UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE MSII

ICube UMR 7357

THÈSE présentée par :

Devesh ADLAKHA

soutenue le : 12 décembre 2022

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Image et Vision

Exploitation de connaissances partielles sur le mouvement et la géométrie des caméras en vision 3D non calibrée

THÈSE dirigée par :

M. DE MATHELIN Michel	Professeur, Université de Strasbourg

CO-ENCADRANTS :

M. HABED Adlane	Maître de Conférences, Université de Strasbourg
M. MORBIDI Fabio	Maître de Conférences, Université de Picardie Jules Verne
M. DEMONCEAUX Cédric	Professeur, Université de Bourgogne Franche-Comté

RAPPORTEURS :

M. MONASSE Pascal	Professeur, Ecole des Ponts Paris Tech
M. FUSIELLO Andrea	Professeur, Université d'Udine

EXAMINATEURS :

M. STURM Peter	Directeur de Recherche, Inria Grenoble Rhône-Alpes
M. VASSEUR Pascal	Professeur, Université de Picardie Jules Verne

Acknowledgements

Several people contributed to making this journey a fun and enriching experience. First and foremost, I was fortunate to have a great group of advisors. Thanks to Michel and Cédric for creating this opportunity for me. I especially appreciate the freedom I had in my research and the place of work. I have also received a lot of advice and support over the years from Cédric, for which I am really grateful. My research has been in collaboration with Adlane, and his passion and dedication have been very inspiring. Our discussions on research, writing, presentations, and everything else, are what I'll cherish the most from this experience. Finally, thanks to Fabio for all the time he spent with me discussing my research, among other things. His feedback and encouragement have been a constant source of motivation.

I am grateful to my thesis committee members: Andrea Fusiello, Pascal Monasse, Peter Sturm, and Pascal Vasseur, for the discussions and their insightful comments. Thanks to Andrea and Pascal Monasse for carefully reviewing my manuscript. Thanks also to Pascal Vasseur and Pascal Monasse for their feedback during my mid-thesis presentation.

During my stay in Strasbourg and Le Creusot, I received valuable administrative support from the labs and universities in both places. I especially thank the Euraxess office for their help. I also gratefully acknowledge the funding support from the ICube lab, the ANR SUMUM project (grant ANR-17-CE38-0004), and the Interreg VA France (Channel) England ADAPT project.

Finally, thanks to my friends and family for their positivity and kindness. I have been fortunate to have some incredible friendships during these years and the constant love and support from my parents and brother.

Abstract

This thesis concerns 3D computer vision, where the fundamental problem is reconstructing a scene in 3D from multiple images captured from different viewpoints. Known as Structure-from-Motion (SfM), this problem has various applications, such as in cultural heritage and augmented reality. We investigate uncalibrated SfM, where a reconstruction only up to a projective transformation can be obtained from feature correspondences across images. The goal is to recover a metric representation of the scene from the projective one. This involves locating the so-called Absolute Conic on the plane at infinity that acts as a virtual calibration object analogous to a physical calibration pattern used in calibrated SfM.

The main contributions of this thesis are two-fold. The first contribution exploits partial knowledge of the camera geometry, specifically that the camera has square pixels. This assumption is satisfied by most modern cameras. We formulate a new polynomial constraint on the plane at infinity under this assumption and propose a method for an affine upgrade that relies on polynomial optimization using the so-called Lasserre's hierarchy of convex relaxations. The metric reconstruction is then recovered by solving linear equations. The second contribution exploits a vague knowledge of the camera motion that the viewpoint is typically changed mildly between images to ensure sufficient overlap to match features. We show that bounds on the relative rotation angle between camera pairs can be used to constrain the plane at infinity to a convex set. Based on this constraint, we show the existence of a new quasi-affine reconstruction of a scene with respect to the *Hodographs of the Horopter*, new geometric objects that we introduce in this thesis. We propose a semidefinite programming problem to recover such reconstruction from a projective one and present a constrained Levenberg-Marquardt optimization method to upgrade it to affine. Experiments with synthetic data and real images validate our proposed methods.

Contents

List of Figures

List of Tables

Acronyms

AC	Absolute Conic
ALQ	Absolute Line Quadric
DAQ	Dual Absolute Quadric
DIAC	Dual Image of the Absolute Conic
IAC	Image of the Absolute Conic
ICT	Infinite Cayley Transform
IMU	Inertial Measurement Unit
LM	Levenberg-Marquardt
LMI	Linear Matrix Inequality
LP	Linear Programming
RMS	Root-Mean-Square
SDP	Semidefinite Programming
SfM	Structure-from-Motion
SLAM	Simultaneous Localization and Mapping

Symbols

$\mathrm{sgn}(\mathsf{a})$	the sign of a real number a: $\mathrm{sgn}(\mathsf{a}) = -1$ if $\mathsf{a} < 0$, $\mathrm{sgn}(\mathsf{a}) = 0$ if $\mathsf{a} = 0$, $\mathrm{sgn}(\mathsf{a}) = 1$ if $\mathsf{a} > 0$
$(\mathsf{A})_{hk}$	the element in the hth row and kth column of any matrix A
A^*	the adjugate of a square matrix A
I_n	the $n \times n$ identity matrix
$\mathbf{0_n}$	an n-vector of zeros
$\simeq$	the equality up to scale
$\mathbb{P}^n$	the projective n-space
$\mathbb{P}^{*n}$	the dual projective n-space
${}^h p$	a homogeneous polynomial

1 Introduction

While we effortlessly perceive the three-dimensional (3D) world around us through visual stimuli, this task is not trivial for a computer. By relying on a camera for a visual input, the 3D information is lost in the projection of a 3D scene to a 2D image. Recovering the 3D scene structure from multiple images is a central problem in computer vision. Research efforts to address this problem have resulted in the development of the theoretical foundations and computational algorithms to obtain 3D reconstructions from images. In this thesis, we are interested in the uncalibrated 3D reconstruction approach, where the camera calibration is not known. In particular, we present contributions in addressing the camera autocalibration problem that forms one component of the uncalibrated 3D reconstruction pipeline.

1.1 Context and motivation

3D scene modeling: 3D models of scenes and objects in the world are now an integral part of a variety of applications. In Augmented Reality (AR), one's view of the physical world is enhanced by digital models that are overlaid virtually in the scene. For instance, the mobile application Snapchat introduced an AR feature called Local Lenses[1]. With this feature, users can engage with virtual content, such as a shared AR world across Carnaby street in London. Local Lenses relies on a 3D reconstruction of the street and buildings that was computed using images, including public snaps, captured around the street. In robotics, building and maintaining a 3D map of an unknown environment is often crucial for a robot to interact with and perform tasks in its surroundings. These tasks may include inspection and monitoring [Maurer et al. 2017], search and rescue operations [Delmerico et al.

[1] `https://ar.snap.com/lens-studio`

Figure 1.1: An application of 3D scene modeling in digital heritage. Images of the interactive visit of the Villa of Diomedes in Pompeii: a view of the present state of the site (left) and a rendering of its possible historical state (right). Images reproduced with permission from the Villa Diomedes Project[3].

2019], and autonomous driving [Bresson et al. 2017]. For example, the autonomous car from Cruise[2] uses a 3D laser scan of its surroundings, among other sensor information, to navigate in its environment. In digital heritage, reconstructing 3D models of cultural sites and objects is a means to preserve them digitally. These digital models can further be used for analysis, such as change detection, and in interactive tours. Figure 1.1 shows screenshots of an interactive visit of the Villa of Diomedes [Dessales 2020] in Pompeii, where a rendering of the possible historical state of the site can be superimposed over a 3D model of its present state. This 3D model was obtained using images and laser scans of the site. There are numerous other applications of 3D scene modeling, for instance, in gaming [Statham 2020], forensics [Galanakis et al. 2021], and paleontology [Falkingham et al. 2020].

3D sensing: Non-contact methods to capture the 3D structure of a scene can be broadly categorized as active or passive [Liu et al. 2020]. Active methods interfere with the scene by introducing controlled electromagnetic radiation using, for instance, structured light projectors (*e.g.* Microsoft Kinect V1, Intel RealSense SR300) or Time-of-Flight scanners (*e.g.* Microsoft Kinect V2, Intel RealSense L515). In the structured light approach, a pattern is projected on the scene, and one or more cameras (CCD/CMOS sensors) are used to image the illuminated scene. As the pattern is known, the correspondences between it and the image projections can be easily determined, and the 3D structure is recovered through trian-

[2] `https://www.getcruise.com/`
[3] `http://villadiomede.huma-num.fr/3dproject/`

gulation. In the Time-of-Flight approach, light is emitted on the scene, and the 3D structure is determined more directly by measuring its reflection. Passive methods, on the other hand, use a camera to capture the scene under ambient conditions in 2D images. The 3D scene structure is retrieved by processing the images using methods of photogrammetry/computer vision. The choice of the acquisition method and sensors used for a particular application is influenced by several factors, including the scene (its size, shape, texture), the acquisition conditions (ambient lighting, accessibility and setup), the required accuracy, and the involved cost. Cameras generally offer more flexibility than an active system because of their compact size, power efficiency, and easy setup and acquisition. The ubiquity of digital cameras (*e.g.* in smartphones) and the constant improvement in their image quality has led to their widespread adoption for 3D acquisition.

Image-based 3D reconstruction: The problem of reconstructing a scene in 3D from multiple images has been studied extensively in computer vision, with its origins being in photogrammetry [Hartley and Mundy 1993; Sturm 2011]. Known as Structure-from-Motion (SfM) in computer vision, it involves recovering both the 3D scene structure and the camera motion from the images. The underlying theory of the geometric relations between multiple views, or multi-view geometry, is now well established [Faugeras et al. 2001; Hartley and Zisserman 2004]. A typical SfM approach involves solving several problems that include determining feature points and their correspondences across the images [Lowe 2004], computing the relative camera poses [Longuet-Higgins 1981; Nistér 2004a], and retrieving the 3D scene structure through triangulation [Hartley and Sturm 1997]. A robust estimation of the camera parameters using methods such as RANSAC [Fischler and Bolles 1981] and a joint refinement of the estimated parameters and the 3D structure through bundle adjustment [Triggs et al. 1999] form key components of this approach. Progress in solving these problems led to the development of automated SfM systems, such as Photo Tourism [Snavely et al. 2008], that take as input a set of images and compute the camera poses and a sparse 3D reconstruction (a point cloud) of the imaged scene. SfM has since evolved into a mature technology, and there are now several SfM packages, including commercial software, such as Agisoft MetaShape[4],

[4] `https://www.agisoft.com/`

3DF Zephyr[5], and RealityCapture[6], and non-commercial software, such as VisualSfM [C. Wu 2013], OpenMVG [Moulon et al. 2016], and COLMAP [Schönberger and Frahm 2016]. A sparse 3D reconstruction obtained from SfM can be processed to obtain a richer representation of the scene, for example, a dense point cloud or a textured mesh. Some SfM packages have such functionality built-in or they can be used with other dedicated software [Furukawa et al. 2010]. The availability and ease of use of SfM packages combined with the high fidelity of the resulting 3D reconstructions have spurred their use for 3D scene modeling.

Calibrated SfM: SfM implementations generally assume the camera intrinsic parameters to be known. These parameters correspond to the internal geometry of the camera, for example, its focal length. Together with the camera pose, *i.e.* the camera extrinsic parameters, they describe the perspective projection model used in computer vision to model the mapping of a 3D point in the world to its 2D pixel location in the image. The intrinsic parameters are obtained by calibrating the camera. The standard camera calibration method [Zhang 2000] involves taking multiple images of a known planar pattern, such as a checkerboard. This procedure is rather tedious and, despite the development of guidance systems [Peng and Sturm 2019], requires some technical expertise to obtain an accurate calibration. In SfM applications, typically the camera is calibrated prior to the image acquisition and its zoom and focus are kept fixed to ensure that the internal camera calibration remains unchanged. If the internal camera calibration does change, the intrinsic parameters must be reestimated, for instance, by recalibrating the camera. SfM pipelines can often benefit from Exif (Exchangeable image file format) metadata to approximate the intrinsic parameters and thereby circumvent the cumbersome offline calibration step. Exif tags are embedded in some image file formats, such as JPEG. They contain information about the camera, notably its focal length. The other intrinsic parameters are usually approximated based on some heuristics. The pipeline then relies on a nonlinear refinement of the camera parameters and scene structure in bundle adjustment to converge to an accurate calibration. However, the Exif metadata is not

[5] https://www.3dflow.net/3df-zephyr-photogrammetry-software/
[6] https://www.capturingreality.com/

always available. In this case, SfM packages resort to loose heuristics to initialize the focal length or require the intrinsic parameters as input.

Uncalibrated SfM: Unlike calibrated SfM, the camera intrinsic parameters are not known in an uncalibrated approach. With feature correspondences across uncalibrated images alone, only a reconstruction up to a projective ambiguity can be recovered. Such a reconstruction is related to a metric representation of the imaged scene by a projective transformation, and it is referred to as a projective reconstruction. A projective reconstruction does not preserve metric properties, such as angles and length ratios, but preserves only projective invariants that include collinearity and coplanarity. It thus appears distorted. For most applications, a more faithful representation of the observed scene is required. A projective reconstruction can be upgraded to its metric counterpart by determining the projective transformation that relates the two reconstructions. This transformation involves the camera intrinsic parameters, and recovering them from multiple uncalibrated images is the camera autocalibration problem in computer vision. Uncalibrated SfM with camera autocalibration is more flexible than the calibrated approach as it avoids the manual camera calibration step. Thus, this problem has received significant attention in the literature [Fusiello 2000].

Camera autocalibration: The process of retrieving the camera intrinsic parameters from multiple uncalibrated images of an unknown scene is known as camera autocalibration. Camera autocalibration does not require a physical calibration pattern to be observed in the images. Instead, it relies on the omnipresence of a virtual one: the so-called *Absolute Conic* (AC). The AC is a particular conic lying on the *plane at infinity* that projects on the image plane as an imaginary conic whose matrix representation embeds the camera intrinsic parameters. Thus, the intrinsic parameters can be retrieved by locating the AC or its image (IAC). The flexibility of working directly with uncalibrated images, however, comes at the cost of autocalibration being a challenging problem. The equations involved are nonlinear and difficult to solve reliably and efficiently. In particular, locating the plane at infinity is the main challenge in camera autocalibration. Once it is identified, the IAC and thus the intrinsic parameters can be determined by solving linear equations. For this reason,

some autocalibration methods first estimate the plane at infinity and then retrieve the intrinsic parameters in a second step. This two-step approach is referred to as *stratified* autocalibration. In *direct* autocalibration, on the other hand, the plane at infinity and the intrinsic parameters are estimated simultaneously. Direct methods generally rely on special virtual quadrics, the *Dual Absolute Quadric* (DAQ) or the *Absolute Line Quadric* (ALQ), that encode both the plane at infinity and the IAC or its dual (DIAC). In either case, all autocalibration methods rely on some prior knowledge of the camera, its internal geometry or motion, to retrieve the intrinsic parameters. For instance, several methods consider the case of a moving camera with constant but unknown intrinsic parameters. Other methods exploit knowledge of the camera motion, such as planar motion in the case of a ground vehicle. A general guideline is to use as much information as possible about the camera [Hartley and Zisserman 2004, Sec. 19.11] to deal with the inherent challenges in camera autocalibration.

1.2 Scope

In this thesis, we consider the problem of 3D reconstruction of a rigid scene from multiple images captured by an uncalibrated perspective camera. In particular, we investigate camera knowledge that is generally available in this context but is yet to be fully exploited in camera autocalibration. Such camera knowledge may arise from priors on the internal geometry of the camera or its motion in the image acquisition process. We describe these priors in the following.

Sensor geometry priors: While the camera intrinsic parameters are not known in the uncalibrated SfM approach, reasonable assumptions can usually be made about some of them. Their physical interpretation allows us to distinguish such parameters. The focal length corresponds to the distance between the optical center and the image plane. This parameter varies with a change in zoom or focus and is unknown in the absence of calibration. The principal point gives the coordinates of the intersection of the optical axis with the image plane. This point is not necessarily at the image center due to optical misalignments and other manufacturing defects. Its location is also known to vary with the focal length. Therefore, the principal

point is also generally not known. The remaining two parameters, the skew and the aspect ratio, describe the pixel shape. These parameters remain quite stable despite a change in the zoom or focus. Moreover, most cameras in the market are manufactured such that the pixels are usually (close to) square, *i.e*. the skew is zero and the aspect ratio is one. Thus, these two parameters can safely be assumed to be known in practice.

Image acquisition priors: When capturing images for 3D reconstruction, the camera motion is implicitly constrained by the requirement to match feature points; there must be sufficient overlap between the images to compute feature correspondences. To this end, SfM packages generally recommend capturing images smoothly in a loop around the scene and changing the viewpoint mildly[7]. Usually, such camera motion is naturally performed when capturing the scene in a video. Thus, some knowledge, albeit vague, of the camera motion is implicitly available from the image acquisition process.

Based on the priors described above, partial knowledge of the camera geometry or motion is generally available in the uncalibrated SfM context. While we were quite specific about the partial camera geometry knowledge, the motion knowledge is yet to be made more explicit. In this thesis, we aim to identify and exploit such partial camera knowledge in camera autocalibration. We detail our contributions in this regard in the following section.

1.3 Contributions

We describe here the contributions of this thesis in exploiting partial knowledge of the camera geometry and motion in camera autocalibration.

[7] `https://www.3dflow.net/technology/documents/photogrammetry-how-to-acquire-pictures/`

1.3.1 Exploiting partial camera geometry knowledge

From the discussion in Section 1.2, most modern cameras have square pixels. Such cameras are said to have a Euclidean Image Plane (EIP) [Heyden and Åström 1997]. Hence, a camera can safely be assumed to have an EIP. While this practical assumption has been used in several camera autocalibration methods, it is yet to be fully exploited in stratified autocalibration. Our main contribution is to formulate a new quartic polynomial equation in the coordinates of the plane at infinity assuming a camera with EIP and constant intrinsic parameters. This polynomial, referred to as the EIP polynomial, is obtained for each image pair. Combined with an existing pairwise quartic equation, the so-called modulus constraint [Pollefeys and Van Gool 1999], each image pair thus provides two constraints on the location of the plane at infinity. Two image pairs are then sufficient to obtain a unique solution in general. We propose a stratified camera autocalibration method that uses the EIP polynomial in conjunction with the modulus constraint to estimate the plane at infinity. Our experiments with synthetic data and real images show the benefits of the EIP polynomial in our proposed method and the improved reliability compared to the state-of-the-art methods, especially with short image sequences. Our work on exploiting the EIP assumption in stratified camera autocalibration has been published in the British Machine Vision Conference (BMVC) 2020:

– Devesh Adlakha, Adlane Habed, Fabio Morbidi, Cédric Demonceaux, and Michel de Mathelin (2020). “Stratified autocalibration of cameras with Euclidean Image Plane". In: *British Machine Vision Conference (BMVC)*.

1.3.2 Exploiting partial camera motion knowledge

As discussed in Section 1.2, there is an implicit knowledge of the camera motion in the image-based 3D reconstruction context, specifically, that the viewpoint is changed mildly when capturing images. We formalize this vague camera motion knowledge and present some contributions in exploiting it in stratified camera autocalibration. We consider the case of a moving camera with constant intrinsic parameters and assume that a vague knowledge of the relative camera orientation angle is available. Under these assumptions, we derive two new sets of pairwise

constraints on the location of the plane at infinity. The constraints in the first set are convex and formulated as Linear Matrix Inequalities (LMIs). We show that the plane at infinity belongs to one of two convex sets, defined by the LMIs, depending on whether the relative orientation angle[8] θ_{ij} between a camera pair (i, j) is either $|\theta_{ij}| \leq 120^\circ$ or $|\theta_{ij}| \geq 120^\circ$. These two convex sets describe the geometric relationship between the plane at infinity and new geometric objects, the *hodographs of the horopter*. Furthermore, a vague knowledge of the relative orientation angle, *i.e.* $|\theta_{ij}| \leq 120^\circ$ or $|\theta_{ij}| \geq 120^\circ$, between a set of a camera pairs can be exploited with the LMI constraints to recover a new quasi-affine reconstruction of a scene that we refer to as a QUARCH: QUasi-Affine Reconstruction with respect to Camera centers and the Hodographs of horopters. A QUARCH is a specialization of the existing QUARC (QUasi-Affine Reconstruction with respect to Camera centers) [Nistér 2004b]. In the considered scenario of image-based 3D reconstruction, the assumption of $|\theta_{ij}| \leq 120^\circ$ is typically satisfied between consecutive views, and we exploit this knowledge to obtain a QUARCH.

The second set of constraints on the plane at infinity are non-convex and formulated as polynomial inequalities. These exploit tighter knowledge of the relative orientation angle, *i.e.* $|\theta_{ij}| < 90^\circ/k$, where k is a positive integer. Unlike the convex constraints, these are derived purely algebraically using the *Hurwitz stability criterion*. The polynomial inequalities are of degree $2k$, thus increasingly tighter orientation knowledge leads to higher degree polynomials. When combined with the convex constraints, these inequalities allow us to recover a specialized QUARCH. To obtain this reconstruction tighter yet still vague knowledge of $|\theta_{ij}| < 90^\circ/k$ for some positive integer k is required for a set of camera pairs.

We propose a stratified camera autocalibration method that exploits a vague camera orientation knowledge to locate the plane at infinity. Our method relies on a QUARCH as a first step towards recovering the affine and metric reconstructions. The final main contribution is a constrained Levenberg-Marquardt (LM) method for nonlinear optimization subject to LMI constraints. This method ensures that the QUARCH constraints are satisfied during the local optimization used to locate the plane at infinity. Our experiments confirm the benefits of incorporating the

[8] The relative camera orientation angle θ_{ij} is about an arbitrary axis in the axis-angle representation and $\theta_{ij} \in [-180^\circ, 180^\circ]$.

orientation-based constraints. We present experimental evaluations using both synthetic data and real images.

Part of the contributions described above have been published in the IEEE/CVF International Conference on Computer Vision (ICCV) 2019:

– Devesh Adlakha, Adlane Habed, Fabio Morbidi, Cédric Demonceaux, and Michel de Mathelin (2019). “QUARCH: A new quasi-affine reconstruction stratum from vague relative camera orientation knowledge". In: *International Conference on Computer Vision (ICCV)*.

These include the convex constraints, their geometric interpretation through the hodographs, the QUARCH stratum, and the constrained LM method. The non-convex constraints and the specialized QUARCH extend the published results. A paper synthesizing these contributions is currently in preparation and planned to be submitted to either the International Journal of Computer Vision (IJCV) or the IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI).

1.4 Organization

The rest of this manuscript is organized as follows.

Chapter 2 reviews some background concepts of multi-view geometry and surveys the relevant camera autocalibration literature.

Chapter 3 presents our work on exploiting a partial knowledge of the camera geometry in camera autocalibration.

Chapter 4 presents our work on exploiting a partial knowledge of the camera motion in camera autocalibration.

Chapter 5 discusses the research contributions of this thesis and some future research directions.

Appendix A provides an overview of some optimization tools that we rely on throughout this thesis.

Appendix B provides the expression of a line projection matrix.

2 Uncalibrated 3D Vision

This chapter reviews some relevant background concepts in 3D computer vision. We describe the pinhole model of a camera, followed by the geometric relations between two views of a scene, including a planar one. We then discuss the recovery of a 3D reconstruction from multiple uncalibrated images of a scene. The different reconstructions that can be obtained are detailed. Finally, we describe the camera autocalibration problem and the geometric entities involved and then survey the relevant literature.

2.1 Introduction

A central problem in computer vision is to recover the 3D structure and cameras from multiple images of a scene. With calibrated cameras, a reconstruction up to some unknown displacement and scale can be obtained. With uncalibrated cameras, a key result is that the 3D points and cameras can be recovered but only up to a projective ambiguity. With additional information, the ambiguity can be reduced to a more specialized transformation. These aspects of multi-view geometry are described in this chapter. We then detail the camera autocalibration problem that aims to retrieve the camera calibration from uncalibrated images. The autocalibration literature is also surveyed.

The theory of multi-view geometry has been studied extensively in computer vision in the last few decades. This is now a mature subject, and the fundamental results are documented in several texts [Faugeras et al. 2001; Hartley and Zisserman 2004; Ma et al. 2004]. This chapter aims to review the results from the literature that are relevant to our contributions (presented in Chapters 3 and 4). We assume a basic working knowledge of projective geometry. The books mentioned above are an excellent reference for such background material. The notation and presentation in this chapter are also inspired by [Hartley and Zisserman 2004].

Organization: The rest of this chapter is divided into two main sections. Section 2.2 reviews some notions of multi-view geometry, and Section 2.3 is devoted to camera autocalibration. Section 2.4 concludes this chapter.

2.2 Multi-view 3D reconstruction

Starting from the perspective projection model, we progress toward the two-view relationships and, eventually, the multi-view reconstruction problem.

2.2.1 Perspective projection

We follow the pinhole model of a camera. In this model, a camera is given by its center of projection and an image plane. A scene point X with homogeneous co-

ordinates X (a 4-vector) projects to a point x in the image plane with homogeneous coordinates x (a 3-vector). This projection is given by,

$$\mathsf{x} \simeq \mathsf{PX}, \tag{2.1}$$

where P is a 3×4 projection matrix that represents the camera and $\simeq$ denotes the equality up to scale. The perspective image projection is thus a linear transformation. The projection matrix P can be factored as follows,

$$\mathsf{P} = [\, \mathsf{KR} \mid -\mathsf{KRt} \,], \tag{2.2}$$

where the rotation matrix R and the translation vector t relate the world reference frame to the camera reference frame. These are known as the camera extrinsic parameters. The matrix K is upper-triangular and contains the camera intrinsic parameters,

$$\mathsf{K} = \begin{bmatrix} f_x & \gamma & u \\ 0 & f_y & v \\ 0 & 0 & 1 \end{bmatrix}, \tag{2.3}$$

that include the focal lengths (f_x, f_y), the principal point coordinates $(u,\, v)$, and the skew γ. These parameters depend only on the internal geometry of the camera. The ratio $\tau = f_x/f_y$ is known as the aspect ratio. The projection matrix P is of row rank 3. Any nonzero 4-vector C of its right null space represents the camera center C (*i.e.* the center of projection). Writing $\mathsf{P} = [\mathsf{p}_1\ \mathsf{p}_2\ \mathsf{p}_3\ \mathsf{p}_4]$, where p_i is the ith column of P, an algebraic expression of C is given by,

$$\mathsf{C} = \begin{bmatrix} \det([\mathsf{p}_2\ \mathsf{p}_3\ \mathsf{p}_4]) \\ -\det([\mathsf{p}_1\ \mathsf{p}_3\ \mathsf{p}_4]) \\ \det([\mathsf{p}_1\ \mathsf{p}_2\ \mathsf{p}_4]) \\ -\det([\mathsf{p}_1\ \mathsf{p}_2\ \mathsf{p}_3]) \end{bmatrix}. \tag{2.4}$$

By C we hereafter refer to this exact expression.

2.2.2 Planar projection

Points on a scene plane $\mathit{\Pi}$ map to an image plane via a projective transformation called a (plane) homography. A homography is represented by a nonsingular 3×3 matrix H. In the following, we give the exact expressions of homography matrices that we use throughout.

Inter-image homography: An inter-image homography maps image projections of co-planar scene points from one image to another. Consider projection matrices of the form, $\mathsf{P}_i = [\ \mathsf{A}_i \mid \mathsf{a}_i\ ]$ with $\mathsf{A}_1 = \mathsf{I}_3$ and $\mathsf{a}_1 = \mathbf{0}_3$, where A_i is a 3×3 matrix, a_i is a 3-vector, I_3 is the 3×3 identity matrix, and $\mathbf{0}_3$ the 3-vector of zeros. It was shown in [Habed et al. 2012] that given P_i, all inter-image homographies induced by planes not containing the origin of the reference frame are linear functions of a real 3-vector π,

$$\mathsf{H}_{ij} = \mathsf{A}_j\mathsf{A}_i^* - \mathsf{A}_j[\ \pi\ ]_\times\, \mathsf{A}_i^\intercal[\ \mathsf{a}_i\ ]_\times^\intercal - \mathsf{a}_j\pi^\intercal\mathsf{A}_i^* \ \text{ for all } \ i \neq j, \tag{2.5}$$

where A^* is the adjugate of A (the transpose of the cofactor matrix) and $[\ \pi\ ]_\times$ denotes the skew-symmetric matrix associated with the cross-product of vector π. The more usual forward, $\mathsf{H}_{1i} = \mathsf{A}_i - \mathsf{a}_i\pi^\intercal$, and inverse, $\mathsf{H}_{i1} = \mathsf{A}_i^* - [\ \pi\ ]_\times\mathsf{A}_i^\intercal[\ \mathsf{a}_i\ ]_\times^\intercal$, mappings relating any view i and the reference image can be extracted from (2.5). Note that,

$$\mathsf{H}_{i1} = \mathsf{H}_{1i}^* = \det(\mathsf{H}_{1i})\mathsf{H}_{1i}^{-1}, \tag{2.6}$$

where $\det(\mathsf{H}_{1i})$ is the determinant of H_{1i}, and that H_{ij} in (2.5) is obtained through $\mathsf{H}_{ij} = \mathsf{H}_{1j}\mathsf{H}_{i1}$.

Infinite homography: For a fixed π, all H_{ij} represent inter-image homographies induced by the same plane. In particular, for some appropriate $\pi = \pi_\infty$, the inter-image homograhies expressed by (2.5) and denoted hereafter by $\mathsf{H}_{\infty ij}$, are those induced by the plane at infinity $\mathit{\Pi}_\infty$. Such a homography $\mathsf{H}_{\infty ij}$ is referred to as an

infinite homography. It is distinctively independent of the camera translation. In particular, when the camera intrinsic parameters are constant,

$$\mathsf{H}_{\infty ij} = \mathsf{A}_j \mathsf{A}_i^* - \mathsf{A}_j [\, \pi_\infty \,]_\times \mathsf{A}_i^\intercal [\, \mathsf{a}_i \,]_\times^\intercal - \mathsf{a}_j \pi_\infty^\intercal \mathsf{A}_i^*, \tag{2.7a}$$

$$= \lambda_i^2 \lambda_j \, \mathsf{K} \mathsf{R}_{ij} \, \mathsf{K}^{-1}, \tag{2.7b}$$

for all $i \neq j$, where R_{ij} is the rotation matrix between cameras i and j, and K is the intrinsic parameters matrix. The scalar λ_i is such that $\mathsf{H}_{\infty 1i} = \lambda_i \mathsf{K} \mathsf{R}_{1i} \, \mathsf{K}^{-1}$ with $\lambda_1 = 1$.

2.2.3 Epipolar geometry

The epipolar geometry describes the geometric relations between two views of the same scene. This geometry is encapsulated in the so-called Fundamental and Essential matrices.

Fundamental matrix: Suppose a scene point X projects to a point x_1 with coordinates x_1 in image 1, and to a point x_2 with coordinates x_2 in image 2. The corresponding image projections x_1 and x_2 satisfy the so-called epipolar constraint, given by,

$$\mathsf{x}_2^\top \mathsf{F}_{12} \mathsf{x}_1 = 0, \tag{2.8}$$

where F_{12} is a 3×3 matrix of rank 2 known as the Fundamental matrix. Geometrically, equation (2.8) imposes the constraint that the corresponding point of x_1 in the second image lies on a line. This line, known as the epipolar line, is given by $\mathsf{F}_{12}\mathsf{x}_1$. A similar relationship holds for the corresponding point of x_2 in the first image, given by $\mathsf{F}_{21} = \mathsf{F}_{12}^\top$. Each image point correspondence provides one linear equation in the entries of F_{12} from (2.8). With at least eight such correspondences, F_{12} can be estimated linearly. Note that only the image correspondences are used, *i.e.* no knowledge of the cameras or the scene is required. The details of this linear eight-point algorithm are given in [Hartley and Zisserman 2004, Ch. 11]. Finally, the Fundamental matrix can be expressed as [Hartley and Zisserman 2004, Result 9.1],

$$\mathsf{F}_{12} \simeq [\mathsf{e}_{12}]_\times \mathsf{H}_{12}, \tag{2.9}$$

where e_{12} is the epipole in the second image, *i.e.* the projection of the first camera on the second image plane, and H_{12} is the inter-image homography induced by any plane $\mathit{\Pi}$, including the plane at infinity $\mathit{\Pi}_\infty$.

Essential matrix: When the cameras are calibrated, the epipolar geometry is described by a specialization of the Fundamental matrix known as the Essential matrix. The Essential matrix E_{12} is related to the Fundamental matrix F_{12} as,

$$\mathsf{E}_{12} \simeq \mathsf{K}_2^\top \mathsf{F}_{12} \mathsf{K}_1, \tag{2.10}$$

where K_1 and K_2 are the calibration matrices of the first and second cameras, respectively. The Essential matrix inherits the rank 2 property from the Fundamental matrix. Its distinguishing property is that its two nonzero singular values are equal. The Essential matrix verifies the epipolar constraint for image points expressed in normalized image coordinates, *i.e.* for $\mathsf{K}_1^{-1}\mathsf{x}_1$ and $\mathsf{K}_2^{-1}\mathsf{x}_2$.

2.2.4 Projective reconstruction

We are interested here in recovering the camera projection matrices and the scene structure from multiple uncalibrated images. We describe a factorization approach to this end.

Factorization-based reconstruction: Consider m scene points X_j imaged by n cameras with projection matrices P_i. The image projections x_{ij} are given by,

$$\lambda_{ij}\mathsf{x}_{ij} = \mathsf{P}_i\mathsf{X}_j, \qquad i = 1, \ldots, n, \quad j = 1, \ldots, m, \tag{2.11}$$

where λ_{ij} are the projective scale factors, referred to as the projective depths. The goal is to estimate the projection matrices P_i and the coordinates X_j of the scene

points from the image measurements x_{ij}. The projection equations can be written in matrix form as folllows,

$$\begin{bmatrix} \lambda_{11}\mathsf{x}_{11} & \lambda_{12}\mathsf{x}_{12} & \dots & \lambda_{1m}\mathsf{x}_{1m} \\ \lambda_{21}\mathsf{x}_{21} & \lambda_{22}\mathsf{x}_{22} & \dots & \lambda_{2m}\mathsf{x}_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{n1}\mathsf{x}_{n1} & \lambda_{n2}\mathsf{x}_{n2} & & \lambda_{nm}\mathsf{x}_{nm} \end{bmatrix} = \begin{bmatrix} \mathsf{P}_1 \\ \mathsf{P}_2 \\ \vdots \\ \mathsf{P}_n \end{bmatrix} \begin{bmatrix} \mathsf{X}_1 & \mathsf{X}_2 & \dots & \mathsf{X}_m \end{bmatrix}. \tag{2.12}$$

The matrix on the left-hand side in (2.12) is known as the measurement matrix, and we denote it by W. Assuming that W is known, it can be factored, as in (2.12), into the product of a $3n \times 4$ matrix stacking all P_i and a $4 \times m$ matrix stacking the coordinate vectors X_j of all the scene points. We see then that W is of rank 4. With noisy measurements, this will usually not be the case, so we can manually set all but the first four singular values to zero. By the singular value decomposition of this rank-4 W, we have $\mathsf{W} = \mathsf{U}\mathsf{D}\mathsf{V}^\top$. The camera projection matrices and scene points can then be recovered from this decomposition as,

$$\begin{bmatrix} \mathsf{P}_1^\top & \mathsf{P}_2^\top & \dots & \mathsf{P}_n^\top \end{bmatrix}^\top = \mathsf{U}\mathsf{D}, \quad \begin{bmatrix} \mathsf{X}_1 & \mathsf{X}_2 & \dots & \mathsf{X}_m \end{bmatrix} = \mathsf{V}^\top. \tag{2.13}$$

Note that the cameras and the points are reconstructed only up to a projective transformation. This means that a reconstruction comprising $\mathsf{P}_i\mathsf{H}$ and $\mathsf{H}^{-1}\mathsf{X}_j$, where H is a projective transformation, is also a valid factorization of W. In other words, as the image projections remain the same with projection matrices $\mathsf{P}\mathsf{H}$ and point coordinates $\mathsf{H}^{-1}\mathsf{X}_j$, this reconstruction is an equally valid factorization of the measurement matrix. Such a reconstruction is related to the actual scene by a projective transformation, hence it is referred to as a projective reconstruction.

Estimating the projective depths: The projective depths λ_{ij} are, in fact, not known and must be estimated. Different strategies have been proposed in the literature to recover coherent projective depths. In [Sturm and Triggs 1996], they are estimated using Fundamental matrices. A projective reconstruction can then be obtained through factorization. Conversely, when P_i and X_j are known, the depths λ_{ij} can be recovered linearly. This prompted an iterative version of [Sturm and Triggs 1996] that alternates between computing the projective reconstruction and

the depths [Triggs 1996]. The projective depths can be initialized using Fundamental matrices or by setting them arbitrarily, typically as $\lambda_{ij} = 1$ for all i and j. Several variants of this iterative algorithm have been proposed [Mahamud and Hebert 2000; Mahamud et al. 2001]. The convergence of such algorithms was analyzed in [Oliensis and Hartley 2007]. The authors showed that these algorithms converge to the trivial solution. They also proposed an extension of the original algorithm, minimizing a regularized cost, that was proven to converge and avoid trivial solutions.

Practical considerations: Appropriately conditioning the data is essential for factorization-based methods. Different normalizations have been used in the literature, and some details can be found in the works cited above. The estimated projective reconstruction is usually refined through bundle adjustment [Triggs et al. 1999]. Factorization is an elegant solution to compute an initial projective reconstruction. Its severe limitation, however, is the visibility constraint that all the points are viewed in all the images. More recent methods [Magerand and Bue 2018; Kasten et al. 2019] deal with missing data, outliers, and large-scale reconstructions.

2.2.5 Reconstruction strata

The previous section showed a key result in uncalibrated computer vision: from point correspondences across multiple uncalibrated images alone, a reconstruction up to a projective transformation can be recovered. This result is for the most general case, where only the point matches between images are known. More specialized reconstructions can be recovered with additional information. In this section, we describe the different reconstruction strata, their properties, and the information required to compute the reconstructions. We focus, in particular, on the forms of the projection matrices in the different strata. In the following, we describe the projective stratum and then progress to the more specialized affine and metric strata.

The projective stratum

A projective reconstruction of a scene differs from a metric and an affine one by a projective transformation. Therefore, it only preserves properties invariant to pro-

jective transformations, such as collinearity and coplanarity. The Euclidean notions of angles and lengths are lost. In a projective reference frame, the canonical form of projection matrices is,

$$\begin{aligned} \mathsf{P}_1 &\simeq [\, \mathsf{I}_3 \mid \mathbf{0}_3 \,], \\ \mathsf{P}_i &\simeq [\, \mathsf{H}_{1i} \mid \mathsf{e}_{1i} \,], \qquad i = 1, \ldots, n, \end{aligned} \tag{2.14}$$

where H_{1i} is a homography between images 1 and i induced by an arbitrary plane $\mathit{\Pi}$, and e_{1i} is the epipole in the ith image. Projection matrices of the form (2.14) can be computed from point correspondences across uncalibrated images. Note that the reference frame is attached to the first camera. We assume this to be the case throughout the present section.

The affine stratum

Given a projective reconstruction computed from uncalibrated images, an affine reconstruction can be recovered by locating the plane at infinity $\mathit{\Pi}_\infty$. In an affine reference frame, $\mathit{\Pi}_\infty$ is at its canonical position given by the coordinates $(\mathbf{0}_3^\top, 1)^\top$. By identifying $\mathit{\Pi}_\infty$ in a projective reconstruction and bringing it to its canonical position, the projective reconstruction can be transformed to affine. Assuming $\mathit{\Pi}_\infty$ has the coordinates $\left(\pi_\infty^\top, 1\right)^\top$ in a projective reference frame, this transformation is of the form,

$$\mathsf{H}_{\mathsf{PA}} = \begin{bmatrix} \mathsf{I}_3 & \mathbf{0}_3 \\ -\pi_\infty^\top & 1 \end{bmatrix}. \tag{2.15}$$

By applying this transformation to the projection matrices (2.14), we obtain the canonical form of the affine projection matrices,

$$\begin{aligned} \mathsf{P}_1^{\mathsf{A}} &\simeq [\, \mathsf{I}_3 \mid \mathbf{0}_3 \,], \\ \mathsf{P}_i^{\mathsf{A}} &\simeq [\, \mathsf{H}_{\infty 1i} \mid \mathsf{e}_{1i} \,], \qquad i = 1, \ldots, n, \end{aligned} \tag{2.16}$$

where $\mathsf{H}_{\infty 1i}$ is the inter-image homography induced by $\mathit{\Pi}_\infty$, see (2.7). The scene points transform via $\mathsf{H}_{\mathsf{PA}}^{-1}$. An affine reconstruction of a scene differs from a metric one by an affine transformation. It preserves properties invariant to affine transformations, such as parallelism.

The metric stratum

A metric reconstruction differs from the actual scene by a metric transformation. It preserves metric properties, including angles and length ratios. Given affine projection matrices (2.16), the transformation taking them to a metric reference frame is of the form,

$$\mathsf{H}_{\mathsf{AM}} = \begin{bmatrix} \mathsf{K}_1 & \mathbf{0}_3 \\ \mathbf{0}_3^\top & 1 \end{bmatrix}. \tag{2.17}$$

where K_1 is the calibration matrix of the reference camera. The canonical form of the metric projection matrices is then,

$$\begin{aligned} \mathsf{P}_1^{\mathsf{M}} &\simeq [\, \mathsf{K}_1 \mid \mathbf{0}_3 \,], \\ \mathsf{P}_i^{\mathsf{M}} &\simeq [\, \mathsf{K}_i \mathsf{R}_{1i} \mid \mathsf{e}_{1i} \,], \qquad i = 1, \ldots, n, \end{aligned} \tag{2.18}$$

where R_{1i} is the rotation matrix between the reference and the ith camera. A further specialization is the Euclidean stratum, where additional properties such as lengths are preserved. Without information on the scale of the scene, only a metric reconstruction can be recovered (with a calibrated camera or via autocalibration).

2.2.6 Quasi-affine reconstruction

The previous section presented the three main reconstruction strata, *i.e.* projective, affine, and metric. In this section, we present another one that lies between the projective and affine strata, the so-called quasi-affine stratum. This stratum is particularly relevant to our contributions in Chapter 4, and so we describe it in more detail here. Roughly speaking, a quasi-affine reconstruction is a projective reconstruction that satisfies some affine properties. Hence, it is closer to an affine reconstruction than an arbitrary projective one. We describe two quasi-affine reconstructions that have been proposed in the literature. The first one [Hartley and Zisserman 2004] preserves the convex hull of the camera centers and that of the scene points, while the second one [Nistér 2004b] preserves the convex hull of the camera centers alone.

Quasi-affine with respect to cameras and scene points

Visible points in an image originate from scene points lying in front of the camera. This constraint is not a given in a projective reconstruction. Hartley introduced the theory of chirality [Hartley and Zisserman 2004, Ch. 21], *i.e*. the property that distinguishes the points lying in front of a camera from those lying behind it. By taking into account the chirality of visible points, he introduced a specialized projective reconstruction, referred to as a quasi-affine reconstruction. Such a reconstruction can be obtained by solving a linear programming problem. In the following, we provide an overview of the computation of a quasi-affine reconstruction. We also describe how the location of $\mathit{\Pi}_\infty$ can be bounded in a quasi-affine reconstruction.

Sign correction: Consider a projective reconstruction with projection matrices P_i and scene points with coordinates X_j, where $i = 1, \ldots, n$ and $j = 1, \ldots, m$. The image projections x_{ij} are given by, $\lambda_{ij}\mathsf{x}_{ij} = \mathsf{P}_i\mathsf{X}_j$, where λ_{ij} is the projective scale factor, as in (2.11). The goal is to upgrade this projective reconstruction to quasi-affine. To do so, the first step is to correct the signs of the projection matrices and the points. Assuming all the scene points are visible in all the images[1], *i.e*. all the points lie in front of all the cameras, λ_{ij} should have the same sign for all i and j, say $\lambda_{ij} > 0$. This is not guaranteed to be the case in an arbitrary projective reconstruction. The signs of each P_i and X_j can be corrected such that $\lambda_{ij} > 0$. This is only a matter of multiplying some P_i and X_j by -1, as described next. Let $\widetilde{\mathsf{P}}_i$ and $\widetilde{\mathsf{X}}_j$ be the sign-corrected projection matrices and point coordinates. These can be obtained from P_i and X_j as follows. Fixing one camera, say $\widetilde{\mathsf{P}}_1 = \mathsf{P}_1$, set $\widetilde{\mathsf{X}}_j = \mathsf{X}_j$ or $\widetilde{\mathsf{X}}_j = -\mathsf{X}_j$ such that $\widetilde{\mathsf{P}}_1\widetilde{\mathsf{X}}_j = \lambda_{1j}\mathsf{x}_{1j}$ with $\lambda_{1j} > 0$ for all j. Similarly, a point can be used to correct the signs of all P_i to obtain $\widetilde{\mathsf{P}}_i$. In the following, we drop the tilde notation and use P_i and X_j as the sign-corrected quantities.

Chirality inequalities: To transform a (sign-corrected) projective reconstruction to quasi-affine, a plane not cutting through the set of camera centers and the set of scene points must be identified. Such a plane $\mathit{\Pi}$ with coordinates Π acts as a

[1] We make this assumption only for convenience in the explanation.

surrogate for the unknown $\mathit{\Pi}_\infty$, and it satisfies [Hartley and Zisserman 2004, Sec. 21.7],

$$\alpha\big(\Pi^\top \mathsf{C}_i\big)\big(\Pi^\top \mathsf{X}_j\big) > 0, \quad i = 1, \ldots, n, \quad j = 1, \ldots, m, \tag{2.19}$$

where C_i is the coordinate vector of the ith camera's center, and α is the sign of the determinant of the quasi-affine upgrade transformation. There are two possible realizations as $\alpha = \pm 1$. As Π is homogeneous, we can assume $\alpha\big(\Pi^\top \mathsf{C}_1\big) > 0$. Thus, the following inequalities are obtained,

$$\begin{aligned} \alpha\Pi^\top \mathsf{C}_i > 0, \quad & i = 1, \ldots, n \\ \Pi^\top \mathsf{X}_j > 0, \quad & j = 1, \ldots, m. \end{aligned} \tag{2.20}$$

Inequalities (2.20) are known as the chirality inequalities. A plane satisfying these inequalities preserves the convex hull of the set of camera centers and the set of scene points, though not necessarily of their union. The value of α is not known, and so solutions for the two cases $\alpha = 1$ and $\alpha = -1$ must be determined.

Quasi-affine upgrade: By identifying a plane satisfying the chirality inequalities (2.20), a quasi-affine reconstruction can be recovered from a projective one. As the inequalities are linear, a Linear Programming (LP) problem can be formulated to determine such a plane. Hartley proposed the following LP problem,

$$\begin{aligned} \max_{\Pi,\,\delta} \quad & \delta \\ \text{s.t.} \quad & \alpha\Pi^\top \mathsf{C}_i > \delta, && i = 1, \ldots, n, \\ & \Pi^\top \mathsf{X}_j > \delta, && j = 1, \ldots, m, \\ & -1 \leq (\Pi)_k \leq 1, && k = 1, \ldots, 4, \end{aligned} \tag{2.21}$$

where δ is a scalar. Problem (2.21) maximizes the extent to which each of the chirality inequalities is satisfied. This prevents the plane from being arbitrarily close to the camera centers and the scene points. The last inequality in (2.21) ensures that the problem is bounded. The LP problem (2.21) is solved for both $\alpha = 1$

and $\alpha = -1$. Provided the last coordinate of the estimated plane is nonzero, the following transformation,

$$\mathsf{H}_\alpha = \begin{bmatrix} \pm\mathsf{I}_3 & \mathbf{0}_3 \\ \Pi^\top & \end{bmatrix}, \tag{2.22}$$

upgrades the scene points to a quasi-affine reference frame, where H_α is such that the sign of its determinant is α. The projection matrices are transformed using its inverse. If problem (2.21) provides a solution for both $\alpha = 1$ and $\alpha = -1$, then two realizations are obtained.

Bounding the plane at infinity: In a quasi-affine reconstruction, bounds on the coordinates of $\mathit{\Pi}_\infty$ can be obtained. To do so, the points and cameras are first translated such that the origin is inside their convex hull. As a result, $\mathit{\Pi}_\infty$ cannot pass through the origin. Otherwise, it crosses the convex hull. The last coordinate of $\mathit{\Pi}_\infty$ is then nonzero and can be fixed to one. Upper and lower bounds on the other coordinates can be obtained by minimizing and maximizing each one subject to the chirality inequalities, *i.e.* by solving six LP problems. For the kth coordinate of $\mathit{\Pi}_\infty$ ($k = 1, 2, 3$), the LP problems are,

$$\begin{aligned} &\min / \max\ (\Pi)_k \\ &\text{s.t.} \qquad \Pi^\top \mathsf{C}_i > 0, \qquad i = 1, \ldots, n, \\ &\qquad\quad \Pi^\top \mathsf{X}_j > 0, \qquad j = 1, \ldots, m. \end{aligned} \tag{2.23}$$

The location of $\mathit{\Pi}_\infty$ is thus confined to a bounded volume. Since all the points and camera centers are on one side with respect to the sought plane, note that the chirality inequalities (2.20) are written as in (2.23). Details of the recommended data preprocessing can be found in [Hartley and Zisserman 2004].

Quasi-affine with respect to cameras

Nistér proposed to compute a quasi-affine reconstruction with respect to cameras alone [Nistér 2004b]. The resulting reconstruction is referred to as a QUARC: QUasi-Affine Reconstruction with respect to Camera centers. A QUARC preserves the convex hull of the set of camera centers. The argument for using only the cameras is that they are more reliable than the scene points. Projection matrices

are generally estimated using many feature correspondences and robust algorithms. Relying on every single scene point when solving the chirality inequalities, on the other hand, can lead to an infeasible problem due to erroneous or inaccurate points. In [Nistér 2004b], a robust method for correcting the signs of the cameras was also introduced. The steps to compute a QUARC from a projective reconstruction are summarized as follows:

(i) Correct the signs of the projection matrices as $\widetilde{\mathsf{P}}_i = \zeta_i \mathsf{P}_i$, where $\zeta_i \in \{-1, 1\}$ and $i = 1, \ldots, n$, such that visible scene points in pairs of images lie on the same side with respect to the cameras.

(ii) Determine a plane $\mathit{\Pi}$ that satisfies the chirality inequalities for the camera centers, *i.e.* $\Pi^\top \mathsf{C}_i > 0$ for $i = 1, \ldots, n$.

(iii) Upgrade the projective reconstruction to a QUARC using the estimated $\mathit{\Pi}$. The scene points are transformed using (2.22) with I_3 in the top-left block. The projection matrices are transformed by the inverse of this transformation.

The robust method for sign correction and the plane computation step, *i.e.* steps (i) and (ii) above, are detailed next.

Robust sign correction: Rather than relying on a single point to correct the signs of the cameras, as was described previously, Nistér proposed a robust method that corrects them such that a majority of points lie on the same side. Setting $\zeta_1 = 1$, the signs for the remaining cameras are computed as,

$$\zeta_i = \zeta_{i-1} \operatorname{sgn}\left(\frac{1}{2} + \sum_{j=1}^{m} \operatorname{sgn}\big((\mathsf{P}_i \mathsf{X}_j)_3 (\mathsf{P}_{i-1} \mathsf{X}_j)_3\big)\right), \qquad i = 2, \ldots, n, \tag{2.24}$$

where we added the $1/2$ term in (2.24) to prevent ζ_i from being zero. Note that we assumed, again, that all the points are visible in all images, but this assumption is not necessary. The sign-corrected projection matrices are then obtained as $\widetilde{\mathsf{P}}_i = \zeta_i \mathsf{P}_i$ for all i. We again drop the tilde notation in the following.

Computing a QUARC plane: Given sign-corrected projection matrices, a plane that does not cut through the convex hull of the camera centers can be used to re-

cover a QUARC. Such a plane is referred to as a QUARC plane. It can be computed by solving the following LP problem,

$$\begin{aligned} \max_{\Pi,\,\delta} \quad & \delta \\ \text{s.t.} \quad & \frac{\Pi^\top \mathsf{C}_i}{\|\mathsf{C}_i\|} > \delta, \qquad i = 1, \ldots, n, \\ & -1 \leq (\Pi)_k \leq 1, \qquad k = 1, \ldots, 4, \end{aligned} \tag{2.25}$$

where the vectors C_i are normalized to unit norm. Note the similarity between problems (2.25) and (2.21).

2.3 Camera autocalibration

Camera autocalibration entails retrieving the camera intrinsic parameters from multiple uncalibrated images of an unknown scene. Once the intrinsic parameters are obtained, a metric reconstruction of the imaged scene can be computed. Alternatively, from image point correspondences across uncalibrated images alone, a projective reconstruction can first be computed. The task at hand is then to recover a metric reconstruction from the projective one. Methods have been proposed for both these approaches. In this section, we review the basics of camera autocalibration and survey some of the relevant results in the literature. In Section 2.3.1, we establish the autocalibration problem statement in terms of recovering a metric reconstruction from a projective one. In Section 2.3.2, we describe the key geometric object in autocalibration, the so-called Absolute Conic, and review some fundamental relations. In Sections 2.3.3 and 2.3.4, we discuss two broad categories of autocalibration methods: direct and stratified methods. Finally, in Section 2.3.5 we survey autocalibration methods incorporating knowledge of the camera motion.

2.3.1 Problem statement

We consider a projective reconstruction with projection matrices P_i, $i = 1, \ldots, n$ and scene points X_j, $j = 1, \ldots, m$. Such a reconstruction can be computed from multiple uncalibrated images, as described in Section 2.2.4. We assume the refer-

ence frame is attached to the first camera, *i.e.* $\mathsf{P}_1 = [\ \mathsf{I}_3 \mid \mathbf{0}_3\ ]$. This can be done by post-multiplying each projection matrix P_i by the inverse of the following matrix,

$$\begin{bmatrix} \mathsf{P}_1 \\ \mathbf{0}_3^\top \quad 1 \end{bmatrix}, \tag{2.26}$$

To upgrade P_i and X_j to their metric counterparts P_i^M and X_j^M, respectively, we seek a projective transformation H such that,

$$\begin{aligned} \mathsf{P}_i^\mathsf{M} &= \mathsf{P}_i\mathsf{H}^{-1}, & i &= 1, \ldots, n, \\ \mathsf{X}_j^\mathsf{M} &= \mathsf{H}\mathsf{X}_j, & j &= 1, \ldots, m. \end{aligned} \tag{2.27}$$

where H is a nonsingular 4×4 matrix. It has the following form [Hartley and Zisserman 2004, Result 19.1],

$$\mathsf{H} = \begin{bmatrix} \mathsf{K}^{-1} & \mathbf{0}_3 \\ \pi_\infty^\top & \lambda \end{bmatrix}, \tag{2.28}$$

where K is the calibration matrix of the first camera, π_∞ contains the first three coordinates of the plane at infinity $\mathit{\Pi}_\infty$, *i.e.* $\Pi_\infty = \left(\pi_\infty^\top, 1\right)^\top$, and λ is a nonzero scalar that fixes the scale of the reconstruction. As the scale of a metric reconstruction can be fixed arbitrary, we set $\lambda = 1$ in the following. Thus, upgrading a projective reconstruction to metric involves estimating $\mathit{\Pi}_\infty$ and the intrinsic parameters of the reference camera.

2.3.2 Absolute conic

Geometrically, upgrading a projective reconstruction to metric is equivalent to identifying a special virtual conic on $\mathit{\Pi}_\infty$ that is referred to as the Absolute Conic (AC) and denoted by $\mathit{\Omega}_\infty$. In a metric reference frame, a point X with coordinates $\mathsf{X} = (\mathsf{X}_1, \mathsf{X}_2, \mathsf{X}_3, 0)^\top$ lying on $\mathit{\Omega}_\infty$ verifies the equation,

$$\mathsf{X}^\top \Omega_\infty \mathsf{X} = \mathsf{X}_1^2 + \mathsf{X}_2^2 + \mathsf{X}_3^2 = 0, \tag{2.29}$$

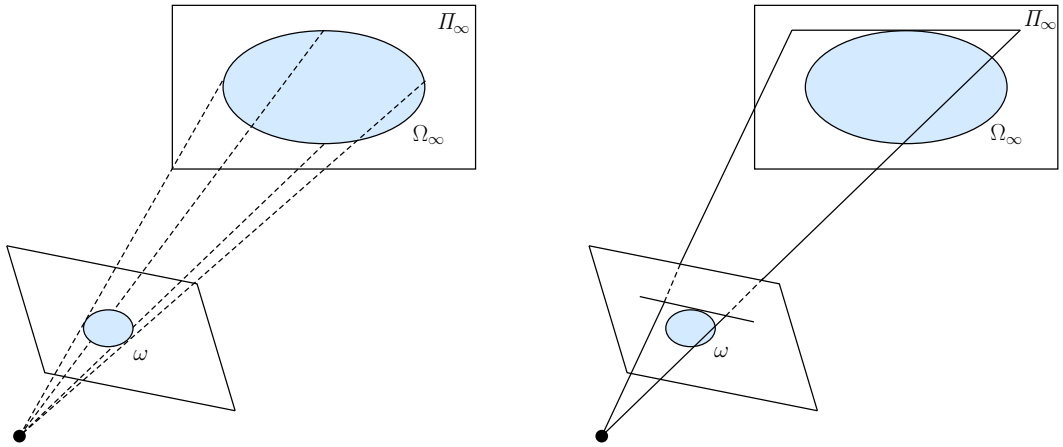


Figure 2.1: The Absolute Conic (AC) and its image projection. Left: the projection of the AC on the image plane. Right: A line tangent to the image of the AC (IAC). The lines tangent to the IAC comprise the dual image of the AC (DIAC).

where Ω_∞ is a 3×3 symmetric matrix representing the AC locally on Π_∞. The metric representation of the AC is therefore, $\Omega_\infty = \mathsf{I}_3$. Furthermore, X projects on the image as a point x with coordiantes x that verifies,

$$\mathsf{x}^\top \omega \mathsf{x} = 0, \tag{2.30}$$

where ω is a 3×3 symmetric matrix given as,

$$\omega = \mathsf{K}^{-\top}\mathsf{K}^{-1}. \tag{2.31}$$

The matrix ω represents the image of the AC (IAC). The IAC is a virtual image conic. Crucially, its matrix ω depends solely on the intrinsic parameters of the imaging camera. Figure 2.1 (left) shows an illustration of the AC and its image projection. The dual of the IAC (DIAC) is an imaginary line conic formed by the set of lines tangent to w. Figure 2.1 (right) shows an illustration of one such tangent line that back-projects to a plane that is tangent to the AC. The DIAC is denoted by ω^* and represented by the matrix ω^{-1}, and as such,

$$\omega^* = \mathsf{K}\mathsf{K}^\top. \tag{2.32}$$

The matrices ω and ω^* are symmetric and positive definite. The intrinsic parameters matrix K can be extracted from ω^* through its Cholesky factorization. Thus, locating the IAC or the DIAC is equivalent to recovering the camera intrinsic parameters.

As Ω_∞ lies on Π_∞, its projections transfer from one image to another via the inter-image homography induced by Π_∞. The IAC ω_1 in the first image is mapped to ω_i in image i via the infinite homography $\mathsf{H}_{\infty 1i}$, and similarly the DIAC ω_1^* in the first image is mapped to ω_i^* in image i. These relations are given as,

$$\mathsf{H}_{\infty 1i}^{-\top}\,\omega_1\,\mathsf{H}_{\infty 1i}^{-1} \;\simeq\; \omega_i, \tag{2.33a}$$

$$\mathsf{H}_{\infty 1i}\,\omega_1^*\,\mathsf{H}_{\infty 1i}^{\top} \;\simeq\; \omega_i^*. \tag{2.33b}$$

Note that (2.33b) is simply the inverse of (2.33a). Given sufficiently many images and known infinite homographies (*i.e.* Π_∞ is identified), the equations (2.33) can be used to linearly estimate the IAC or the DIAC and thereby recover the intrinsic parameters. We give further details of this approach in Section 2.3.4. In general, the infinite homographies are not known and must be determined by locating Π_∞.

2.3.3 Direct autocalibration

There are two main approaches in direct camera autocalibration. In the first approach, methods require the epipolar geometry between pairs of views to be known, and they solve directly for the intrinsic parameters. The earliest autocalibration methods in computer vision belong to this approach. In the second approach, methods instead require a projective reconstruction to be computed. They then directly upgrade this reconstruction to a metric one by simultaneously estimating Π_∞ and the camera intrinsic parameters. In this section, we describe the key autocalibration constraints that have been used in these two approaches and discuss some relevant methods in the literature. We present the so-called Kruppa's equations and some methods relying on them. These belong to the first kind of approach. We then present two geometric objects that have been used in direct autocalibration, the Dual Absolute Quadric (DAQ) and the Absolute Line Quadric (ALQ). We discuss their properties and some methods relying on them. These fit into the second approach discussed above. There are other well-known methods, notably those based on the

essential matrix [Mendonça and Cipolla 1999; Fusiello 2001; Fusiello et al. 2004], that we do not discuss here.

Kruppa's equations

The Kruppa's equations refer to two-view autocalibration constraints introduced in [Faugeras et al. 1992; Maybank and Faugeras 1992]. The equations are derived from the work of Erwin Kruppa[2] [Kruppa 1913], and are generally considered the first autocalibration constraints used in computer vision. For images 1 and 2, the Kruppa's equations are,

$$\mathsf{F}_{12}\omega_1^*\mathsf{F}_{12}^\top \simeq [\mathsf{e}_{12}]_\times \omega_2^* [\mathsf{e}_{12}]_\times, \tag{2.34}$$

where F_{12} is the Fundamental matrix between the two views, and e_{12} and ω_2^* are the epipole and DIAC, respectively, in the second image. Equation (2.34) can be obtained from (2.33b) by multiplying on the left and right by the matrix $[\mathsf{e}_{12}]_\times$, noting that $\mathsf{F}_{12} = [\mathsf{e}_{12}]_\times \mathsf{H}_{\infty 12}$. As such, only the epipolar geometry is required to use these equations. Roughly speaking, they represent the correspondence of epipolar lines tangent to ω_1^* and ω_2^*. A more precise description is given in [Hartley and Zisserman 2004, Sec. 19.4]. Hartley presented a simplified form of (2.34) in terms of the singular value decomposition of the Fundamental matrix [Hartley 1997a].

To use the Kruppa's equations or their simplified form for autocalibration, most methods assume a moving camera with constant intrinsic parameters. Each image pair provides two independent quadratic equations in ω^*. Hence, at least three images are required to obtain a finite number of solutions for the five unknowns in ω^*. In [Faugeras et al. 1992], a square system of equations was solved using homotopy continuation [Verschelde 1999; Sommese et al. 2005]. The overdetermined system of equations was solved using nonlinear least-squares optimization in [Zeller and Faugeras 1996]. For two views with an unknown and varying focal length but otherwise known intrinsic parameters, a closed-form solution for the focal length was derived in [Bougnoux 1998]. The more general varying parameters case was addressed using the simplified Kruppa's equations in [Lourakis 2000]. With three or more views, a major drawback of using these equations is that there

[2] An English translation of his paper is also available [Gallego et al. 2018].

are degenerate configurations specific to them [Sturm 2000]. The reason is that by eliminating $\mathit{\Pi}_\infty$ (*i.e.* $\mathsf{H}_{\infty 1i}$) from (2.33b), these equations do not enforce a common support plane for the AC. Hence, the transfer of the DIAC from one image to another via (2.34) is not unique, and so they provide weaker constraints than (2.33b) and suffer from additional degeneracies.

Dual Absolute Quadric

The *Dual Absolute Quadric* (DAQ), denoted by $\mathit{\Omega}^*_\infty$, is the dual of the AC $\mathit{\Omega}_\infty$. It is a degenerate quadric in $\mathbb{P}^{*3}$ consisting of the set of planes tangent to the AC. Figure 2.1 (right) shows one such tangent plane. The DAQ projects to the DIAC in the image as,

$$\mathsf{P}_i \Omega^*_\infty \mathsf{P}_i^\top \simeq \omega_i^*, \tag{2.35}$$

where P_i is the projection matrix of the ith image and Ω^*_∞ is a symmetric 4×4 matrix of rank 3 representing the DAQ. In an affine reference frame, the projection equation (2.35) is equivalent to (2.33b). The matrix Ω^*_∞ is of the form,

$$\Omega^*_\infty = \begin{bmatrix} \omega_1^* & -\omega_1^* \pi_\infty \\ -\pi_\infty^\top \omega_1^* & \pi_\infty^\top \omega_1^* \pi_\infty \end{bmatrix}, \tag{2.36}$$

where ω_1^* is the DIAC in the reference image and the coordinate vector of $\mathit{\Pi}_\infty$ is $\Pi_\infty = \left(\pi_\infty^\top, 1\right)^\top$. Depending on its scale, Ω^*_∞ is either positive or negative semidefinite. The appealing property of $\mathit{\Omega}^*_\infty$ is that its matrix encodes both $\mathit{\Pi}_\infty$ and the DIAC ω_1^*. In fact, the coordinate vector of $\mathit{\Pi}_\infty$ is the null space of Ω^*_∞. Hence, determining $\mathit{\Omega}^*_\infty$ is equivalent to determining both $\mathit{\Pi}_\infty$ and the camera calibration.

The DAQ was first used for camera autocalibration in [Heyden and Åström 1996; Triggs 1997]. Methods based on the DAQ rely on its projection equation (2.35). To estimate the DAQ, constraints on the intrinsic parameters, *i.e.* on ω_i^*, are translated to constraints on Ω^*_∞. The case of constant intrinsic parameters has been considered in [Heyden and Åström 1996; Triggs 1997; Bocquillon et al. 2007; Habed et al. 2014]. The projection equation (2.35) can also be exploited in the varying parameters case [Pollefeys et al. 1999; Chandraker et al. 2007]. In [Pollefeys et al. 1999], the problem is linearized by assuming an image-centered principal point and neglecting the rank-3 and semidefiniteness constraints in the initial estimation.

However, linearization leads to artificial degeneracies [Gurdjos et al. 2009]. More recent works [Bocquillon et al. 2007; Chandraker et al. 2007; Habed et al. 2014] therefore include the nonlinear rank and semidefiniteness constraints in the optimization. Once Ω_∞^* is estimated, $\mathit{\Pi}_\infty$ and ω_1^* can be recovered and the projective reconstruction upgraded to a metric one.

Absolute Line Quadric

The set of 3-space lines intersecting the AC form a special object, denoted by $\widetilde{\mathit{\Omega}}_\infty$, that has been referred to as the Absolute Line Quadric (ALQ) [Valdés et al. 2006] or as the Absolute Quadratic Complex [Ponce et al. 2005]. The lines in the ALQ intersect an image plane i in ω_i. Figure 2.1 (left) shows some of these lines. This leads to the projection equation,

$$\mathsf{M}_i \widetilde{\Omega}_\infty \mathsf{M}_i^\top \simeq \omega_i, \tag{2.37}$$

where $\widetilde{\Omega}_\infty$ is a 6×6 symmetric matrix representing $\widetilde{\mathit{\Omega}}_\infty$, and M_i is a 3×6 line projection matrix that maps a scene line to its image. A line projection matrix M_i and the more familiar point projection matrix P_i are conveniently related, and one can be obtained from the other as, $\mathsf{M}_i \simeq [\, \det(\mathsf{H}_{1i})\mathsf{H}_{1i}^{-\top} \mid -[\mathsf{e}_{1i}]_\times \mathsf{H}_{1i} \,]$, where $\mathsf{P}_i = [\, \mathsf{H}_{1i} \mid \mathsf{e}_{1i} \,]$ and H_{1i} is a 3×3 matrix (see Appendix B for this expression). Line correspondences are therefore not needed to obtain M_i since P_i can be computed from point correspondences. The matrix $\widetilde{\Omega}_\infty$ has the structure,

$$\widetilde{\Omega}_\infty = \begin{bmatrix} \omega & \omega[\pi_\infty]_\times \\ [\pi_\infty]_\times^\top \omega & [\pi_\infty]_\times^\top \omega [\pi_\infty]_\times \end{bmatrix}, \tag{2.38}$$

where the 4-vector $\left(\pi_\infty^\top, 1\right)^\top$ represents the coordinates of $\mathit{\Pi}_\infty$. It is of rank 3 and satisfies $\widetilde{\Omega}_\infty \succeq 0$. Furthermore, considering the partitioning of $\widetilde{\Omega}_\infty$ into 3×3 matrix blocks $\widetilde{\Omega}_{11}$, $\widetilde{\Omega}_{21}$, and $\widetilde{\Omega}_{22}$ as,

$$\widetilde{\Omega}_\infty = \begin{bmatrix} \widetilde{\Omega}_{11} & \widetilde{\Omega}_{21}^\top \\ \widetilde{\Omega}_{21} & \widetilde{\Omega}_{22} \end{bmatrix}, \tag{2.39}$$

it was shown in [Ponce et al. 2005] that $\widetilde{\Omega}_\infty$ carries the additional property,

$$\mathrm{tr}(\widetilde{\Omega}_{21}) = 0. \tag{2.40}$$

Further properties and results on the ALQ are given in [Ponce et al. 2005; Valdés et al. 2006].

Should $\widetilde{\Omega}_\infty$ be estimated, both ω and π_∞ can be extracted from it, making a metric upgrade possible. Considering camera i with no skew and unit aspect ratio, ω_i exhibits a simple and convenient form, which with adequate scaling, is given by

$$\omega_i = \begin{bmatrix} 1 & 0 & -u_i \\ 0 & 1 & -v_i \\ -u_i & -v_i & \beta \end{bmatrix}, \tag{2.41}$$

where $\beta = f_i^2 + u_i^2 + v_i^2$. The following constraints enforcing unit aspect ratio and the absence of skew are then obtained,

$$\begin{aligned} (\mathsf{M}_i \widetilde{\Omega}_\infty \mathsf{M}_i{}^\top)_{11} &= (\mathsf{M}_i \widetilde{\Omega}_\infty \mathsf{M}_i{}^\top)_{22}, \\ (\mathsf{M}_i \widetilde{\Omega}_\infty \mathsf{M}_i{}^\top)_{12} &= 0, \end{aligned} \tag{2.42}$$

where the focal length and the principal point are unknown and free to vary. A linear least-squares solution for $\widetilde{\Omega}_\infty$ may be obtained from (2.42) and the trace property (2.40) using at least 10 images. The rank-3 condition can be enforced subsequently. However, the solution is not guaranteed to satisfy $\widetilde{\Omega}_\infty \succeq 0$.

2.3.4 Stratified autocalibration

Unlike direct autocalibration methods, stratified methods first recover an affine reconstruction from a projective one by locating $\mathit{\Pi}_\infty$. Once an affine reconstruction is obtained, the camera intrinsic parameters can be retrieved by solving linear equations for the IAC or the DIAC. Locating $\mathit{\Pi}_\infty$ reliably and accurately, however, has proven to be challenging in stratified autocalibration. The equations involved are highly nonlinear and difficult to solve reliably and efficiently. In this section, we describe some of the constraints that have been used to estimate $\mathit{\Pi}_\infty$ and discuss some methods in the literature that rely on these constraints. Most stratified meth-

ods assume constant camera intrinsic parameters, and we focus on such methods and the involved constraints. Towards the end of this section, we detail how the intrinsic parameters can be recovered once Π_∞ is located. The discussion in this section is particularly relevant to our contributions presented in Chapters 3 and 4, where we introduce new constraints on Π_∞ and propose stratified methods to exploit them.

Modulus constraint

When the camera intrinsic parameters are constant, the inter-image infinite homography $\mathsf{H}_{\infty ij}$ is of the form,

$$\mathsf{H}_{\infty ij} = \lambda_i^2 \lambda_j \mathsf{K} \mathsf{R}_{ij} \mathsf{K}^{-1}, \tag{2.43}$$

where λ_i and λ_j are scalars. As such, $\mathsf{H}_{\infty ij}$ is similar to a scaled rotation matrix, and its eigenvalues are,

$$\lambda_i^2 \lambda_j \{1,\ \cos\theta_{ij} + i \sin\theta_{ij},\ \cos\theta_{ij} - i \sin\theta_{ij}\}, \tag{2.44}$$

where θ_{ij} is the relative orientation angle (about an arbitrary axis in the axis-angle representation) between the two cameras. Since a rotation matrix has eigenvalues with unit modulus, those of $\mathsf{H}_{\infty ij}$ necessarily have equal moduli. In [Pollefeys and Van Gool 1999], the authors derived a necessary condition on Π_∞ for $\mathsf{H}_{\infty ij}$ to carry this so-called modulus constraint property. For any two views i and j, this condition involves the coefficients of the equation,

$$\det(\lambda \mathsf{H}_{\infty 1i} - \mathsf{H}_{\infty 1j}) = c_i(\pi_\infty)\lambda^3 - t_{ij}(\pi_\infty)\lambda^2 + t_{ji}(\pi_\infty)\lambda - c_j(\pi_\infty) = 0, \tag{2.45}$$

where c_i and t_{ij} are affine functions of Π_∞. For $\mathsf{H}_{\infty ij}$ to satisfy the modulus constraint, it was shown in [Pollefeys and Van Gool 1999] that Π_∞ must satisfy the following quartic polynomial equation,

$$m_{ij}(\pi_\infty) = c_i(\pi_\infty) t_{ji}^3(\pi_\infty) - c_j(\pi_\infty) t_{ij}^3(\pi_\infty) \ = \ 0, \quad \text{for all } \ i \neq j. \tag{2.46}$$

This polynomial can also be derived from (2.7a) and (2.7b). From (2.45), note that,

$$c_i(\pi_\infty) = \det(\mathsf{H}_{\infty 1i}), \qquad i = 1, 2, \ldots, n, \tag{2.47a}$$

$$t_{ij}(\pi_\infty) = \mathrm{tr}(\mathsf{H}_{\infty ij}), \qquad \text{for all } \; i \neq j. \tag{2.47b}$$

Furthermore, c_i and t_{ij} are related to the scaling of the infinite homography as,

$$\det(\mathsf{H}_{\infty 1i}) = \lambda_i^3, \tag{2.48a}$$

$$\mathrm{tr}(\mathsf{H}_{\infty ij}) = \lambda_i^2 \lambda_j \mathsf{a}_{ij}, \tag{2.48b}$$

where $\mathsf{a}_{ij} = 1 + 2\cos\theta_{ij}$. Equation (2.46) is then obtained from (2.47) and (2.48) by equating the scalars.

Each image pair provides a quartic polynomial equation in $\mathit{\Pi}_\infty$ from the modulus constraint. A finite number of candidate solutions for $\mathit{\Pi}_\infty$ can be obtained when three such polynomials, *i.e*. as many images, are available. In [Pollefeys and Van Gool 1999], the polynomials from three image pairs were solved using homotopy continuation. With more images, an initial solution from homotopy continuation was refined through nonlinear least-squares optimization. While there are as many as 64 possible solutions for $\mathit{\Pi}_\infty$ in the minimal case of three views, Schaffalitzky showed that these can be divided into three sets of 21 and one solution that represents the trifocal plane, where only one of the three sets is feasible [Schaffalitzky 2000]. This work also showed the connection between the modulus constraint and the *horopter* of a camera pair. We discuss the horopter curve and its link with the modulus constraint in detail in Chapter 4. More recently, a globally optimal method was presented in [Chandraker et al. 2010] that uses a Branch-and-Bound search to locate $\mathit{\Pi}_\infty$. Chirality constraints were used to confine the search space for $\mathit{\Pi}_\infty$. This method also solved the metric upgrade step using a Branch-and-Bound search, where the positive definiteness of the DIAC was enforced.

Infinite Cayley Transform

For a camera pair (i, j) with constant intrinsic parameters, the matrix defined as

$$\mathsf{Q}_{\infty ij} = \lambda_j \mathsf{H}_{\infty ij} - \lambda_i \mathsf{H}_{\infty ji}, \tag{2.49a}$$

$$= \lambda_i^2 \lambda_j^2 \, \mathsf{K}[\mathsf{r}_{ij}]_\times \mathsf{K}^{-1}, \tag{2.49b}$$

was introduced in [Habed et al. 2012] and [F. Wu et al. 2013] in the context of stratified camera autocalibration. This matrix, referred to as the *Infinite Cayley Transform* (ICT) in [F. Wu et al. 2013], is similar to the skew-symmetric matrix $[\mathsf{r}_{ij}]_\times$, defined as,

$$[\mathsf{r}_{ij}]_\times = \mathsf{R}_{ij} - \mathsf{R}_{ij}^\top. \tag{2.50}$$

The matrix $[\mathsf{r}_{ij}]_\times$ is equal up to a scale to the Cayley transform of the rotation matrix R_{ij} [F. Wu et al. 2013]. Thus, while the inter-image infinite homography $\mathsf{H}_{\infty ij}$ is similar (up to a scale) to R_{ij}, the ICT $\mathsf{Q}_{\infty ij}$ is similar (up to a scale) to the Cayley transform of R_{ij}. Hence, it has been termed as the Infinite Cayley Transform.

The ICT carries interesting properties that allow the derivation of constraints on $\mathit{\Pi}_\infty$ that are complementary to the modulus constraint [Pollefeys and Van Gool 1999]. For instance, the inequality,

$$\mathrm{tr}(\mathsf{Q}_{\infty ij}^*) > 0, \tag{2.51}$$

combined with the modulus constraint (2.46) are necessary and sufficient conditions for $\mathsf{Q}_{\infty ij}$ to be similar to a skew-symmetric matrix [F. Wu et al. 2013]. Note that, using (2.7a) and (2.48), inequality (2.51) is a polynomial in $\mathit{\Pi}_\infty$. Furthermore, in [Habed et al. 2012] the authors showed that for cameras with zero skew, *i.e.* $\gamma = 0$, the coordinates $(u,\, v)$ of the principal point can be expressed as follows,

$$u = \frac{(\mathsf{Q}_{\infty ij})_{11}}{(\mathsf{Q}_{\infty ij})_{31}}, \quad v = \frac{(\mathsf{Q}_{\infty ij})_{22}}{(\mathsf{Q}_{\infty ij})_{32}}. \tag{2.52}$$

New polynomials in $\mathit{\Pi}_\infty$, enforcing the constancy of $(u,\, v)$ across images, were derived from image triplets in [Habed et al. 2012; F. Wu et al. 2013]. These polynomials were solved along with those from the modulus constraint using homotopy

continuation, and the solution was refined through nonlinear optimization. This is the same approach as in [Pollefeys and Van Gool 1999] but with additional constraints. Further properties of the ICT are given in [F. Wu et al. 2013].

Determining the calibration: Once Π_∞ is located, the infinite homographies between images are known, and the intrinsic parameters can be estimated linearly using (2.33). In the case of constant intrinsic parameters, the IAC and the DIAC remain invariant throughout the sequence of images. Considering the DIAC, this means that ω^* is the same for all images. Moreover, in this case, the infinite homography $\mathsf{H}_{\infty ij}$, for any two views i and j, is known to exact scale as, with an appropriate normalization, it verifies $\det(\mathsf{H}_{\infty ij}) = 1$. Thus, from (2.33b) each image pair provides six linear equations in the ω^*, of which only four are linearly independent. Given three or more images, ω^* can be uniquely estimated so long as the camera motion is sufficiently general. The intrinsic parameters are obtained through its Cholesky factorization. In the approach described above, some camera motions lead to ambiguous solutions for particular intrinsic parameters. Further constraints are usually required to resolve the ambiguities. Such motions and their ambiguities and resolution are reported in [Demirdjian et al. 1998; Hartley and Zisserman 2004].

2.3.5 Motion constraints

In many applications, the camera motion is either inherently restricted or partially known. For example, a Pan-Tilt (PT) or Pan-Tilt-Zoom (PTZ) surveillance camera only rotates about its center. Such constraints on the camera motion have been exploited in camera autocalibration, and they often result in a simpler problem formulation. Methods based on restricted camera motions have also been found to be more reliable than those based on general motion [Hartley and Zisserman 2004, Sec. 19.11]. In this section, we describe some constrained camera motions and survey relevant methods in the literature. The survey includes both direct and stratified methods, the interest being in the motion constraints and their applicability. While there are numerous methods that can be included in such a survey, we select a few that may be considered representative of similar works. Where possible, we also

Constraint	Method	Direct/Stratified	Constant/Varying intrinsic parameters
Pure translation	[Armstrong et al. 1994]	Stratified	Constant
	[Pollefeys et al. 1996]	Stratified	Varying
Pure rotation	[Hartley 1997b]	Stratified	Constant
	[Agapito et al. 2001; Li and C. Shen 2006; Rameau et al. 2012]	Stratified	Varying
Planar motion	[Armstrong et al. 1996; Faugeras et al. 2000; Espuny 2007]	Stratified	Constant
Partially known motion			
Known rotation	[Frahm and Koch 2003]	Direct	Varying
Known rotation angle	[Martyushev 2018]	Direct	Constant
Small rotation	[F. Shen and Wang 2002]	Stratified	Constant
Rotation angle $\leq 90^\circ$	[Habed et al. 2014]	Direct	Constant

Table 2.1: Autocalibration methods exploiting camera motion constraints. Note that only an affine reconstruction can be recovered from a pure translation. Similarly, with a purely rotating camera, only the intrinsic parameters can be obtained and a reconstruction is not possible. Refer to the text for such details.

pick more recent methods. Table 2.1 summarizes the survey. The discussion in this section is pertinent to our contributions in Chapter 4 on exploiting partial camera motion knowledge. Towards the end of this section, we very briefly discuss motion sequences that prevent camera autocalibration.

Pure translation: One of the earlier works exploiting camera motion constraints considered a translation motion (of the camera or equivalently of an object imaged by a stationary camera). With this motion, *i.e*. no rotation, an affine reconstruction can be recovered from two views [Van Gool et al. 1994]. Upgrading to metric, however, requires further views where the camera also rotates. This approach was followed in [Armstrong et al. 1994] assuming constant intrinsic parameters. The case of a varying focal length, *i.e*. a zooming camera, was addressed in [Pollefeys et al. 1996]. The pure translation constraint is suited to robotics applications [Beardsley et al. 1994] and controlled settings, where such a motion can be easily performed.

Pure rotation: Several methods deal with the case of a rotating camera, where the camera rotates about its center but does not translate. This configuration has received considerable attention as it applies to many applications, such as surveillance and broadcast cameras and panoramic mosaics. Though a reconstruction cannot be obtained without a translation motion, the recovery of the intrinsic parameters is simplified. For a non-translating camera, the different views are related by inter-image infinite homographies that can be computed from feature correspondences across the images. Autocalibration then reduces to solving a linear problem for the DIAC/IAC. To estimate the DIAC, the intrinsic parameters were assumed to be constant in [Hartley 1994b]. They were allowed to vary in [Agapito et al. 2001], and the IAC was estimated instead using priors on some of the parameters. In [Li and C. Shen 2006], the authors incorporated the positive definiteness of the DIAC/IAC as a constraint by reformulating the problem as an LMI optimization problem. In [Rameau et al. 2012], additional constraints on the intrinsic parameters, such as restricting the principal point to lie within the image, were included in the LMI formulation of [Li and C. Shen 2006].

Planar motion: A camera undergoing planar motion translates in a plane and rotates about an axis perpendicular to the plane. This motion is typical of a camera mounted on a ground vehicle. Methods exploiting planar motion [Armstrong et al. 1996; Faugeras et al. 2000; Espuny 2007] usually assume a camera with constant intrinsic parameters and rely on certain fixed image points, the so-called circular points, under this constrained motion. The methods differ mainly in the computation of the circular points. When at least three images are available, $\mathit{\Pi}_\infty$ can be identified and the intrinsic parameters then recovered linearly. Since the rotations are all about the same axis, constraints on some of the intrinsic parameters are required to disambiguate the solutions.

Partially known motion: A partial knowledge of the camera motion is often available, for instance, from the measurements of an external sensor such as an Inertial Measurement Unit (IMU). A known rotation was exploited in [Frahm and Koch 2003]. Given the Fundamental matrices between view pairs, linear equations in the unknown and possibly varying intrinsic parameters were derived. The additional

orientation information reduces the number of images/constraints on the intrinsic parameters required for autocalibration. The particular case of a non-translating camera with varying intrinsic parameters and known orientation was also addressed in this work. The entry in Table 2.1 for this method corresponds to the general motion case. In [Martyushev 2018], the relative orientation angle (about an arbitrary axis) was assumed to be known. For two views with square pixels (*i.e*. zero skew and unit aspect ratio) and otherwise unknown but constant intrinsic parameters, a new quartic polynomial constraint was formulated on the Essential matrix based on the known orientation angle. A minimal solver using Gröbner bases was proposed to estimate the unknown focal length and principal point. The use case, in particular, is where a camera and an IMU are rigidly attached. In this case, an extrinsic calibration between the two sensors is not required, and the relative orientation angle from the IMU can be used directly. In contrast to the exact motion information described above, some methods exploit a more vague motion knowledge. In [F. Shen and Wang 2002], the authors assumed a small rotation between consecutive views, *i.e*. the rotation angle about each axis being within 10°. This assumption is usually implicitly satisfied in a video. By approximating the camera rotation and some of the intrinsic parameters, linear equations were obtained in the coordinates of Π_∞. In [Habed et al. 2014], new LMI constraints on the DAQ were formulated assuming constant intrinsic parameters and the relative orientation angle between consecutive views being within 90°. This assumption applies more broadly to image sequences captured for 3D reconstruction.

Critical motion sequences: Some camera motions lead to ambiguities in recovering the intrinsic parameters or a metric reconstruction. An example is a camera undergoing a pure translation (discussed above), where only an affine reconstruction is possible. In other cases, there is no unique solution for some of the intrinsic parameters. Such degeneracies are inherent to the motion and require additional information (*e.g*. more constraints) for its resolution. The existence of motions resulting in these degeneracies, referred to as *Critical Motion Sequences* (CMS), is one of the main pitfalls of camera autocalibration. Thus, knowledge of these motions is essential to be able to avoid them when using autocalibration methods,

especially those based on restricted camera motions. A detailed account of CMS is given in [Sturm 1997a; Sturm 1997b; Kahl et al. 2000; Sturm 2002].

2.4 Conclusion

In this chapter, we reviewed some key results in computer vision, focusing on recovering a 3D reconstruction of a scene from multiple images. We also described the camera autocalibration problem and surveyed the relevant literature. In the following two chapters, we present the contributions of this thesis. We rely on some of the results summarized in this chapter, and the literature review lends context to our contributions.

3 Exploiting camera geometry

This chapter presents our contributions in exploiting a partial knowledge of the camera geometry in camera autocalibration. We consider the case of a camera with a Euclidean image plane (EIP), *i.e*. zero skew and unit aspect ratio. Most cameras in the market today fit the EIP model. While this assumption has often been exploited in direct autocalibration methods, stratified methods have yet to fully benefit from it. Assuming a moving camera with EIP and constant intrinsic parameters, we show that a new quartic polynomial equation in the unknown plane at infinity is obtained for each image pair in addition to the polynomial from the modulus constraint. We propose a stratified autocalibration method that relies on these two polynomials and solves a constrained polynomial optimization problem to locate the plane at infinity. We report the experiments that we conducted using both synthetic data and real images to evaluate our proposed algorithm and compare it with existing methods.

3.1 Introduction

EIP cameras: Retrieving the camera calibration parameters from feature correspondences across images, *i.e*. camera autocalibration, is a prerequisite to recover the metric structure of an unknown scene imaged by uncalibrated perspective cameras. Autocalibration methods rely on some assumptions on the calibration parameters, such as constant [Heyden and Åström 1996; Luong and Faugeras 1997; Triggs 1997; Pollefeys and Van Gool 1999] or partially known intrinsic parameters [Heyden and Åström 1999; Pollefeys et al. 1999; Ponce et al. 2005; Valdés et al. 2006; Gherardi and Fusiello 2010]. When the images are captured by the same moving camera, its internal geometry remains unchanged in the absence of zooming and focusing. The sensor's aspect ratio and skew factor also remain quite stable despite a change in the focus or zoom. Moreover, modern cameras commonly have square pixels, *i.e*. zero skew and unit aspect ratio. Such cameras are said to have a *Euclidean Image Plane* (EIP) [Heyden and Åström 1997].

EIP constraint in direct autocalibration: The EIP assumption has often been exploited in direct autocalibration methods, which simultaneously estimate the plane at infinity ($\mathit{\Pi}_\infty$) and the intrinsic parameters. Direct methods rely mainly on either the *Dual Absolute Quadric* (DAQ) [Heyden and Åström 1996; Triggs 1997] or the *Absolute Line Quadric* (ALQ) [Ponce et al. 2005; Valdés et al. 2006] formulations (see Section 2.3.3) that encode both $\mathit{\Pi}_\infty$ and the intrinsic parameters. These formulations are not equally suited to exploit the EIP constraint. The DAQ-based autocalibration equations, for instance, involve the *Dual Image of the Absolute Conic* (DIAC). Due to the coupling of the intrinsic parameters in the DIAC, exploiting the EIP assumption requires the principal point to be known. The problem is then linear. To this end, an image-centered principal point is assumed in [Pollefeys et al. 1999]. However, the location of the principal point can be far from the image center, and it also varies with zooming or focusing [Sturm 1996]. The ALQ-based autocalibration equations instead involve the *Image of the Absolute Conic* (IAC). These are simpler to use under the EIP constraint as linear equations are obtained without assuming the principal point to be known. However, a practical difficulty in using either the DAQ or the ALQ is to enforce the nonlinear rank-3 con-

straint in their estimation. Linearization, on the other hand, may result in artificial degeneracies [Gurdjos et al. 2009].

EIP constraint in stratified autocalibration: Unlike direct methods, a stratified approach first tackles the more challenging problem of estimating Π_∞. Once Π_∞ is located, the intrinsic parameters can be retrieved by solving linear equations for the IAC or the DIAC [Hartley and Zisserman 2004]. The advantages of this approach over a direct quadric-based one are that the nonlinearity is confined to fewer unknowns and a rank condition is not required. Most stratified methods assume constant intrinsic parameters [Pollefeys and Van Gool 1999; Chandraker et al. 2010; Habed et al. 2012; F. Wu et al. 2013; Adlakha et al. 2019] and rely on the polynomial derived from the modulus constraint [Pollefeys and Van Gool 1999]. Some methods use properties of the so-called *Infinite Cayley Transform* (ICT) [Habed et al. 2012; F. Wu et al. 2013] to further exploit a partial knowledge of the intrinsic parameters (see Section 2.3.4 for more details on the ICT). For instance, the zero-skew assumption was exploited in [Habed et al. 2012; F. Wu et al. 2013] to derive quartic polynomials in Π_∞ for image triplets. However, using triplets introduces several unknown projective scale factors that may lead to numerically ill-conditioned polynomial systems. To the best of our knowledge, the assumption of a camera with EIP and constant intrinsic parameters has not been exploited so far in stratified camera autocalibration.

Global optimization: The camera autocalibration problem is inherently nonlinear, and methods have traditionally relied on local optimization to obtain the calibration parameters. While local optimization methods are typically fast and broadly applicable, they only provide a locally optimal solution with no guarantee of global optimality. The solution then depends crucially on the initialization used in the optimization. Autocalibration methods often solve a simpler, usually linear or convex, problem and then refine the estimated solution using local optimization of a preferred nonlinear cost. Such an approach is followed in [Heyden and Åström 1999; Pollefeys et al. 1999; Nistér 2004b; Adlakha et al. 2019]. Alternatively, some methods [Pollefeys and Van Gool 1999; Habed et al. 2012; F. Wu et al. 2013; Martyushev 2018] solve a system of polynomial equations using homotopy continuation [Ver-

schelde 1999] or Gröbner bases [Cox et al. 2005]. These techniques compute all the solutions of a square polynomial system. However, they do not scale well to overdetermined systems, and it is difficult to incorporate additional inequality constraints. More recent works in camera autocalibration have also investigated globally optimal optimization methods, either based on Branch-and-Bound algorithms [Fusiello et al. 2004; Bocquillon et al. 2007; Chandraker et al. 2010; Habed et al. 2014; Paudel and Van Gool 2018] or on polynomial optimization [Chandraker et al. 2007] using Lasserre's hierarchy of semidefinite relaxations [Lasserre 2001; Henrion et al. 2009]. Lasserre's method deals with nonlinear and nonconvex optimization problems where the cost function and constraints are multivariate scalar polynomials. It does so by constructing a hierarchy of convex Linear Matrix Inequality (LMI) relaxations of increasing order that are underestimators of the original problem. This hierarchy of convex relaxations is provably convergent to the global optimum at a finite order of relaxation. In [Chandraker et al. 2007], Lasserre's hierarchy has been used to estimate the DAQ under rank, semidefiniteness, and chirality [Hartley and Zisserman 2004] constraints.

Contributions: In this chapter, we present a stratified autocalibration method for a moving camera with EIP and constant intrinsic parameters. Our key contribution is in the formulation of a new quartic polynomial in the unknown Π_∞ under the assumed camera model. This polynomial is obtained for each image pair in addition to the polynomial from the modulus constraint. It is derived using a yet unexploited property of the ICT. For three or more images, estimating Π_∞ is stated as a constrained polynomial optimization problem that is solved using Lasserre's hierarchy. Polynomial inequality constraints on Π_∞, such as those derived from the ICT as well as the chirality inequalities, can be easily incorporated in Lasserre's method. We also derive a new pair of inequalities in Π_∞ that restrict the principal point to lie within the image bounds. The estimated Π_∞ is refined using local optimization of a normalized cost, and the intrinsic parameters are subsequently retrieved by estimating the DIAC. Experiments with synthetic data and real images show that the new polynomial in our proposed algorithm leads to a more reliable performance than existing methods, especially for short sequences.

Organization: The rest of this chapter is organized as follows. In Section 3.2, we derive the new EIP-based quartic polynomial constraint on $\mathit{\Pi}_\infty$. In Section 3.3, we describe some existing polynomial inequalities in $\mathit{\Pi}_\infty$ and derive new inequalities that constrain the location of the principal point. We detail our stratified autocalibration method in Section 3.4 and report the experiments that we conducted to evaluate it in Section 3.5. Finally, Section 3.6 concludes this chapter.

3.2 EIP-based polynomial constraint

We consider a camera with a Euclidean image plane, EIP (*i.e.* zero skew, $\gamma = 0$, and unit aspect ratio, $f_x/f_y = 1$) whose focal length and principal point coordinates are constant but unknown. We show that the ICT carries a yet unexploited property under these assumptions (Proposition 3.1), and we use it to derive a new quartic polynomial constraint on $\mathit{\Pi}_\infty$ (Proposition 3.2). The results presented in this section also apply to a non-EIP camera with fixed intrinsic parameters should the skew and aspect ratio be known. The projection matrix of such a camera can be transformed to the EIP model by an appropriate change of coordinates [Heyden and Åström 1997]. We now define a matrix operator that we use to derive the EIP-based constraint.

Definition 3.1. *Given a* 3×3 *matrix* B*, we define the matrix operator* Φ *as,*

$$\Phi(\mathsf{B}) = (\mathsf{B}^* \circ \mathsf{B})_{31} + (\mathsf{B}^* \circ \mathsf{B})_{32}, \tag{3.1}$$

where $\circ$ *denotes the Hadamard (elementwise) product,* i.e. $(\mathsf{B} \circ \mathsf{C})_{hk} = (\mathsf{B})_{hk}(\mathsf{C})_{hk}$ *for any two matrices* B *and* C *of the same dimensions.*

The following proposition reveals a yet unexploited property of the ICT under the EIP assumption.

Proposition 3.1. *Consider an image pair* (i, j) *captured by a moving camera with EIP and constant intrinsic parameters. The ICT* $\mathsf{Q}_{\infty ij}$ *of these images satisfies,*

$$\Phi(\mathsf{Q}_{\infty ij}) = 0. \tag{3.2}$$

Proof. From (2.49b), the ICT $\mathsf{Q}_{\infty ij}$ is given as $\mathsf{Q}_{\infty ij} \simeq \mathsf{K}[\mathsf{r}_{ij}]_\times \mathsf{K}^{-1}$. As such, the matrix $\mathsf{K}^{-1}\mathsf{Q}_{\infty ij}\mathsf{K}$ is skew-symmetric. Given a camera with EIP, the 2×2 matrix,

$$\begin{bmatrix} (\mathsf{Q}_{\infty ij})_{11} - u(\mathsf{Q}_{\infty ij})_{31} & (\mathsf{Q}_{\infty ij})_{12} - u(\mathsf{Q}_{\infty ij})_{32} \\ (\mathsf{Q}_{\infty ij})_{21} - v(\mathsf{Q}_{\infty ij})_{31} & (\mathsf{Q}_{\infty ij})_{22} - v(\mathsf{Q}_{\infty ij})_{32} \end{bmatrix}, \tag{3.3}$$

obtained by eliminating the third row and column of $\mathsf{K}^{-1}\mathsf{Q}_{\infty ij}\,\mathsf{K}$ is also skew-symmetric. Enforcing the diagonal entries of (3.3) to be zero leads to the expressions of u and v given in (2.52), as obtained in [Habed et al. 2012] for cameras with zero skew. Furthermore, the sum of the off-diagonal elements of (3.3) also being zero yields:

$$(\mathsf{Q}_{\infty ij})_{12} + (\mathsf{Q}_{\infty ij})_{21} - u(\mathsf{Q}_{\infty ij})_{32} - v(\mathsf{Q}_{\infty ij})_{31} = 0. \tag{3.4}$$

Substituting the expressions of u and v from (2.52) in (3.4) leads to (3.2). ■

Proposition 3.1 allows us to formulate a new constraint on $\mathit{\Pi}_\infty$. We consider hereafter that the coordinates of $\mathit{\Pi}_\infty$ are $\Pi_\infty = \left(\pi_\infty^\top, 1\right)^\top$. Using (2.49a) and (2.7a), we observe that $\Phi(\mathsf{Q}_{\infty ij})$ expands as:

$$\Phi(\mathsf{Q}_{\infty ij}) = a_{ij}(\pi_\infty)\lambda_j^3 - b_{ij}(\pi_\infty)\lambda_i\lambda_j^2 + b_{ji}(\pi_\infty)\lambda_i^2\lambda_j - a_{ji}(\pi_\infty)\lambda_i^3, \tag{3.5}$$

where the coefficients a_{ij} and b_{ij}, for any combination of i and j, are cubic polynomials in $\mathit{\Pi}_\infty$. Substituting for λ_i and λ_j using (2.48) leads to a polynomial of degree six at best. Polynomials of such high degree are challenging to solve using current techniques, be it homotopy continuation, Gröbner bases, or Lasserre's hierarchy, as they often lead to numerically ill-conditioned or even intractable systems of equations. We show, however, that equation (3.2) can be used to derive a quartic polynomial in $\mathit{\Pi}_\infty$, *i.e.* a polynomial of the same degree as the modulus constraint (2.46). Although a_{ij} and b_{ij} are fully defined through the expansion in (3.5), it is interesting to note that:

$$a_{ij}(\pi_\infty) = \Phi(\mathsf{H}_{\infty ij}) \quad \text{and} \quad a_{ji}(\pi_\infty) = \Phi(\mathsf{H}_{\infty ji}). \tag{3.6}$$

As such, we show that the terms involving $a_{ij}(\pi_\infty)$ and $a_{ji}(\pi_\infty)$ can be eliminated from (3.5) so long as the modulus constraint is satisfied. To show this, we recall here some properties of the Hadamard product and adjugate matrices.

Property 3.1. *Let* B *and* C *be two* 3×3 *matrices and* λ *a scalar. We have that:*

$$\mathsf{B} \circ \mathsf{C} = \mathsf{C} \circ \mathsf{B}, \tag{3.7a}$$

$$(\lambda\mathsf{B}) \circ \mathsf{C} = \mathsf{B} \circ (\lambda\mathsf{C}) = \lambda(\mathsf{B} \circ \mathsf{C}), \tag{3.7b}$$

$$(\mathsf{BC})^* = \mathsf{C}^*\mathsf{B}^*, \tag{3.7c}$$

$$(\mathsf{B}^*)^* = \det(\mathsf{B})\mathsf{B}. \tag{3.7d}$$

Using properties of adjugate matrices, we can deduce that,

$$\mathsf{H}^*_{\infty ij} = \lambda_i^3 \, \mathsf{H}_{\infty ji}. \tag{3.8}$$

To see this, from property (3.7c), we have that $\mathsf{H}^*_{\infty ij} = (\mathsf{H}_{\infty 1j}\mathsf{H}_{\infty i1})^* = \mathsf{H}^*_{\infty i1}\mathsf{H}^*_{\infty 1j}$. In addition, using (2.6), $\mathsf{H}^*_{\infty ij} = (\mathsf{H}^*_{\infty 1i})^*\mathsf{H}_{\infty j1}$. Now, property (3.7d) leads to (3.8). Similarly, we obtain, $\mathsf{H}^*_{\infty ji} = \lambda_j^3 \, \mathsf{H}_{\infty ij}$. We can deduce, using (3.8) and property (3.7b), that $\mathsf{H}^*_{\infty ij} \circ \mathsf{H}_{\infty ij} = \lambda_i^3(\mathsf{H}_{\infty ji} \circ \mathsf{H}_{\infty ij})$ and also that $\mathsf{H}^*_{\infty ji} \circ \mathsf{H}_{\infty ji} = \lambda_j^3(\mathsf{H}_{\infty ij} \circ \mathsf{H}_{\infty ji})$. It must be clear now, with property (3.7a), that $\lambda_j^3(\mathsf{H}^*_{\infty ij} \circ \mathsf{H}_{\infty ij}) = \lambda_i^3(\mathsf{H}^*_{\infty ji} \circ \mathsf{H}_{\infty ji})$. As a consequence, we have that,

$$\lambda_j^3 \, a_{ij}(\pi_\infty) = \lambda_i^3 \, a_{ji}(\pi_\infty). \tag{3.9}$$

This constitutes the proof to our main result stated in the following proposition.

Proposition 3.2. *Consider an image pair* (i, j) *captured by a moving camera with EIP and constant intrinsic parameters. The plane at infinity* $\mathit{\Pi}_\infty$ *satisfies the quartic polynomial equation:*

$$p_{ij}(\pi_\infty) = -b_{ij}(\pi_\infty)t_{ji}(\pi_\infty) + b_{ji}(\pi_\infty)t_{ij}(\pi_\infty) = 0, \tag{3.10}$$

for all $i \neq j$*. The expressions* b_{ij} *and* b_{ji} *are cubic polynomials in* $\mathit{\Pi}_\infty$ *defined by the expansion in* (3.5) *while* t_{ij} *and* t_{ji} *are linear functions of* $\mathit{\Pi}_\infty$ *defined by* (2.45)*.*

We refer to the polynomial p_{ij} as the EIP polynomial. It is obtained from (3.2) by substituting for λ_i and λ_j in (3.5) using the expressions in (2.48) and eliminating the a_{ij} and a_{ji} terms.

3.3 Inequality constraints on infinity

With the assumption of a camera with EIP and constant intrinsic parameters, each image pair provides two quartic polynomial equations in $\mathit{\Pi}_\infty$, *i.e.* the modulus constraint and the EIP polynomial. In Section 3.4, we present a stratified autocalibration method, where we estimate $\mathit{\Pi}_\infty$ by minimizing a cost that is constructed using these two polynomials. In such an approach, multiple solutions for $\mathit{\Pi}_\infty$ may persist, especially when few images are used. It is then useful to incorporate additional inequality constraints on $\mathit{\Pi}_\infty$ to steer the optimization towards the sought solution. In this section, we describe some existing inequality constraints on $\mathit{\Pi}_\infty$ and derive a new pair of inequalities that restrict the location of the principal point.

3.3.1 Chirality inequalities

Hartley's chirality inequalities [Hartley and Zisserman 2004] can be used to preserve the convex hull of the camera centers. They impose that all $c_i(\pi_\infty) = \det(\mathsf{H}_{\infty 1i})$, $i = 1, 2, \ldots, n$, defined in (2.47), carry the same sign, provided that all the projection matrices are sign-corrected (see Section 2.2.6). Although the same can be done for scene points, these are generally considered to be less reliable [Nistér 2004b].

3.3.2 ICT-based inequality

The inequality given in (2.51) is a necessary condition for the ICT $\mathsf{Q}_{\infty ij}$ to be similar to a skew-symmetric matrix. This inequality can be expressed in terms of $\mathit{\Pi}_\infty$. There are different ways to do this. For instance, using the expressions in (2.48), we can consider a scaled ICT matrix, $\widetilde{\mathsf{Q}}_{\infty ij} \simeq \mathsf{Q}_{\infty ij}$, given by,

$$\widetilde{\mathsf{Q}}_{\infty ij} = \operatorname{tr}(\mathsf{H}_{\infty ji})\mathsf{H}_{\infty ij} - \operatorname{tr}(\mathsf{H}_{\infty ij})\mathsf{H}_{\infty ji}, \tag{3.11}$$

and thereby obtain the constraint,

$$q_{ij}(\pi_\infty) \;=\; \mathrm{tr}(\widetilde{\mathsf{Q}}^{*}_{\infty ij}) > 0. \tag{3.12}$$

From (2.7a), $\mathsf{H}_{\infty ij}$ is linear in $\mathit{\Pi}_\infty$. Therefore, each entry of $\widetilde{\mathsf{Q}}_{\infty ij}$ is a quadratic polynomial in $\mathit{\Pi}_\infty$ and inequality (3.12) is a quartic polynomial constraint on $\mathit{\Pi}_\infty$. Note that the sign of the unknown scale in (3.11) does not affect the sign of $q_{ij}(\pi_\infty)$ since it is squared in the adjugate matrix.

3.3.3 Principal point bounds

We can exploit the expressions in (2.52) to restrict the principal point to lie within the image bounds. Assuming a $2\overline{u} \times 2\overline{v}$ image and an image-centered frame, the principal point $(u,\, v)$ is within the image bounds if the following inequalities are satisfied:

$$\begin{aligned} u_{ij}(\pi_\infty) \;&=\; \overline{u}^2(\mathsf{Q}_{\infty ij})^2_{31} - (\mathsf{Q}_{\infty ij})^2_{11} \geq 0, \\ v_{ij}(\pi_\infty) \;&=\; \overline{v}^2(\mathsf{Q}_{\infty ij})^2_{32} - (\mathsf{Q}_{\infty ij})^2_{22} \geq 0. \end{aligned} \tag{3.13}$$

These inequalities have not been used so far in stratified camera autocalibration. Using the ICT expression in (3.11), they are quartic polynomial constraints on $\mathit{\Pi}_\infty$. Note that, in this way, the location of the principal point can also be restricted more stringently to a rectangular region around the image center.

3.4 EIP-based stratified autocalibration

Our stratified camera autocalibration method relies on polynomial optimization using Lasserre's hierarchy to estimate $\mathit{\Pi}_\infty$. Lasserre's method (see Appendix A.3 for more details) allows us to obtain the global optimum of a least-squares cost and thereby gracefully handle an overdetermined polynomial system. It also allows us to easily incorporate additional polynomial (in)equality constraints, such as those described in Section 3.3. We use the modulus constraint (2.46) and the EIP polynomial (3.10) to define a suitable cost function. Since these two polynomials are obtained for each image pair, our method can be used with three or more images. The estimated $\mathit{\Pi}_\infty$ is refined using local optimization to fit a normalized cost. The

camera intrisic parameters are subsequently recovered by solving a system of linear equations. The steps of our algorithm are summarized in Section 3.4.2. We first detail the polynomial and local optimization problems to locate $\mathit{\Pi}_\infty$ in Section 3.4.1.

3.4.1 Locating infinity

A normalized cost function is preferred in uncalibrated computer vision problems to eliminate the effect of projective scale factors. This usually leads to a cost in the form of a sum of rational functions. The resulting optimization problem is difficult to solve globally and optimally [Bugarin et al. 2015] and is not handled well by Lasserre's method. We deal with this limitation in two steps. First, to reduce the effect of scaling with an unnormalized cost, we propose to use homogenized polynomials [Cox et al. 2005]. This allows us to include an additional constraint to impose some global scaling. Though this problem formulation is not equivalent to the normalized case, it works well in practice. Second, we refine the estimated $\mathit{\Pi}_\infty$ using local optimization of a normalized cost function. We describe these two steps in the following.

Polynomial optimization: We use homogeneous polynomials in this step. The homogeneous counterpart of a polynomial p of degree d in π, denoted ${}^h p$, is defined by introducing an additional variable π_4 such that ${}^h p(\pi, \pi_4) = \pi_4^d p(\pi/\pi_4)$. Note that the degree of ${}^h p$ remains the same as that of p. We then solve the following polynomial optimization problem to estimate $\mathit{\Pi}_\infty$,

$$\min_{\pi,\, \pi_4} \quad \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} {}^h m_{ij}^2(\pi, \pi_4) \; + \; {}^h p_{ij}^2(\pi, \pi_4) \tag{3.14a}$$

$$\text{s.t.} \quad {}^h c_i(\pi, \pi_4) > 0, \qquad i = 1, \ldots, n, \tag{3.14b}$$

$${}^h q_{ij}(\pi, \pi_4) > 0, \qquad i = 1, \ldots, n-1, \tag{3.14c}$$

$${}^h u_{ij}(\pi, \pi_4) \geq 0, \quad {}^h v_{ij}(\pi, \pi_4) \geq 0, \qquad j = i+1, \ldots, n, \tag{3.14d}$$

$${}^h c_1(\pi, \pi_4)\, {}^h c_n(\pi, \pi_4) \; + \; \frac{1}{n-1} \sum_{i=1}^{n-1} {}^h c_i(\pi, \pi_4)\, {}^h c_{i+1}(\pi, \pi_4) = 1, \tag{3.14e}$$

where we use the homogenized polynomials of,

m_{ij}: the polynomial derived from the modulus constraint (2.46),

p_{ij}: the EIP polynomial (3.10),

c_i: the polynomial in (2.47a) that is used in the chirality inequalities,

q_{ij}: the polynomial from the ICT-based inequality (3.12),

u_{ij}, v_{ij}: the polynomials from the principal point-based inequalities (3.13).

Equation (3.14e) is the global scaling constraint that we have found the most appropriate. We have observed that this constraint improves the numerical stability and leads more often to a certified optimal solution with the minimal relaxation order of four. We suggest using the pairwise constraints (3.14c)-(3.14d) only between consecutive views so that their number grows linearly and not quadratically with the number of views. For short sequences, constraints between all image pairs can be included. With long sequences, all but (3.14e) are optional and can be dropped.

Nonlinear refinement: We refine the $\mathit{\Pi}_\infty$ estimated from problem (3.14) using local optimization of the normalized cost,

$$\min_{\pi} \quad \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{m_{ij}^2(\pi) + p_{ij}^2(\pi)}{\left(c_i(\pi)\, c_j(\pi)\right)^4}. \tag{3.15}$$

Though the constraint in (3.14e) scales the polynomials suitably, it is recommended, particularly with high levels of noise, to refine the solution to fit a normalized cost. Note that even if the EIP assumption is not fully satisfied, problems (3.14) and (3.15) are still suitable to obtain, in practice, a satisfactory solution for $\mathit{\Pi}_\infty$. In such a case, this solution can be further refined, for instance, without using the EIP constraint.

3.4.2 Autocalibration algorithm

We consider a projective reconstruction with sign-corrected projection matrices P_i, $i = 1, \ldots, n$, and scene points X_j, $j = 1, \ldots, m$. The reference frame is attached to

the first camera, and the projection matrices are normalized such that their associated images are centered at the origin. Our autocalibration algorithm then proceeds as follows:

(i) estimate $\mathit{\Pi}_\infty$ by solving problem (3.14) using Lasserre's hierarchy,

(ii) refine the estimated $\mathit{\Pi}_\infty$ using (3.15) and upgrade the projective reconstruction to an affine one as $\mathsf{P}_i^{\mathsf{A}} = \mathsf{P}_i \mathsf{H}_{\mathsf{A}}^{-1}$ and $\mathsf{X}_j^{\mathsf{A}} = \mathsf{H}_{\mathsf{A}} \mathsf{X}_j$, where

$$\mathsf{H}_{\mathsf{A}} = \begin{bmatrix} \mathsf{I}_3 & \mathbf{0}_3 \\ \Pi_\infty^\top & \end{bmatrix},$$

(iii) compute the camera intrinsic parameters matrix K by solving linear equations for the DIAC and upgrade the affine reconstruction to a metric one as $\mathsf{P}_i^{\mathsf{M}} = \mathsf{P}_i^{\mathsf{A}} \mathsf{H}_{\mathsf{M}}^{-1}$ and $\mathsf{X}_j^{\mathsf{M}} = \mathsf{H}_{\mathsf{M}} \mathsf{X}_j^{\mathsf{A}}$, where

$$\mathsf{H}_{\mathsf{M}} = \begin{bmatrix} \mathsf{K}^{-1} & \mathbf{0}_3 \\ \mathbf{0_3}^\top & 1 \end{bmatrix}.$$

In the following section, we report the experiments that we conducted to evaluate our camera autocalibration algorithm.

3.5 Experimental results

We tested our camera autocalibration method using both synthetic data and real images. We computed the 3D Root-Mean-Square (RMS) error and the following calibration error metrics to assess our results:

$$\Delta f = \sqrt{\frac{(f_x - \widehat{f}_x)^2 + (f_y - \widehat{f}_y)^2}{{f_x}^2 + {f_y}^2}}, \quad \Delta uv = \sqrt{\frac{(u - \widehat{u})^2 + (v - \widehat{v})^2}{u^2 + v^2}}, \quad \Delta\gamma = |\gamma - \widehat{\gamma}|, \tag{3.16}$$

where $(\widehat{f}_x, \widehat{f}_y)$, $(\widehat{u}, \widehat{v})$, and $\widehat{\gamma}$ are the estimated focal lengths, principal point, and skew, respectively. The 3D RMS error was computed after aligning the estimated metric point cloud to the ground truth Euclidean point cloud by a best-fit similarity

transformation in the least-squares sense. Our algorithm was implemented in MATLAB R2018b. We used GloptiPoly [Henrion et al. 2009] to solve problem (3.14) and set a relaxation order of $d = 4$ in all the experiments. We used MOSEK[1] as the SDP solver and set `MSK_DPAR_INTPNT_CO_TOL_{P|D}FEAS` $= 10^{-20}$. We used the Levenberg-Marquardt (LM) algorithm implemented in the MATLAB Optimization Toolbox to solve problem (3.15). All the experiments were conducted on an i7 3.10 GHz 32 GB RAM computer.

We denote our algorithm (see Section 3.4.2) by EIP*. We denote the same algorithm without the EIP polynomial, which then relies only on the modulus constraint in (3.14a) and (3.15), by MODULUS*. In addition, the two approaches EIP* and MODULUS* excluding the inequality constraints in problem (3.14) are denoted by EIP and MODULUS, respectively. In these experiments, we used the inequalities in problem (3.14) only between consecutive views. Furthermore, we estimated all five intrinsic parameters in step (iii) of our algorithm. This ensured a fair comparison with the selected existing methods, all of which solve for five intrinsic parameters.

3.5.1 Synthetic data experiments

Each synthetic scene consisted of 200 points sampled randomly from the surface of the unit sphere. The cameras were positioned at a distance of 3.5–4 units from the sphere center, and oriented such that their optical axes passed close to the sphere center. All cameras were simulated to have an EIP with focal length $f_x = f_y = 800$, and an image-centered principal point, $(u,\, v) = (256,\, 256)$, in pixels. Noise, modeled as a zero-mean Gaussian distribution with standard deviation in the range $[0,\, 2]$ pixels, was added to the pixel coordinates in increments of 0.5 pixels. Projective reconstructions were obtained using the factorization approach in [Oliensis and Hartley 2007] implemented in the VSfM toolbox[2], followed by a projective bundle adjustment [Lourakis and Argyros 2009]. We report the statistics collected over 100 generated scenes.

[1] `https://docs.mosek.com/8.1/toolbox/index.html`
[2] `http://github.com/vrabaud/sfm_toolbox`

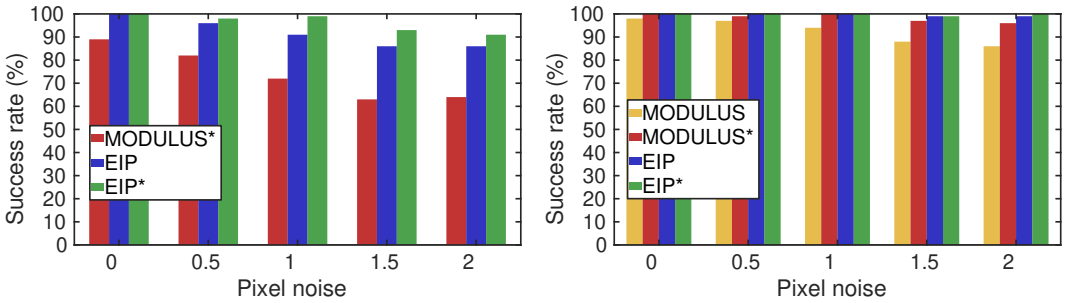


Figure 3.1: Benefits of the EIP polynomial. Success rate of EIP and EIP* compared with MODULUS and MODULUS* using 3 views (left) and 4 views (right) with varying pixel noise levels.

Benefits of the EIP polynomial: We assessed the contribution of the EIP polynomial (see Proposition 3.2) in our algorithm's performance by comparing the reliability of EIP and EIP* in obtaining a metric reconstruction with that of MODULUS and MODULUS*. We focused on short sequences and considered a 3D error above 0.25 as a failed metric upgrade. Figure 3.1 shows the success rate using 3 and 4 views for varying noise levels. With 3 views, MODULUS failed most of the time (result not shown) as multiple solutions exist using the modulus constraint alone. A higher success rate was obtained using MODULUS* due to the inequality constraints, but it declined considerably with an increasing amount of noise. On the other hand, EIP led to a reliable metric upgrade even with high levels of noise, and the inequalities in EIP* further improved the success rate. With 4 views, there are sufficient polynomials from the modulus constraint to obtain a unique solution. Even so, the success rate of EIP was significantly higher than that of MODULUS in the presence of noise. With additional views, all the approaches performed reliably. For the successful trials, the estimated plane at infinity and, consequently, the 3D errors were similar using all the approaches. These results show that the EIP polynomial is especially useful for short sequences.

Effect of refinement: We analyzed the impact of the refinement step, *i.e.* step (ii), in our algorithm. Figure 3.2 shows the 3D error distribution using EIP and EIP* with and without refinement for 4 views. The box plots shown here and throughout this section follow the convention in MATLAB. We clipped errors above the 0.25 threshold to the axis limit in Figure 3.2. The errors decreased overall after

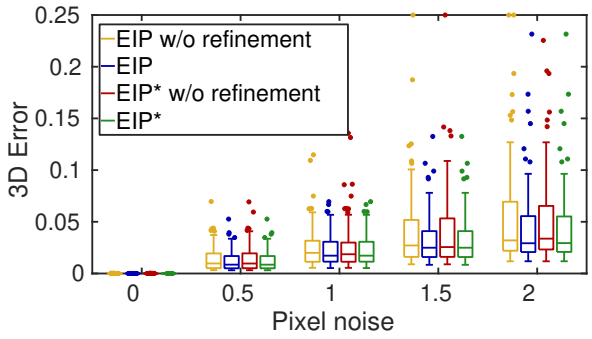


Figure 3.2: Effect of refinement. Distribution of 3D errors with and without refinement using 4 views and varying pixel noise levels.

refinement, particularly for high levels of noise. Moreover, a few reconstructions that failed in the metric upgrade without refinement were recovered after refinement. The observations were similar when varying the number of views. Thus, refinement using a normalized cost improves the accuracy of our algorithm.

Comparisons with the state of the art: We compared EIP and EIP* with two stratified methods, GO-Stratified [Chandraker et al. 2010] and QUARCH*M [Adlakha et al. 2019], and a DAQ-based method, GO-DAQ [Chandraker et al. 2007]. GO-Stratified and GO-DAQ are briefly discussed in Section 2.3. QUARCH*M will be presented in Chapter 4. For GO-Stratified, we computed solutions for both signs of chirality and retained the one with lower calibration error. We used the authors' implementation[3] of this method. For GO-DAQ, we set a relaxation order of $d = 2$. The rotation angle assumption of QUARCH*M is satisfied in our simulations. Figure 3.3 shows the success rate (left column) of the tested methods. With three views, the success rate of GO-DAQ was lower than that of EIP* and it dropped considerably with increasing noise. EIP also generally outperformed GO-DAQ in terms of success rate. Although GO-DAQ uses additional priors on the location of the principal point, the results are inferior because our simulated cameras are close to an artificial degenerate configuration for the DAQ estimation (all the optical axes intersect in one point) [Gurdjos et al. 2009]. GO-DAQ then fails when its rank-3 constraint is not well enforced due to numerical scaling issues, as has also been reported in [Adlakha et al. 2019]. With four views, EIP and EIP* outperformed

[3] `https://cseweb.ucsd.edu/~mkchandraker/stratum.html`

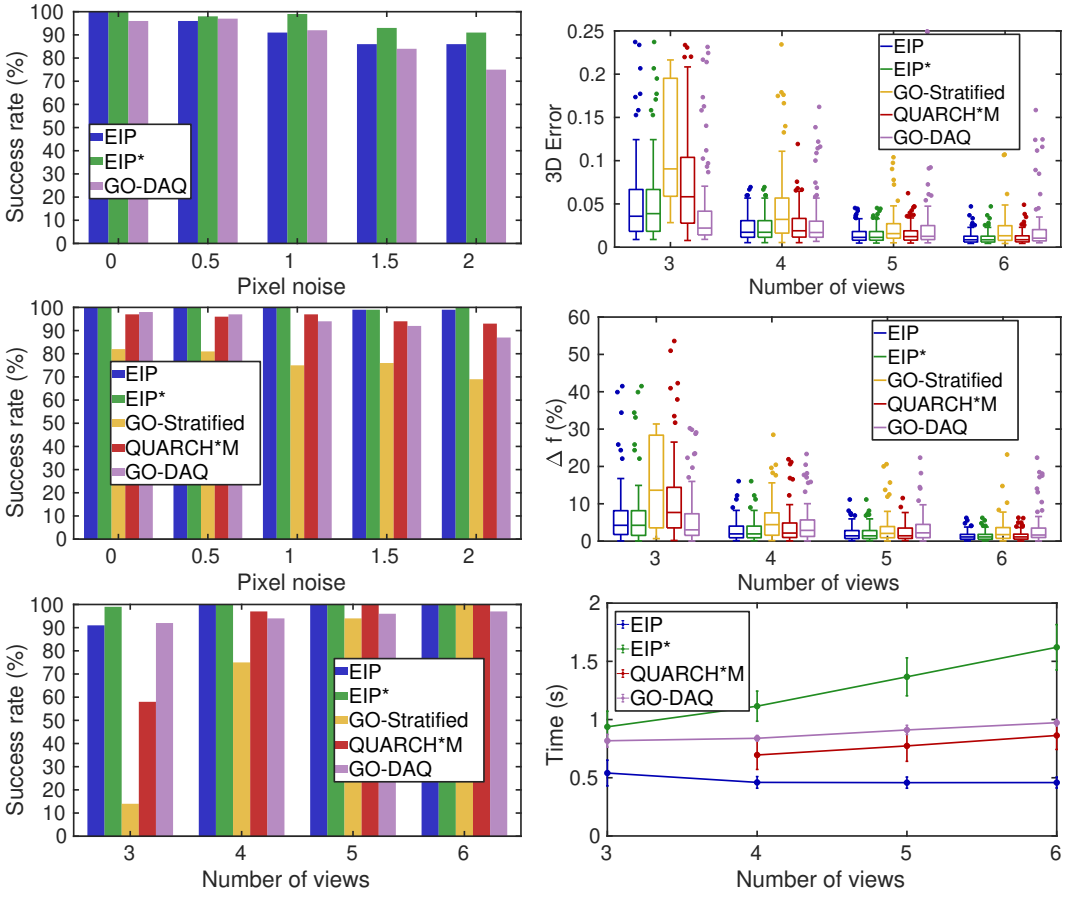


Figure 3.3: Comparisons with the state of the art. Left: success rate using 3 views (top) and 4 views (middle) with varying noise levels, and using 3–6 views with 1 pixel noise level (bottom). Right: distribution of 3D errors (top) and focal length errors (middle), and runtime results (bottom) using 3–6 views with 1 pixel noise.

GO-Stratified and QUARCH*M as well due to the additional EIP constraint. With more views, all the methods succeeded most of the time. The 3D and focal length error distributions as well as the runtime results are shown in Figure 3.3 (right column). GO-Stratified obtained relatively higher 3D and calibration errors and was also two orders of magnitude slower (not shown) than the other methods. For our algorithm, we report the computation time excluding the problem modeling overhead in GloptiPoly. The time complexity of EIP is constant with respect to the number of images as it does not use the pairwise inequalities. From four views onward, EIP can thus be used instead of EIP* for a speedup as it achieved a similar success rate and calibration accuracy in this test.

Sequence	Method	$\Delta f(\%)$	$\Delta uv(\%)$	$\Delta\gamma$	Time (s)
fountain-P11	MODULUS	62.90	66.39	2172.12	0.91
	EIP	0.08	0.25	1.06	0.59
	GO-Stratified	0.10	0.19	1.08	302.90
	QUARCH*M	0.05	0.23	1.05	2.44
	GO-DAQ	0.36	1.26	0.01	1.49
Herz-Jesu-P8	MODULUS	0.89	3.12	2.16	0.82
	EIP	0.55	2.84	3.98	0.57
	GO-Stratified	43.86	31.13	157.31	243.18
	QUARCH*M	0.88	3.11	2.03	1.26
	GO-DAQ	1.43	1.27	0.05	1.53
City hall Leuven	MODULUS	2.96	6.73	5.90	0.62
	EIP	0.78	0.72	2.80	0.56
	GO-Stratified	7.09	10.10	25.85	169.21
	QUARCH*M	2.94	6.70	5.81	1.02
	GO-DAQ	9.93	7.68	9.70	1.38

Table 3.1: Quantitative assessment. Autocalibration results on the real image sequences from [Strecha et al. 2003; Strecha et al. 2008].

3.5.2 Real image experiments

We used four image sequences, *fountain-P11*, *Herz-Jesu-P8*, *Herz-Jesu-P25* [Strecha et al. 2008], and *City hall Leuven* [Strecha et al. 2003], with known ground truth calibration to quantitatively compare our algorithm with existing methods. The ground truth intrinsic parameters come close to satisfying the EIP assumption for these sequences. We also qualitatively assessed the metric reconstructions obtained with our algorithm using other image sequences. We used P2SfM [Magerand and Bue 2018] to compute the projective reconstructions and COLMAP [Schönberger and Frahm 2016] to obtain the feature matches.

Quantitative assessment: Table 3.1 reports the calibration errors from MODULUS, EIP, and state-of-the-art methods on three tested sequences. With *fountain-P11*, MODULUS led to large calibration errors and failed to obtain a metric upgrade. MODULUS* provided a calibration similar to that from EIP, but it required 10 times the computation time. MODULUS and EIP otherwise yielded the same calibration

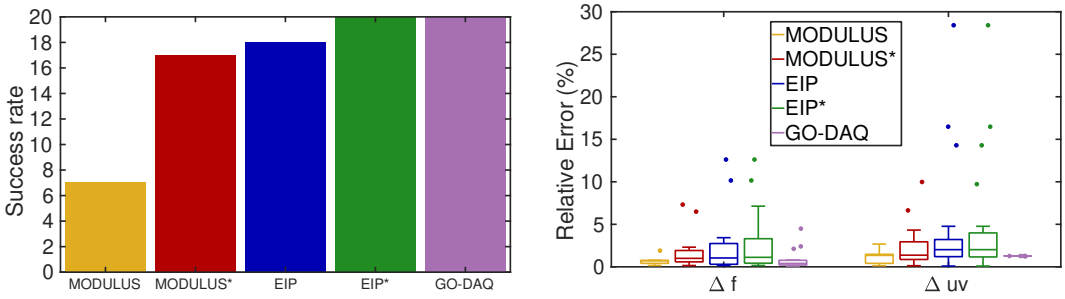
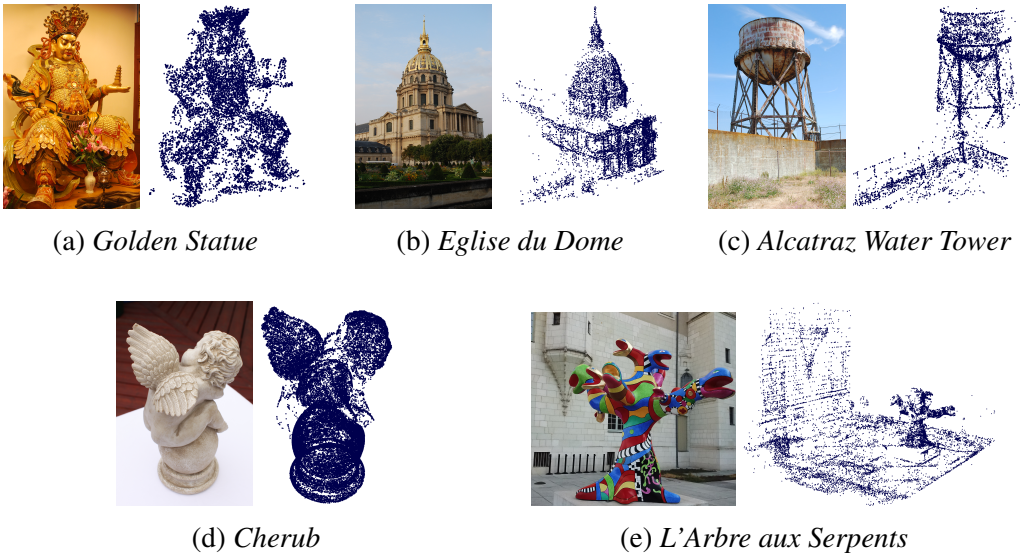


Figure 3.4: Quantitative assessment using 3 views. Success rate (left) and distribution of calibration errors (right) over 20 image triplets sampled from *Herz-Jesu-P25*.

as MODULUS* and EIP*, respectively, in this experiment. With *Herz-Jesu-P8*, GO-Stratified failed to obtain a metric upgrade. Moreover, while the results improved with the longer *Herz-Jesu-P25* sequence for all the methods (not shown), GO-Stratified still led to an erroneous calibration. This is due to the method relying on scene points that prove unreliable with noise and outliers. With *City hall Leuven*, the reference calibration parameters do not fit the assumptions of GO-DAQ as closely as those of the previous sequences. The principal point is farther from the image center and the skew is not zero. As a result, the errors are larger using GO-DAQ. In contrast, EIP provided an accurate calibration for all the sequences and required only around half a second of computation time. The errors from MODULUS and QUARCH*M are similar as they rely on the same cost.

Quantitative assessment using 3 views: To test the minimal case of three views, we sampled image triplets sequentially from the *Herz-Jesu-P25* sequence, discarding those with insufficient feature matches, leaving a set of 20 triplets. Figure 3.4 shows the results from MODULUS, EIP, their inequality-constrained counterparts, and GO-DAQ on this set. We considered a focal length error above 25% as a failure in this experiment. From our tests, the quality of the metric reconstruction was mostly influenced by the estimated focal length, and errors above this threshold corresponded to distorted reconstructions. The results in Figure 3.4 are consistent with those on the synthetic data as MODULUS failed most of the time and EIP performed reliably. Both EIP* and GO-DAQ succeeded with all the triplets. GO-DAQ also consistently provided an accurate calibration as its assumptions on the intrinsic parameters are closely satisfied in this sequence. The skew parameter was also ac-



(a) *Golden Statue* (b) *Eglise du Dome* (c) *Alcatraz Water Tower*

(d) *Cherub* (e) *L'Arbre aux Serpents*

Figure 3.5: Qualitative assessment. Sample images and metric 3D reconstructions obtained using EIP. (e) Image courtesy of Renato Saleri. *L'Arbre aux Serpents de Niki de Saint Phalle* © Musées d'Angers, Niki Charitable Art Foundation.

curately estimated by all the methods (not shown), resulting in less than 1° deviation from a rectangular image plane on average.

Qualitative assessment: We visually assessed the metric 3D reconstructions obtained with our algorithm using several other image sequences. These sequences include (number of images given in parentheses): *Golden Statue* (18), *Eglise du Dome* (85), and *Alcatraz Water Tower* (173) from [Olsson and Enqvist 2011], *Cherub* (65) from 3DFlow[4], as well as one from a digital cultural heritage application, *L'Arbre aux Serpents* (154). The last sequence was captured by a camera drone for digital preservation of the imaged sculpture as part of a French national research project[5]. The projective reconstructions contained a maximum of 66 cameras for the tested sequences. Figure 3.5 shows the metric 3D reconstructions obtained using EIP. The recovered metric structure faithfully represents the imaged scene. Thus, the EIP assumption was applicable for all these sequences that have been captured using dif-

[4] `https://www.3dflow.net/3df-zephyr-reconstruction-showcase/`
[5] `https://anr-sumum.fr/`

ferent cameras. These results were obtained in just over one second of computation time on average.

3.6 Conclusion

In this chapter, we addressed the stratified camera autocalibration problem for a moving camera with EIP and otherwise constant but unknown intrinsic parameters. We showed that, with these assumptions, each image pair provides a new quartic polynomial in Π_∞ in addition to the polynomial from the modulus constraint. We proposed a stratified autocalibration method that uses the new EIP-based polynomial in conjunction with the modulus constraint to estimate Π_∞. For three or more images, we formulated the problem of estimating Π_∞ as a constrained polynomial optimization problem that is solved using Lasserre's hierarchy of semidefinite relaxations. Our experiments with synthetic data and real images show the benefits of the EIP polynomial in reliably retrieving a metric reconstruction from uncalibrated images, especially for short sequences. In the future, it may be worth investigating other ways the EIP polynomial can be exploited. For instance, given the EIP polynomial and the modulus constraint, autocalibration is possible using only two images with one additional equation (from a vanishing point, for example). This is also potentially interesting for the uncalibrated view synthesis problem [Canclini et al. 2019].

4 Exploiting camera motion

This chapter presents our contributions in exploiting a partial knowledge of the camera motion in camera autocalibration. We consider the case of a moving camera with constant intrinsic parameters where some vague knowledge of its relative orientation is available. We derive two new sets of constraints on the plane at infinity based on such knowledge. The constraints in the first set are convex and can be exploited when the orientation angle between camera pairs is known to be either under or over 120°. They describe the relationship between the plane at infinity and new geometric objects, the hodographs of the horopter. The constraints in the second set are non-convex and exploit a tighter camera orientation knowledge. Using the orientation-based constraints, we show that a new quasi-affine reconstruction of a scene can be recovered, referred to as a QUARCH. We propose a stratified autocalibration method that relies on a QUARCH to recover an affine and a metric reconstruction from a projective one. We finally report the experiments conducted to evaluate our method using both synthetic data and real images.

4.1 Introduction

Camera motion constraints: Camera autocalibration is a highly nonlinear and challenging problem. To cope with the involved challenges, methods based on restricted camera motions have been found to be more reliable than those based on general motion. The camera motion is inherently restricted in some applications. For instance, a camera mounted on a ground vehicle is constrained to a planar motion, while a pan-tilt-zoom (PTZ) camera may only rotate about its center. Autocalibration methods have been proposed to exploit different constrained camera motions, including pure translation [Armstrong et al. 1994; Moons et al. 1996; Pollefeys et al. 1996], pure rotation [Hartley 1997b; Agapito et al. 2001; Li and C. Shen 2006; Rameau et al. 2012], and planar motion [Armstrong et al. 1996; Faugeras et al. 2000; Espuny 2007]. The camera orientation may also be known, for example, from the measurements of an external sensor such as an Inertial Measurement Unit (IMU). Methods exploiting a known [Frahm and Koch 2003] or partially known [Martyushev 2018] camera orientation have been proposed in the literature. Such prior knowledge of the camera motion often leads to a simpler problem or to additional constraints that can be exploited in autocalibration (see Section 2.3.5 for a more detailed discussion).

Vague motion knowledge: In the general context of image-based 3D modeling (*e.g.* with a hand-held camera), the camera motion is usually not as restricted as described above and exact motion knowledge is also not available. However, there is some implicit constraint on the camera motion that stems from the way images are typically captured for 3D reconstruction. The camera is typically moved smoothly around the scene with a mild change in viewpoint between the images to ensure a sufficient image overlap for feature matching. Such a motion is also generally recommended by 3D reconstruction software[1]. Thus, some vague knowledge of the camera motion is usually available in applications of image-based 3D modeling. Efforts have been made to formalize and exploit such vague knowledge in camera autocalibration. In [F. Shen and Wang 2002], the authors assumed a small rotation

[1] `https://www.3dflow.net/technology/documents/photogrammetry-how-to-acquire-pictures/`

between consecutive views, *i.e*. the rotation angle about each axis being within 10°, to obtain linear equations in the coordinates of the plane at infinity (Π_∞). These equations are derived by approximating the rotation matrix and some of the camera intrinsic parameters. While the small rotation assumption is reasonable for a video, it is not broadly applicable to image sequences. In [Habed et al. 2014], the authors introduced Linear Matrix Inequality (LMI) constraints in the entries of the DAQ assuming that the relative orientation angle between consecutive views is within 90°. This is a mild assumption that is often implicitly satisfied in practice. In fact, from our inspection of numerous ordered image sequences, the orientation angle between consecutive views tends to be much smaller than 90°. It is interesting to note that the orientation angle is encoded in the eigenvalues of the inter-image homography induced by Π_∞. Therefore, one may wonder if such orientation constraints can be imposed on Π_∞ alone and used in a stratified autocalibration method.

Affine and quasi-affine strata: In stratified camera autocalibration, reliably locating Π_∞ to recover an affine reconstruction has proved to be challenging due to the nonlinearity of the problem [Hartley and Zisserman 2004]. Once it is located, the camera intrinsic parameters, and thereby a metric reconstruction, can be obtained by solving linear equations. For the case of constant intrinsic parameters, stratified methods usually rely on the so-called modulus constraint [Pollefeys and Van Gool 1999] that is a necessary condition on Π_∞ for the eigenvalues of its inter-image homography matrices to have equal moduli. Further knowledge of the camera geometry has been exploited in [Habed et al. 2012; F. Wu et al. 2013; Adlakha et al. 2020]. Other methods [Hartley 1994a; Hartley et al. 1999; Nistér 2004b; Chandraker et al. 2010] locate Π_∞ by first upgrading a projective reconstruction to a quasi-affine one based on Hartley's chirality theory [Hartley and Zisserman 2004]. In a quasi-affine reference frame, the chirality inequalities can be used to obtain bounds on the coordinates of Π_∞ (see Section 2.2.6). The sought Π_∞ can then be located within these bounds through random sampling followed by a local optimization [Hartley 1994a], an exhaustive search [Hartley et al. 1999], or a search via a Branch-and-Bound algorithm [Chandraker et al. 2010]. The quasi-affine reconstruction in these methods is with respect to the set of camera centers and that of scene points: the sets whose respective convex hulls are preserved. Nistér pointed out that scene points may not

be reliable and therefore sought a quasi-affine reconstruction with respect to camera centers (QUARC) alone [Nistér 2004b]. A QUARC is then upgraded to a metric reconstruction through nonlinear optimization of a geometrically meaningful cost function derived from priors on the intrinsic parameters.

Contributions: In this chapter, we present some contributions in exploiting a vague knowledge of the camera orientation in stratified camera autocalibration. Our contributions address the case of a moving camera with constant intrinsic parameters and are stated in the following.

(i) Orientation-based convex constraints: We show that a vague knowledge of the relative camera orientation can be exploited to constrain the location of Π_∞. For a camera pair with relative orientation angle[2] θ, we show that Π_∞ belongs to one of two convex sets depending on whether $|\theta| \leq 120^\circ$ or $|\theta| \geq 120^\circ$. These convex sets are defined by new quadratic constraints on Π_∞ that we formulate as LMIs. We show that the LMI constraints describe the relationship between Π_∞ and new geometric objects associated to a camera pair termed the *hodographs of the horopter*.

(ii) Orientation-based non-convex constraints: We show that a tighter relative camera orientation knowledge of $|\theta| < 90^\circ/k$, where k is a positive integer, can be exploited to further constrain Π_∞. We derive a set of constraints for this case using the Hurwitz stability criterion. These constraints are non-convex, and they are formulated as polynomial inequalities of degree $2k$.

(iii) QUARCH: We show the existence of a new quasi-affine reconstruction of a scene that can be obtained when a vague knowledge of the relative camera orientation angle between camera pairs is available. We refer to this reconstruction as a QUARCH: QUasi-Affine Reconstruction with respect to Camera centers and the Hodographs of horopters. A QUARCH is a QUARC that additionally preserves affine properties with respect to the hodographs of horopters. We propose a Semidefinite Programming (SDP) formulation to efficiently compute a QUARCH.

[2] The relative orientation angle θ is about an arbitrary axis in the axis-angle representation, where $\theta \in [-180^\circ, 180^\circ]$.

(iv) LMI-constrained optimization method: We propose a constrained Levenberg-Marquardt method for nonlinear optimization subject to LMI constraints. We use this method to upgrade a QUARCH to an affine reconstruction.

(v) Stratified autocalibration method: We propose a stratified autocalibration algorithm that recovers a QUARCH as an initial step towards obtaining an affine and a metric reconstruction from a projective one. Experiments with synthetic data and real images show that our algorithm performs more reliably than existing methods, especially for short sequences.

Organization: The rest of this chapter is organized as follows. In Section 4.2, we introduce the hodographs of the horopter and derive the set of orientation-based convex constraints on Π_∞ using them. We also present the QUARCH stratum based on these convex constraints. In Section 4.3, we derive the set of orientation-based non-convex constraints on Π_∞ using the Hurwitz stability criterion . We further present a specialized QUARCH using these constraints. In Section 4.4, we detail the LMI-constrained optimization method and our proposed stratified autocalibration algorithm. We report the experimental evaluation of our algorithm in Section 4.5. Finally, Section 4.6 concludes this chapter.

4.2 Orientation-based convex constraints

In this section, we present a new set of camera orientation-based constraints on Π_∞. These constraints are convex, formulated as LMIs, and they are derived from the relations between Π_∞ and new geometric objects associated to a camera pair, the hodographs of the horopter. We first review the horopter curve and some of its properties from the literature in Sections 4.2.1 and 4.2.2. Then, we define the hodographs of the horopter in Section 4.2.3 and study their geometric relationship with Π_∞ in Section 4.2.4, culminating in the new set of camera orientation-based constraints on Π_∞. Finally, we show the existence of a new quasi-affine reconstruction stratum, the QUARCH stratum, using these constraints in Section 4.2.5.

4.2.1 Horopters

In the study of binocular vision, the horopter is the locus of points in space that project to identical points in both retinas [Von Helmholtz 1867]. In computer vision, Maybank studied the horopter curves in the context of ambiguous surfaces for 3D reconstruction [Maybank 1990; Maybank 1993]. Horopters were then used in camera autocalibration [Armstrong et al. 1996]. Schaffalitzky showed the connection between the horopter curves and the modulus constraint [Pollefeys and Van Gool 1999] and derived a cubic polynomial constraint on Π_∞ for three views based on some of their properties [Schaffalitzky 2000]. Further properties of horopters as well as camera autocalibration algorithms based on them have been presented in [Ronda et al. 2004]. Other works on camera autocalibration have used horopters for certain constrained camera motions (primarily planar motion) [Armstrong et al. 1996; Faugeras et al. 2000; Espuny 2007; Espuny et al. 2011]. We now formally define the horopter associated to two cameras. To do so, we rely on an algebraic expression of the nullspace of a camera projection matrix defined in [Schaffalitzky 2000].

Definition 4.1 (Algebraic nullspace)**.** *Given a full row rank camera projection matrix* P*, the algebraic nullspace* $\mathcal{N}(\mathsf{P})$ *of* P *is defined by the equation,*

$$\det\begin{pmatrix} \mathsf{P} \\ \Pi^\top \end{pmatrix} = \Pi^\top \mathcal{N}(\mathsf{P}), \tag{4.1}$$

for any plane Π *in* $\mathbb{P}^{*3}$*.*

The algebraic nullspace is a cubic function of its argument. Its scale is defined by the scale of the matrix P. This representation is especially useful as we will often be interested in the exact scale of homogeneous entities in this chapter. The vector $\mathcal{N}(\mathsf{P})$ gives the coordinates of the optical center of the camera represented by the matrix P. We are now ready to define the horopter associated to a camera pair as follows.

Definition 4.2 (Horopter)**.** *The horopter* $\mathcal{H}$ *of a camera pair* (i, j) *with identical intrinsic parameters is the locus of points in* $\mathbb{P}^3$ *that are imaged at the same coordinates by both cameras. Let* P_i *and* P_j *be the projection matrices of the two*

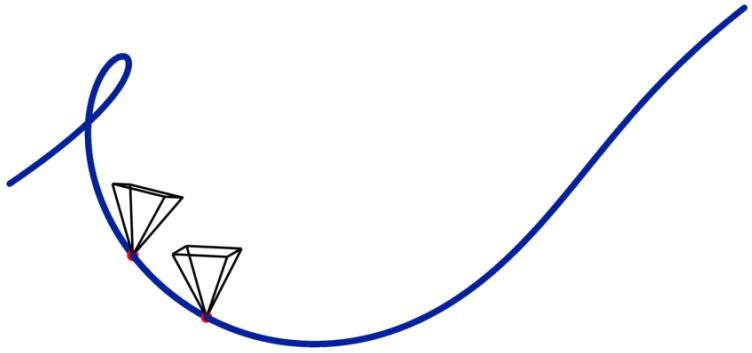


Figure 4.1: The horopter of a camera pair. The horopter (shown in blue) is a twisted cubic that passes through both camera centers (shown in red).

cameras. A point X lies on $\mathcal{H}$ if $\mathsf{P}_i\mathsf{X} \simeq \mathsf{P}_j\mathsf{X}$, and the locus of these points is $\mathcal{H}(s,t) = \mathcal{N}(s\mathsf{P}_i - t\mathsf{P}_j)$ for parameters s and t. The parametric representation of the horopter is given by the expansion,

$$\mathcal{H}(s,t) = s^3\,\mathsf{C}_i - s^2t\,\mathsf{T}_{ij} + st^2\,\mathsf{T}_{ji} - t^3\,\mathsf{C}_j, \tag{4.2}$$

where $\mathsf{C}_i = \mathcal{N}(\mathsf{P}_i)$ and $\mathsf{C}_j = \mathcal{N}(\mathsf{P}_j)$ are the coordinates of the two camera centers, $\mathsf{T}_{ij} = \mathcal{T}(\mathsf{P}_i, \mathsf{P}_j)$, $\mathsf{T}_{ji} = \mathcal{T}(\mathsf{P}_j, \mathsf{P}_i)$, and operator $\mathcal{T}$ is defined by this expansion.

Note that we represent the horopter in homogeneous form by using two parameters in Definition 4.2. From its parametric representation (4.2), we observe that the horopter of a camera pair is a twisted cubic in $\mathbb{P}^3$ that passes through both camera centers. Figure 4.1 shows a simulated camera pair and its associated horopter. The vectors C_i and C_j in (4.2) give the coordinates of the camera centers C_i and C_j, respectively. The points with coordinates T_{ij} and T_{ji}, however, appear less familiar at this point. These points feature throughout our discussion of the orientation-based constraints on Π_∞, and we define the following term to refer to them.

Definition 4.3 (Control points of the horopter)**.** *The virtual points T_{ij} and T_{ji} with coordinate vectors T_{ij} and T_{ji}, respectively, in the parametric representation of the horopter $\mathcal{H}$ are the control points of $\mathcal{H}$.*

This terminology is inspired from its use in computer graphics and geometric modeling, and in particular, from the representation of parametric curves through a set of control points [Agoston 2005]. We will have more to say about the control

points of the horopter in Section 4.2.4. In the following section, we present the connection between the horopter curves and the modulus constraint.

4.2.2 Horopters and the modulus constraint

We review here some results on the relationship of Π_∞ with the horopter curves. The following lemma shows the connection between their incidence and the inter-image infinite homography, as stated in [Schaffalitzky 2000]. For the sake of completeness, we also provide the proof.

Lemma 4.1. *For a camera pair (i, j), let P_i and P_j be the projection matrices, and let $\mathcal{H}$ be the attached horopter such that $\mathcal{H}(s,t) = \mathcal{N}(s\mathsf{P}_i - t\mathsf{P}_j)$. Let Π_∞ be the plane at infinity with coordinates $\Pi_\infty = \left(\pi_\infty^\top, 1\right)^\top$, and let $\mathsf{H}_{\infty ij}$ be its induced inter-image homography, where $\mathsf{H}_{\infty 1i} = \mathsf{P}_i \begin{bmatrix} \mathsf{I}_3 & -\pi_\infty \end{bmatrix}^\top$ and $\mathsf{H}_{\infty ij} = \mathsf{H}_{\infty 1j}\mathsf{H}_{\infty 1i}^*$. We have that,*

$$\det(s\mathsf{H}_{\infty 1i} - t\mathsf{H}_{\infty 1j}) = \Pi_\infty^\top \mathcal{H}(s,t). \tag{4.3}$$

Proof. The proof is along the lines of that given in [Schaffalitzky 2000]. Assume that the cameras are embedded in an affine reference frame, where the projection matrices are of the form $\mathsf{P}_i^{\mathsf{A}} = [\, \mathsf{H}_{\infty 1i} \mid \mathsf{e}_{1i} \,]$, e_{1i} being the epipole, and Π_∞ is at its canonical position with coordinates $\left[\mathbf{0}^\top, 1\right]^\top$. Now, the left-hand side of (4.3) can be written as,

$$\det(s\mathsf{H}_{\infty 1i} - t\mathsf{H}_{\infty 1j}) = \det \begin{pmatrix} s\mathsf{H}_{\infty 1i} - t\mathsf{H}_{\infty 1j} & s\mathsf{e}_{1i} - t\mathsf{e}_{1j} \\ \mathbf{0}_3^\top & 1 \end{pmatrix} \tag{4.4a}$$

$$= \det \begin{pmatrix} s\mathsf{P}_i^{\mathsf{A}} - t\mathsf{P}_j^{\mathsf{A}} \\ \mathbf{0}_3^\top \quad 1 \end{pmatrix} \tag{4.4b}$$

$$= \det \begin{pmatrix} s\mathsf{P}_i - t\mathsf{P}_j \\ \pi_\infty^\top \quad 1 \end{pmatrix} \det \begin{pmatrix} \mathsf{I}_3 & \mathbf{0}_3 \\ -\pi_\infty^\top & 1 \end{pmatrix} \tag{4.4c}$$

$$= \Pi_\infty^\top \mathcal{N}(s\mathsf{P}_i - t\mathsf{P}_j), \tag{4.4d}$$

thus leading to (4.3). ∎

Lemma 4.1 links the characteristic polynomial of the inter-image infinite homography with the vanishing polynomial representing the intersection of Π_∞ and the

horopter of the considered camera pair. Note that in [Schaffalitzky 2000], this result is stated more generally for any plane Π and its induced inter-image homography. The following properties [Schaffalitzky 2000; Ronda et al. 2004] of the intersection of Π_∞ with the horopter curves follow from Lemma 4.1.

Property 4.1. *The plane at infinity Π_∞ meets the horopter $\mathcal{H}$ of a camera pair at three points whose parameter values are related to the eigenvalues of its induced inter-image homography matrix.*

Property 4.2. *Consider a camera pair (i, j) with projection matrices P_i and P_j, and attached horopter $\mathcal{H}$ with parametric representation $\mathcal{H}(s,t) = s^3\,\mathsf{C}_i - s^2t\,\mathsf{T}_{ij} + st^2\,\mathsf{T}_{ji} - t^3\,\mathsf{C}_j$. The plane at infinity Π_∞, with coordinates $\Pi_\infty = \left(\pi_\infty^\top, 1\right)^\top$, induces an inter-image homography $\mathsf{H}_{\infty ij}$, where $\mathsf{H}_{\infty 1i} = \mathsf{P}_i \begin{bmatrix} \mathsf{I}_3 & -\pi_\infty \end{bmatrix}^\top$ and $\mathsf{H}_{\infty ij} = \mathsf{H}_{\infty 1j}\mathsf{H}_{\infty 1i}^*$. The coefficients of the polynomial $\Pi_\infty^\top \mathcal{H}(s,t)$ satisfy:*

$$\Pi_\infty^\top \mathsf{C}_i = \det(\mathsf{H}_{\infty 1i}) = \lambda_i^3, \qquad \Pi_\infty^\top \mathsf{C}_j = \det(\mathsf{H}_{\infty 1j}) = \lambda_j^3, \tag{4.5a}$$

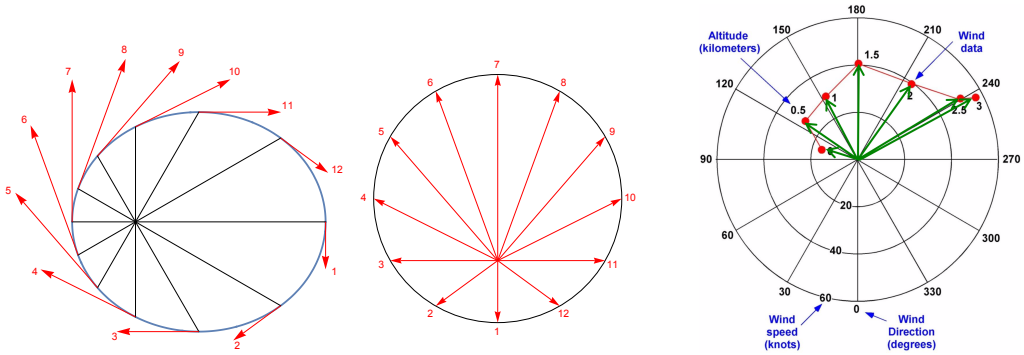
$$\Pi_\infty^\top \mathsf{T}_{ij} = \mathrm{tr}(\mathsf{H}_{\infty ij}) = \lambda_i^2\lambda_j \mathsf{a}_{ij}, \qquad \Pi_\infty^\top \mathsf{T}_{ji} = \mathrm{tr}(\mathsf{H}_{\infty ji}) = \lambda_i\lambda_j^2 \mathsf{a}_{ij}, \tag{4.5b}$$

where $\mathsf{a}_{ij} = 1 + 2\cos\theta_{ij}$, with θ_{ij} the relative orientation angle between the two cameras, and λ_i and λ_j are two scalars.

As the horopter curve is a twisted cubic, it intersects any plane, including Π_∞, at three points. From Property 4.1, the points of intersection with Π_∞ are related to the eigenvalues of the inter-image infinite homography matrix. These eigenvalues have equal moduli and a necessary condition for this so-called modulus constraint [Pollefeys and Van Gool 1999] is that a quartic polynomial equation in the coordinates of Π_∞ is satisfied (see Section 2.3.4). Schaffalitzky showed that this polynomial constraint can be expressed in terms of the coefficients in Property 4.2 as follows [Schaffalitzky 2000].

Proposition 4.2 (Modulus constraint)**.** *For a camera pair (i, j), let $\mathcal{H}$ be the attached horopter with parametric representation $\mathcal{H}(s,t) = s^3\,\mathsf{C}_i - s^2t\,\mathsf{T}_{ij} + st^2\,\mathsf{T}_{ji} - t^3\,\mathsf{C}_j$. A necessary condition for a plane Π in $\mathbb{P}^{*3}$ to be the plane at infinity Π_∞ is that it satisfies the following quartic equation,*

$$\mathcal{M}_{ij}(\Pi) = \Pi^\top \mathsf{C}_i (\Pi^\top \mathsf{T}_{ji})^3 - \Pi^\top \mathsf{C}_j (\Pi^\top \mathsf{T}_{ij})^3 = 0. \tag{4.6}$$



(a) Classical mechanics (b) Meteorology

Figure 4.2: Some applications of hodographs. (a) The elliptical orbit of a body and its instantaneous velocity vectors (left). The velocity vectors are translated to a common origin and the traced hodograph is a circle for this trajectory (right). Hodographs serve as a geometric tool in classical mechanics. Figure from the Wolfram Demonstrations Project[3] (S. M. Blinder, "Hodographs for Kepler Orbits"). (b) The hodograph (red) traced by the wind velocity vectors (green) is plotted on a polar diagram. Such diagrams are used in meteorology to analyze the behaviour of wind. Figure courtesy of the National Weather Service Louisville, Kentucky[4].

The polynomial constraint in Proposition 4.2 is obtained from the coefficients in Property 4.2 by equating the scalars. This completes our review of horopters and their relevant properties. In the next section, we introduce new geometric objects derived from the horopter, the hodographs of the horopter.

4.2.3 Hodographs of the horopter

The hodograph refers to the curve traced by the velocity vectors of a moving body. The term was coined by William R. Hamilton, who used this curve to study orbital motion [Hamilton 1847]. Hodographs have been used since then as a geometric tool in classical mechanics [Derbes 2001; Apostolatos 2003]. The application of hodographs is perhaps better known though from meteorology, where the hodograph traced by the wind velocity vectors is used to analyze the behavior of wind [Pucik et al. 2021]. Figure 4.2 shows these two applications of hodographs.

[3] https://demonstrations.wolfram.com/HodographsForKeplerOrbits/
[4] https://www.weather.gov/lmk/

(a) A planar curve (blue) and its tangent vectors. (b) The hodograph (red) of the curve.

Figure 4.3: A parametric curve and its hodograph. In (b), the tangent vectors of the curve in (a) are translated to a common origin (and scaled for visualization), and the hodograph is traced out by joining their tips. Figure adapted from [Agoston 2005, p. 460].

More formally, the hodograph of a parametric curve is the locus defined by its derivative [Perwass et al. 2009, Sec. 9.2]. Figure 4.3 shows an example of a parametric curve and its hodograph. The curve in Figure 4.3(a) can be interpreted as the trajectory of a moving point, where the tangent vectors represent its instantaneous velocity at different instants. Hodographs are used in geometric modeling to characterize the properties of various curves [Agoston 2005]. Typically, in such a context, the considered curve is embedded in a Cartesian space and represented using a single parameter. As such, it admits a single hodograph. However, in the context of the present work, we are interested in the horopter curve as an object embedded in a projective space and defined using two parameters. It hence admits two hodographs that are defined as follows.

Definition 4.4 (Hodographs of the horopter)**.** *For a camera pair (i, j) with attached horopter $\mathcal{H}$, the hodographs $\mathcal{H}_s$ and $\mathcal{H}_t$ of $\mathcal{H}$ are the curves defined by the partial derivatives of the parametric horopter. Let $\mathcal{H}_s(s,t) = \frac{\partial}{\partial s}\mathcal{H}(s,t)$ and $\mathcal{H}_t(s,t) = \frac{\partial}{\partial t}\mathcal{H}(s,t)$. The parametric forms of these curves are:*

$$\begin{aligned} \mathcal{H}_s(s,t) &= 3s^2\,\mathsf{C}_i - 2st\,\mathsf{T}_{ij} + t^2\,\mathsf{T}_{ji}, \\ \mathcal{H}_t(s,t) &= -3t^2\,\mathsf{C}_j + 2st\,\mathsf{T}_{ji} - s^2\,\mathsf{T}_{ij}. \end{aligned} \tag{4.7}$$

The hodographs of the horopter are two conics in $\mathbb{P}^3$. Observe that $\mathcal{H}_s$ passes through the points C_i and T_{ji}, while $\mathcal{H}_t$ passes through C_j and T_{ij}. Hereafter, we

refer to the hodographs of the horopter simply as the hodographs. In the following section, we study the geometric relationship between Π_∞ and these curves, and we show that their incidence is described by a new set of convex constraints on Π_∞.

4.2.4 Hodographs-based constraints on infinity

Our interest in this section is to characterize the geometric relationship between Π_∞ and the hodograph curves. We will show that the hodographs act as positioning objects to locate Π_∞ when a vague knowledge of the relative camera orientation is available. In particular, their positioning with respect to Π_∞ is described by new quadratic convex inequalities in Π_∞ that are formulated as LMIs. Our main result then in this section is that Π_∞ is constrained to one of two convex sets, defined by the new LMI constraints, depending on whether $|\theta_{ij}| \leq 120^\circ$ or $|\theta_{ij}| \geq 120^\circ$, where θ_{ij} is the orientation angle between the two cameras. The results for these two cases are stated in Propositions 4.6 and 4.7. To meaningfully describe the relative positioning of Π_∞ and the hodographs in projective space, we consider a fixed representation of the horopter (and thereby its associated hodographs) that we define as the *oriented horopter*.

Definition 4.5 (Oriented horopter)**.** *The oriented horopter of a camera pair is the horopter $\mathcal{H}$ that is obtained using sign-corrected projection matrices.*

It follows from Definition 4.5 that for an oriented horopter, both the camera centers lie on one side with respect to Π_∞. Using this representation, we first characterize the geometric relationship between Π_∞ and the control points of an oriented horopter. In the following lemma, we show that the positioning of these points with respect to Π_∞ depends on the relative camera orientation angle and that this relationship is described by a new set of linear constraints on Π_∞.

Lemma 4.3. *For a camera pair (i, j), let θ_{ij} be the relative camera orientation angle and $\mathcal{H}$ the attached oriented horopter. With respect to the plane at infinity Π_∞, the control points of $\mathcal{H}$ lie: (i) on the same side as the camera centers if $|\theta_{ij}| < 120^\circ$, and (ii) on the opposite side as the camera centers if $|\theta_{ij}| > 120^\circ$. If $|\theta_{ij}| = 120^\circ$, they lie on Π_∞.*

Proof. From Property 4.2, the incidence of Π_∞ with the control points T_{ij} and T_{ji} of $\mathcal{H}$ is given as $\Pi_\infty^\top \mathsf{T}_{ij} = \lambda_i^2 \lambda_j \mathsf{a}_{ij}$ and $\Pi_\infty^\top \mathsf{T}_{ji} = \lambda_i \lambda_j^2 \mathsf{a}_{ij}$, where $\mathsf{a}_{ij} = 1 + 2\cos\theta_{ij}$. Thus, $\Pi_\infty^\top \mathsf{T}_{ij}$ and $\Pi_\infty^\top \mathsf{T}_{ji}$ depend on the angle θ_{ij}. For an oriented horopter, the camera centers C_i and C_j lie on the same side with respect to Π_∞, their incidence arbitrarily being, $\Pi_\infty^\top \mathsf{C}_i > 0$ and $\Pi_\infty^\top \mathsf{C}_j > 0$. From Property 4.2, we also have that $\Pi_\infty^\top \mathsf{C}_i = \lambda_i^3$ and $\Pi_\infty^\top \mathsf{C}_j = \lambda_j^3$, hence λ_i and λ_j are positive. As a result, $\operatorname{sgn}(\Pi_\infty^\top \mathsf{T}_{ij}) = \operatorname{sgn}(\Pi_\infty^\top \mathsf{T}_{ji}) = \operatorname{sgn}(\mathsf{a}_{ij})$, where $\operatorname{sgn}(\cdot)$ is the sign function and a_{ij} lies in the interval:

$$0 \leq \mathsf{a}_{ij} \leq 3 \;\text{ if }\; |\theta_{ij}| \leq 120^\circ, \tag{4.8a}$$

$$-1 \leq \mathsf{a}_{ij} \leq 0 \;\text{ if }\; |\theta_{ij}| \geq 120^\circ, \tag{4.8b}$$

and $\mathsf{a}_{ij} = 0$ if $|\theta_{ij}| = 120^\circ$. Thus, Π_∞ satisfies the following inequalities:

$$\Pi_\infty^\top \mathsf{T}_{ij} \geq 0 \;\text{ and }\; \Pi_\infty^\top \mathsf{T}_{ji} \geq 0 \;\text{ if }\; |\theta_{ij}| \leq 120^\circ, \tag{4.9a}$$

$$\Pi_\infty^\top \mathsf{T}_{ij} \leq 0 \;\text{ and }\; \Pi_\infty^\top \mathsf{T}_{ji} \leq 0 \;\text{ if }\; |\theta_{ij}| \geq 120^\circ, \tag{4.9b}$$

where equality holds if $|\theta_{ij}| = 120^\circ$. Hence, the control points and the camera centers are on the same side with respect to Π_∞ when $|\theta_{ij}| < 120^\circ$, whereas they lie on opposite sides with respect to Π_∞ when $|\theta_{ij}| > 120^\circ$. When $|\theta_{ij}| = 120^\circ$, the control points lie on Π_∞. ■

Lemma 4.3 introduces a new camera orientation-based constraint on Π_∞. The linear inequalities (4.9) can be used to constrain the location of Π_∞ when a vague knowledge of the relative orientation angle between camera pairs, *i.e.* $|\theta_{ij}| \leq 120^\circ$ or $|\theta_{ij}| \geq 120^\circ$, is available. Note that when $|\theta_{ij}| = 120^\circ$, Π_∞ passes through the control points of the horopter and we obtain two linear equations in the coordinates of Π_∞. Such precise camera rotation may be performed in a controlled setting, where these equations can be used to estimate Π_∞. However, this case is quite restrictive. We are interested here in the more general context of image-based 3D reconstruction, where such exact camera orientation information is typically not available. On the other hand, we often have a vague knowledge of the camera orientation in this context, and exploiting such knowledge is the main focus of this chapter. Lemma 4.3 provides a glimpse of how the location of Π_∞ is constrained

depending on the relative camera orientation. Propositions 4.6 and 4.7 present the complete picture. To reach there, we rely on a couple of intermediate results, including the following lemma.

Lemma 4.4. *For a camera pair (i, j), let $\mathcal{H}$ be the attached (not necessarily oriented) horopter with parametric representation $\mathcal{H}(s,t) = s^3\,\mathsf{C}_i - s^2t\,\mathsf{T}_{ij} + st^2\,\mathsf{T}_{ji} - t^3\,\mathsf{C}_j$, and let $\mathcal{H}_s$ and $\mathcal{H}_t$ be the hodographs of $\mathcal{H}$. Let Π be a plane in $\mathbb{P}^{*3}$ not containing the camera centers C_i and C_j, and consider:*

- *$\Pi^\top\mathcal{H}_s(s,t)$ is a quadratic polynomial in s whose coefficients depend on t and Π. Let $\Delta_s(\Pi)$ be its associated discriminant.*

- *$\Pi^\top\mathcal{H}_t(s,t)$ is a quadratic polynomial in t whose coefficients depend on s and Π. Let $\Delta_t(\Pi)$ be its associated discriminant.*

- *$\mathcal{S}_{ij}(\Pi)$ and $\mathcal{S}_{ji}(\Pi)$ are the Schur complements (see Appendix A.1) of $\Pi^\top\mathsf{C}_i$ and $\Pi^\top\mathsf{C}_j$, respectively, in the two matrices*

$$\begin{bmatrix} \Pi^\top\mathsf{C}_i & \Pi^\top\mathsf{T}_{ij} \\ \Pi^\top\mathsf{T}_{ij} & 3\Pi^\top\mathsf{T}_{ji} \end{bmatrix} \text{ and } \begin{bmatrix} \Pi^\top\mathsf{C}_j & \Pi^\top\mathsf{T}_{ji} \\ \Pi^\top\mathsf{T}_{ji} & 3\Pi^\top\mathsf{T}_{ij} \end{bmatrix}. \tag{4.10}$$

Then, the signs of $\Delta_s(\Pi)$ and $\Delta_t(\Pi)$ are given by:

$$\begin{aligned} \operatorname{sgn}(\Delta_s(\Pi)) &= -\operatorname{sgn}\big(\Pi^\top\mathsf{C}_i\big)\operatorname{sgn}(\mathcal{S}_{ij}(\Pi)), \\ \operatorname{sgn}(\Delta_t(\Pi)) &= -\operatorname{sgn}\big(\Pi^\top\mathsf{C}_j\big)\operatorname{sgn}(\mathcal{S}_{ji}(\Pi)). \end{aligned} \tag{4.11}$$

Proof. The polynomials $\Pi^\top\mathcal{H}_s(s,t)$ and $\Pi^\top\mathcal{H}_t(s,t)$ are expressed as:

$$\begin{aligned} \Pi^\top\mathcal{H}_s(s,t) &= 3s^2\Pi^\top\mathsf{C}_i - 2st\Pi^\top\mathsf{T}_{ij} + t^2\Pi^\top\mathsf{T}_{ji}, \\ \Pi^\top\mathcal{H}_t(s,t) &= -3t^2\Pi^\top\mathsf{C}_j + 2st\Pi^\top\mathsf{T}_{ji} - s^2\Pi^\top\mathsf{T}_{ij}. \end{aligned} \tag{4.12}$$

These polynomials are quadratic in s and t, respectively. The associated discriminants $\Delta_s(\Pi)$ and $\Delta_t(\Pi)$ correspond to:

$$\begin{aligned} \Delta_s(\Pi) &= -4t^2\big(3(\Pi^\top\mathsf{T}_{ji})(\Pi^\top\mathsf{C}_i) - (\Pi^\top\mathsf{T}_{ij})^2\big), \\ \Delta_t(\Pi) &= -4s^2\big(3(\Pi^\top\mathsf{T}_{ij})(\Pi^\top\mathsf{C}_j) - (\Pi^\top\mathsf{T}_{ji})^2\big). \end{aligned} \tag{4.13}$$

As $\mathit{\Pi}$ does not contain the camera centers C_i and C_j, we have that $\Pi^\top \mathsf{C}_i \neq 0$ and $\Pi^\top \mathsf{C}_j \neq 0$, and therefore the Schur complements $\mathcal{S}_{ij}(\Pi)$ and $\mathcal{S}_{ji}(\Pi)$ are defined as (see Lemma A.1):

$$\begin{aligned} \mathcal{S}_{ij}(\Pi) = \; & 3\Pi^\top \mathsf{T}_{ji} - (\Pi^\top \mathsf{T}_{ij})^2 (\Pi^\top \mathsf{C}_i)^{-1}, \\ \mathcal{S}_{ji}(\Pi) = \; & 3\Pi^\top \mathsf{T}_{ij} - (\Pi^\top \mathsf{T}_{ji})^2 (\Pi^\top \mathsf{C}_j)^{-1}. \end{aligned} \tag{4.14}$$

From (4.13) and (4.14), it can be observed that:

$$\begin{aligned} \Delta_s(\Pi) = \; & -4t^2 \, \Pi^\top \mathsf{C}_i \, \mathcal{S}_{ij}(\Pi), \\ \Delta_t(\Pi) = \; & -4s^2 \, \Pi^\top \mathsf{C}_j \, \mathcal{S}_{ji}(\Pi). \end{aligned} \tag{4.15}$$

Thus, the signs of these discriminants can be expressed as in (4.11). ■

The intersection of the considered plane $\mathit{\Pi}$ with the hodographs $\mathcal{H}_s$ and $\mathcal{H}_t$ is defined by the vanishing polynomials $\Pi^\top \mathcal{H}_s(s,t)$ and $\Pi^\top \mathcal{H}_t(s,t)$, respectively. To meaningfully characterize this intersection, we consider hereafter the hodographs associated to an oriented horopter (see Definition 4.5). With this representation, we consider that the discriminants $\Delta_s(\Pi)$ and $\Delta_t(\Pi)$ in Lemma 4.4 are negative for no real points of intersection, positive for two distinct real points of intersection, and zero for one repeated real point of intersection with each of the hodographs. Before focusing our attention on $\mathit{\Pi}_\infty$ and its geometric relationship with the hodographs, we first characterize more generally the set of QUARC planes that either do not intersect the hodographs or are, at most, tangent to each hodograph. In the following proposition, we show that this is a convex set defined by a pair of LMIs.

Proposition 4.5. *For a camera pair (i,j), let $\mathcal{H}$ be the attached oriented horopter with parametric representation $\mathcal{H}(s,t) = s^3 \, \mathsf{C}_i - s^2 t \, \mathsf{T}_{ij} + s t^2 \, \mathsf{T}_{ji} - t^3 \, \mathsf{C}_j$, and let $\mathcal{H}_s$ and $\mathcal{H}_t$ be the hodographs of $\mathcal{H}$. Let $\mathit{\Pi}$ be a plane in $\mathbb{P}^{*3}$ such that $\Pi^\top \mathsf{C}_i > 0$ and $\Pi^\top \mathsf{C}_j > 0$. The following statements are equivalent:*

- *(i) The hodograph $\mathcal{H}_s$, with respect to $\mathit{\Pi}$, lies on the same side as the camera center C_i, while the hodograph $\mathcal{H}_t$ lies on the opposite side, with each hodograph intersecting $\mathit{\Pi}$ in at most one real point.*
- *(ii) $\mathit{\Pi}$ belongs to a convex set $\mathcal{A}_{ij}$ that is defined as,*

$$\mathcal{A}_{ij} = \left\{ \mathit{\Pi} \in \mathbb{P}^{*3} : \begin{bmatrix} \Pi^\top \mathsf{C}_i & \Pi^\top \mathsf{T}_{ij} \\ \Pi^\top \mathsf{T}_{ij} & 3\Pi^\top \mathsf{T}_{ji} \end{bmatrix} \succeq 0, \quad \begin{bmatrix} \Pi^\top \mathsf{C}_j & \Pi^\top \mathsf{T}_{ji} \\ \Pi^\top \mathsf{T}_{ji} & 3\Pi^\top \mathsf{T}_{ij} \end{bmatrix} \succeq 0 \right\}. \quad (4.16)$$

Proof. The proof relies on the Schur complement lemma (Lemma A.1). We first prove the assertion $(i) \Rightarrow (ii)$. From statement (i), we have that, with respect to $\mathit{\Pi}$, the hodograph $\mathcal{H}_s$ lies on the same side as the camera center C_i, while $\mathcal{H}_t$ lies on the opposite side, with each hodograph intersecting $\mathit{\Pi}$ in at most one real point. Since $\mathcal{H}_s$ and $\mathcal{H}_t$ each intersect $\mathit{\Pi}$ in at most one real point, the discriminants $\Delta_s(\Pi)$ and $\Delta_t(\Pi)$ given in Lemma 4.4 are nonpositive. As $\Pi^\top \mathsf{C}_i > 0$ and $\Pi^\top \mathsf{C}_j > 0$, we can deduce from Lemma 4.4 that $\mathcal{S}_{ij}(\Pi) \geq 0$ and $\mathcal{S}_{ji}(\Pi) \geq 0$ for $\Delta_s(\Pi)$ and $\Delta_t(\Pi)$ to be nonpositive. From Lemma A.1, since $\Pi^\top \mathsf{C}_i > 0$ and $\Pi^\top \mathsf{C}_j > 0$, the LMIs in (4.16) hold if and only if $\mathcal{S}_{ij}(\Pi)$ and $\mathcal{S}_{ji}(\Pi)$ are nonnegative. Thus, $\mathit{\Pi}$ satisfies the LMIs in (4.16) and thereby belongs to the set $\mathcal{A}_{ij}$. This proves the forward implication $(i) \Rightarrow (ii)$.

We now prove the reverse implication. From statement (ii), we have that $\mathit{\Pi}$ belongs to the set $\mathcal{A}_{ij}$ defined by the two LMIs in (4.16). From Lemma A.1, these LMIs imply that $\mathcal{S}_{ij}(\Pi) \geq 0$ and $\mathcal{S}_{ji}(\Pi) \geq 0$ since $\Pi^\top \mathsf{C}_i > 0$ and $\Pi^\top \mathsf{C}_j > 0$. We can deduce from Lemma 4.4 that $\Delta_s(\Pi)$ and $\Delta_t(\Pi)$ are nonpositive. Thus, $\mathit{\Pi}$ intersects each of the hodographs $\mathcal{H}_s$ and $\mathcal{H}_t$ in at most one real point. Now, from (4.7), we observe that the camera center C_i (with coordinates C_i) lies on $\mathcal{H}_s$, whereas the point with coordinates $-\mathsf{C}_j$ lies on $\mathcal{H}_t$. As $\Pi^\top \mathsf{C}_i > 0$ and $\Pi^\top \mathsf{C}_j > 0$, and $\mathit{\Pi}$ is at most tangent to each hodograph, it can be deduced that $\mathcal{H}_s$ is on the same side as C_i with respect to $\mathit{\Pi}$, while $\mathcal{H}_t$ is on the opposite side. This proves the reverse implication $(ii) \Rightarrow (i)$. ■

We observe from Lemma 4.4 that the discriminants $\Delta_s(\Pi)$ and $\Delta_t(\Pi)$ being nonpositive, *i.e.* $\mathit{\Pi}$ meeting the hodographs in at most one real point, implies a pair of quadratic constraints on $\mathit{\Pi}$. Proposition 4.5 shows that these constraints are, in fact, convex as they can be formulated as LMIs (see Appendix A.1). Now, within the set of QUARC planes, our main interest is in $\mathit{\Pi}_\infty$, and we aim to characterize its geometric relationship with the hodographs. In the following proposition, we show that the intersection of $\mathit{\Pi}_\infty$ with the hodographs depends on the relative camera orientation angle and that $\mathit{\Pi}_\infty$ is constrained to the set $\mathcal{A}_{ij}$ when $|\theta_{ij}| \leq 120^\circ$.

Proposition 4.6. *For a camera pair (i, j), let θ_{ij} be the relative camera orientation angle. Let $\mathcal{H}$ be the oriented horopter attached to the camera pair, and let $\mathcal{H}_s$ and $\mathcal{H}_t$ be the hodographs of $\mathcal{H}$. If $|\theta_{ij}| \leq 120°$, the plane at infinity $\mathit{\Pi}_\infty \in \mathcal{A}_{ij}$.*

Proof. We prove here that $\mathit{\Pi}_\infty$ satisfies the first statement in Proposition 4.5 when $|\theta_{ij}| \leq 120°$. The result that $\mathit{\Pi}_\infty \in \mathcal{A}_{ij}$ then follows from the equivalence of the two statements. For an oriented horopter, recall that the camera centers C_i and C_j lie on the same side with respect to $\mathit{\Pi}_\infty$. To be consistent with Proposition 4.5, we consider a homogeneous representation of $\mathit{\Pi}_\infty$ such that $\Pi_\infty^\top \mathsf{C}_i > 0$ and $\Pi_\infty^\top \mathsf{C}_j > 0$. The intersection of $\mathit{\Pi}_\infty$ with the hodographs $\mathcal{H}_s$ and $\mathcal{H}_t$ is represented by the vanishing polynomials $\Pi_\infty^\top \mathcal{H}_s(s,t)$ and $\Pi_\infty^\top \mathcal{H}_t(s,t)$, respectively, and characterized by the signs of their associated discriminants $\Delta_s(\Pi_\infty)$ and $\Delta_t(\Pi_\infty)$. In particular, $\Delta_s(\Pi_\infty)$ and $\Delta_t(\Pi_\infty)$ ought to be nonpositive if the intersection with each hodograph is in at most one real point. To prove that these discriminants are so when $|\theta_{ij}| \leq 120°$, consider the expressions in (4.11). By substituting the values from Property 4.2 in these expressions, we have that:

$$\begin{aligned} \operatorname{sgn}(\Delta_s(\Pi_\infty)) &= -\operatorname{sgn}(\mathsf{a}_{ij}(3 - \mathsf{a}_{ij})), \\ \operatorname{sgn}(\Delta_t(\Pi_\infty)) &= -\operatorname{sgn}(\mathsf{a}_{ij}(3 - \mathsf{a}_{ij})), \end{aligned} \tag{4.17}$$

where $\mathsf{a}_{ij} = 1 + 2\cos\theta_{ij}$. The signs of $\Delta_s(\Pi_\infty)$ and $\Delta_t(\Pi_\infty)$ depend on a_{ij} and thus on θ_{ij}. From the values that a_{ij} can take in (4.8), we can deduce that:

$$\Delta_s(\Pi_\infty) \leq 0 \;\text{ and }\; \Delta_t(\Pi_\infty) \leq 0 \;\text{ if }\; |\theta_{ij}| \leq 120°, \tag{4.18a}$$

$$\Delta_s(\Pi_\infty) \geq 0 \;\text{ and }\; \Delta_t(\Pi_\infty) \geq 0 \;\text{ if }\; |\theta_{ij}| \geq 120°, \tag{4.18b}$$

where $\Delta_s(\Pi_\infty) = 0$ and $\Delta_t(\Pi_\infty) = 0$ for $|\theta_{ij}| \in \{0°, 120°\}$ since $\mathsf{a}_{ij} = 0$ if $|\theta_{ij}| = 120°$ and $\mathsf{a}_{ij} = 3$ if $\theta_{ij} = 0°$. Thus, $\mathit{\Pi}_\infty$ intersects each of the hodographs $\mathcal{H}_s$ and $\mathcal{H}_t$ in at most one real point when $|\theta_{ij}| \leq 120°$. This proves one part of statement (i) in Proposition 4.5. The other part is that $\mathcal{H}_s$ is on the same side as C_i with respect to $\mathit{\Pi}_\infty$, while $\mathcal{H}_t$ is on the opposite side. From (4.7), we observe that $\mathcal{H}_s$ contains C_i and $\mathcal{H}_t$ contains the point with coordinates $-\mathsf{C}_j$. As $\Pi_\infty^\top \mathsf{C}_i > 0$ and $\Pi_\infty^\top \mathsf{C}_j > 0$, and $\mathit{\Pi}_\infty$ is at most tangent to each hodograph, it can be deduced that the

first statement in Proposition 4.5 is satisfied. Hence, it follows that $\mathit{\Pi}_\infty \in \mathcal{A}_{ij}$ when $|\theta_{ij}| \leq 120°$. ■

From the proof of Proposition 4.6, we observe that when $|\theta_{ij}| \leq 120°$, $\mathit{\Pi}_\infty$ either does not intersect the hodographs in real points or it is simultaneously tangent to both of them. Note that the tangency occurs when $|\theta_{ij}| \in \{0°, 120°\}$, the points of tangency being the control points of the horopter when $|\theta_{ij}| = 120°$ (see Lemma 4.3). Thus, Proposition 4.6 extends the result in Lemma 4.3 on the incidence of $\mathit{\Pi}_\infty$ with the control points to the hodographs for the case of $|\theta_{ij}| \leq 120°$. The corresponding constraints so imposed on $\mathit{\Pi}_\infty$ are formulated as linear inequalities in Lemma 4.3, whereas they are formulated as LMIs in Proposition 4.6. We consider now the case of $|\theta_{ij}| \geq 120°$. In the following proposition, we show that $\mathit{\Pi}_\infty$ is constrained in this case to a convex set $\mathcal{A}'_{ij}$ that is a counterpart of $\mathcal{A}_{ij}$.

Proposition 4.7. *For a camera pair (i, j), let θ_{ij} be the relative camera orientation angle. Let $\mathcal{H}$ be the attached oriented horopter with parametric representation $\mathcal{H}(s,t) = s^3\,\mathsf{C}_i - s^2t\,\mathsf{T}_{ij} + st^2\,\mathsf{T}_{ji} - t^3\,\mathsf{C}_j$, and let $\mathcal{H}_s$ and $\mathcal{H}_t$ be the hodographs of $\mathcal{H}$. Let $\mathit{\Pi}$ be a plane in $\mathbb{P}^{*3}$ such that $\Pi^\top \mathsf{C}_i > 0$ and $\Pi^\top \mathsf{C}_j > 0$, and consider the convex set $\mathcal{A}'_{ij}$ defined as:*

$$\mathcal{A}'_{ij} = \left\{ \mathit{\Pi} \in \mathbb{P}^{*3} : \begin{bmatrix} \Pi^\top \mathsf{C}_i & \Pi^\top \mathsf{T}_{ij} \\ \Pi^\top \mathsf{T}_{ij} & -\Pi^\top \mathsf{T}_{ji} \end{bmatrix} \succeq 0, \quad \begin{bmatrix} \Pi^\top \mathsf{C}_j & \Pi^\top \mathsf{T}_{ji} \\ \Pi^\top \mathsf{T}_{ji} & -\Pi^\top \mathsf{T}_{ij} \end{bmatrix} \succeq 0 \right\}. \quad (4.19)$$

If $|\theta_{ij}| \geq 120°$, the plane at infinity $\mathit{\Pi}_\infty$ intersects each of the hodographs $\mathcal{H}_s$ and $\mathcal{H}_t$ in at least one real point and $\mathit{\Pi}_\infty \in \mathcal{A}'_{ij}$.

Proof. First we prove that $\mathit{\Pi}_\infty$ intersects each of the hodographs $\mathcal{H}_s$ and $\mathcal{H}_t$ in at least one real point when $|\theta_{ij}| \geq 120°$. Note that the proof of Proposition 4.6 already showed this result in (4.18b). We provide an alternative reasoning here that adds to the subsequent discussion on the geometric constraint imposed by the LMIs in (4.19). As usual, we consider the homogeneous representation of $\mathit{\Pi}_\infty$ such that its incidence with the camera centers C_i and C_j is $\Pi_\infty^\top \mathsf{C}_i > 0$ and $\Pi_\infty^\top \mathsf{C}_j > 0$. From Lemma 4.3, we have that the camera centers C_i and C_j and the control points T_{ij} and T_{ji} of the horopter $\mathcal{H}$ lie on opposite sides with respect to $\mathit{\Pi}_\infty$ when $|\theta_{ij}| > 120°$. As a result, $\mathit{\Pi}_\infty$ does not satisfy the LMIs in Proposition 4.5 in this

case and thereby does not belong to the set $\mathcal{A}_{ij}$. From the equivalent statements in Proposition 4.5, we thus have that $\mathit{\Pi}_\infty$ intersects each of the hodographs $\mathcal{H}_s$ and $\mathcal{H}_t$ in two real points when $|\theta_{ij}| > 120^\circ$. It can also be easily verified from Lemmas 4.3 and 4.4 that $\mathit{\Pi}_\infty$ is simultaneously tangent to both the hodographs when $|\theta_{ij}| = 120^\circ$. Hence, $\mathit{\Pi}_\infty$ intersects each of the hodographs in at least one real point when $|\theta_{ij}| \geq 120^\circ$.

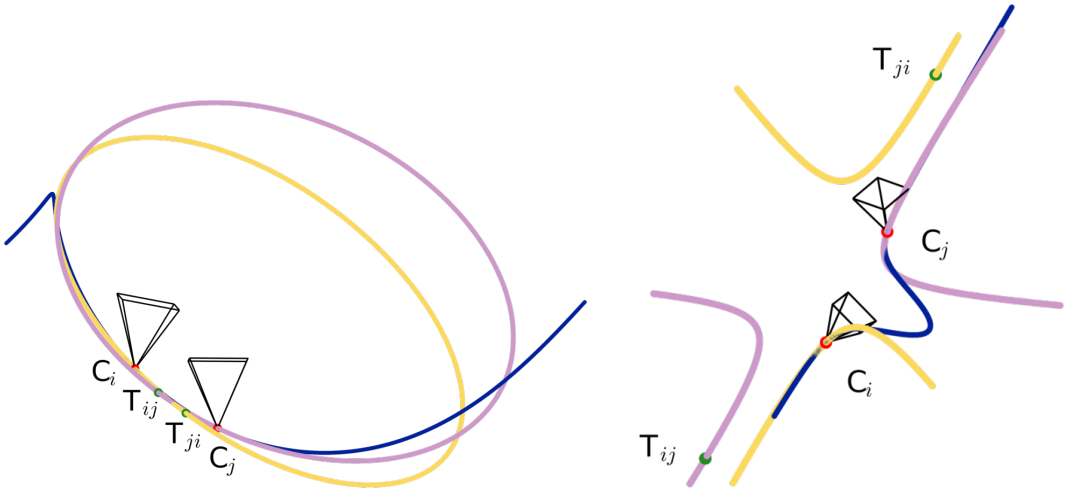
We now prove that $\mathit{\Pi}_\infty$ is constrained more stringently to the set $\mathcal{A}'_{ij}$. As $\mathit{\Pi}_\infty$ intersects each of the hodographs in at least one real point when $|\theta_{ij}| \geq 120^\circ$, the discriminants $\Delta_s(\Pi_\infty)$ and $\Delta_t(\Pi_\infty)$ in Lemma 4.4 are nonnegative. Since we considered the sign of $\mathit{\Pi}_\infty$ such that $\Pi_\infty^\top \mathsf{C}_i > 0$ and $\Pi_\infty^\top \mathsf{C}_j > 0$, the functions $\mathcal{S}_{ij}(\Pi_\infty)$ and $\mathcal{S}_{ji}(\Pi_\infty)$ are nonpositive. Recall that $\mathcal{S}_{ij}(\Pi_\infty)$ and $\mathcal{S}_{ji}(\Pi_\infty)$, defined in (4.14), are the Schur complements of $\Pi_\infty^\top \mathsf{C}_i$ and $\Pi_\infty^\top \mathsf{C}_j$, respectively, in the first and second matrix in (4.10). Also, these two matrices appear in the LMIs in Proposition 4.5. Substituting the values from Property 4.2 in (4.14), $\mathcal{S}_{ij}(\Pi_\infty)$ and $\mathcal{S}_{ji}(\Pi_\infty)$ can be expressed as:

$$\begin{aligned} \mathcal{S}_{ij}(\Pi_\infty) &= \Pi_\infty^\top \mathsf{T}_{ji}(3 - \mathsf{a}_{ij}), \\ \mathcal{S}_{ji}(\Pi_\infty) &= \Pi_\infty^\top \mathsf{T}_{ij}(3 - \mathsf{a}_{ij}), \end{aligned} \tag{4.20}$$

where $\mathsf{a}_{ij} = 1 + 2\cos\theta_{ij}$. From Lemma 4.3, we have that $\Pi_\infty^\top \mathsf{T}_{ij} \leq 0$ and $\Pi_\infty^\top \mathsf{T}_{ji} \leq 0$ when $|\theta_{ij}| \geq 120^\circ$. Moreover, from (4.8b), a_{ij} is in the interval $-1 \leq \mathsf{a}_{ij} \leq 0$ when $|\theta_{ij}| \geq 120^\circ$. Since $\mathcal{S}_{ij}(\Pi_\infty)$ and $\mathcal{S}_{ji}(\Pi_\infty)$ are nonpositive and a_{ij} is bounded, we observe that $\mathcal{S}_{ij}(\Pi_\infty)$ and $\mathcal{S}_{ji}(\Pi_\infty)$ are bounded from below as:

$$\begin{aligned} \mathcal{S}_{ij}(\Pi_\infty) - 4\Pi_\infty^\top \mathsf{T}_{ji} \geq 0, \\ \mathcal{S}_{ji}(\Pi_\infty) - 4\Pi_\infty^\top \mathsf{T}_{ij} \geq 0. \end{aligned} \tag{4.21}$$

Note that the left-hand sides of the two inequalities in (4.21) are the Schur complements of $\Pi_\infty^\top \mathsf{C}_i$ and $\Pi_\infty^\top \mathsf{C}_j$, respectively, in the first and second matrix in (4.19). From the Schur complement lemma, since $\Pi_\infty^\top \mathsf{C}_i > 0$ and $\Pi_\infty^\top \mathsf{C}_j > 0$, the LMIs in (4.19) hold for $\mathit{\Pi}_\infty$ if and only if these Schur complements are nonnegative, *i.e.* the inequalities in (4.21) are satisfied. Thus, $\mathit{\Pi}_\infty$ satisfies the LMIs in (4.19) and thereby belongs to the set $\mathcal{A}'_{ij}$ when $|\theta_{ij}| \geq 120^\circ$. ■



(a) $0° < |\theta_{ij}| < 120°$ (b) $|\theta_{ij}| > 120°$

Figure 4.4: The hodographs of the horopter and Π_∞. The horopter curve is shown in blue and its hodographs in yellow and purple. (a) When $0° < |\theta_{ij}| < 120°$, Π_∞ does not intersect the hodographs in real points. They are therefore ellipses in this case. (b) When $|\theta_{ij}| > 120°$, Π_∞ cuts through each of the hodographs in a defined region. They are therefore hyperbolas in this case. (Not shown: when $\theta_{ij} \in \{0°, 120°\}$, Π_∞ is simultaneously tangent to both hodographs, which are therefore parabolas in this case.)

From the reasoning in the proof of Proposition 4.7, the linear inequalities in (4.9b) alone ensure that Π_∞ intersects each of the hodographs in at least one real point when $|\theta_{ij}| \geq 120°$. The LMIs in Proposition 4.7 further characterize the region of this intersection. To see this, first note that the region of the intersection depends on the discriminants $\Delta_s(\Pi_\infty)$ and $\Delta_s(\Pi_\infty)$ and thereby on the functions $\mathcal{S}_{ij}(\Pi_\infty)$ and $\mathcal{S}_{ji}(\Pi_\infty)$ in Lemma 4.4. These functions, and thereby the discriminants, are constrained by the inequalities in (4.21) and hence by the LMIs in (4.19). Thus, the convex set $\mathcal{A}'_{ij}$ delineates a region where Π_∞ intersects each of the hodographs when $|\theta_{ij}| \geq 120°$. Proposition 4.7 then extends the result in Lemma 4.3 for $|\theta_{ij}| \geq 120°$. This completes the set of orientation-based LMI constraints on Π_∞.

To summarize our findings in this section, we showed that the relative positioning of Π_∞ and the hodographs $\mathcal{H}_s$ and $\mathcal{H}_t$ is characterized by the relative camera orientation angle. In particular, Π_∞ is constrained to one of two convex sets depending on whether $|\theta_{ij}| \leq 120°$ or $|\theta_{ij}| \geq 120°$. When $|\theta_{ij}| \leq 120°$, Π_∞ partitions the hodographs such that $\mathcal{H}_s$ lies on the same side as the camera centers with re-

spect to Π_∞ while $\mathcal{H}_t$ lies on the opposite side, with Π_∞ intersecting each of the hodographs in at most one real point. This implies that it belongs to the set $\mathcal{A}_{ij}$. When $|\theta_{ij}| \geq 120^\circ$, on the other hand, Π_∞ intersects each of the hodographs in at least one real point and its location is constrained to a region defined by the set $\mathcal{A}'_{ij}$. Each of these convex sets, $\mathcal{A}_{ij}$ and $\mathcal{A}'_{ij}$, is defined by a pair of LMI constraints on Π_∞. Finally, Π_∞ is simultaneously tangent to both the hodographs when $|\theta_{ij}| \in \{0^\circ, 120^\circ\}$. The hodographs thus act as virtual positioning objects for Π_∞. Figure 4.4 shows a simulated camera pair and its associated hodographs to illustrate this incidence relationship.

The orientation-based LMIs are new necessary conditions for a given plane to be Π_∞. They are complementary to existing constraints on Π_∞, such as the modulus constraint [Pollefeys and Van Gool 1999] and the chirality conditions [Hartley and Zisserman 2004]. It is worth delving a little deeper into the algebraic constraint imposed by these LMIs to appreciate their role. We delay this discussion till Section 4.3.3, where we consider the constraint imposed by these LMIs on the inter-image homography. From a computational viewpoint, the appeal of these new constraints is that they are convex and formulated as LMIs. LMI problems in convex optimization can be solved reliably and efficiently using interior-point methods [Boyd et al. 1994; Boyd and Vandenberghe 2004]. In the following section, we show how a vague knowledge of the relative orientation angle between camera pairs, *i.e.* $|\theta_{ij}| \leq 120^\circ$ or $|\theta_{ij}| \geq 120^\circ$, can be exploited with these LMI constraints to obtain a new quasi-affine reconstruction of a scene: a QUARCH.

4.2.5 QUARCH: A new quasi-affine reconstruction stratum

The orientation-based LMI constraints on Π_∞ presented in the previous section allow us to define a new quasi-affine reconstruction of a scene that we refer to as QUARCH: QUasi-Affine Reconstruction with respect to Camera centers and the Hodographs of horopters. A QUARCH is a specialization of a QUARC that is additionally quasi-affine with respect to the hodographs of a set of camera pairs. It is thus one step closer to the sought affine and metric reconstructions, as illustrated in Figure 4.5. A QUARCH can be recovered from a projective reconstruction when the camera intrinsic parameters are constant and a vague knowledge of the relative

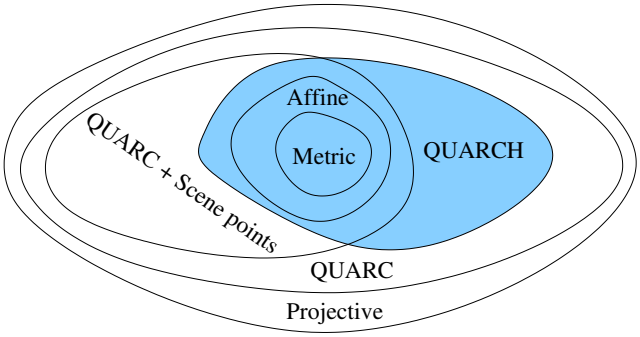


Figure 4.5: QUARCH in the 3D reconstruction hierarchy. QUARCH is a specialization of the QUARC stratum and is therefore one step closer to the affine stratum.

camera orientation angle, *i.e.* either $|\theta_{ij}| \leq 120^\circ$ or $|\theta_{ij}| \geq 120^\circ$, is available for a set of camera pairs. Upgrading a projective reconstruction to a QUARCH involves the same steps outlined for a QUARC in Section 2.2.6 except for locating a QUARCH plane instead in the second step. For ease of presentation, we split the definition of a QUARCH plane for the two cases of $|\theta_{ij}| \leq 120^\circ$ and $|\theta_{ij}| \geq 120^\circ$.

Definition 4.6 (QUARCH plane, $|\theta_{ij}| \leq 120^\circ$)**.** *For a camera pair (i, j) with relative orientation angle $|\theta_{ij}| \leq 120^\circ$ and attached oriented horopter $\mathcal{H}$, a QUARCH plane Π_Q with respect to the camera pair is a QUARC plane that partitions the hodographs $\mathcal{H}_s$ and $\mathcal{H}_t$ of $\mathcal{H}$ such that $\mathcal{H}_s$ lies on the same side as the camera centers while $\mathcal{H}_t$ lies on the opposite side, with Π_Q being at most tangent to each hodograph. Equivalently, it is a QUARC plane such that $\Pi_\mathsf{Q} \in \mathcal{A}_{ij}$.*

Definition 4.7 (QUARCH plane, $|\theta_{ij}| \geq 120^\circ$)**.** *For a camera pair (i, j) with relative orientation angle $|\theta_{ij}| \geq 120^\circ$ and attached oriented horopter $\mathcal{H}$, a QUARCH plane Π_Q with respect to the camera pair is a QUARC plane that either intersects the hodographs $\mathcal{H}_s$ and $\mathcal{H}_t$ of $\mathcal{H}$ in a region defined by $\mathcal{A}'_{ij}$ or is tangent to each hodograph.*

We also define a strict QUARCH plane as follows.

Definition 4.8 (Strict QUARCH plane)**.** *A strict QUARCH plane is a QUARCH plane that satisfies the strict form of the LMIs in Propositions 4.5 and 4.7.*

Compared to a QUARC plane, locating a QUARCH plane additionally requires the camera intrinsic parameters to be constant as well as a knowledge of either

$|\theta_{ij}| \leq 120^\circ$ or $|\theta_{ij}| \geq 120^\circ$ for a set of camera pairs. Mapping this plane to the canonical position with coordinates $(\mathbf{0}_3^\top, 1)^\top$ upgrades the projective reconstruction to a QUARCH, which we formally define as follows.

Definition 4.9 (QUARCH)**.** *A QUARCH is a projective reconstruction with sign-corrected projection matrices where the canonical plane is a QUARCH plane with respect to a set of camera pairs.*

The remaining detail is the computation of a QUARCH plane. As the involved constraints are convex and formulated as LMIs, it can be computed by solving an SDP problem. The benefit of an SDP formulation is that it can be efficiently solved using interior-point methods that are conveniently implemented in off-the-shelf solvers (see Appendix A.2). Assuming that a vague knowledge of the relative orientation angle, *i.e*. either $|\theta_{ij}| \leq 120^\circ$ or $|\theta_{ij}| \geq 120^\circ$, is available for all camera pairs, the following SDP problem can be solved to obtain a QUARCH plane:

$$
\begin{aligned}
\max_{\Pi,\,\delta} \quad & \delta \\
\text{s.t.} \quad & \frac{\Pi^\top \mathsf{C}_l}{\|\mathsf{C}_l\|} > \delta, && l = 1, \ldots, n, \\
& -1 \leq (\Pi)_k \leq 1, && k = 1, \ldots, 4, \\
& \mathit{\Pi} \in \mathcal{A}_{ij}, && \forall (i,j) : |\theta_{ij}| \leq 120^\circ, \\
& \mathit{\Pi} \in \mathcal{A}'_{ij}, && \forall (i,j) : |\theta_{ij}| \geq 120^\circ, \\
& && i = 1, \ldots, n-1, \\
& && j = i+1, \ldots, n.
\end{aligned}
\tag{4.22}
$$

Note that without the orientation-based constraints, the SDP problem (4.22) reduces to the QUARC LP problem (2.25). A QUARCH may be extended to preserve the convex hull of the set of scene points by simply augmenting problem (4.22) with the corresponding linear inequalities for the scene points.

We use a QUARCH in a camera autocalibration algorithm as an initial step towards recovering a metric reconstruction from a projective one. To compute a QUARCH, the SDP problem in (4.22) assumes that a vague knowledge of the relative orientation angle is available for all camera pairs. Such knowledge may be available, for instance, using an external sensor (*e.g*. a gyroscope). While the mea-

surements provided by such a sensor may be too noisy to be used as exact orientation information, they can instead be exploited to automatically assign the appropriate LMI constraints for each camera pair. However, such knowledge is not always available in the image-based 3D reconstruction scenario. In the proposed algorithm, we instead assume that $|\theta_{ij}| \leq 120^\circ$ between consecutive views of an ordered image sequence. As discussed earlier, this mild assumption is often implicitly satisfied when capturing images for 3D reconstruction. We propose a dedicated SDP problem to compute a QUARCH under this assumption. We present this SDP problem and our autocalibration algorithm in Section 4.4. In the following section, we investigate if a tighter yet still vague knowledge of the relative camera orientation angle allows us to further constrain the location of Π_∞.

4.3 Orientation-based non-convex constraints

In the previous section, we derived a new set of camera orientation-based constraints on Π_∞ from its geometric relationship with the hodographs. These constraints are convex, formulated as LMIs, and can be exploited when a vague knowledge of the relative camera orientation angle, *i.e.* either $|\theta_{ij}| \leq 120^\circ$ or $|\theta_{ij}| \geq 120^\circ$, is available for a set of camera pairs. Using these constraints, we defined the QUARCH stratum and indicated that we can safely assume $|\theta_{ij}| \leq 120^\circ$ between consecutive views of an ordered sequence to obtain a QUARCH in practice. A natural question to ask now is if a tighter knowledge of the orientation angle, though still vague, can be exploited to further constrain the location of Π_∞. Besides the theoretical interest, this question stems from practice. Images for 3D reconstruction are typically captured such that the relative orientation angle is considerably smaller than 120°, at least between consecutive views, to ensure sufficient overlap between them for an effective feature matching.

In this section, we show that a tighter knowledge of the relative camera orientation angle can indeed be exploited to constrain Π_∞. Specifically, we derive new inequality constraints on Π_∞ from a vague knowledge of the relative camera orientation angle being $|\theta_{ij}| < 90^\circ/k$, where k is a positive integer. These constraints are non-convex and formulated as polynomials of degree $2k$. Hence, polynomial inequalities of an increasingly higher degree are obtained with tighter orientation

knowledge. We derive these constraints using the *Hurwitz stability criterion*, a standard tool in control theory for stability analysis. Unlike the LMI constraints, which have a geometric interpretation through the hodographs, the development of the non-convex constraints is purely algebraic and relies on inter-image homographies alone. We further show that the LMI constraints for $|\theta_{ij}| \leq 120^\circ$ can also be derived by using the Hurwitz stability criterion. Finally, we show that the LMIs and the non-convex inequalities can be used together to recover a specialized QUARCH.

In Section 4.3.1, we state the Hurwitz stability criterion. In Section 4.3.2, we show how this criterion can be used to derive a new set of orientation-based constraints on $\mathit{\Pi}_\infty$ for $|\theta_{ij}| < 90^\circ/k$ (where k is a positive integer). In Section 4.3.3, we then show the link with the LMIs for $|\theta_{ij}| \leq 120^\circ$ and present a specialized QUARCH leveraging the LMIs and the non-convex polynomial constraints.

4.3.1 Hurwitz stability criterion

In control system theory, the stability of a linear time-invariant dynamical system can be determined by analyzing the roots of its characteristic polynomial. In particular, all the roots of the characteristic polynomial of a stable system have negative real parts [Golnaraghi and Kuo 2009, Sec. 2.10]. In this context, a stable polynomial is defined as follows.

Definition 4.10 (Stable polynomial)**.** *A real polynomial (i.e. a polynomial with real coefficients)* $p(\lambda) = \sum_{k=0}^{n} a_k \lambda^k$ *is stable if all of its roots have negative real parts.*

Several algebraic tests have been devised in the control literature to determine if the characteristic polynomial of a given system is stable without explicitly computing its roots. One such test was proposed by Hurwitz, now referred to as the Hurwitz stability criterion, that gives a necessary and sufficient condition for a real polynomial p of degree n to be stable [Hurwitz 1895]. This condition involves the leading

principal minors of the so-called *Hurwitz matrix* M of p, where M is a square matrix of size n that is constructed using the coefficients of p as:

$$\mathsf{M} = \begin{bmatrix} a_{n-1} & a_{n-3} & a_{n-5} & \cdots \\ a_n & a_{n-2} & a_{n-4} & \cdots \\ 0 & a_{n-1} & a_{n-3} & \cdots \\ 0 & a_n & a_{n-2} & \cdots \\ 0 & 0 & a_{n-1} & \cdots \\ 0 & 0 & a_n & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \tag{4.23}$$

with $a_{n-k} = 0$ if $k > n$. The leading principal minors of M,

$$\Delta_1 = a_{n-1}, \quad \Delta_2 = \det\begin{pmatrix} a_{n-1} & a_{n-3} \\ a_n & a_{n-2} \end{pmatrix}, \quad \ldots, \quad \Delta_n = \det(\mathsf{M}), \tag{4.24}$$

are referred to as the *Hurwitz determinants*. We are now ready to state the Hurwitz stability criterion.

Proposition 4.8 (Hurwitz stability criterion [Lancaster and Tismenetsky 1985, Sec. 13.9])**.** *A real polynomial $p(\lambda) = \sum_{k=0}^{n} a_k \lambda^k$ of degree n with $a_n > 0$ is stable if and only if the leading principal minors of the Hurwitz matrix of p are positive.*

Determining if a given real polynomial is stable thereby reduces to verifying if the inequalities given by the criterion in Proposition 4.8 are satisfied. We conclude this overview of the Hurwitz stability criterion by defining the notion of stability for square matrices.

Definition 4.11 (Stable matrix)**.** *A square matrix is stable if its characteristic polynomial is stable. Equivalently, it is stable if the real parts of all of its eigenvalues are negative.*

4.3.2 Stability-based constraints on infinity

We use the Hurwitz stability criterion to derive a new set of orientation-based constraints on Π_∞. These constraints exploit a vague knowledge of the relative camera

orientation angle being $|\theta_{ij}| < 90^\circ/k$, where k is a positive integer. As with the LMI constraints presented in Section 4.2.4, they are derived under the assumption of a moving camera with constant intrinsic parameters. Unlike those LMIs, however, the stability-based constraints are non-convex and formulated as polynomial inequalities of degree $2k$. We thus obtain polynomials of an increasingly higher degree with tighter orientation knowledge.

To formulate the new constraints, we analyze the stability of the inter-image homography matrix $\mathsf{H}_{\infty ij}$ induced by $\mathit{\Pi}_\infty$. In particular, we show that when $|\theta_{ij}| < 90^\circ/k$, the matrix $-\mathsf{H}_{\infty ij}^k$ is stable. Using the Hurwitz criterion, we can derive necessary and sufficient conditions for this matrix to be stable. The resulting inequalities form the new set of orientation-based constraints on $\mathit{\Pi}_\infty$ for $|\theta_{ij}| < 90^\circ/k$. For clarity and ease of presentation, we first state the result for $k = 1$, *i.e.* using $\mathsf{H}_{\infty ij}$ (Proposition 4.11) and then generalize it to higher powers, *i.e.* using $\mathsf{H}_{\infty ij}^k$ (Proposition 4.12). In the following proposition, we use the Hurwitz criterion to derive necessary and sufficient conditions for an arbitrary inter-image homography matrix to be stable. We then generalize this result to its kth power in Proposition 4.10.

Proposition 4.9. *For a camera pair (i, j), let P_i and P_j be the sign-corrected projection matrices. Let $\mathit{\Pi}$ be a plane in $\mathbb{P}^{*3}$ with coordinates $\Pi = \left(\pi^\top, 1\right)^\top$ and H_{ij} be its induced inter-image homography, where $\mathsf{H}_{1i} = \mathsf{P}_i \begin{bmatrix} \mathsf{I} & -\pi \end{bmatrix}^\top$ and $\mathsf{H}_{ij} = \mathsf{H}_{1j}\mathsf{H}_{1i}^*$. Provided that $\det(\mathsf{H}_{1i}) > 0$, the matrix $-\mathsf{H}_{ij}$ is stable if and only if:*

$$\begin{aligned} &\det(\mathsf{H}_{1j}) > 0, \quad \operatorname{tr}(\mathsf{H}_{ij}) > 0, \\ &\operatorname{tr}(\mathsf{H}_{ij}) \operatorname{tr}(\mathsf{H}_{ji}) - \det(\mathsf{H}_{1i}) \det(\mathsf{H}_{1j}) > 0. \end{aligned} \tag{4.25}$$

Proof. From Definition 4.11, the matrix $-\mathsf{H}_{ij}$ is stable if its characteristic polynomial is stable. Its characteristic equation,

$$\det(\lambda \mathsf{I} + \mathsf{H}_{ij}) = 0, \tag{4.26}$$

is cubic in λ. Multiplying both sides of (4.26) by $\det(\mathsf{H}_{1i})$, we obtain,

$$\det(\lambda \mathsf{H}_{1i} + \mathsf{H}_{1j} \det(\mathsf{H}_{1i})) = 0. \tag{4.27}$$

Provided that $\det(\mathsf{H}_{1i})$ is positive, the stability of the characteristic polynomial in (4.26) is not affected by this multiplication. The Hurwitz stability criterion gives necessary and sufficient conditions for a real polynomial to be stable. By applying this criterion to the polynomial in (4.27), we obtain the inequalities (4.25). ■

Recall from (2.5) that H_{ij} is linear in the coordinates of $\mathit{\Pi}$ for any i and j, and so are $\det(\mathsf{H}_{1i})$ and $\mathrm{tr}(\mathsf{H}_{ij})$. Hence, Proposition 4.9 provides one quadratic and a set of linear inequalities in $\mathit{\Pi}$ as necessary and sufficient conditions for $-\mathsf{H}_{ij}$ to be stable. Note that $\mathrm{tr}(\mathsf{H}_{ji}) > 0$ is implied by these inequalities. In the following proposition, we consider the stability of the powers of H_{ij}.

Proposition 4.10. *For a camera pair (i, j), let P_i and P_j be the sign-corrected projection matrices. Let $\mathit{\Pi}$ be a plane in $\mathbb{P}^{*3}$ with coordinates $\Pi = \begin{pmatrix}\pi^\top, 1\end{pmatrix}^\top$ and H_{ij} be its induced inter-image homography, where $\mathsf{H}_{1i} = \mathsf{P}_i \begin{bmatrix}\mathsf{I} & -\pi\end{bmatrix}^\top$ and $\mathsf{H}_{ij} = \mathsf{H}_{1j}\mathsf{H}_{1i}^*$. Provided that $\det(\mathsf{H}_{1i}) > 0$, for a positive integer k, the matrix $-\mathsf{H}_{ij}^k$ is stable if and only if:*

$$\begin{aligned} & \mathrm{tr}\big(\mathsf{H}_{ij}^k\big) > 0, \\ & \mathrm{tr}\big(\mathsf{H}_{ij}^k\big)\,\mathrm{tr}\big(\mathsf{H}_{ji}^k\big) - \big(\det(\mathsf{H}_{1i})\det(\mathsf{H}_{1j})\big)^k > 0, \\ \textit{and} \quad & \det(\mathsf{H}_{1j}) > 0 \quad \textit{if } k \textit{ is odd.} \end{aligned} \tag{4.28}$$

Proof. The proof is similar to the one of Proposition 4.9. The matrix $-\mathsf{H}_{ij}^k$ is stable if its characteristic polynomial is stable. Its characteristic equation,

$$\det\big(\lambda\mathsf{I} + \mathsf{H}_{ij}^k\big) = 0, \tag{4.29}$$

is cubic in λ. Multiplying both sides of (4.29) by $\det(\mathsf{H}_{1i})$, we obtain,

$$\det\big(\lambda\mathsf{H}_{1i} + \mathsf{H}_{ij}^{k-1}\mathsf{H}_{1j}\det(\mathsf{H}_{1i})\big) = 0. \tag{4.30}$$

The stability of the characteristic polynomial in (4.29) is unaffected by this multiplication, provided that $\det(\mathsf{H}_{1i})$ is positive. Compared with the standard form, $p(\lambda) = a_3\lambda^3 + a_2\lambda^2 + a_1\lambda + a_0$, the coefficients of the polynomial in (4.30) are:

$$\begin{aligned} a_3 &= \det(\mathsf{H}_{1i}), \\ a_2 &= \det(\mathsf{H}_{1i})\,\mathrm{tr}\big(\mathsf{H}_{ij}^k\big), \\ a_1 &= (\det(\mathsf{H}_{1i}))^{k+1}\,\mathrm{tr}\big(\mathsf{H}_{ji}^k\big), \\ a_0 &= (\det(\mathsf{H}_{1i}))^{2k+1}(\det(\mathsf{H}_{1j}))^k. \end{aligned} \tag{4.31}$$

We then obtain the inequalities (4.28) by applying the Hurwitz stability criterion to this polynomial. ■

Proposition 4.10 generalizes the result given in Proposition 4.9 to (positive integer) powers of H_{ij}. For a given power k, we obtain polynomial inequalities of maximum degree $2k$ as necessary and sufficient conditions for $-\mathsf{H}_{ij}^k$ to be stable. Note that $\mathrm{tr}\big(\mathsf{H}_{ji}^k\big) > 0$ is implied by these inequalities. We now show how these stability conditions can be used to constrain $\mathit{\Pi}_\infty$ when a vague knowledge of the relative camera orientation angle is available. Similar to the presentation of the stability conditions, we first state the constraints for $|\theta_{ij}| < 90°$, followed by the more general result for $|\theta_{ij}| < 90°/k$.

Proposition 4.11. *For a camera pair (i, j) with constant intrinsic parameters, let θ_{ij} be the relative camera orientation angle, and P_i and P_j be the sign-corrected projection matrices. Let $\mathit{\Pi}$ be a plane in $\mathbb{P}^{*3}$ with coordinates $\Pi = \big(\pi^\top, 1\big)^\top$, and let H_{ij} be its induced inter-image homography, where $\mathsf{H}_{1i} = \mathsf{P}_i\begin{bmatrix}\mathsf{I} & -\pi\end{bmatrix}^\top$ and $\mathsf{H}_{ij} = \mathsf{H}_{1j}\mathsf{H}_{1i}^*$. If $|\theta_{ij}| < 90°$, a necessary condition for $\mathit{\Pi}$ to be the plane at infinity $\mathit{\Pi}_\infty$, and hence for H_{ij} to be the infinite homography $\mathsf{H}_{\infty ij}$, is that $-\mathsf{H}_{ij}$ is stable.*

Proof. When $\Pi = \Pi_\infty$, the homography H_{ij} is the inter-image infinite homography $\mathsf{H}_{\infty ij}$ in (2.7). For constant intrinsic parameters, recall from (2.44) that the eigenvalues of $\mathsf{H}_{\infty ij}$ are a scaled version of those of the rotation matrix between the two cameras. The real parts of the eigenvalues of $-\mathsf{H}_{\infty ij}$ are,

$$\big\{-\lambda_i^2\lambda_j,\ -\lambda_i^2\lambda_j\cos\theta_{ij},\ -\lambda_i^2\lambda_j\cos\theta_{ij}\big\}, \tag{4.32}$$

where $\det(\mathsf{H}_{\infty 1i}) = \lambda_i^3$ (see Property 4.2). For sign-corrected projection matrices, we have that $\lambda_i > 0$ and $\lambda_j > 0$. When $|\theta_{ij}| < 90^\circ$, $\cos\theta_{ij} > 0$, and hence the real parts of all the eigenvalues are negative, *i.e.* $-\mathsf{H}_{\infty ij}$ is stable (see Definition 4.11). Thus, a necessary condition for $\mathit{\Pi}$ to be $\mathit{\Pi}_\infty$, and thereby for its induced homography H_{ij} to be the infinite homography $\mathsf{H}_{\infty ij}$, is that $-\mathsf{H}_{ij}$ is stable. ■

The polynomial inequalities in Proposition 4.9 are necessary and sufficient conditions for a homography matrix to be stable. From Proposition 4.11, they are, in turn, necessary conditions for it to be the infinite homography, *i.e.* for the reference plane $\mathit{\Pi}$ to be $\mathit{\Pi}_\infty$, when $|\theta_{ij}| < 90^\circ$. Hence, these inequalities form a new set of orientation-based constraints on $\mathit{\Pi}_\infty$ that can be exploited when the relative orientation angle is known to satisfy $|\theta_{ij}| < 90^\circ$ for a set of camera pairs. In the following proposition, we state the more general result for $|\theta_{ij}| < 90^\circ/k$.

Proposition 4.12. *For a camera pair (i, j) with constant intrinsic parameters, let θ_{ij} be the relative camera orientation angle, and P_i and P_j be the sign-corrected projection matrices. Let $\mathit{\Pi}$ be a plane in $\mathbb{P}^{*3}$ with coordinates $\Pi = \left(\pi^\top, 1\right)^\top$, and let H_{ij} be its induced inter-image homography, where $\mathsf{H}_{1i} = \mathsf{P}_i \begin{bmatrix} \mathsf{I} & -\pi \end{bmatrix}^\top$ and $\mathsf{H}_{ij} = \mathsf{H}_{1j}\mathsf{H}_{1i}^*$. If $|\theta_{ij}| < 90^\circ/k$, where k is a positive integer, a necessary condition for $\mathit{\Pi}$ to be the plane at infinity $\mathit{\Pi}_\infty$ is that $-\mathsf{H}_{ij}^k$ is stable.*

Proof. The proof is along the lines of the one in Proposition 4.11. It relies, in addition, on the De Moivre's theorem [Abramowitz and Stegun 1964], which states that the following identity holds,

$$(\cos\theta + \mathrm{i}\sin\theta)^k = \cos k\theta + \mathrm{i}\sin k\theta, \tag{4.33}$$

for any positive integer k, where $\mathrm{i}^2 = -1$. Now, when $\Pi = \Pi_\infty$, H_{ij} is the inter-image infinite homography $\mathsf{H}_{\infty ij}$ in (2.7). The kth power of $\mathsf{H}_{\infty ij}$, given by,

$$\mathsf{H}_{\infty ij}^k = \lambda_i^{2k}\lambda_j^k \mathsf{K}\mathsf{R}_{ij}^k\mathsf{K}^{-1}, \tag{4.34}$$

is similar, up to scale, to R_{ij}^k, where R_{ij} is the rotation matrix between the two cameras. In (4.34), K is the intrinsic parameters matrix and $\det(\mathsf{H}_{\infty 1i}) = \lambda_i^3$ (see Property 4.2). As such, the eigenvalues of $-\mathsf{H}_{\infty ij}^k$ are,

$$\left\{-\lambda_i^{2k}\lambda_j^k,\ -\lambda_i^{2k}\lambda_j^k(\cos\theta_{ij} + i\sin\theta_{ij})^k,\ -\lambda_i^{2k}\lambda_j^k(\cos\theta_{ij} - i\sin\theta_{ij})^k\right\}. \quad (4.35)$$

Using De Moivre's theorem, the real parts of these eigenvalues are,

$$\left\{-\lambda_i^{2k}\lambda_j^k,\ -\lambda_i^{2k}\lambda_j^k\cos k\theta_{ij},\ -\lambda_i^{2k}\lambda_j^k\cos k\theta_{ij}\right\}. \quad (4.36)$$

For sign-corrected projection matrices, we have that $\lambda_i > 0$ and $\lambda_j > 0$. When $|\theta_{ij}| < 90^\circ/k$, we have $\cos k\theta_{ij} > 0$ and hence $-\mathsf{H}_{\infty ij}^k$ is stable. Thus, a necessary condition for $\mathit{\Pi}$ to be the plane at infinity $\mathit{\Pi}_\infty$ is that the matrix $-\mathsf{H}_{ij}^k$ is stable. ■

The polynomial inequalities in Proposition 4.10 are necessary and sufficient conditions for the matrix $-\mathsf{H}_{ij}^k$ to be stable. From Proposition 4.12, these form necessary conditions for the reference plane $\mathit{\Pi}$ to be $\mathit{\Pi}_\infty$ when $|\theta_{ij}| < 90^\circ/k$. Hence, we obtain a set of polynomial constraints on $\mathit{\Pi}_\infty$ that can be exploited when the relative orientation angle is known to be $|\theta_{ij}| < 90^\circ/k$ for a set of a camera pairs. These constraints are non-convex and complementary to the LMIs for $|\theta_{ij}| \leq 120^\circ$. We elaborate on this latter point in the next section. For a given k, we propose to simultaneously use the constraints corresponding to $k,\ k-1,\ \ldots, 1$. We explain the rationale behind doing so with the help of some examples.

- Case of $k = 2$: this case corresponds to the constraints for $|\theta_{ij}| < 45^\circ$. The matrix $-\mathsf{H}_{\infty ij}^2$ is stable in this interval as $\cos 2\theta_{ij}$ is positive (see the proof of Proposition 4.12). Note, however, that $\cos 2\theta_{ij}$ is also positive when $|\theta_{ij}| > 135^\circ$. Therefore, $-\mathsf{H}_{\infty ij}^2$ is stable in this interval as well. To restrict ourselves to $|\theta_{ij}| < 45^\circ$, we can additionally consider the stability conditions for $-\mathsf{H}_{\infty ij}$, corresponding to $|\theta_{ij}| < 90^\circ$ and $\cos\theta_{ij} > 0$.

- Case of $k = 3$: following the same reasoning as for $k = 2$, we observe that besides the considered interval $|\theta_{ij}| < 30^\circ$, $\cos 3\theta_{ij}$ is also positive when $90^\circ < |\theta_{ij}| < 150^\circ$. Thus, $-\mathsf{H}_{\infty ij}^3$ is stable in this interval as well. To restrict ourselves to $|\theta_{ij}| < 30^\circ$, we can consider, in addition, the stability conditions

for $-\mathsf{H}_{\infty ij}$, corresponding to $|\theta_{ij}| < 90^\circ$ and $\cos\theta_{ij} > 0$. Note that we do not simultaneously require the conditions for $-\mathsf{H}_{\infty ij}^2$ in this case. Conversely, the conditions for $-\mathsf{H}_{\infty ij}^2$, without those for $-\mathsf{H}_{\infty ij}$, are not sufficient to restrict the considered interval to $|\theta_{ij}| < 30^\circ$.

We can continue this reasoning for higher powers. From these examples, we conclude that for a given k, a subset of the inequalities for $k, k-1, \ldots, 1$ ought to be used simultaneously for the orientation constraint $|\theta_{ij}| < 90^\circ/k$. We propose to simultaneously use all the inequalities. While this may introduce some redundant constraints, the redundancy can, in fact, be beneficial in numerical optimization. This is notably the case with polynomial optimization problems.

4.3.3 A more specialized QUARCH

The stability-based constraints presented in the previous section exploit a knowledge of $|\theta_{ij}| < 90^\circ/k$ for a positive integer k. It may seem, at first sight, that they replace the QUARCH LMIs for $|\theta_{ij}| \leq 120^\circ$. In this section, we show that these constraints are, in fact, complementary and can be used together to recover a specialized QUARCH. This section is divided into two parts. In the first part, we establish the complementary roles played by the non-convex polynomial constraints and the QUARCH LMIs. To do so, we discuss the algebraic constraint imposed by the LMIs on the inter-image homography. This discussion complements their presentation in Section 4.2.4. In the second part, we focus on recovering a specialized QUARCH using the LMIs and the non-convex constraints together. This involves estimating bounds on the coordinates of $\mathit{\Pi}_\infty$ using the orientation-based constraints in conjunction with the chirality inequalities.

Complementary orientation-based constraints

Before diving into the details, we illustrate the independence of the stability-based polynomial constraints and the QUARCH LMIs through a simple numerical example. Consider the following diagonal matrix,

$$\mathsf{H} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

The eigenvalues of H are all real and positive. Hence, $-\mathsf{H}$ is stable, and the inequalities in Proposition 4.9 are satisfied for this matrix. Furthermore, as $\det(\mathsf{H}) = 1$ and $\operatorname{tr}(\mathsf{H}^*)/\operatorname{tr}(\mathsf{H}) = 1$, H satisfies the modulus constraint (2.46). Now, consider the QUARCH LMIs in Proposition 4.5. Using Property 4.2, we have that the two matrices,

$$\begin{aligned} \begin{bmatrix} \det(\mathsf{I}) & \operatorname{tr}(\mathsf{H}) \\ \operatorname{tr}(\mathsf{H}) & 3\operatorname{tr}(\mathsf{H}^*) \end{bmatrix} &= \begin{bmatrix} 1 & 3.5 \\ 3.5 & 10.5 \end{bmatrix}, \\ \begin{bmatrix} \det(\mathsf{H}) & \operatorname{tr}(\mathsf{H}^*) \\ \operatorname{tr}(\mathsf{H}^*) & 3\operatorname{tr}(\mathsf{H}) \end{bmatrix} &= \begin{bmatrix} 1 & 3.5 \\ 3.5 & 10.5 \end{bmatrix}, \end{aligned} \tag{4.37}$$

ought to be positive semidefinite for H to satisfy these LMIs. These two matrices are identical, and their eigenvalues are -0.15 and 11.65. Hence, they are not positive semidefinite, and H does not satisfy the QUARCH LMIs. If H is a homography between two images, then it satisfies the stability-based inequalities for $|\theta_{ij}| < 90°$ and the modulus constraint, yet it does not satisfy the QUARCH LMIs for $|\theta_{ij}| \leq 120°$. Thus, these LMIs are not implied by the stability-based inequalities or by the modulus constraint polynomial (or their combination). The following proposition and the ensuing discussion shed more light on the complementary constraint imposed by the LMIs.

Proposition 4.13. *For a camera pair (i, j) with constant intrinsic parameters, let θ_{ij} be the relative camera orientation angle, and P_i and P_j be the sign-corrected projection matrices. Let Π be a plane in $\mathbb{P}^{*3}$ with coordinates $\Pi = \left(\pi^\top, 1\right)^\top$, and let H_{ij} be its induced inter-image homography, where $\mathsf{H}_{1i} = \mathsf{P}_i \begin{bmatrix} \mathsf{I} & -\pi \end{bmatrix}^\top$ and*

$\mathsf{H}_{ij} = \mathsf{H}_{1j}\mathsf{H}_{1i}^{*}$. *Consider the shifted matrix* $\mathsf{D}_{ij} = \mathsf{H}_{ij} - \alpha_{ij}\mathsf{I}$, *where* α_{ij} *is a real eigenvalue of* H_{ij}. *If* $|\theta_{ij}| < 120^{\circ}$ *and the modulus constraint is satisfied, the following statements are equivalent:*

(i) The two nonzero eigenvalues of D_{ij} *have negative real parts and* $\operatorname{tr}(\mathsf{H}_{ij}) > 0$.

(ii) Π *is a strict QUARCH plane (see Definition 4.8).*

Proof. For simplicity and without loss of generality, we consider a homography H_{1j} relating image j and the reference view. Its shifted matrix is $\mathsf{D}_{1j} = \mathsf{H}_{1j} - \alpha_{1j}\mathsf{I}$, where α_{1j} is a real eigenvalue of H_{1j}. We now prove the assertion $(i) \Rightarrow (ii)$. Consider the characteristic equation of D_{1j},

$$\det(\lambda\mathsf{I} + \alpha_{1j}\mathsf{I} - \mathsf{H}_{1j}) = 0, \tag{4.38}$$

which expands as $p(\lambda) = a_3\lambda^3 + a_2\lambda^2 + a_1\lambda + a_0 = 0$, where $a_0, \ldots, a_3$ are real coefficients. The tail coefficient, $a_0 = \det(\alpha_{1j}\mathsf{I} - \mathsf{H}_{1j})$, is zero as α_{1j} is an eigenvalue of H_{1j}. The roots of the quadratic polynomial $a_3\lambda^2 + a_2\lambda + a_1$ correspond to the two nonzero eigenvalues of D_{1j}. The coefficients of this polynomial are,

$$\begin{aligned} a_1 &= \operatorname{tr}((\alpha_{1j}\mathsf{I} - \mathsf{H}_{1j})^{*}) = 3\alpha_{1j}^2 - 2\alpha_{1j}\operatorname{tr}(\mathsf{H}_{1j}) + \operatorname{tr}(\mathsf{H}_{j1}), \\ a_2 &= \operatorname{tr}(\alpha_{1j}\mathsf{I} - \mathsf{H}_{1j}) = 3\alpha_{1j} - \operatorname{tr}(\mathsf{H}_{1j}), \\ a_3 &= \det(\mathsf{I}) = 1. \end{aligned} \tag{4.39}$$

The expression of a_1 can be obtained using a symbolic software, such as Maple[5]. From statement (i), we have that the real parts of the two nonzero eigenvalues of D_{1j} are negative. Thus, the polynomial $a_3\lambda^2 + a_2\lambda + a_1$ is stable. From the Hurwitz stability criterion, a necessary and sufficient condition for this polynomial to be stable is that $a_1 > 0$ and $a_2 > 0$, provided that $a_3 > 0$. Since the modulus constraint (2.46) is satisfied, we also have $\alpha_{1j} = \operatorname{tr}(\mathsf{H}_{j1})/\operatorname{tr}(\mathsf{H}_{1j})$. By substituting for α_{1j} in a_1 and a_2, we obtain the following inequalities,

$$\begin{aligned} \operatorname{tr}(\mathsf{H}_{1j})\big(3\operatorname{tr}(\mathsf{H}_{j1}) - \operatorname{tr}^2(\mathsf{H}_{1j})\big) > 0, \\ \operatorname{tr}(\mathsf{H}_{j1})\big(3\operatorname{tr}(\mathsf{H}_{j1}) - \operatorname{tr}^2(\mathsf{H}_{1j})\big) > 0. \end{aligned} \tag{4.40}$$

[5] `https://www.maplesoft.com/products/Maple/`

From statement (i), we also have that $\mathrm{tr}(\mathsf{H}_{1j}) > 0$. Thus, we obtain the following quadratic inequality from (4.40),

$$3\,\mathrm{tr}(\mathsf{H}_{j1}) - \mathrm{tr}^2(\mathsf{H}_{1j}) > 0. \tag{4.41}$$

Using the Schur complement lemma (Lemma A.1), inequality (4.41) can be reformulated into the following LMI,

$$\begin{bmatrix} \det(\mathsf{I}) & \mathrm{tr}(\mathsf{H}_{1j}) \\ \mathrm{tr}(\mathsf{H}_{1j}) & 3\,\mathrm{tr}(\mathsf{H}_{j1}) \end{bmatrix} \succ 0. \tag{4.42}$$

It can be seen using Property 4.2, that the LMI (4.42) is the strict form of the first LMI in (4.16), *i.e.* the strict version of the first QUARCH LMI for $|\theta_{ij}| \leq 120^\circ$. To obtain the second LMI, we observe that the following inequality is implied by (4.41) and the modulus constraint (2.46),

$$3\,\mathrm{tr}(\mathsf{H}_{1j})\det(\mathsf{H}_{1j}) - \mathrm{tr}^2(\mathsf{H}_{j1}) > 0. \tag{4.43}$$

Using Lemma A.1, this inequality can be reformulated into the strict form of the second QUARCH LMI for $|\theta_{ij}| \leq 120^\circ$. Note that $\det(\mathsf{H}_{1j}) > 0$, required to apply Lemma A.1, is ensured by the modulus constraint, $\mathrm{tr}(\mathsf{H}_{1j}) > 0$, and inequality (4.41) being satisfied simultaneously. Hence, $\mathit{\Pi}$ satisfies the strict form of the QUARCH LMIs for $|\theta_{ij}| \leq 120^\circ$ and is therefore a strict QUARCH plane. This proves $(i) \Rightarrow (ii)$. The reverse implication follows from Lemma A.1, and we omit here the detailed steps. ■

Proposition 4.13 shows that the (strict) QUARCH LMIs for $|\theta_{ij}| \leq 120^\circ$ can be derived using the Hurwitz stability criterion. While the geometric interpretation of these LMIs was described in Section 4.2.4, we now discuss the algebraic constraint imposed by them. In Proposition 4.13, we showed that the strict QUARCH LMIs are necessary and sufficient conditions for the nonzero eigenvalues of the shifted matrix $\mathsf{D}_{ij} = \mathsf{H}_{ij} - \alpha_{ij}\mathsf{I}$ to have negative real parts. In turn, these are conditions for the real parts of the eigenvalues of H_{ij} to be bounded from above by α_{ij}, where $\alpha_{ij} = \mathrm{tr}(\mathsf{H}_{ji})/\,\mathrm{tr}(\mathsf{H}_{ij})$. The inequalities in Proposition 4.9, on the other hand, are necessary and sufficient conditions for $-\mathsf{H}_{ij}$ to be stable, *i.e.* for the real parts of

the eigenvalues of H_{ij} to be positive. Hence, we conclude that the stability-based inequalities for $|\theta_{ij}| < 90^\circ$ and the QUARCH LMIs for $|\theta_{ij}| \leq 120^\circ$ play complementary roles in constraining the eigenvalues of H_{ij}. As illustrated by the example at the beginning of this section, these constraints thus ought to be used together.

Orientation-based bounds on the plane at infinity

We now turn our attention to recovering a specialized QUARCH using the polynomial constraints for $|\theta_{ij}| < 90^\circ/k$ together with the QUARCH LMIs. By a specialized QUARCH, we mean a QUARCH that is closer to the affine reconstruction than the one defined in Section 4.2.5. In other words, we identify a stratum that is a specialization of the QUARCH shown in Figure 4.5. We proceed by deriving bounds on the coordinates of $\mathit{\Pi}_\infty$ using the orientation-based constraints and the chirality inequalities [Hartley and Zisserman 2004]. These bounds are expressed as linear inequalities in the coordinates of $\mathit{\Pi}_\infty$, and they define a convex set. The intersection of QUARCH and this set defines a specialized QUARCH. A projective reconstruction can be upgraded to a specialized QUARCH by identifying a plane within this subset. We describe next the computation of the bounds on $\mathit{\Pi}_\infty$ using the orientation-based constraints.

In a centered quasi-affine reference frame, bounds on the first three coordinates of $\mathit{\Pi}_\infty$ (the fourth being one) can be obtained by minimizing/maximizing each coordinate subject to the chirality inequalities, *i.e.* by solving six linear programming problems. The location of $\mathit{\Pi}_\infty$ is thereby restricted to a bounded volume, see Section 2.2.6 for further details. We can additionally use the orientation-based constraints to compute such bounds by including the corresponding polynomial inequalities in the optimization problem. The linear programs are then turned into polynomial optimization problems that can be solved, for instance, using Lasserre's

hierarchy [Lasserre 2001]. To estimate the bounds on the lth coordinate of $\mathit{\Pi}_\infty$ (where $l = 1, 2, 3$), the following polynomial optimization problem can be solved,

$$\begin{aligned} &\min \,/\, \max\ (\Pi)_l \\ &\text{s.t.} \qquad \mathit{\Pi} \in \mathcal{B}, \\ &\qquad\quad \mathit{\Pi} \text{ is a QUARCH plane with respect to camera pairs } (i,j), \qquad (4.44) \\ &\qquad\quad -\mathsf{H}_{ij}^{k}, -\mathsf{H}_{ij}^{k-1}, \ldots, -\mathsf{H}_{ij} \text{ are stable,} \qquad i = 1, \ldots, n-1, \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad j = i+1, \end{aligned}$$

where $\mathcal{B}$ is the set of chirality inequalities (2.20), and the orientation knowledge $|\theta_{ij}| < 90^\circ/k$, for a particular k, is assumed between consecutive views. Problem (4.44) optimizes a linear objective subject to linear (chirality) and polynomial (orientation-based) inequalities. Note that the QUARCH constraints are given by quadratic inequalities, as expressed in (4.41) and (4.43). By solving problem (4.44) for $l = 1, 2, 3$, we obtain lower and upper bounds b and b', respectively, on the first three coordinates of $\mathit{\Pi}_\infty$. These bounds define the following convex set,

$$\mathcal{B}' = \left\{ \mathit{\Pi} \in \mathbb{P}^{*3} : \ (\mathsf{b})_l \leq (\Pi)_l \leq (\mathsf{b}')_l, \ \ (\Pi)_4 = 1, \quad l = 1, 2, 3 \right\}. \qquad (4.45)$$

The intersection of $\mathcal{B}'$ with QUARCH gives rise to a specialized QUARCH. In practice, the chirality inequalities for the scene points usually far outnumber the other constraints. A large number of constraints can significantly burden polynomial optimization methods. Therefore, rather than using the chirality inequalities directly, we propose to instead incorporate bounds computed using them in (4.44). In other words, we define the set $\mathcal{B}$ in (4.44) by the bounds on $\mathit{\Pi}_\infty$ computed using the chirality constraints alone (see Section 2.2.6). While this is not as strict as using all the constraints simultaneously, it is a practical workaround to solve the polynomial optimization problems efficiently.

4.4 Orientation-based stratified autocalibration

In this section, we present a stratified camera autocalibration method that relies on the orientation-based constraints introduced in Sections 4.2 and 4.3 to estimate $\mathit{\Pi}_\infty$.

We use these constraints to compute a QUARCH that serves as an initial step towards recovering the affine and metric reconstructions from a projective one. To estimate $\mathit{\Pi}_\infty$, we rely on a QUARCH plane as an initialization for local optimization of a suitable cost function. We also propose a constrained Levenberg-Marquardt method for nonlinear optimization subject to LMI constraints. This method ensures that the iterating plane during the optimization remains in the subset of QUARCH planes defined by the orientation constraints. Once $\mathit{\Pi}_\infty$ is estimated using local optimization, the camera intrinsic parameters are retrieved by solving linear equations for the DIAC. In the rest of this section, we fill in the details of our stratified method. In Section 4.4.1, we describe the QUARCH computation. In Section 4.4.2, we present the constrained optimization method used to locate $\mathit{\Pi}_\infty$. Finally, in Section 4.4.3 we summarize the steps of our autocalibration algorithm.

4.4.1 Computing a QUARCH

To compute a QUARCH, we assume that the orientation angle θ_{ij} between consecutive views satisfies, $|\theta_{ij}| \leq 120^\circ$. As already discussed, this is a mild assumption that is often implicitly satisfied in ordered image sequences captured for 3D reconstruction. Further knowledge of the relative orientation angle being $|\theta_{ij}| < 90^\circ/k$, where k is a positive integer, can be used to obtain a specialized QUARCH (see Section 4.3.3). We solve the following SDP problem to obtain a QUARCH plane,

$$
\begin{aligned}
\max_{\Pi,\,\mathsf{Z}} \quad & \log\det \mathsf{Z} \\
\text{s.t.} \quad & \mathsf{Z} \succeq 0, \\
& (\mathsf{b})_l(\Pi)_4 \leq (\Pi)_l \leq (\mathsf{b}')_l(\Pi)_4, \qquad l = 1, \ldots, 3, \\
& (\Pi)_4 \leq 1, \\
& \begin{bmatrix} \Pi^\top \mathsf{C}_i & \Pi^\top \mathsf{T}_{ij} \\ \Pi^\top \mathsf{T}_{ij} & 3\Pi^\top \mathsf{T}_{ji} \end{bmatrix} \succeq \mathsf{Z}, \qquad i = 1, \ldots, n-1, \\
& \begin{bmatrix} \Pi^\top \mathsf{C}_j & \Pi^\top \mathsf{T}_{ji} \\ \Pi^\top \mathsf{T}_{ji} & 3\Pi^\top \mathsf{T}_{ij} \end{bmatrix} \succeq \mathsf{Z}, \qquad j = i+1,
\end{aligned}
\tag{4.46}
$$

where the values for $(\mathsf{b})_l$ and $(\mathsf{b}')_l$ depend on whether the orientation constraints for $|\theta_{ij}| < 90^\circ/k$ are used. If so, $(\mathsf{b})_l$ and $(\mathsf{b}')_l$ are the bounds in (4.45) for $l = 1, 2, 3$. We obtain a specialized QUARCH in this case. Otherwise, we set all $(\mathsf{b})_l = -1$ and $(\mathsf{b}')_l = 1$, as in problem (4.22). The problem is then bounded. Note that maximizing $\log\det \mathsf{Z}$ (Z here is a symmetric 2×2 matrix) is a convex optimization problem. We use this cost function rather than the one in (4.22) as maximizing $\log\det \mathsf{Z}$ prevents the determinants of the two matrices in the LMIs in (4.46) from being arbitrarily close to zero. This is the case for $\mathit{\Pi}_\infty$ as $|\theta_{ij}|$ approaches 0° or 120°. We assume that the camera motion between consecutive views is not close to either of these configurations, *i.e.* a pure translation or a large rotation. Through empirical tests, we have found that when using our algorithm, a QUARCH plane computed via (4.46) converges more reliably to the sought $\mathit{\Pi}_\infty$ than the one determined via (4.22).

4.4.2 Locating infinity using LMI-constrained optimization

To estimate $\mathit{\Pi}_\infty$, we use the QUARCH plane from problem (4.46) as an initialization for local optimization of a suitable cost function. We require the constraints used in (4.46) to be enforced during the optimization so that the iterating plane remains in the QUARCH set. We propose solving the following problem to locate $\mathit{\Pi}_\infty$,

$$\begin{aligned}
\min_{\Pi} \quad & \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \frac{\mathcal{M}_{ij}^2(\Pi)}{(\Pi^\top \mathsf{C}_i)^4\,(\Pi^\top \mathsf{C}_j)^4} \\
\text{s.t.} \quad & (\mathsf{b})_l \le (\Pi)_l \le (\mathsf{b}')_l, && l = 1, \ldots, 3, \\
& \begin{bmatrix} \Pi^\top \mathsf{C}_i & \Pi^\top \mathsf{T}_{ij} \\ \Pi^\top \mathsf{T}_{ij} & 3\Pi^\top \mathsf{T}_{ji} \end{bmatrix} \succeq 0, && i = 1, \ldots, n-1, \\
& \begin{bmatrix} \Pi^\top \mathsf{C}_j & \Pi^\top \mathsf{T}_{ji} \\ \Pi^\top \mathsf{T}_{ji} & 3\Pi^\top \mathsf{T}_{ij} \end{bmatrix} \succeq 0, && j = i+1,
\end{aligned} \tag{4.47}$$

where $\mathcal{M}_{ij}$ is the polynomial from the modulus constraint in (4.6), and the normalization eliminates the projective scale factors from the cost. The vectors b and b' are, again, the bounds in (4.45). The bound inequalities are used if they were also used in the QUARCH plane computation in (4.46), *i.e.* depending on the relative orientation knowledge. Note that we optimize for the first three coordinates of $\mathit{\Pi}_\infty$,

Algorithm 1 Constrained Levenberg-Marquardt Method [Kanzow et al. 2004]

1: Choose $\mathsf{x}_0 \in \mathcal{C}$, $\mu > 0$, and set $k = 0$.

2: If $\mathsf{F}(\mathsf{x}_k) = 0$, stop.

3: Choose J_k, set $\mu_k = \mu\|\mathsf{F}(\mathsf{x}_k)\|^2$, and compute d_k as the solution of (4.49).

4: Set $\mathsf{x}_{k+1} = \mathsf{x}_k + \mathsf{d}_k$, $k \leftarrow k + 1$, and go to step 2.

fixing the fourth to one. Problem (4.47) is a nonlinear optimization problem subject to LMI constraints (linear inequalities can also be written as an LMI). Unfortunately, there do not seem to be tailored methods for this problem in current solvers. While implementations of unconstrained optimization methods are ubiquitous, an unconstrained optimization of problem (4.47) may converge to a solution that is not a QUARCH plane. Such a solution is therefore not the sought Π_∞. To solve problem (4.47), we propose a constrained Levenberg-Marquardt (LM) method for nonlinear optimization subject to LMI constraints. The rest of this section details this constrained optimization method.

Our method builds upon previous work on LM optimization subject to convex constraints [Kanzow et al. 2004]. The general optimization problem is stated as,

$$\begin{aligned} &\min_{\mathsf{x}} f(\mathsf{x}) \\ &\text{s.t. } \mathsf{x} \in \mathcal{C}, \end{aligned} \tag{4.48}$$

where $f(\mathsf{x}) = \|\mathsf{F}(\mathsf{x})\|^2$ is the natural merit function corresponding to the mapping $\mathsf{F}(\mathsf{x})$ and $\mathcal{C}$ is a nonempty closed convex set. In [Kanzow et al. 2004], the authors solved the following regularized form of problem (4.48),

$$\begin{aligned} &\min_{\mathsf{d}} \|\mathsf{F}(\mathsf{x}_k) + \mathsf{J}_k\,\mathsf{d}\|^2 + \mu_k\|\mathsf{d}\|^2 \\ &\text{s.t. } \mathsf{x}_k + \mathsf{d} \in \mathcal{C}, \end{aligned} \tag{4.49}$$

that involves iteratively computing a step d such that the iterate $\mathsf{x}_k + \mathsf{d}$ is in the set $\mathcal{C}$. In problem (4.49), J_k is the Jacobian of $\mathsf{F}(\mathsf{x}_k)$ and μ_k is a positive parameter at iteration k. Note that the quadratic objective function in (4.49) is strictly convex. Thus, a convex minimization problem is solved at each iteration. The steps of the opti-

mization algorithm are summarized in Algorithm 1. This constrained LM method was shown to converge locally quadratically under a local error bound condition. While the authors considered box constraints in their work, efficiently dealing with general convex constraints was left as an open question.

Our contribution is to formulate the step computation in (4.49) as an SDP problem, where the QUARCH LMIs and bound inequalities can be easily incorporated. To do so, we first observe that the objective function in (4.49) expands as follows,

$$\|\mathsf{F}_k + \mathsf{J}_k\mathsf{d}\|^2 + \mu_k\|\mathsf{d}\|^2 = \mathsf{F}_k^\top\mathsf{F}_k + 2\mathsf{F}_k^\top\mathsf{J}_k\mathsf{d} + \mathsf{d}^\top(\mathsf{J}_k^\top\mathsf{J}_k + \mu_k\mathsf{I})\mathsf{d}, \tag{4.50}$$

where F_k is a shorthand for $\mathsf{F}(\mathsf{x}_k)$. Hence, problem (4.49) is equivalent to,

$$\begin{aligned} \min_{\mathsf{d},\,\delta} \quad & \delta \\ \text{s.t.} \quad & \mathsf{x}_k + \mathsf{d} \in \mathcal{C}, \\ & \delta - \mathsf{F}_k^\top\mathsf{F}_k - 2\mathsf{F}_k^\top\mathsf{J}_k\mathsf{d} - \mathsf{d}^\top\big(\mathsf{J}_k^\top\mathsf{J}_k + \mu_k\mathsf{I}\big)\mathsf{d} \geq 0. \end{aligned} \tag{4.51}$$

The inequality (4.51) is quadratic in d and can be reformulated into an LMI by applying the Schur complement lemma. The step d can then be computed by solving the following SDP problem,

$$\begin{aligned} \min_{\mathsf{d},\,\delta} \quad & \delta \\ \text{s.t.} \quad & \begin{bmatrix} \mathsf{J}_k^\top\mathsf{J}_k + \mu_k\mathsf{I} & (\mathsf{J}_k^\top\mathsf{J}_k + \mu_k\mathsf{I})\mathsf{d} \\ \mathsf{d}^\top(\mathsf{J}_k^\top\mathsf{J}_k + \mu_k\mathsf{I}) & \delta - \mathsf{F}_k^\top\mathsf{F}_k - 2\mathsf{F}_k^\top\mathsf{J}_k\mathsf{d} \end{bmatrix} \succeq 0, \\ & \mathsf{x}_k + \mathsf{d} \in \mathcal{C}. \end{aligned} \tag{4.52}$$

Note that the term $\mathsf{J}_k^\top\mathsf{J}_k + \mu_k\mathsf{I}$ in the LMI (4.52) is positive definite by construction. Our LMI-constrained method thus solves the SDP problem (4.52) in the third step of Algorithm 1. In our case, x_0 is the QUARCH plane from the SDP problem (4.46), $\mathcal{C}$ is the subset of QUARCH planes defined by the QUARCH LMIs and optionally the bound inequalities, and $f(\mathsf{x})$ is the cost function used to locate $\mathit{\Pi}_\infty$ in (4.47). We optimize for the first three coordinates of $\mathit{\Pi}_\infty$, fixing the fourth to 1. Thus, we compute the first three coordinates of the step d, the fourth being 0. Our constrained LM method ensures that the iterating plane remains in the subset of QUARCH

planes to which $\mathit{\Pi}_\infty$ belongs. This constraint also prevents the iterating plane from crossing the camera centers, which could be fatal for the cost function in (4.47). As a local optimization method, it remains susceptible to converge to a local minimum of the cost function, albeit one that is also a QUARCH plane.

We conclude this section by discussing the choice of the parameter μ_k in Algorithm 1. This parameter influences the optimization, and some authors have pointed out that the choice $\mu_k = \mu\|\mathsf{F}(\mathsf{x}_k)\|^2$ is not the most appropriate [Yu 2004; Fan and Yuan 2005]. When x_k is far from the solution set, μ_k may be quite large, leading to a small step d and hindering the optimization progress. Conversely, as x_k approaches the solution set, μ_k may be quite small and thereby no longer play a part in the cost. We set $\mu_k = \mu\|\mathsf{F}(\mathsf{x}_k)\|$, rather than using the squared norm, following [Yu 2004]. As shown in [Fan 2013], the same quadratic convergence rate is obtained with this choice. In the following section, we summarize the steps of our stratified camera autocalibration algorithm.

4.4.3 Autocalibration algorithm

We consider a projective reconstruction with sign-corrected projection matrices P_i, $i = 1, \ldots, n$, where the reference frame is attached to the first camera, and scene points X_j, $j = 1, \ldots, m$. Our autocalibration algorithm then proceeds as follows:

(i) **QUARCH:** compute a QUARCH plane $\mathit{\Pi}_\mathsf{Q}$ by solving the SDP problem (4.46), and upgrade the projective reconstruction to a QUARCH as $\mathsf{P}_i^\mathsf{Q} = \mathsf{P}_i\mathsf{H}_\mathsf{Q}^{-1}$ and $\mathsf{X}_j^\mathsf{Q} = \mathsf{H}_\mathsf{Q}\mathsf{X}_j$, where,

$$\mathsf{H}_\mathsf{Q} = \begin{bmatrix} \mathsf{I}_3 & \mathbf{0}_3 \\ \Pi_\mathsf{Q}^\top \end{bmatrix}.$$

(ii) **Affine:** locate $\mathit{\Pi}_\infty$ by solving problem (4.47) using the LMI-constrained optimization method with $\mathit{\Pi}_\mathsf{Q}$ as the initialization. Upgrade the QUARCH to affine as $\mathsf{P}_i^\mathsf{A} = \mathsf{P}_i^\mathsf{Q}\mathsf{H}_\mathsf{A}^{-1}$ and $\mathsf{X}_j^\mathsf{A} = \mathsf{H}_\mathsf{A}\mathsf{X}_j^\mathsf{Q}$, where,

$$\mathsf{H}_\mathsf{A} = \begin{bmatrix} \mathsf{I}_3 & \mathbf{0}_3 \\ \Pi_\infty^\top \end{bmatrix}.$$

(iii) **Metric:** compute the camera calibration matrix K by solving linear equations for the DIAC. Upgrade the affine reconstruction to a metric one as $\mathsf{P}_i^{\mathsf{M}} = \mathsf{P}_i^{\mathsf{A}}\mathsf{H}_{\mathsf{M}}^{-1}$ and $\mathsf{X}_j^{\mathsf{M}} = \mathsf{H}_{\mathsf{M}}\mathsf{X}_j^{\mathsf{A}}$, where,

$$\mathsf{H}_{\mathsf{M}} = \begin{bmatrix} \mathsf{K}^{-1} & \mathbf{0}_3 \\ \mathbf{0}_3^\top & 1 \end{bmatrix}.$$

In the following section, we report the experiments that we conducted to evaluate our autocalibration algorithm and the benefits of the orientation-based constraints.

4.5 Experimental results

We tested our camera autocalibration method using both synthetic data and real images. In the experiments, we primarily evaluated our algorithm using the QUARCH LMIs. In one experiment with synthetic data, we evaluated the benefits of additionally using the non-convex polynomial constraints. For a quantitative evaluation, we computed the 3D RMS error and the following calibration error metrics:

$$\begin{aligned} \Delta f = \left|f_x - \widehat{f}_x\right| + \left|f_y - \widehat{f}_y\right|, \quad \Delta uv = |u - \widehat{u}| + |v - \widehat{v}|, \\ \Delta\gamma = |\gamma - \widehat{\gamma}|, \end{aligned} \tag{4.53}$$

where $(\widehat{f}_x,\, \widehat{f}_y)$, $(\widehat{u},\, \widehat{v})$, and $\widehat{\gamma}$ are the estimated focal lengths, principal point, and skew, respectively. We computed the 3D RMS error after aligning the recovered metric point cloud to the ground truth Euclidean one by a best-fit similarity transformation in the least-squares sense. Each point cloud was first scaled such that the mean distance of all the points from the center was one unit. We set $\mu = 0.5$, $\mu_0 = \mu\|\mathsf{F}(\mathsf{x}_k)\|$ and used the update $\mu_{k+1} = \min\{\mu_k,\, \mu_k\|\mathsf{F}(\mathsf{x}_{k+1})\|\}$ in the LMI-constrained optimization. Our algorithm was implemented in MATLAB R2017b. We used YALMIP [Löfberg 2004] to model the SDP problems, with MOSEK[6] as the solver. GloptiPoly [Henrion et al. 2009] was used to solve the polynomial optimization problems. All the experiments were conducted on an Intel Core i7 3.10 GHz 32GB RAM computer.

[6] `https://docs.mosek.com/8.1/toolbox/index.html`

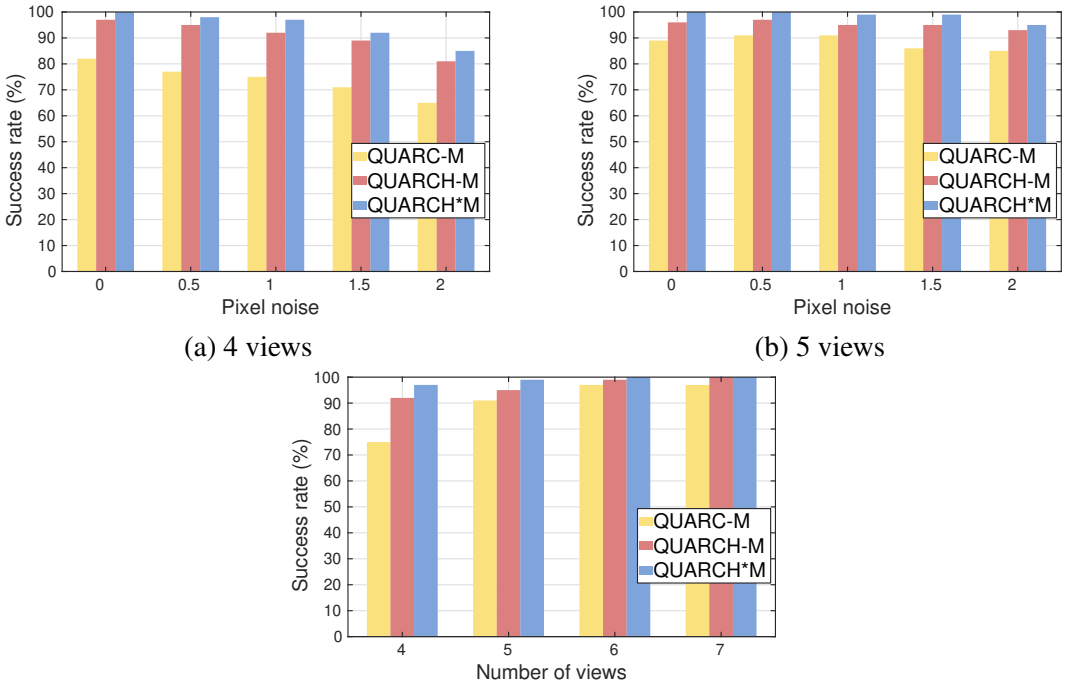
4.5.1 Synthetic data experiments

Each synthetic scene in our tests consisted of 500 points scattered randomly within the unit sphere. The cameras were placed at a distance of 2.75–3.45 units from the sphere center and facing towards it. Their positions were then perturbed by a small random translation. The rotation angle θ_{ij} between consecutive views was sampled randomly from the range $[20°, 60°]$ to satisfy the assumption $|\theta_{ij}| \leq 120°$. All cameras had (in pixels) focal length $f_x = f_y = 300$, zero skew, *i.e.* $\gamma = 0$, and an image-centered principal point, $(u_0, v_0) = (128, 128)$. Zero-mean Gaussian noise with standard deviation in the $[0, 2]$ pixel range was added to pixel coordinates in increments of 0.5 pixels. The length of the sequence was varied from 4 to 16 views, and 100 trials were run for each sequence length and image noise level. Projective reconstructions were computed using a factorization method [Oliensis and Hartley 2007] implemented in the VSfM toolbox[7], followed by a projective bundle adjustment [Lourakis and Argyros 2009].

In the first couple of experiments described in the following, we assess the benefits of the orientation-based constraints in our algorithm. We then compare the results obtained from our algorithm with those from existing methods. In these experiments, we used only the QUARCH LMIs as the orientation-based constraints in the first two steps of our algorithm. In the final experiment reported in this section, we evaluate the benefits of incorporating further orientation knowledge using the non-convex polynomial constraints.

Benefits of QUARCH: We compared a QUARCH plane initialization with a QUARC by solving an unconstrained version of problem (4.47) in step (ii) of our algorithm. We denote these two unconstrained variants of our algorithm by QUARCH-M and QUARC-M, respectively. We compared their reliability in recovering a metric reconstruction from a projective one and focused on short sequences. We considered a 3D error of 0.02 as a threshold for a successful metric upgrade. Figure 4.6 shows the success rate of QUARC-M and QUARCH-M for a varying number of views and pixel noise levels. For 4 and 5 views in Figures 4.6(a) and 4.6(b), respectively, QUARCH-M obtained a significantly higher success rate

[7] `http://github.com/vrabaud/sfm_toolbox`



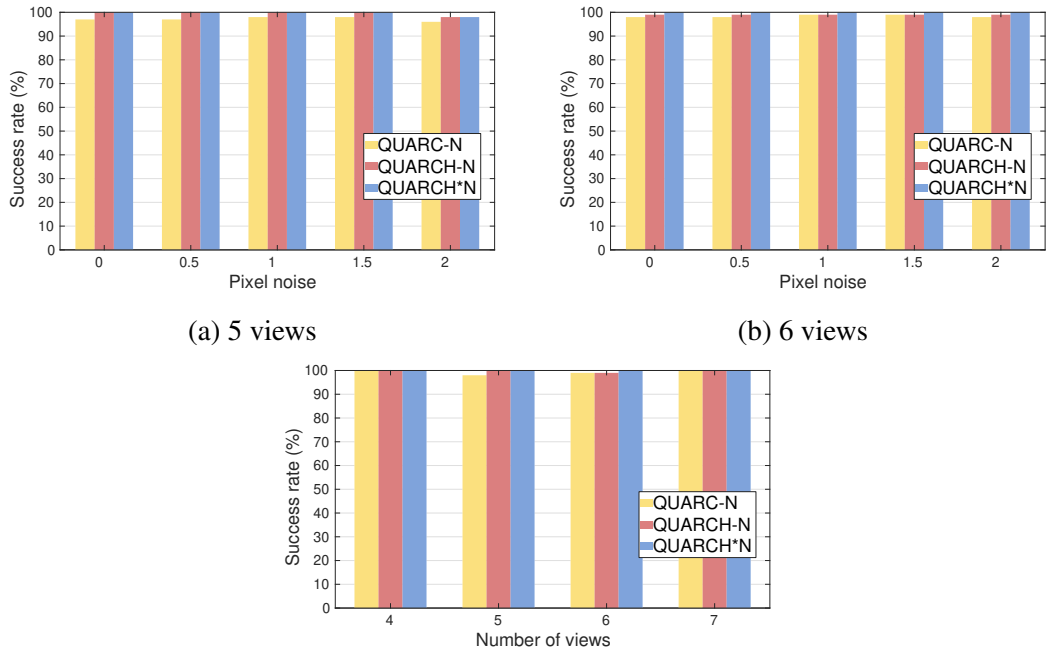
(a) 4 views

(b) 5 views

(c) 1 pixel noise

Figure 4.6: Benefits of QUARCH and the LMI-constrained optimization. QUARCH*M denotes our algorithm. QUARCH-M and QUARC-M are its variants using an unconstrained optimization. QUARC-M uses a QUARC plane as initialization.

across the different pixel noise levels. These results show that the unconstrained optimization converged more reliably to Π_∞ when starting from a QUARCH plane, thereby leading to a successful metric upgrade more often. As the sequence length increased to 6 and 7 views in Figure 4.6(c), both approaches succeeded most of the time, though QUARCH-M still outperformed QUARC-M. We also compared the reliability using the cost function in [Nistér 2004b]. The corresponding variants of our algorithm are similarly denoted as QUARCH-N and QUARC-N. Figure 4.7 shows the results using this cost function. Here as well, a QUARCH plane led to a more reliable metric upgrade, though the difference is less pronounced. This is because Nistér's cost function is based on strong priors on the camera intrinsic parameters, such as zero skew, unit aspect ratio, and an image-centered principal point, that are fully satisfied by our simulated cameras. Thus, both plane initializations converged to Π_∞ most of the time, the QUARCH plane being slightly more reliable. We also performed some tests using a randomly selected plane as an ini-



(a) 5 views

(b) 6 views

(c) 1 pixel noise

Figure 4.7: Benefits of QUARCH and the LMI-constrained optimization when using Nistér's cost function. QUARCH*N is a variant of our algorithm using Nistér's cost function. QUARCH-N and QUARC-N are unconstrained versions of it, with QUARC-N using a QUARC plane as initialization.

tialization. For 4–5 views, this approach very often failed to obtain a metric upgrade (not shown), confirming that it is imperative to start such an optimization from at least a QUARC plane or preferably a QUARCH.

Benefits of the LMI-constrained optimization: We assessed the contribution of the LMI-constrained optimization in our algorithm's performance. We compared the reliability of our algorithm, denoted by QUARCH*M, with that of its unconstrained variant, QUARCH-M. From the results in Figure 4.6, QUARCH*M successfully upgraded several projective reconstructions that had otherwise failed using the unconstrained optimization in QUARCH-M. This shows that enforcing the QUARCH LMIs during the optimization led the QUARCH plane to reliably converge to Π_∞ while avoiding succumbing to a non-QUARCH local minimum of the cost function. In Figure 4.7, using Nistér's cost function, we observed that only

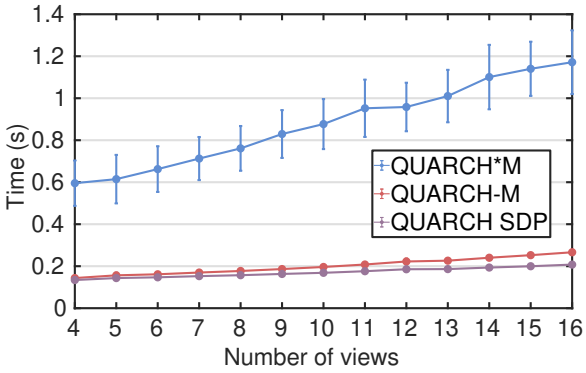
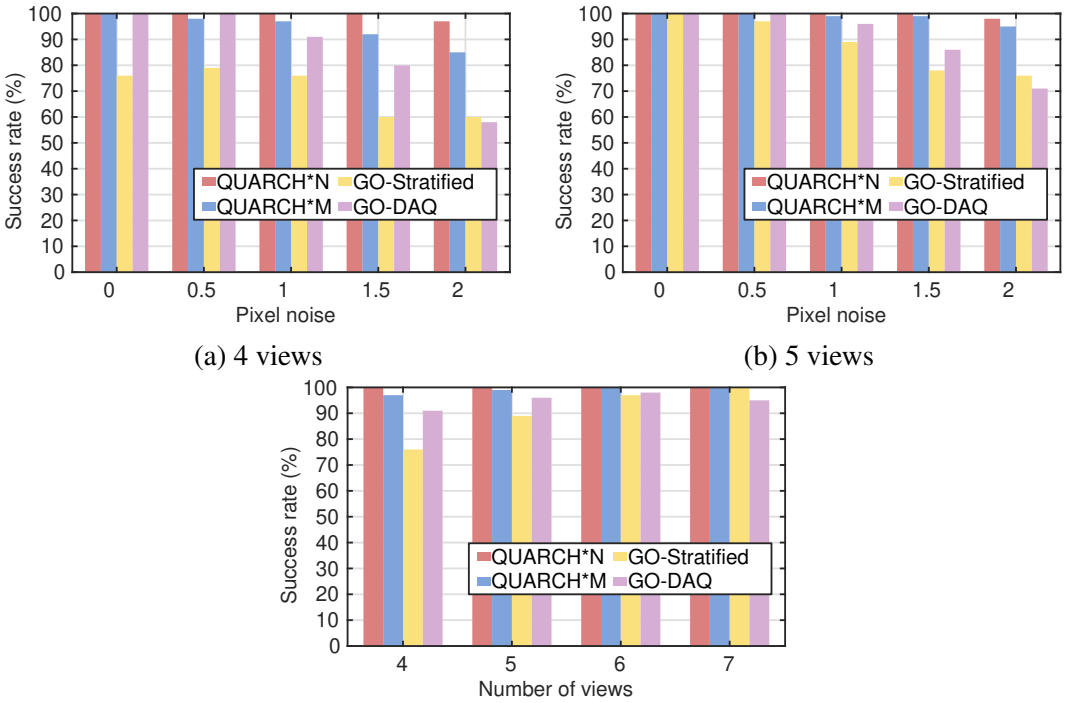


Figure 4.8: Runtime (seconds) with an increasing number of views. The error bars show the mean $\pm$ the standard deviation. The noise level is 1 pixel.

one additional projective reconstruction (for 6 views) was successfully upgraded to metric using the constrained optimization approach, denoted as QUARCH*N. The improvement is less striking, again, due to the strong camera priors used in this cost function that are fully satisfied by our simulated cameras, thus curbing the impact of the LMIs in the optimization. These results show the benefits of enforcing the QUARCH LMIs during the local optimization, particularly for short sequences and when using the modulus constraint polynomials. In terms of the computation cost, the constrained optimization is, of course, more expensive than the unconstrained optimization as it solves an SDP problem at each iteration. Figure 4.8 shows the runtime of QUARCH*M, QUARCH-M, as well as of the QUARCH SDP problem for an increasing number of views and a fixed 1 pixel noise level. Since we use the LMIs only for consecutive views, their number grows linearly with the number of views. This is reflected in Figure 4.8 where the runtime also increases linearly. The constrained optimization generally converged in a few iterations, and it is particularly suited to short sequences. Beyond six views, QUARCH-M sufficed for a successful metric upgrade, and it can be used instead for a speedup.

Comparisons with the state of the art: We compared the results obtained from QUARCH*M and QUARCH*N with those from two globally optimal methods, GO-Stratified [Chandraker et al. 2010] and GO-DAQ [Chandraker et al. 2007]. QUARCH*M and GO-Stratified rely on the modulus constraint, whereas QUARCH*N and GO-DAQ use strong priors on the intrinsic parameters to formulate their cost



(a) 4 views

(b) 5 views

(c) 1 pixel noise

Figure 4.9: Comparisons with the state of the art. GO-Stratified and GO-DAQ are globally optimal methods. GO-Stratified and QUARCH*M rely on the modulus constraint. GO-DAQ and QUARCH*N use priors on the intrinsic parameters.

functions. For GO-Stratified, we computed the solutions for both signs of chirality and picked the one with lower calibration error. We used the authors' implementation[8] of this method. For GO-DAQ, we fixed the relaxation order to 2. Figure 4.9 shows the success rate obtained with these methods. With 4 and 5 views in Figures 4.9(a) and 4.9(b), respectively, QUARCH*M significantly outperformed GO-Stratified, especially for high levels of pixel noise. With such short sequences, the modulus constraint polynomials are likely to admit multiple global solutions for Π_∞. The additional constraints in QUARCH*M then help to isolate the sought solution. GO-Stratified also relies on scene points that may prove unreliable in the presence of noise. QUARCH*N and GO-DAQ fared better than the other two methods for low levels of pixel noise. However, GO-DAQ suffered a drastic decline in the success rate as the noise level increased to 1 pixel or more. The explanation

[8] `https://cseweb.ucsd.edu/~mkchandraker/stratum.html`

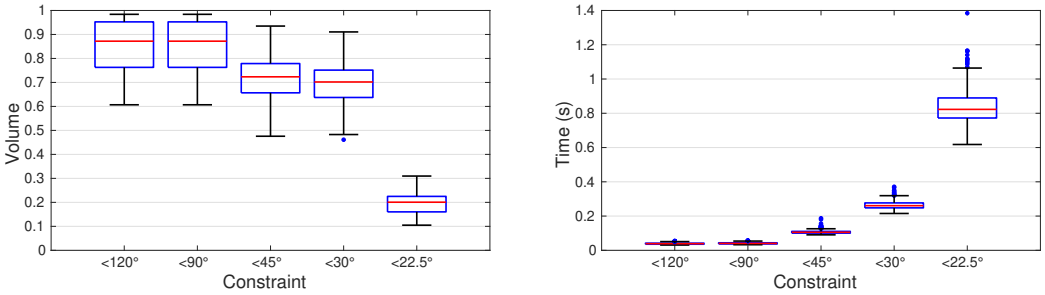


Figure 4.10: Benefits of a specialized QUARCH. Left: volume of the search space for Π_∞ using progressively tighter orientation knowledge. The reference volume (one unit) is computed using the chirality constraints alone. Right: the time taken (seconds) to compute the volume.

for this result is that our simulated cameras approach a known artificial degenerate configuration for estimating the DAQ. This degenerate configuration occurs when all the optical axes pass through a common point and the rank of the DAQ is not enforced [Gurdjos et al. 2009]. Our cameras approach this configuration in the presence of noise. GO-DAQ is then likely to fail as the rank-3 constraint is not earnestly enforced due to numerical scaling issues (the cost function and constraints are unnormalized). QUARCH*N, on the other hand, performed reliably with high levels of noise as well. As the sequence length increased to seven, all the methods barring GO-DAQ succeeded most of the time.

Benefits of a specialized QUARCH: We assessed the benefits of incorporating a tighter orientation knowledge using the non-convex polynomial constraints on Π_∞ to recover a specialized QUARCH. We compared the bounds on Π_∞ obtained using the constraints for $|\theta_{ij}| < 90^\circ/k$, with $k = 1, \ldots, 4$. Recall that for a given k, we use the QUARCH LMIs and the constraints for all $k = 1, \ldots, k-1$ to compute the bounds on Π_∞ (see Section 4.3.3). We solved the involved polynomial optimization problems using the minimum relaxation order, *i.e.* k. For this experiment, we generated synthetic data where the rotation angle between consecutive views was sampled randomly from the range $[10^\circ, 20^\circ]$ to satisfy the assumption $|\theta_{ij}| < 22.5^\circ$ (for $k = 4$). Figure 4.10 (left) shows the resulting volume of the search space for Π_∞ computed from the obtained bounds (the box plots follow the convention in MATLAB). This result is for four views and no noise was added to the pixel coordinates.

The volume is relative to the one obtained with the chirality constraints alone (that corresponds to a unit volume). Figure 4.10 also shows the result obtained by using only the QUARCH LMIs, *i.e*. by exploiting only $|\theta_{ij}| \leq 120^\circ$ between consecutive views. The results show the QUARCH LMIs reduced the search space by as much as 40% compared to the chirality constraints alone, while the median decrease is just over 10%. Using the quadratic inequalities for $|\theta_{ij}| < 90^\circ$ did not result in any noticeable difference. The constraints for $|\theta_{ij}| < 45^\circ$ led to a smaller volume, and the constraints for $|\theta_{ij}| < 30^\circ$ further restricted it. A substantially reduced search space was obtained with the constraints for $|\theta_{ij}| < 22.5^\circ$. The resulting volume is around 20%, in median terms, of the one obtained with chirality constraints alone. Hence, these results show that incorporating a tighter orientation knowledge can significantly narrow the location of Π_∞. The price to pay of using the higher degree polynomial constraints is a longer computation time, as shown in Figure 4.10 (right). The reported results refer to the time required for solving one polynomial optimization problem, *i.e*. each box plot corresponds to 600 measurements as six polynomial optimization problems are solved for each trial. The results in this experiment indicate that the higher degree polynomial constraints are useful to restrict the location of Π_∞. Their behaviour in the presence of noise and impact in our autocalibration algorithm remains to be evaluated through further experiments.

4.5.2 Real image experiments

In this section, we present results on some real image sequences: *fountain-P11*, *Herz-Jesu-P8*, and *Herz-Jesu-P25* from [Strecha et al. 2008], *Vercingetorix* and *Alcatraz water tower* from [Olsson and Enqvist 2011], *Cherub* from 3DFlow[9], and *L'Arbre aux Serpents* from a cultural heritage application. The first three provide the ground truth calibration with focal lengths $f_x = 2759.48$ and $f_y = 2764.16$, principal point $(u, v) = (1520.69, 1006.81)$, and skew $\gamma = 0$, all in pixels. We quantitatively compared our algorithm with existing methods using these sequences. For the remaining sequences, we analyzed the recovered metric reconstructions qualitatively. Our experiments also served to verify the practical applicability of our assumption that $|\theta_{ij}| \leq 120^\circ$ for consecutive views. For all the sequences, we com-

[9] `https://www.3dflow.net/3df-zephyr-reconstruction-showcase/`

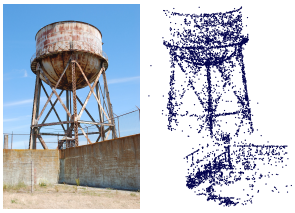
Sequence	Method	Δf	Δuv	$\Delta\gamma$	Time (s)
fountain-P11	QUARCH*M	1.91	4.01	0.99	2.71
	QUARC-M	2.44	4.30	0.99	0.09
	QUARCH*N	42.92	28.29	0.71	1.47
	QUARC-N	43.73	28.61	0.69	0.10
	GO-DAQ	76.15	31.92	0.10	1.27
	GO-Stratified	12.64	9.75	1.17	449.47
Herz-Jesu-P8	QUARCH*M	53.49	78.68	1.56	1.32
	QUARC-M	4114.66	101.16	586.29	0.07
	QUARCH*N	83.61	33.93	1.24	1.90
	QUARC-N	76.81	34.22	1.21	0.09
	GO-DAQ	66.10	33.84	0.27	1.85
	GO-Stratified	2552.62	1006.05	132.85	154.74
Herz-Jesu-P25	QUARCH*M	34.89	23.71	2.52	2.04
	QUARC-M	34.75	23.70	2.55	0.45
	QUARCH*N	59.04	31.44	1.40	2.08
	QUARC-N	2812.61	185.73	21.86	0.46
	GO-DAQ	4.40	33.90	0.60	1.60
	GO-Stratified	52.94	32.16	1.69	893.76

Table 4.1: Quantitative evaluation. Autocalibration results on the real image sequences from [Strecha et al. 2008].

puted the projective reconstructions using P2SfM [Magerand and Bue 2018] with feature matches obtained using COLMAP [Schönberger and Frahm 2016].

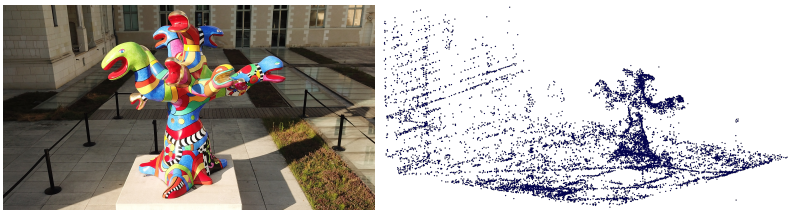
Quantitative evaluation: Table 4.1 reports the calibration errors from the tested methods on the three sequences with known ground truth calibration. From these results, we observe that an erroneous calibration was obtained using QUARC-M and GO-Stratified on the *Herz-Jesu-P8* sequence, and using QUARC-N on the *Herz-Jesu-P25* sequence. Upon inspection, the corresponding reconstructions failed to achieve a metric upgrade and remained projectively distorted. Except for these failures, all the methods otherwise led to a calibration close to the ground truth and thereby to a successful metric upgrade. From our observations, the reconstruction quality was primarily affected by the focal length and skew errors. Note that QUARCH*M succeeded on the *Herz-Jesu-P8* sequence, whereas the other two

(a) *Cherub*



(b) *Alcatraz water tower*

(c) *Vercingetorix*



(d) *L'Arbre aux Serpents*

Figure 4.11: Qualitative evaluation. Sample images and metric 3D reconstructions obtained using QUARCH*M. (d) Image courtesy of Renato Saleri. *L'Arbre aux Serpents de Niki de Saint Phalle* © Musées d'Angers, Niki Charitable Art Foundation.

methods that also rely on the modulus constraint, *i.e*. QUARC-M and GO-Stratified, failed. This corroborates our results with the synthetic data, that the QUARCH LMIs aid in reliably locating Π_∞ and that GO-Stratified often fails for short sequences. From the computation time reported in Table 4.1, the methods based on unconstrained local optimization were considerably faster than the others. The constrained optimization-based methods solve an SDP problem at each iteration and are therefore slower than their unconstrained counterparts. While GO-DAQ took a similar amount of time as QUARCH*M and QUARCH*N, GO-Stratified was significantly slower for all the sequences.

Qualitative evaluation: We visually evaluated the 3D reconstructions obtained with our method using four image sequences: *Cherub*, *Vercingetorix*, *Alcatraz water tower*, and *Arbre aux Serpents*. These sequences have 65, 69, 173, and 154 images, respectively. Their corresponding projective reconstructions contained a maximum of 66 cameras. The last sequence, *Arbre aux Serpents*, was captured using a camera drone for the purpose of digital preservation of the imaged sculpture. The drone operator did not have any knowledge of the present work and the involved

orientation constraints. Figure 4.11 shows the metric 3D reconstructions obtained using QUARCH*M. The recovered reconstructions closely resemble the captured scenes. Similar metric reconstructions were obtained using QUARCH*N. We observed that several points were poorly estimated in these reconstructions. These led to failures with GO-Stratified, as it relies on the scene points. These results also confirm the applicability of the assumption of $|\theta_{ij}| \leq 120^\circ$ for consecutive views. In fact, this assumption might as well be extended to every other view in these sequences.

4.6 Conclusion

In this chapter, we showed how a vague knowledge of the camera orientation can be exploited in stratified autocalibration. Assuming a moving camera with constant intrinsic parameters, we introduced new pairwise constraints on Π_∞ based on a vague knowledge of the relative camera orientation angle. We showed that Π_∞ belongs to one of two convex sets depending on whether $|\theta_{ij}| \leq 120^\circ$ or $|\theta_{ij}| \geq 120^\circ$, with θ_{ij} the orientation angle between a camera pair. These two convex sets are defined by LMIs, and they describe the relations between Π_∞ and new geometric objects, the hodographs of the horopter. We then showed that a tighter orientation knowledge of $|\theta_{ij}| < 90^\circ/k$, for a positive integer k, can be exploited to further constrain Π_∞. The resulting constraints are non-convex polynomial inequalities of degree $2k$. They were derived algebraically using the Hurwitz stability criterion. Unlike the convex constraints, their geometric interpretation remains to be uncovered. Based on the new orientation-based constraints, we showed the existence of a new quasi-affine reconstruction of a scene that we called a QUARCH. We used a QUARCH in a stratified camera autocalibration algorithm as an intermediate step towards recovering an affine reconstruction from a projective one. We also proposed an LMI-constrained Levenberg-Marquardt method that ensures that the QUARCH constraints are satisfied during the local optimization used to locate Π_∞ in our algorithm.

Our experiments showed the benefits of a QUARCH and the orientation-based constraints in reliably recovering a metric reconstruction from a projective one. These experiments primarily evaluated the convex constraints in our algorithm. A

preliminary test using the non-convex constraints indicated the potential benefits of incorporating tighter orientation knowledge. Further experiments are necessary to evaluate the impact of these constraints in our algorithm, and this forms the near-term future work. The orientation-based constraints can also be useful in other auto-calibration methods. Search-based methods [Hartley et al. 1999; Chandraker et al. 2010], in particular, can benefit from tighter bounds on Π_∞. Our contributions can also be potentially beneficial in other problems. For instance, similar to previous work on exploiting oriented constraints [Werner and Pajdla 2001], the QUARCH constraints could be useful in the feature matching phase to eliminate infeasible realizations of a scene. Finally, the LMI-constrained optimization method can be applied more broadly as LMIs appear in numerous computer vision problems [Speciale et al. 2017; Speciale et al. 2018].

5 Conclusion

In this final chapter, we discuss the main contributions of this thesis and some perspectives for future research.

5.1 Discussion of contributions

In this thesis, we presented some contributions in uncalibrated 3D computed vision. In particular, we investigated camera knowledge available in the image-based 3D reconstruction scenario that has not been fully exploited in camera autocalibration. Our contributions address stratified autocalibration, where reliably and accurately locating the plane at infinity is challenging. We derived new constraints on this plane by exploiting priors on the camera geometry and motion. First, we exploited partial knowledge of the camera geometry, specifically that it has square pixels. We derived a new quartic equation in the plane at infinity assuming such a camera with constant intrinsic parameters. Compared to existing constraints under similar assumptions, this polynomial has the advantages of being concentrated to only three unknowns, a reasonably low degree polynomial, and a pairwise constraint. Our experiments showed the benefits of using the new polynomial in reliably estimating the plane at infinity.

Second, we exploited a vague knowledge of the relative camera orientation. We thoroughly studied the constraints that can be derived when a vague knowledge of the relative camera orientation angle between camera pairs is available. One set of constraints is convex and can be exploited when the relative orientation angle between camera pairs is known to be either under or over 120°. These constraints have a geometric interpretation through the hodographs of the horopter. They also define the QUARCH stratum. The non-convex constraints exploit tighter orientation knowledge and can be used in addition to obtain a specialized QUARCH. We

proposed a Levenberg-Marquardt method subject to Linear Matrix Inequality constraints that is tailored to these constraints. Our experiments showed an improvement in reliably obtaining a metric reconstruction when using the orientation-based constraints, and incorporating tighter knowledge could further improve it.

We devised autocalibration methods dedicated to the different constraints and evaluated the benefits of incorporating the additional camera knowledge. Our contributions, however, are complementary and can be used together. For instance, the quartic equation can be used in the cost of the QUARCH method. Similarly, the orientation-based constraints can be introduced (as polynomial inequalities) in the camera geometry-based method. The LMI-constrained optimization can be used in both cases for local refinement. These constraints may be used in other autocalibration methods and potentially in other computer vision problems as well. We briefly discussed such potential applications in Section 4.6.

5.2 Perspectives

Finally, we discuss some potential directions for future research.

Exploiting external sensor information: Images are increasingly captured using a smartphone/tablet or a camera mounted on a robot (*e.g.* a drone), where other sensors, such as an IMU, are also available. An IMU combines a gyroscope and an accelerometer and provides partial motion information in the sensor reference frame. In particular, the gyroscope measures the angular velocity, whereas the accelerometer measures the gravity direction. The measurements from the gyroscope are known to be noisy, and integrating them to obtain the relative orientation induces a drift. The measurements from the accelerometer, on the other hand, are generally more reliable. To use the accelerometer readings as camera motion information, the rigid transformation between the camera and IMU must be known. Fortunately, for smartphones/tablets, the relative transformation is standard across devices and can thus be assumed to be known. Knowledge of the gravity direction has been exploited before in some computer vision problems [Kukelova et al. 2010; Ding et al. 2020]. However, it is yet to be fully exploited in camera autocalibration.

We have studied this problem and obtained some promising preliminary results that we plan to develop further.

Exploiting semantic priors: Thanks to recent advances in machine learning, particularly deep learning [LeCun et al. 2015], higher-level features, such as objects, regions, and their semantic labels, can now be efficiently obtained across different modalities, including images, videos, RGB-D, and 3D data. Exploring the use of such semantic information in camera autocalibration and other geometric computer vision problems is an interesting direction for future research. Semantic priors may be especially useful when low-level features cannot be easily obtained, for instance, due to a lack of texture or severe appearance changes in the scene, or when dealing with different modalities. Pursuing research in this direction entails several potential challenges, including identifying useful semantic priors, mathematically modeling the optimization problem to account for such priors, and dealing with erroneous semantic information. Some relevant recent works in computer vision and robotics address registration [Paudel et al. 2017; Speciale et al. 2018], visual localization [Toft et al. 2018], visual odometry [Lianos et al. 2018], and SLAM [Rosinol et al. 2020].

A Optimization tools

This appendix provides an overview of some optimization tools that are widely used in this thesis. The aim is to summarize the relevant results. In Appendices A.1 and A.2, we provide the basics of Linear Matrix Inequalities and semidefinite programming, respectively. In Appendix A.3, we give an overview of Lasserre's polynomial optimization method.

A.1 Linear Matrix Inequalities

A *Linear Matrix Inequality* (LMI) is a constraint on a real vector $\mathsf{x} = (\mathsf{x}_1, \ldots, \mathsf{x}_n)^\top$ such that $\mathsf{F}(\mathsf{x}) \succeq 0$, where

$$\mathsf{F}(\mathsf{x}) \;=\; \mathsf{F}_0 + \sum_{i=1}^{n} \mathsf{x}_i \, \mathsf{F}_i \tag{A.1}$$

is an affine function of x involving symmetric matrices $\mathsf{F}_0, \mathsf{F}_1, \ldots, \mathsf{F}_n$. The LMI $\mathsf{F}(\mathsf{x}) \succeq 0$ means that $\mathsf{F}(\mathsf{x})$ is positive semidefinite. The LMI may also be strict, *i.e.* $\mathsf{F}(\mathsf{x}) \succ 0$, in which case $\mathsf{F}(\mathsf{x})$ is positive definite. The LMI (A.1) is a convex constraint on x, *i.e.* the set $\{\mathsf{x} : \mathsf{F}(\mathsf{x}) \succ 0\}$ is convex. Convex quadratic inequalities can be reformulated into LMIs using the Schur complement lemma [Boyd and Vandenberghe 2004, Sec. A.5.5].

Lemma A.1 (Schur complement lemma)**.** *Consider a real symmetric matrix* D *that is partitioned into matrix blocks as:*

$$\mathsf{D} = \begin{bmatrix} \mathsf{A} & \mathsf{B} \\ \mathsf{B}^\top & \mathsf{C} \end{bmatrix}, \tag{A.2}$$

where A *is a symmetric matrix. If* A *is nonsingular, the matrix,*

$$\mathsf{S} = \mathsf{C} - \mathsf{B}^\top \mathsf{A}^{-1} \mathsf{B}, \tag{A.3}$$

is defined as the Schur complement of A *in* D*. The positive (semi)definiteness of* D *is characterized by the following conditions:*

(i) $\mathsf{D} \succ 0$ *if and only if* $\mathsf{A} \succ 0$ *and* $\mathsf{S} \succ 0$*.*

(ii) If $\mathsf{A} \succ 0$*, then* $\mathsf{D} \succeq 0$ *if and only if* $\mathsf{S} \succeq 0$*.*

Further details on LMIs can be found in [Boyd et al. 1994].

A.2 Semidefinite programming

A *Semidefinite Programming* (SDP) problem is a convex optimization problem of the form,

$$\begin{aligned} \min_{\mathsf{x}} \quad & \mathsf{c}^\top \mathsf{x} \\ \text{s.t.} \quad & \mathsf{F}(\mathsf{x}) \succeq 0, \end{aligned} \tag{A.4}$$

where x is the sought vector of variables, c and x are both real vectors, and the matrix $\mathsf{F}(\mathsf{x})$ is defined in (A.1). In other words, an SDP problem minimizes (or maximizes) a linear objective function subject to LMI constraints. SDPs generalize *Linear Programming* (LP) problems. An important property of SDP problems is that they can be solved efficiently in polynomial time using interior-point methods [Boyd and Vandenberghe 2004; Wright 2005]. These methods are conveniently implemented in several off-the-shelf solvers, such as MOSEK[1] and CVXOPT[2].

A.3 Polynomial optimization

Consider the, generally nonconvex, polynomial optimization problem of the form,

$$\begin{aligned} \min_{\mathsf{x}} \quad & f(\mathsf{x}) \\ \text{s.t.} \quad & g_i(\mathsf{x}) \geq 0, \qquad i = 1, 2, \ldots, m, \end{aligned} \tag{A.5}$$

[1] `https://docs.mosek.com/latest/intro/index.html`
[2] `http://cvxopt.org/userguide/intro.html`

where $f(\mathsf{x})$ and $g_1(\mathsf{x}), \ldots, g_m(\mathsf{x})$ are all multivariate scalar polynomials in x, an n-vector. In [Lasserre 2001], it has been shown that problem (A.5) can be solved through a hierarchy of convex LMI relaxations of increasing order $d = 1, 2, \ldots,$ yielding monotonically non-decreasing lower bounds on the original problem and converging to its global minimum. At each order d, the problem is linearized and the following surrogate SDP is solved,

$$\begin{aligned} \min_{\mathsf{y}} \quad & f(\mathsf{y}) \\ \text{s.t.} \quad & \mathsf{M}_{d-1}(g_i(\mathsf{y})) \succeq 0, \qquad i = 1, 2, \ldots, m, \\ & \mathsf{M}_d(\mathsf{y}) \succeq 0. \end{aligned} \tag{A.6}$$

Denoting by $\mathsf{v}_d(\mathsf{x})$ the vector containing all monomials up to degree d, the linearization (A.6) is obtained from (A.5) by:

(i) replacing the scalar polynomial constraints $g_i(\mathsf{x}) \geq 0$ by matrix constraints $g_i(\mathsf{x})\mathsf{v}_{d-1}(\mathsf{x})\mathsf{v}_{d-1}(\mathsf{x})^\top \succeq 0$,

(ii) introducing the constraint $\mathsf{v}_d(\mathsf{x})\mathsf{v}_d(\mathsf{x})^\top \succeq 0$,

(iii) substituting the monomials $\mathsf{x}_1^{k_1}\mathsf{x}_2^{k_2}\ldots\mathsf{x}_n^{k_n}$ up to degree $2d$ by new variables $y_{k_1 k_2 \ldots k_n}$ in the objective and the resulting constraints. In (A.6), the linearized matrices $g_i(\mathsf{x})\mathsf{v}_{d-1}(\mathsf{x})\mathsf{v}_{d-1}(\mathsf{x})^\top$ and $\mathsf{v}_d(\mathsf{x})\mathsf{v}_d(\mathsf{x})^\top$ are respectively denoted by $\mathsf{M}_{d-1}(g_i(\mathsf{y}))$ and $\mathsf{M}_d(\mathsf{y})$, and they are known as the moment matrices.

Note that linearization is possible at a starting relaxation order d, in which no monomial in the problem is of a degree higher than $2d$. At a given relaxation order, the global optimality of the solution(s) can be certified by a rank condition on the moment matrices. Though the relaxation order at which global optima are obtained is not known in advance, it has been observed that this occurs at a relatively small order of relaxation for many problems [Henrion and Lasserre 2005]. In practice, even if the global optimality is not certified at the first order of relaxation, the corresponding solution can be used as an approximation to save the computational effort of solving the problem at higher orders of relaxation. Lasserre's method, often referred to as Lasserre's hierarchy, is conveniently implemented in GloptiPoly [Henrion and Lasserre 2003; Henrion et al. 2009]. It has been used to solve several poly-

nomial optimization problems in computer vision [Chandraker et al. 2007; Kahl and Henrion 2007; Magerand et al. 2012; Bugarin et al. 2016; Parashar et al. 2017; Trutman et al. 2022]. The interested reader may refer to [Lasserre 2001; Lasserre 2015] for more details on this method.

B Line projection matrix

This appendix provides the expression of a line projection matrix in terms of its point projection counterpart. This expression is given in (B.3).

A line L in 3-space is represented by its Plücker coordinates encapsulated in a 6-vector L. The coordinates of L may be obtained from those of two points, say Y and Z, on this line. Suppose Y and Z have the coordinates, $\mathsf{Y} \simeq (\mathsf{y}^\top, \mathsf{y}_4)^\top$ and $\mathsf{Z} \simeq (\mathsf{z}^\top, \mathsf{z}_4)^\top$, respectively. L is constructed by stacking two 3-vectors, the so-called moment vector $\mathsf{y} \times \mathsf{z}$ and displacement vector $\mathsf{y}_4\mathsf{z} - \mathsf{z}_4\mathsf{y}$. Arranged as in [Bartoli and Sturm 2004], $\mathsf{L}^\top$ is given by,

$$\mathsf{L}^\top \simeq (\mathsf{y}^\top \times \mathsf{z}^\top,\, \mathsf{y}_4\mathsf{z}^\top - \mathsf{z}_4\mathsf{y}^\top). \tag{B.1}$$

A camera is generally described by its point projection matrix P, where P is a 3×4 matrix, that maps a scene point Y to its pixel projection PY. A similar relationship, governed by a 3×6 matrix M, was shown in [Ponce et al. 2005] to exist between a scene line L and its image,

$$\mathsf{l} \simeq \mathsf{ML}, \tag{B.2}$$

where L and l are the coordinates of L and its image, respectively. A line projection matrix M is related to its point projection counterpart P, and each one can be obtained from the other. Indeed, if P is partitioned as $\mathsf{P} = [\,\mathsf{H} \mid \mathsf{e}\,]$, with H a 3×3 matrix, then,

$$\mathsf{M} \simeq [\,\det(\mathsf{H})\mathsf{H}^{-\top} \mid -[\mathsf{e}]_\times \mathsf{H}\,]. \tag{B.3}$$

This expression of M can be obtained by using the fact that the image of a scene line L is a line containing the image projections of two points Y and Z on L, *i.e.* $\mathsf{ML} \simeq \mathsf{PY} \times \mathsf{PZ}$. Since P can be computed from point correspondences and it is related to M via (B.3), line correspondences are not needed to estimate M.

C Résumé

Alors que l'œil humain perçoit le monde en trois dimensions (3D) sans effort, notamment grâce à des stimuli visuels, cette capacité est loin d'être triviale pour un ordinateur. En utilisant une caméra pour les données visuelles, l'information 3D est perdue lors de la projection d'une scène 3D sur une image 2D. Retrouver la structure 3D de la scène à partir de plusieurs images est un problème majeur dans le domaine de la vision par ordinateur. Les nombreux travaux de recherche sur ce problème ont conduit au développement d'une base et d'un cadre théoriques mais aussi d'algorithmes pour le calcul de structures 3D à partir d'images bidimensionnelles calibrées et non calibrées. Dans cette thèse, nous nous intéressons au problème de reconstruction 3D non calibrée, où aussi bien les géométries interne qu'externe de la caméra sont inconnues. Nous présentons des contributions au problème dit d'autocalibrage des caméras qui constitue une étape aussi cruciale que difficile dans le processus de reconstruction 3D à partir d'images non calibrées.

C.1 Contexte et motivation

Modélisation de scènes en 3D : Les modèles 3D de scènes et d'objets dans le monde font désormais partie intégrante de diverses applications. En réalité augmentée (RA), notre vue du monde physique est enrichie par des modèles virtuels qui se superposent à la scène. Par exemple, l'application mobile Snapchat[1] intègre une fonction de réalité augmentée appelée "Local Lenses". Grâce à cette fonctionnalité, les utilisateurs peuvent s'engager dans un contenu virtuel, tel qu'un monde en RA partagé à travers la rue Carnaby à Londres. Local Lenses dépend d'une reconstruction 3D de la rue et des bâtiments qui a été réalisée à partir d'images, y

[1] `https://ar.snap.com/lens-studio`

compris des photos publiques, capturées autour de la rue. En robotique, construire et maintenir une carte 3D d'un environnement inconnu est souvent cruciale pour un robot afin d'interagir et d'effectuer des tâches dans son environnement. Ces tâches peuvent inclure l'inspection et la surveillance [MAURER et al. 2017], les opérations de recherche et de sauvetage [DELMERICO et al. 2019], et les véhicules autonomes [BRESSON et al. 2017]. Par exemple, la voiture autonome de Cruise[2] utilise un scan laser 3D de son environnement, entre autres informations de capteurs, pour naviguer dans son environnement. Dans le patrimoine numérique, les modèles 3D de sites et d'objets culturels offrent un moyen de les conserver numériquement. Ces modèles numériques peuvent ensuite être utilisés pour l'analyse, tel que la détection des changements, et pour des visites virtuelles interactives. Par exemple, dans une visite interactive de la Villa de Diomède [DESSALES 2020] à Pompéi, un rendu de l'état historique du site peut être superposé sur un modèle 3D de son état actuel. Le modèle 3D a été obtenu à partir d'images et de scans laser du site. Il existe de nombreuses autres applications de la modélisation 3D : dans les jeux [STATHAM 2020], la police scientifique [GALANAKIS et al. 2021] et la paléontologie [FALKINGHAM et al. 2020].

Détection 3D : Les méthodes d'acquisition 3D sans contact sont généralement classées comme actives ou passives [LIU et al. 2020]. Les méthodes actives interfèrent avec la scène par un rayonnement électromagnétique contrôlé au moyen de projecteurs de lumière structurée (Microsoft Kinect V1, Intel RealSense SR300) ou de scanners à temps de vol (Microsoft Kinect V2, Intel RealSense L515). Dans l'approche par lumière structurée, un motif est projeté sur la scène, et une ou plusieurs caméras (capteurs CCD/CMOS) sont employées pour imager la scène éclairée. Le motif projeté étant connu, les correspondances entre celui-ci et les projections sur l'image peuvent être établies aisément, et la structure 3D est obtenue par triangulation. Dans l'approche par temps de vol, la lumière est émise sur la scène, et la structure 3D est déterminée plus directement en mesurant sa réflexion. Contrairement aux approches actives, les méthodes passives reposent uniquement sur l'usage d'une caméra pour imager la scène en 2D dans des conditions ambiantes. La structure 3D est déterminée en traitant les images obtenues par des méthodes de photo-

[2] https://www.getcruise.com/

grammétrie/vision par ordinateur. Le choix de la méthode d'acquisition et des capteurs utilisés pour une application spécifique dépend de plusieurs critères, comme la scène (sa taille, sa forme, sa texture), les conditions d'acquisition (éclairage ambiant, accessibilité), la qualité des modèles et le coût d'acquisition. En général, les caméras sont plus flexibles qu'un système actif grâce à leur taille compacte, leur efficacité énergétique et leur simplicité d'installation et d'utilisation. L'omniprésence des appareils photo (par exemple dans les smartphones) et les progrès réguliers de leur qualité d'image ont favorisé leur adoption pour l'acquisition de données 3D.

Reconstruction 3D à partir d'images : Le problème de la reconstruction d'une scène en 3D à partir de plusieurs images a été largement étudié dans le domaine de la vision par ordinateur, et trouve son origine dans la photogrammétrie [HARTLEY et MUNDY 1993 ; STURM 2011]. Appelé Structure-from-Motion (SfM) en vision par ordinateur, il s'agit de récupérer la structure de la scène en 3D et le mouvement de la caméra à partir de plusieurs images. Les relations géométriques entre plusieurs vues, ou géométrie multi-vues, est désormais bien établie [FAUGERAS et al. 2001 ; HARTLEY et ZISSERMAN 2004]. Une approche SfM implique la résolution de plusieurs problèmes, y compris la détermination des points caractéristiques et de leurs correspondances dans les images [LOWE 2004], le calcul des poses relatives de caméra [LONGUET-HIGGINS 1981 ; NISTÉR 2004a], et l'extraction de la structure de la scène en 3D par triangulation [HARTLEY et STURM 1997]. Une estimation robuste des paramètres de la caméra à l'aide de méthodes telles que RANSAC [FISCHLER et BOLLES 1981] et un raffinement conjoint des paramètres estimés et de la structure 3D par ajustement de faisceaux [TRIGGS et al. 1999] sont très importants dans cette approche. Grâce aux progrès réalisés dans ces problèmes, des systèmes automatisés de SfM ont été développés, tels que Photo Tourism [SNAVELY et al. 2008], qui, à partir d'un ensemble d'images, calculent les poses de l'appareil photo et une reconstruction 3D éparse (un nuage de points) de la scène imagée. L'approche SfM est devenue une technologie mature et il existe aujourd'hui plusieurs progiciels de SfM, y compris des logiciels commerciaux (Agisoft Meta-

Shape[3], 3DF Zephyr[4], RealityCapture[5]) et non commerciaux (VisualSfM [C. WU 2013], OpenMVG [MOULON et al. 2016], COLMAP [SCHÖNBERGER et FRAHM 2016]). Une reconstruction 3D éparse obtenue par SfM peut être traitée pour obtenir une représentation plus riche de la scène, comme par exemple un nuage de points dense ou un maillage texturé. Certains logiciels SfM intègrent cette fonctionnalité ou peuvent être utilisés avec d'autres logiciels dédiés [FURUKAWA et al. 2010]. La disponibilité et la simplicité d'utilisation des logiciels SfM, ainsi que la qualité des reconstructions 3D obtenues, justifient leur utilisation pour la modélisation 3D.

SfM calibré : Les logiciels de reconstruction 3D à partir d'images (SfM) font généralement l'hypothèse que les paramètres intrinsèques de la caméra sont connus. Ces paramètres sont ceux du modèle décrivant la géométrie interne de la caméra (longueur focale, centre optique, etc.). Avec la pose de la caméra, c'est-à-dire les paramètres extrinsèques de la caméra, ils définissent le modèle de projection perspective considéré généralement en vision par ordinateur pour décrire les caméras conventionnelles. Les paramètres intrinsèques peuvent être obtenus en calibrant la caméra. La méthode classique de calibrage [ZHANG 2000] est la plus répandue. Elle consiste à prendre plusieurs images d'un motif planaire connu. Malgré le développement de systèmes guidés [PENG et STURM 2019], cette procédure est plutôt laborieuse et suppose une certaine expertise technique pour obtenir un calibrage précis. Dans les applications SfM, la caméra est typiquement calibrée avant l'acquisition des images et son zoom et sa mise au point sont maintenus fixes pour s'assurer que la calibration interne de la caméra reste inchangée. Si la calibration interne de la caméra change, les paramètres intrinsèques doivent être réestimés, par exemple en calibrant la caméra à nouveau. Les approches SfM peuvent souvent bénéficier des métadonnées Exif (Exchangeable image file format) pour approximer les paramètres intrinsèques de la caméra et ainsi éviter l'étape laborieuse de la calibration hors ligne. Les étiquettes Exif sont intégrées dans certains formats de fichiers d'images, comme JPEG. Ils contiennent des informations sur la caméra, notamment sa longueur focale. Les autres paramètres intrinsèques sont généralement

[3] https://www.agisoft.com/
[4] https://www.3dflow.net/3df-zephyr-photogrammetry-software/
[5] https://www.capturingreality.com/

approximés sur la base d'une heuristique. L'approche SfM s'appuie ensuite sur un ajustement de faisceaux : une optimisation non linéaire impliquant les paramètres de la caméra et la structure de la scène pour converger vers un calibrage précis. Cependant, les métadonnées Exif ne sont pas toujours disponibles. Dans ce cas, les logiciels de SfM ont recours à des heuristiques peu rigoureuses pour initialiser la longueur focale ou demandent les paramètres intrinsèques à l'utilisateur.

SfM non calibré : Contrairement au SfM calibré, les paramètres intrinsèques de la caméra ne sont pas connus dans une approche non calibrée. Cependant, avec des correspondances entre images non calibrées, seule une structure à une ambiguïté projective près peut être obtenue. Cette reconstruction, dite projective, diffère de la vraie représentation métrique de la scène par une transformation projective. Une reconstruction projective ne préserve pas les caractéristiques métriques, comme les angles et les proportions de longueur, mais préserve des invariants projectifs, y compris la colinéarité et la coplanarité. La scène semble donc distordue. Dans la plupart des applications, une représentation plus fiable de la scène observée est souhaitée. Une reconstruction projective peut être transformée en sa contrepartie métrique en déterminant la transformation projective entre les deux reconstructions. Cette transformation comprend les paramètres intrinsèques de la caméra et leur détermination à partir de plusieurs images non calibrées constitue le problème de l'autocalibrage de la caméra en vision par ordinateur. L'approche SfM non calibrée avec autocalibrage est plus flexible que l'approche calibrée puisque l'étape de calibrage manuel de la caméra n'est pas effectuée. Par conséquent, ce problème a été largement étudié dans la littérature [FUSIELLO 2000].

L'autocalibrage des caméras : L'autocalibrage est le processus d'estimation des paramètres intrinsèques d'une caméra à partir de plusieurs images non calibrées d'une scène inconnue. Ce processus ne nécessite pas la présence d'un objet physique de calibrage. En revanche, il se base sur l'omniprésence d'un objet virtuel : la conique absolue (AC). L'AC est une conique particulière située sur le plan à l'infini et qui se projette sur le plan image sous la forme d'une conique imaginaire dont la représentation matricielle intègre les paramètres intrinsèques de la caméra. Les paramètres intrinsèques peuvent donc être déduits en localisant l'AC ou son image

(IAC). Cependant, la flexibilité de travailler directement avec des images non calibrées entraîne la nécessité de résoudre le problème difficile d'autocalibrage. En effet, les équations d'autocalibrage sont non linéaires et difficiles à résoudre de manière fiable et efficace. En particulier, cette difficulté est d'autant plus importante lorsqu'il s'agit, en préalable, de localiser le plan à l'infini. Une fois que ce dernier est localisé, l'IAC, et donc les paramètres intrinsèques, peuvent être déterminés par la résolution d'équations linéaires. Pour cette raison, certaines méthodes d'autocalibrage estiment d'abord le plan à l'infini, puis estiment les paramètres intrinsèques dans une seconde étape. Cette approche en deux étapes est appelée l'autocalibrage stratifié. Contrairement à une approche stratifiée, l'autocalibrage direct consiste à estimer simultanément le plan à l'infini et les paramètres intrinsèques. Les méthodes directes se basent en général sur des quadriques virtuelles spéciales, la quadrique absolue duale (DAQ) ou encore le complexe quadratique absolu (ALQ), qui encodent le plan à l'infini ainsi que l'IAC ou son image duale (DIAC). Dans les deux cas, toutes les méthodes d'autocalibrage se basent sur une connaissance a priori de la caméra, de sa géométrie interne ou de son mouvement, pour estimer les paramètres intrinsèques. Par exemple, plusieurs méthodes traitent le cas d'une caméra en mouvement dont les paramètres intrinsèques sont constants mais inconnus. D'autres méthodes utilisent la connaissance du mouvement de la caméra, comme le mouvement planaire dans le cas d'un véhicule terrestre. En général, il est conseillé d'utiliser autant d'informations que possible sur la caméra [HARTLEY et ZISSERMAN 2004, Sec. 19.11] afin de relever les difficultés intrinsèques à l'autocalibrage.

C.2 Portée de la thèse

Dans cette thèse, nous nous intéressons au problème de reconstruction 3D d'une scène rigide à partir de plusieurs images prises par une caméra perspective non calibrée. En particulier, nous examinons l'utilisation de connaissances a priori, souvent disponibles dans ce contexte, mais qui n'ont pas encore été entièrement exploitées pour résoudre le problème d'autocalibrage des caméras. Ces connaissances, présentées dans ce qui suit, sont issues de la géométrie interne de la caméra ou de son mouvement lors d'acquisition d'images.

A priori sur la géométrie interne du capteur : Même si les paramètres intrinsèques de la caméra ne sont pas connus dans l'approche SfM non calibrée, nous pouvons généralement faire des hypothèses raisonnables sur certains d'entre eux. Leurs notions physiques nous permettent de distinguer ces paramètres. La longueur focale est la distance entre le centre optique et le plan de l'image. Elle varie en fonction du zoom ou de la mise au point et n'est pas connue en l'absence de calibration. Le point principal donne les coordonnées de l'intersection de l'axe optique avec le plan de l'image. Ce point n'est pas toujours au centre de l'image en raison de défauts d'alignement optique et d'autres défauts de fabrication. Sa position varie également avec la longueur focale. Ainsi, le point principal n'est pas non plus connu en général. Les deux autres paramètres, le facteur de décentrage (skew factor) et le rapport d'aspect d'un pixel, définissent la forme du pixel. Ces paramètres sont assez stables même si le zoom ou la mise au point changent. De plus, la plupart des caméras modernes ont généralement des pixels carrés (ou presque), c'est-à-dire que le facteur de décentrage est nul et que le rapport d'aspect est unitaire. Nous pouvons donc supposer que ces deux paramètres sont connus en pratique.

A priori sur l'acquisition d'images : Quand on capture des images pour la reconstruction 3D, le déplacement de la caméra est contraint implicitement par le besoin de déterminer les correspondances. En effet, les images doivent suffisamment se chevaucher pour établir les correspondances. Dans ce but, les logiciels SfM conseillent typiquement de capturer des images dans une boucle autour de la scène, en changeant légèrement le point de vue[6]. En général, ce type de mouvement de caméra est réalisé par défaut dans une vidéo. Une sorte de connaissance, bien que vague, du mouvement de la caméra est donc implicitement disponible dans la procédure d'acquisition de l'image.

Une connaissance partielle de la géométrie ou du mouvement de la caméra est en principe disponible dans le contexte SfM non calibré, au vu de la description ci-dessus. Nous avons été assez précis sur la connaissance partielle de la géométrie de la caméra, mais la connaissance du mouvement n'a pas encore été explicitée.

[6] https://www.3dflow.net/technology/documents/photogrammetry-how-to-acquire-pictures/

Dans cette thèse, notre but est d'identifier et d'exploiter cette connaissance partielle de la caméra pour l'autocalibrage de la caméra. Nous présentons ici en détail notre contribution à ce sujet dans la section suivante.

C.3 Contributions

Nous présentons ci-après les contributions de cette thèse concernant l'exploitation des connaissances partielles de la géométrie et du mouvement de la caméra dans l'autocalibrage de la caméra.

C.3.1 Exploitation de connaissances partielles de la géométrie de la caméra

Suite à la discussion dans la Section C.2, la majorité des caméras modernes ont des pixels carrés. Ces caméras sont communément appelées caméras à plan d'image euclidien (EIP) [HEYDEN et ÅSTRÖM 1997]. Nous pouvons donc supposer en toute confiance que la caméra a un plan d'image euclidien (EIP). Cette hypothèse pratique a déjà été utilisée par plusieurs méthodes d'autocalibrage, mais il reste encore à l'exploiter dans l'autocalibrage stratifié. Notre contribution principale est de formuler une nouvelle équation polynomiale quartique en les coordonnées du plan à l'infini pour une caméra EIP et des paramètres intrinsèques constants. Ce polynôme, qu'on dénomme polynôme EIP, est obtenu pour chaque paire d'images. En combinaison avec une équation quartique classique pour chaque paire, la contrainte de module [POLLEFEYS et VAN GOOL 1999], chaque paire d'images donne donc deux contraintes sur la localisation du plan à l'infini. Deux paires d'images sont par conséquent suffisantes en général pour obtenir une solution unique. En conséquence, nous proposons une méthode d'autocalibrage stratifiée qui utilise le polynôme EIP en combinaison avec la contrainte de module pour estimer le plan à l'infini. Les expérimentations réalisées avec des données synthétiques et des images réelles montrent les avantages qu'apporte le polynôme EIP dans la méthode proposée, notamment sur de courtes séquences d'images. Ce travail sur l'exploitation de l'hypothèse EIP dans l'autocalibrage de caméra stratifiée a été publié dans la British Machine Vision Conference (BMVC) 2020 :

– Devesh Adlakha, Adlane Habed, Fabio Morbidi, Cédric Demonceaux, and Michel de Mathelin (2020). “Stratified autocalibration of cameras with Euclidean Image Plane". In : *British Machine Vision Conference (BMVC)*.

C.3.2 Exploitation de la connaissance partielle du mouvement de la caméra

Comme indiqué dans la Section C.2, une connaissance implicite du mouvement de la caméra est disponible dans le contexte de la reconstruction 3D basée sur l'image, en particulier que le point de vue change légèrement lors de la capture d'images. Nous formalisons cette connaissance vague du mouvement de la caméra et présentons quelques contributions pour l'exploiter dans l'autocalibrage stratifié. Nous considérons le cas d'une caméra en mouvement dont les paramètres intrinsèques sont constants et nous supposons qu'une connaissance vague de l'angle d'orientation relatif de la caméra est disponible. Sous ces hypothèses, nous dérivons deux nouveaux groupes de contraintes entre des paires de caméras sur la localisation du plan à l'infini. Les contraintes du premier groupe sont convexes et formulées sous forme d'inégalités matricielles linéaires (LMI). Nous prouvons que le plan à l'infini fait partie de l'un des deux groupes convexes définis par les LMI, selon que l'angle d'orientation relatif[7] entre la paire de caméras (i, j) est soit $|\theta_{ij}| \leq 120^\circ$ soit $|\theta_{ij}| \geq 120^\circ$. Ces deux groupes convexes définissent la relation géométrique entre le plan à l'infini et de nouveaux objets géométriques, les *hodographes de l'horoptère*. De plus, une vague connaissance de l'angle d'orientation relatif, c'est-à-dire $|\theta_{ij}| \leq 120^\circ$ ou $|\theta_{ij}| \geq 120^\circ$, entre un ensemble de paires de caméras peut être exploitée avec les contraintes LMI pour obtenir une nouvelle reconstruction quasi-affine d'une scène que nous appelons QUARCH : “QUasi Affine Reconstruction with respect to Camera centers and the Hodographs of horopters”. Une reconstruction QUARCH est une spécialisation du QUARC (QUasi-Affine Reconstruction with respect to Camera centers) [NISTÉR 2004b]. Dans le scénario de reconstruction 3D basée sur l'image, l'hypothèse de $|\theta_{ij}| \leq 120^\circ$ est typiquement satisfaite

[7] L'angle d'orientation relatif de la caméra θ_{ij} est autour d'un axe arbitraire dans la représentation axe-angle et $\theta_{ij} \in [-180^\circ, 180^\circ]$.

entre des vues consécutives, et nous exploitons cette connaissance pour obtenir une reconstruction QUARCH.

La seconde série de contraintes sur le plan à l'infini n'est pas convexe et est formulée sous forme d'inégalités polynomiales. Elles exploitent une connaissance plus fine de l'angle d'orientation relatif, c'est-à-dire $|\theta_{ij}| < 90°/k$, où k est un entier positif. Ces contraintes, contrairement aux contraintes convexes, sont dérivées de manière purement algébrique à l'aide du critère de *stabilité de Hurwitz*. Les inégalités polynomiales sont de degré $2k$, donc une connaissance de l'orientation de plus en plus fine se traduit en des polynômes de degré plus élevé. Quand elles sont combinées avec les contraintes convexes, ces inégalités nous permettent de calculer une QUARCH spécialisée. Pour obtenir cette reconstruction, une connaissance plus fine mais toujours vague de $|\theta_{ij}| < 90°/k$ pour un entier positif k est nécessaire pour un ensemble de paires de caméras.

Enfin, nous proposons une méthode d'autocalibrage de caméra stratifiée qui exploite une connaissance vague de l'orientation de la caméra pour localiser le plan à l'infini. Notre méthode se base sur une QUARCH comme première étape vers le calcul des reconstructions affines et métriques. Notre dernière contribution majeure est une méthode contrainte de Levenberg-Marquardt (LM) pour l'optimisation non linéaire soumise à des contraintes LMI. Cette méthode garantit que les contraintes QUARCH sont satisfaites pendant l'optimisation locale utilisée pour localiser le plan à l'infini. Nos expérimentations montrent les avantages de l'incorporation des contraintes basées sur l'orientation. Nous présentons les résultats d'expérimentations conduites aussi bien sur des données synthétiques que sur des images réelles.

Une partie des contributions mentionnées ci-dessus a été publiée dans l'IEEE/CVF International Conference on Computer Vision (ICCV) 2019 :

– Devesh Adlakha, Adlane Habed, Fabio Morbidi, Cédric Demonceaux, and Michel de Mathelin (2019). "QUARCH : A new quasi-affine reconstruction stratum from vague relative camera orientation knowledge". In : *International Conference on Computer Vision (ICCV)*.

Elles comprennent des contraintes convexes, leur interprétation géométrique par les hodographes, la strate QUARCH et la méthode LM avec contraintes. Les contraintes non convexes et le QUARCH spécialisé complètent les résultats publiés. Un article

synthétisant ces contributions est actuellement en préparation et devrait être soumis à l'International Journal of Computer Vision (IJCV) ou à l'IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI).

C.4 Structure du manuscrit

Le manuscrit de thèse est organisé comme suit.

Le chapitre 2 passe en revue certains concepts de base de la géométrie multi-vues et rassemble la littérature pertinente sur l'autocalibrage des caméras.

Le chapitre 3 présente nos travaux sur l'exploitation d'une connaissance partielle de la géométrie de la caméra pour l'autocalibrage.

Le chapitre 4 présente nos travaux sur l'exploitation de connaissances partielles du mouvement de la caméra pour l'autocalibrage.

Le chapitre 5 revient sur les contributions de cette thèse et sur les pistes de recherche pour l'avenir.

L'annexe A présente quelques outils d'optimisation utilisés dans cette thèse.

L'annexe B donne l'expression de la matrice de projection de ligne.

References

[Abramowitz and Stegun 1964] Milton Abramowitz and Irene Stegun (1964). *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. Vol. 55. US Government printing office.

[Adlakha et al. 2019] Devesh Adlakha, Adlane Habed, Fabio Morbidi, Cédric Demonceaux, and Michel de Mathelin (2019). "QUARCH: A New Quasi-Affine Reconstruction Stratum From Vague Relative Camera Orientation Knowledge". In: *International Conference on Computer Vision (ICCV)*, pp. 1082–1090.

[Adlakha et al. 2020] Devesh Adlakha, Adlane Habed, Fabio Morbidi, Cédric Demonceaux, and Michel de Mathelin (2020). "Stratified Autocalibration of Cameras with Euclidean Image Plane". In: *British Machine Vision Conference (BMVC)*.

[Agapito et al. 2001] Lourdes Agapito, Eric Hayman, and Ian Reid (2001). "Self-calibration of rotating and zooming cameras". In: *International Journal of Computer Vision (IJCV)* 45.2, pp. 107–127.

[Agoston 2005] Max K. Agoston (2005). *Computer Graphics and Geometric Modelling: Implementation & Algorithms*. Springer.

[Apostolatos 2003] Theocharis A. Apostolatos (2003). "Hodograph: A useful geometrical tool for solving some difficult problems in dynamics". In: *American Journal of Physics* 71.3, pp. 261–266.

[Armstrong et al. 1994] Martin Armstrong, Andrew Zisserman, and Paul Beardsley (1994). "Euclidean Structure from Uncalibrated Images". In: *British Machine Vision Conference (BMVC)*. Vol. 2, pp. 509–518.

[Armstrong et al. 1996] Martin Armstrong, Andrew Zisserman, and Richard Hartley (1996). "Self-calibration from image triplets". In: *European Conference on Computer Vision (ECCV)*. Springer, pp. 1–16.

[Bartoli and Sturm 2004] Adrien Bartoli and Peter Sturm (2004). "The 3D Line Motion Matrix and Alignment of Line Reconstructions". In: *International Journal of Computer Vision (IJCV)* 57, pp. 159–178.

[Beardsley et al. 1994] Paul Beardsley, Andrew Zisserman, and David Murray (1994). "Navigation using affine structure from motion". In: *European Conference on Computer Vision (ECCV)*, pp. 85–96.

[Bocquillon et al. 2007] Benoît Bocquillon, Adrien Bartoli, Pierre Gurdjos, and Alain Crouzil (2007). "On Constant Focal Length Self-Calibration From Multiple Views". In: *Conference on Computer Vision and Pattern Recognition (CVPR)*.

[Bougnoux 1998] Sylvain Bougnoux (1998). "From projective to euclidean space under any practical situation, a criticism of self-calibration". In: *International Conference on Computer Vision (ICCV)*. IEEE, pp. 790–796.

[Boyd et al. 1994] Stephen Boyd, Laurent El Ghaoui, Eric Feron, and Venkataramanan Balakrishnan (1994). *Linear matrix inequalities in system and control theory*. SIAM.

[Boyd and Vandenberghe 2004] Stephen Boyd and Lieven Vandenberghe (2004). *Convex Optimization*. Cambridge University Press.

[Bresson et al. 2017] Guillaume Bresson, Zayed Alsayed, Li Yu, and Sébastien Glaser (2017). "Simultaneous localization and mapping: A survey of current trends in autonomous driving". In: *IEEE Transactions on Intelligent Vehicles* 2.3, pp. 194–220.

[Bugarin et al. 2016] Florian Bugarin, Adrien Bartoli, Didier Henrion, Jean-Bernard Lasserre, Jean-José Orteu, and Thierry Sentenac (2016). "Rank-Constrained Fundamental Matrix Estimation by Polynomial Global Optimization Versus the Eight-Point Algorithm". In: *Journal of Mathematical Imaging and Vision (JMIV)* 53.1, pp. 42–60.

[Bugarin et al. 2015] Florian Bugarin, Didier Henrion, and Jean-Bernard Lasserre (2015). "Minimizing the sum of many rational functions". In: *Mathematical Programming Computation* 8.1, pp. 83–111.

[Canclini et al. 2019] Antonio Canclini, Francesco Malapelle, Marco Marcon, Stefano Tubaro, and Andrea Fusiello (2019). "View-synthesis from uncalibrated cameras and parallel planes". In: *Signal Processing: Image Communication* 79, pp. 40–53.

[Chandraker et al. 2007] Manmohan Chandraker, Sameer Agarwal, Fredrik Kahl, David Nistér, and David Kriegman (2007). "Autocalibration via Rank-Constrained Estimation of the Absolute Quadric". In: *Conference on Computer Vision and Pattern Recognition (CVPR)*.

[Chandraker et al. 2010] Manmohan Chandraker, Sameer Agarwal, David Kriegman, and Serge Belongie (2010). "Globally Optimal Algorithms for Stratified Autocalibration". In: *International Journal of Computer Vision (IJCV)* 90.2, pp. 236–254.

[Cox et al. 2005] David A. Cox, John Little, and Donal O'Shea (2005). *Using Algebraic Geometry*. Graduate Texts in Mathematics. Springer New York.

[Delmerico et al. 2019] Jeffrey Delmerico, Stefano Mintchev, Alessandro Giusti, Boris Gromov, Kamilo Melo, Tomislav Horvat, Cesar Cadena, Marco Hutter, Auke Ijspeert, Dario Floreano, et al. (2019). "The current state and future outlook of rescue robotics". In: *Journal of Field Robotics (JFR)* 36.7, pp. 1171–1191.

[Demirdjian et al. 1998] David Demirdjian, Gabriela Csurka, and Radu Horaud (1998). "Autocalibration in the presence of critical motions". In: *British Machine Vision Conference (BMVC)*. Vol. 2. The British Machine Vision Association (BMVA), pp. 751–759.

[Derbes 2001] David Derbes (2001). "Reinventing the wheel: Hodographic solutions to the Kepler problems". In: *American Journal of Physics* 69.4, pp. 481–489.

[Dessales 2020] Hélène Dessales (2020). *The Villa of Diomedes: The Making of a Roman Villa in Pompeii*. Collection Histoire et archéologie. Hermann Editeurs.

[Ding et al. 2020] Yaqing Ding, Jian Yang, Jean Ponce, and Hui Kong (2020). "Minimal solutions to relative pose estimation from two views sharing a common direction with unknown focal length". In: *Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7045–7053.

[Espuny 2007] Ferran Espuny (2007). "A new linear method for camera self-calibration with planar motion". In: *Journal of Mathematical Imaging and Vision (JMIV)* 27.1, pp. 81–88.

[Espuny et al. 2011] Ferran Espuny, Joan Aranda, and José I. Burgos Gil (2011). "Camera self-calibration with parallel screw axis motion by intersecting imaged horopters". In: *Scandinavian Conference on Image Analysis*. Springer, pp. 1–12.

[Falkingham et al. 2020] Peter L. Falkingham, Morgan L. Turner, and Stephen M. Gatesy (2020). "Constructing and testing hypotheses of dinosaur foot motions from fossil tracks using digitization and simulation". In: *Palaeontology* 63.6, pp. 865–880.

[Fan 2013] Jinyan Fan (2013). "On the Levenberg-Marquardt methods for convex constrained nonlinear equations". In: *Journal of Industrial and Management Optimization* 9.1, pp. 227–241.

[Fan and Yuan 2005] Jinyan Fan and Ya-xiang Yuan (2005). "On the quadratic convergence of the Levenberg-Marquardt method without nonsingularity assumption". In: *Computing* 74.1, pp. 23–39.

[Faugeras et al. 1992] Olivier Faugeras, Quang-Tuan Luong, and Stephen Maybank (1992). "Camera self-calibration: Theory and experiments". In: *European Conference on Computer Vision (ECCV)*, pp. 321–334.

[Faugeras et al. 2001] Olivier Faugeras, Quang-Tuan Luong, and Théodore Papadopoulo (2001). *The Geometry of Multiple Images: The Laws that Govern the Formation of Multiple Images of a Scene and Some of Their Applications*. MIT Press.

[Faugeras et al. 2000] Olivier Faugeras, Quang-Tuan Luong, and Peter Sturm (2000). "Self-calibration of a 1D projective camera and its application to the self-calibration of a 2D projective camera". In: *Transactions on Pattern Analysis and Machine Intelligence (TPAMI)* 22.10, pp. 1179–1185.

[Fischler and Bolles 1981] Martin A. Fischler and Robert C. Bolles (1981). "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography". In: *Communications of the ACM* 24.6, pp. 381–395.

[Forsyth and Ponce 2003] David Forsyth and Jean Ponce (2003). *Computer Vision: A Modern Approach*. Prentice Hall.

[Frahm and Koch 2003] Jan-Michael Frahm and Reinhard Koch (2003). "Camera calibration with known rotation". In: *International Conference on Computer Vision (ICCV)*. Vol. 3. IEEE Computer Society, pp. 1418–1418.

[Furukawa et al. 2010] Yasutaka Furukawa, Brian Curless, Steven M. Seitz, and Richard Szeliski (2010). "Towards internet-scale multi-view stereo". In: *Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, pp. 1434–1441.

[Fusiello 2000] Andrea Fusiello (2000). "Uncalibrated Euclidean reconstruction: a review". In: *Image and Vision Computing (IVC)* 18.6-7, pp. 555–563.

[Fusiello 2001] Andrea Fusiello (2001). "A new autocalibration algorithm: Experimental evaluation". In: *International Conference on Computer Analysis of Images and Patterns*. Springer, pp. 717–724.

[Fusiello et al. 2004] Andrea Fusiello, Arrigo Benedetti, Michela Farenzena, and Alessandro Busti (2004). "Globally convergent autocalibration using interval analysis". In: *Transactions on Pattern Analysis and Machine Intelligence (TPAMI)* 26.12, pp. 1633–1638.

[Galanakis et al. 2021] George Galanakis, Xenophon Zabulis, Theodore Evdaimon, Sven-Eric Fikenscher, Sebastian Allertseder, Theodora Tsikrika, and Ste-

fanos Vrochidis (2021). “A Study of 3D Digitisation Modalities for Crime Scene Investigation”. In: *Forensic Sciences* 1.2, pp. 56–85.

[Gallego et al. 2018] Guillermo Gallego, Elias Mueggler, and Peter Sturm (2018). *Translation of “Zur Ermittlung eines Objektes aus zwei Perspektiven mit innerer Orientierung" by Erwin Kruppa (1913).*

[Gherardi and Fusiello 2010] Riccardo Gherardi and Andrea Fusiello (2010). “Practical Autocalibration”. In: *European Conference on Computer Vision (ECCV)*, pp. 790–801.

[Golnaraghi and Kuo 2009] Farid Golnaraghi and Benjamin Kuo (2009). *Automatic control systems*. 9th ed. Wiley.

[Gurdjos et al. 2009] Pierre Gurdjos, Adrien Bartoli, and Peter Sturm (2009). “Is dual linear self-calibration artificially ambiguous?” In: *International Conference on Computer Vision (ICCV)*, pp. 88–95.

[Habed et al. 2012] Adlane Habed, Kassem Al Ismaeil, and David Fofi (2012). “A New Set of Quartic Trivariate Polynomial Equations for Stratified Camera Self-calibration under Zero-Skew and Constant Parameters Assumptions”. In: *European Conference on Computer Vision (ECCV)*, pp. 710–723.

[Habed et al. 2014] Adlane Habed, Danda Pani Paudel, Cédric Demonceaux, and David Fofi (2014). “Efficient pruning lmi conditions for branch-and-prune rank and chirality-constrained estimation of the dual absolute quadric”. In: *Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 493–500.

[Hamilton 1847] William Rowan Hamilton (1847). “A New Method of Expressing, in Symbolical Language, the Newtonian Law of Attraction”. In: *Proceedings of the Royal Irish Academy* 3, pp. 344–382.

[Hartley 1994a] Richard Hartley (1994a). “Euclidean reconstruction from uncalibrated views”. In: *Joint European-US Workshop on Applications of Invariance in Computer Vision*, pp. 235–256.

[Hartley 1994b] Richard Hartley (1994b). "Self-calibration from multiple views with a rotating camera". In: *European Conference on Computer Vision (ECCV)*. Springer, pp. 471–478.

[Hartley 1997a] Richard Hartley (1997a). "Kruppa's equations derived from the fundamental matrix". In: *Transactions on Pattern Analysis and Machine Intelligence (TPAMI)* 19.2, pp. 133–135.

[Hartley 1997b] Richard Hartley (1997b). "Self-calibration of stationary cameras". In: *International Journal of Computer Vision (IJCV)* 22.1, pp. 5–23.

[Hartley 1998] Richard Hartley (1998). "Chirality". In: *International Journal of Computer Vision (IJCV)* 26.1, pp. 41–61.

[Hartley et al. 1999] Richard Hartley, Eric Hayman, Lourdes de Agapito, and Ian Reid (1999). "Camera calibration and the search for infinity". In: *International Conference on Computer Vision (ICCV)*, pp. 510–517.

[Hartley and Mundy 1993] Richard Hartley and Joseph Mundy (1993). "Relationship between photogrammmetry and computer vision". In: *Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision*. Vol. 1944. International Society for Optics and Photonics. SPIE, pp. 92–105.

[Hartley and Sturm 1997] Richard Hartley and Peter Sturm (1997). "Triangulation". In: *Computer Vision and Image Understanding (CVIU)* 68.2, pp. 146–157.

[Hartley and Zisserman 2004] Richard Hartley and Andrew Zisserman (2004). *Multiple View Geometry in Computer Vision*. 2nd. Cambridge University Press.

[Henrion and Lasserre 2003] Didier Henrion and Jean-Bernard Lasserre (2003). "GloptiPoly: Global optimization over polynomials with Matlab and SeDuMi". In: *ACM Transactions on Mathematical Software* 29.2, pp. 165–194.

[Henrion and Lasserre 2005] Didier Henrion and Jean-Bernard Lasserre (2005). "Detecting global optimality and extracting solutions in GloptiPoly". In: *Positive polynomials in control*. Springer, pp. 293–310.

[Henrion et al. 2009] Didier Henrion, Jean-Bernard Lasserre, and Johan Löfberg (2009). "GloptiPoly 3: Moments, Optimization and Semidefinite Programming". In: *Optimization Methods and Software* 24.4–5, pp. 761–779.

[Heyden and Åström 1996] Anders Heyden and Kalle Åström (1996). "Euclidean reconstruction from constant intrinsic parameters". In: *International Conference on Pattern Recognition (ICPR)*. Vol. 1, pp. 339–343.

[Heyden and Åström 1997] Anders Heyden and Kalle Åström (1997). "Euclidean Reconstruction from Image Sequences with Varying and Unknown Focal Length and Principal Point". In: *Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 438–443.

[Heyden and Åström 1999] Anders Heyden and Kalle Åström (1999). "Flexible Calibration: Minimal Cases for Auto-calibration". In: *International Conference on Computer Vision (ICCV)*.

[Hurwitz 1895] Adolf Hurwitz (1895). "Ueber die Bedingungen, unter welchen eine Gleichung nur Wurzeln mit negativen reellen Theilen besitzt". In: *Mathematische Annalen* 46, pp. 273–284.

[Kahl and Henrion 2007] Fredrik Kahl and Didier Henrion (2007). "Globally Optimal Estimates for Geometric Reconstruction Problems". In: *International Journal of Computer Vision (IJCV)* 74.1, pp. 3–15.

[Kahl et al. 2000] Fredrik Kahl, Bill Triggs, and Kalle Åström (2000). "Critical motions for auto-calibration when some intrinsic parameters can vary". In: *Journal of Mathematical Imaging and Vision (JMIV)* 13.2, pp. 131–146.

[Kanzow et al. 2004] Christian Kanzow, Nobuo Yamashita, and Masao Fukushima (2004). "Levenberg-Marquardt methods with strong local convergence properties for solving nonlinear equations with convex constraints". In: *Journal of Computational and Applied Mathematics* 172.2, pp. 375–397.

[Kasten et al. 2019] Yoni Kasten, Amnon Geifman, Meirav Galun, and Ronen Basri (2019). "GPSfM: Global Projective SFM Using Algebraic Constraints on

Multi-View Fundamental Matrices". In: *Conference on Computer Vision and Pattern Recognition (CVPR)*.

[Kruppa 1913] Erwin Kruppa (1913). *Zur Ermittlung eines Objektes aus zwei Perspektiven mit innerer Orientierung*. Vol. 122. Sitzungsberichte der Akademie der Wissenschaften, Wien, Mathematisch-Naturwissenschaftlichen Klasse, Abteilung IIa, pp. 1939–1948.

[Kukelova et al. 2010] Zuzana Kukelova, Martin Bujnak, and Tomas Pajdla (2010). "Closed-form solutions to minimal absolute pose problems with known vertical direction". In: *Asian Conference on Computer Vision (ACCV)*. Springer, pp. 216–229.

[Lancaster and Tismenetsky 1985] Peter Lancaster and Miron Tismenetsky (1985). *The Theory of Matrices: With Applications*. Computer Science and Scientific Computing. Elsevier Science.

[Lasserre 2001] Jean-Bernard Lasserre (2001). "Global Optimization with Polynomials and the Problem of Moments". In: *SIAM Journal on Optimization* 11.3, pp. 796–817.

[Lasserre 2015] Jean-Bernard Lasserre (2015). *An introduction to polynomial and semi-algebraic optimization*. Vol. 52. Cambridge University Press.

[LeCun et al. 2015] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton (2015). "Deep learning". In: *Nature* 521.7553, pp. 436–444.

[Li and C. Shen 2006] Hongdong Li and Chunhua Shen (2006). "An LMI approach for reliable PTZ camera self-calibration". In: *International Conference on Video and Signal Based Surveillance*. IEEE, pp. 79–79.

[Lianos et al. 2018] Konstantinos-Nektarios Lianos, Johannes L. Schonberger, Marc Pollefeys, and Torsten Sattler (2018). "VSO: Visual semantic odometry". In: *European Conference on Computer Vision (ECCV)*, pp. 234–250.

[Liu et al. 2020] Yonghuai Liu, Nick Pears, Paul Rosin, and Patrik Huber (2020). *3D Imaging, Analysis and Applications*. Springer.

[Löfberg 2004] J. Löfberg (2004). "YALMIP: A Toolbox for Modeling and Optimization in MATLAB". In: *International Conference on Robotics and Automation (ICRA)*, pp. 284–289.

[Longuet-Higgins 1981] Christopher Longuet-Higgins (1981). "A computer algorithm for reconstructing a scene from two projections". In: *Nature* 293.5828, pp. 133–135.

[Lourakis 2000] Manolis Lourakis (2000). "Camera Self-Calibration Using the Kruppa Equations and the SVD of the Fundamental Matrix: The Case of Varying Intrinsic Parameters". In: *Research Report* 3911.

[Lourakis and Argyros 2009] Manolis Lourakis and Antonis Argyros (2009). "SBA: A Software Package for Generic Sparse Bundle Adjustment". In: *ACM Transactions on Mathematical Software* 36.1, pp. 1–30.

[Lowe 2004] David G. Lowe (2004). "Distinctive image features from scale-invariant keypoints". In: *International Journal of Computer Vision (IJCV)* 60.2, pp. 91–110.

[Luong and Faugeras 1997] Quang-Tuan Luong and Olivier Faugeras (1997). "Self-Calibration of a Moving Camera from Point Correspondences and Fundamental Matrices". In: *International Journal of Computer Vision (IJCV)* 22.3, pp. 261–289.

[Ma et al. 2004] Yi Ma, Stefano Soatto, Jana Košecká, and Shankar Sastry (2004). *An Invitation to 3-D Vision: From Images to Geometric Models*. Vol. 26. Springer.

[Magerand et al. 2012] Ludovic Magerand, Adrien Bartoli, Omar Ait-Aider, and Daniel Pizarro (2012). "Global Optimization of Object Pose and Motion from a Single Rolling Shutter Image with Automatic 2D-3D Matching". In: *European Conference on Computer Vision (ECCV)*, pp. 456–469.

[Magerand and Bue 2018] Ludovic Magerand and Alessio Del Bue (2018). "Revisiting Projective Structure from Motion: A Robust and Efficient Incremental Solution". In: *Transactions on Pattern Analysis and Machine Intelligence (TPAMI)* 42.2, pp. 430–443.

[Mahamud and Hebert 2000] Shyjan Mahamud and Martial Hebert (2000). "Iterative projective reconstruction from multiple views". In: *Conference on Computer Vision and Pattern Recognition (CVPR)*. Vol. 2. IEEE, pp. 430–437.

[Mahamud et al. 2001] Shyjan Mahamud, Martial Hebert, Yasuhiro Omori, and Jean Ponce (2001). "Provably-convergent iterative methods for projective structure from motion". In: *Conference on Computer Vision and Pattern Recognition (CVPR)*. Vol. 1. IEEE, pp. I–I.

[Martyushev 2018] Evgeniy Martyushev (2018). "Self-calibration of cameras with Euclidean image plane in case of two views and known relative rotation angle". In: *European Conference on Computer Vision (ECCV)*, pp. 415–429.

[Maurer et al. 2017] Michael Maurer, Manuel Hofer, Friedrich Fraundorfer, and Horst Bischof (2017). "Automated inspection of power line corridors to measure vegetation undercut using UAV-based images". In: *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pp. 33–40.

[Maybank 1990] Stephen Maybank (1990). "The Projective Geometry of Ambiguous Surfaces". In: *Philosophical Transactions: Physical Sciences and Engineering* 332.1623, pp. 1–47.

[Maybank 1993] Stephen Maybank (1993). *Theory of Reconstruction from Image Motion*. Springer Series in Information Sciences. Springer Berlin Heidelberg.

[Maybank and Faugeras 1992] Stephen Maybank and Olivier Faugeras (1992). "A theory of self-calibration of a moving camera". In: *International Journal of Computer Vision (IJCV)* 8.2, pp. 123–151.

[Mendonça and Cipolla 1999] Paulo R.S. Mendonça and Roberto Cipolla (1999). "A simple technique for self-calibration". In: *Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 500–505.

[Moons et al. 1996] Theo Moons, Luc Van Gool, Marc Proesmans, and Eric Pauwels (1996). "Affine reconstruction from perspective image pairs with a relative object-camera translation in between". In: *Transactions on Pattern Analysis and Machine Intelligence (TPAMI)* 18.1, pp. 77–83.

[Moulon et al. 2016] Pierre Moulon, Pascal Monasse, Romuald Perrot, and Renaud Marlet (2016). "OpenMVG: Open multiple view geometry". In: *International Workshop on Reproducible Research in Pattern Recognition*. Springer, pp. 60–74.

[Nistér 2004a] David Nistér (2004a). "An efficient solution to the five-point relative pose problem". In: *Transactions on Pattern Analysis and Machine Intelligence (TPAMI)* 26.6, pp. 756–770.

[Nistér 2004b] David Nistér (2004b). "Untwisting a Projective Reconstruction". In: *International Journal of Computer Vision (IJCV)* 60.2, pp. 1–33.

[Oliensis and Hartley 2007] John Oliensis and Richard Hartley (2007). "Iterative Extensions of the Sturm/Triggs Algorithm: Convergence and Nonconvergence". In: *Transactions on Pattern Analysis and Machine Intelligence (TPAMI)* 29.12, pp. 2217–2233.

[Olsson and Enqvist 2011] Carl Olsson and Olof Enqvist (2011). "Stable Structure from Motion for Unordered Image Collections". In: *Scandinavian Conference on Image Analysis*, pp. 524–535.

[Parashar et al. 2017] Shaifali Parashar, Daniel Pizarro, and Adrien Bartoli (2017). "Isometric non-rigid shape-from-motion with riemannian geometry solved in linear time". In: *Transactions on Pattern Analysis and Machine Intelligence (TPAMI)* 40.10, pp. 2442–2454.

[Paudel et al. 2017] Danda Pani Paudel, Adlane Habed, and Luc Van Gool (2017). "Optimal transformation estimation with semantic cues". In: *International Conference on Computer Vision (ICCV)*, pp. 4658–4667.

[Paudel and Van Gool 2018] Danda Pani Paudel and Luc Van Gool (2018). "Sampling algebraic varieties for robust camera autocalibration". In: *European Conference on Computer Vision (ECCV)*, pp. 265–281.

[Peng and Sturm 2019] Songyou Peng and Peter Sturm (2019). "Calibration Wizard: A Guidance System for Camera Calibration Based on Modelling Geomet-

ric and Corner Uncertainty". In: *International Conference on Computer Vision (ICCV)*.

[Perwass et al. 2009] Christian Perwass, Herbert Edelsbrunner, Leif Kobbelt, and Konrad Polthier (2009). *Geometric Algebra with Applications in Engineering*. Vol. 4. Springer.

[Pollefeys et al. 1999] Marc Pollefeys, Reinhard Koch, and Luc Van Gool (1999). "Self-Calibration and Metric Reconstruction Inspite of Varying and Unknown Intrinsic Camera Parameters". In: *International Journal of Computer Vision (IJCV)* 32.1, pp. 7–25.

[Pollefeys and Van Gool 1999] Marc Pollefeys and Luc Van Gool (1999). "Stratified self-calibration with the modulus constraint". In: *Transactions on Pattern Analysis and Machine Intelligence (TPAMI)* 21.8, pp. 707–724.

[Pollefeys et al. 1996] Marc Pollefeys, Luc Van Gool, and Marc Proesmans (1996). "Euclidean 3D reconstruction from image sequences with variable focal lengths". In: *European Conference on Computer Vision (ECCV)*. Springer, pp. 31–42.

[Ponce et al. 2005] Jean Ponce, Kenton McHenry, Théodore Papadopoulo, Monique Teillaud, and Bill Triggs (2005). "On the absolute quadratic complex and its application to autocalibration". In: *Conference on Computer Vision and Pattern Recognition (CVPR)*.

[Pucik et al. 2021] Tomas Pucik, Pieter Groenemeijer, and Ivan Tsonevsky (2021). *Vertical wind shear and convective storms*.

[Rameau et al. 2012] Francois Rameau, Adlane Habed, Cédric Demonceaux, Désiré Sidibé, and David Fofi (2012). "Self-Calibration of PTZ Camera using New LMI Constraints". In: *Asian Conference on Computer Vision (ACCV)*. Springer, pp. 297–308.

[Ronda et al. 2004] José Ignacio Ronda, Antonio Valdés, and Fernando Jaureguizar (2004). "Camera Autocalibration and Horopter Curves". In: *International Journal of Computer Vision (IJCV)* 57.3, pp. 219–232.

[Rosinol et al. 2020] Antoni Rosinol, Marcus Abate, Yun Chang, and Luca Carlone (2020). "Kimera: an open-source library for real-time metric-semantic localization and mapping". In: *International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 1689–1696.

[Schaffalitzky 2000] Frederik Schaffalitzky (2000). "Direct Solution of Modulus Constraint". In: *Indian Conference Computer Vision, Graphics and Image Processing*, pp. 314–321.

[Schönberger and Frahm 2016] Johannes Lutz Schönberger and Jan-Michael Frahm (2016). "Structure-from-Motion Revisited". In: *Conference on Computer Vision and Pattern Recognition (CVPR)*.

[F. Shen and Wang 2002] Fei Shen and Han Wang (2002). "A new stratified self-calibration algorithm under small camera rotations". In: *Image and Vision Computing New Zealand (IVCNZ)*.

[Snavely et al. 2008] Noah Snavely, Steven M. Seitz, and Richard Szeliski (2008). "Modeling the world from internet photo collections". In: *International Journal of Computer Vision (IJCV)* 80.2, pp. 189–210.

[Sommese et al. 2005] Andrew Sommese, Jan Verschelde, and Charles Wampler (2005). *Solving Polynomial Equations: Foundations, Algorithms, and Applications*. Springer Berlin Heidelberg.

[Speciale et al. 2017] Pablo Speciale, Danda Pani Paudel, Martin Oswald, Till Kroeger, Luc Van Gool, and Marc Pollefeys (2017). "Consensus maximization with linear matrix inequality constraints". In: *Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4941–4949.

[Speciale et al. 2018] Pablo Speciale, Danda Pani Paudel, Martin Oswald, Hayko Riemenschneider, Luc Van Gool, and Marc Pollefeys (2018). "Consensus maximization for semantic region correspondences". In: *Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7317–7326.

[Statham 2020] Nataska Statham (2020). "Use of Photogrammetry in Video Games: A Historical Overview". In: *Games and Culture: A Journal of Interactive Media* 15.3, pp. 289–307.

[Strecha et al. 2003] Christoph Strecha, Tinne Tuytelaars, and Luc Van Gool (2003). "Dense Matching of Multiple Wide-baseline Views". In: *International Conference on Computer Vision (ICCV)*. Vol. 2, pp. 1194–1201.

[Strecha et al. 2008] Christoph Strecha, Wolfgang von Hansen, Luc Van Gool, Pascal Fua, and Ulrich Thoennessen (2008). "On Benchmarking Camera Calibration and Multi-View Stereo for High Resolution Imagery". In: *Conference on Computer Vision and Pattern Recognition (CVPR)*.

[Sturm 1996] Peter Sturm (1996). "Self-calibration of a moving camera by pre-calibration". In: *British Machine Vision Conference (BMVC)*. Vol. 2, pp. 675–684.

[Sturm 1997a] Peter Sturm (1997a). "Critical motion sequences for monocular self-calibration and uncalibrated Euclidean reconstruction". In: *Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, pp. 1100–1105.

[Sturm 1997b] Peter Sturm (1997b). "Vision 3D non calibrée: contributions à la reconstruction projective et étude des mouvements critiques pour l'auto-calibrage". PhD thesis. Institut National Polytechnique de Grenoble-INPG.

[Sturm 2000] Peter Sturm (2000). "A case against Kruppa's equations for camera self-calibration". In: *Transactions on Pattern Analysis and Machine Intelligence (TPAMI)* 22.10, pp. 1199–1204.

[Sturm 2002] Peter Sturm (2002). "Critical motion sequences for the self-calibration of cameras and stereo systems with variable focal length". In: *Image and Vision Computing (IVC)* 20.5-6, pp. 415–426.

[Sturm 2011] Peter Sturm (2011). "A historical survey of geometric computer vision". In: *International Conference on Computer Analysis of Images and Patterns*. Springer, pp. 1–8.

[Sturm and Triggs 1996] Peter Sturm and Bill Triggs (1996). "A factorization based algorithm for multi-image projective structure and motion". In: *European conference on computer vision*. Springer, pp. 709–720.

[Szeliski 2022] Richard Szeliski (2022). *Computer vision: algorithms and applications*. Springer Nature.

[Toft et al. 2018] Carl Toft, Erik Stenborg, Lars Hammarstrand, Lucas Brynte, Marc Pollefeys, Torsten Sattler, and Fredrik Kahl (2018). "Semantic match consistency for long-term visual localization". In: *European Conference on Computer Vision (ECCV)*, pp. 383–399.

[Triggs 1996] Bill Triggs (1996). "Factorization methods for projective structure and motion". In: *Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, pp. 845–851.

[Triggs 1997] Bill Triggs (1997). "Autocalibration and the absolute quadric". In: *Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 609–614.

[Triggs et al. 1999] Bill Triggs, Philip McLauchlan, Richard Hartley, and Andrew Fitzgibbon (1999). "Bundle adjustment—a modern synthesis". In: *International Workshop on Vision Algorithms*. Springer, pp. 298–372.

[Trutman et al. 2022] Pavel Trutman, Mohab Safey El Din, Didier Henrion, and Tomas Pajdla (2022). "Globally optimal solution to inverse kinematics of 7DOF serial manipulator". In: *Robotics and Automation Letters (RAL)* 7.3, pp. 6012–6019.

[Valdés et al. 2006] Antonio Valdés, José Ignacio Ronda, and Guillermo Gallego (2006). "The Absolute Line Quadric and Camera Autocalibration". In: *International Journal of Computer Vision (IJCV)* 66.3, pp. 283–303.

[Van Gool et al. 1994] Luc Van Gool, Theo Moons, Marc Proesmans, and Marc Van Diest (1994). "Affine reconstruction from perspective image pairs obtained by a translating camera". In: *International Conference on Pattern Recognition (ICPR)*. Vol. 1, pp. 290–294.

[Verschelde 1999] Jan Verschelde (1999). "PHCpack: A general-purpose solver for polynomial systems by homotopy continuation". In: *ACM Transactions on Mathematical Software* 25.2, pp. 251–276.

[Von Helmholtz 1867] Hermann Von Helmholtz (1867). *Handbuch der physiologischen Optik: mit 213 in den Text eingedruckten Holzschnitten und 11 Tafeln*. Vol. 9. Voss.

[Werner and Pajdla 2001] Tomas Werner and Tomas Pajdla (2001). "Oriented Matching Constraints." In: *British Machine Vision Conference (BMVC)*, pp. 1–10.

[Wright 2005] Margaret Wright (2005). "The interior-point revolution in optimization: history, recent developments, and lasting consequences". In: *Bulletin of the American Mathematical Society* 42.1, pp. 39–56.

[C. Wu 2013] Changchang Wu (2013). "Towards linear-time incremental structure from motion". In: *International Conference on 3D Vision (3DV)*, pp. 127–134.

[F. Wu et al. 2013] Fuchao Wu, Ming Zhang, and Zhanyi Hu (2013). "Self-Calibration Under the Cayley Framework". In: *International Journal of Computer Vision (IJCV)* 103, pp. 372–398.

[Yu 2004] Zhensheng Yu (2004). "On the global convergence of a Levenberg-Marquardt method for constrained nonlinear equations". In: *Journal of Applied Mathematics and Computing* 16.1, pp. 183–194.

[Zeller and Faugeras 1996] Cyril Zeller and Olivier Faugeras (1996). "Camera self-calibration from video sequences: the Kruppa equations revisited". In: *Research Report*.

[Zhang 2000] Zhengyou Zhang (2000). "A flexible new technique for camera calibration". In: *Transactions on Pattern Analysis and Machine Intelligence (TPAMI)* 22.11, pp. 1330–1334.

Devesh ADLAKHA

Exploitation de connaissances partielles sur le mouvement et la géométrie des caméras en vision 3D non calibrée

Résumé

Reconstruire la structure 3D de la scène à partir de plusieurs images est un problème fondamental de la vision par ordinateur, appelé Structure-from-Motion (SfM). Nous nous intéressons au problème de SfM non calibrée, où seule une structure à une ambiguïté projective peut être obtenue. Le but est de transformer la reconstruction projective en une reconstruction métrique, ce qui consiste à localiser la conique absolue sur le plan à l'infini. Cette thèse présente deux contributions principales. La première exploite une connaissance partielle de la géométrie de la caméra, en particulier que la caméra a des pixels carrés. La plupart des caméras modernes satisfont cette hypothèse. Nous formulons une nouvelle contrainte polynomiale sur le plan à l'infini sous cette hypothèse. La deuxième contribution exploite une vague connaissance du mouvement de la caméra, que le point de vue change légèrement lors de la capture d'images pour établir des correspondances entre les images. Nous prouvons que le plan à l'infini est confiné à un groupe convexe en exploitant les limites de l'angle de rotation relatif entre les paires de caméras. Nous proposons des méthodes dédiées à chaque contribution et présentons les résultats d'expérimentations conduites aussi bien sur des données synthétiques que sur des images réelles.

Mots-clés: vision par ordinateur, Structure-from-Motion, autocalibrage de la caméra

Summary

Reconstructing a scene in 3D from multiple images is a fundamental problem in computer vision known as Structure-from-Motion (SfM). We investigate uncalibrated SfM, where a reconstruction only up to a projective transformation can be obtained. The goal is to recover a metric reconstruction from the projective one that involves locating the so-called Absolute Conic on the plane at infinity. The main contributions of this thesis are twofold. The first contribution exploits partial knowledge of the camera geometry, specifically that the camera has square pixels. This assumption is satisfied by most modern cameras. We formulate a new polynomial constraint on the plane at infinity under this assumption. The second contribution exploits a vague knowledge of the camera motion that the viewpoint is typically changed mildly between images to ensure sufficient overlap to match features. We show that bounds on the relative rotation angle between camera pairs can be used to constrain the plane at infinity to a convex set. We propose dedicated methods for each contribution and report the experimental evaluation conducted using synthetic and real data.

Keywords: computer vision, Structure-from-Motion, camera autocalibration