Posudek diplomové práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce	Martin Cífka		
Název práce	6D Pose Estimation of Objects in Images		
Rok odevzdání	2024		
Studijní program	Computer Science	**Studijní obor**	Visual Computing and Game Development

Autor posudku	Josef Šivic	**Role**	Vedoucí práce
Pracoviště	CIIRC ČVUT		

Posudek byl připraven ve spolupráci s Vlado Petríkem (CIIRC ČVUT), který práci spoluškolil.

Text posudku:

Thesis overview. The objective of the thesis is to develop new methods for 6D object pose estimation from images. 6D pose means here 3D translation and 3D rotation with respect to the camera. This is one of the oldest computer vision problems, which has important applications. For example, the knowledge of the accurate position and orientation of objects is important to manipulate the scene in robotics. Estimating 6D pose of objects in images could also enable new applications in computer graphics allowing the user, for example, to manipulate the 3D orientation of the object in the 2D image. The 6D pose estimation problem is, however, still very hard, especially in uncontrolled conditions, where the internal camera calibration (mainly focal length) may not be known or when we do not have the exact CAD model of the object available. These challenging setups are addressed in this thesis.

Content of the thesis and its main contributions. The thesis is divided into an introduction chapter (not numbered), describing the motivation, challenges and the objectives of thesis, and four content chapters. Chapter 1 gives an overview of the main related work for the 6D object pose estimation problem. The main two technical contributions are described in chapters 2 and 3.

Chapter 2 considers the problem of 6D pose estimation without known camera calibration. The thesis builds on the existing approach FocalPose (published at CVPR 2022), which it reviews, and extends the approach with several innovations. First, the thesis derives an improved 6D pose and focal length update rule. It compares the new rule with the rule used in the FocalPose approach. It shows the FocalPose rule is only an approximation to the correct (derived) update and demonstrates experimentally the benefits of the correct update rule. Second, this thesis investigates different distributions of 6D poses and focal lengths for synthesizing training data. It shows that a parametric distribution estimated from the training data works best and experimentally improves the pose estimation performance. Finally, the thesis shows that replacing the model retrieval method with a different approach (taken from prior work) also improves results. All these innovations result in improvements of the measured metrics on three considered datasets by almost 10% on average over the original FocalPose method. The results of Chapter 2 have been submitted to IEEE Transactions on Pattern Analysis and Machine Intelligence as a part of an extended version of the FocalPose method.

Chapter 3 then considers the problem of object detection for 6D object pose estimation, when the CAD model of the object is not known at training time, only at test time. The goal is to produce a 2D bounding box of the object in the input image given the CAD model of the object at test time (but not training time). This is a scenario often occurring in robotics, where CAD models of the manipulated objects are often available, but there is no time to re-train the system. The contribution of the thesis is a set of experiments probing the state-of-the-art method (CNOS) for this problem. The thesis experimentally evaluates the state-of-the-art method in different settings of its parameters and analyzes its strengths and weaknesses. This chapter is contributing towards a submission (still in preparation) to ECCV 2024.

Finally, conclusions and future work are discussed in Chapter 5.

Evaluation. Overall, the student has demonstrated independent work on challenging research problems. The thesis was set up as an open-ended research-oriented project. The student has mastered the state-of-the-art in 6D object pose estimation and designed, implemented and validated several innovative improvements of the existing 6D pose estimation method (Chapter 2) as well as experimentally probed another state-of-the-art approach (Chapter 3). From this perspective, this work was more akin to the level of difficulty of the first year of Phd research. The thesis is well written and the writing has required relatively minor feedback mainly on the structural side. The obtained results reach (in Chapter 3) and outperform (in Chapter 2) the current state-of-the-art in this area. The work from Chapter 2 has been submitted to IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI) as a part of an extended version of the FocalPose approach. PAMI is one of the top two journals in computer vision (together with IJCV) with an impact factor of 24.3. The work from Chapter 3 is contributing towards a submission (still in preparation) to ECCV 2024. Such levels of innovation and results are not very common for a MSc thesis.

Práci doporučuji k obhajobě.

Práci nenavrhuji na zvláštní ocenění.

Datum **Podpis**

4.2.2024