Review of the PhD dissertation

"Video based Object Retrieval and Recognition using Lightweight Devices" submitted by Metwally Rashad Metwally Omar

András László Majdik

The dissertation addresses the fundamental problem of video-based 3D object retrieval and recognition. The proposed algorithms and methods contribute to the state of the art. The presentation of the different approaches is clear with convincing experimental evaluation and validation.

1. PRESENTATION

The dissertation consists of 7 chapters. The introduction (Chapter 1) describes the motivation and the previous work related to the presented research work. Chapter 2 presents the problem formulation and describes briefly the main theoretical concepts used in the dissertation. Chapter 3 displays the benchmark datasets and the test scenarios used in order to evaluate the main scientific contributions, namely the video-based 3D object retrieval and recognition algorithms presented in Chapter 4-6. Finally, Chapter 7 summarizes the main achievements, the new scientific results, and publications. In general, the manuscript is well organize and clearly written. For the workplace discussion of the thesis I raised a number of 14 questions and remarks in order to clarify certain issues and to improve this work. The author has carefully answered all my questions and remarks, and updated the dissertation accordingly. Hereby, I accept all his answers and adjustments.

2. SCIENTIFIC CONTENT

The main scientific contributions are summarized in the form of three theses. All these contribution contain new findings as well as have important applications in video based object retrieval and recognition using handheld lightweight devices.

Thesis 1:

A novel algorithm design is proposed, implemented, and evaluated using benchmark image datasets for a video-based object retrieval system, where a mobile camera is rotated around a set of known objects. I consider that the main contribution is the Hough framework to evaluate candidate object matches, in order to reliably select the correct one even in case the query image is affected by strong image blur or additive noise. In order

to keep the algorithm lightweight the CEDD compact image descriptors is used and to increase speed the KD-tree indexing is applied. However, the Hough voting of candidate index scheme can be easily applied to any other image descriptor and nearest neighbor search technique. This thesis was presented and discussed at several peer-reviewed international conferences in the field of computer vision. These findings were also part of two international journal paper publications.

Thesis 2:

The algorithm presented in the previous thesis is extended to efficiently integrate the IMU data that is available in all modern mobile devices in use today. It is shown that the same Hough paradigm can be extended to incorporate the orientation-dependent hypothesis for the 3D object recognition system. Consequently, the retrieval performance is significantly increased by extending the independent retrieval lists of possible candidates based on the orientation of the images. This thesis was presented and discussed at a peer-reviewed international conference and two international journal paper. Both journal publication had an impact factor of above two in 2016. One is well known in the field computer vision and pattern recognition, and the other one in the field of multimodal sensor data processing.

Thesis 3:

Another novel algorithm is introduced for 3D object retrieval, using a Hidden Markov Model (HMM) framework, where 2D object views correspond to states, observations are coded by compact image feature descriptors, and the orientation sensor is used to secure temporal inference by estimating transition probabilities between states. The algorithm outperforms the previous ones in them of hit-rate and computational time. This thesis was presented and discussed at a peer-reviewed international conference. I would strongly recommend extending this work and publishing the results in form of an international journal paper.

3. CONCLUSION

The author has numerous publications, in particular 2 international journal papers, 4 conference papers published at international conferences, and 1 conference paper presented at the KÉPAF'17, that is the premier venue and conference of the Hungarian computer vision and image processing research community. In summary, the scientific content, presentation and supporting publications are up to the standards required by a PhD degree. Therefore, **I support the defense of the thesis and agree to award the candidate the PhD degree.**

Budapest, May 14, 2018

András László Majdik

Majdik András