Advances in Applied Mathematics 88 (2017) 62–91

Contents lists available at ScienceDirect

Advances in Applied Mathematics

www.elsevier.com/locate/yaama

# Congruences and concurrent lines in multi-view geometry

Jean Ponce [a,b], Bernd Sturmfels [c,*], Mathew Trager [b]

[a] École Normale Supérieure/PSL Research Univ., Paris, France
[b] INRIA, Paris, France
[c] University of California, Berkeley, USA

A R T I C L E   I N F O

*Article history:*
Received 21 August 2016
Received in revised form 31 December 2016
Accepted 3 January 2017
Available online xxxx

*MSC:*
14Q15
51N15
65D18
68W30

A B S T R A C T

We present a new framework for multi-view geometry in computer vision. A camera is a mapping between $\mathbb{P}^3$ and a line congruence. This model, which ignores image planes and measurements, is a natural abstraction of traditional pinhole cameras. It includes two-slit cameras, pushbroom cameras, catadioptric cameras, and many more. We study the concurrent lines variety, which consists of $n$-tuples of lines in $\mathbb{P}^3$ that intersect at a point. Combining its equations with those of various congruences, we derive constraints for corresponding images in multiple views. We also study photographic cameras which use image measurements and are modeled as rational maps from $\mathbb{P}^3$ to $\mathbb{P}^2$ or $\mathbb{P}^1 \times \mathbb{P}^1$.



## 1. Introduction

Multi-view geometry lays the foundations for algorithms that reconstruct a scene from multiple images. Developed in the 1980s, building on classical photogrammetry, this subject has had many successful applications in computer vision. The book [16]

* Corresponding author.
*E-mail addresses:* Jean.Ponce@ens.fr (J. Ponce), bernd@berkeley.edu (B. Sturmfels), matthew.trager@inria.fr (M. Trager).

http://dx.doi.org/10.1016/j.aam.2017.01.001
0196-8858/

offers a comprehensive introduction. Recently, on the mathematical side, the field of *algebraic vision* emerged. It studies objects such as the multi-view varieties [2,28] and their moduli in tensor spaces [1,23].

A pinhole camera is typically modeled as a linear map $\mathbb{P}^3 \dashrightarrow \mathbb{P}^2$, described by a $3 \times 4$-matrix up to scale. This has eleven degrees of freedom, three of which describe the center (or pinhole) in $\mathbb{P}^3$, while the remaining eight degrees of freedom account for the choice of image coordinates. In this paper we distinguish between traditional *photographic* cameras that use image measurements, and *geometric* ones that do not require fixing coordinate systems, but map points onto the corresponding viewing rays. We work with a generalized notion of camera, where the family of these rays is not necessarily focused at a pinhole. This includes several practical devices, such as pushbroom, panoramic and catadioptric cameras [27].

The main requirement for any camera model is that the fibers of all image points must be lines. This is essential since light travels along lines. With this condition, a *photographic camera* is for us a map $\mathbb{P}^3 \dashrightarrow \mathbb{P}^2$ or $\mathbb{P}^3 \dashrightarrow \mathbb{P}^1 \times \mathbb{P}^1$, where $\mathbb{P}^2$ or $\mathbb{P}^1 \times \mathbb{P}^1$ is the space of image measurements. A *geometric camera* is instead a map $\mathbb{P}^3 \dashrightarrow \mathrm{Gr}(1, \mathbb{P}^3)$ from 3-space into the Grassmannian of lines. The latter is an abstraction of a physical camera, which ignores part of the image formation process, namely the mapping from viewing rays to coordinates. In this paper, we focus mostly on this type of geometric cameras. We will also assume that the coordinates of the map from points to lines are algebraic functions. A geometric camera is always associated with a *congruence* of lines [17], i.e., a two-dimensional family of lines, that is the image of the camera in the Grassmannian $\mathrm{Gr}(1, \mathbb{P}^3)$. Indeed, it has already been argued that congruences should play a central role in multi-view geometry, e.g., [5,24,25]. In this setting, congruences of order one [19] are of particular interest. These define *rational* geometric cameras, where the map from points to image lines is given by rational functions. For example, a pinhole camera is associated with the bundle of lines passing through a fixed point in $\mathbb{P}^3$, and the action of camera takes a point in $\mathbb{P}^3$ to the line joining it to the pinhole. A *two-slit camera* is associated with the common transversals of two lines $\ell_1$ and $\ell_2$ in $\mathbb{P}^3$ (the slits), and taking the picture of a world point $x$ now means mapping $x$ to the line through $x$ that intersects both $\ell_1$ and $\ell_2$. Other rational cameras arise from the common transversals to an algebraic space curve $C$ of degree $d$ and a line $\ell$ meeting $C$ in $d-1$ points.

Taking pictures with $n$ rational cameras for congruences $C_1, \ldots, C_n$ defines a rational map

$$\phi : \mathbb{P}^3 \dashrightarrow C_1 \times C_2 \times \cdots \times C_n \subset (\mathrm{Gr}(1, \mathbb{P}^3))^n \subset (\mathbb{P}^5)^n. \tag{1}$$

The rightmost inclusion is the Plücker embedding of the Grassmannian. The surface $C_i$ now plays the role of the $i$-th image plane $\mathbb{P}^2$ in classical multi-view geometry [2,16]. Our main object of study in this paper is the image of the map $\phi$. This lives in $(\mathrm{Gr}(1, \mathbb{P}^3))^n$ and hence in $(\mathbb{P}^5)^n$. The Zariski closure of this image is an irreducible projective variety of dimension 3. We call this variety the *multi-image variety* of the $n$-tuple of cameras