New Methods for 3D Object Reconstruction and Camera Auto-Calibration

Ákos Pernek

PhD Theses

Supervisors:
Dr. Levente Hajder (MTA SZTAKI)
Dr. István Vajk (BME AAIT)

Computer and Automation Research Institute
Hungarian Academy of Sciences

Department of Automation and Applied Informatics
Budapest University of Technology and Economics

Budapest, 2015
1 Introduction

Computer vision is an important area of science aiming to entail computers with vision capabilities. These enhanced computers are able to perform tasks that require understanding images: detecting, tracking and identifying objects in an image sequence, understanding the relation between these objects and estimating their position and 3D structure. The applicability of computer vision is thus enormous including tasks such as monitoring and controlling of industrial procedures, aiding object modelling and supporting interaction between the human and the computer. To sum up, computer vision is a complex and challenging field of science with plenty of problems to be solved, continuously attracting the attention of researchers.

A major and probably the most dynamically improving field of this research area is 3D computer vision. This field has gone through a vast evolution in the last years powered by the recent revolution of computer science: rapid performance increase of computers alongside with quick decrease of costs. This revolution boosted the development and spreading of computationally expensive 3D algorithms and this progress is still ongoing. The demand for 3D content is continuously rising. Representative applications of 3D vision can be realistic 3D modelling, supporting cinematic effects or improving the human-computer interaction.

The main topic of the dissertation focuses on 3D structure reconstruction, a common task of 3D computer vision. Although researchers have achieved tremendous progress in the area, there is still place for improvement of 3D methods. They suffer from com-
mon problems such as non-optimal solutions, lack of desired precision, not fully automatic execution and inadequate performance. The introduced theses are focusing on these areas.

The first thesis of the dissertation proposes improvements on the well-known reconstruction technique: structure from motion (SfM). SfM aims to estimate the camera parameters and the 3D structure from a set of 2D features tracked across an image sequence. It is a so-called passive reconstruction technique requiring only photographs as input that can be taken by a regular camera. The advantage of SfM is that it does not require dedicated and (often) expensive equipment such as many of the active methods (for example laser scanning). The proposed improvements focus on SfM reconstruction under special conditions such as (i) reconstruction in the presence of missing data: not all tracked features are visible in all images, (ii) reconstruction of non-rigid objects where flexibility is simulated as a weighted sum of online/offline learnt rigid base objects and (iii) reconstruction of symmetric objects. For this purpose, I propose three novel, efficient, accurate and fast alternation-based optimization methods operating under weak-perspective (or orthographic) projection.

The next contribution of the dissertation is a new camera auto-calibration method for perspective cameras. The task closely relates to the previous thesis: our cameras are an integral part of the reconstruction pipeline and thus they need to be calibrated to obtain an Euclidean reconstruction. The proposed method is an auto-calibration approach which relies on a set of tracked and correlated feature points as opposed to classical calibration algorithms.
which utilize special calibration objects or special properties of the scene. The proposed method converts the camera auto-calibration problem of semi-calibrated cameras (all intrinsic parameters but the focal length(s) are known) into the generalized eigenvalue problem class. The provided formalization is compact, efficient, easy to implement, and provides an accurate solution for multi image camera auto-calibration. The method has two variants applicable for fixed and variable focal length cameras.

The third thesis introduces our novel eye corner detector which benefits from 3D head pose information and improves the accuracy of eye corner detection. The method can also support gaze detection. It utilizes the constrained local model (CLM) formalization to track a set of facial features and to obtain 3D head pose estimates needed for our algorithm. Test evaluation shows that the proposed method delivers more accurate eye corner locations than the CLM itself, especially at extreme head poses. The thesis also includes the demonstration of the applicability of our method: we combine the precise eye corners with a facial feature set of CLM and perform 3D structure reconstruction on this set. Results illustrate the capabilities of non-rigid reconstruction and improve the accuracy of eye corner detection and 3D head pose. This allows precise gaze detection. Although we do not evaluate the gaze detection problem within the dissertation, the results are promising and we expect that our achievements can give a greater depth to human-computer interaction (HCI).

The 3D reconstruction pipeline we utilize is displayed in Figure 1. The colored boxes visualize the center of interest in the dissertation.
Figure 1: 3D reconstruction pipeline
Outline of the Dissertation

After the introduction, this dissertation is structured into three main chapters.

In Chapter 2 I introduce the fundamental method of solving the SfM problem and evaluate its limitations. Furthermore I also introduce my three proposed algorithms that overcome these limitations and perform the evaluation of the rigid reconstruction method that handles missing data.

In Chapter 3 I describe the camera calibration problem, explain my proposed camera auto-calibration method, and perform its evaluation on synthetic and real test sequences.

In Chapter 4 I give an overview of important eye corner detector and face tracker methods, then my proposed 3D eye corner detector is introduced. I also show how to apply this method in order to improve the accuracy of 3D head pose estimation, and perform the evaluation of eye corner detection and non-rigid reconstruction.
2 New Scientific Results

The main contributions of the dissertation are summarized in the following three theses.

Thesis 1 - Precise Motion-Based Object Reconstruction

Novel down-hill alternation methods optimizing structure and motion.

1.a Rigid Reconstruction with Missing Data

A novel method is introduced that solves the reconstruction problem under missing data. The method is a down-hill alternation built from a set of closed-form and (in least squares sense) optimal steps. The key idea of the proposed method is that the parameters of the optimization are separated into distinct groups and refined step-by-step. The algorithm minimizes the reprojection error and guarantees convergence.

1.b Non-Rigid Reconstruction with Online Flexible Shape

The method has been extended to cope with non-rigid objects. The flexible shape is modelled as a linear combination of a set of key objects that are learnt during the reconstruction procedure (online learning). The method can be easily improved to handle missing data, but that is beyond the scope of this work.

1.c Non-Rigid Reconstruction with Offline Flexible Shape
The method has been developed further to perform the non-rigid reconstruction based on a statistical shape model learnt from a training database (offline learning). Statistical shape models are efficient representations of flexible objects and are advantageous for optimization. On the other hand, the fact that they need a labelled training database reduces their applicability.

1.d Non-Rigid Reconstruction with Object Symmetry
Online learning has the disadvantage that the key objects introduce ambiguities into the reconstruction process. Here we show that a special formalization of the optimization parameters enables us to implicitly enforce object symmetry. The symmetry constraint is proven to stabilize the non-rigid reconstruction yielding more accurate results.

Thesis 2 - Camera Auto-Calibration as Rectangular Eigenvalue Problem

Camera auto-calibration is transformed into rectangular generalized eigenvalue problem formalization.

2.a Camera Auto-Calibration for Fixed Focal Length
The camera auto-calibration problem for semi-calibrated cameras with fixed focal length (all intrinsic parameters but the focal length are known) is transformed into the rectangular generalized eigenvalue problem (RGEP) class. The method utilizes the fundamental matrices of image pairs to
derive the compact RGEP formula, which is converted into square shape and solved via robust and efficient conventional methods. The initial estimates are then refined via alternation to achieve accurate results.

2.b Camera Auto-Calibration for Variable Focal Lengths
The proposed method is extended to allow changes of the focal length. The main idea of the method is similar to its fixed focal length variant. The algorithm utilizes the fundamental matrices of image pairs and formulates its RGEP equation, which is afterwards solved as explained previously. However, in this case the RGEP formalization has a completely different form, and both the eigenvalues and eigenvectors are important to obtain the focal lengths.

Thesis 3 - 3D Enhanced Eye Corner Detection and Non-Rigid Reconstruction

Eye corner detection is enhanced by 3D feedback.

3.a 3D Enhanced Eye Corner Detection
A novel eye corner detector method is proposed. The main idea of the method is that 3D head pose estimates can increase the accuracy of the eye corner detection. We propose an eye model which is continuously being aligned with the head pose calculated by a 3D constrained local model (CLM). The eye model is utilized to provide precise estimates for eye corner selection features. The approach is
efficient and outperforms the CLM eye corners especially at extreme head poses (large head movements).

3.b Precise 3D Head Pose utilizing Non-Rigid Reconstruction

The 3D enhanced eye corner detector is combined with non-rigid structure reconstruction to achieve precise 3D head pose. We utilize the eye corner estimates of the proposed method to update the facial feature set tracked by the CLM. The updated feature set is the basis of 3D structure and camera motion estimation by symmetric and non-rigid reconstruction. It is shown that the head pose estimates are more accurate than that of CLM.
3 Applications

In this section the proposed methods are demonstrated by a representative set of examples.

First of all, the capabilities of rigid reconstruction under missing data introduced in Thesis 1 are shown. Figures 2 and 3 illustrate the reconstruction of the ’cat’ sequence. Feature points of the ’cat’ were detected using the widely used KLT algorithm and tracked by a correlation-based template matching method. The measurement matrix of the sequence consists of 2290 points and 92 frames. The missing data ratio is 82%.

![Figure 2: Original images of ‘Cat’ sequence.](image1)

![Figure 3: Reconstructed 3D points of ‘Cat’ sequence.](image2)

Next, the application of the camera auto-calibration method in-
roduced in Thesis 2 is shown. Figures 4 and 5 visualize the reconstruction utilizing focal length estimates obtained using the proposed camera auto-calibration method. The Wadham college is a fixed focal length test sequence consisting of 1331 tracked points in 5 images with missing data ratio of 55%. The House model is a paper model constructed by the author. The sequence consists of 268 manually tracked points in 4 images with missing data ratio of 20%.

Figure 4: Reconstruction of the Wadham college
Finally, a few examples illustrate the eye corner detector method introduced in Thesis 3. Figures 6 and 7 display the eye corner detection results of a synthetic and a real test sequence, respectively. The yellow crosses denote the eye corner estimates provided by the proposed 3D enhanced eye corner detector. The green crosses denote the points of the fitted CLM model. The relevant eye regions and the curves of the eyelids are indicated in blue.
Figure 6: Synthetic test sequence

Figure 7: Real test sequence
4 List of the Author’s Publications


