Methods for the Rectification of Imaged Coplanar Repeated Patterns

A dissertation presented to the Faculty of Electrical Engineering at Czech Technical University in Prague in partial fulfillment of the requirements for the Ph.D. degree in study program No. P 2612 - Electrotechnics and Informatics, branch No. 3902V035 - Artificial Intelligence and Biocybernetics, by

James Pritts

Prague, January 2020

Thesis Advisor

Doc. Mgr. Ondřej Chum, Ph.D.

Thesis Co-advisor

RNDr. Zuzana Kúkelová, Ph.D.

Visual Recognition Group
Department of Cybernetics
Faculty of Electrical Engineering
Czech Technical University in Prague
Karlovo náměstí 13, 121 35 Prague 2, Czech Republic
phone: +420 224 357 666
http://cmp.felk.cvut.cz

Abstract

This thesis presents novel, general and automated methods for the detection, rectification, and segmentation of imaged coplanar repeated patterns. The only assumption made of the scene geometry is that repeated scene elements are mapped onto each other by planar rigid transformations. The class of patterns covered is broad and includes nearly all planar man-made repeated patterns.

Novel minimal solvers are used to hypothesize lens undistortion and plane rectification parameters. A stratum of constraints is derived that defines the necessary configurations of coplanar repeats for each successive level of rectification. The methods work on scenes without straight lines and, in general, relax strong assumptions about scene content made by the state of the art. A synthetic fronto-parallel view of an imaged scene plane (equivalently metric rectification) can be estimated with as few a three coplanar repeats from an image taken with a rectilinear lens or with as few as four coplanar repeats from a radially-distorted lens.

The thesis introduces the first minimal solvers that jointly estimate lens undistortion and affine rectification from imaged coplanar repeated texture. Even with imagery from moderately distorted lenses, plane rectification using the pinhole camera model is inaccurate or invalid. The proposed solvers incorporate lens distortion into the camera model and extend accurate rectification to wide-angle imagery, which is now common from consumer cameras. Accurate rectifications on imagery taken with narrow focal lengths to fisheye lenses demonstrate the broad applicability of the proposed solvers.

In addition, a multi-model estimator is proposed to solve the ill-posed problem of jointly segmenting repetitive texture and regressing the rectification. The estimator encodes a discriminative model in an energy functional that captures global interactions between distinct coplanar repeated patterns and scene planes and combines several features that model how planes with coplanar repeats are projected into images. Energy minimization is achieved by alternately solving labeling and regression problems, which correspond to repetitive texture and plane segmentation and scene geometry estimation, respectively.

Keywords: rectification, repeated patterns, minimal problems, radial distortion, minimal solvers, symmetry, local features

Abstrakt

Disertační práce prezentuje nové a obecné metody pro automatickou detekci, narovnání a segmentaci opakujících se rovinných vzorů. Jediný předpoklad o geometrii scény je ten, že opakující se prvky je možné na sebe transformovat pomocí eukleidovské rovinné transformace. Tuto podmínku splňuje široká škála vzorů, do které spadají téměř všechny lidmi vytvořené opakující se rovinné obrazce.

V práci jsou navržený nové algoritmy pro řešení některých minimálních problémů - výpočet parametrů zkreslení objektivu a parametrů narovnání roviny. Jsou odvozeny nezbytné podmínky pro konfigurace koplanárních opakujících se vzorů pro jednotlivé druhy narovnání. Navržené metody nevyžadují silné předpoklady, jako například stávajícími metodami vyžadovaná přítomnost přímek ve scéně. Nalezené parametry umožňují vygenerovat syntetický čelní pohled snímané scény, a to ze třech korespondujících koplanárních oblastí v případě prosté projektivní kamery a ze čtyřech oblastí v případě radiálně zkreslených snímků.

Tato disertační práce prezentuje první algoritmy pro minimální problémy, které z obrazu opakujícího rovinného vzoru počítají současně parametry radiálního zkreslení a parametry afinního narovnání roviny. I v případě malého radiálního zkreslení narovnání roviny za použití prostého projektivního modelu kamery vede k nepřesným či chybným výsledkům. Navržené algoritmy proto začleňují radiální zkreslení do modelu kamery a poskytují přesné odzkreslení roviny i pro širokoúhlé kamery, které jsou v současné době běžné používané. Přesné narovnání roviny z kamer s úzkým zorným polem stejně tak jako širokoúhlých kamer s velkým radiálním zkreslením demonstruje širokou použitelnost navržených metod.

V práci je dále navržena metoda pro souběžný odhad několika modelů, řešící špatně podmíněný problém současné segmentace opakujících se textur a odhadu parametrů narovnání. Navržený estimátor obsahuje diskriminativní model v objektivní funkci, který zachycuje globální vazby mezi různými koplanárními opakujícími se vzory, různými rovinami ve scéně, a jejich projekcemi do obrazů. Funkce je minimalizována střídavým opakovaným řešením problémů značkování a regrese, což odpovídá problémům segmentace rovin s opakující se texturou a odhadování geometrie scény.

Klíčová slova narovnání roviny, opakující se vzory, minimální problémy, radiální zkreselní, systémy polynomiálních rovnic, symetrie

Acknowledgements

James Pritts was supported by the European Regional Development Fund under the project Robotics for Industry 4.0 (reg. no. CZ.02.1.01/0.0/0.0/15_003/0000470) and the grants MSMT LL1303 ERC-CZ, GACR P103/12/2310, and SGS17/185/OHK3/3T/13.

Contents

List of Figures

1 Introduction

Scene-plane recti cation is a fundamental task of computer vision and is a prerequisite for many classic computer-vision tasks. Recti cation restores properties of the scene that make it possible to detect parallel scene lines, measure ratios of areas, angles and ratios of lengths. The task of restoring these properties is a gateway to many sophisticated computer-vision applications. In particular, this thesis proposes methods for the robust estimation of recti cation from imaged coplanar repeated texture. The importance of detecting and modeling imaged repeated scene elements grows with the increasing usage of scene-understanding systems in urban settings, where man-made objects predominate and coplanar repeated structures are common. Methods that jointly detect coplanar repeated texture and estimate imaged scene-plane recti cation serve as powerful tools for scene understanding.

There are several tasks where recti cation is essential and coplanar repetitions are assumed. E.g., single-view 3D reconstruction of buildings and facade parsing relies on comparing metric properties of matched features on the building facade [96]. Inpainting and image beauti cation are symmetry-aware image editing tasks that bene t from planar symmetry labeling, which is performed in a recti ed space [61].

State-of-the-art recti cation methods that use coplanar repeated texture as input ignore lens distortion. However, wide-angle imagery that has signi cant lens distortion is common since consumer photography is now dominated by mobile-phone and GoPro-type cameras. High-accuracy recti cation from wide-angle imagery is not possible with only rectilinear camera models [45, 94]. Lens distortion can be estimated by performing a camera calibration a priori, but a fully automated method for scene-plane recti cation is desirable. Furthermore, in the case of Internet imagery, the camera and its metadata are often unavailable for use with off-line calibration techniques. A primary goal of the proposed methods in this thesis is to extend accurate recti cation to lens-distorted images containing coplanar repeated texture.

Augmented reality applications require calibrated cameras to place virtual objects in the imaged scene, and calibration data may not be readily available for Internet imagery or from mobile phones. Recti cation is a necessary task of single-view auto-calibration methods. The state-of-the-art single-view auto-calibration methods for lens-distorted images rely on the presence of parallel scene lines to estimate vanishing points [93, 94, 3]. Expanding recti cation to lens-distorted images with sparse scene lines but available coplanar repeated scene texture also expands the class of images that can be auto-calibrated.

In particular, the af ne recti cation of a scene plane transforms the camera's principal plane so that it is parallel to the scene plane. This restores the af ne invariants of the imaged scene plane, which include parallelism of lines and translational symmetries [32, 74]. There is only an af ne transformation between the af ne-recti ed imaged scene plane and its real-world counterpart.

Chapters 5 and 6 propose minimal solvers that jointly estimate lens undistortion and affine rectification from local features extracted from repeating coplanar texture. The proposed solvers are the first solvers that can directly affinely rectify from the radially-distorted image of points or regions extracted from coplanar repeated texture. The input to the solvers are intra-image correspondences of local features. Geometrically, the local features are represented by **local affine frames**, that is, by triplets of (semi-) locally measured image points (see Section 3.2).

The solvers can be differentiated by the assumptions made with respect to the configurations of the inputted local features. Chapter 5 introduces solvers that jointly undistort and affinely rectify from the imaged translations and reflections. This feature configuration is shown in Figure 1.1a. Chapter 6 generalizes joint undistortion and affine rectification to work with the images of rigidly-transformed local features. This feature configuration is shown in Figure 1.1b. All of the proposed solvers eliminate the intermediary undistortion step that is required by the state-of-the-art solvers using repeated texture as input. The best solver can be chosen based on the expected scene content or speed requirements of the application. In general, the solvers are fast and robust to image noise, so they work well in robust estimation frameworks like RANSAC [24].

Metric rectification restores the metric invariants of the imaged scene plane, which include length ratios and angles [32, 74]. In general, the removal of the effects of perspective imaging is helpful for understanding the geometry of the scene plane, and the recovery of metric invariants greatly helps with tasks such as detecting symmetries and repeated image content. Metric rectification is used throughout the thesis to synthesize fronto-parallel views of scene planes. Section 2.10.3 introduces linear minimal solvers to estimate either a semi-metric upgrade from the affine-rectified images of glide-reflected coplanar repeated texture (see Figure 1.1a) or a metric upgrade from the affine-rectified images of rigidly-transformed coplanar repeated texture (see Figure 1.1b).

Imaged scene plane rectification is a poorly constrained problem and verifying the restoration of affine or metric invariants is typically insufficient to correctly assign measurements extracted from imaged scene planes to the model that generated it in the multi-plane setting. Furthermore, good feature coverage over large spans of the imaged scene plane is necessary to properly constrain rectification estimation. Chapter 7 proposes an energy functional that combines several features that model how planes with coplanar repeats are projected into images and captures global interactions between different coplanar repeated texture and scene planes. In particular, regularization terms are incorporated that encourage the assignment of measurements to models such that they conform to the expectations of how a physical scene containing planes must look. These scene prior terms benefit rectification estimation by assuring smooth and dense coverage of measurements over contiguous spans of the imaged scene plane. Minimal solvers for rectification, e.g., proposed in Chapters 5 and 6, are easily plugged into the energy minimization framework to provide scene plane proposals. The model proposals are jointly and globally evaluated, which prevents a model's validity from being biased by the order of its evaluation, which is a common problem with greedy methods like sequential RANSAC. These properties of the energy function and minimization proposed in Chapter 7 enable rectification solvers to be used in difficult scenes with multiple planes lacking a dominant plane.

(a) Translations and Reflections

(b) Rigid Transformations

Figure 1.1: Rectifications of Coplanar Repeated Texture. The top row is a scene plane with (a) translated and reflected regions, which is the assumed configuration for the solvers of Chapter 5, and (b) rigidly-transformed regions, which is the assumed configuration for the solvers of Chapter 6. The state-of-the art requires an intermediate undistortion estimation, while the proposed solvers can directly affine-rectify from the distorted image of coplanar repeated texture. Affine-rectifications are metrically upgraded with the linear solvers introduced in Section 2.10.3.

1.1 Contributions

This thesis introduces the first minimal solvers of polynomial systems of equations for single-view geometry. In particular, the thesis introduces several novel methods for rectifying imaged scene planes from coplanar repeated texture.

The thesis derives novel constraints on lens undistortion and scene-plane rectification parameters using different configurations of radially-distorted conjugately-translated and reflected texture. The complexities of the generated solvers are compared with respect to the choice to eliminate particular unknowns from the polynomial system of equations in their derivations. One of the proposed solver variants can jointly undistort and rectify in only 0.5 microseconds.

In addition, the thesis generalizes the problem of lens undistortion and imaged scene-plane rectification to admit imaged rigidly-transformed coplanar repeats. In particular, derivations of constraints on rectification parameters that either directly use the undistorting and rectifying transform or its linearization are given. The solvers are generated with either elementary methods or the Gröbner basis method. A method adapted from [54] is used to sample feasible monomial bases to maximize the numerical stability of the solvers generated with the Gröbner basis method. The constraints derived from the linearized rectifying transform are used to calculate the dense relative change-of-scale due to the imaging of a scene plane, which gives the relative change of scale at any point of the imaged scene plane.

The code repository associated with this thesis https://github.com/prittjam/repeats provides solvers that cover all minimal configurations of these problems.

The thesis proposes several adaptations to the RANSAC robust estimation framework for the problem of rectifying imaged coplanar repeats [24]. In particular, a criterion for pre-empting consensus set construction is introduced for candidate solutions that are generated from measurements that provide redundant constraints. The pre-emptive strategy, called best minimal solution selection, eliminates the need to construct the consensus sets for all but one candidate solution. Best minimal solution selection significantly increases rectification accuracy compared to a strategy of random selection. In addition, a sequential two-stage RANSAC verification strategy is proposed that: (i) verifies that affine-rectified coplanar repeats are consistent with affine invariants, and (ii) metrically-upgraded coplanar repeats respect metric invariants. The combination of tests for scale consistency and congruence greatly increases the precision of covariant regions that are labeled as coplanar repeats and rectification accuracy.

The methods proposed in this thesis extend accurate rectification to a new class of imagery. Accurate rectificatoins of imaged scene planes from challenging wide-angle and fisheye images are achieved using the new solvers.

The thesis proposes and novel energy function for modeling scenes containing imaged coplanar repeated texture. The energy function is designed such that efficient inference can be achieved with state-of-the-art methods in discrete optimization. The global context of the energy function enables accurate scene plane segmentation and rectification of scene containing multiple planes and lacking a dominant plane. Rectifying solvers can easily be integrated into the minimization framework to provide scene plane models. A challenging dataset is introduced that is used for quantitative evaluation against the state of the art.

1.2 Publications

The content of this thesis is based on the material from the following articles published during the time of the PhD candidacy,

- [74] J. Pritts, O. Chum, and J. Matas. Detection, rectification and segmentation of coplanar repeated patterns. In *CVPR*, 2014.

  The article *Detection, Rectification and Segmentation of Coplanar Repeated Patterns* won the Computer Vision Winter Workshop 2014 *Best Presentation Award*.

- [77] J. Pritts, D. Rozumnyi, M. P. Kumar, and O. Chum. Coplanar repeats by energy minimization. In *BMVC*, 2016.

- [75] J. Pritts, Z. Kukelova, V. Larsson, and O. Chum. Radially-distorted conjugate translations. In *CVPR*, 2018.

- [76] J. Pritts, Z. Kukelova, V. Larsson, and O. Chum. Rectification from radially-distorted scales. In *ACCV*, 2018.

  The article *Rectification from Radially-Distorted Scales* won the Saburo Tsuji *Best Paper Award* at the 14th Asian Conference on Computer Vision (ACCV) 2018.

- [80] J. Pritts, Z. Kukelova, V. Larsson, Y. Lochman, and O. Chum. Minimal solvers for rectifying from radially-distorted scales and change of scales, 2019. arXiv `1907.11539` [cs.CV].

  The journal article *Minimal Solvers for Rectifying from Radially-Distorted Scales and Change of Scales* is accepted to The International Journal of Computer Vision (IJCV).

- [79] J. Pritts, Z. Kukelova, V. Larsson, Y. Lochman, and O. Chum. Minimal solvers for rectifying from radially-distorted conjugate translations. In 2019. arXiv `1911.01507` [cs.CV].

  The journal article *Minimal Solvers for Rectifying from Radially-Distorted Conjugate Translations* was submitted for review to IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI).