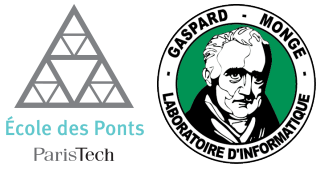
Towards high precision internal camera calibration

Victoria Rudakova



LABORATOIRE D'INFORMATIQUE
GASPARD MONGE

Sous la co-tutelle de :
CNRS
ÉCOLE DES PONTS PARISTECH
ESIEE PARIS
UPEM • UNIVERSITÉ PARIS-EST MARNE-LA-VALLÉE

UNIVERSITÉ
PARIS-EST
MARNE-LA-VALLÉE

École Doctorale Paris-Est
Mathématiques & Sciences et Technologies
de l'Information et de la Communication

Thèse de doctorat
de l'UNIVERSITÉ PARIS EST

Domaine : Informatique

présentée par **Victoria Rudakova**

pour obtenir le grade de
Docteur de l'UNIVERSITÉ PARIS EST

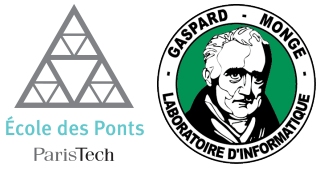
Vers l'étalonnage interne de caméra à haute précision

Soutenue publiquement le XXI janvier 2014 devant le jury composé de :

Rapporteurs :	Andrès Almansa	-	Telecom ParisTech
	David Fofi	-	Université de Bourgogne
Directeur :	Pascal Monasse	-	École des Ponts ParisTech
Examinateurs :	Marc Pierrot-Deseilligny	-	ENSG Paris-Est
Président :	Jean-Michel Morel	-	ENS Cachan

École des Ponts ParisTech
LIGM-IMAGINE
6, Av Blaise Pascal - Cité Descartes
Champs-sur-Marne
77455 Marne-la-Vallée cedex 2
France

Université Paris-Est Marne-la-Vallée
École Doctorale Paris-Est MSTIC
Département Études Doctorales
6, Av Blaise Pascal - Cité Descartes
Champs-sur-Marne
77454 Marne-la-Vallée cedex 2
France



LABORATOIRE D'INFORMATIQUE
GASPARD MONGE

Sous la co-tutelle de :
CNRS
ÉCOLE DES PONTS PARISTECH
ESIEE PARIS
UPEM • UNIVERSITÉ PARIS-EST MARNE-LA-VALLÉE

UNIVERSITÉ
PARIS-EST
MARNE-LA-VALLÉE

École Doctorale Paris-Est
Mathématiques & Sciences et Technologies
de l'Information et de la Communication

PHD thesis
of UNIVERSITÉ PARIS EST

Speciality : Informatics

presented by **Victoria Rudakova**

to obtain the title of

PhD of Science of UNIVERSITÉ PARIS EST

Towards high precision internal camera calibration

Presented on XXI january 2014 before the jury composed of :

Reviewers :	Andrès Almansa	-	Telecom ParisTech
	David Fofi	-	Université de Bourgogne
Advisor :	Pascal Monasse	-	École des Ponts ParisTech
Examinators :	Marc Pierrot-Deseilligny	-	ENSG Paris-Est
President :	Jean-Michel Morel	-	ENS Cachan

École des Ponts ParisTech
LIGM-IMAGINE
6, Av Blaise Pascal - Cité Descartes
Champs-sur-Marne
77455 Marne-la-Vallée cedex 2
France

Université Paris-Est Marne-la-Vallée
École Doctorale Paris-Est MSTIC
Département Études Doctorales
6, Av Blaise Pascal - Cité Descartes
Champs-sur-Marne
77454 Marne-la-Vallée cedex 2
France

Acknowledgements

The thesis was conducted at Ecole des Ponts Paris-TECH, and was funded by the Agence Nationale de la Recherche, Callisto project (ANR-09-CORD-003). I would like to thank ANR for giving me such an opportunity to come and do my studies in France.

I am very grateful to Pascal Monasse who supervised me during these three years, for his explanations, patience and availability. His remarkable scientific and personal qualities helped me a lot in understanding, analysing, and solving the encountered problems during my thesis. I appreciate and feel very grateful for this. Also, I was happy to belong to the IMAGINE lab during last years and enjoyed occurred interactions - whether scientific or not.

I also appreciate the jury members of the thesis: Andrès Almansa, David Fofi, Jean-Michel Morel and Marc Pierrot-Deseilligny. Thank you for being interested in my work, for your comments and questions.

For the constant support I thank my friends and family (in alphabetical order): Anna, Cécile, Natalia Lyubova, Olga, Q.J., Steve P., Yumiko; and I thank all my Parisian friends and people from club de Neuilly-sur-Marne for sharing great time in the city of lights and Ile-de-France region.

Titre Vers l'étalonnage interne de caméra à haute précision

Établissment École des Ponts ParisTech
IMAGINE / CERTIS Nobel B006
6, Av Blaise Pascal - Cité Descartes, Champs-sur-Marne
77455 Marne-la-Vallée cedex 2 - France

Resumé Cette thèse se concentre sur le sujet de l'étalonnage interne de la caméra et, en particuler, sur les aspects de haute précision. On suit et examine deux fils principaux: la correction d'une aberration chromatique de lentille et l'estimation des paramétres intrinsèques de la caméra.

Pour la problème de l'aberration chromatique, on adopte une méthode de post-traitement numérique de l'image, afin de se débarrasser des artéfacts de couleur provoqués par le phénomène de dispersion du système d'objectif de la caméra, ce qui produit une désalignement perceptible des canaux couleur. Dans ce contexte, l'idée principale est de trouver un modéle de correction plus général pour réaligner les canaux de couleur que ce qui est couramment utilisé – différentes variantes du polynôme radial. Celui-ci peut ne pas être suffisamment général pour assurer la correction précise pour tous les types de caméras. En combinaison avec une détection précise des points clés, la correction la plus précise de l'aberration chromatique est obtenue en utilisant un modèle polynomial qui est capable de capter la nature physique du décalage des canaux couleur. Notre détection de points clés donne une précision allant jusqu'à 0,05 pixels, et nos expériences montrent sa grande résistance au bruit et au flou. Notre méthode de correction de l'aberration, par opposition aux logiciels existants, montre une erreur géométrique résiduelle inférieure à 0,1 pixels, ce qui est la limite de la perception de la vision humaine.

En ce qui concerne l'estimation des paramètres intrinsèques de la caméra, la question est de savoir comment éviter la compensation d'erreur résiduelle inhérente aux méthodes globales d'étalonnage, dont le principe fondamental consiste à estimer tous les paramètres de la caméra ensemble – l'ajustement de faisceaux. Détacher les estimations de la distorsion de la caméra et des paramètres intrinséques devient possible lorsque la distorsion est compensée séparément. Cela peut se faire au moyen de la harpe d'étalonnage, récemment développée, qui calcule le champ de distorsion en utilisant la mesure de la rectitude de fils tendus dans différentes orientations. Une autre difficulté, étant donnée une image déjà corrigée de la distorsion, est de savoir comment éliminer un biais perspectif. Ce biais dû à la perspective est présent quand on utilise les centres de cibles circulaires comme points clés, et il s'amplifie avec l'augmentation de l'angle de vue. Afin d'éviter la modélisation de chaque cercle par une fonction conique, nous intégrons plutôt une fonction de transformation affine conique dans la procédure de minimisation pour l'estimation de l'homographie.

Nos expériences montrent que l'élimination séparée de la distorsion et la correction du biais perspectif sont efficaces et plus stables pour l'estimation des paramètres intrinsèques de la caméra que la méthode d'étalonnage globale.

Mots clés Étalonnage interne, matrice de caméra, aberration chromatique latérale, étalonnage de haute précision, correction de la distorsion, biais perspectif, points de contrôle circulaires.

Title Towards high precision internal camera calibration

Institution École des Ponts ParisTech
IMAGINE / CERTIS Nobel B006
6, Av Blaise Pascal - Cité Descartes, Champs-sur-Marne
77455 Marne-la-Vallée cedex 2 - France

Abstract This dissertation focuses on internal camera calibration and, especially, on its high-precision aspects. Two main threads are followed and examined: lens chromatic aberration correction and estimation of camera intrinsic parameters.

For the chromatic aberration problem, we follow a path of digital post-processing of the image in order to get rid of the color artifacts caused by dispersion phenomena of the camera lens system, leading to a noticeable color channels misalignment. In this context, the main idea is to search for a more general correction model to realign color channels than what is commonly used – different variations of radial polynomial. The latter may not be general enough to ensure stable correction for all types of cameras. Combined with an accurate detection of pattern keypoints, the most precise chromatic aberration correction is achieved by using a polynomial model, which is able to capture physical nature of color channels misalignment. Our keypoint detection yields an accuracy up to 0.05 pixels, and our experiments show its high resistance to noise and blur. Our aberration correction method, as opposed to existing software, demonstrates a final geometrical residual error of less than 0.1 pixels, which is at the limit of perception by human vision.

When referring to camera intrinsics calculation, the question is how to avoid residual error compensation which is inherent for global calibration methods, the main principle of which is to estimate all camera parameters simultaneously - the bundle adjustment. Detachment of the lens distortion from camera intrinsics becomes possible when the former is compensated separately, in advance. This can be done by means of the recently developed calibration harp, which captures distortion field by using the straightness measure of tightened strings in different orientations. Another difficulty, given a distortion-compensated calibration image, is how to eliminate a perspective bias. The perspective bias occurs when using centers of circular targets as keypoints, and it gets more amplified with increase of view angle. In order to avoid modelling each circle by a conic function, we rather incorporate conic affine transformation function into the minimization procedure for homography estimation. Our experiments show that separate elimination of distortion and perspective bias is effective and more stable for camera's intrinsics estimation than global calibration method.

Keywords Internal calibration, camera matrix, lateral chromatic aberration, high precision calibration, lens distortion correction, perspective bias, circular con-

trol points.

Contents

Chapter 1

Introduction - version française

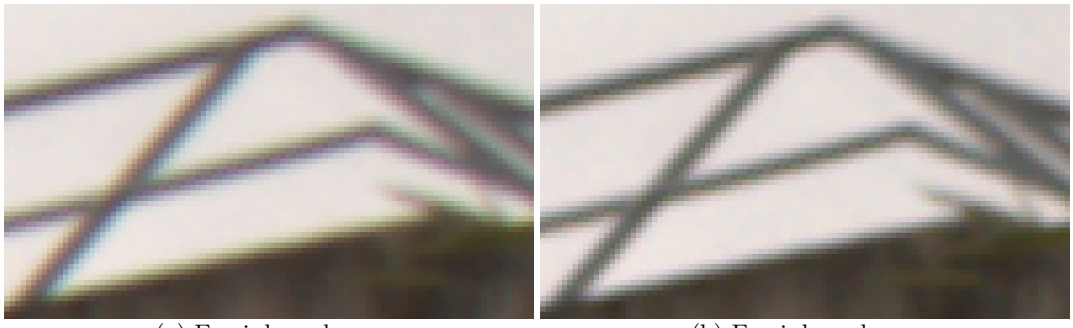
Cette thèse se concentre sur les aspects de précision de l'étalonnage interne de l'appareil photo, et il appartient au projet de recherche CALLISTO (Calibration en vision stéréo par méthodes statistiques) financé par l'ANR (Agence Nationale de la Recherche), dont l'objectif final est la reconstruction des scènes 3D en haute précision. Pour la dissertation deux directions principales d'étalonnage ont été choisis: la correction de l'aberration chromatique et l'extraction de paramétres internes de camèra.

La raison principale pour laquelle nous nous référons à l'étalonnage de l'appareil photo dans le contexte de scènes 3D de haute précision parce qu'elle est la première étape dans une chaîne de reconstruction 3D, et si l'étalonnage n'est pas fait correctement, il va ruiner les étapes suivantes, peu importe le degré de précision ils ont; l'erreur sera propagée, amplifiée ou mélangée avec les erreurs suivantes. En conséquence, cela conduira á un modèle 3D imprécis. Bien qu'il ne semble pas possible d'améliorer directement la précision globale utilisant les données imprécises obtenus, la bonne façon est de se référer à chaque composante séparément et étudier sa précision. Par ailleurs, le calibrage de la caméra doit être fait une fois, jadis les réglages de l'appareil sont fixes.

Lorsqu'on référe au problème d'étalonnage, une question se pose - est-ce que le sujet peut être considéré complet et résolu, ou s'il y a toujours de travail qui peut être fait dans la domain. La question n'est pas simple et dépend de ce qui est considéré comme une contribution à la recherche de valeur et aussi si la solution actuelle satisfait les résultats requis. Par exemple, les méthodes d'étalonnage et de modèles qui étaient valables pour les exigences de précision dernières, ne sont plus satisfaisant pour les nouveaux appareils photo numériques avec une résolution plus élevée, ce qui signifie que le sujet n'est pas entièrement fermé. L'augmentation de résolution du capteur concerne également le problème de l'aberration chromatique. Des tests de perception visuelle ont été effectués afin de voir que les solutions existantes ne sont plus si efficaces.

1.1 Correction de l'aberration chromatique

La première partie de la thèse est consacrée à la méthode précise pour la correction de l'aberration chromatique. En raison de l'évolution plus rapide de la technologie des capteurs par rapport à la technologie optique pour les systèmes d'imagerie, la perte de qualité de résultat qui se produit en raison de l'aberration chromatique latérale est de plus en plus importantes pour l'augmentation de la résolution du



(a) Essai du cadrage, non corrigée

(b) Essai du cadrage, corrigée

Figure 1.1: Recadrée et agrandie dans l'image de la caméra Canon EOS 40D, avant (a) et après (b) la correction de l'aberration chromatique par notre méthode. Remarquez les franges de couleurs atténuées au niveau des bords entre (a) et (b) des images.

capteur. Nous cherchons à trouver de nouvelles façons de surmonter les limites de qualité de l'image obtenue pour des raisons de performances et systèmes de lentilles plus légers.

La principale raison de l'aberration chromatique est le phénomène physique qui s'appelle réfraction. C'est la cause pour laquelle les canaux de couleur se concentrer un peu différemment. En raison de l'aberration, les couches de couleur sont légèrement décalée lorsque l'image numérique est récupérée, ce qui conduit à des franges de couleur au niveau des zones de contraste élevés et les bords de l'image. Pour les applications de haute précision, lorsque l'utilisation d'informations de couleur devient importante, il est nécessaire de corriger ces défauts avec précision. Figure 1.1 montre l'effet de notre correction de l'aberration chromatique sur une image réelle.

D'une manière générale, l'aberration peut être classé en deux types: axiale et latérale. La première se produit lorsque différentes longueurs d'onde se concentrent à différentes distances de la lentille - en images numériques, il produit effet de flou depuis canaux bleu et rouge sont vaporisé (en supposant que le canal vert est au point). Un défaut latérale se produit lorsque les longueurs d'onde se concentrent sur differénts points du plan focal et donc géométriques désalignements de plans couleurs se produisent qui se manifeste sous forme de franges colorées autour des zones à fort contraste, comme il est indiqué sur la Figure 1.1 (a).

Avant de procéder à la tâche, nous cherchons à définir une magnitude de correction de haute précision qui sera notre précision de but. Pour ce faire, une expérience de perception visuelle a été réalisée pour les différents niveaux de désalignement, en unités de pixels. Les tests ont indiqué que 0.1 pixel désalignement est une limite lorsque aberration devient juste perceptible, tandis que désalignements de 0.3 pixel et plus sont tout à fait perceptible.

Comme la plupart des autres approches, nous abordons le problème numériquement, ce qui signifie se référant uniquement à l'aberration chromatique latérale;

l'aide d'un seul coup l'image de modéle de l'étalonnage. La correction est formulé comme un problème de gauchissement d'image, ce que veut réalignement des canaux de couleur numérique. La principale différence avec les solutions existantes, c'est que nous cherchons un modèle de correction plus générale que les différents types de polynômes radiaux habituelles - ils ne peuvent pas également corriger la distorsion pour tous les types de caméras. A son tour, le bidimensionnelle modèle de correction de polynôme choisi, associé à une détection précise des points clés de patron, est capable de saisir la nature physique des canaux de couleurs non alignées, ce qui conduit à une erreur résiduelle géométrique finale inférieure à 0.1 pixels, ce qui est à la limite de perception par la vision humaine.

1.2 Extraction de caméra matrice

Compte tenu de la méthode d'étalonnage global par [ZHANG 2000], théoriquement, on peut prétendre que le calibrage de la caméra est un sujet fermé. Dans le même temps, lors de l'étalonnage d'un appareil photo, la principale difficulté réside dans la distorsion optique, sa correction est une étape nécessaire pour obtenir des résultats de haute précision. L'approche de calibrage global visé mêle les paramètres de distorsion avec les autres paramètres de la caméra et leur calcul est détenu par minimisation simultanée. Toutefois, cela pourrait conduire à une compensation d'erreur résiduelle de paramètres de distorsion et d'autres paramètres de l'appareil photo qui réduiraient la stabilité de l'étalonnage, depuis la physique de champ de distorsion ne seraient pas exploités correctement. En outre, la compensation d'erreur ne peut être éliminé dans le cadre de méthodes globales, et par conséquent, la compensation de distorsion doit être tenue séparément, comme une étape préliminaire à toute nouvelle calibration.

La méthode récemment développée en s'appuyant sur la harpe d'étalonnage par [TANG 2011] permet de calculer un champ de distorsion séparément des autres paramètres. Son idée principale est basée sur la mesure de la rectitude des cordes bien tendues, des images qui doivent être prises dans des orientations différentes. À cet égard, elle se situe dans la catégorie des méthods fil à plomb. Bien qu'il nécessite l'aide d'un modèle de calibrage supplémentaire, le compromis est que nous sommes en mesure de contrôler l'ampleur des erreurs de distorsion résiduelle en plus d'avoir distorsion détaché des autres paramètres de la caméra. Cette séparation devrait également permettre produire des résultats plus fiables, car il résout le problème de la compensation d'erreur résiduelle.

Une autre question que nous abordons, donné un image d'étalonnage compensé de distorsion, est de savoir comment éliminer un biais en perspective. Puisque nous traitons avec des patrons circulaires et des centres d'ellipse comme les points-clés (car il est plus précis que l'utilisation de patrons carrés), les points de contrôle détectés peuvent potentiellement être corrompues par le biais de perspective. Il peut être décrit par la fait que l'image du centre de l'ellipse ne correspond pas au centre de l'image d'ellipse. Par conséquent, nous essayons de compenser le biais

de perspective en tenant compte plutôt cercle-ellipse transformation affine que la transformation du point, puis on utilise la correspondance de points-clés détectés avec des points-clés donné à la patron, donné la transformation affine conique.

Pour utiliser la transformation conique pour le calcul de la matrice d'étalonnage, nous le faisons en intégrant la transformation affine conique dans l'étape de minimisation pour l'estimation de l'homographie. La fonction de transformation est capable de faire correspondre le centre du cercle de la tendance avec le centre de l'ellipse dans l'image. Par conséquent, la fonction de détection principal reste toujours un centre de l'ellipse, il n'est pas nécessaire d'extraire d'ellipse contour. La fonction mentionnée ci-dessus permet d'éliminer la distorsion en perspective, donc, il produit des résultats plus précis pour l'estimation de matrice homographie, et dans le contexte de l'extraction de la matrice de calibrage, il conduit à des résultats plus stables.

1.3 Les chapitres de la thèse

Chapitre 3 montre un choix du patron d'étalonnage qui est représenté par un plan 2D avec des cercles noirs imprimés sur elle, et aussi la façon de détecter les points clés qui sont les centres des cercles. L'objectif est d'estimer avec précision les points-clés et leurs positions relatives par rapport à un modèle donné, laissant le processus entièrement automatique. La méthode de détection des points-cleès s'avère être robuste contre le bruit de l'image et flou, et que les expériences ont montré, la précision de détection reste à 0.05 pixels de la réalité de terrain.

Chapitre 4 montre une méthode robuste pour minimiser l'aberration chromatique latérale, la récupération de la perte de qualité d'image en utilisant une seule image encerclée comme un patron. Différentes séries de tests et de mesures sont utilisées pour valider l'algorithme. Pour l'évaluation de la performance, nous avons utilisé deux données - synthétiques et réelles.

Chapitre 5 couvre le sujet de l'étalonnage précis de la caméra à l'aide des points de contrôle circulaires. Il est obtenu en faisant référence à deux aspects. Tout d'abord, la séparation des paramétres de distorsion de lentille provenant d'autres paramétres de la camèra et le calcul du champ de déformation à l'avance sont effectuées. Deuxièmement, la compensation du biais de perspective qui est susceptible de se produire lors de l'utilisation encerclé modèle est expliqué. Cela se fait en intégrant transformation affine conique dans l'erreur de minimisation lors du calcul de l'homographie, tandis que toutes les autres étapes de calibrage sont laissées telles qu'elles sont utilisées dans la littérature. Nos deux expériences synthétiques et réelles ont montré des résultats plus stables que l'état de l'art - méthode d'étalonnage globale.

Chapitre 6 tire quelques conclusions et expose les avantages et les limites des méthodes utilisées.

1.4 Les contributions principales

- Description détaillée pour la détection de points-clés automatiquement et la commande du patron encerclé qui est précis (moins de 0.05 pixels), même pour les petites rayon du cercle.

- Un algorithme efficace pour corriger l'aberration chromatique latérale robuste à travers la couleur des plans déformation de haute précision (largement sous-pixel) réalignement des canaux de couleur. L'installation ne nécessite qu'une configuration de disques noirs sur papier blanc et un seul cliché. Mesure d'erreur est effectuée en termes de géométrie et de la couleur avec des résultats quantitatifs sur des images réelles. L'examen de l'objectif de précision est fournie en termes de perception visuelle humaine.

- La méthode de calibrage précis de la caméra avec l'aide de points de contrôle circulaires. Le détachement des paramètres de distorsion se fait au moyen de harpe d'étalonnage récemment développé [TANG 2011], puis distorsion d'images compensées sont traités pour l'extraction de caméra matrice. La compensation du biais de perspective est réalisée en intégrant la fonction de transformation conique dans l'estimation de l'homographie.

- Mise en oeuvre de la méthode de correction de la distorsion optique dans la langue C++, ainsi que des améliorations des formules pour des raisons de simplicité et de gain en temps de calcul.

Chapter 2

Introduction

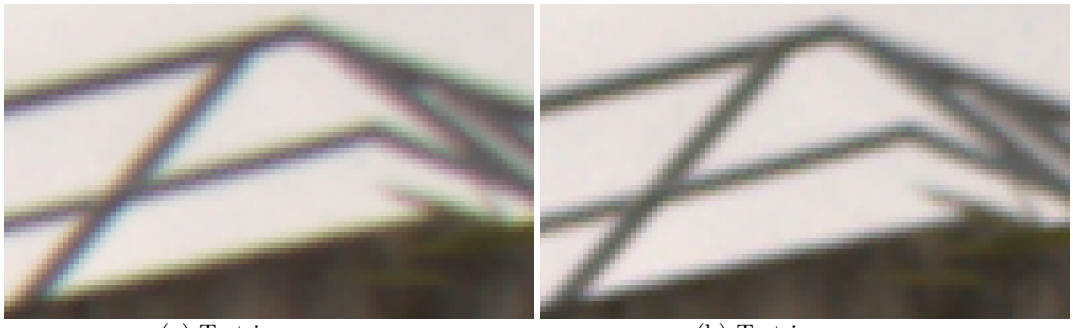
This thesis focuses on precision aspects of internal camera calibration, and it belongs to the research project CALLISTO (Calibration en vision stéréo par méthodes statistiques) funded by ANR (Agence Nationale de la Recherche), whose final aim is to reconstruct 3D scenes with high precision. For the dissertation two main calibration directions are chosen: correction of chromatic aberration and camera internal parameters extraction.

The main reason why we refer to camera calibration in the context of high precision 3D scenes is because it is the first step in a 3D reconstruction chain, and if the calibration is not done accurately, it will ruin the following steps, no matter how accurate they are; the error will be propagated, amplified or mixed with the following errors. As a result, it will lead to an imprecise 3D model. While it does not seem possible to directly improve the overall precision of the obtained imprecise data, the proper way is to refer to each component separately and study its precision. Besides, the camera calibration needs to be done one time, once the camera settings are fixed.

When referring to the calibration problem, one question arises – whether or not the topic can be considered complete and solved, or if there is more work that can be done in the area. The question is not simple and depends on what is considered as a valuable research contribution and also if current solution satisfies the required outcomes. For example, the calibration methods and models that were valid for past precision requirements, are becoming unsatisfying for new digital cameras with higher resolution, which means the topic is not entirely closed. The increasing sensor resolution also concerns the chromatic aberration problem. Visual perceptual tests were performed in order to see that existing solutions are not so effective anymore.

2.1 Chromatic aberration correction

The first part of the thesis is dedicated to the precise method for chromatic aberration correction. Due to the more rapid development of the sensor technology in comparison with the optical technology for imaging systems, the result quality loss that occurs because of the lateral chromatic aberration is becoming more significant for the increased sensor resolution. We aim at finding new ways to overcome resulting image quality limitations for the sake of higher performance and lighter lens systems.



(a) Test image crop, uncorrected

(b) Test image crop, corrected

Figure 2.1: Cropped and zoomed-in image from camera Canon EOS 40D, before (a) and after (b) chromatic aberration correction by our method. Notice the attenuated color fringes at edges between (a) and (b) images.

The main reason of the chromatic aberration is the physical phenomenon of refraction. It is the cause why color channels focus slightly differently. As a result of the aberration, the color channels are slightly misaligned when digital image is retrieved, which leads to color fringes at the high contrast areas and image edges. For high precision applications, when usage of color information becomes important, it is necessary to accurately correct such defects. Figure 2.1 shows the effect of our chromatic aberration correction on a real image.

Generally speaking, the aberration can be classified into two types: axial and lateral. The former occurs when different wavelengths focus at different distances from the lens - in digital images it produces blurring effect since blue and red channels are defocused (assuming the green channel is in focus). A lateral defect occurs when the wavelengths focus at different points on the focal plane and thus geometrical color plane misalignments occur which manifests itself as colorful fringes around high-contrast areas, as it is shown on Figure 2.1 (a).

Before proceeding to the task, we aim to define a magnitude of high precision correction which will be our goal precision. For this, a visual perception experiment was done for different misalignment levels, in pixel units. The tests stated that 0.1 pixel misalignment is a borderline when aberration becomes just-noticeable, while misalignments of 0.3 pixel and higher are quite perceptible.

Like most other approaches, we address the problem digitally, which means referring only to the lateral chromatic aberration; using a single shot of pattern image for calibration. The correction is formulated as an image warping problem, which means re-aligning color channels digitally. The main difference from existing solutions is that we search for a more general correction model than the usual different types of radial polynomials – they cannot equally correct the distortion for all types of cameras. In turn, the chosen bivariate polynomial correction model, combined with an accurate detection of pattern keypoints, is able to capture physical nature of the misaligned color channels, leading to a final geometrical residual error

of less than 0.1 pixels, which is at the limit of perception by human vision.

2.2 Camera matrix extraction

Considering the global calibration method by [ZHANG 2000], theoretically one can claim that camera calibration is a closed topic. At the same time, when calibrating a camera, the major difficulty lies in optical distortion; its correction is a necessary step for high precision results. The mentioned global calibration approach mixes the distortion parameters with other camera parameters and their calculation is held by simultaneous minimization. However, this could potentially lead to residual error compensation of distortion parameters and other camera parameters that would decrease the calibration stability, since the physics of distortion field would not be captured correctly. Moreover, the error compensation cannot be eliminated in the framework of global methods, and therefore, the distortion compensation must be held separately, as a preliminary step to any further calibration.

The recently developed method relying on the calibration harp by [TANG 2011] allows calculating a distortion field separately from other parameters. Its main idea is based on straightness measure of tightly stretched strings, pictures of which must be taken in different orientations. In that respect, it lies in the category of plumb-line methdods. While it requires using an additional calibration pattern, the trade-off is that we are able to control the residual distortion error magnitude in addition to having distortion detached from other camera parameters. This separation should also allow producing more reliable results since it solves the problem of residual error compensation.

Another questions we address, given distortion compensated calibration image, is how to eliminate a perspective bias. Since we deal with circular patterns and ellipse centers as keypoints (as it is more precise than using square patterns), the detected control points can potentially be corrupted by perspective bias. It can be described by fact that image of the ellipse center does not correspond to the center of the ellipse image. Therefore, we try to compensate for the perspective bias by taking into account rather circle-ellipse affine transformation than point transformation and then use correspondence of detected keypoints with pattern keypoints given the affine conic transform.

In order to use the conic transform for the calibration matrix calculation, we do it by incorporating the conic affine transformation into the minimization step for homography estimation. The transformation function is able to match center of the circle of the pattern with center of the ellipse in the image. Therefore, the main detection feature still remains an ellipse center, there is no need for ellipse contour extraction. The aforementioned function allows eliminating the perspective bias, thus, it produces more precise results for homography matrix estimation, and in the context of the calibration matrix extraction it leads to more stable results.

2.3 The thesis chapter by chapter

Chapter 3 shows a choice of the calibration pattern which is represented by a 2D plane with printed black circles on it, and also how to detect the keypoints which are the circles' centers. The aim is to accurately estimate those keypoints and their relative positions with respect to a given pattern, leaving the process fully automatic. The method for keypoint detection is shown to be robust against image noise and blur, and, as experiments showed, the detection precision stays within 0.05 pixels from the groundtruth.

Chapter 4 demonstrates a robust method to minimize the lateral chromatic aberration, recovering the loss of image quality by using a single circled pattern image. Different series of tests and measures are used to validate the algorithm. For the performance evaluation, we used both synthetic and real data.

Chapter 5 covers the topic of the precise camera calibration using circular control points. It is achieved by referring to two aspects. First, separation of the lens distortion parameters from other camera parameters and calculation of the distortion field in advance are done. Second, the compensation for perspective bias that is prone to occur when using circled pattern is explained. This is done by incorporating conic affine transformation into minimization error when calculating the homography, while all the other calibration steps are left as they are used in the literature. Both our synthetic and real experiments demonstrated more stable results than state-of-art global calibration method.

Chapter 6 draws some conclusions and exposes advantages and limitations of the used methods.

Appendix A is mainly based on the work of [TANG 2011] and is a preliminary part of calibration process described in Chapter 5. Some minor alterations were incorporated and exposed in order to improve computational results.

2.4 Main contributions

- Detailed description for automatic keypoint detection and ordering of the circled pattern which is precise (less than 0.05 pixels) even for small circle radius.
- An effective algorithm to robustly correct lateral chromatic aberration through color planes warping of high precision (largely sub-pixel) realignment of color channels. The setup requires only a pattern of black discs on white paper and a single snapshot. Error measurement is performed in both geometry and color terms with quantitative results on real images. The examination of the precision goal is provided in terms of human visual perception.

- The precise camera calibration method with using circular control points. The detachment of distortion parameters is done by means of recently developed calibration harp [TANG 2011], and then distortion-compensated images are treated for camera matrix extraction. The compensation for the perspective bias is carried out by incorporating the conic transform function into homography estimation.
- Implementation of the optical distortion correction method in C++, as well as improvements of the formulas for the sake of simplicity and gain in computational time.

Chapter 3

Robust and precise feature detection of a pattern plane

The aim of the chapter is to accurately estimate the keypoints from an image and their relative positions with respect to a given pattern. The calibration pattern is represented by a $2D$ plane with black circles printed on it. The process is fully automatic and is robust against image noise and aliasing, leaving the detected keypoints at deviation of average 0.05 pixels from the groundtruth.

Keywords. Precise keypoints, feature detection, pattern plane, circle center, ellipse center, keypoint ordering.

Contents

3.1 Introduction

In the context of high precision camera calibration, we are interested in precise allocation and detection of the keypoints which would ensure dense and consistent field registrations, as well as robustness to noise and aliasing. The notion *high precision* often means the residual error between the camera and its obtained numerical model is far smaller than a pixel size. For example, a calibration of lateral chromatic aberrations requires a correction model, residual of which is to stay within 0.1 pixels in order not to be visually perceptible (more details on this experiment are given in the Chapter 4), therefore, our main goal will be to detect the keypoints with deviation no more than 0.1 pixels from the groundtruth.

One of the most common types of keypoints are feature based interest points. Such local image descriptors do not require any type of calibration pattern, and they have quite a broad range of applications – from object recognition [LOWE 2004] to image retrieval [NISTER 2006, SIVIC 2006], and similar. The most famous local feature extractor is Scale-Invariant Feature Transform (SIFT) algorithm [LOWE 1999], further developed into [LOWE 2004]. For the mentioned applications the precision of spatial position may appear less important. Often the relative spatial layout of interest points are used together with a tolerance for large variations in the corresponding points relative position [SIVIC 2005].

An alternative to feature based interest points would be to pick the interest points at random, but it will be unlikely to obtain precise spatial correspondence between a sparse set of randomly picked points. The ability to detect corresponding interest points, in a precise and repeatable manner, is a desirable property for obtaining geometric scene structure. Therefore, when it concerns applications of $3D$ reconstruction and camera calibration from interest points, it is of high importance to have precise point correspondence [SNAVELY 2008, TORR 2000, FURUKAWA 2010] which assumes using some kind of calibration pattern to ensure spatial consistency.

Different types of planar charts exist for the purpose of camera calibration as sources of both $2D$ and $3D$ control points. Normally, these points are constructed on a planar surface by creating some high contrast pattern. The pattern also facilitates the recovery of the control points projections on the image plane. The most common pattern are: squares [WENG 1992, ZHANG 2000], chekerboards [LUCCHESE 2002], circles [ASARI 1999, HEIKKILÄ 2000, KANNALA 2006]. Those became popular as they can be always manufactured to a sufficient precision, and their data points are recoverable through the use of standard image processing techniques.

When choosing a plane calibration pattern, it is important to consider an aspect for invariance to the potential bias from projective transformations and nonlinear distortions. [MALLON 2007b] provides a comparative study on the use of planar patterns in the generations of control points for camera calibration. There, a circled pattern is compared to a checkerboard, and it is theoretically and experimentally shown that the former can potentially be affected by bias sources. As a contrast,

appropriate checkerboard pattern detection is shown to be bias free.

At the same time [MALLON 2007a] provides results for sub-pixel detection error of the keypoints which are the intersections of a chessboard pattern. The extraction is done automatically using standard corner detector such as those described by [LUCCHESE 2002, JAIN 1995] with sub-pixel refinement step of saddle points. Those results expose an accuracy magnitude of about 0.1-0.3 pixels, depending on the camera. Such precision result would not be sufficient for high precision calibration and could be potentially improved if we utilise higher precision detector of the circled pattern with the condition of compensation for distortion (see Appendix A) and perspective bias (refer to Chapter 5) beforehand.

Under perspective transformation circles are observed as ellipses, therefore our main interest lies into ellipse center detection. One of the common ways to detect ellipse is through Hough transform [PRINCEN 1992] - it is based on voting system for some ellipse parameters using contribution of contour pixels; as an example, [ÁLVAREZ LEÓN 2007] detect ellipses using Hough transform with parameter space reduction and a final Least-Square Minimization refinement. While Hough transform is a good tool in applications like pattern recognition, it may not be sufficient since it has its limitations like dependence on the results from edge detector and might be less precise in noisy and blurry environments.

Other types of estimation of an ellipse rely on accurate extraction of the contour points with subpixel precision and then fitting ellipse parameters on the obtained set of points. Numerous methods exist for fitting ellipse parameters from a given set of contour points [GANDER 1994, KANATANI 1994, FITZGIBBON 1995, CABRERA 1996, FITZGIBBON 1999, KANATANI 2006, KANATANI 2011]. They differ from each other depending on their precision, accuracy, robustness to outliers, etc. All those rely on a set of contour points that are extracted beforehand. Nevertheless, extracting contour points usually subsumes multiple stages including gradient estimation, non-maximum suppression, thresholding, and subpixel estimation. Extracting contour points imposes making a decision for each of those points based on neighbourhood pixels in the image. The processing of low contrast images would be quite challenging where each contour point can hardly be extracted along the ellipse, therefore, it is better to refer to the information encoded in all pixels in the ellipse surrounding. By eliminating the participation of contour points, the method would be greatly simplified and the uncertainty on the recovered ellipse parameters will be assessed more closely to the image data.

The current chapter presents a method for high precision keypoint detection for the purpose of camera calibration which takes into account both geometrical and color information. The method is based on defining intensity and affine parameters that describe an ellipse, followed by minimization of those parameters so as to fit the observed image of the ellipse and its surrounding pixels. No contour information is necessary. It allows a detection of maximum accuracy 0.05 pixels, and, as experiments show, it is resistant to noise and aliasing.

The rest of chapter is organized as follows: Section 3.2 gives a description of

the method, Section 3.3 demonstrates a simple way how the ordering of keypoints was performed, which is a necessary step for automatic camera matrix calibration. Finally Section 3.4 includes synthetic experiments for detection precision against noise and aliasing.

3.2 Sub-pixel keypoint detection

The calibration pattern is represented by disks and their keypoints are the centers. A *disk* means a circle (or ellipse) filled with black color. Therefore, our interest lies in the precise center detection of a disk, which has an elliptic form considering a camera tilt with respect to the pattern plane normal. The precise ellipse center detection is obtained by an adjustment of a *parametric model* simulating a CCD response using an iterative optimization process. The parametric model takes into account both color (intensity) and geometry aspects.

3.2.1 Geometrical model

A general affine transformation A:

$$X' = AX, \tag{3.1}$$

which describes the relationship between model point $X = (x, y)$ of the circular patch and the image point $X' = (x', y')$ of the corresponding elliptic patch can be written as follows:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} l_1 \cos\theta & l_2 \sin\theta & t_u \\ -l_1 \sin\theta & l_2 \cos\theta & t_v \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{3.2}$$

In order to avoid color interpolation of the pixels on CCD matrix, the inverse transformation A^{-1}:

$$X = A^{-1}X' \tag{3.3}$$

is introduced since it allows to obtain a model point X having an image point X' as it can be seen on Figure 3.1. For an elliptic patch with (t_u, t_v) representing its sub-pixel center position, h_1, h_2 - elongation factors of the major and minor axes and θ - an angle between the major axis and abscissa, the inverse transform is formulated as:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} h_1 \cos\theta & -h_2 \sin\theta & 0 \\ h_1 \sin\theta & h_2 \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x' - t_u \\ y' - t_v \\ 1 \end{bmatrix} \tag{3.4}$$

Therefore, there are five parameters which describe the model geometrically and which will be used in the minimization process:

- h_1 and h_2 are the elongation axes of the ellipse,
- θ is the inclination angle,

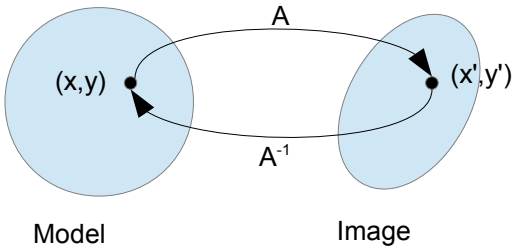


Figure 3.1: Affine transformation A and its inverse A^{-1} for model point (x, y) and its corresponding image point (x', y').

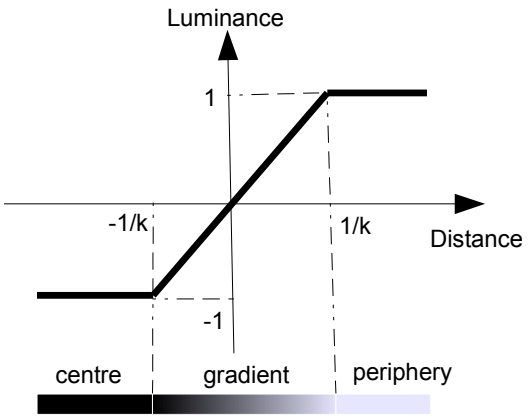


Figure 3.2: The luminance transition model of parameter k.

- t_u and t_v are the coordinates of the ellipse centers; these are the two parameters that represent keypoint coordinates.

3.2.2 Intensity model

The model assumes constant intensity in the disk center and in the periphery with a linear transition between both. For the given luminance levels L_1 at the center of the patch and L_2 at its periphery, assuming the patch is darker than background $L_1 < L_2$, the luminance transition is represented by three line segments as shown on Figure 3.2 with the gradient part being linear with slope k. The distances $-\frac{1}{k}$ and $\frac{1}{k}$ define the border of the center, gradient and periphery areas.

For each pixel point (x', y') there is a model point (x, y) obtained by (3.4) which lies at distance $d_c = \sqrt{x^2 + y^2}$ from model's center for a circle with normalized radius 1; if we place the origin at distance 1 from circle center as displayed on Figure 3.2, the model point distance will be defined as $d = d_c - 1$. The model point has its corresponding normalized luminance level $\tilde{L}(x', y')$ (normalized on the

interval $[\tilde{L}_1 = -1, \tilde{L}_2 = 1]$) which is assigned depending on the interval:

$$\tilde{L}(x', y') = \begin{cases} -1, & d \leq -\frac{1}{k} \\ kd, & -\frac{1}{k} < d < \frac{1}{k} \\ 1, & d \geq \frac{1}{k} \end{cases} \tag{3.5}$$

The denormalization of $\tilde{L}(x', y')$ is to be done:

$$L(x', y') = L_1 + \frac{\tilde{L}(x', y') + 1}{2}(L_2 - L_1) \tag{3.6}$$

Therefore, there are three parameters which describe the color model and which will be used in the minimization process:

- L_1 - luminance level at the center area of the patch
- L_2 - luminance level at the periphery
- k - slope which defines the transition from the center to periphery areas

3.2.3 Parametric model estimation through minimization

Levenberg-Marquardt algorithm (LMA) is chosen to minimize the sum of squared differences of the gray levels between each pixel (x', y') of the elliptic patch in the image I and corresponding point (x, y) obtained by (3.4) of the theoretical CCD model with intensity L as in (3.6). The model is represented by a set of parameters $\{h_1, h_2, \theta, t_u, t_v, k, L_1, L_2\}$ that comprises both geometry and color properties. The following distance function is to be minimized by the LMA:

$$\underset{h_1,h_2,\theta,t_u,t_v,k,L_1,L_2}{\arg\min} \sum_{x'}^{w} \sum_{y'}^{h} (I(x', y') - L(x', y'))^2. \tag{3.7}$$

After the minimization process is terminated, among the set of obtained model parameters, there is a sub-pixel center coordinate (t_u, t_v) which represents a key-point.

3.2.3.1 Parameter initialization

Given three channels of the image, the very first step is to estimate initial position of each disk and the size of its enclosing sub-image. This is done by proceeding:

Step 1. Binarization of each channel.

Step 2. Finding connected components of a black color for each binarized channel.

Step 3. Calculating moments for each component i:

(a) radii r_{i_x}, r_{i_y},

(b) compactness measure $C_i = \frac{4\pi S_i}{P_i^2}$, where P_i is a closed curve of the connected component area S_i,

(c) centroid (t_{i_u}, t_{i_v}).

Step 4. Noise elimination by meeting the conditions:

$$\begin{cases} \min(r_{i_x}, r_{i_y}) > 8 \\ 1 - \delta \leq C_i \leq 1 + \delta, \delta \in [0.2, 0.4] \\ S_i \in \{S\}_{freq}, \end{cases} \tag{3.8}$$

where $\{S\}_{freq}$ is a histogram area with most frequent connected component sizes

As a result we obtain initial positions of each disk (t_{i_u}, t_{i_v}) and its enclosing sub-image with size $w_i = h_i = r_i \frac{5}{2}$, where $r_i = \frac{1}{2}(r_{i_x} + r_{i_y})$.

The initialization of other geometric parameters h_1, h_2 and θ is done with help of principle component analysis. If we represent an ellipse by its covariance matrix

$$Cov = \begin{bmatrix} (\sigma_x)^2 & \sigma_{xy} \\ \sigma_{xy} & (\sigma_y)^2 \end{bmatrix}, \tag{3.9}$$

where σ_x - one-sigma uncertainty in x direction, σ_y - in y direction and σ_{xy} - covariance between x and y. The axes and their lengths are represented by eigenvectors and eigenvalues accordingly. We are interested in eigenvalues in order to initialize h_1 and h_2. Given matrix Cov, a characteristic equation can be written:

$$|Cov - \lambda I| = \left| \begin{bmatrix} (\sigma_x)^2 & \sigma_{xy} \\ \sigma_{xy} & (\sigma_y)^2 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right| = 0. \tag{3.10}$$

The determinant calculation will lead to a quadratic equation:

$$\lambda^2 - ((\sigma_x)^2 + (\sigma_y)^2)\lambda + (\sigma_x)^2(\sigma_y)^2 - (\sigma_{xy})^2 = 0, \tag{3.11}$$

which has roots

$$\lambda_{1,2} = \frac{(\sigma_x)^2 + (\sigma_y)^2 \pm \sqrt{((\sigma_x)^2 + (\sigma_y)^2)^2 - 4((\sigma_x\sigma_y)^2 - (\sigma_{xy})^2)}}{2} \tag{3.12}$$

The lengths of the ellipse axes are square root of eigenvalues λ_1, λ_2 of covariance matrix Cov and since parameters h_1 and h_2 represent semi-axes, we can initialize them as

$$\begin{aligned} h_1 &= \frac{\sqrt{\lambda_1}}{2} \\ h_2 &= \frac{\sqrt{\lambda_2}}{2} \end{aligned} \tag{3.13}$$

The counter-clockwise rotation θ of the ellipse then can be deduced from the first column of 3.10, which also means $[\cos\theta \sin\theta]^T$, and therefore we can write

$$\theta = \text{atan2}((\sigma_x)^2 - \lambda_1, \sigma_{xy}). \tag{3.14}$$

The other model parameters are initialized: $\{k = 2, L_1 = black^*, L_2 = white^*\}$, where $black^*, white^*$ are global maximum and minimum intensities for the given sub-image $[w_i \times h_i]$.

3.2.3.2 Error

The element of an error vector $E = (e_{(0,0)}, e_{(0,1)}, \cdots, e_{(w,h)})$ for a set of pixels of the image with size $w \times h$ of the elliptic patch is given by:

$$e_{(x',y')} = I(x', y') - L(x', y') \tag{3.15}$$

3.2.3.3 Jacobian matrix

The Jacobian matrix is determined as a matrix of all first-order partial derivatives of the vector function $\{\lambda_1, \lambda_2, \theta, t_u, t_v, k, L_1, L_2\}$ with respect to data vector. The generic formulations for the geometry variables (not including k, L_1, L_2 variables) for a given image pixel (x', y') are:

$$\begin{aligned} \frac{\partial e_{(x',y')}}{\partial \bullet} &= -\frac{\partial L(x', y')}{\partial \bullet} \\ \frac{\partial L(x', y')}{\partial \bullet} &= \tfrac{1}{2}(L_2 - L_1)\frac{\partial \tilde{L}(x', y')}{\partial \bullet} \\ \frac{\partial \tilde{L}(x', y')}{\partial \bullet} &= k\frac{\partial d}{\partial \bullet}, \end{aligned} \tag{3.16}$$

where the formulas of partial derivatives $\dfrac{\partial \tilde{L}(x', y')}{\partial \bullet}$ for each variable are given only for the gradient interval $-\frac{1}{k} < d < \frac{1}{k}$ (since the derivatives will be zeros at periphery and at the center areas, see (3.5)), and further on the derivatives are shown only for this interval.

The derivatives for color variables k, L_1, L_2 are straightforward:

$$\begin{aligned} \frac{\partial e_{(x',y')}}{\partial \bullet} &= -\frac{\partial L(x', y')}{\partial \bullet} \\ \frac{\partial L(x', y')}{\partial k} &= (L_2 - L_1)\frac{d}{2} \\ \frac{\partial L(x', y')}{\partial L_1} &= 1 - \frac{\tilde{L}(x', y') + 1}{2} \\ \frac{\partial L(x', y')}{\partial L_2} &= \frac{\tilde{L}(x', y') + 1}{2} \end{aligned} \tag{3.17}$$

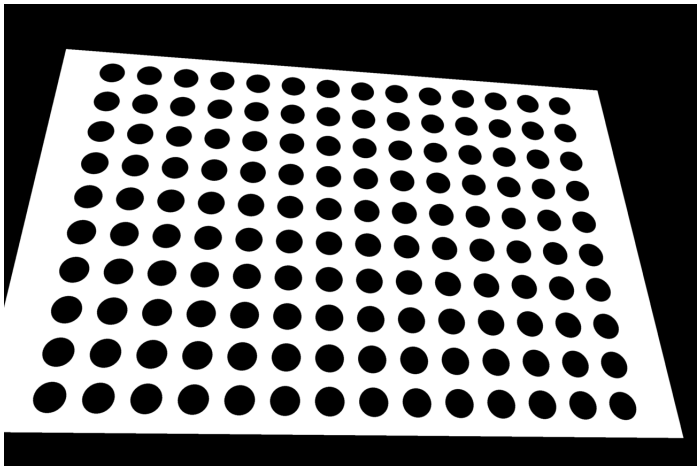


Figure 3.3: Example of the pattern image.

The formulas of partial derivatives $\frac{\partial d}{\partial \bullet}$ for each geometric variable are:

$$\begin{aligned}
\frac{\partial d}{\partial \bullet} &= \frac{1}{d}(x\frac{\partial x}{\partial \bullet} + y\frac{\partial y}{\partial \bullet}) \\
\frac{\partial x}{\partial \lambda_1} &= (x' - t_u)\cos\theta, & \frac{\partial y}{\partial \lambda_1} &= (x' - t_u)\sin\theta \\
\frac{\partial x}{\partial \lambda_2} &= -(y' - t_v)\sin\theta, & \frac{\partial y}{\partial \lambda_2} &= (y' - t_v)\cos\theta \\
\frac{\partial x}{\partial \theta} &= -(\lambda_1(x' - t_u)\sin\theta + \lambda_2(y' - t_v)\cos\theta), & \frac{\partial y}{\partial \theta} &= (\lambda_1(x' - t_u)\cos\theta - \lambda_2(y' - t_v)\sin\theta) \\
\frac{\partial x}{\partial t_u} &= -\lambda_1\cos\theta, & \frac{\partial y}{\partial t_u} &= -\lambda_1\sin\theta \\
\frac{\partial x}{\partial t_v} &= \lambda_2\sin\theta, & \frac{\partial y}{\partial t_v} &= -\lambda_2\cos\theta
\end{aligned} \tag{3.18}$$

The resulting Jacobian matrix has the form:

$$J = \begin{pmatrix}
\frac{\partial e_{(0,0)}}{\partial \lambda_1} & \frac{\partial e_{(0,0)}}{\partial \lambda_2} & \cdots & \frac{\partial e_{(0,0)}}{\partial L_1} & \frac{\partial e_{(0,0)}}{\partial L_2} \\
\frac{\partial e_{(0,1)}}{\partial \lambda_1} & \frac{\partial e_{(0,1)}}{\partial \lambda_2} & \cdots & \frac{\partial e_{(0,1)}}{\partial L_1} & \frac{\partial e_{(0,1)}}{\partial L_2} \\
\vdots & & & \ddots & \\
\frac{\partial e_{(w,h)}}{\partial \lambda_1} & \frac{\partial e_{(w,h)}}{\partial \lambda_2} & \cdots & \frac{\partial e_{(w,h)}}{\partial L_1} & \frac{\partial e_{(w,h)}}{\partial L_2}
\end{pmatrix} \tag{3.19}$$

3.3 Keypoint ordering

The algorithm is fully automatic and does not require any user interaction. We aim at a set of very simple steps that help to order keypoints. The pattern sample is displayed on Figure 3.3

In order to process the set of keypoints to the algorithm, it is important to order them exactly same way as pattern keypoints are, for example from top to bottom column-wise. Simple sorting techniques such as ordering according u and then v

coordinate may not be efficient since we deal with the image of the pattern which was rotated, translated and then projected into camera image. The simplest way was to determine approximate homography H_{app} using match of corner keypoints of the image and the pattern, and then order the rest of image keypoints with help of H_{app}. Therefore, we are interested in selecting the four ellipses that are located at the corners of the calibration pattern and then putting them in correspondence with the corners of model pattern. If we are able to estimate this transformation, then we can easily estimate the correspondence for the rest of the points. This is carried out by means of homography.

The similar idea is usually applied in human-assisted semi-automatic environment where a user selects the four corners and the algorithm manages to do the rest, for example, Matlab Calibration Toolbox [BOUGUET 2000]. Our goal is to have fully automatic software.

The principle for determination of corners from a given set of keypoints is displayed in Figure 3.4. We followed these steps to order all the keypoints:

1. For each of the keypoint c (potentially corner) do:

   - find its first three neighbours n_1, n_2 and n_3 based on Euclidean distance;
   - calculate the angles $\angle n_1 c n_2$, $\angle n_1 c n_3$, $\angle n_2 c n_3$ and pick the maximum $\angle_{max} = \angle n_{i1} c n_{i2}$ with i_1 and i_2 corresponding keypoint indices that satisfy maximum angle condition;
   - if the angle is more than 180° then c is not a corner;
   - otherwise for all the rest of the keypoints $\{n_i\}_{i=3,\cdots,N}$, make sure they are located within the maximum angle $\angle n_{i1} c n_{i2}$ (within small tolerance ε, for example, 2°) and if this condition holds, c is a corner:

$$\begin{aligned} \angle n_{i1} c n_{i2} &> \angle n_i c n_{i1} - \varepsilon \\ \angle n_{i1} c n_{i2} &> \angle n_i c n_{i2} - \varepsilon \end{aligned} \tag{3.20}$$

2. Sort the four corners by coordinate - firstly by u, then by v.

3. Calculate approximate homography H_{app} using correspondence of corner keypoints c_1, c_2, c_3, c_4, and corresponding pattern corner keypoint locations.

4. Sort all the keypoints same way as they are sorted for the pattern by calculating approximate location of the keypoint in image $x_{app} = H_{app} X$ and then identifying the closest keypoint x to x_{app}.

3.4 Sub-pixel ellipse center detection accuracy

Experiments are performed to measure the ellipse center detection accuracy at different ellipse sizes against noise and aliasing, and also against different angles

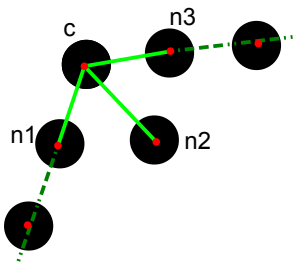


Figure 3.4: Defining the corner keypoint: when the maximum angle $\angle n_1 c n_3$ between corner's c first three neighbours n_1, n_2, n_3 stays maximum for the rest of the keypoints (within small tolerance) and at the same time less than 180°.

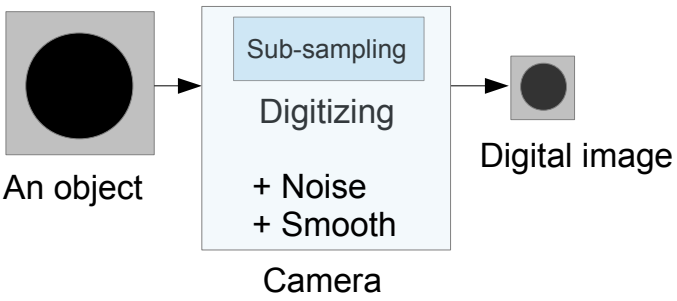


Figure 3.5: Digitizing of the pattern.

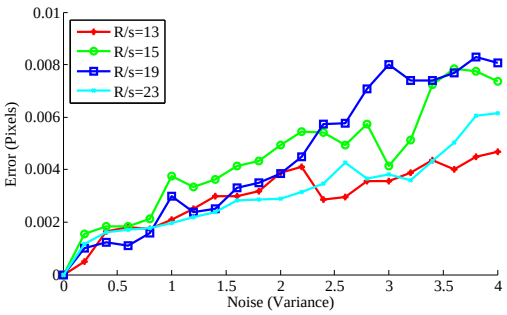
of view of the circle. The synthetic data was generated according a principle of digitizing shown on Figure 3.5.

Synthetic disk (filled circle) 8-bit images were generated on a large image of size $W \times W$ pixels ($W = 1000$ for the first set and $W = 2000$ for the second), which was blurred, downsampled and finally Gaussian noise was added. Subpixel disk center location is used as ground truth and compared to detected disk center.

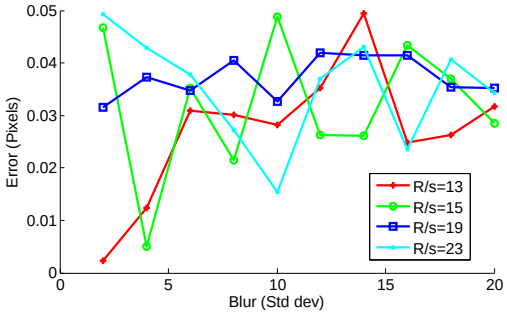
Each set includes four 8-bit images with a disk on each of a different radius size. That is, an ellipse $\{x, y\}$ is drawn with the fixed center (x_c, y_c), radius R and rotation angle φ (fixed to 30°) along z axis and changing angle of view θ as described:

$$\begin{aligned} &\frac{((x - x_c)\cos\varphi - (y - y_c)\sin\varphi)^2}{R^2} + \frac{((x - x_c)\sin\varphi + (y - y_c)\cos\varphi)^2}{(R\cos\theta)^2} \leq 1, \\ &(x_c, y_c) = (0.5W + \delta_x, 0.5W + \delta_y), \\ &R = n\frac{W}{2}, \end{aligned} \tag{3.21}$$

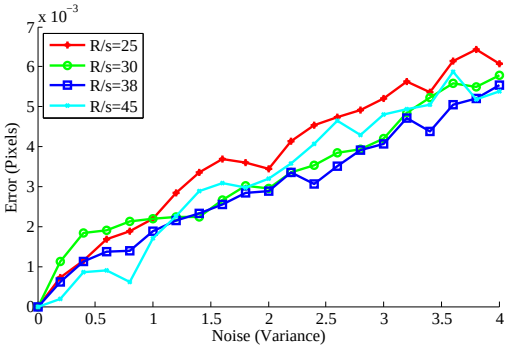
where (δ_x, δ_y) are shifts from image center for x, y directions, and $n = [0.45, 0.6, 0.75, 0.9]$ is assigned so that to see if the ratio between image area and disk area influences the detection accuracy.



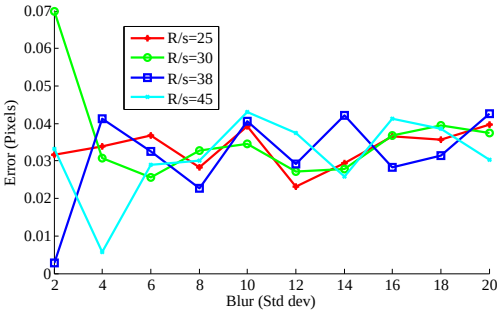
(a) Gaussian noise, $W = 1000$



(b) Aliasing, $W = 1000$



(c) Gaussian noise, $W = 2000$



(d) Aliasing, $W = 2000$

Figure 3.6: Keypoint detection precision performance for normal angle view $\theta = 0°$ with subsampling rate $s = 20$, original image resolution W and subsampled circle radii $\frac{R}{s}$, pixels.

The ground-truth circle centers for each sub-sampled image are found:

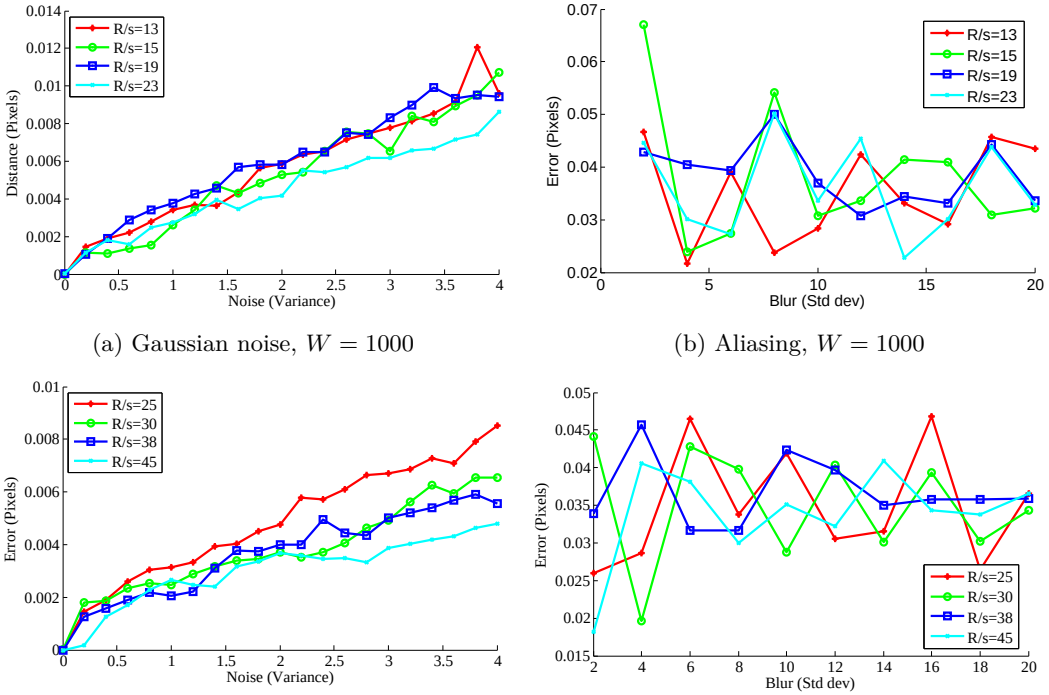
$$(x_{g.t}, y_{g.t}) = (0.5\frac{W}{s} + \frac{\delta_x}{s}, 0.5\frac{W}{s} + \frac{\delta_y}{s}), \tag{3.22}$$

with s being a downsampling rate ($s = 20$ for both image sets). Figure 3.6 and Figure 3.7 show the performance of the algorithm against Gaussian noise level (median error out of 100 iterations) and amount of aliasing for different disk radii (the view angle θ is set to 0° on the first figure and 55° on the second). Similar experiments were performed for other view angles, up to 70° and it was found that error always has the same tendency as in shown figures.

3.5 Conclusion

The described algorithm allows automatic keypoint detection and ordering of the circled pattern. From the graphs presented at experiment section it can be also concluded:

1. In all the cases the average error is less than 0.05 pixels, even for small disk radius (11 pixels). And considering the fact that ratio of the disk area and its



(a) Gaussian noise, $W = 1000$

(b) Aliasing, $W = 1000$

(c) Gaussian noise, $W = 2000$

(d) Aliasing, $W = 2000$

Figure 3.7: Keypoint detection precision performance for $\theta = 55°$ angle view with subsampling rate $s = 20$, original image resolution W and subsampled circle radii $\frac{R}{s}$, pixels.

enclosing image area has little impact on the precision detection, allows us to pack more keypoints in a given pattern area.

2. The angle of view has little influence on precision even for large angles, which means the detection stays robust even when pattern plane is viewed from a large angle of view (here we don't consider perspective bias which is a subject in Chapter 5, but rather affine transform of the elliptic parameters).

3. As expected, the error increases linearly with noise level, but remains constant under aliasing. This is important because in a Bayer pattern image, red and blue channels are notoriously aliased. Figure 3.6 (b,d) and Figure 3.7 (b,d) show that this does not affect the disk center detection.

Chapter 4

High-precision correction of lateral chromatic aberration in digital images

Nowadays digital image sensor technology continues to develop much faster than optical technology for imaging systems. The result quality loss is due to lateral chromatic aberration and it is becoming more significant with the increase of sensor resolution. For the sake of higher performance and lighter lens systems, especially in the field of computer vision, we aim to find new ways to overcome resulting image quality limitations.

This chapter demonstrates a robust method to minimize the lateral chromatic aberration, recovering the loss of image quality by using a single circled pattern image. Different series of tests and measures are used to validate the algorithm. For the performance evaluation we used both synthetic and real data.

The primary contribution of this work is an effective algorithm to robustly correct lateral chromatic aberration through color planes warping. We aim at high precision (largely sub-pixel) realignment of color channels. This is achieved thanks to two ingredients: high precision keypoint detection, which in our case are circle centers, and more general correction model than what is commonly used in the literature, radial polynomial. The setup is quite easy to implement, requiring a pattern of black disks on white paper and a single snapshot.

We perform the error measurements in terms of geometry and of color. Quantitative results on real images show that our method allows alignment of average 0.05 pixel of color channels and residual color error divided by a factor 3 to 6. Finally, the performance of the system is compared and analysed against three different software programs in terms of geometrical misalignments.

Keywords Chromatic aberration, image warping, camera calibration, polynomial model, image enhancement

Contents

4.1 Introduction

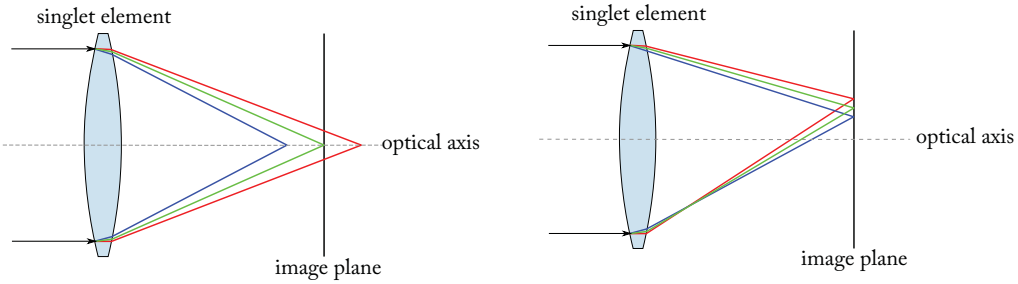
Every optical system that uses lenses suffers from aberrations that occur due to the refractive characteristics of the lenses [JÜRGEN 1995]. Loss in image accuracy because of a known set of aberrations can be classified as either monochromatic or chromatic types of aberration. They occur due to physical interaction of light with materials, lens design constraints and manufacture limitations. The five Seidel monochromatic aberrations include spherical aberration, comatic, astigmatic, curvature of field and distortion [SEIDEL 1856]. The chromatic type of aberration is independent from monochromatic one.

A type of distortion named Chromatic Aberration (CA) is inherent for any optical system due to result of the different refractive indices of the lens medium (typically some form of glass) for varying wavelength of the transmitted light, and such phenomena is termed *dispersion* [NEWTON 1704] and explains prismatic behaviour. From aesthetic point of view, CA gives overall impression of poor quality and definition, while from the view of computer vision application algorithms - may reduce stability and precision for the application when color information matters (i.e. deteriorating of details on the texture or edge areas). The main classification of CA is two categories [SLAMA 1980]: *Axial* (or longitudinal) Chromatic Aberration - ACA, and *Lateral* (or transverse) Chromatic Aberration - LCA.

The ACA occurs when different wavelengths focus at different distances from the lens, i.e., different points on the optical axis as shown on Figure 4.1 (a). It causes a failure of all the wavelengths to be focused at the same convergence point, and as a result, as light strikes the sensor panel, out of focus rays contribute to a circle of confusion, or bokeh. In digital images it manifests as a subtle coloured halo around the boundary of an object, especially in the circumstances like lens widest aperture setting. Such image artifacts might be decreased when the lens aperture is stopped down or reduced due to the increase in depth of field which brings the axially-misaligned focal points nearer. Many modern digital cameras, when in fully-automatic mode, are able to balance the aperture size preventing significant spacial frequency loss due to photon diffraction [MIELENZ 1999] and a shallow depth of field, increasing focus selectivity and focus error as a side effect. Thanks to camera's automatic mode, ACA is nominally minimised to imperceptible levels.

The LCA happens when the wavelengths focus at different points on the focal plane and thus geometrical color plane misalignments occur as shown on Figure 4.1 (b). It is relative and non-linear displacement of the three color planes; on the obtained digital images the channels are misaligned with respect to each other. This source of degradation manifests itself as fringes of color at edges and high contrast areas which leads to a perceived detail loss. Less perceptible impact in lower-contrast areas reduces texture detail and generally tends to reduce the perception that LCA compromises image quality.

An example of CA is shown on Figure 4.2. It can be seen that CA limits image detail. Also the present CA has assymetric nature which means the type LCA is the dominant aberration and attenuating quality. ACA occurs symmetrically to



(a) Axial chromatic aberration (b) Lateral chromatic aberration

Figure 4.1: Occurence of axial (a) and lateral (b) chromatic aberrations.

Figure 4.2: Example of LCA presence in digital image crop taken by camera Sony DSLR-A200 with focal length $f = 24mm$. The color channel misalignments are noticeable due to LCA and are mostly highlighted in the regions of edges.

image feaure and thus much less significant. Based on this our main focus will be LCA type of distortion.

Some work had been done in this field: some included optical and hardware solutions, some only used digital manipulations to combat the problem. Optimal lens design techniques such as [POWELL 1981, FARRAR 2000, MILLÁN 2006a, MILLÁN 2006b, FANGA 2006] have been proposed to compensate the aberration affect. These approaches alleviate CA by using multiple lenses with different optical indices, i.e., the achromatic doublet [KINGSLAKE 1978], or by improving the lens characteristics [MILLÁN 2006b]. Those corrections, however, have been effective for only zones near the optical axis, and the hardware solutions increase the system cost and can be only used in large cameras.

Another way to approach the problem is through modification of camera settings (magnification, focus, sensor shifting) for each color channel so that to eliminate the effects introduced by chromatic aberration, i.e., active lens system introduced by [WILLSON 1991, WILLSON 1994a]. The purpose of the active lens control is to get best focused image for each color component by slightly shifting the image plane backward and forward for each color. The solution requires *a priori* know-

ledge about magnification factor and the image plane shift degree [WILLSON 1991]. Such approach may not be practical since it requires taking three pictures for each channel under different camera settings. The better alternative is to refer to a post processing of a digital image which is less applicable for the correction of axial aberration, but can perform well for the lateral one. The post processing not only allows to save the high image quality by much cheaper cost, but also remains more practical.

The algorithms that perform *automatic* compensation of both types of CA do not require any *a priori* information and do the correction on the given image directly. They are fast and practical for the purposes of art photography when we only want to reduce the effects of CA in areas of interested (high contrast areas, textures) for a selected image. One example is [KANG 2007] which corrects CA from a single degraded image: it first models the whole process of the imaging pipeline and then estimates every simplified step of the pipeline to correct both LCA and ACA. Another example, [CHUNG 2010] suggests a fast and simple algorithm for low-cost digital cameras or cellular phones with built-in cameras; its main idea is to perform a correction so that to adjust color difference signals in areas where they are greatly misaligned. While such method might be effective for photographic purposes, it is not possible to integrate it if we aim for camera calibration where some kind of *a priory* (i.e., calibration pattern) is necessary.

Given calibration pattern, [BOULT 1992] refers to the LCA and formulates its compensation as an *image warping* problem which means digital manipulation of each color channel so that to re-align them. The method is focused solely on the warping problem and does not employ any aberration model, performing the correction by interpolating the control points. [KOZUBEK 2000] shows how to do the correction for both types of aberrations in the environment of fluorescent microscopy. However, this method as well as others that work for microscopic devices [MANDERS 1997, SHI 2004, FREITAG 2005] can not be applied for the camera imaging systems, therefore, our focus is solely into the model-based warping methods that only deal with the correction of lateral type of aberration. [MATSUOKA 2012] provides an evaluation of the chromatic aberration correction methods based on calibration, all of which assume radial nature of the distortion and thus employ different radial models for the correction, i.e., [REBIAI 1992, KAUFMANN 2005, JOHNSON 2006, LUHMANN 2006, MALLON 2007a]. However, it is important to keep in mind that not all the cameras can be satisfactorily corrected by the same radial model [TANG 2011]. Besides, they require estimation of the center point of the aberration for accurate correction which is often different from image center due to complexity of multi-lens system [WILLSON 1994b].

Therefore, our main objective is to achieve correction of LCA using a single shot of a calibration pattern by typical camera imaging system using digital post-processing steps. We want to address LCA by a generalized distortion model, and we expect it to be robust and high-precision. In order to define a magnitude of *high precision* correction, a visual perception experiment was done for different misalignment levels (in pixel units). A synthetic disk (black filled circle) was generated on

(a) $d = 0.05$ (b) $d = 0.1$ (c) $d = 0.2$ (d) $d = 0.3$ (e) $d = 0.5$ (f) $d = 1$

Figure 4.3: Visual perception tests for chromatic aberration on synthetic disk image for different values of displacements d (in pixels). Note that a displacement value of 0.1 pixel is just noticeable, while 0.3 pixel displacement is quite perceptible.

a large image with misalignment between channels introduced, then the image was blurred in order to avoid aliasing, and downsampled. A part of the downsampled image was cropped and zoomed-in in Figure 4.3. It can be noted that 0.1 pixel misalignment is a borderline when aberration becomes just-noticeable, while misalignments of 0.3 pixel and higher are quite perceptible.

Consequently, the primary objective is to develop a robust, generalized algorithm which leaves final correction within the range of 0.1 pixels. Secondary objectives are to develop some metrics to tangibly quantify corrected image quality and demonstrate the quality improvement after CA correction.

This work demonstrates how the lateral chromatic aberration in digital images can be modelled and corrected by a polynomial model which remains general for any kind of distortion. The correction is achieved through image warping by realigning the red and blue color planes with respect to a reference green plane. An important role of the high precision correction plays a sub-pixel keypoint detection which is described in Chapter 3 and also its dense field registration; one keypoint of the calibration pattern is represented by a center of filled black circle (dics).

Further Section 4.2 demonstrates how the geometrical displacements that occur between the keypoints of the color planes are modelled and corrected by a general polynomial model. Section 4.3 exposes the result measurements of the lateral aberration in terms of both color and geometry misalignments, as well as provides real scene examples. Finally, Section 4.4 concludes.

4.2 Calibration and correction

Lateral chromatic aberration can be considered as geometrical distortion which occurs with respect to color channels. Assuming the distortion field is smooth by nature, the main idea is to model the distorted fields, normally red and blue channels, by *polynomials* for x and y directions, while keeping one as a reference (green). By comparison with many other distortion models (radial, division, rational function), the polynomial model is general and can perfectly approximate the former models as stated by [GROMPONE VON GIOI 2010, TANG 2011]. To achieve high precision the degree 11 was chosen for both x and y, since it is the order when residual error stabilizes for the most of the images.

For a given calibration pattern image, the detected keypoints are represented in pixel coordinates as (x_f, y_f) for a certain color plane f which is either red (r),

blue (b) or green (g). The lateral misalignments between the red (or blue) and the green planes are corrected by identifying the parameters of polynomial p_{fx}, p_{fy} approximating at best the equations

$$\begin{aligned} x_{gi} &= p_{fx}(x_{fi}, y_{fi}) \\ y_{gi} &= p_{fy}(x_{fi}, y_{fi}), \end{aligned} \tag{4.1}$$

with the target colors $f = r$ or $f = b$ and i describing keypoint index.

The polynomial model p_{fx}, p_{fy} with orders m, n and polynomial coefficients $\left\{p_{x_0}, p_{x_1}, \cdots, p_{x_{\frac{(m+1)(m+2)}{2}-1}}\right\}$, $\left\{p_{y_0}, p_{y_1}, \cdots, p_{y_{\frac{(n+1)(n+2)}{2}-1}}\right\}$ can be expanded as:

$$\begin{aligned} x_g = {} & p_{x_0}x_f^m + p_{x_1}x_f^{m-1}y_f + p_{x_2}x_f^{m-2}y_f^2 + \ldots + p_{x_m}y_f^m + p_{x_{m+1}}x_f^{m-1} \\ & + p_{x_{m+2}}x_f^{m-2}y_f + \ldots + p_{x_{2m}}y_f^{m-1} + \ldots + p_{x_{\frac{(m+1)(m+2)}{2}-3}}x_f \\ & + p_{x_{\frac{(m+1)(m+2)}{2}-2}}y_f + p_{x_{\frac{(m+1)(m+2)}{2}-1}} \\ y_g = {} & p_{y_0}x_f^n + p_{y_1}x_f^{n-1}y_f + p_{y_2}x_f^{n-2}y_f^2 + \ldots + p_{y_n}y_f^n + p_{y_{n+1}}x_f^{n-1} \\ & + p_{y_{n+2}}x_f^{n-2}y_f + \ldots + p_{y_{2n}}y_f^{n-1} + \ldots + p_{y_{\frac{(n+1)(n+2)}{2}-3}}x_f \\ & + p_{y_{\frac{(n+1)(n+2)}{2}-2}}y_f + p_{y_{\frac{(n+1)(n+2)}{2}-1}} \end{aligned} \tag{4.2}$$

The unknowns are the parameters of polynomials p_{fx} and p_{fy}. These are $\frac{(m+1)(m+2)}{2} = 78$ for degrees $m = n = 11$ for each polynomial. Our pattern is composed of about 1000 disks (we already showed in Chapter 3 that we can pack so many disks of small diameter on our pattern without deterioration of precision), so there is no risk of overfitting. The main principle is based on how well we can approach the reference field (x_g, y_g) by $p_{fx}(x_f, y_f)$ and $p_{fy}(x_f, y_f)$. For a number K of keypoints $(x_{fi}, y_{fi}), i = 1, \cdots, K$ distributed all over the image, the polynomial coefficients are computed by minimizing the difference of displacements between the reference and distorted channels:

$$E = \sum_{i=1}^{K} (p_{fx}(x_{fi}, y_{fi}) - x_{gi})^2 + (p_{fy}(x_{fi}, y_{fi}) - y_{gi})^2 \tag{4.3}$$

The solution vector p of this least-square problem satisfies a linear system:

$$Ap = v \tag{4.4}$$

with A the coefficient matrix built from the x_{fi}, y_{fi}, and p the unknown coefficient vector. For favorable numerical conditioning of A, these pixel coordinates need to be normalized between 0 and 1.

The optimal solution minimizing the norm $\|Ap - v\|$ is $p = (A^T A)^{-1} A^T v$. In practice, the coefficient matrix A is ill-conditioned and can make the solution unstable. The following normalization is necessary in order to make the linear system

more stable. A is multiplied by normalization matrices T_1 and T_2 so that the entries of the normalized matrix $\hat{A}$ do not vary a lot.

$$\hat{A}\hat{p} = T_2 A T_1 (T_1^{-1} p) = T_2 v \tag{4.5}$$

chosen so that the entries of T_2AT_1 get closer to each other. The solution is $p = T_1(\hat{A}^T\hat{A})^{-1}\hat{A}^T T_2 v$.

By some simple computations, the linear system in Eq. (4.4) is explicitly written as:

$$\begin{aligned} Ap = &\begin{bmatrix} x_{r_1}^m & x_{r_1}^{m-1}y_{r_1} & \cdots & y_{r_1} & 1 & & & & & \\ x_{r_2}^m & x_{r_2}^{m-1}y_{r_2} & \cdots & y_{r_2} & 1 & & & & & \\ \vdots & & & \ddots & & & & 0 & & \\ x_{r_M}^m & x_{r_M}^{m-1}y_{r_M} & \cdots & y_{r_M} & 1 & & & & & \\ & & & & & x_{r_1}^n & x_{r_1}^{n-1}y_{r_1} & \cdots & y_{r_1} & 1 \\ & & & & & x_{r_2}^n & x_{r_2}^{n-1}y_{r_2} & \cdots & y_{r_2} & 1 \\ & & 0 & & & \vdots & & & \ddots & \\ & & & & & x_{r_M}^n & x_{r_M}^{n-1}y_{r_M} & \cdots & y_{r_M} & 1 \end{bmatrix} \begin{pmatrix} p_{x_0} \\ p_{x_1} \\ \vdots \\ p_{x_{\frac{(m+1)(m+2)}{2}-1}} \\ p_{y_0} \\ p_{y_1} \\ \vdots \\ p_{y_{\frac{(n+1)(n+2)}{2}-1}} \end{pmatrix} \\ &= v = \begin{pmatrix} x_{g_1} \\ x_{g_2} \\ \vdots \\ x_{g_M} \\ y_{g_1} \\ y_{g_2} \\ \vdots \\ y_{g_M} \end{pmatrix} \end{aligned} \tag{4.6}$$

It can be seen that for polynomials of a high degree, i.e. 11, the entries of A differ by a big ratio $\frac{x_{r_i}^{11}}{1}$ for x variable and $\frac{y_{r_i}^{11}}{1}$ for y variable, and this causes a numerical instability of the linear system. The normalization matrices T_1 and T_2

are used to lessen the instability in Eq. (4.5) and are computed as:

$$T_1 = \begin{bmatrix} \frac{1}{x_{r_1}^m} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{x_{r_1}^{m-1} y_{r_1}} & \cdots & & & & \\ \vdots & & \ddots & & & & \vdots \\ & & & & \frac{1}{x_{r_1}^n} & & \\ & & & & & \ddots & 0 \\ 0 & & \cdots & & & 0 & 1 \end{bmatrix}$$

$$T_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{x_{r_1}^m}{x_{r_1}^{m-1} y_{r_1}} & \cdots & & & & \\ \vdots & & \ddots & & & & \vdots \\ & & & & 1 & & \\ & & & & & \ddots & 0 \\ 0 & & \cdots & & & 0 & x_{r_1}^n \end{bmatrix} \tag{4.7}$$

When the calibration is done, it is now straightforward to build a corrected image for any distorted image which was taken under the same fixed camera settings. The polynomials p_{fx} and p_{fy} calculate the corrected pixel coordinates for each distorted pixel coordinate (x_f, y_f) as in Eq. (4.1), and then the corrected color value is obtained by interpolation from the corrected coordinates.

4.3 Experiments

To evaluate the CA correction method we use two types of metrics: geometry and color. Comparison to existing commercial software is done as well.

4.3.1 Chromatic aberration correction accuracy with reflex cameras

To evaluate the correction method three types of images are used: calibration pattern, test pattern and real images. Two types of calibration patterns are considered in order to show the importance of having precise keypoints: with circles where each keypoint is a center, and with noise where keypoints are detected using SIFT algorithm. The circled pattern is printed on $A3$ format paper and there, $37 \times 26 = 962$ black disks are drawn of the radius $0.4cm$ and separation of $1.1cm$ between consecutive disks. The picture of calibration pattern is taken with chosen fixed settings, after that a test image is taken of the same pattern. After the calibration is complete, the correction method is applied to the test image to evaluate precision there; after, some outdoor images were taken under the same camera settings as before in order to demonstrate the image quality improvement.

Abbreviation	EOS 5D	EOS 40D	A200
Camera	Canon EOS 5D	Canon EOS 40D	Sony DSLR A200
Sensor size	$35.8mm \times 23.9mm$	$22.2mm \times 14.8mm$	$23.6mm \times 15.8mm$
Pixel size on the focal plane	$8.2\mu m$	$5.7\mu m$	$6.05\mu m$
Number of recording pixels	5634×3752	3888×2592	3898×2616
Lens	Canon EF $24-70mm$ $f2.8-f22$ 16 elements in 13 groups	Canon EFS $18-55mm$ $f3.5-f5.6$ 11 elements in 9 groups	Sony DT $18-70mm$ $f3.5-f5.6$ 11 elements in 9 groups

Table 4.1: Cameras specifications.

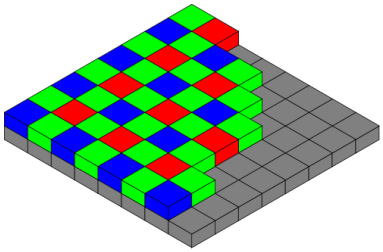


Figure 4.4: The Bayer arrangement of color filters on the pixel array of an image sensor

4.3.1.1 Camera specifications

Three digital cameras with interchangeable lenses were used to capture the test images. Their specifications are given in the Table 4.1.

4.3.1.2 Demosaicking

The demosaicking is performed so as to keep the original raw image resolution. The camera's sensor is overlaid with a color filter array, normally a Bayer filter which consists of a mosaic of a 2×2 matrix of red, blue, green and green filters as shown in Figure 4.4. Normally, to reconstruct a full color image from the data collected by a color filter array, a form of interpolation is needed to fill in the blanks for each channel.

Considering the aberration correction algorithm, the main steps of demosaicking are:

Step 1. RAW image is separated in its three channels.

Step 2. The green channel is kept at original resolution by averaging the four nearest neighbours for each empty pixel.

Step 3. Blue and red channels are kept in aliased, half-dimension images.

Step 4. For all the channels the keypoint detection is performed.

Step 5. The correction polynomials for red and blue images are obtained after matching the keypoints green-blue and green-red geometrically (for this, the red/blue keypoint coordinates are multiplied by 2 to be at the same resolution as green keypoint coordinates).

Step 6. The corrected and zoomed-in images for blue and red are calculated by first obtaining the correction coordinates, and then interpolating them from given blue/red distorted images.

The crude demosaicking of the green channel (bilinear interpolation) creates an aliased image (step 1). This may be true, but earlier it is already shown (refer to Figure 3.6 (b,d)) that blurring does not affect the ellipse center detection precision.

4.3.1.3 Geometrical misalignments for a circled pattern

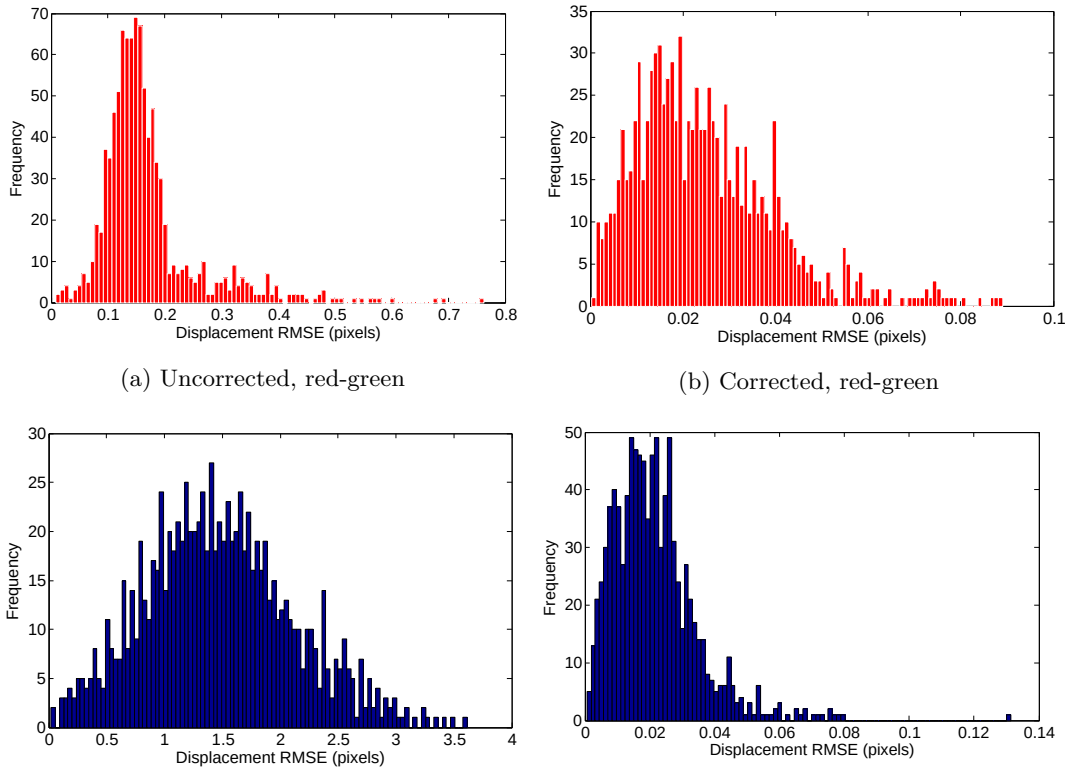
The color plane keypoint displacement RMSE and maximum distances (Euclidean) before and after the correction are presented in Table 4.2. From the table it is seen the significant reduction in the displacements. It can be noted that the residuals are of similar magnitude as the sub-pixel keypoint detection, which emphasize the importance of having precise keypoints.

	Uncorrected		Corrected	
Cameras' shots	R/G	B/G	R/G	B/G
EOS 5D				
Calib $f_1 = 24mm$	0.1917 (0.7632)	1.6061 (3.6154)	0.0291 (0.0889)	0.0249 (0.1323)
Test $f_1 = 24mm$	0.1659 (0.5818)	1.5129 (3.2057)	0.0492 (0.0880)	0.0551 (0.0904)
Calib $f_2 = 70mm$	0.5547 (0.7720)	1.4087 (1.8920)	0.0292 (0.0800)	0.0291 (0.07511)
Test $f_2 = 70mm$	0.5321 (0.7864)	1.4140 (1.8044)	0.0352 (0.0917)	0.0331 (0.1024)
EOS 40D				
Calib $f_1 = 18mm$	0.6546 (0.9784)	1.4190 (3.3588)	0.0298 (0.1136)	0.0584 (0.1531)
Test $f_1 = 18mm$	0.6713 (0.9916)	1.4133 (3.2901)	0.0487 (0.1408)	0.0341 (0.0917)
Calib $f_2 = 55mm$	0.4590 (0.8794)	1.5242 (2.4967)	0.0447 (0.1233)	0.0398 (0.0922)
Test $f_2 = 55mm$	0.4391 (0.8029)	1.5231 (2.5574)	0.0522 (0.1666)	0.0564 (0.1921)
A200				
Calib $f_1 = 18mm$	0.9106 (1.1422)	1.5371 (3.4125)	0.0344 (0.1037)	0.0373 (0.0882)
Test $f_1 = 18mm$	0.9127 (1.3381)	1.5371 (3.4092)	0.0504 (0.1356)	0.0490 (0.0916)
Calib $f_2 = 70mm$	0.2502 (0.5382)	1.7066 (2.4355)	0.0492 (0.1249)	0.0429 (0.1502)
Test $f_2 = 70mm$	0.2627 (0.5521)	1.7184 (2.5051)	0.0495 (0.1845)	0.0624 (0.1890)

Table 4.2: Color plane keypoint Euclidean displacements in RMSE (maximum error) format in pixels before and after the calibration for three different cameras, each with two different focal lengths. R/G and B/G correspond to red and blue misalignments with reference to the green channel.

Further details on keypoint displacements are presented for one case, (camera EOS 5D with $f_1 = 24mm$) in Figure 4.5. From the histograms it can be seen that error distribution decreases and stays within 0.05 pixels; and this numerical result holds for the most of our tests. The field of geometrical displacements was built

before and after the correction and is shown in Figure 4.6. It also demonstrates how the character of the field had changed: before the correction it had radial nature, after correction it is less structured.



(a) Uncorrected, red-green

(b) Corrected, red-green

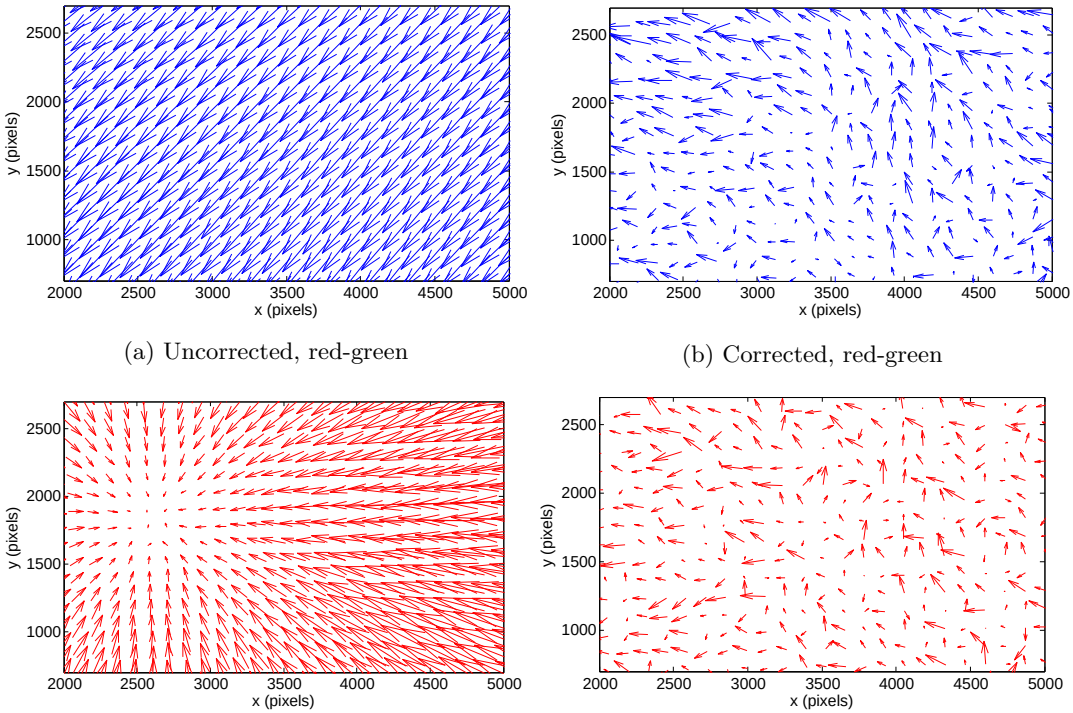
(c) Uncorrected, blue-green

(d) Corrected, blue-green

Figure 4.5: Histograms of Euclidean displacements of the keypoints: before the correction (a,c) and after (b,d). The color of the graph corresponds to color of the channel - (a,b) for red and (c,d) for blue.

4.3.1.4 Geometrical misalignments for a noise pattern using SIFT keypoints

In order to see the influence of keypoint detection precision on the correction accuracy, another type of calibration pattern was used in addition - noise pattern. The keypoints were detected using standard Scale-Invariant Feature Transform (SIFT) algorithm [LOWE 2004] which allowed to have around 9000 keypoints for each channel. The precision results are demonstrated for camera EOS 40D in Table 4.3. As it can be seen the corrected precision is about 5-10 times larger than for circled pattern tests, which means that it is not the number of keypoints that matters, but the detection precision.



(a) Uncorrected, red-green

(b) Corrected, red-green

(c) Uncorrected, blue-green

(d) Corrected, blue-green

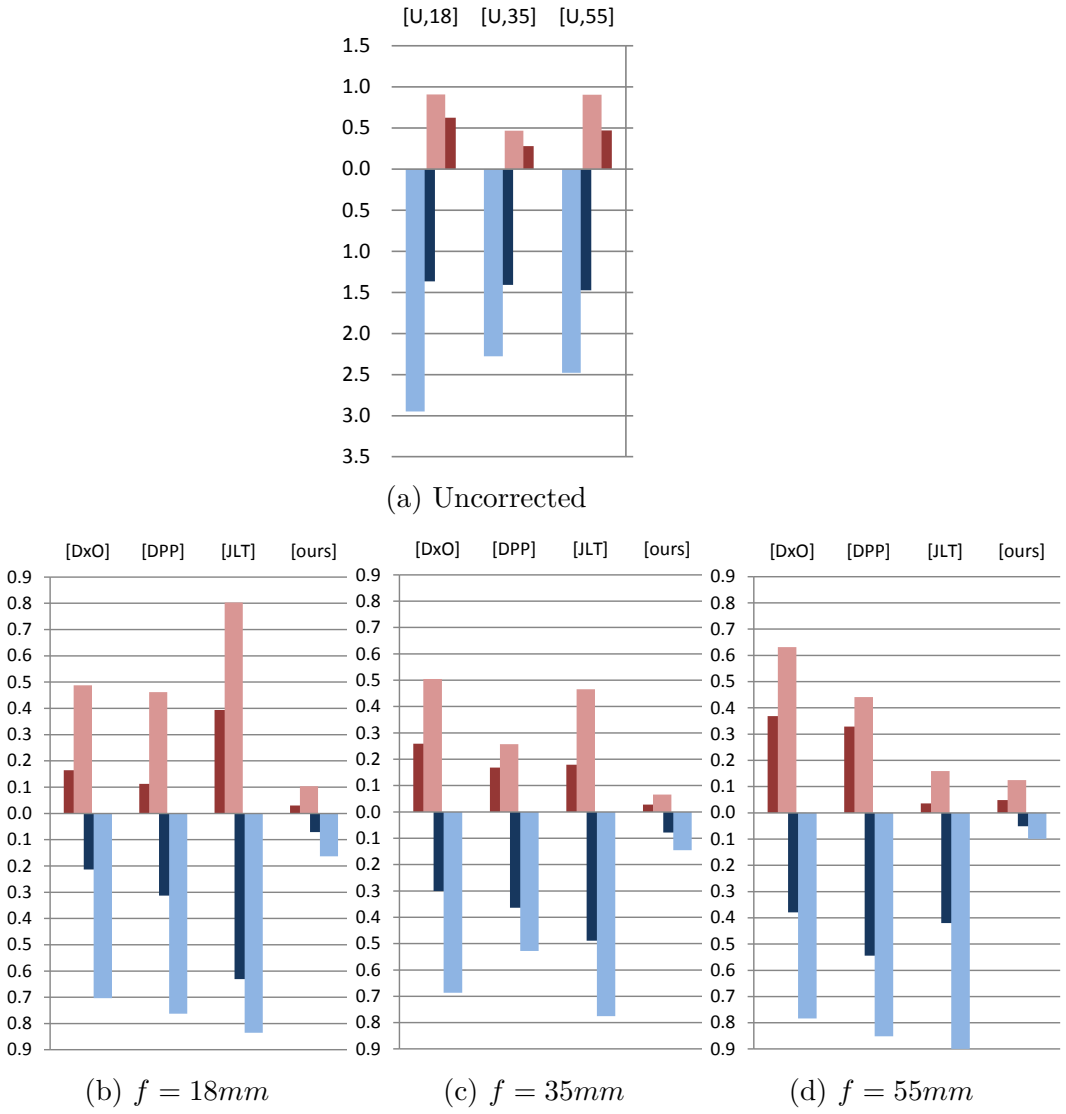
Figure 4.6: Vector fields of Euclidean displacements of the keypoints: before the correction (a,c) and after (b,d). The color of the arrows corresponds to color of the channel - (a,b) for blue and (c,d) for red.

4.3.1.5 Geometrical misalignments comparison with existing software corrections

Three software solutions were chosen to perform a comparison with our method: DxO Optics Pro (noted as 'DxO') [DxO], Canon DPP ('DPP') [CANON] and Jaliko lenstool ('JLT') [DUNNE 2010]. The first two use lab-computed calibration database for each camera and each lens, and perform correction based on the database, with possibility of manual readjustment for each image. The CA correction by JLT method is fully automatic and for any kind of camera (no database) but requires several images. Cam 2 was chosen for this experiment and the keypoint displacement results are shown in Figure 4.7 for different focal lengths. The results demonstrate that only our method achieves precision where defects are not visible anymore (see Figure 4.3): the mean always stays within $0.05 - 0.1$ pixel while for other methods the average remains around 0.4 pixels.

4.3.1.6 Color misalignments

To visualize the effect of chromatic aberration correction, color based $3D$ histogram is built for the circled calibration image. In ideal case such histogram would consist



(a) Uncorrected

(b) $f = 18mm$ (c) $f = 35mm$ (d) $f = 55mm$

Figure 4.7: Comparison of our method to other software for cam 2 and different focal lengths f, the comparison is in terms of mean (dark colored bars) and maximum (light colored bars) misalignments for both red (positive axis) and blue (negative axis) channels. (a) provides information on the initial uncorrected keypoint displacements for the same focal lengths (notice the different scale).

Camera shots	Uncorrected R/G	Uncorrected B/G	Corrected R/G	Corrected B/G
EOS 40D				
Calib $f_1 = 18mm$	0.8909	1.3388	0.2115	0.1697
Test $f_1 = 18mm$	0.9129	1.3404	0.2025	0.1627
Calib $f_2 = 55mm$	0.5129	1.5359	0.3801	0.3695
Test $f_2 = 55mm$	0.4517	1.5326	0.3316	0.3363

Table 4.3: Color plane keypoint Euclidean displacements RMSE for noise pattern before and after the calibration for the camera EOS 40D with two different focal lengths. R/G and B/G correspond to red and blue misalignments with reference to the green channel.

of only two distinct spots: $R_b = G_b = B_b = 0$ (disks) and $R_w = G_w = B_w = 255$ (background). At the same time, an ideal case of the $3D$ color histogram for a digitized image would be represented by an *idealized gray line* connecting black and white spots: $R_i = G_i = B_i$, where $i = [0, \cdots, 255]$.

Consequently, it is expected that $3D$ color cloud of the calibration image would lie along this gray line. If chromatic aberration is present, the color channels are not aligned and so, occurring blue-red hues come out at significant distance from the idealized gray line, thus creating a bulged $3D$ color cloud. Figure 4.8 shows two color clouds built in RGB space and their corresponding calibration images (close-ups only are shown) before and after the correction for the camera EOS 5D at $f_1 = 24mm$.

The shape of such color histograms can be captured with a $3D$ ellipsoid whose axes are the eigenvectors of the covariance matrix of the point cloud, and their lengths correspond to the eigenvalues.

For the point cloud which is defined as

$$X = \{f(x) - m\}_x \subset \Re^3, \tag{4.8}$$

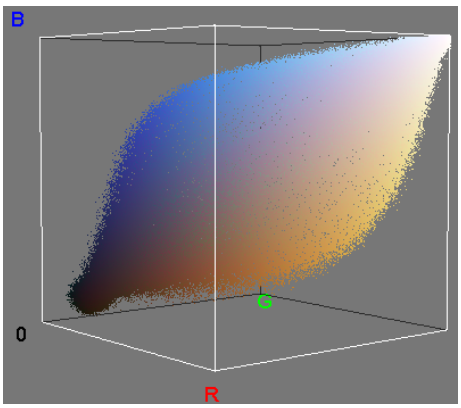
where the point index $x = [1, \cdots, N]$ and $m = \frac{1}{N}\sum_x f(x) \subseteq \Re^3$ is the cloud's centroid, the $3D$ covariance matrix of the point cloud formed by the image pixels:

$$C = \frac{1}{N} X^T X \subseteq \Re^{3\times 3}. \tag{4.9}$$

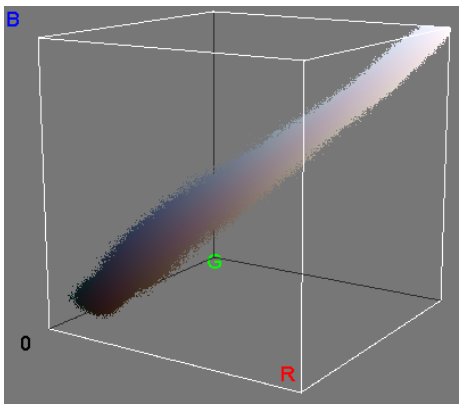
The orthogonal principal directions of the cloud are represented by eigenvectors $U \subseteq \Re^{3\times 3}$ and obtained with the eigen-decomposition of the symmetric covariance matrix C so that

$$C = UVU', \tag{4.10}$$

where $V = \begin{pmatrix} v_1 & 0 & 0 \\ 0 & v_2 & 0 \\ 0 & 0 & v_3 \end{pmatrix}$ is a diagonal matrix with eigenvalues $v_1 \geq v_2 \geq v_3$.

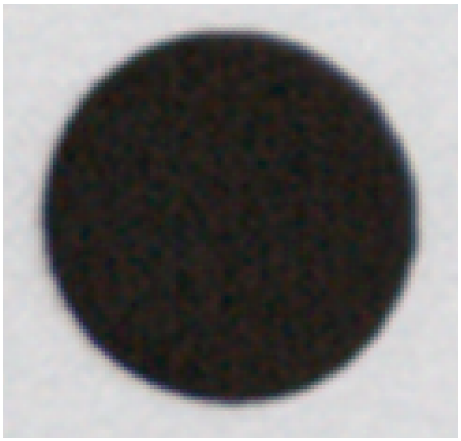


(a) Color histogram, uncorrected



(b) Color histogram, corrected

(c) Calibration image, uncorrected



(d) Calibration image, corrected

Figure 4.8: $3D$ color clouds before the correction (a) and after (b) for the calibration image and image's close-ups before the correction (c) and after (d) taken by camera EOS 5D at $f_1 = 24mm$.

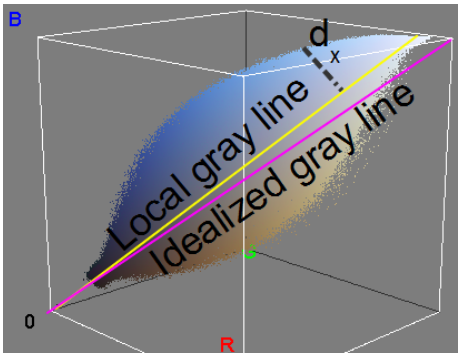


Figure 4.9: The local gray line and idealized gray line. For every point x the local gray line minimizes distance d_x from this point to the line.

The first principal component corresponds to a line that passes through multi-dimensional mean which is the *local gray line*. Normally the local gray line is not aligned with idealized gray line due to illumination presence as shown in Figure 4.9. The obtained local gray line (first component) minimizes the sum of squared distances $D = \sum_x d_x^2$ of the points from the line. Each eigenvalue is the length of principal axes, and thus is proportional to D which is correlated with each eigenvector.

To get a quantitative representation of the cloud X, we use statistics on the distances d_x from each point to the local gray line. However, not all the points are subject for consideration. Figure 4.10 shows the frequency histogram for each color pixel from the calibration image. Obviously, for the image with white background and black disks on it there are a lot of pixels of whitish and blackish colors. At the same time, red-blue hues which occurred as a result of chromatic aberration have very little frequencies. Therefore, to give more significant statistical description of the cloud point, we eliminate all the majority pixels. The distances of the remaining points to the local gray line are used to get a quantitative measure:

$$S = \sqrt{\frac{\sum_x d_x^2}{N}} \tag{4.11}$$

The elimination of local black and white spots from point cloud X is done by following:

Step 1. Calculate local centroids $\bar{X}_b, \bar{X}_w$ for black and white areas by averaging the coordinates of all the points lying below centroid (black) or above centroid (white).

Step 2. Define the range Δ_b, Δ_w where the majority points are located by calculating variances σ_b^2, σ_w^2 for black and white areas, and multiply by some threshold value, i.e. $\Delta = 2\sigma$.

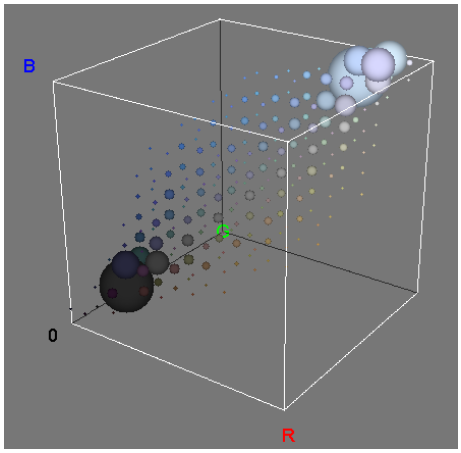


Figure 4.10: The color cloud is represented by frequency histogram: bigger spots stand for higher frequency of the corresponding color.

Step 3. When collecting statistics on the point cloud, take into account only those cloud points X that lie in the range far enough from the black and white centroids: $\bar{X}_b + \Delta_b < X < \bar{X}_w - \Delta_w$.

Assuming the gray line passes through any two points $X_1^{(g)}$, $X_2^{(g)}$, the distance from a color point x to the gray line is expressed as:

$$d_0 = \frac{|(x - X_1^{(g)}) \times (x - X_2^{(g)})|}{|X_2^{(g)} - X_1^{(g)}|} \tag{4.12}$$

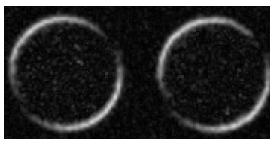
The calculation of points $X_1^{(g)}$, $X_2^{(g)}$ is straightforward considering the point cloud has a zero centroid (see eq.(4.8)) and knowing the eigenvector U_1 which corresponds to the largest eigenvalue $v1$ (from eq.(4.10)):

$$X^{(g)}(t) = tU_1, \tag{4.13}$$

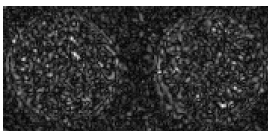
where any point on the line can be given by parameter t. It can also be noted that local gray line can be obtained from calculated local black and white centroids $\bar{X}_b, \bar{X}_w$ since they would most likely to lie on it. There was no significant difference between the results we obtained by using eigen-decomposition or local centroids. Table 4.4 shows the RMSE and maximum distances for the calibration image and three cameras. The color misalignments can be reduced upto 6 times, depending on the camera. The plot of color error around two disks of the pattern image for camera EOS 40D is shown on Figure 4.11 as gray level values (white is maximum, black is zero). After correction, the error magnitude for chromatic aberration is now comparable with noise level, as opposed to the plot before correction, where noise is masked by the amplitude of color error.

Camera	Uncorrected, $S_u(d_{u,max})$	Corrected, $S_c(d_{c,max})$	Ratio, $\frac{S_u}{S_c}(\frac{d_{u,max}}{d_{c,max}})$
EOS 5D			
Calib $f_1 = 24mm$	41.54 (78.21)	6.91 (22.87)	6.01 (3.41)
Test $f_1 = 24mm$	39.53 (75.55)	8.26 (26.93)	4.78 (2.80)
Calib $f_2 = 70mm$	36.18 (73.45)	5.50 (17.43)	6.57 (4.21)
Test $f_2 = 70mm$	29.87 (59.79)	5.46 (19.45)	5.46 (3.06)
EOS 40D			
Calib $f_1 = 18mm$	32.35 (66.07)	9.46 (34.08)	3.41 (1.93)
Test $f_1 = 18mm$	22.02 (46.97)	8.76 (27.55)	2.51 (1.70)
Calib $f_2 = 55mm$	36.48 (78.73)	7.35 (31.32)	4.96 (2.51)
Test $f_2 = 55mm$	21.91 (50.81)	5.93 (19.29)	3.69 (2.63)
A200			
Calib $f_1 = 18mm$	28.32 (76.43)	7.09 (58.45)	3.99 (1.31)
Test $f_1 = 18mm$	19.99 (50.18)	5.75 (24.45)	3.47 (2.05)
Calib $f_2 = 70mm$	28.51 (94.54)	7.22 (76.76)	3.94 (1.23)
Test $f_2 = 70mm$	13.08 (35.32)	7.10 (24.77)	1.84 (1.42)

Table 4.4: Table of quantitative error measures for color distribution: RMSE and maximal distances from cloud point to the local gray line.



(a) Error distances, uncorrected



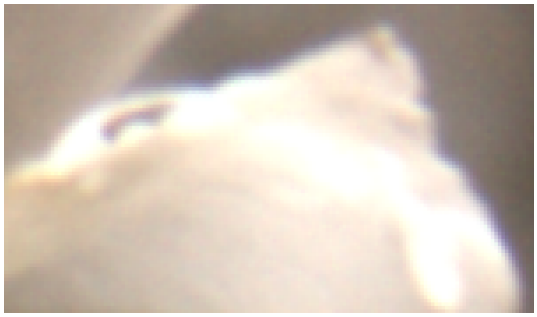
(b) Error distances, corrected

Figure 4.11: Error distances (eq.4.11) to the local gray line for the calibration image before the correction (a) and after (b) for camera EOS 40D.

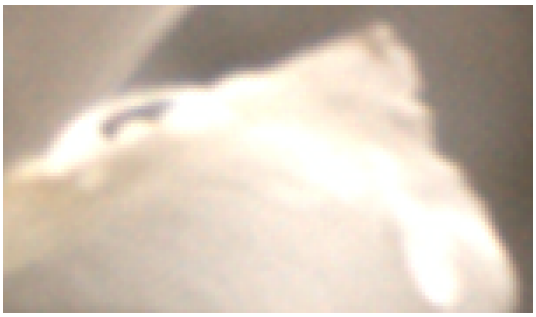
4.3.2 Visual improvement for real scenes

The fact that ellipse center precision is constant under the influence of blur (Figure 3.6 (b,d)) allows to take images of really far located objects like buildings by using small size calibration patter ($A3$ format). For this purpose, it is necessary to fix focal length and focus distance for farther distance and then take the calibration image with those settings. Obviously, the disks on the pattern will be smoothed out. But taken under presence of good illumination, their centers are detected with the same sub-pixel accuracy. To see the improvement in image quality, a zoomed-in crops of outdoor images are shown in Figure 4.12, Figure 4.13 and Figure 4.14. Red and blue fringes can be seen in original images, they disappear after correction.

(a) Test image



(b) Test image crop, uncorrected



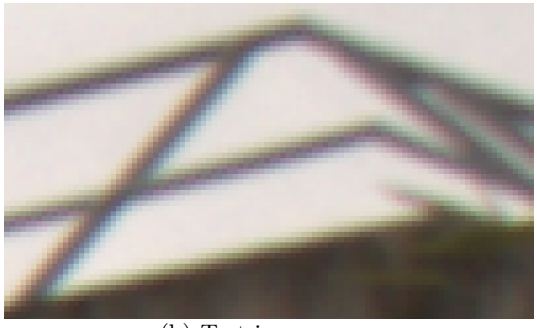
(c) Test image crop, corrected

Figure 4.12: Test image (a) and its crops before the correction (b) and after (c) for camera EOS 5D.

4.3.3 Experiments with compact digital cameras

The compact cameras always output images in JPEG format. Even when using the lossless type of compression, certain amount of information is still lost and can not be restored. Even if two image are taken under the same camera settings, it is not ensured that JPEG transformations applied to both will be identical due to different compositions. We decided to try our correction method for three digital cameras (Canon PowerShot SX100, Canon PowerShot A470 and Olimpus) on calibration circled pattern and see what kind of result we will obtain. Figure 4.15 displays cropped and zoomed in images before and after correction for the first compact

(a) Test image



(b) Test image crop, uncorrected

(c) Test image crop, corrected

Figure 4.13: Test image (a) and its crops before the correction (b) and after (c) for camera EOS 40D.

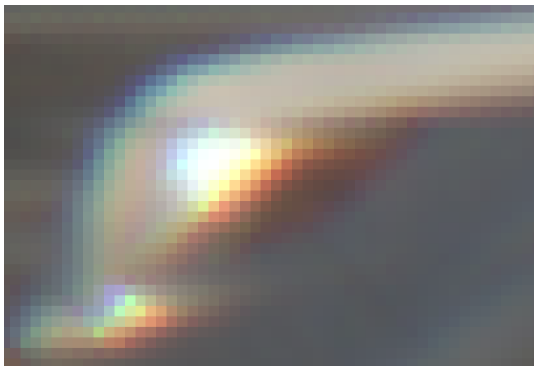
Canon camera. While the obtained geometrical misalignment were quite small (0.01-0.02 pixels), as it can be seen, the corrected image still have color fringes round the circle edge. However, the color is distributed around quite uniformly which simply means that JPEG compression results in different circle sizes, so, even if the disk centers are aligned along the channels, the aberration will take place due to the different circle sizes. This observation holds for all the tested compact cameras.

4.4 Conclusion

This chapter presents a high precision chromatic aberration correction method, using a single snapshot of a pattern of black disks. Disks' centers are used as keypoints and aligned between different channels. A dense correction vector field is then deduced by a general polynomial model which offers an alternative to existing radial-based models since no assumption must be made on the nature of aberration.

The calibration technique is easy to use, it is practical and is based on a single shot of a circled pattern. The correction is performed by re-sampling the originals based on the obtained calibrated model, and resulting into no aberration, high quality images. Numerical results expose high-precision correction leaving the corrected geometrical misalignments within 0.05 pixel range and improving the corrected color misalignments in $3-6$ times. Such magnitude of the result geometrical misalignment

(a) Test image

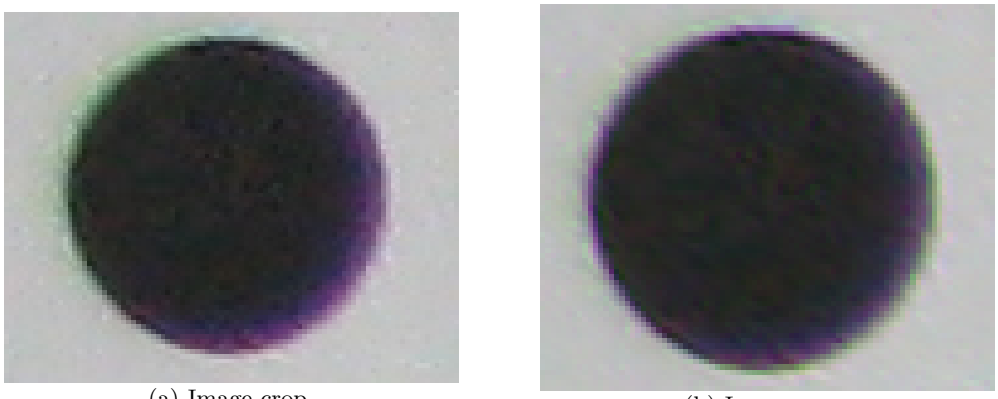


(b) Test image crop, uncorrected

(c) Test image crop, corrected

Figure 4.14: Test image (a) and its crops before the correction (b) and after (c) for camera A200.

also demonstrates the absence of aberration in sense of human visual perception, since it becomes visible at the magnitude of about 0.2 pixels. Real scenes examples are provided and corrected, their visual results demonstrate the elimination of the additional colors introduced by lateral chromatic aberration, thus producing quality color images.



(a) Image crop, uncorrected

(b) Image crop, corrected

Figure 4.15: Calibration image crops before (a) and after (b) the correction for compact camera Canon PowerShot SX100.

Chapter 5

Camera matrix calibration using circular control points

Abstract

The precise camera calibration using circular control points is achieved by, firstly, separation the lens distortion parameters from other camera parameters and calculation of the distortion field in advance by using the calibration harp. Secondly, in order to compensate for perspective bias which is prone to occur when using circled pattern, we incorporate conic affine transformation into minimization error when calculating the homography, and leave all the other calibration steps as they are used in the literature. Such error function allows to compensate for the perspective bias. Combined with precise keypoint detection, the approach showed to be more stable than existing state-of-art global calibration method.

Keywords Camera calibration, camera intrinsics, perspective bias, distortion bias, high precision calibration, planar homography, conic-based affine transformation.

Contents

5.1 Introduction

Camera calibration is the process of finding the true parameters of a camera given an image (or video), and it is a first step towards computational computer vision. Although some information concerning the measuring of scenes can be obtained by uncalibrated cameras [HARTLEY 1994], calibration is an essential step when metric information is required.

During past decades a lot of work had been done on subject of camera calibration for different kinds of applications – starting from photogrammetry [BROWN 1971, FAIG 1975], and more recent in computer vision, for example, dual space geometry based method [BOUGUET 1998a], flexible technique by [ZHANG 2000], plumb-line method [DEVERNAY 2001], plane-based calibration [STURM 1999, GURDJOS 2003], automatic calibration method by [JIANG 2005], pattern-free method [HARTLEY 2007]. The software packages are provided as well: calibration toolbox by [BOUGUET 1998b] and automatic bundle adjustment calibration software [PIERROT DESEILLIGNY 2011]. All of the methods allow camera calibration and 3D scene reconstruction with no or little human interaction.

Naturally, many works in camera calibration have focused especially on achieving high calibration accuracy and stability. These studies are mainly based on high precision control points of either 2D or 3D nature, and the accurate detection of their projections. The points are generally constructed on a planar surface by means of some high contrast pattern. Linear and least-square techniques for calibration are built upon [TSAI 1992] and [WENG 1992] who improve the calibration accuracy by thoroughly modelling lens distortion and further optimising with other parameters together.

For a real camera, an image of the calibration pattern is likely to be a subject of two types of transformations: a projective transformation as a result of relative 3D position, and a nonlinear transformation due to various lens distortions. The control point invariance to errors resulting from these two transformations is based on combination of the pattern employed, and the detection method used. Therefore, there exist two possible sources of bias in control point recovery which are named [MALLON 2007b]: *perspective bias* and *distortion bias*. The main goal would be to obtain bias free data, as this is clearly necessary for obtaining unbiased estimates for calibration algorithms.

The most famous planar calibration techniques were presented by [STURM 1999] and [ZHANG 2000]. They only require the camera to observe the pattern shown at a few (at least three) different orientations to get a unique solution up to scale factor. Either the pattern or camera are moved and the motion need not to be known. The pattern can be printed on a laser printer and attached to planar surface. The method places the world coordinate system on the calibration object and thus require only arbitrary scaled 2D coordinates. While trying to eliminate distortion bias during optimization, the mentioned works assume that the detected image points have zero mean Gaussian distributions in order to correctly converge to the optimal solution. However, it is not always the case. The bias does not have

the same magnitude for all types of patterns.

[MALLON 2007b] show that detected control points obtained by using centroid recovery principle, can potentially be corrupted by both perspective bias and distortion bias, with the likelihood of greater distortion bias magnitude in a typical camera. They also show that the compensation of distortion bias from such circular pattern points is not possible without knowing the distortion field. At the same time, Chapter 3 shows that most precise keypoints can be obtained by using circular points (conics), therefore we will proceed with them by aiming at how to eliminate distortion and perspective bias.

Regarding the distortion bias, most common approaches mix distortion parameters with other camera parameters (both intrinsic and extrinsic) and optimize them simultaneously. This could potentially lead to *residual error compensation*, thus, decreasing calibration stability. This error compensation can not be eliminated in the framework of global methods, and, therefore, compensation for distortion correction must be held separately, as a preliminary step to any calibration. Additionally, lens distortion introduces a nonlinear shape warping to the area of the conic and so it is no longer a true conic with biased center. The recently proposed *calibration harp* [TANG 2011] allows to solve distortion bias problem by compensation of the lens distortion field separately. Its main idea is based on straightness measure of stretched strings, pictures of which are taken in different orientations. More details, some minor improvements and results on this step are presented in Appendix A. Thanks to calibration harp, we can eliminate distortion bias from calibration calculation before the main optimization.

When considering perspective bias, the calibration accuracy can be improved by its correction, i.e. [HEIKKILÄ 2000] and [KANNALA 2006] which describe calibration techniques with use of circular points. Or, it can be avoided by using the projection of the conic contour, rather than its center [OUELLET 2008]. As an example, variation of Zhang's method is represented in [YANG 2000] where instead of point features three or more conics are used to obtain homography matrix and then deduce calibration parameters.

The main idea of elimination the perspective bias is based on the fact that the projection of the center of a circle does not correspond to the center of the resulting ellipse in the image. We compensate the perspective bias by taking into account circle-ellipse affine transformation and correspondence of detected keypoints with pattern keypoints. This is done by incorporating the conic affine transformation into minimization step for homography calculation. The transformation function matches center of the circle of the pattern with center of the ellipse in the image. The main feature still remains a conic center, therefore, no need for conic contour extraction, and the perspective bias is taken cared by the error function.

The Subsection 5.1.2 provides an overview and equations of the main calibration steps which are mainly based on the method of [ZHANG 2000], including the homography estimation. The modelling of its error function is given in detail in Section 5.2.

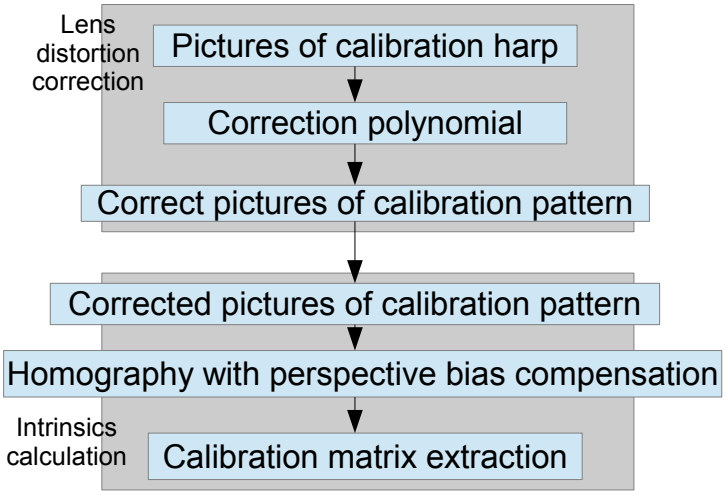


Figure 5.1: Camera calibration workflow: distortion parameters are calculated separately from other parameter.

5.1.1 Camera calibration workflow

Figure 5.1 demonstrates the sequence of main calibration steps. From there, it can be clearly seen the distortion correction phase is performed separately. The distortion compensation using calibration harp is explained in Appendix A.

5.1.2 Camera calibration basic steps and equations

We denote a $2D$ point as $m = [u, v]^T$. A $3D$ point is denoted by $M = [X, Y, Z]^T$. We use $\tilde{x}$ to indicate the augmented vector by adding 1 as the last element, i.e., $\tilde{m} = [u, v, 1]^T$ and $\tilde{M} = [X, Y, Z, 1]^T$. The camera model is usual pinhole and the relationship between a $3D$ point M and its image projection m is given by

$$s\tilde{m} = K \begin{bmatrix} R & t \end{bmatrix} \tilde{M}, \tag{5.1}$$

where s is an arbitrary scale factor, (R, t) are the extrinsic parameters (rotation and translation) which relates the world coordinate system to camera coordinate system, and K is the intrinsic camera matrix given by

$$K = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{5.2}$$

with (u_0, v_0) being the coordinates of the principal point, α and β the scale factors in image with respect to u and v axes, and γ the parameter describing the skewness of the image axes.

5.1.2.1 Homography between model plane and its image

The i^{th} column of the rotation matrix R is denoted by r_i. Assuming the model plane is located at $Z = 0$, we can write (5.1) as

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K \begin{bmatrix} r_1 & r_2 & t \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}. \tag{5.3}$$

Since Z is always equal to zero, the $3D$ point M will be denoted as $M = [X, Y]^T$ and its corresponding $\tilde{M} = [X, Y, 1]$ (same holds for m and $\tilde{m}$). The relation between model point M and image point m is described by homography H:

$$s\tilde{m} = H\tilde{M}, \tag{5.4}$$

or, written more precisely

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = H \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}, \tag{5.5}$$

with homography defined up to scale factor

$$H = K \begin{bmatrix} r_1 & r_2 & t \end{bmatrix}. \tag{5.6}$$

5.1.2.2 Estimation of the homography between the model plane and its image

The Direct Linear Transform (DLT) algorithm is a simple algorithm used to solve for the homography matrix H given sufficient number of point correspondences, namely 4 (explained in chapter 4.1 of [HARTLEY 2004]).

We represent matrix $H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$ and by dividing the first two rows of (5.5) by the third row, we get

$$\begin{aligned} -h_{11}X - h_{12}Y - h_{13} + (h_{31}X + h_{32}Y + h_{33})u = 0 \\ -h_{21}X - h_{22}Y - h_{23} + (h_{31}X + h_{32}Y + h_{33})v = 0. \end{aligned} \tag{5.7}$$

Eqs. (5.7) can be written in a matrix form as

$$A_i h = 0,$$

where $A_i = \begin{bmatrix} -X & -Y & -1 & 0 & 0 & 0 & uX & uY & u \\ 0 & 0 & 0 & -X & -Y & -1 & vX & vY & v \end{bmatrix}$ and $h = \begin{bmatrix} h_{11} & h_{12} & h_{13} & h_{21} & h_{22} & h_{23} & h_{31} & h_{32} & h_{33} \end{bmatrix}^T$.

Each point correspondence provides 2 equations, therefore, 4 correspondences are sufficient to solve for the 8 degrees of freedom of H. The only restrictions is that

no 3 points can be collinear. Four 2×9 A_i matrices (one for each correspondence) are stacked on top of one another to get an 8×9 matrix A, or if we deal with more than four points, the dimension for A will be $2n \times 9$. The solution space for h is right singular vector of A associated with smallest singular value.

In many cases more that 4 correspondences are used to ensure a more robust solution. A single and exact solution is possible only if all of the correspondences are exact, however, in practice there will always be some uncertainty (influence by noise), so there will be no exact solution. The problem then is to solve for a vector h that minimizes a chosen cost function.

Normalization. The DLT algorithm result presented above is dependent on the origin and scale of the coordinate system in the image as shown in chapter 4.4 of [HARTLEY 2004]. This is an undesirable property since it makes the algorithm unstable. The reason for this non-invariance is in how the DLT method uses SVD of A to obtain a solution to the over-determined set of equations $Ah = 0$ (explained in detail by [HARTLEY 1997a]). If we deal with exact data and infinite precision, the result is fine, but in the presence of noise the solution typically diverges from the correct result.

In order to ensure the numerical accuracy of the solution, [HARTLEY 2004] propose a normalization step. The main steps of DLT algorithm are then:

1. Compute a similarity transform T that takes points M_i to a new set of points M_i' such that the centroid of the points M_i' is the coordinate origin and their average distance from origin is $\sqrt{2}$.
2. Compute a similar transformation T' transforming points m_i to m_i'.
3. Apply the DLT agorithm from above using M_i' and m_i' to obtain homography matrix H'
4. Set $H = (T')^{-1} H' T$.

Precise estimation: dealing with noise. The homography estimation algorithm requires a set of correspondence as input. The real data always introduces some noise, therefore, it is necessary to take this into account. The estimate of the parameters is based on maximum likelihood criterion. If M_i and m_i are the model and image points respectively, then they should satisfy (5.4), but they do not for real data because of the noise in extracted image points. If assume m_i is corrupted by Gaussian noise with mean 0 and covariance matrix Λ_{m_i}, then maximum likelihood estimation of H is obtained by minimizing the following functional

$$\sum_i (m_i - \hat{m}_i)^T \Lambda_{m_i}^{-1} (m_i - \hat{m}_i), \tag{5.8}$$

where in the literature $\hat{m}_i$ is estimated as

$$\hat{m}_i = \frac{1}{\bar{h_3}^T M_i} \begin{bmatrix} \bar{h_1}^T M_i \\ \bar{h_2}^T M_i \end{bmatrix} \tag{5.9}$$

with $\bar{h_i}$, the i^{th} row of H. Later we will see that (5.9) cannot be applied for circled pattern due to perspective bias, and it will be replaced by applying homography to a conic, rather than to its center.

In practice it is assumed $\Lambda_{m_i} = \sigma^2 I$ for all i. This is reasonable if points are extracted independently with the same procedure. In this case the problem becomes a non-linear-least-square:

$$H = \arg\min_H \sum_i \|m_i - \hat{m}_i\|^2 . \tag{5.10}$$

The nonlinear minimization is conducted by using Levenberg-Marquardt Algorithm [MARQUARDT 1963]. It requires an initial guess which is estimated by DLT algorithm as shown at the beginning of this sub-section.

5.1.2.3 Constraints on the intrinsic parameters

If we write the homography column-wise as $H = \begin{bmatrix} h_1 & h_2 & h_3 \end{bmatrix}$, 5.6 will be

$$\begin{bmatrix} h_1 & h_2 & h_3 \end{bmatrix} = \lambda K \begin{bmatrix} r_1 & r_2 & t \end{bmatrix}, \tag{5.11}$$

where λ is an arbitrary scale. Using condition that r_1 and r_2 are orthonormal, we get

$$\begin{aligned} h_1^T K^{-T} K^{-1} h_2 &= 0 \\ h_1^T K^{-T} K^{-1} h_1 &= h_2^T K^{-T} K^{-1} h_2. \end{aligned} \tag{5.12}$$

Given one homography, these are two basic constraints on the intrinsic parameters. Because a homography has 8 degrees of freedom and there are 6 extrinsic parameters (3 for rotation and 3 for translation), we can only obtain 2 constraints on the intrinsic parameters.

5.1.2.4 Solving camera calibration

Let

$$B = K^{-T} K^{-1} = \begin{bmatrix} \frac{1}{\alpha^2} & -\frac{\gamma}{\alpha^2\beta} & \frac{v_0\gamma - u_0\beta}{\alpha^2\beta} \\ -\frac{\gamma}{\alpha^2\beta} & \frac{\gamma^2}{\alpha^2\beta^2} + \frac{1}{\beta^2} & -\frac{\gamma(v_0\gamma - u_0\beta)}{\alpha^2\beta^2} - \frac{v_0}{\beta^2} \\ \frac{v_0\gamma - u_0\beta}{\alpha^2\beta} & -\frac{\gamma(v_0\gamma - u_0\beta)}{\alpha^2\beta^2} - \frac{v_0}{\beta^2} & \frac{(v_0\gamma - u_0\beta)^2}{\alpha^2\beta^2} + \frac{v_0^2}{\beta^2} + 1 \end{bmatrix} \tag{5.13}$$

The matrix B is symmetric and so defined by a $6D$ vector

$$b = [B_{11}, B_{12}, B_{22}, B_{13}, B_{23}, B_{33}]^T. \tag{5.14}$$

If we note the i^{th} column vector of H as $h_i = [h_{1i}, h_{2i}, h_{3i}]^T$, then we can write

$$h_i^T B h_j = v_{ij}^T b \tag{5.15}$$

with

$$v_{ij} = [h_{1i}h_{1j}, h_{1i}h_{2j} + h_{2i}h_{1j}, h_{2i}h_{2j}, h_{3i}h_{1j} + h_{1i}h_{3j}, h_{3i}h_{2j} + h_{2i}h_{3j}, h_{3i}h_{3j}]^T. \tag{5.16}$$

Then the two fundamental constraints (5.12), from a given homography, can be rewritten as 2 homogeneous equations in b:

$$\begin{bmatrix} v_{12}^T \\ (v_{11} - v_{22})^T \end{bmatrix} b = 0. \tag{5.17}$$

If n images of the model plane are observed, by stacking n such equations as5.17 we have

$$Vb = 0, \tag{5.18}$$

where V is a $2n \times 6$ matrix. If $n \geq 3$, then there will be in general a unique solution b defined up to scale factor. In fact, the solution to (5.18) is well known as the eigenvector of V^TV associated with the smallest eigenvalue (or, the right singular vector of V associated with the smallest singular value).

Extraction of intrinsic parameters from matrix B. Once b is estimated, one can compute all the camera intrinsic parameters, the matrix K. The matrix B is estimated up to a scale factor, i.e., $B = \lambda K^{-T}K$ with λ an arbitrary scale. The parameters are uniquely extracted as follows:

$$\begin{aligned} v_0 &= \frac{B_{12}B_{13} - B_{11}B_{23}}{B_{11}B_{22} - B_{12}^2} \\ \lambda &= B_{33} - \frac{B_{13}^2 + v_0(B_{12}B_{13} - B_{11}B_{23})}{B_{11}} \\ \alpha &= \sqrt{\frac{\lambda}{B_{11}}} \\ \beta &= \sqrt{\frac{\lambda B_{11}}{B_{11}B_{22} - B_{12}^2}} \\ \gamma &= -\frac{B_{12}\alpha^2\beta}{\lambda} \\ u_0 &= \frac{\gamma v_0}{\beta} - \frac{B_{13}\alpha^2}{\lambda}. \end{aligned} \tag{5.19}$$

Extraction of extrinsic parameters. After K is obtained, the extrinsic parameters for each image are from (5.6) computed as follows:

$$\begin{aligned} r_1 &= \lambda K^{-1}h_1 \\ r_2 &= \lambda K^{-1}h_2 \\ r_3 &= r_1 \times r_2 \\ t &= \lambda K^{-1}h_3 \end{aligned} \tag{5.20}$$

with $\lambda = \frac{1}{\|K^{-1}h_1\|} = \frac{1}{\|K^{-1}h_2\|}$.

5.2 Incorporation of conic transform into homography estimation as perspective bias compensation

When we take a photo of a circle in $3D$, it becomes an ellipse in the projection plane. Both circle and ellipse shapes can be generalized as conics. A general conic is a curve which can be described by a second-degree equation on the plane. In $2D$ projective geometry all non-degenerate conics are equivalent under projective transformations. The equation of conics in inhomogeneous coordinates is

$$s_1x^2 + s_2xy + s_3y^2 + s_4x + s_5y + s_6 = 0 \tag{5.21}$$

that is, a polynomial of degree 2. When using homogeneous coordinates and replacing $x \to \frac{x_1}{x_3}$, $y \to \frac{x_2}{x_3}$, we obtain

$$s_1x_1^2 + s_2x_1x_2 + s_3x_2^2 + s_4x_1x_3 + s_5x_2x_3 + s_6x_3^2 = 0. \tag{5.22}$$

Using the notations for homogeneous coordinates $\tilde{M} = [X, Y, 1]^T$ and setting $X = x_1$ and $Y = x_2$, the matrix form of conics is

$$\tilde{M}^T S \tilde{M} = 0, \tag{5.23}$$

where the conic coefficient matrix S is given by

$$S = \begin{bmatrix} s_1 & s_2/2 & s_4/2 \\ s_2/2 & s_3 & s_5/2 \\ s_4/2 & s_5/2 & s_6 \end{bmatrix}. \tag{5.24}$$

The conic coefficient matrix is always symmetric and matrix S is its homogeneous representation.

Transformation of conics. Under a point transformation (5.4), (5.23) becomes

$$\tilde{m}^T H^{-T} S H^{-1} \tilde{m} = 0 \tag{5.25}$$

so the transformed conic or image of conic S is the conic

$$H(S) = H^{-T} S H^{-1}. \tag{5.26}$$

5.2.1 Center of conic's image vs. image of conic's center

Given (5.23) of a conic in homogeneous coordinates with S a 3×3 symmetric matrix, let's define its center by operator

$$C(S) = -S_{2\times2}^{-1} S_3, \tag{5.27}$$

with $S_{2\times2}$ the top left 2×2 block of S and S_3 the 2×1 top part of its last column:

$$S = \begin{pmatrix} S_{2\times2} & S_3 \\ S_3^T & \lambda \end{pmatrix}. \tag{5.28}$$

For conic S we can compute its image $H(S)$ as in (5.26) and now we wish to compare image of conic center $HC(S)$ and center of image conic $CH(S)$. We will see it will not be the same point and the difference will be quantified.

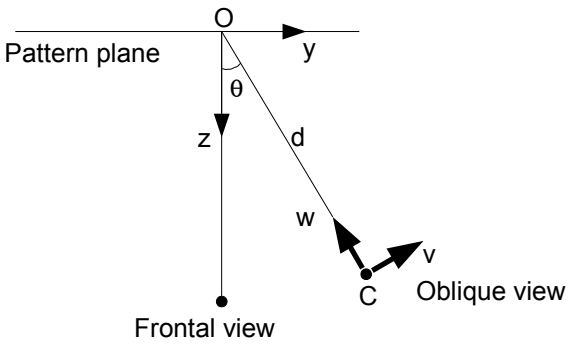


Figure 5.2: View of pattern plane along an angle θ.

Affine transformation. Let's take a specific case when we deal with large focal length and so H is an affine transformation ($h_{31} = h_{32} = 0$ and $h_{33} = 1$); we will have $HC(S)$ coincide with $CH(S)$. Indeed, we may write

$$H = \begin{pmatrix} H_{2\times 2} & H_3 \\ 0_2^T & 1 \end{pmatrix} \text{ and } H^{-1} = \begin{pmatrix} H_{2\times 2}^{-1} & -H_{2\times 2}^{-1} H_3 \\ 0_2^T & 1 \end{pmatrix}. \tag{5.29}$$

We then get

$$H(S) = \begin{pmatrix} H_{2\times 2}^{-T} S_{2\times 2} H_{2\times 2}^{-1} & -H_{2\times 2}^{-T} S_{2\times 2} H_{2\times 2}^{-1} H_3 + H_{2\times 2}^{-T} S_3 \\ \cdots & \cdots \end{pmatrix} \tag{5.30}$$

and therefore

$$\begin{aligned} CH(S) &= -(H_{2\times 2}^{-T} S_{2\times 2} H_{2\times 2}^{-1})^{-1} (-H_{2\times 2}^{-T} S_{2\times 2} H_{2\times 2}^{-1} H_3 + H_{2\times 2}^{-T} S_3) \\ &= H_3 - H_{2\times 2} S_{2\times 2}^{-1} S_3, \end{aligned} \tag{5.31}$$

which equals $HC(S)$.

At first order, the homography can be approximated by an affine transform, so the difference we get is second order which is still needed to be quantified.

Transformation for change of viewpoint. Suppose the calibration pattern is planar and the camera points at it from an angle $\theta \in [0, \pi/2)$ as on Fig. 5.2. Take as origin of the world coordinate system the intersection O of the principal ray of the camera with the pattern plane. Suppose also the camera has square pixels. Take the v-axis of the image in the plane determined by the principal ray and the normal vector to the pattern plane. We also take as origin in the image the principal point. Let us write d the distance of the optical center to the pattern plane along the principal ray.

In the front view at the same distance, we can write the projection matrix

$$P_0 = \begin{pmatrix} f & & \\ & f & \\ & & 1 \end{pmatrix} \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & d \end{pmatrix} \tag{5.32}$$

and in the oblique view

$$P = \begin{pmatrix} f & & \\ & f & \\ & & 1 \end{pmatrix} \begin{pmatrix} 1 & & & \\ & \cos\theta & -\sin\theta & \\ & \sin\theta & \cos\theta & d \end{pmatrix}. \tag{5.33}$$

Since the plane has equation $z = 0$, the homography between front view and oblique view is

$$H = K \begin{pmatrix} 1 & & \\ & \cos\theta & 0 \\ & \sin\theta & d \end{pmatrix} \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1/d \end{pmatrix} K^{-1} \sim \begin{pmatrix} 1 & & \\ & \cos\theta & \\ & \frac{\sin\theta}{f} & 1 \end{pmatrix}. \tag{5.34}$$

Given a circle with center $M = (X\,Y)^T$ and radius r in the front view, its matrix representation is

$$S = \begin{pmatrix} 1 & 0 & -X \\ 0 & 1 & -Y \\ -X & -Y & \|M\|^2 - r^2 \end{pmatrix}. \tag{5.35}$$

After computations, we get

$$CH(S) = \frac{1}{(1 + Y\frac{\sin\theta}{f})^2 - r^2\frac{\sin^2\theta}{f^2}} \begin{pmatrix} X(1 + Y\frac{\sin\theta}{f}) \\ Y\cos\theta(1 + Y\frac{\sin\theta}{f}) - r^2\frac{\cos\theta\sin\theta}{f} \end{pmatrix}, \tag{5.36}$$

whereas

$$HC(S) = \frac{1}{1 + Y\frac{\sin\theta}{f}} \begin{pmatrix} X \\ Y\cos\theta \end{pmatrix}. \tag{5.37}$$

Notice that the term $1 + Y\frac{\sin\theta}{f}$ vanishes for points in front view that map to the line at infinity of the plane (horizon).

Numerical results. For small r, the distance is proportional to r^2. If $r = 0$ (point), we get coinciding points, as for $\theta = 0$. Distance in pixels between $HC(S)$ and $CH(S)$ for different values of θ are displayed on Fig. 5.3.

5.2.2 Recovering homography by conic transform cost function

Considering (5.9), we can re-write it in the context of conic transformation, therefore we get

$$\begin{aligned} [\hat{u}, \hat{v}, \hat{w}]^T &= H\tilde{M}_i = HC(S_i) \\ \hat{m_i} &= \left[\tfrac{\hat{u}}{\hat{w}} \quad \tfrac{\hat{v}}{\hat{w}}\right]^T \end{aligned} \tag{5.38}$$

for the pattern point M_i which is a center of circle S_i. The detected 2D keypoint m_i which corresponds to the center of projected circle is

$$m_i = CH_0(S_i), \tag{5.39}$$

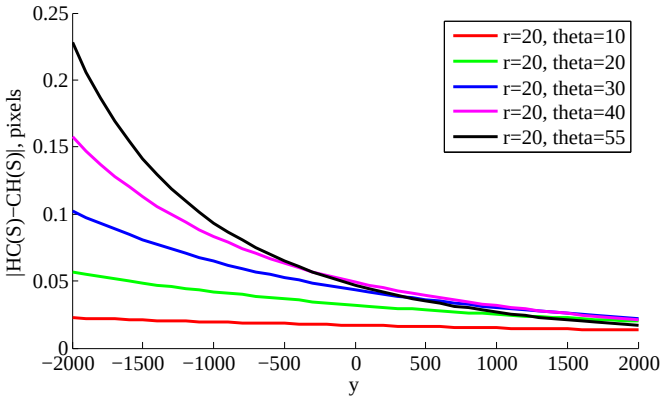


Figure 5.3: Difference in pixels between $HC(S)$ and $CH(S)$ for $f = 4000$, $x = 0$, $r = 20$ as a function of y for different angles θ (in degrees).

where H_0 is the ground truth homography. Therefore, (5.10) becomes

$$H_{\texttt{point}} = \arg\min_H \sum_i \|CH_0(S_i) - HC(S_i)\|^2 . \tag{5.40}$$

In order to compensate for perspective bias, we have to take into account conic transform, that is to minimize with respect to center of projected circle, not to projection of the circle center:

$$H_{\texttt{conic}} = \arg\min_H \sum_i \|CH_0(S_i) - CH(S_i)\|^2 . \tag{5.41}$$

As it can be seen, $H_{\texttt{conic}}$ does not require conic contour extraction, but only the operator which allows to extract conic center. In case of synthetic data, the performance can be evaluated by error function

$$E = \|H_0 C(S_i) - HC(S_i)\|^2 . \tag{5.42}$$

Levenberg-Marquardt algorithm is used to calculate the homography matrices. For simplicity, each new estimate is calculated using finite-difference, not Jacobian. The error function for conic based transform method is calculated as

$$e_i = \sqrt{\|m_i - CH(S_i)\|^2}. \tag{5.43}$$

5.3 Experiments

5.3.1 Homography estimation precision

We aim to compare the performance of $H_{\texttt{point}}$ (5.40) with $H_{\texttt{conic}}$ (5.41) against factors such as noise and view angle. To generate synthetic data, we use ground truth homography obtained by (5.34) with $f = 4000$. The model plane is represented by circle pattern and it consists of 10×14 circles, each of radius 20 pixels, and

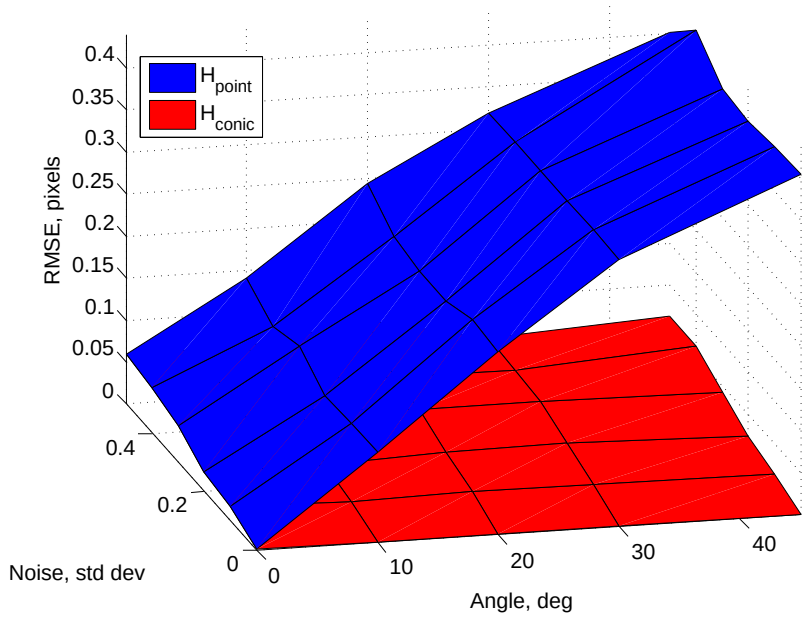


Figure 5.4: Comparison of two methods for homography estimation for different levels of noise and view angles. Blue color stands for $H_{\texttt{point}}$ and denotes a point-based feature minimization; the red color graph is for $H_{\texttt{conic}}$ – takes into account conic affine transform. Note the increasing difference of final RMSE due to perspective bias for point method when the angle increases.

that gives us 140 keypoints in total. The pattern is always rotated over axis y on a given angle view $[0, \cdots, 45]^\circ$. No physical image is generated, we only deal with keypoint coordinates which are obtained based on camera settings, for example, for circle S_i of the pattern, its projected image is $H(S_i)$, therefore, the extracted keypoint has coordinates $CH(S_i)$. Gaussian noise with 0 mean and standard deviation 0.5 pixels is added to the projected image keypoints $CH(S_i)$. For each noise level, 25 independent trials are performed and the results are displayed as an average.

Figure 5.4 shows a 3D comparison graph, where the error measure is a root mean square of distance described by (5.42). From now on, we denote conic based minimization as 'conic method' and standard minimization as 'point method'. As it can be seen from the graph, the conic method is invariant with respect to the change of view angle, which indicates that it does not suffer from perspective bias; on the contrary, the point method is prone to perpective bias.

In order to see the homography estimation performance for different camera views, we generate synthetic data based on synthetic camera with following settings: $\alpha = 1250$, $\beta = 900$, $\gamma = 1.09083$, $u_0 = 255$ and $v_0 = 255$. The result image has resolution of 512×512. The model plane is represented by circle pattern and it consists of 10×14 circles. The pattern size is set to $18cm \times 25cm$, and the radius of each circle is $\frac{1}{3}$ of the distance between consecutive circle centers. The camera is set at three views: $r_1 = [20^\circ, 0, 0]^T$, $t_1 = [-9, 12.5, 50.0]^T$, $r_2 = [0, 20^\circ, 0]^T, t_2 =$

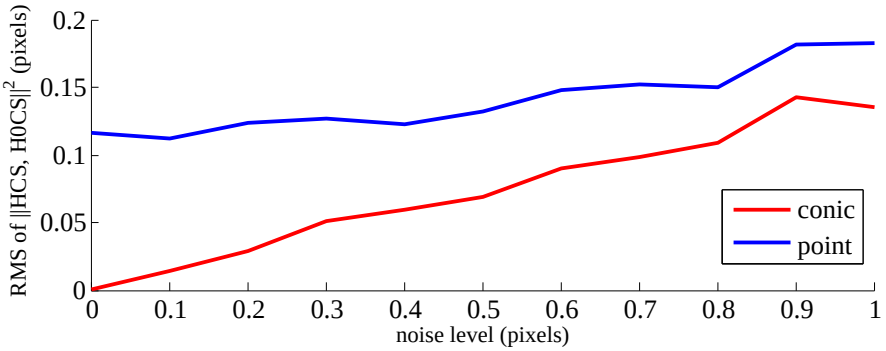
$[-9, -12.5, 51.0]^T$, $r_3 = \frac{1}{\sqrt{5}}[-30°, -30°, -15°]^T$, $t_3 = [-10.5, -12.5, 52.5]^T$. The estimated homographies are compared to the ground truth which is obtained as in (5.6). The noise level varies from 0 to 1 pixels. For each noise level, 25 independent trials are performed, and the results are displayed as average in Figure 5.5. It can be seen from the graph that homography matrix which takes into account conic transform is closer to ground truth (smaller residual) since it allows compensation for perspective bias.

5.3.2 Calibration matrix stability for synthetic data

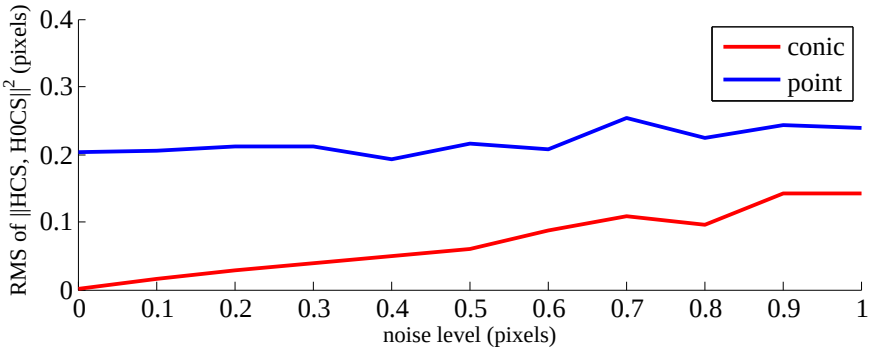
Pattern and camera views synthesis The image resolution is set to 1296×864 pixels. The number of circles is 10×14, each of radius $1cm$, and consecutive circles have $3cm$ separation between each other. The pattern has resolution $42cm \times 30cm$. The synthetic camera has following parameters: $\alpha = 1250$, $\beta = 1250$, $\gamma = 1.09083$, $u_0 = 648$, $v_0 = 432$. For the high quality images, we firstly generate high resolution pattern image and then subject it to the geometric displacement (all distortion is eliminated), Gaussian blurring and then down-sampling to the image resolution. Geometric image re-sampling is carried out by mapping from the transformed image to the original pattern. This involves calculating for every pixel in the transformed image, the corresponding pixel coordinate in the original image, which requires an inverse mapping. The transformed image intensity is then calculated based on the standard linear interpolation around the corresponding coordinate of the original pattern.

Pattern positioning In order to test calibration matrix stability, we generated 5 sets of images, each set included 5 images (different views on the pattern). This allows to extract 5 calibration matrices so that to see stability of its parameters along the sets. The generated image views are simulated by using pseudo-randomly generated homographies which consist of $3D$ rotation and translation whose values are drawn randomly from a specific range. This range limit ensures that the transformed image lies roughly within an image window. Meanwhile, there is always variance of rotations and translation along the sets. The roation angles always lie whithin the range $[15°, -45°]$. For the re-sampling of the transformed image, an inverse of homography matrices is used.

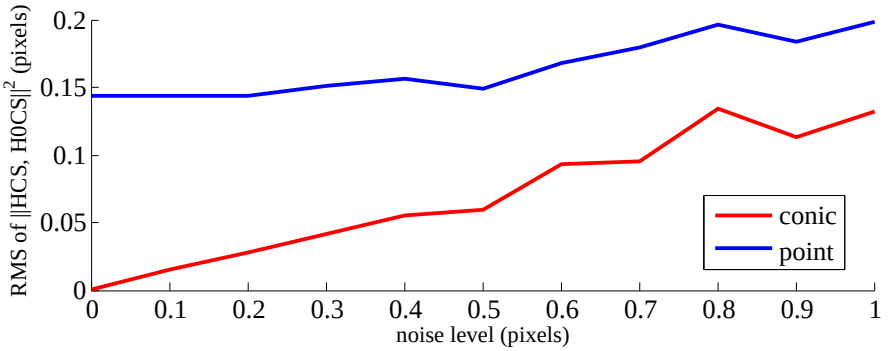
K matrix stability To compare our method with state-of-art method, we chose Software Package for Precise Camera Calibration [HIGUCHI 2012] which is built upon Matlab Calibration Toolbox [BOUGUET 2000] with the difference that circle pattern can be used for the calibration, plus some minor improvements of the original software. The generated images are treated as distortion free for both software. Table 5.1 provides the standard deviation results for both methods (we denote state-of-art method as 'SoA'). As we had set both noise and distortion to zero, we can clearly see the improvement in calibration stability based on having more precise homography matrices for our method.



(a) Camera view 1



(b) Camera view 2



(c) Camera view 3

Figure 5.5: Comparison of homography precision for two methods and three camera views (a), (b) and (c) against different noise levels.

Parameter	SoA	ours
dev(α)	0.10	0.008
dev(β)	0.11	0.008
dev(u_0)	0.02	0.006
dev(v_0)	0.14	0.014

Table 5.1: Standard deviations of camera calibration parameters for five image sets – comparison of the state-of-art method with ours. Any noise and distortion are eliminated. Thanks to improved precision of homography calculation, we notice less deviation for our method.

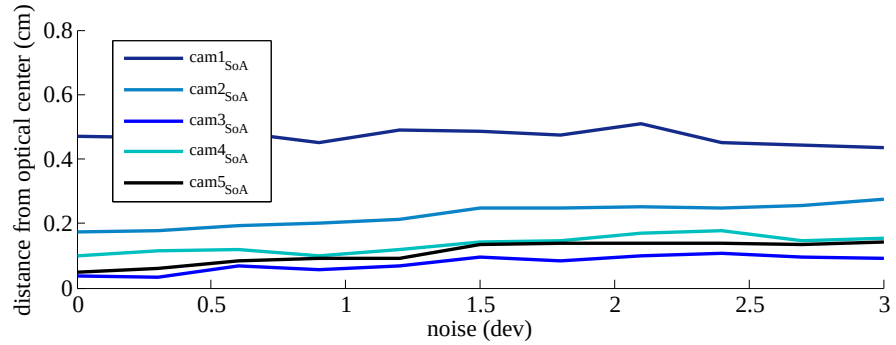
Deviation from camera center Calculated homography and camera matrix allow to obtain 3D coordinates of the camera center. If the ground truth is known, then it is possible to get an error measure in centimeters (the pattern unit for our synthetic tests). For this, we extract rotation and translation parameters, using homography matrix as described by (5.20). For the camera center ground truth C_0, it is easy to deduce from the ground truth rotation and translation by inverting them as

$$C_0 = -R_0^{-1} T_0. \tag{5.44}$$

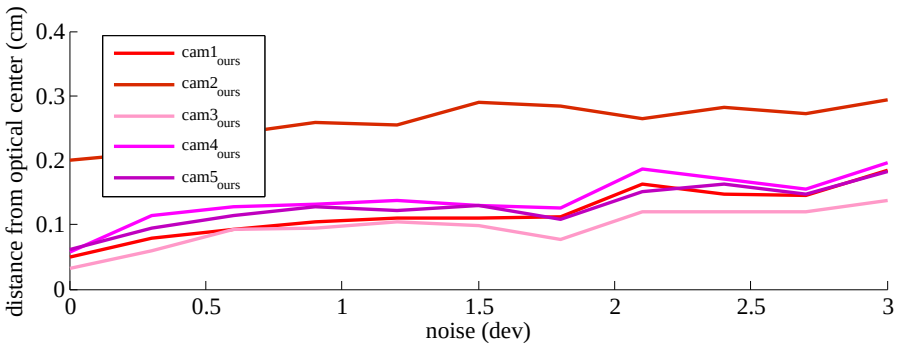
For each method, we extract the R and T parameters from the homography, and then we use eq (5.44) to calculate camera centers for the two methods, so as to compare the magnitude of the deviation from the ground truth. The Euclidean distances from C_0 to the obtained centers for each camera view are displayed on Figure 5.6: (a) is for the state-of-art method for each camera view against different noise level, (b) is for our method for each camera view against the noise level, and on (c) the average of each method is plotted one against another for different noise level. As expected, the graphs show that our method has lower magnitude residual than for the state-of-art global calibration method.

5.3.3 Calibration matrix stability for real data

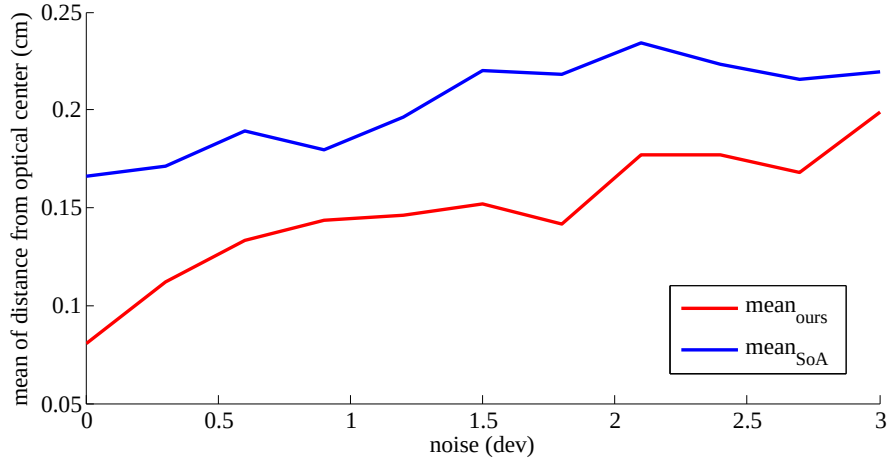
The experiments for real data were performed for the Camera Canon EOS40D with lens Canon EF 18-55mm. Several tests were performed – for different focal lengths. For comparison we use same state-of-art software as in synthetic tests. In order to treat the distortion for our method, it was necessary to use two different patterns – calibration harp and circled pattern (the main workflow is shown in Figure 5.1). The input images for state-of-art method remained the same as input images for our method for the extraction of calibration matrix, but not distortion-compensated. For each focal lengths we took 6 datasets, each contained 5 images of circled pattern under different camera orientation. K matrix was extracted for each set and then standard deviation was taken so as to measure the result stability. The comparison of two method is shown in Table 5.2. As it can be seen from the table, our method



(a) Software of [HIGUCHI 2012]



(b) Our method software



(c) Comparison of the mean values of the two methods

Figure 5.6: Euclidean distance from the ground truth optical center, in cm, to the obtained center for [HIGUCHI 2012] (a), our method (b) for the used cameras and their mean values for the two methods (c). Experiments are displayed against different noise level introduced in the synthetic images.

Parameter	f18mm		f27mm		f55mm	
	SoA	ours	SoA	ours	SoA	ours
$\texttt{dev}(\alpha)$	0.583	0.245	2.394	1.727	8.810	2.668
$\texttt{dev}(\beta)$	0.547	0.175	2.360	1.721	7.644	2.641
$\texttt{dev}(u_0)$	0.494	0.135	1.732	0.840	3.565	1.297
$\texttt{dev}(v_0)$	0.787	0.332	1.273	0.935	2.845	1.314

Table 5.2: Standard deviations of the camera Canon EOS5D parameters for real data – comparison state-of-art method (noted as 'SoA') with ours (noted as 'ours').

was able to achieve more stable results since deviation is smaller than for state-of-art method. The results validate an assumption that separation of distortion from other camera parameters help to avoid residual error compensation by leading to more precise camera calibration.

5.4 Conclusion

In order to improve the precision of camera calibration, we first performed a detachment of lens distortion from other camera parameters. As an advantage, it helps to avoid residual error compensation which is inherent for global calibration methods and which cannott be solved in its framework. As a disadvantage, it requires building and using of additional calibration pattern – named calibration harp for the distortion calculation and only then standard circled pattern – for the camera calibration. However, after the distortion is calculated for the fixed camera settings, we can treat the processed images as distortion-compensated, and this allows to calculated other camera parameters separately from distortion; as well as using the circled pattern which is prone to distortion bias, and, therefore, using high-precision keypoint detector.

The second aspect that was addressed was a compensation of the perspective bias for the circled pattern which was achieved by incorporation of conic affine transformation into homography estimation. The function serves as a compensator and at the same time it allows avoiding the use of conic contour detection, the main feature still remains a conic centroid. Of course, this would not be possible without advance distortion compensation, since distortion bias has much higher magnitude than the perspective bias, and elimination is only possible if the distortion field is known in advance, which is possible using calibration harp.

The numerical results for both undistorted synthetic and distorted real data demonstrated that our method allows to get more stable results for camera calibration parameters, meaning lesser magnitude of parameters deviations.

Chapter 6

Thesis conclusions

This thesis had studied precision issues for two problems in internal camera calibration: chromatic aberration correction and camera matrix extraction. Even though it could seem both problems are quite fundamental, the precision aspects were not studied as extensively.

A simple human perception experiment had stated that chromatic aberration in images is not noticeable if the magnitude of color channel geometric misalignment is less than 0.1 pixels. This analysis defined our goal precision for the geometrical color channel misalignment. At the same time most of the existing approaches use different types of radial polynomial to model the geometrical misalignments that is not general enough and cannot provide satisfiable correction for all camera types.

For our method, we combine high-precision keypoint detection of the pattern plane which was explained in Chapter 3 with the bivariate polynomial of degree 11 which serves as a correction model for color plane misalignment. Our method proved to be the most effective among existing state-of-art software - as a result it always leaves final geometric residual less than 0.1 pixels which was our precision goal. The calibration technique itself is easy to use, it is practical and is based on a single shot of circled pattern. It is performed by re-sampling the originals based on the obtained calibrated model, and as a result it allows to obtain no aberration, high quality images.

We also measured the result images in terms of color metric which is based on the color distance from local gray line. It was found this distance may decrease 3-6 times, depending on the camera type.

Real scenes examples taken by three different reflex cameras are provided, calibrated and corrected, their visual results demonstrated the elimination of the color artefacts introduced by lateral chromatic aberration.

Compact digital cameras which can only output images in JPEG format were tested as well. Even though the final geometric misalignment was smaller than 0.1 pixels, the method was not efficient in terms of visual evaluation: the disk size for calibration images taken by a compact camera was different for each color plane, probably, due to compression, therefore, even if the disk centers are aligned, the aberration will take place because of the different disk sizes. From this, we conclude that high precision aberration correction is only possible for RAW images taken by reflex cameras.

One of the future directions could be incorporation and evaluation the precision of optical flow technique and compare it to the current method performance.

Camera calibration is considered to have been solved since a long time. The literature shows a trend towards global camera calibration method from set of images - when optical distortion model is estimated as its part, together with internal and external camera parameters in a single minimization step. Even the precision and stability of the extracted parameters had always been considered enough for practical applications, it was not possible to measure carefully the precision of each set of parameters. Moreover, as it was explained, the main drawback of global camera calibration method is a potential for error compensation.

The solution we used lies in separation of the distortion correction from the camera calibration to avoid error compensation. For the optical distortion correction we used recently developed calibration harp, the main principle of which is described in Appendix A with some minor improvements.

Another issue which was addressed is how to eliminate perspective bias when using circular control points. It was achieved by incorporating of conic affine transformation into homography estimation step which matches center of pattern circle with center of image ellipse, while leaving other calibration steps as it is described in the literature.

As our experiments showed, the conic transform function allows to obtain more precise homography, thus, leading to more stable calibration results. Both synthetic and real experiments showed more stable results for camera calibration which means smaller deviation of camera parameters within several sets of images.

The limitation of the proposed method is that we have to deal with two different calibration patterns which can lead to incremental error, while global calibration requires only one pattern. However, as a result it allows us to directly process the optical distortion before proceeding to camera calibration and thus avoid residual error compensation.

As a future work, it would be interesting to expand experimental dataset (for example, different reflex cameras), and also use a calibration pattern which is not printed on paper and attached to the wall, but rather glued on aluminium sheet. The later could allow to avoid having non-flat areas which are difficult to avoid using a paper pattern.

Other unanswered questions remain, for example, how is it possible to assess the distortion correction by using the circles calibration patter. One possible solution could lie in using calculated homographies. Another questions with regard to compensation of parameter - if there is a way to quantify the occurred compensation. And the third important question - whether or not we achieved the limits for precision.

If considering application future direction, the camera calibration is only one part of the 3D reconstruction chain in high precision. After internal and external camera parameters are extracted, the reconstructed 3D scene can be cross-validated by registering several 3D point clouds reconstructed from different view images. Such precision specification could be a preliminary step to perform 3D points merging. The precision of the obtained data can be compared with precision of the

global calibration method vs. groundtruth data to check the precision gain using non-global calibration method.

Appendix A

High-precision lens distortion correction using calibration harp

The content of this chapter is mainly based on the work of [TANG 2011] and is a preliminary part of calibration process described in Chapter 5. Some minor alterations were incorporated in order to improve results, additionally, the method was implemented on C++ language.

Keywords Optical distortion, lens distortion, error compensation, plumb lines, polynomial model.

Contents

A.1 Introduction

When we refer to a notion *high precision distortion correction*, it is assumed that residual error between the camera and its obtained numerical model is far smaller than a pixel size. The problem seems to be solved adequately by global calibration methods, i.e. flexible pattern based method [ZHANG 2000], plumb-line [DEVERNAY 2001], pattern-free [HARTLEY 2005], dual-spaced geometry based method [BOUGUET 1998b, BOUGUET 1998a], automatic bundle adjustment calibration software [PIERROT DESEILLIGNY 2011], as well as many others [SLAMA 1980, TSAI 1992, WENG 1992, LAVEST 1998] that estimate internal and external parameters simultaneously. For example, the [LAVEST 1998] method is based on measurement the non-flatness of a pattern and results in a remarkably small reprojection error - 0.02 pixels. For the computer vision goals this precision would be sufficient. Yet, according to a measurement tool of lens distortion correction precision developed by [TANG 2011] which is based on plumb line concept as distortion measure, the straightness criterion gives an RMSE as large as 0.2 pixels, which is a contradiction to the obtained 0.02 pixels re-projection accuracy.

This inconsistency have a meaning that, in the global optimization process, errors in the external and internal camera parameters are being compensated by opposite errors in the distortion model. Therefore, an inaccurate distortion model may pass undetected since it can not capture the real physical aspect of real lens distortion. Such facts raise an objection to simultaneous estimation of the lens distortion and camera parameters.

Separation of those parameters one from each other will not only allow to capture the distortion field more precisely, but will also perform more precise calibration of external parameters like it was shown in Chapter 5 since it can be done independently from distortion.

Let's consider the common camera model

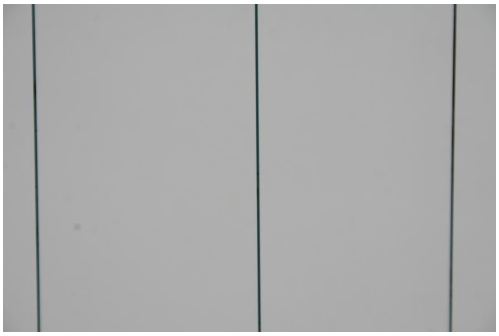
$$C = DKR[I| - T] \tag{A.1}$$

with D non linear operator for lens distortion, K calibration matrix, R camera orientation matrix and vector T the camera optic center in a fixed world frame. Given $3D$ point is first transformed in the camera-based frame by the translation T then rotation R. After, it is projected into the image plane by K, followed by the non-linear lens distortion D. According to this model, camera calibration consists of two parts:

- D and K - internal parameters calibration
- R and T - external parameters calibration

The first parameters should remain constant once the camera is fixed in its configuration. But in the experiments [TANG 2011] they vary from one test to another, even if the same camera settings were used. Thus, the internal calibration is not always reusable for the other datasets. And this is a common drawback for many

(a) The calibration harp



(b) Close-up of calibration harp

Figure A.1: The calibration harp (a) which serves for measuring and correction of lens distortion. (b) shows a close-up of the harp.

global camera calibration methods that calculate internal and external parameters simultaneously. This is why it was decided to address the distortion correction separately and, actually, as a preliminary step to any further calibration.

Let's assume we are given *non-parametric pattern-based method* to achieve a hight precision. For this purpose, a pattern would require a very flat non-deformable surface with highly accurate patterns printed on it. [TANG 2011] showed that a relatively small $100\mu m$ flatness error can nonetheless introduce a 0.3 pixels error in distortion estimation. Fabrication of a sizeable (half meter or more is necessary for camera calibration) very flat pattern, or event to validate their flatness, is a quite difficult physical and technical task. This is why it is suggested a *plumb lines method* [BROWN 1971] for which we can ensure some serious straightness by tighten strings. A *calibration harp* [TANG 2011] as shown on Figure A.1 serves well for this purpose - it is used as both correction and validation tool for lens distortion problem.

The choice of good distortion model associated with the plumb-line method is as well important: it must be able to capture different types of realistic lens distortions. For this purpose, the polynomial model [KILPELÄ 1980] seems more adapted to correct real distortion [TANG 2011] by linear approximation - whether it is complex radial or non-radial, provided its degree is high enough. In addition, it is invariant to any translation to the distortion center, therefore, it can be fixed anywhere without being estimated. It is an advantage if compare to other models such as radial [BROWN 1971], rational [HARTLEY 1997b, MA 2004, CLAUS 2005], division [FITZGOBBON 2001], FOV [DEVERNAY 2001].

The calibration procedure consists of taking photographs of different orientations to estimate the best coefficients of the polynomial model so that to correct the lens distortion. The method is a direct non-parametric, non-iterative and model-free. By non-parametric and model-free it means that the distortion model allows for any diffeomorphism. The obtained precision compares favorably to the distortion given by state-of-art global calibration and reaches RMSE of 0.02 pixels for calib-

rated data. The rest of the appendix is organizes as follows: Section A.2 provides an overview on the calibration method, giving more details on the minor simplifications that were incorporated for the software implementation. The Chapter A.3 shows real data experiments and results that we obtained as a conjunction of the experiments in Chapter 5.

A.2 The harp calibration method

A.2.1 Main equations

The main idea is to combine the advantage of plumb-line methods with the universality of polynomial approximation. The plumb-line method consists in correcting the distorted points which are supposed to be on a straight line, by minimizing the average squared distance from the corrected points to their corresponding regression lines.

Polynomial model It has the form

$$\begin{aligned} x = b_0 x_d^p + b_1 x_d^{p-1} y_d + b_2 x_d^{p-2} y_d^2 + \cdots + b_p y_d^p + b_{p+1} x_d^{p-1} + b_{p+2} x_d^{p-2} y_d + \cdots \\ + b_2 y_d^{p-1} + \cdots + b_{\frac{(p+1)(p+2)}{2}-3} x_d + b_{\frac{(p+1)(p+2)}{2}-2} y_d + b_{\frac{(p+1)(p+2)}{2}-1} \\ y = c_0 x_d^q + c_1 x_d^{q-1} y_d + c_2 x_d^{q-2} y_d^2 + \cdots + c_q y_d^q + c_{q+1} x_d^{q-1} + c_{q+2} x_d^{q-2} y_d + \cdots \\ + c_2 y_d^{q-1} + \cdots + c_{\frac{(q+1)(q+2)}{2}-3} x_d + c_{\frac{(q+1)(q+2)}{2}-2} y_d + c_{\frac{(q+1)(q+2)}{2}-1} \end{aligned} \tag{A.2}$$

with (x, y) undistorted point, and (x_d, y_d) distorted point. This polynomial approximation is translation invariant, and the origin is arbitrarily fixed at the image center. The orders for the x and y components are respectively p and q. The number of parameters for x and y is respectively $\frac{(p+1)(p+2)}{2}$ and $\frac{(q+1)(q+2)}{2}$. If orders are set so that $p = q = 3$, such model is called bicubic.

Regression line. The integration of the polynomial model into the plumb-line method is done through the calculation of the linear regression line. Given a set of corrected points $(x_i, y_i)_{i=1,\cdots,N}$ which are supposed to be on a line, at first we compute the linear regression line which has a form

$$\alpha x_i + \beta y_i - \gamma = 0 \tag{A.3}$$

with $\alpha = \sin\theta$, $\beta = \cos\theta$, $\tan 2\theta = -\frac{2(A_{xy} - A_x A_y)}{V_{xx} - V_{yy}}$, $A_x = \frac{1}{N}\sum_{i=1}^N x_i$, $A_y = \frac{1}{N}\sum_{i=1}^N y_i$, $A_{xy} = \frac{1}{N}\sum_{i=1}^N x_i y_i$, $V_{xx} = \frac{1}{N}\sum_{i=1}^N (x_i - A_x)^2$, $V_{yy} = \frac{1}{N}\sum_{i=1}^N (y_i - A_y)^2$, and $\gamma = A_x \sin\theta + A_y \cos\theta$. The sum of squared distances from the points to regression line is $\sum_{i=1}^N (\alpha x_i + \beta y_i - \gamma)^2$.

Measuring the straightness The straightness of a line is measured as the root-mean-square (RMS) distance from its edge points to its global regression line. Considering (A.3), for G groups of lines, with L_g the numbers of lines in group g, and N_{gl} the number of points of line l in a group g, a total sum of squared distance is

$$S = \sum_{g=1}^{G} \sum_{l=1}^{L_{gl}} \sum_{i=1}^{N_{gl}} (\alpha_g x_{gli} + \beta_g y_{gli} - \gamma_{gl})^2. \tag{A.4}$$

After, the straight measure is computed as

$$RMSE = \sqrt{\frac{S}{N}}, \tag{A.5}$$

where N - the total number of points $N = \sum_{g=1}^{G} \sum_{l=1}^{L_g} N_{gl}$.

Minimization and linear estimation For the sake of conciseness, we assume that we deal with a bicubic model with $p = q = 3$. Combining (A.2) and (A.4), the energy S becomes

$$S = \sum_{g=1}^{G} \sum_{l=1}^{L_{gl}} \sum_{i=1}^{N_{gl}} (\alpha_g (b_0 x_{d_{gli}}^3 + \cdots + b_9) + \beta_g (c_0 x_{d_{gli}}^3 + \cdots + c_9) - \gamma_{gl})^2. \tag{A.6}$$

It is a non-linear problem to minimize the energy S with respect to the parameters $b_0, \cdots, b_9, c_0, \cdots, c_9$. It is performed by an iterative Levenberg-Marquardt (LM) algorithm with estimation the parameters of polynomials in increasing order, which means that the algorithm starts estimating the parameters of a 3-order polynomial, then the output is used to initialize the 4-order polynomial, and the process continues until 11-order (the choice of the polynomial degree is explained in experimental section).

After non-linear minimization step, the parameters α_g and β_g are estimated, therefore, we can iteratively refine the precision - since the problem becomes linear. The line orientations are first initialized by the orientation of the regression lines. By differentiating S with respect to each parameter, we obtain a linear system

$$At = 0 \tag{A.7}$$

with $t = (b_0, \cdots, b_9, c_0, \cdots, c_9, \gamma_{1l}, \cdots, \gamma_{GL_G})^T$ and the coefficient matrix

$$A = \begin{bmatrix} A_b \\ A_c \\ A_\gamma \end{bmatrix}, \tag{A.8}$$

which composed of 3 submatrices. The iteration is repeated until the result stabilizes, or the required precision is reached.

To obtain a unique non-trivial solution, we always set $b_7 = c_8 = 1$, $b_9 = c_9 = 0$ which fixes scale and translation to the solution. The minimized S can be changed by the introduced scale, but this change is consistent of the distortion center and b_9, c_9, b_7 and c_8 are fixed.

Implementation and computational time It is straightforward to implement the method with iterative LM minimization and then linear estimation. However, we had encountered the problem of low computational time due to the calculation of derivatives to each of the variable. A simplification of the formulas presented below helped to solve this problem by reducing the time from more than a minute to seconds. It is mainly based on extraction of the value for line inclination angle, θ, as well as its sine and cosine based on the retrieved statistical values (average and variance). This allows to simplify the minimization formula and get rid from the calculation of γ coefficients.

A.2.2 A solution for the line angle θ

Let's consider an individual line for which we aim to define an inclination angle. The angle is defined as a slope of the regression line, the squared distances sum of which we have to minimize

$$S = \frac{1}{2}\sum_i (x_i \sin\theta + y_i \cos\theta - \gamma)^2 \tag{A.9}$$

Taking partial derivatives to θ and γ yields

$$\begin{aligned} &\frac{\partial}{\partial\theta} : \sum_i (x_i \cos\theta - y_i \sin\theta)(x_i \sin\theta + y_i \cos\theta - \gamma) = 0 \\ &\frac{\partial}{\partial\gamma} : -\sum_i (x_i \sin\theta + y_i \cos\theta - \gamma) = 0 \end{aligned} \tag{A.10}$$

Therefore, we can formulate γ as

$$\gamma = \bar{x}\sin\theta + \bar{y}\cos\theta \tag{A.11}$$

where $\bar{x} = \frac{1}{N}\sum_i x_i$ and $\bar{y} = \frac{1}{N}\sum_i y_i$. After we can deduce the angle θ through following steps using obtained equations:

$$\begin{aligned} &\sum_i (x_i \cos\theta - y_i \sin\theta)((x_i - \bar{x})\sin\theta + (y_i - \bar{y})\cos\theta) = 0 \\ &\sum_i x_i(y_i - \bar{y})\cos^2\theta - \sum_i y_i(x_i - \bar{x})\sin^2\theta + \langle \sum_i x_i(x_i - \bar{x}) - \sum_i y_i(y_i - \bar{y})\rangle \cos\theta \sin\theta = 0 \end{aligned} \tag{A.12}$$

The double-angle formulas for sin and cos simplifies it to

$$(\sum_i x_i y_i - N\bar{x}\bar{y})\cos 2\theta + \frac{1}{2}\langle \sum_i {x_i}^2 - N\bar{x}^2 - \sum_i \bar{y_i}^2 + N\bar{y}^2\rangle \sin 2\theta = 0. \tag{A.13}$$

By putting $\tilde{x}_i = x_i - \bar{x}$ and $\tilde{y}_i = y_i - \bar{y}$, the derivative to θ from (A.10) can be written as

$$\frac{\partial}{\partial\theta} : \sum_i \tilde{x}_i\tilde{y}_i \cos 2\theta + \frac{1}{2}(\sum_i \tilde{x}_i^2 - \sum_i \tilde{y}_i^2)\sin 2\theta = 0 \tag{A.14}$$

There are two possible solutions which we denote as θ_0 and θ_1 and which are related as $\theta_1 = \theta_0 + \frac{\pi}{2}$, but only one of them that provides a minimum will be of the interest:

$$\begin{aligned} 2\theta_0 &= atan2(-2\sum_i \tilde{x}_i\tilde{y}_i, \sum_i \tilde{x}_i^2 - \sum_i \tilde{y}_i^2) \\ 2\theta_1 &= atan2(2\sum_i \tilde{x}_i\tilde{y}_i, -(\sum_i \tilde{x}_i^2 - \sum_i \tilde{y}_i^2)) \end{aligned} \tag{A.15}$$

To define which of them is a solution, a second derivative to θ is taken:

$$\begin{aligned} \frac{\partial^2 S}{\partial\theta^2} &= -2\sum_i \tilde{x}_i\tilde{y}_i \sin 2\theta + (\sum_i \tilde{x}_i^2 \sum_i \tilde{y}_i^2)\cos 2\theta \\ &= k\sin 2\theta_0 \sin 2\theta + k\cos 2\theta_0 \cos 2\theta = k\cos 2(\theta - \theta_0) \end{aligned} \tag{A.16}$$

where k is a normalization constant

$$k = \left| \begin{pmatrix} \sum_i \tilde{x}_i^2 - \sum_i \tilde{y}_i^2 \\ -2\sum_i \tilde{x}_i\tilde{y}_i \end{pmatrix} \right|, \tag{A.17}$$

which is always positive $k \geq 0$.

From (A.16) we can write

$$\begin{cases} \text{for } \theta = \theta_0 \text{ we get } k(\geq 0) \\ \text{for } \theta = \theta_1 \text{ we get } -k(\leq 0) \end{cases} \tag{A.18}$$

The first solution for θ_0 represents a minimum in which we are interested in, since the (A.16) is a second derivative and the first solution of (A.15) is an according concave-up. Thus, the only one solution for the line angle

$$\theta = \frac{1}{2} atan2(-2\sum_i \tilde{x}_i\tilde{y}_i, \sum_i \tilde{x}_i^2 - \sum_i \tilde{y}_i^2) \tag{A.19}$$

Moreover, by taking into consideration the above notations for $\tilde{x}$ and $\tilde{y}$, as well as the (A.11) for γ, now the squared distance sum to minimize, (A.9), will have a simpler look of

$$S = \sum_i (\tilde{x}_i \sin\theta + \tilde{y}_i \cos\theta)^2 \tag{A.20}$$

A.2.3 Simplified minimization for obtaining the polynomial coefficients

For one line with number of points $i = 1, \cdots, N$, given a set of distorted points $(x_{d,i}, y_{d,i})$ and a set of corresponding corrected points (x_i, y_i) which are supposed to lie on a linear regression line

$$\alpha\tilde{x}_i + \beta\tilde{y}_i = 0 \tag{A.21}$$

where $\tilde{x}_i = x_i - \bar{x}$, $\tilde{y}_i = y_i - \bar{y}$ with $\bar{x} = \frac{\sum x_i}{N}$ and $\bar{y} = \frac{\sum y_i}{N}$; and $\alpha = \sin\theta$, $\beta = \cos\theta$. The previous subsection gives a description for obtaining an angle θ for the selected line. We denote $V_{xx} = \frac{\sum \tilde{x}_i^2}{N}$, $V_{yy} = \frac{\sum \tilde{y}_i^2}{N}$, $V_{xy} = \frac{\sum \tilde{x}_i \tilde{y}_i}{N}$ and $A_x = \bar{x}$ and $A_y = \bar{y}$ for further simplicity.

Supposing, we deal with degree-3 polynomial correction model which is represented as

$$\begin{cases} x_i = b_1 {x_{c,i}}^3 + b_2 {x_{c,i}}^2 y_{c,i} + \cdots + b_{10} + x_c \\ y_i = c_1 {x_{c,i}}^3 + c_2 {x_{c,i}}^2 y_{c,i} + \cdots + c_{10} + y_c \end{cases} \tag{A.22}$$

where $x_{c,i} = \frac{x_{d,i} - x_c}{S}$ and $y_{c,i} = \frac{y_{d,i} - y_c}{S}$ with (x_c, y_c) being a center of the image and S a normalization factor.

The sum of squared distances from the points $(x_{d,i}, y_{d,i})$ of a line to the corresponding regression line is presented as an equation (A.20).

Error Given a set of lines $k = 1, \cdots, K$, the error vector for LM algorithm is $E = (e_1 e_2 \cdots e_K)^T$. The root mean square distance e_k which represents an error for one line k as in (A.20), is given by

$$e_k = \sqrt{\frac{\sum_{i=1}^{N} (\alpha \tilde{x}_i + \beta \tilde{y}_i)^2}{N}} \tag{A.23}$$

In the above formula, we have to know what are the constants α and β for each line which are sine and cosine of the angle θ. It is possible to avoid calculation of the angle itself for simplicity and an advantage in running time by further expansion

$$e_k = \sqrt{\frac{\sum_N (\alpha^2 \tilde{x}_i^2 + \beta^2 \tilde{y}_i^2 + 2\alpha\beta \tilde{x}_i \tilde{y}_i)}{N}} \tag{A.24}$$

From the first solution in (A.15), we can deduce the expressions for cosine and sine of the angle 2θ:

$$\begin{aligned} \cos 2\theta &= \frac{1}{c_{norm}} (\sum \tilde{x}_i^2 - \sum \tilde{y}_i^2) = \frac{V_{xx} - Vyy}{c_{norm}} \\ \sin 2\theta &= -\frac{2}{c_{norm}} \sum \tilde{x}_i \tilde{y}_i = -\frac{2V_{xy}}{c_{norm}}, \end{aligned} \tag{A.25}$$

where $c_{norm} = \sqrt{(V_{xx} - V_{yy})^2 + 4V_{xy}^2}$ is a normalization coefficient.

Applying formulas for double angles $\cos 2\theta = 2\cos^2\theta - 1 = 1 - 2sin^2\theta$, we can get an expressions for each constant

$$\begin{aligned} \alpha^2 = \sin^2\theta &= \frac{1}{2}(1 - cos2\theta) = \frac{1}{2}(1 - \frac{V_{xx} - V_{yy}}{c_{norm}}) \\ \beta^2 = \cos^2\theta &= \frac{1}{2}(cos2\theta + 1) = \frac{1}{2}(1 + \frac{V_{xx} - V_{yy}}{c_{norm}}) \\ \alpha\beta = \sin\theta\cos\theta &= \frac{1}{2}\sin 2\theta = -\frac{V_{xy}}{c_{norm}}. \end{aligned} \tag{A.26}$$

Therefore, using the above expressions, we can expand (A.24) into

$$e_k = \sqrt{\frac{\sum(0.5(1-\cos 2\theta)\tilde{x}_i^2 + 0.5(\cos 2\theta + 1)\tilde{y}_i^2 + \sin 2\theta \tilde{x}_i\tilde{y}_i)}{N}}$$
$$= \sqrt{\frac{1}{2}((1-\frac{V_{xx}-V_{yy}}{c_{norm}})V_{xx} + (1+\frac{V_{xx}-V_{yy}}{c_{norm}})V_{yy} - \frac{4V_{xy}^2}{c_{norm}})} = \sqrt{\frac{1}{2}\langle V_{xx}+V_{yy}-c_{norm}\rangle} \tag{A.27}$$

As a result

$$e_k = \sqrt{\frac{1}{2}\left(V_{xx}+V_{yy}-\sqrt{(V_{xx}-V_{yy})^2+4V_{xy}^2}\right)} \tag{A.28}$$

Jacobian The Jacobian matrix is determined as a matrix of all first-order partial derivatives of the vector function which is unknown parameters in our case $(b_1, b_2, \cdots, b_{10}, c_1, c_2, \cdots, c_{10})$. For a given line k the general formulation is expressed as

$$\frac{\partial e_k}{\partial \bullet} = \frac{1}{4e_k}\left[\frac{\partial V_{xx}}{\partial \bullet} + \frac{\partial V_{yy}}{\partial \bullet} - \frac{4V_{xy}\frac{\partial V_{xy}}{\partial \bullet} + (V_{xx}-V_{yy})(\frac{\partial V_{xx}}{\partial \bullet} - \frac{\partial V_{yy}}{\partial \bullet})}{\sqrt{(V_{xx}-V_{yy})^2+4V_{xy}^2}}\right] \tag{A.29}$$

The partial derivatives $\frac{\partial V_{xx}}{\partial \bullet}$, $\frac{\partial V_{yy}}{\partial \bullet}$ and $\frac{\partial V_{xy}}{\partial \bullet}$ are calculated as follows:

$$\begin{aligned}
&\frac{\partial V_{xx}}{\partial b_1} = \frac{1}{N}\frac{\partial \sum(x_i - A_x)^2}{\partial b_1} = \frac{1}{N}\sum 2(x_i - A_x)(\frac{\partial x_i}{\partial b_1} - \frac{\partial A_x}{\partial b_1}), \cdots, \\
&\frac{\partial V_{yy}}{\partial b_1} = \frac{1}{N}\frac{\partial \sum(y_i - A_y)^2}{\partial b_1} = \frac{1}{N}\sum 2(y_i - A_y)(\frac{\partial y_i}{\partial b_1} - \frac{\partial A_x}{\partial b_1}), \cdots, \\
&\frac{\partial V_{xy}}{\partial b_1} = \frac{\partial A_{xy}}{\partial b_1} - A_y\frac{\partial A_x}{\partial b_1}, \cdots, \frac{\partial V_{xy}}{\partial c_1} = \frac{\partial A_{xy}}{\partial c_1} - A_x\frac{\partial A_y}{\partial c_1}, \cdots.
\end{aligned} \tag{A.30}$$

The partial derivatives $\frac{\partial x_i}{\partial \bullet}$ and $\frac{\partial y_i}{\partial \bullet}$ are:

$$\begin{aligned}
&\frac{\partial x_i}{\partial b_1} = x_{c,i}^3, \frac{\partial x_i}{\partial b_2} = x_{c,i}^2 y_{c,i}, \cdots, \frac{\partial x_i}{\partial c_1} = 0, \frac{\partial x_i}{\partial c_2} = 0, \cdots, \\
&\frac{\partial y_i}{\partial b_1} = 0, \frac{\partial y_i}{\partial b_2} = 0, \cdots, \frac{\partial y_i}{\partial c_1} = x_{c,i}^3, \frac{\partial y_i}{\partial c_2} = x_{c,i}^2 y_{c,i}, \cdots.
\end{aligned} \tag{A.31}$$

So, we just need the partial derivatives of A_{xy}, A_x and A_y with respect to the unknown parameters $b_1, b_2, \cdots, b_{10}, c_1, c_2, \cdots, c_{10}$:

$$\begin{aligned}
&\frac{\partial A_{xy}}{\partial b_1} = \frac{1}{N}\frac{\partial \sum x_i y_i}{\partial b_1} = \frac{1}{N}\sum(\frac{\partial x_i}{\partial b_1}y_i + x_i\frac{\partial y_i}{\partial b_1}), \cdots, \\
&\frac{\partial A_x}{\partial b_1} = \frac{1}{N}\frac{\partial \sum x_i}{\partial b_1} = \frac{1}{N}\sum\frac{\partial x_i}{\partial b_1}, \cdots, \frac{\partial A_y}{\partial b_1} = \frac{1}{N}\frac{\partial \sum y_i}{\partial b_1} = \frac{1}{N}\sum\frac{\partial y_i}{\partial b_1}, \cdots.
\end{aligned} \tag{A.32}$$

The result Jacobian matrix has a form of

$$J = \begin{pmatrix} \frac{\partial e_1}{\partial b_1} & \frac{\partial e_1}{\partial b_2} & \cdots & \frac{\partial e_1}{\partial c_1} & \frac{\partial e_1}{\partial c_2} & \cdots \\ \frac{\partial e_2}{\partial b_1} & \frac{\partial e_2}{\partial b_2} & \cdots & \frac{\partial e_2}{\partial c_1} & \frac{\partial e_2}{\partial c_2} & \cdots \\ \vdots & & & \ddots & & \end{pmatrix} \tag{A.33}$$

In practice, $x_c = \frac{w}{2}$ and $y_c = \frac{h}{2}$ are chosen to be the center of the images,

$$\begin{aligned} b_1 = 0, b_2 = 0, \cdots, b_8 = 1, b_9 = 0, b_{10} = 0. \\ c_1 = 0, c_2 = 0, \cdots, c_8 = 0, c_9 = 1, c_{10} = 0. \end{aligned} \tag{A.34}$$

$b_8, b_9, b_{10}, c_8, c_9, c_{10}$ are fixed as above and not estimated.

Denormalization As a final step, a normalization procedure must be performed:

$$\begin{array}{ll} \tilde{b}_1 \leftarrow \frac{b_1}{S^3} & \tilde{c}_1 \leftarrow \frac{c_1}{S^3} \\ \tilde{b}_2 \leftarrow \frac{b_2}{S^3} & \tilde{c}_2 \leftarrow \frac{c_2}{S^3} \\ \tilde{b}_3 \leftarrow \frac{b_3}{S^3} & \tilde{c}_3 \leftarrow \frac{c_3}{S^3} \\ \tilde{b}_4 \leftarrow \frac{b_4}{S^3} & \tilde{c}_4 \leftarrow \frac{c_4}{S^3} \\ \tilde{b}_5 \leftarrow \frac{b_5}{S^2} & \tilde{c}_5 \leftarrow \frac{c_5}{S^2} \\ \vdots & \vdots \end{array} \tag{A.35}$$

So, for the new distorted point the correction model is expressed as:

$$\begin{cases} x_i = \tilde{b}_1(x_{c,i} - x_c)^3 + \tilde{b}_2 x_{c,i} - {x_c}^2(y_{c,i} - y_c) + \cdots + x_c \\ y_i = \tilde{c}_1(x_{c,i} - x_c)^3 + \tilde{c}_2 x_{c,i} - {x_c}^2(y_{c,i} - y_c) + \cdots + y_c \end{cases} \tag{A.36}$$

A.3 Experiments

A.3.1 Choice of polynomial degree

A synthetic test was performed to see how the degree of the polynomial influences the obtained precision. We used 8 groups of lines with orientations θ as 10°, 20°, 30°, 40°, 50°, 60°, 70° and 80° to estimate the correction parameters by minimizing S (see (A.20)). We also chose another independent group of lines with orientation 55° for verification. The ideal lines are plotted in an image with resolution 1761×1174 pixels. The sampling step for each line is set to 30 pixels, and the number of samples on each line is always more than 15. The step between each consecutive lines is 30 pixels as well. The ideal lines are distorted by some radial distortion with tangential component, equation of which is given by

$$\begin{aligned} x_d &= x(k_0 + k_1 r + k_2 r^2) + \cdots + [p_1(r^2 + 2x^2) + 2p_2 xy](1 + p_3 r^2) + s_1 r^2 \\ y_d &= y(k_0 + k_1 r + k_2 r^2) + \cdots + [p_2(r^2 + 2y^2) + 2p_1 xy](1 + p_3 r^2) + s_2 r^2, \end{aligned} \tag{A.37}$$

where $k_0, k_1, \cdots$ are the radial distortion coefficients ($k_0 = 1.0$, $k_1 = 1.0e-4$, $k_2 = -2.0e-7$, $k_3 = -6.0e-14$); p_1, p_2, p_3 - decentering distortion coefficients ($p_1 = 4.0e{-}6$, $p_2 = -2.0e{-}6$), s_1, s_2 thin prism distortion coefficients ($s_1 = 3.0e{-}6$, $s_2 = 1.02{-}6$). The correction results are shown in Table A.1. The residual stabilizes and does not change much after degree 11 is achieved. It was also confirmed by our real tests. And it is the reason why we always chose the degree 11 for calibration of distortion, i.e., in Chapter 4.

order $p = q$	$RMSE$, in pixels, (A.5)	
	Calibration	Test
3	0.6935	0.6239
4	0.6096	0.5312
5	0.2439	0.2093
6	0.2419	0.2064
7	0.1050	0.0879
8	0.1031	0.0870
9	0.0521	0.0512
10	0.0515	0.0509
11	0.0477	0.0449
12	0.0474	0.0444

Table A.1: Synthetic test of lines correction: the residual errors for polynomial degrees of 3 to 12 for different lines data - the ones used for minimization ('Calibration' column) and independent measure ('Test' column).

focal length	distorted $RMSE$	undistorted $RMSE$	# images
$f = 18mm$	6.34 (23.20)	0.086 (0.57)	24
$f = 27mm$	1.39 (14.63)	0.044 (0.39)	29
$f = 55mm$	0.65 (4.26)	0.023 (0.11)	23

Table A.2: Straightness measure $RMSE$, (A.5) (in pixels) before and after the optical distortion correction for different focal lengths ($18mm$, $27mm$, $55mm$) of the camera Canon EOS5D; given in the format *mean (max)*. The last column indicates total number of images used for calibration.

A.3.2 Real data experiments

The real experiments were done with a reflex camera Canon EOS5D and EFS 18-55mm lens. Three focal lengths were chosen $f = 18mm$, $f = 27mm$ and $f = 55mm$ so that to have 3 different datasets. The RAW images were demosaicked same way as it is described in Section 4.3 by leaving same resolution as the raw image (3888×2592 pixels). In order to take different line orientations, the camera was set on a tripod and then rotated from $-90°$ to $90°$ in z plane with about $7°$ step, thus, we obtained 20-30 images for each focal length.

We performed the calibration with participation of different number of images, as well as with a verification images that were not used for calibration. Table A.2 provides correction results for each of three datasets given initial distortion measure; we used all the available images for calibration for each dataset (more than 20 for each). As it can be seen, the mean of corrected residual stays within 0.02-0.08 pixels, depending on the camera setting.

Table A.3 provides only results for the dataset taken with $f = 18mm$ and exposes correction results for test images against different number of images used for minimization. It can be seen from the results that we do not really need the

Image(-s)	# calibration images		
	22	15	8
Calibration	0.086 (0.57)	0.075 (0.48)	0.064 (0.31)
Test 1	0.067 (0.28)	0.064 (0.32)	0.058 (0.27)
Test 2	0.082 (0.43)	0.084 (0.46)	0.087 (0.39)
Test 3	0.066 (0.20)	0.083 (0.26)	0.039 (0.12)

Table A.3: Straightness measure $RMSE$, (A.5) (in pixels) for the corrected test images which did not take participation in minimization process for a dataset $f = 18mm$ of the camera Canon EOS5D; given in the format *mean (max)*. The 'Calibration' row shows mean and maximum result of the corrected calibration images.

number of images for calibration to be large. From our experiments and also with other datasets it was stated that most optimal number in sense of number-quality is 8-10 images with the condition that they cover well enough all different line orientations. The number of images 5-8 still provided meaningful results but were less precise, while the number less than 5 was not able to ensure stable calibration for the test images whose orientation did not participate in calibration process.

A.3.3 Measuring distortion correction of global calibration method

The calibration harp is also a tool to *measure* the optical distortion, since it holds the properties that distortion-free camera projected straight lines always remain straight.

After the calibration is performed for [BOUGUET 2000] software, we took the obtained radial distortion model and applied it to the images of harp taken with the same camera settings. However, the result corrected image did not seem to be corrected, the lens distortion was not compensated successfully; it is possible to visually observe the lines are not straight - see Figure A.2, therefore, we did not even proceed to actually measuring the straightness $RMSE$.

Our conclusion is that the obtained radial model by [BOUGUET 2000] is not able to capture the physical properties of the distortion field, and the high precision calibration results are rather explained by an error compensations of the internal and external parameters.

A.4 Conclusion

This section briefly describes the work done by [TANG 2011] which is a preliminary step to a calibration process given in Chapter 5. The minor contributions introduced: simplification of minimization function for the gain of computational cost. The method is implemented in C++ language. Results for different focal lengths and different number of calibration images are demonstrated. Depending on the focal length, the corrected residual stays within 0.02-0.08 pixels, and the

Figure A.2: A crop of undistorted image of lines calibration pattern by using global calibration radial model of [BOUGUET 2000] software: note, the lines are not even visually straight.

most optimal number of images for calibration is 8-10 with the condition it covers all different line orientations. The calibration harp was also used as a measurement tool for a global calibration method like [ZHANG 2000] and showed that correction by an obtained radial polynomial is not efficient; and this supports the fact of error compensation between internal and external parameters when calculating all the parameters simultaneously.

List of Figures

List of Tables

Bibliography

[ÁLVAREZ LEÓN 2007] Luis ÁLVAREZ LEÓN, Agustín J. SALGADO DE LA NUEZ and Javier Sánchez PÉREZ. *Robust detection and ordering of ellipses on a calibration pattern.* Pattern Recognition and Image Analysis, vol. 17(4), pages 508–522, 2007. (Cited on page 15.)

[ASARI 1999] Vijayan K. ASARI, Sanjiv KUMAR and D. RADHAKRISHNAN. *A new approach for nonlinear distortion correction in endoscopic images based on least squares estimation.* IEEE Trans. Med. Imaging, vol. 18, no. 4, pages 345–354, 1999. (Cited on page 14.)

[BOUGUET 1998a] Jean-Yves BOUGUET. *Camera calibration from points and lines in dual-space geometry.* Rapport technique, California Institute of Technology, 1998. (Cited on pages 52 and 74.)

[BOUGUET 1998b] Jean-Yves BOUGUET. *Closed-form camera calibration in dual-space geometry.* Rapport technique, California Institute of Technology, 1998. (Cited on pages 52 and 74.)

[BOUGUET 2000] Jean-Yves BOUGUET. *Matlab camera calibration toolbox*, 2000. (Cited on pages 22, 64, 84, 85 and 89.)

[BOULT 1992] Terrance E. BOULT and George WOLBERG. *Correcting chromatic aberrations using image warping.* In Computer Vision and Pattern Recognition, 1992. Proceedings CVPR'92., 1992 IEEE Computer Society Conference on, pages 684–687. IEEE, 1992. (Cited on page 31.)

[BROWN 1971] Duane C. BROWN. *Close-range camera calibration.* Photogrammetric Engineering, vol. 37, no. 8, pages 855–866, 1971. (Cited on pages 52 and 75.)

[CABRERA 1996] Javier CABRERA and Peter MEER. *Unbiased Estimation of Ellipses by Bootstrapping.* IEEE Trans. Pattern Anal. Mach. Intell., vol. 18, no. 7, pages 752–756, 1996. (Cited on page 15.)

[CANON] CANON. *Digital photo professional.* http://www.canon.fr/. (Cited on page 39.)

[CHUNG 2010] Soon-Wook CHUNG, BYOUNG-KWANG Kim and WOO-JIN Song. *Removing chromatic aberration by digital image processing.* Optical Engineering, vol. 49(6), page 067002, 2010. (Cited on page 31.)

[CLAUS 2005] David CLAUS and Andrew W. FITZGIBBON. *A Rational Function Lens Distortion Model for General Cameras.* In CVPR (1), pages 213–219. IEEE Computer Society, 2005. (Cited on page 75.)

[DEVERNAY 2001] Frédéric DEVERNAY and Olivier FAUGERAS. *Straight lines have to be straight.* In In SPIE, volume 2567, 2001. (Cited on pages 52, 74 and 75.)

[DUNNE 2010] Aubrey DUNNE and John MALLON. *Jaliko software enhanced optics*, 2010. http://www.jaliko.com/index.php. (Cited on page 39.)

[DxO] DxO. *DxO Optics Pro.* http://www.dxo.com/. (Cited on page 39.)

[FAIG 1975] W. FAIG. *Calibration of close-range photogrammetry systems: mathematical formulation.* Photogrammetric Engineering and Remote Sensing, vol. 41(12), pages 127–140, 1975. (Cited on page 52.)

[FANGA 2006] Yi-Chin FANGA, Tung-Kuan LIUA, John MACDONALD, Jyh-Horng CHOUA, Bo-Wen WUA, Hsien-Lin TSAIA and En-Hao CHANGA. *Optimizing chromatic aberration calibration using a novel genetic algorithm.* Journal of Modern Optics, vol. 53(10), pages 1411–1427, 2006. (Cited on page 30.)

[FARRAR 2000] Nigel R. FARRAR, Adlai H. SMITH, Daniel R. BUSATH and Dennis TAITANO. *In-situ measurement of lens aberrations.* In Optical Microlithography, volume XIII, 2000. (Cited on page 30.)

[FITZGIBBON 1995] Andrew FITZGIBBON and Robert B. FISHER. *A Buyer's Guide to Conic Fitting.* In In British Machine Vision Conference, pages 513–522, 1995. (Cited on page 15.)

[FITZGIBBON 1999] M. FITZGIBBON A. W.and Pilu and R. B. FISHER. *Direct least-squares fitting of ellipses.* vol. 21, no. 5, pages 476–480, May 1999. (Cited on page 15.)

[FITZGOBBON 2001] Andrew W. FITZGOBBON. *Simultaneous linear estimation of multiple view geometry and lens distortion.* In CVPR (1), pages 125–132. IEEE Computer Society, 2001. (Cited on page 75.)

[FREITAG 2005] B. FREITAG, S. KUJAWA, P.M. MUL, J. RINGNALDA and P.C. TIEMEIJER. *Breaking the spherical and chromatic aberration barrier in transmission electron microscopy.* Ultramicroscopy, vol. 102(3), pages 209–214, 2005. (Cited on page 31.)

[FURUKAWA 2010] Yasutaka FURUKAWA and Jean PONCE. *Accurate, Dense, and Robust Multiview Stereopsis.* IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 8, pages 1362–1376, August 2010. (Cited on page 14.)

[GANDER 1994] Walter GANDER, GENE H. Golub and Rolf STREBEL. *Least Squares Fitting of Circles and Ellipses*, 1994. (Cited on page 15.)

[GROMPONE VON GIOI 2010] R. GROMPONE VON GIOI, MONASSE P., J.-M. MOREL and Z. TANG. *Self-consistency and universality of camera lens distortion models*. CMLA Preprint, 2010. (Cited on page 32.)

[GURDJOS 2003] Pierre GURDJOS and Peter F. STURM. *Methods and Geometry for Plane-Based Self-Calibration.* In CVPR (1), pages 491–496. IEEE Computer Society, 2003. (Cited on page 52.)

[HARTLEY 1994] Richard I. HARTLEY. *Euclidean Reconstruction from Uncalibrated Views*. In Proceedings of the Second Joint European - US Workshop on Applications of Invariance in Computer Vision, pages 237–256, London, UK, UK, 1994. Springer-Verlag. (Cited on page 52.)

[HARTLEY 1997a] Richard I. HARTLEY. *In defense of the eight-point algorithm*. IEEE Trans. Pattern Anal. Mach. Intell., vol. 19, no. 6, pages 580–593, June 1997. (Cited on page 56.)

[HARTLEY 1997b] Richard I. HARTLEY and Tushar SAXENA. *The cubic rational polynomial camera model*. In In Image Understanding Workshop, pages 649–653, 1997. (Cited on page 75.)

[HARTLEY 2004] Richard HARTLEY and Andrew ZISSERMAN. Multiple view geometry in computer vision. Cambridge University Press, ISBN 0521540518, New York, NY, USA, second édition, 2004. (Cited on pages 55 and 56.)

[HARTLEY 2005] Richard HARTLEY and Sing Bing KANG. *Parameter-free radial distortion correction with centre of distortion estimation*. In In ICCV, pages 1834–1841, 2005. (Cited on page 74.)

[HARTLEY 2007] Richard HARTLEY and Sing Bing KANG. *Parameter-Free Radial Distortion Correction with Center of Distortion Estimation*. IEEE Trans. Pattern Anal. Mach. Intell., vol. 29, no. 8, pages 1309–1321, August 2007. (Cited on page 52.)

[HEIKKILÄ 2000] Janne HEIKKILÄ. *Geometric camera calibration using circular control points*. IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 10, pages 1066–1077, October 2000. (Cited on pages 14 and 53.)

[HIGUCHI 2012] Mirai HIGUCHI, Ankur DATTA and Takeo KANADE. *Software package for precise camera calibration*, 2012. (Cited on pages 64, 67 and 88.)

[JAIN 1995] Ramesh C. JAIN, Rangachar KASTURI and Brian G. SCHUNCK. Machine vision. McGraw-Hill, 1995. (Cited on page 15.)

[JIANG 2005] Guang JIANG and Long QUAN. *Detection of Concentric Circles for Camera Calibration*. In Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1 - Volume 01, ICCV

'05, pages 333–340, Washington, DC, USA, 2005. IEEE Computer Society. (Cited on page 52.)

[JOHNSON 2006] M.K. JOHNSON and H. FARID. *Exposing digital forgeries through chromatic aberration.* In ACM Multimedia and Security Workshop, Geneva, Switzerland, 2006. (Cited on page 31.)

[JÜRGEN 1995] R. Meyer-Arendt JÜRGEN. Introduction to classical and modern optics. Prentice-Hall International, 1995. (Cited on page 29.)

[KANATANI 1994] Kenichi KANATANI. *Statistical Bias of Conic Fitting and Renormalization.* IEEE Trans. Pattern Anal. Mach. Intell., vol. 16, no. 3, pages 320–326, 1994. (Cited on page 15.)

[KANATANI 2006] Kenichi KANATANI. *Ellipse Fitting with Hyperaccuracy.* IEICE Transactions, vol. 89-D, no. 10, pages 2653–2660, 2006. (Cited on page 15.)

[KANATANI 2011] Kenichi KANATANI and Prasanna RANGARAJAN. *Hyper least squares fitting of circles and ellipses.* Computational Statistics & Data Analysis, vol. 55, no. 6, pages 2197–2208, 2011. (Cited on page 15.)

[KANG 2007] Sing Bing KANG. *Automatic removal of chromatic aberration from a single image.* In IEEE Conference on Computer Vision and Pattern Recognition, pages 1–8, 2007. (Cited on page 31.)

[KANNALA 2006] Juho KANNALA and Sami S. BrANDT. *A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses.* IEEE TRANS. PATTERN ANALYSIS AND MACHINE INTELLIGENCE, vol. 28, pages 1335–1340, 2006. (Cited on pages 14 and 53.)

[KAUFMANN 2005] V. KAUFMANN and R. LADSTÄDTER. *Elimination of color fringes in digital photographs caused by lateral chromatic aberration.* In CIPA XX International Symposium, 2005. (Cited on page 31.)

[KILPELÄ 1980] E. KILPELÄ. *Compensation of systematic errors of image and model coordinates.* Photogrammetria, vol. 37(1), pages 15–44, 1980. (Cited on page 75.)

[KINGSLAKE 1978] Rudolf KINGSLAKE. Lens design fundamentals. Academic Press, New York, 1978. (Cited on page 30.)

[KOZUBEK 2000] M. KOZUBEK and P. MATULA. *An efficient algorithm for measurement and correction of chromatic aberrations in fluorescence microscopy.* Journal of Microscopy, vol. 200, no. 3, pages 206–217, 2000. (Cited on page 31.)

[LAVEST 1998] Jean-Marc LAVEST, Marc VIALA and Michel DHOME. *Do we really need an accurate calibration pattern to achieve a reliable camera calibration?* In Hans Burkhardt and Bernd Neumann, editeurs, ECCV (1), volume 1406 of *Lecture Notes in Computer Science*, pages 158–174. Springer, 1998. (Cited on page 74.)

[LOWE 1999] David G. LOWE. *Object Recognition from Local Scale-Invariant Features.* In Proceedings of the International Conference on Computer Vision-Volume 2 - Volume 2, ICCV '99, pages 1150–, Washington, DC, USA, 1999. IEEE Computer Society. (Cited on page 14.)

[LOWE 2004] David G. LOWE. *Distinctive image features from scale-invariant keypoints.* Int. J. Comput. Vision, vol. 60, no. 2, pages 91–110, November 2004. (Cited on pages 14 and 38.)

[LUCCHESE 2002] Luca LUCCHESE and Sanjit K. MITRA. *Using saddle points for subpixel feature detection in camera calibration targets.* In Asia-Pacific Conference on Circuits and Systems, volume 2, pages 191–195, 2002. (Cited on pages 14 and 15.)

[LUHMANN 2006] Thomas LUHMANN, Heidi HASTEDT and Werner TECKLENBURG. *Modelling of chromatic aberration for high precision photogrammetry.* In Commission V Symp. on Image Engineering and Vision Metrology, Proc. ISPRS, volume 36, pages 173–178, 2006. (Cited on page 31.)

[MA 2004] Lili MA, Yangquan Chen and Kevin L. Moore. *Rational Radial Distortion Models of Camera Lenses with Analytical Solution for Distortion Correction.* I. J. Information Acquisition, vol. 1, no. 2, pages 135–147, 2004. (Cited on page 75.)

[MALLON 2007a] John MALLON and Paul F. WHELAN. *Calibration and removal of lateral chromatic aberration in images.* Pattern Recognition Letters, vol. 28, no. 1, pages 125–135, January 2007. (Cited on pages 15 and 31.)

[MALLON 2007b] John MALLON and Paul F. WHELAN. *Which pattern? biasing aspects of planar calibration patterns and detection methods.* Pattern Recognition Letters, vol. 28(8), pages 921–930, 2007. (Cited on pages 14, 52 and 53.)

[MANDERS 1997] E. M. M. MANDERS. *Chromatic shift in multicolour confocal microscopy.* Journal of Microscopy, vol. 185(3), pages 312–328, 1997. (Cited on page 31.)

[MARQUARDT 1963] Donald W. MARQUARDT. *An algorithm for least-squares estimation of nonlinear parameters.* SIAM Journal on Applied Mathematics, vol. 11, no. 2, pages 431–441, 1963. (Cited on page 57.)

[MATSUOKA 2012] Ryuji MATSUOKA, Kazuyoshi ASONUMA, Genki TAKAHASHI, Takuya DANJO and Kayoko HIRANA. *Evaluation of correction methods of chromatic aberration in digital camera images.* In ISPRS Photogrammetric image analysis, volume I-3, 2012. (Cited on page 31.)

[MIELENZ 1999] K. MIELENZ. *On the diffraction limit for lensless imaging.* Journal of Research of the National Institute of Standards and Technology, vol. 104(5), pages 479–485, 1999. (Cited on page 29.)

[MILLÁN 2006a] Maria S. MILLÁN, Joaquin OTÓN and Elisabet PÉREZ-CABRÉ. *Dynamic compensation of chromatic aberration in a programmable diffractive lens.* Optics Express, vol. 14(20), pages 9103–9112, 2006. (Cited on page 30.)

[MILLÁN 2006b] Maria S. MILLÁN, Joaquin OTÓN and Elisabet PÉREZ-CABRÉ. *Dynamic compensation of chromatic aberration in a programmable fresnel lenses.* Optics Express, vol. 14(13), pages 6226–6242, 2006. (Cited on page 30.)

[NEWTON 1704] Isaak NEWTON. Opticks. William Innys, 1704. (Cited on page 29.)

[NISTER 2006] David NISTER and Henrik STEWENIUS. *Scalable Recognition with a Vocabulary Tree.* In Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 2, CVPR '06, pages 2161–2168, Washington, DC, USA, 2006. IEEE Computer Society. (Cited on page 14.)

[OUELLET 2008] Jean-Nicolas OUELLET, Felix ROCHETTE and Patrick HEBERT. *Geometric calibration of a structured light system using circular control points.* In 3D Data Processing, Visualization and Transmission, pages 183–190, 2008. (Cited on page 53.)

[PIERROT DESEILLIGNY 2011] M. PIERROT DESEILLIGNY and I. CLERY. *APERO, an open source bundle adjusment software for automatic calibration and orientation of set of images.* ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXVIII-5/W16, pages 269–276, 2011. (Cited on pages 52 and 74.)

[POWELL 1981] Ian POWELL. *Lenses for correcting chromatic aberration of the eye.* Applied Optics, vol. 20(24), pages 4152–4155, 1981. (Cited on page 30.)

[PRINCEN 1992] John PRINCEN, John ILLINGWORTH and Josef KILTTLER. *A formal definition of the Hough transform: Properties and relationships.* Journal of Mathematical Imaging and Vision, vol. 1, no. 2, pages 153–168, 1992. (Cited on page 15.)

[REBIAI 1992] M. REBIAI, S. MANSOURI, F. PINSON and B. TICHIT. *Image distortion from zoom lenses: modeling and digital correction.* In Broadcasting Convention, pages 438–441, 1992. (Cited on page 31.)

[SEIDEL 1856] L. SEIDEL. *Astronomische nachrichten*, 1856. (Cited on page 29.)

[SHI 2004] Kebin SHI, Peng LI, Shizhuo YIN and Zhiwen LIU. *Chromatic confocal microscopy using supercontinuum light.* Optics Express, vol. 12(10), pages 2096–2101, 2004. (Cited on page 31.)

[SIVIC 2005] J. SIVIC, B. C. RUSSELL, A. A. EFROS, A. ZISSERMAN and W. T. FREEMAN. *Discovering objects and their location in images.* volume 1, pages 370–377, 2005. (Cited on page 14.)

[SIVIC 2006] Josef SIVIC and Andrew ZISSERMAN. *Video Google: Efficient Visual Search of Videos.* In Toward Category-Level Object Recognition, pages 127–144, 2006. (Cited on page 14.)

[SLAMA 1980] Chester C. SLAMA. Manual of photogrammetry. American Society of Photogrammetry, Falls Church, VA, 4th ed. édition, 1980. (Cited on pages 29 and 74.)

[SNAVELY 2008] Noah SNAVELY, Steven M. SEITZ and Richard SZELISKI. *Modeling the World from Internet Photo Collections.* Int. J. Comput. Vision, vol. 80, no. 2, pages 189–210, November 2008. (Cited on page 14.)

[STURM 1999] Peter F. STURM and Stephen J. MAYBANK. *On Plane-Based Camera Calibration: A General Algorithm, Singularities, Applications.* In CVPR, pages 1432–1437. IEEE Computer Society, 1999. (Cited on page 52.)

[TANG 2011] Z. TANG. *High accuracy measurement in 3d stereo reconstruction.* PhD thesis, ENS Cachan, France, 2011. (Cited on pages 3, 5, 9, 10, 11, 31, 32, 53, 73, 74, 75 and 84.)

[TORR 2000] P.H.S. TORR and A. ZISSERMAN. *Feature Based Methods for Structure and Motion Estimation.* In Vision Algorithms: Theory and Practice, number 1883 in LNCS, pages 278–295. Springer-Verlag, 2000. (Cited on page 14.)

[TSAI 1992] Roger Y. TSAI. *A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses.* IEEE Journal on Robotics and Automation, vol. RA-3, pages 323–44, 1992. (Cited on pages 52 and 74.)

[WENG 1992] Juyang WENG, Paul COHEN and Marc HERNIOU. *Camera calibration with distortion models and accuracy evaluation.* IEEE Trans. Pattern Anal. Mach. Intell., vol. 14, no. 10, pages 965–980, October 1992. (Cited on pages 14, 52 and 74.)

[WILLSON 1991] Reg WILLSON and Steven SHAFER. *Active lens control for high precision computer imaging.* In Proc. IEEE ICRA, volume 3, pages 2063 – 2070, April 1991. (Cited on pages 30 and 31.)

[WILLSON 1994a] Reg WILLSON. *Modelling and calibration of automated zoom lenses.* PhD thesis, Carnegie Mellon University, 1994. (Cited on page 30.)

[WILLSON 1994b] Reg WILLSON and Steven SHAFER. *What is the center of the image?* JOSA A, vol. 11(11), pages 2946–2955, 1994. (Cited on page 31.)

[YANG 2000] Changjiang YANG, Fengmei SUN and Zhanyi HU. *Planar conic based camera calibration.* In ICPR, pages 1555–1558, 2000. (Cited on page 53.)

[ZHANG 2000] Zhengyou ZHANG. *A flexible new technique for camera calibration.* IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 11, pages 1330–1334, November 2000. (Cited on pages 3, 9, 14, 52, 53, 74 and 85.)