



HAL
open science

On point configurations, Carlsson-Weinshall duality, and multi-view geometry

Matthew Trager, Martial Hebert, Jean Ponce

► **To cite this version:**

Matthew Trager, Martial Hebert, Jean Ponce. On point configurations, Carlsson-Weinshall duality, and multi-view geometry. 2018. hal-01676732v1

HAL Id: hal-01676732

<https://inria.hal.science/hal-01676732v1>

Preprint submitted on 6 Jan 2018 (v1), last revised 15 Jun 2019 (v4)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On point configurations, Carlsson-Weinshall duality, and multi-view geometry

Matthew Trager^{1,2}, Martial Hebert³, and Jean Ponce^{1,2}

¹Inria

²École Normale Supérieure, CNRS, PSL Research University

³Carnegie Mellon University

Abstract

This paper proposes projective point configurations as a natural setting for studying perspective projection in a geometric, coordinate-free manner. We show that classical results on the effect of permutations on point configurations give a purely synthetic formulation of the well known analytical Carlsson-Weinshall duality between camera pinholes and scene points. We further show that the natural parameterizations of configurations in terms of subsets of their points provides a new and simple analytical formulation of Carlsson-Weinshall duality in any scene and image coordinate systems, not just in the reduced coordinate frames used traditionally. When working in such reduced coordinate systems, we give a new and complete characterization of multi-view geometry in terms of a reduced joint image and its dual. We also introduce a new parametrization of trinocular geometry in terms of reduced trilinearities, and show that, unlike trifocal tensors, these are not subject to any nonlinear internal constraints. This leads to purely linear primal and dual structure-from-motion algorithms, that we demonstrate with a preliminary implementation on real data.

1. Introduction

As shown in [16, 20] for example, point correspondences across multiple images can be characterized by studying incidence relations among the corresponding visual rays. This approach has the merit of making explicit the geometric constraints defining correspondences, which are often hidden behind algebra in the traditional multilinear approaches to structure from motion [2, 8, 6, 7, 11, 12, 13, 18, 22]. In particular, Ponce, Sturmfels and Trager introduce in [16] the *concurrent lines variety* V_n formed by all n -tuples of lines in \mathbb{P}^3 that meet at some point, and completely characterize it in algebraic geometry terms by its dimension $(2n + 3)$

and the generators of its prime ideal. The concurrent lines variety can be used to study multi-view geometry for arbitrary linear (e.g., pinhole or two-slit) cameras and non-linear (e.g., catadioptric) ones, modeled by so-called line congruences. As shown in [16], constraining the lines in each tuple to pass through n fixed and distinct points \mathbf{x}_1 to \mathbf{x}_n amounts to intersecting V_n with $3n$ hyperplanes, yielding a sub-variety $M(\mathbf{x}_1, \dots, \mathbf{x}_n)$ of V_n of dimension 3, the *multi-image variety*, which is isomorphic to Triggs’s joint image [21]. Interestingly, the multi-image variety can either be seen as the set of all possible images taken by n fixed pinhole cameras, or as the set of all possible perspective images of n fixed points, revealing a profound *geometric duality* between camera pinholes and scene points.

Unfortunately, this duality collapses when one introduces image measurements, since the retinal plane of a camera (or equivalently, the line bundle of its pinhole) must be equipped with a coordinate system for these measurements to take place. Individual scene points, on the other hand, are not equipped with physically meaningful coordinate systems. A different, mostly *analytical* take on duality was introduced by Carlsson and Weinshall [2], who showed that, given four fiducial scene points observed by all cameras, the role of camera pinholes and (other) scene points in multi-view geometry could be swapped through appropriate choices of coordinate systems and algebraic manipulations (see [9, 11, 12] for related work). Briefly, assuming that four fiducial points are observed by all cameras of interest, it was shown in [2, 9] that the projection equations reduce to $\mathbf{u} = P_{\mathbf{c}}\mathbf{x} = P_{\hat{\mathbf{x}}}\hat{\mathbf{c}}$, where $P_{\mathbf{c}}$ is a 3-parameter, reduced form of the projection matrix associated with a camera depending only on the position of its pinhole \mathbf{c} , and $z \mapsto \hat{z}$ is an involution ($\hat{\hat{z}} = z$) with a simple analytical form. In particular, as argued in [2, 11], it follows from the reduced projection equations that any algorithm for solving the structure-from-motion (SFM) problem from m images of n scene points also provides a (dual) solution to the SFM problem from $n - 4$ images and $m + 4$ scene points.

1.1. Proposed approach and contributions

The rest of this presentation bridges the gap between this approach and the geometric view advocated at the beginning of the introduction using projective *point configurations* [3] and line geometry as our main tools. Our main contributions can be summarized as follows.

- (1) We introduce in Sect. 2 a new, coordinate-free characterization of perspective projection and Carlsson-Weinshall (or CW for short) duality (Prop. 2.4) based on classical results dating back to a 1915 paper by Coble [3] on the effect of permutations on projective point configurations.
- (2) We show in Sect. 2 that the natural parameterizations of configurations in terms of subsets of their points provides a new and simple analytical formulation of CW duality in *any* scene and image coordinate systems (Prop. 2.7), not just in the reduced coordinate frames used traditionally [2, 9, 11].
- (3) When working in such reduced coordinate systems, we use the simple form of the line projection matrix in this case to give in Sect. 3 a new and complete characterization of multi-view geometry in terms of a *reduced joint image* and its dual (Prop. 3.4).
- (4) We also introduce a new parametrization of trinocular geometry in terms of both primal and dual *reduced trilinearities*, and show in Proposition 3.8 that, unlike trifocal tensors [10, 18, 22], these are not subject to any nonlinear internal constraints [7, 8, 11].
- (5) This paves the way toward a truly linear approach to structure from motion from primal or dual trilinearities. We introduce such an algorithm in Sect. 4, and demonstrate its implementation on real data.

1.2. Background and notation

We assume that the reader is familiar with elementary notions of projective geometry such as projective transformations and homogeneous coordinates. Much of the presentation will distinguish purely geometric, coordinate-free properties of point configurations from analytical properties established in some coordinate system. To avoid confusion, we will thus use a teletype font to designate points in \mathbb{P}^n , *e.g.*, \mathbf{x} , \mathbf{y} , and a bold italic font to designate their homogeneous coordinates in some coordinate frame, *e.g.*, \mathbf{x} , \mathbf{y} . Whether we speak of points or their homogeneous coordinates should thus be clear, and we will often call both representations *points* for simplicity. We will call the first $n + 1$ points of any projective basis $(\mathbf{x}_1, \dots, \mathbf{x}_{n+1}, \mathbf{x}_{n+2})$, with coordinates $(1, 0, \dots, 0)^T$ to $(0, \dots, 0, 1)^T$, the *coordinate points*. The last one, \mathbf{x}_{n+2} , with coordinates $(1, \dots, 1)^T$, is called the *unit point*.

To make the presentation self-contained, let us also recall here some basic concepts of line geometry. The *join operator* associates with two distinct points \mathbf{x} and \mathbf{y} the unique line $\mathbf{x} \vee \mathbf{y}$ passing through these points. Given some coordinate system for \mathbb{P}^3 , this purely geometric operator admits

an analytical counterpart, and the line $\mathbf{l} = \mathbf{x} \vee \mathbf{y}$ joining two points with coordinates $\mathbf{x} = (x_1, \dots, x_4)^T$ and $\mathbf{y} = (y_1, \dots, y_4)^T$ has *Plücker coordinates*

$$\mathbf{l} = \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} \text{ with } \mathbf{u} = \begin{bmatrix} x_4 y_1 - x_1 y_4 \\ x_4 y_2 - x_2 y_4 \\ x_4 y_3 - x_3 y_4 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{bmatrix}. \quad (1)$$

Plücker coordinates are homogeneous. In addition, the vectors \mathbf{u} and \mathbf{v} in (1) are orthogonal by construction, so Plücker coordinates identify the four-dimensional set of lines in \mathbb{P}^3 with a quadratic hypersurface of \mathbb{P}^5 . We refer to the space of lines with this algebraic structure as the Grassmannian of lines $\text{Gr}(1, \mathbb{P}^3)$. Given two lines with Plücker coordinates $\mathbf{l} = (\mathbf{u}; \mathbf{v})$ and $\mathbf{l}' = (\mathbf{u}'; \mathbf{v}')$,¹ one can define an inner product on the Grassmannian as $(\mathbf{l} | \mathbf{l}') = \mathbf{u} \cdot \mathbf{v}' + \mathbf{u}' \cdot \mathbf{v}$, and a necessary and sufficient condition for the two lines to intersect (or, equivalently, be coplanar) is that $(\mathbf{l} | \mathbf{l}') = 0$. The *line bundle* associated with a point \mathbf{x} in \mathbb{P}^3 is the set of lines passing through that point. It forms a two-dimensional projective subspace of $\text{Gr}(1, \mathbb{P}^3)$, (projectively) isomorphic to any plane π not passing through \mathbf{x} , each line in the bundle being associated with the point where it intersects π .

Finally, as shown in [15], a necessary (and generically sufficient) condition for three lines with Plücker coordinates \mathbf{l} , \mathbf{m} , \mathbf{n} to intersect is that the four minors of the 6×3 matrix $[\mathbf{l}, \mathbf{m}, \mathbf{n}]$ defined by

$$T_1 = \begin{vmatrix} l_2 m_2 n_2 \\ l_3 m_3 n_3 \\ l_4 m_4 n_4 \end{vmatrix}, T_2 = \begin{vmatrix} l_3 m_3 n_3 \\ l_1 m_1 n_1 \\ l_5 m_5 n_5 \end{vmatrix}, T_3 = \begin{vmatrix} l_1 m_1 n_1 \\ l_2 m_2 n_2 \\ l_6 m_6 n_6 \end{vmatrix}, T_4 = \begin{vmatrix} l_4 m_4 n_4 \\ l_5 m_5 n_5 \\ l_6 m_6 n_6 \end{vmatrix} \quad (2)$$

all vanish. These minors have been used in [20] to derive trilinear relations among image point coordinates that characterize the fact that the corresponding visual rays intersect, or equivalently, that the points correspond. We take this idea one step further in this presentation.

2. Point configurations and cameras

2.1. Geometric point of view

Consider a finite set of points, observed by a finite set of cameras. Moving together the points and the cameras to any place without changing their relative positions (and for that matter, choosing or changing a coordinate system for \mathbb{P}^3) will not change the images of the scene recorded by the cameras. This is sometimes referred to as the projective ambiguity of structure from motion, but we propose here instead to capture the underlying “projective rigidity” in terms of *projective point configurations*.

Definition 2.1. *Two k -tuples $(\mathbf{x}_1, \dots, \mathbf{x}_k)$ and $(\mathbf{y}_1, \dots, \mathbf{y}_k)$ of points in \mathbb{P}^n are said to be projectively equivalent when*

¹Here $(\mathbf{u}; \mathbf{v})$ is used, Matlab style, to denote the element of \mathbb{R}^6 obtained by stacking the two vectors \mathbf{u} and \mathbf{v} of \mathbb{R}^3 .

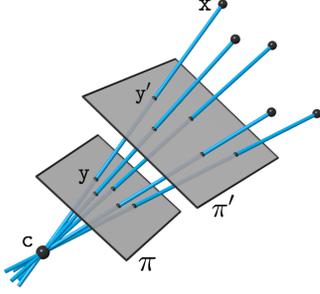


Figure 1: The configuration of a set of image points (or the corresponding visual rays) is uniquely determined by the configuration of the corresponding scene points and the pinhole, independently of the choice of retinal plane.

there exists a projective transformation that maps x_i to y_i for $i = 1, \dots, k$. This is indeed an equivalence relation, and an equivalence class under this relation is called a k -configuration. The configuration associated with k points x_1 to x_k is denoted by $\langle x_1, \dots, x_k \rangle$.

Note that for $k \leq n + 2$, generic point configurations are always equivalent, and we will thus assume $k \geq n + 2$ from now on. The space of point configurations is a classical object of study in mathematics. The first detailed analysis can be found in the work of Coble from 1915 [3] (see also [4] for a more modern account of his work). We will not attempt to summarize this rich and complex theory, but we will draw inspiration from some of Coble’s results to shed new light on Carlsson-Weinshall duality.

Given some pinhole c in \mathbb{P}^3 and some retinal plane π not passing through c , the corresponding perspective projection can be defined in a purely geometric manner as the mapping that associates with any point $x \neq c$ in \mathbb{P}^3 the point y where the viewing ray joining c to x intersects π . As shown by the following lemma, perspective projections can also be seen as mappings between configurations.

Lemma 2.2 (Figure 1). *Let us consider the perspective projection $P_{c,\pi} : \mathbb{P}^3 \setminus \{c\} \rightarrow \pi$ associated with pinhole c and retinal plane π . Given k points x_j ($j = 1, \dots, k$) in $\mathbb{P}^3 \setminus \{c\}$ and their projections $y_j = P_{c,\pi}(x_j)$ into π , the configuration of \mathbb{P}^3 formed by the image points $\langle y_1, \dots, y_k \rangle$ is isomorphic to the configuration of the Grassmannian formed by the visual rays $\langle c \vee x_1, \dots, c \vee x_k \rangle$. In particular, it is independent of π and is uniquely determined by the configuration $\langle x_1, \dots, x_k, c \rangle$.*

This lemma is of course just a retelling of the line bundle properties mentioned in Sect. 1.2 in the language of configurations. It shows however that we may view perspective projection as a way of associating *scene* configurations $\langle x_1, \dots, x_k, c \rangle$ with their *image* counterparts $\langle y_1, \dots, y_k \rangle$. We will sometimes write a scene configura-

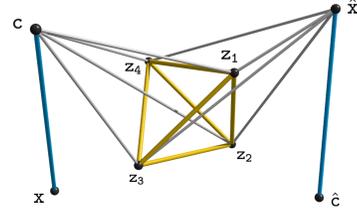


Figure 2: Geometric Carlsson-Weinshall duality: the points z_1, z_2, z_3, z_4, c, x and $z_1, z_2, z_3, z_4, \hat{x}, \hat{c}$ are in the same projective configuration. In particular, the projections of z_1, z_2, z_3, z_4, x from c are equivalent to the projections of $z_1, z_2, z_3, z_4, \hat{x}$ from \hat{c} .

tion as $\langle x_1, \dots, x_k | c \rangle$ instead of $\langle x_1, \dots, x_k, c \rangle$ to emphasize that the last point is viewed as a pinhole. In this setting, swapping the roles of pinhole and scene point results in permuting the corresponding elements of a scene configuration. Studying the effect of permutations on configurations is in fact one of the fundamental questions investigated by Coble in [3]. *Cremona transformations* play a prominent role in this general setting, as illustrated in the particular case of \mathbb{P}^3 most relevant to us by the following result.

Proposition 2.3. *If $Z = (z_1, \dots, z_4)$ is a quadruple of fixed points of \mathbb{P}^3 in general position, there exists a family of birational Cremona involutions $T_Z : x \mapsto \hat{x}$ relative to Z , defined on a dense open set of \mathbb{P}^3 , such that for any points x and y in that set $\langle z_1, z_2, z_3, z_4, x, y \rangle = \langle z_1, z_2, z_3, z_4, \hat{y}, \hat{x} \rangle$. Any two Cremona involutions relative to Z are related to each other by a projective transformation of \mathbb{P}^3 .*

Proof. This follows from the discussions in [3, 4] on the action of the “Weyl group” as birational automorphisms of the point configurations in \mathbb{P}^3 . See in particular [3, Sect. 7] or [4, Chap. 6]. \square

The transformations T_Z are indeed involutions, that is, $\hat{\hat{x}} = x$ for any x where they are defined. Explicit analytical expressions for Cremona involutions will be given in the next section. However, we can already deduce as an immediate corollary of Lemma 2.2 and Prop. 2.3 a coordinate-free form of the Carlsson-Weinshall duality.

Proposition 2.4 (Geometric Carlsson-Weinshall duality, Figure 2). *Denoting as before by \hat{x} the image of a point x by any Cremona involution relative to Z , the two configurations $\langle z_1, z_2, z_3, z_4, x | c \rangle$ and $\langle z_1, z_2, z_3, z_4, \hat{x} | \hat{c} \rangle$ are equal, and thus give rise to the same image configuration $\langle y_1, \dots, y_4, y \rangle$, where y can be thought of as either the projection of x from c or that of \hat{x} from \hat{c} .*

2.2. Analytical point of view

We have so far considered point configurations as purely geometric objects, independent of any choice of coordinates. In order to express our results in analytical form,

we introduce a simple, local parameterization of the space of configurations of k points in \mathbb{P}^n (with $k > n + 1$): we pick $n + 2$ of these points and assign them arbitrary but fixed homogeneous coordinates (it is often convenient, but by no means necessary, to choose these points as a basis for \mathbb{P}^n). Assuming that the points are in general position and the coordinates assigned to any $n + 1$ of them are linearly independent, this uniquely defines a coordinate system for \mathbb{P}^n , dependent on the choice of the $n + 2$ points, but *intrinsic* to the whole configuration. In particular, the coordinates of the $k - n - 2$ remaining points can be used to parameterize the configuration.

In our setting, this translates into assigning arbitrary coordinates to the four fixed points z_1 to z_4 in the form of a 4×4 matrix $Z = [z_1, z_2, z_3, z_4]$, and assigning to the pinhole arbitrary coordinates c . This freezes the coordinate system of \mathbb{P}^3 and provides a parameterization of the configurations $\langle z_1, \dots, z_4, x | c \rangle$ through the coordinates x of the point x . We also pick the four visual rays l_1 to l_4 joining the pinhole to the points z_1 to z_4 as reference points for the corresponding bundle, and assign them arbitrary coordinates in the form of a 3×4 matrix $U = [u_1, u_2, u_3, u_4]$. This freezes the coordinate frame for the bundle and provides a parameterization for the configurations $\langle l_1, \dots, l_4, l \rangle$ of lines in the bundle by the coordinates u of the line l . This also provides, of course, a parameterization of the configurations $\langle y_1, \dots, y_4, y \rangle$ of the corresponding image points through the coordinates u of the point y . We obtain the following result, where $\mathbf{1} = (1, 1, 1, 1)^T$.

Proposition 2.5. *When $Z = Id_4$ and $U = [Id_3, \mathbf{1}]$, the perspective projection associated with the pinhole c can be represented analytically as the projective map from \mathbb{P}^3 to \mathbb{P}^2 , $x \mapsto u = P_C x$, where*

$$P_C = \begin{bmatrix} \frac{1}{c_1} & 0 & 0 & -\frac{1}{c_4} \\ 0 & \frac{1}{c_2} & 0 & -\frac{1}{c_4} \\ 0 & 0 & \frac{1}{c_3} & -\frac{1}{c_4} \end{bmatrix}, \text{ or } u = P_1 \begin{bmatrix} x_1/c_1 \\ x_2/c_2 \\ x_3/c_3 \\ x_4/c_4 \end{bmatrix}. \quad (3)$$

This is the *reduced camera model* appearing in different guises in [2, 9, 11], and it corresponds to taking the reference points z_1 to z_4 as coordinate points for \mathbb{P}^3 and the corresponding visual rays as a basis for the bundle of lines passing through c . The reader is referred to these sources for the proof, which immediately follows from our choice of coordinate systems. Let us emphasize again that the *reduced camera relative to Z* mapping $\langle z_1, \dots, z_4, x | c \rangle$ onto $\langle y_1, \dots, y_4, y \rangle$ is entirely determined by the pinhole c and the four reference points z_1, \dots, z_4 , and can be represented analytically in *any* coordinate system, as shown by the following result.

Proposition 2.6. *Given arbitrary (general) choices for Z and U , let \bar{U} denote the unique projective transform mapping the standard basis of \mathbb{P}^2 onto U , and \bar{Z} denote the*

unique projective transform mapping Z and c onto the standard basis of \mathbb{P}^3 . Then the perspective projection associated with pinhole c can be represented analytically as the projective map from \mathbb{P}^3 to \mathbb{P}^2 , $x \mapsto u = \bar{P}_C x$, where $\bar{P}_C = \bar{U} P_1 \bar{Z} c$.²

This proposition is a direct corollary of Prop. 2.5 under appropriate image and scene coordinate changes. Note that $\bar{P}_C = \bar{P}_1 \bar{Z} c$. Simple, explicit forms for \bar{P}_1 and $\bar{Z} c x$, and thus for the mapping $x \mapsto u$, are easily obtained as

$$\bar{P}_1 = U \text{diag} \begin{bmatrix} -\frac{|u_2, u_3, u_4|}{|u_3, u_4, u_1|} \\ -\frac{|u_4, u_1, u_2|}{|u_1, u_2, u_3|} \end{bmatrix}, \quad \bar{Z} c x = \begin{bmatrix} \frac{z_2 z_3 z_4 x}{z_2 z_3 z_4 c} \\ \frac{z_3 z_4 z_1 x}{z_3 z_4 z_1 c} \\ \frac{z_4 z_1 z_2 x}{z_4 z_1 z_2 c} \\ \frac{z_1 z_2 z_3 x}{z_1 z_2 z_3 c} \end{bmatrix}. \quad (4)$$

We now turn to Carlsson-Weinshall duality and the Cremona involution. Equation (3) is symmetric in x and \hat{c} , where $y \mapsto \hat{y} = (1/y_1, 1/y_2, 1/y_3, 1/y_4)^T$ is the *standard Cremona involution*. It is easy to verify that this is in fact a Cremona involution, in the sense of Prop. 2.3, relative to Z when the points z_j are chosen as coordinate points for \mathbb{P}^3 . See the supplemental material for a proof. More generally, applying an appropriate change of coordinates shows that the Cremona involution for arbitrary values of Z is

$$\hat{y} = Z \begin{bmatrix} \frac{1}{|y, z_2, z_3, z_4|} \\ \frac{|z_1, y, z_3, z_4|}{|z_1, z_2, z_3, y|} \\ \frac{|z_1, z_2, y, z_4|}{|z_1, z_2, z_3, y|} \end{bmatrix}. \quad (5)$$

It is also a simple matter to verify that $\bar{Z} c x = \bar{Z} \hat{x} \hat{c}$. The following proposition immediately follows.

Proposition 2.7 (Analytical Carlsson-Weinshall duality). *If $y \mapsto \hat{y}$ denotes the Cremona transformation relative to Z of (5), and \bar{P}_C is the projection matrix of Prop. 2.6, then the Carlsson-Weinshall duality relation $\bar{P}_C(x) = \bar{P}_{\hat{x}}(\hat{c})$ holds.*

This novel and fully general formulation of Carlsson duality is one of the main contributions of this presentation.

3. Multi-view geometry

We now apply our results to multi-view geometry. As before, we assume that four fixed reference points $Z = (z_1, \dots, z_4)$ are observed by all cameras, but, for simplicity, we restrict our study to the case where $Z = Id_4$ and $U = [Id_3, \mathbf{1}]$ so that all cameras can be represented by projection matrices in “standard” reduced form (3).³ We thus identify from now on scene points x and their images y with their coordinates x and u in \mathbb{P}^3 and \mathbb{P}^2 .

²Although \bar{P}_C obviously depends on U and Z , we leave this dependency implicit in the notation to avoid clutter.

³This choice yields simpler expressions for multilinearities but do not affect in any way the qualitative nature of our results.

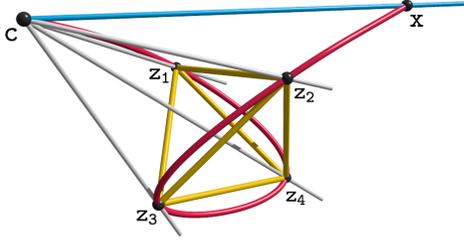


Figure 3: A twisted cubic through \mathbf{x} , \mathbf{z}_1 , \mathbf{z}_2 , \mathbf{z}_3 , \mathbf{z}_4 contains points \mathbf{c} such that the line $\mathbf{c} \vee \mathbf{x}$ has the same coordinates \mathbf{u} with respect to $\mathbf{c} \vee \mathbf{z}_1$, $\mathbf{c} \vee \mathbf{z}_2$, $\mathbf{c} \vee \mathbf{z}_3$, $\mathbf{c} \vee \mathbf{z}_4$.

3.1. Reduced joint images

Let S denote the subset $\mathbb{P}^3 \times \mathbb{P}^3 \times \mathbb{P}^2$ formed by triples $(\mathbf{c}, \mathbf{x}, \mathbf{u})$ such that $P_{\mathbf{c}}(\mathbf{x}) = \mathbf{u}$ where $P_{\mathbf{c}}$ is a reduced camera as in (3). For given values of \mathbf{c} and \mathbf{x} , there is of course a single value of \mathbf{u} such that $(\mathbf{c}, \mathbf{x}, \mathbf{u})$ belongs to S . More generally, we have the following result.

Lemma 3.1. *The following properties of the set S hold:*

(1) *The set S is characterized analytically by the relation*

$$\begin{bmatrix} c_1x_4 - c_4x_1 \\ c_2x_4 - c_4x_2 \\ c_3x_4 - c_4x_3 \end{bmatrix} \times \begin{bmatrix} u_1c_1 \\ u_2c_2 \\ u_3c_3 \end{bmatrix} = 0. \quad (6)$$

(2) *For fixed \mathbf{c} and \mathbf{u} , the set of points \mathbf{x} such that $(\mathbf{c}, \mathbf{x}, \mathbf{u})$ belongs to S is a line (this is obvious from $P_{\mathbf{c}}(\mathbf{x}) = \mathbf{u}$). The Plücker coordinates of this line are given by*

$$\xi = Q_{\mathbf{c}}\mathbf{u} \text{ where } Q_{\mathbf{c}} = \begin{bmatrix} c_1c_4 & 0 & 0 \\ 0 & c_2c_4 & 0 \\ 0 & 0 & c_3c_4 \\ 0 & -c_2c_3 & c_2c_3 \\ c_1c_3 & 0 & -c_1c_3 \\ -c_1c_2 & c_1c_2 & 0 \end{bmatrix}. \quad (7)$$

(3) *For fixed \mathbf{x} and \mathbf{u} , the set of points \mathbf{c} such that $(\mathbf{c}, \mathbf{x}, \mathbf{u})$ belongs to S is a twisted cubic passing through $\mathbf{z}_1, \dots, \mathbf{z}_4$ and \mathbf{x} (Figure 3).*

Variants of this lemma can be found in the literature [2, 9, 11]. The explicit formula for $Q_{\mathbf{c}}$ is of course an instance of the classical (transposed) line projection matrix (cf. the supplemental material for a derivation), although we are not aware of this expression appearing before. Equation (7) will play a key role in the rest of this presentation.

We now consider n pinholes $\mathbf{c}_1, \dots, \mathbf{c}_n$, and the associated system of reduced cameras $P_{\mathbf{c}_1}, \dots, P_{\mathbf{c}_n}$ as in (3). Following [19, 21], we describe the multi-view geometry of these cameras using the *joint image* in $(\mathbb{P}^2)^n$.

Definition 3.2. *The reduced joint image $V_Z(\mathbf{c}_1, \dots, \mathbf{c}_n)$ associated with n fixed pinholes \mathbf{c}_1 to \mathbf{c}_n is the set of tuples $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ in $(\mathbb{P}^2)^n$ such that the corresponding visual*

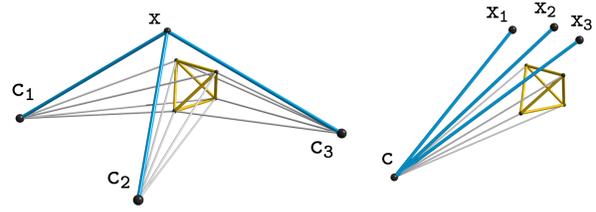


Figure 4: A reduced joint image $V_Z(\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3)$ characterizes all converging visual rays from three pinholes $\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3$, using four fixed points to determine reference frames in each bundle (left). A dual reduced joint image $\hat{V}_Z(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$, characterizes all perspective images of three scene points $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$, using four fixed points for reference (right).

rays for $P_{\mathbf{c}_i}$ are concurrent or, equivalently, there exists some point \mathbf{x} such that $(\mathbf{c}_i, \mathbf{x}, \mathbf{u}_i)$ belongs to S .

Note that a reduced joint image is a special case of Triggs's joint image [21], with the additional property that the basis points of the image (or bundle) coordinates are in correspondence. Moreover, projective transformations in \mathbb{P}^3 do not affect a joint image, so a reduced joint image is completely determined by the point configuration $\langle \mathbf{c}_1, \dots, \mathbf{c}_n, \mathbf{z}_1, \dots, \mathbf{z}_4 \rangle$. See Figure 4 (left). Let us now consider instead n fixed scene points $\mathbf{x}_1, \dots, \mathbf{x}_n$.

Definition 3.3. *The dual reduced joint image $\hat{V}_Z(\mathbf{x}_1, \dots, \mathbf{x}_n)$ associated with n scene points \mathbf{x}_1 to \mathbf{x}_n is the set of tuples $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ in $(\mathbb{P}^2)^n$ that are image coordinates for the points \mathbf{x}_1 to \mathbf{x}_n for some reduced camera $P_{\mathbf{c}}$ with (unknown) pinhole \mathbf{c} or, equivalently, there exists some point \mathbf{x} such that $(\mathbf{c}, \mathbf{x}_i, \mathbf{u}_i)$ belongs to S . Note that this condition imposes that n twisted cubics passing through $\mathbf{z}_1, \dots, \mathbf{z}_4$ intersect at a point \mathbf{c} .*

The set $\hat{V}_Z(\mathbf{x}_1, \dots, \mathbf{x}_n)$ can be thought of as characterizing all perspective images of n fixed points (Figure 4, right). The dual reduced joint image is invariant to projective transformations of \mathbb{P}^3 , and is completely determined by the point configuration $\langle \mathbf{x}_1, \dots, \mathbf{x}_n, \mathbf{z}_1, \dots, \mathbf{z}_4 \rangle$ in \mathbb{P}^3 . Algebraic characterizations of dual multi-view constraints are mostly absent from the literature (see [12] for an exception), but the following result is an immediate corollary of Carlsson-Weinshall duality.

Proposition 3.4. *The dual reduced joint image associated with n scene points \mathbf{x}_1 to \mathbf{x}_n is the reduced joint image associated with their images $\hat{\mathbf{x}}_1$ to $\hat{\mathbf{x}}_n$ under any Cremona involution relative to $\mathbf{z}_1, \dots, \mathbf{z}_4$.*

In particular, constraints for dual joint images are the same as for (reduced) joint images: these are the well known multilinearities, as studied for example in [1, 19].

3.2. Reduced multilinearity

Let us now apply the general approach presented so far to the specific case of bilinear and trilinear constraints associated with point correspondences for reduced cameras of the form (3). Given two points \mathbf{u} and \mathbf{u}' associated with cameras with pinhole $\mathbf{c} = (1, 1, 1, 1)^T$ and $\mathbf{c}' = (c'_1, c'_2, c'_3, c'_4)^T$, a necessary and sufficient condition for the points \mathbf{u} and \mathbf{u}' to form a correspondences is that the corresponding visual rays $\mathbf{l} = Q_{\mathbf{c}}\mathbf{u}$ and $\mathbf{l}' = Q_{\mathbf{c}'}\mathbf{u}'$ intersect each other, or equivalently, that $(\mathbf{l} | \mathbf{l}') = 0$. This immediately yields the bilinear relation $\mathbf{u}^T F \mathbf{u}' = 0$, where F is the *reduced fundamental matrix*

$$F = \begin{bmatrix} 0 & c'_2(c'_4 - c'_3) & c'_3(c'_2 - c'_4) \\ c'_1(c'_3 - c'_4) & 0 & c'_3(c'_4 - c'_1) \\ c'_1(c'_4 - c'_2) & c'_2(c'_1 - c'_4) & 0 \end{bmatrix} \quad (8)$$

of [2, 11] (see also [9]). In turn, substituting $\hat{\mathbf{x}}', \hat{\mathbf{x}}''$ for $\mathbf{c}', \mathbf{c}''$ in (8), we also obtain the expression for the *dual fundamental matrix*, which characterizes the dual reduced joint image $\hat{V}_Z(\mathbf{x}, \mathbf{x}')$ for two fixed scene points \mathbf{x}, \mathbf{x}' .

We can use the same approach to characterize correspondences in three images. Indeed, using the expression (7) for $Q_{\mathbf{c}}$ and the trilinearities from (2) to write four necessary conditions $T_1 = T_2 = T_3 = T_4 = 0$ for three lines to converge at a point yields the following result.

Lemma 3.5. *Let $\mathbf{c} = (1, 1, 1, 1)^T$, $\mathbf{c}' = (c'_1, c'_2, c'_3, c'_4)^T$ and $\mathbf{c}'' = (c''_1, c''_2, c''_3, c''_4)^T$. A necessary (and generically sufficient) condition for a triple $\mathbf{u} = (u_1, u_2, u_3)$, $\mathbf{u}' = (u'_1, u'_2, u'_3)$, and $\mathbf{u}'' = (u''_1, u''_2, u''_3)$ to form a correspondence for the reduced cameras $P_{\mathbf{c}}, P_{\mathbf{c}'}, P_{\mathbf{c}''}$, is that $T_1 = T_2 = T_3 = T_4 = 0$, where*

$$\begin{aligned} T_1 &= \begin{vmatrix} u_2 & c'_3 u'_2 & c''_3 u''_2 \\ u_3 & c'_2 u'_3 & c''_2 u''_3 \\ v_1 & c'_4 v'_1 & c''_4 v''_1 \end{vmatrix}, T_2 = \begin{vmatrix} u_3 & c'_1 u'_3 & c''_1 u''_3 \\ u_1 & c'_3 u'_1 & c''_3 u''_1 \\ v_2 & c'_4 v'_2 & c''_4 v''_2 \end{vmatrix}, \\ T_3 &= \begin{vmatrix} u_1 & c'_2 u'_1 & c''_2 u''_1 \\ u_2 & c'_1 u'_2 & c''_1 u''_2 \\ v_3 & c'_4 v'_3 & c''_4 v''_3 \end{vmatrix}, T_4 = \begin{vmatrix} v_1 & c'_1 v'_1 & c''_1 v''_1 \\ v_2 & c'_2 v'_2 & c''_2 v''_2 \\ v_3 & c'_3 v'_3 & c''_3 v''_3 \end{vmatrix}. \end{aligned} \quad (9)$$

where $\hat{x} = 1/x$, and $v_i = u_{i+2} - u_{i+1}$ (index addition is done modulo 3), and similarly for v'_i and v''_i .

The lemma follows immediately from (2) and the form of the matrix $Q_{\mathbf{c}}$. The dual of this statement involves three scene points $\mathbf{x}, \mathbf{x}', \mathbf{x}''$ instead of three pinholes.

Lemma 3.6. *Let $\mathbf{x} = (1, 1, 1, 1)^T$, $\mathbf{x}' = (x'_1, x'_2, x'_3, x'_4)^T$ and $\mathbf{x}'' = (x''_1, x''_2, x''_3, x''_4)^T$. A necessary (and generically sufficient) condition for a triple $\mathbf{u}, \mathbf{u}', \mathbf{u}''$ to be projections of $\mathbf{x}, \mathbf{x}', \mathbf{x}''$ for a reduced camera $P_{\mathbf{c}}$ (for some unknown \mathbf{c})*

is that $\hat{T}_1 = \hat{T}_2 = \hat{T}_3 = \hat{T}_4 = 0$, where

$$\begin{aligned} \hat{T}_1 &= \begin{vmatrix} u_2 & x'_3 u'_2 & x''_3 u''_2 \\ u_3 & x'_2 u'_3 & x''_2 u''_3 \\ v_1 & x'_4 v'_1 & x''_4 v''_1 \end{vmatrix}, \hat{T}_2 = \begin{vmatrix} u_3 & x'_1 u'_3 & x''_1 u''_3 \\ u_1 & x'_3 u'_1 & x''_3 u''_1 \\ v_2 & x'_4 v'_2 & x''_4 v''_2 \end{vmatrix}, \\ \hat{T}_3 &= \begin{vmatrix} u_1 & x'_2 u'_1 & x''_2 u''_1 \\ u_2 & x'_1 u'_2 & x''_1 u''_2 \\ v_3 & x'_4 v'_3 & x''_4 v''_3 \end{vmatrix}, \hat{T}_4 = \begin{vmatrix} v_1 & x'_1 v'_1 & x''_1 v''_1 \\ v_2 & x'_2 v'_2 & x''_2 v''_2 \\ v_3 & x'_3 v'_3 & x''_3 v''_3 \end{vmatrix}. \end{aligned} \quad (10)$$

These expressions can be obtained by substituting $\hat{\mathbf{x}}$ for \mathbf{c} in the primal trilinearities from Lemma 3.5.

3.3. Main properties of reduced trilinearities

All (primal or dual) reduced trilinearities have the same form. Let us start by focusing on T_4 . From (9), we see that the determinant is the sum of *six* terms that are, up to sign, given by $a_j b_k v_i v'_j v''_k$, for distinct i, j, k indices, and where $a_i = c'_i$, $b_i = c''_i$. This allows us to view T_4 as a trilinear polynomial in v_i, v'_i, v''_i , or as a linear form in the six coefficients $\rho_{ij} = a_i b_j$ with $i, j = 1, 2, 3$ and $i \neq j$ (which in turn are obviously bilinear in a_i, b_j). In the following, we write $\boldsymbol{\rho}$ for the vector of \mathbb{R}^6 with entries ρ_{ij} in lexicographic order and $\mathbf{1}$ for the vector of \mathbb{R}^6 whose components are all equal to one.

Lemma 3.7. *A necessary and sufficient condition for the components ρ_{ij} of a vector $\boldsymbol{\rho}$ in \mathbb{R}^6 to factor as $\rho_{ij} = a_i b_j$ for some vectors $\mathbf{a} = (a_1, a_2, a_3)^T$ and $\mathbf{b} = (b_1, b_2, b_3)^T$ is that $\rho_{12}\rho_{23}\rho_{31} = \rho_{13}\rho_{21}\rho_{32}$. When this condition is satisfied, \mathbf{a} and \mathbf{b} can uniquely be determined up to independent scales. When v_i, v'_i, v''_i are obtained from image coordinates as in Lemma 3.5 (i.e., are of the form $v_i = u_{i+2} - u_{i+1}$ for some u_i , and similarly for v'_i, v''_i), the equation $T_4 = 0$ in $\boldsymbol{\rho}$ always admits $\mathbf{1}$ as a solution.*

Proof. The first part of the proposition follows from direct computations. For the second part, we observe that $\boldsymbol{\rho} = \mathbf{1}$ implies, up to scale, $\mathbf{a} = (1, 1, 1)^T$, and $\mathbf{b} = (1, 1, 1)^T$. We can now verify that T_4 is always zero when we substitute $\hat{c}'_i = \hat{c}''_i = 1$ and $v_i = u_{i+2} - u_{i+1}$ (because we obtain the determinant of a matrix of rank two). \square

Similar statements can be obtained for all other reduced trilinearities, with cubic constraints, and $\mathbf{1}$ as a trivial solution. In the next section, we will use this result and combine the linear constraints on ρ_{ij} arising from all four (primal or dual) trilinearities. Despite the cubic constraint, we will show that it is possible to give a purely linear method for recovering the pinholes (or their duals) from these equations.

The trilinear constraints T_i can also be viewed as polynomials in the raw image coordinates u_i, u'_i, u''_i , without introducing the auxiliary variables v_i, v'_i, v''_i . In this case, it is less obvious to spell out the coefficients in closed form, and the terms of T are in general $p_{ijk}(a_j, b_k)u_i u'_j u''_k$ (but

i, j, k need not be distinct) where $p_{ijk}(a_j, b_k)$ is a quadratic polynomial in a_i, b_i . These (a priori 27) polynomials are different for each T_i . In fact, by explicit calculation, we verify that T_1, T_2, T_3 (and their duals) have six non-zero (coefficient) polynomials $p_{ijk}(a_j, b_k)$ in a_i, b_i (different from the ones considered above), except for T_4 with 24 non-zero terms in this case. The following result is analogous to Lemma 3.7 in this setting.

Proposition 3.8. *For each T_i , the set of coefficients $\tau = (\tau_{ijk})$ in \mathbb{R}^{27} that can be expressed in the form $\tau_{ijk} = p_{ijk}(a_j, b_k)$ for some vectors \mathbf{a}, \mathbf{b} of \mathbb{R}^3 forms a vector space of dimension five.*

For example, for T_1, T_2, T_3 , the (appropriate non-zero) six coefficients τ_{ijk} satisfy $\tau_{ijk} = p_{ijk}(a_j, b_k)$ for some \mathbf{a}, \mathbf{b} if and only if they sum to zero. For T_4 , the 24 coefficients need to satisfy 19 linear constraints. These facts can be proven by direct computation with a computer algebra system, but they also admit a geometric justification, briefly sketched in the rest of this section, which can also be used to deduce all of the formulae and results from this section and the previous one. The reader is referred to the supplementary material for additional details.

The basic observation is that the six nonzero coefficients of T_1 (for example) can be viewed as the coefficients of a 2D reduced trifocal tensor. As explained for example in [17], a 2D trifocal tensor is a $2 \times 2 \times 2$ tensor that encodes correspondence for three projections from \mathbb{P}^2 to \mathbb{P}^1 . Algebraic properties of 2D trifocal tensors are well understood, and much simpler than for 3D trifocal tensors. In particular, 2D trifocal tensors satisfy *no internal constraints*, i.e., any $2 \times 2 \times 2$ tensor arises from a configuration of projections.

It can be shown that each T_i encodes the convergence of three coplanar lines passing through three fixed points (cf. supplementary material). These lines and points are obtained by considering a projection with center z_i (Figure 5). This leads to a 2D SFM problem, which is “reduced” because we can use the projections of the remaining three reference points z_j as reference points in \mathbb{P}^2 (shown as yellow triangles in Figure 5, in analogy with Figure 4). This situation is described by a “reduced” trifocal tensor, which has three “synthetic” correspondences, and has thus five linearly independent parameters. As a final note, we wish to remark that Prop. 3.8 appears to be a significant advantage of reduced trilinearities over (traditional) trifocal tensors [10, 18, 22], since they are not subject to the complex algebraic *internal constraints* of the latter [7, 8, 11].⁴ This paves the way toward truly linear approaches to structure from motion, as illustrated in the next section.

⁴Note that even (reduced) primal or dual fundamental matrices obey such a constraint, i.e., that their determinant must be zero.

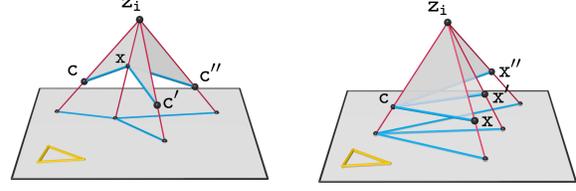


Figure 5: The trilinearities T_i (or \hat{T}_i) express the condition that three coplanar lines intersect. The three lines are projections from z_i of three viewing rays converging at some (unknown) scene point x in the primal case (left), and some (unknown) pinhole c in the dual one (right). The situation on the image plane is a lower-dimensional analogue of the reduced joint image and dual reduced joint image shown in Figure 4. The yellow triangle plays the same role as the tetrahedron in the 3D case.

4. Linear SFM from reduced trilinearities

Let us show that the four trilinearities (either primal or dual) can be used to solve the structure-from-motion problem when sufficient correspondences is available. Let us start with the primal trilinearities. As argued in the previous section, we can substitute triplets of correspondences in each trilinearity and obtain linear equations in the vector $\rho = (\rho_{12}, \rho_{13}, \dots, \rho_{43})^T$ of \mathbb{R}^{12} where $\rho_{ij} = \tilde{c}'_i \tilde{c}'_j$. More precisely, given p correspondences we can stack the four equations (trilinearities) in ρ associated with each correspondence to form a $4p \times 12$ matrix T and write all the constraints as $T\rho = 0$. However, it follows from Prop. 3.7 that the vector $\mathbf{1} = (1, \dots, 1)^T$ of \mathbb{R}^{12} always satisfies $T\mathbf{1} = 0$, independently of the image coordinate values. In the absence of noise, the kernel of T is thus always (at least) two-dimensional, since it contains $\mathbf{1}$ and the “true” solution ρ to our problem. We address this issue by exploiting the special form of the vector ρ . In particular, let e be a vector orthogonal to $\mathbf{1}$ in the kernel, so that we may write $\rho = e + \lambda \mathbf{1}$ for some scalar λ , and thus $\rho_{ij} = \tilde{c}'_i \tilde{c}'_j = e_{ij} + \lambda$ for $j \neq i$. We have, for example $\tilde{c}'_3 = (e_{13} + \lambda)/c'_1 = (e_{23} + \lambda)/c'_2$, and $\tilde{c}'_4 = (e_{14} + \lambda)/c'_1 = (e_{24} + \lambda)/c'_2$. This allows us to eliminate λ , obtaining $(e_{23} - e_{24})c'_1 + (e_{14} - e_{13})c'_2 = 0$. Collecting all similar equations finally yields

$$\begin{bmatrix} e_{23} - e_{24} & e_{14} - e_{13} & 0 & 0 \\ e_{32} - e_{34} & 0 & e_{14} - e_{12} & 0 \\ e_{42} - e_{43} & 0 & 0 & e_{13} - e_{12} \\ 0 & e_{31} - e_{34} & e_{24} - e_{21} & 0 \\ 0 & e_{43} - e_{41} & 0 & e_{21} - e_{23} \\ 0 & 0 & e_{41} - e_{42} & e_{32} - e_{31} \end{bmatrix} \tilde{c}' = 0, \quad (11)$$

and

$$\begin{bmatrix} e_{32} - e_{42} & e_{41} - e_{31} & 0 & 0 \\ e_{23} - e_{43} & 0 & e_{41} - e_{21} & 0 \\ e_{24} - e_{34} & 0 & 0 & e_{31} - e_{21} \\ 0 & e_{13} - e_{43} & e_{42} - e_{12} & 0 \\ 0 & e_{14} - e_{34} & 0 & e_{32} - e_{12} \\ 0 & 0 & e_{14} - e_{24} & e_{23} - e_{13} \end{bmatrix} \tilde{c}'' = 0, \quad (12)$$

which are sufficient to determine the two points \hat{c}' and \hat{c}'' and thus c' and c'' . We can thus impose the constraint $\rho \cdot \mathbf{1} = 0$ when solving for ρ . Three correspondences are necessary to obtain a unique solution for the 11 unknowns. Note that although the vector ρ must satisfy certain algebraic constraints in order to be of the form $\rho_{ij} = c'_i c''_j$ (cf. Proposition 3.7), our strategy bypasses this difficulty by directly (linearly) recovering the vectors c' and c'' . In other words, this linear method is always guaranteed to return a valid (approximate in the presence of noise) solution.

In practice, we pick four reference points among the (known) correspondences between these pictures, and apply appropriate image coordinate changes so they become basis points. We then use singular value decomposition and the remaining points to find the least-squares solution e of the system of equation in ρ associated with T which is orthogonal to $\mathbf{1}$, and finally use Eqs. (11) and (12) to compute the position of the dual pinholes and from those the positions of the pinholes themselves. Linear least-squares can then be used again to reconstruct all scene points from the known pinholes and image coordinates. We repeat this process for random quadruples of reference points and a fixed number of iterations, and report the results.

We conducted experiments on the Inria toy house dataset [14], that consists of six images of 38 points. These experiments are of course preliminary, and need to be extended to characterize sensitivity to varying levels of noise, and to evaluate performance on other datasets. Figure 6 (top) shows the results of the primal reconstruction algorithm outlined above. The reconstruction is projectively registered with the ground-truth 3D structure available for this dataset for display purposes. Results are quite reasonable after 50 trials. The figure shows results for 5, 50, 500, and 5,000 different choices of basis points. Here, the criterion for picking the best basis is minimizing the mean of the reprojection errors overs the three images in their original coordinate systems (without of course using 3D ground-truth information). The corresponding mean errors are respectively 1.7, 0.9, 0.9, and 0.8 pixel. Even though the error is noticeably larger after only 5 random picks, the reconstruction itself is quite reasonable. Indeed the mean distance between the ground-truth 3D points and the (registered) reconstructed ones, relative to the radius of the smallest sphere centered at the mean of the ground-truth points and enclosing them is quite small, respectively 1.7%, 1.8%, 0.6% and 0.7% for the various iterations. This error does not decrease monotonically, which is not surprising since the reconstruction error is not used as a criterion to pick the reference points.

The dual algorithm is very similar except that this time we fix three points x , x' and x'' instead of three images. We again repeatedly pick four random references points, and use all images of the three points x , x' and x'' to recon-

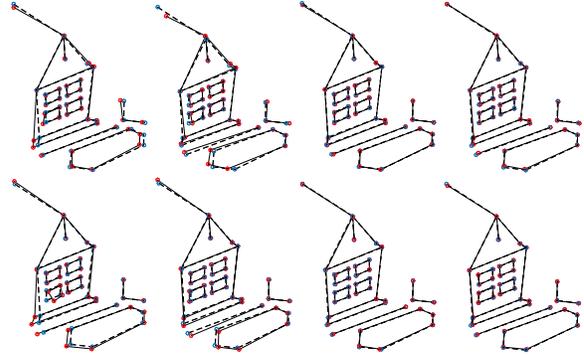


Figure 6: Registered reconstruction (in red) vs. ground truth (in blue) for the primal (top) and dual (bottom) versions of our algorithm. The results are shown, from left to right, for 5, 50, 500, and 5,000 different choices of image basis points. Best seen on a computer screen with zooming.

struct them. Once these have been estimated, it is a simple matter to reconstruct the pinholes and then the scene using linear least squares. Figure 6 (bottom) shows an example of reconstruction obtained with this method for 5, 50, 500 and 5,000 different choices for the four reference points. The corresponding errors are respectively 18.0, 3.8, 2.4 and 1.5 pixel. The three points x , x' and x'' themselves were obtained by minimizing the reprojection error among 5×10 choices of point triplets and image quadruplets. The reconstructions are again rather good even with a small number of random choices, as shown by the figure, and confirmed by computing the relative mean error between the ground-truth 3D points and the reconstructed ones. This time we obtain errors 1.6%, 1.2%, 0.7% and 0.6% for the various iterations.

5. Conclusion

We have presented in this paper a new approach to the study of duality between scene points and pinholes and multi-view geometry in terms of projective configurations, and shown that it led to a new algorithm for primal and dual structure from motion that gives satisfactory solutions on real data. Several natural extensions of our work can be envisioned. From the theoretical standpoint, it would be interesting to strengthen our ties to classical work on point configurations. In particular, Coble describes in detail an “association” between point configurations in different spaces (often known today as “Gale duality” [5]), and we believe that this could be a useful concept for multi-view geometry as well. For example, this association maps configurations of six points in \mathbb{P}^3 (which always lie on a twisted cubic) to configurations of six points on a line, in a way that reflects the relative position of the original points on the twisted cubic (compare this with Figure 3). From a more practical perspective, we hope to further develop our SFM

algorithms. As noted for example in [11], a general difficulty of “reduced” SFM methods is that changes of image coordinates affect the error distribution in measurements, which often causes poor performance in traditional estimation techniques. In our implementation, we have addressed this issue by picking random sets of reference points, and picking the one that minimizes the reprojection error, in a manner reminiscent of RANSAC, and we believe this to be a simple and effective solution in practice. On the other hand, our analysis in Sect. 3 includes formulae for reduced projections in arbitrary coordinate frames (4), which allow in principle to carry out reduced algorithms in original image coordinates. It will be interesting to see whether this can be used to improve the results of the estimation.

Acknowledgments. This work was supported in part by the ERC grant VideoWorld, the Institut Universitaire de France, the Inria-CMU associated team GAYA, and ONR MURI N000141010934.

References

- [1] C. Aholt, B. Sturm, and R. Thomas. A Hilbert scheme in computer vision. *Canadian Journal of Mathematics*, 65:961–988, 2013.
- [2] S. Carlsson and D. Weinshall. Dual computation of projective shape and camera positions from multiple images. *IJCV*, 27(3):227–241, 1998.
- [3] Arthur B Coble. Points sets and allied cremona groups (part i). *Transactions of the American Mathematical Society*, 16(2):155–198, 1915.
- [4] Igor Dolgachev. Point sets in projective spaces and theta functions. *Astérisque*, 165, 1988.
- [5] David Eisenbud and Sorin Popescu. The projective geometry of the gale transform. *Journal of algebra*, 230(1):127–173, 2000.
- [6] O. Faugeras, Q.-T. Luong, and T. Papadopoulos. *The Geometry of Multiple Images*. MIT Press, 2001.
- [7] O. Faugeras and B. Mourrain. On the geometry and algebra of the point and line correspondences between n images. Technical Report 2665, INRIA Sophia-Antipolis, 1995.
- [8] O. Faugeras and T. Papadopoulos. Grassman-Cayley algebra for modeling systems of cameras and the algebraic equations of the manifold of trifocal tensors. Technical Report 3225, INRIA Sophia-Antipolis, 1997.
- [9] O.D. Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig? In G. Sandini, editor, *ECCV*, volume 588 of *Lecture Notes in Computer Science*, pages 563–578, Santa Margherita, Italy, 1992. Springer-Verlag.
- [10] R. Hartley. Lines and points in three views and the trifocal tensor. *IJCV*, 22(2):125–140, 1997.
- [11] R. Hartley and A. Zisserman. *Multiple view geometry in computer vision*. Cambridge University Press, 2nd edition, 2004.
- [12] A. Levin and A. Shashua. Revisiting single-view shape tensors: Theory and applications. In *ECCV*, 2002.
- [13] Q.-T. Luong and O.D. Faugeras. The fundamental matrix: theory, algorithms, and stability analysis. *IJCV*, 17(1):43–76, 1996.
- [14] R. Mohr, L. Quan, F. Veillon, and B. Boufama. Relative 3D reconstruction using multiple uncalibrated images. Technical Report RT 84-IMAG 12-LIFIA, LIFIA-IRIMAG, June 1992.
- [15] J. Ponce, T. Papadopoulos, M. Teillaud, and B. Triggs. The absolute quadratic complex and its application to camera self calibration. In *CVPR*, 2005.
- [16] J. Ponce, B. Sturm, and M. Trager. Congruences and concurrent lines in multi-view geometry. *Advances in Applied Mathematics*, 88:62–91, 2017.
- [17] Long Quan. Two-way ambiguity in 2d projective reconstruction from three uncalibrated 1d images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(2):212–216, 2001.
- [18] A. Shashua. Algebraic functions for recognition. *PAMI*, 17(8):779–789, 1995.
- [19] M. Trager, M. Hebert, and J. Ponce. The joint image handbook. In *ICCV*, 2015.
- [20] M. Trager, J. Ponce, and M. Hebert. Trinocular geometry revisited. *IJCV*, 120(2):134–152, 2016.
- [21] B. Triggs. Matching constraints and the joint image. In *ICCV*, 1995.
- [22] J. Weng, T.S. Huang, and N. Ahuja. Motion and structure from line correspondences: closed-form solution, uniqueness, and optimization. *PAMI*, 14(3):318–336, 1992.