopus. They mostly cover the results presented thesis, although several details concerning the cross-modal bridges are not mentioned in the published papers.

Overall, the thesis describes a novel and interesting approach and generally leaves a good impression. As a reviewer, I certainly recommend this thesis for defense.

Questions:

1. One of the significant problems of the heterogeneous face recognition is the ambiguity problem. For example, multiple different sketches can be drawn for one photo and multiple photos can correspond to a single sketch. How this problem can be addressed in the proposed approach?
2. The experimental comparison of different cross-modal bridges is performed for a single recognition network. Will the results for another recognition network be consistent with the obtained values? What will happen if the recognition network is applied to the sketches directly, without reconstruction? Will it be able to extract usable discriminative features from a sketch image?
3. The dataset used to train the generative network is rather small, which may lead to overfitting. Is there any reason to believe that it is not the case for the conducted experiments?
4. Deep generative and discriminative features are essentially very different. Does the translation process preserve the “discriminative component”, or it is partially eliminated?
5. What FFPS values can be expected for “direct” photo-sketch recognition methods, which do not involve image translation?

Saint Petersburg, Russia

27 November 2019

Andrei Oleinik, Ph.D.

Researcher at Speech Technology Center (STC),

Department of Scientific Development