Coordinate-Free Carlsson-Weinshall Duality and Relative Multi-View Geometry

Supplementary Material

Matthew Trager\textsuperscript{1}, Martial Hebert\textsuperscript{2}, and Jean Ponce\textsuperscript{3,4}

\textsuperscript{1}New York University \textsuperscript{2}Carnegie Mellon University \textsuperscript{3}INRIA, Paris, France
\textsuperscript{4}Département dinformatique de IENS, ENS, CNRS, PSL University, Paris, France

We present some proofs that were not included in the main part of the paper and some additional results of our algorithm on synthetic data.

1. Analytical cameras and duality

The analytical expression for $P_c(x)$ given in Proposition 2.4 follows easily from the following general result.

**Proposition 1.1.** The unique projective transformation of $\mathbb{P}^n$ that maps $n+2$ general points $x_1, \ldots, x_{n+2}$ to the standard basis of $\mathbb{P}^n$ can be described by

$$ y \mapsto \begin{bmatrix} x_2 \cdots x_n \cdots x_{n+2} \\ x_2 \cdots x_{n+1} \cdots x_{n+2} y \\ \vdots \\ x_1 \cdots x_{n+2} \\ x_1 \cdots x_{n+1} y \\ \end{bmatrix}, \quad (1) $$

**Proof.** It is clear that (1) describes a projective transformation, since all the expressions are linear in the coordinates of $y$. It also follows from elementary properties of determinants that (1) maps $x_1, \ldots, x_{n+2}$ to the standard basis. \hfill $\square$

We now turn to Cremona transformations and justify the general formula (5) given in the paper.

**Proposition 1.2.** The rational map of $\mathbb{P}^3$ given by

$$ y = \begin{bmatrix} y_1 x_2 x_3 x_4 \\ y_1 y_2 x_3 x_4 \\ y_2 x_1 x_3 x_4 \\ y_1 y_3 x_2 x_4 \\ \end{bmatrix}, \quad (2) $$

where $Z$ has columns $z_1, \ldots, z_4$, is a Cremona transformation relative to $z_1, z_2, z_3, z_4$ (in the sense of Lemma 2.2 from the main part of the paper).

**Proof.** We need to show that $z_1, \ldots, z_4, y_1, y_2$ are in the same projective configuration as $z_1, \ldots, z_4, y_2, y_1$. Indeed, a projective transformation relating the two sets of points is given by

$$ x \mapsto Z \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ \end{bmatrix}, \quad (3) $$

where $y_1, y_2$ are considered fixed. Note that when $z_1, \ldots, z_4$ are basis points then (2) yields the standard Cremona transformation

$$ y = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ \end{bmatrix}^T, \quad (4) $$

and (3) becomes

$$ x \mapsto \begin{bmatrix} x_1 \\ \frac{x_2}{y_1 y_2} \\ \frac{x_3}{y_1 y_2 y_3} \\ \frac{x_4}{y_1 y_2 y_3 y_4} \end{bmatrix}^T, \quad (5) $$

where $x = (x_1, x_2, x_3, x_4)^T$, $y_1 = (y_1, y_1, y_1, y_1)^T$ and $y_2 = (y_2, y_2, y_2, y_2)^T$.

We next prove Proposition 3.1 from the main part of the paper. We recall that the set $S$ in $\mathbb{P}^1 \times \mathbb{P}^3 \times \mathbb{P}^2$ was defined as the set of triples $(c, x, u)$ such that $P_c(x) = u$ where $P_c$ is a reduced camera.

**Proposition 3.1.** (1) For fixed $c$ and $u$, the set of points $x$ such that $(c, x, u)$ belongs to $S$ is a line with Plücker coordinates

$$ \xi = Q_c u \quad \text{where} \quad Q_c = \begin{bmatrix} c_1 c_4 & 0 & 0 \\ c_2 c_4 & 0 & 0 \\ 0 & c_3 c_4 & 0 \\ 0 & c_2 c_3 & c_2 c_3 \\ -c_1 c_2 & -c_1 c_3 & 0 \end{bmatrix}. \quad (6) $$

(2) For fixed $x$ and $u$, the set of points $c$ such that $(c, x, u)$ belongs to $S$ is a twisted cubic passing through $z_1, \ldots, z_4$ and $x$. 
Proof. To show (1), we define $\xi_j = c \lor z_j$ (j = 1, ..., 4). With our choice of coordinate system, we have

$$
\begin{bmatrix}
  c_4 \\
  0 \\
  0 \\
  -c_2 \\
\end{bmatrix}, \quad 
\begin{bmatrix}
  0 \\
  c_4 \\
  0 \\
  -c_3 \\
\end{bmatrix}, \quad 
\begin{bmatrix}
  0 \\
  0 \\
  c_4 \\
  -c_1 \\
\end{bmatrix}, \quad 
\begin{bmatrix}
  c_1 \\
  c_2 \\
  c_3 \\
  0 \\
\end{bmatrix}.
\end{equation}

(7)

We write $\xi = \rho_1 u_1 \xi_1 + \rho_2 u_2 \xi_2 + \rho_3 u_3 \xi_3$, where the scalars $\rho_1, \rho_2, \rho_3$ have been chosen so that $\xi_4 = \rho_1 \xi_1 + \rho_2 \xi_2 + \rho_3 \xi_3$. A simple calculation shows that $\rho_1 = -c_1/c_4, \rho_2 = -c_2/c_4, \rho_3 = -c_3/c_4$, and (6) immediately follows.

For the second statement, we first note that the set $S$ is characterized algebraically by the relation

$$
\text{rk } \begin{bmatrix}
  c_1 x_4 - c_4 x_1 & u_1 c_1 \\
  c_2 x_4 - c_4 x_2 & u_2 c_2 \\
  c_3 x_4 - c_4 x_3 & u_3 c_3 \\
\end{bmatrix} = 1.
\text{rk } \begin{bmatrix}
  c_1 x_4 - c_4 x_1 & u_1 c_1 \\
  c_2 x_4 - c_4 x_2 & u_2 c_2 \\
  c_3 x_4 - c_4 x_3 & u_3 c_3 \\
\end{bmatrix} = 1
\end{equation}

(8)

This expression follows from

$$
\begin{bmatrix}
  u_1 \\
  u_2 \\
  u_3 \\
\end{bmatrix} \sim P c(x) = \begin{bmatrix}
  x_1/c_1 - x_4/c_4 \\
  x_2/c_2 - x_4/c_4 \\
  x_3/c_3 - x_4/c_4 \\
\end{bmatrix},
\end{equation}

(9)

after we clear denominators and eliminate a factor of $c_4$. We now observe that for fixed $x$ and $u$, the three quadratic equations in $c$ from the minors of (8) define a twisted cubic curve. Indeed, as shown for example in [1, p.14], if $L_1, L_2, L_3, M_1, M_2, M_3$ are linear forms in $c_1, c_2, c_3, c_4$, then the projective set defined by

$$
\text{rk } \begin{bmatrix}
  L_1 \\
  M_1 \\
  L_2 \\
  M_2 \\
  L_3 \\
  M_3 \\
\end{bmatrix} = 1
\end{equation}

(10)

is a twisted cubic if and only if for any $(\lambda, \mu)$ the three linear forms $\lambda L_i + \mu M_i$ ($i = 1, 2, 3$) are independent. It follows that (8) is a twisted cubic in $x$ if and only if

$$
\begin{bmatrix}
  -\lambda x_4 + \mu u_1 & 0 & 0 & \lambda x_1 \\
  0 & -\lambda x_4 + \mu u_2 & 0 & \lambda x_2 \\
  0 & 0 & -\lambda x_4 + \mu u_3 & \lambda x_3 \\
\end{bmatrix}
\end{equation}

(11)

has rank three for all $\lambda, \mu$. This is true if the coordinates $u_1, u_2, u_3$ are all distinct and not zero, and the coordinates $x_1, x_2, x_3$ are not zero, which is indeed the case under our genericity assumptions.

1.1. Reduced 2D trifocal tensors and trilinearities

We begin by presenting some properties of the SFM problem for projections from $\mathbb{P}^2$ to $\mathbb{P}^1$. We represent an analytical projection of this type using a $2 \times 3$ real matrix $P$ defined up to scale. As for traditional cameras, the center of this projection is the point of $\mathbb{P}^2$ associated with the null-space of $P$. Note that if $P, P'$ are two projections with distinct centers, then every pair of points $(t, t')$ in $\mathbb{P}^1 \times \mathbb{P}^1$ will be a “correspondence” for $P, P'$, i.e., there will always exist a point $x$ in $\mathbb{P}^2$ such that $P x \sim t$ and $P' x \sim t'$. This follows from the fact that two lines in $\mathbb{P}^2$ always intersect. Thus, the first interesting case of multi-view geometry is for $n = 3$.

**Proposition 1.3.** Given three projections $P, P', P''$ with disjoint centers, there exists a $2 \times 2 \times 2$ “trifocal tensor” $T$ such that $(t, t', t'')$ in $\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1$ correspond if an only if

$$
T^{ijk} t_i t_j t_k = 0,
\end{equation}

(12)

where we use Einstein notation for summation. Entries for $T$ are given by

$$
T^{ijk} = (-1)^{i+j} |P_{3-i} P'_{3-j} P''_{3-k}|, \quad i, j, k \in \{1, 2\},
\end{equation}

(13)

where $|P_i P' P''|$ denotes the determinant of the $3 \times 3$ matrix obtained by stacking the $i$-th row of $P$, the $j$-th row of $P'$, and the $k$-th row of $P''$. Finally, the trifocal tensor $T$ satisfies the following properties:

1. Any general $2 \times 2 \times 2$ tensor is a “valid” trifocal tensor.
2. Given a general trifocal tensor $T$, there are two (possibly complex) projectively distinct sets of parameters for $P, P', P''$.

**Proof.** These facts are shown in [2], but we give a short proof here for completeness. We first note that for any $2 \times 3$ projection matrix $P$ there is an associated $3 \times 2$ “inverse projection” matrix $Q$ that maps points in $\mathbb{P}^1$ to the corresponding viewing lines in $\mathbb{P}^2$ for $P$ (expressed using dual coordinates). The relation between $P$ and $Q$ is simply

$$
Q = P^T \begin{bmatrix}
  0 & -1 \\
  1 & 0 \\
\end{bmatrix}.
\end{equation}

(14)

Since three lines in $\mathbb{P}^2$ written in dual coordinates as $l, l', l''$ converge if and only if $|l, l', l''| = 0$, we see that $(t, t', t'')$ in $\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1$ form a correspondence for $P, P', P''$ if and only if

$$
|Q^T t Q^T t' Q^T t''| = 0,
\end{equation}

(15)

where $Q, Q', Q''$ denote the inverse projection matrices for $P, P', P''$. The expansion of this determinant, together with (14), immediately yields (12) and (13).

The properties (1) and (2) can be shown computationally. Alternatively, one can explicitly a describe method for reconstructing two projectively distinct projection mappings from a general tensor. We do this below in the case of “reduced” tensors. The fact that every tensor is a valid trifocal tensor can also be argued informally by noting that each projection has 5 degrees of freedom, so after removing projective ambiguity (with 8 parameters) we are left with...
The matrix on the left has rank 5, with a null-space generated by \((1, 1, 1, 1, 1)^T\). This means that we may write \(\rho_{ij} = e_{ij} + t\) where \((e_{12}, e_{13}, e_{21}, e_{23}, e_{31}, e_{32})\) is any vector that satisfies (18) and \(t\) is unknown. However, we may solve for \(t\) using the fact that \(\rho_{ij}\) must satisfy

\[
\rho_{12}\rho_{23}\rho_{31} - \rho_{13}\rho_{21}\rho_{32} = 0.
\]  
(19)

This yields a constraint on \(t\), which is actually quadratic rather than cubic, since the cubic term in \(t\) cancels out from the two summands. Given a valid set of \(\rho_{ij}\), it is straightforward to recover \(\hat{e}_1\) and \(\hat{e}_2\).

Indeed, it is sufficient for example to set

\[
(\hat{e}_1, \hat{e}_2, \hat{e}_3) = (\rho_{12}/\rho_{32}, \rho_{21}/\rho_{31}, 1),
\]

\[
(\hat{e}_1', \hat{e}_2', \hat{e}_3') = (\rho_{21}/\rho_{32}, \rho_{31}/\rho_{21}, 1).
\]  
(20)

Note that this procedure is very similar to the method used in the proof of Proposition 4.3 in the paper. Indeed, the trilinearities \(T_1, T_2, T_3, T_4\) are closely related to the 2D reduced trifocal tensor, and we will now spell out this relation in detail. We first recall the following fact from [3] (see also Proposition 1.1 in the paper).

**Proposition 1.4.** Three image points \(u, u', u''\) satisfy the trilinearity \(T_i\) if and only if the three associated viewing lines \(\xi, \xi', \xi''\) admit a common transversal through the basis point \(z_i\).

The geometric condition from Proposition 1.4 can be expressed by considering an arbitrary projection \(P_{z_i}\) with center \(z_i\), and imposing that the projections \(l, l', l''\) of \(\xi, \xi', \xi''\) under \(P_{z_i}\) are lines in \(\mathbb{P}^2\) that converge at a point. This is why 2D trifocal tensors come into play.

In what follows, we write \(c_i', c_i''\) for the points in \(\mathbb{P}^2\) obtained from \(c_i', c_i''\) by excluding the \(r\)-th coordinate (e.g., \(c_2 = (c_1', c_3', c_4')^T\)).

**Proposition 1.5.** The trilinearity \(T_r (r = 1, 2, 3, 4)\) applied to \(u, u', u''\) can be written as

\[
T_r^{ijk}(M_r u)_j(M_r u')_j(M_r u'')_k = 0,
\]  
(21)

where \(T_r^{ijk}\) is the reduced 2D trifocal tensor associated with \((1, 1, 1)^T, c_i', c_i''\) (see (17) for the coefficients) and

\[
M_1 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \quad M_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad M_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.
\]  
(22)

**Proof.** This property can be verified computationally with a computer algebra system using the expressions for \(T_r\) (given in Proposition 8 of the paper), and for the reduced trifocal tensor (17). A geometric justification based on Proposition 1.4 is as follows.

---

1To make this argument more precise, one needs to observe that projective transformations act freely on triples of cameras (i.e., no projective transformation of \(\mathbb{P}^2\) fixes \(P, P', P''\) simultaneously).
Let us focus on $T_1$, for notational simplicity. The argument is identical for the other trilinearies. We consider the following simple projection with pinhole $z_1$

$$
P_{z_1} = \begin{bmatrix}
0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0
\end{bmatrix}.
$$ (23)

Any triple of image points $u, u', u''$ (in distinct images) determines three lines $l, l', l''$ in the image plane of $P_{z_1}$, which contain the points $(1, 1, l')^T$, $c_1 = (c_1', c_1', c_1')^T$, $c_1'' = (c_1'', c_1'', c_1'')^T$. These three points are in fact the images of $c, c', c''$ under $P_{z_1}$ (i.e., they are three epipoles). We can parameterize the three bundles through $c_1, c_1', c_1''$ using the reference frame induced by the three points $P_{z_1}(z_2), P_{z_1}(z_2), P_{z_1}(z_2)$. This way, converging triplets $l, l', l''$ are described by a reduced 2D trifocal tensor.

Finally, the coordinates in $\mathbb{P}^1$ of the lines associated with $u, u', u''$ are given simply by $M_1u, M_1u', M_1u''$. This is because $M_1$ is the unique projection matrix that maps $(0, 0, 1)^T, (0, 0, 1)^T, (0, 0, 1)^T$ (which are images of $z_2, z_3, z_4$ for all three cameras) to the standard basis of $\mathbb{P}^1$.

In conclusion, a triplet $u, u', u''$ satisfies $T_1$ if and only if the associated lines $l, l', l''$ in the image plane of $P_{z_1}$ converge, and this in turn is equivalent to the fact that $M_1u, M_1u', M_1u''$ are three points in $\mathbb{P}^1$ which satisfy the constraint from the reduced 2D trifocal tensor.

Expanding (21) for $r = 1, 2, 3$ we obtain the expressions for $T_1, T_2, T_3$ given in equation (10) of the main part of the paper. For completeness, we include here the explicit expanded form of all the trilinearies:

$$
T_1 = u_1u_1u_1''(-c_1'c_2'' + c_1''c_1') + u_2u_2u_2''(c_2'c_3'' - c_2''c_3') + u_3u_3u_3''(-c_3'c_2'' + c_3''c_1') + u_1u_2u_3''(-c_1'c_2'' + c_1''c_1') + u_2u_3u_1''(-c_2'c_3'' + c_2''c_1') + u_3u_1u_2''(-c_3'c_2'' + c_3''c_1') + u_1u_3u_2''(-c_1'c_2'' + c_1''c_1') + u_2u_1u_3''(-c_2'c_3'' + c_2''c_1') + u_3u_2u_1''(-c_3'c_2'' + c_3''c_1'),
$$

$$
T_2 = u_1u_1u_1''(-c_1'c_2'' + c_1''c_1') + u_2u_2u_2''(c_2'c_3'' - c_2''c_3') + u_3u_3u_3''(-c_3'c_2'' + c_3''c_1') + u_1u_2u_3''(-c_1'c_2'' + c_1''c_1') + u_2u_3u_1''(-c_2'c_3'' + c_2''c_1') + u_3u_1u_2''(-c_3'c_2'' + c_3''c_1') + u_1u_3u_2''(-c_1'c_2'' + c_1''c_1') + u_2u_1u_3''(-c_2'c_3'' + c_2''c_1') + u_3u_2u_1''(-c_3'c_2'' + c_3''c_1'),
$$

$$
T_3 = u_1u_1u_1''(-c_1'c_2'' + c_1''c_1') + u_2u_2u_2''(c_2'c_3'' - c_2''c_3') + u_3u_3u_3''(-c_3'c_2'' + c_3''c_1') + u_1u_2u_3''(-c_1'c_2'' + c_1''c_1') + u_2u_3u_1''(-c_2'c_3'' + c_2''c_1') + u_3u_1u_2''(-c_3'c_2'' + c_3''c_1') + u_1u_3u_2''(-c_1'c_2'' + c_1''c_1') + u_2u_1u_3''(-c_2'c_3'' + c_2''c_1') + u_3u_2u_1''(-c_3'c_2'' + c_3''c_1'),
$$

$$
T_4 = u_1u_1u_1''(-c_1'c_2'' + c_1''c_1') + u_2u_2u_2''(c_2'c_3'' - c_2''c_3') + u_3u_3u_3''(-c_3'c_2'' + c_3''c_1') + u_1u_2u_3''(-c_1'c_2'' + c_1''c_1') + u_2u_3u_1''(-c_2'c_3'' + c_2''c_1') + u_3u_1u_2''(-c_3'c_2'' + c_3''c_1') + u_1u_3u_2''(-c_1'c_2'' + c_1''c_1') + u_2u_1u_3''(-c_2'c_3'' + c_2''c_1') + u_3u_2u_1''(-c_3'c_2'' + c_3''c_1'),
$$

We can also use Proposition 1.5 to prove theoretical properties of all four trilinearies. The following statement is equivalent to Proposition 4.3 in the main part of the paper.

**Proposition 1.6.** The internal constraints of each trilinearity $T_r$ are linear. More precisely, the coefficients in $\mathbb{R}^{27}$ that are entries of $T_r$ for some choice of $c, c'$ form a linear space, of dimension five. Moreover, the coefficients of $T_r$ characterize $c_r'$ and $c_r''$ up to a two-fold ambiguity.

**Proof.** It follows from (21) that the coefficients of $T_r$ are images of the coefficients of $T_{ijk}$ under an injective linear map (note that this map is completely described by the four matrices (22)). Since the entries of a reduced 2D trifocal are only constrained by three linear conditions, the coefficients of $T_r$ form linear spaces of dimension five. Moreover, the knowledge of the $T_r$ is equivalent to that of $T_{ijk}$, which means that $c_r'$ and $c_r''$ are determined up to a two-fold ambiguity.

Finally, we provide some details on the proof of Proposition 4.2 in the paper.

**Proposition 4.2** A vector $d = (d_{ij})$ in $\mathbb{R}^{12}$ with no zero entries can be written as $d_{ij} = a_i b_j$ for some vectors $a = (a_1, a_2, a_3, a_4)^T, b = (b_1, b_2, b_3, b_4)^T$ in $\mathbb{R}^3$ if and only if $d_1, d_{kl} = d_{il} d_{kj}$ holds for all permutations $(i, j, k, l)$ of $(1, 2, 3, 4)$.

**Proof.** A factorization of $d = (d_{ij})$ exists if and only if there is a rank-1 completion of

$$
\begin{bmatrix}
* & d_{12} & d_{13} & d_{14} \\
& d_{21} & * & d_{24} \\
& d_{31} & d_{32} & * \\
& d_{41} & d_{42} & d_{43}
\end{bmatrix}.
$$ (25)

Indeed, a completion corresponds to $a^T b \in \mathbb{R}^{4 \times 4}$. It is clear that the constraints $d_{ij} d_{kl} = d_{il} d_{kj}$, which correspond
to certain $2 \times 2$ minors, are necessary. Conversely, it is always possible to deduce the diagonal elements, for example:

\[
d_{11} = \frac{d_{13}d_{21}}{d_{23}} = \frac{d_{14}d_{21}}{d_{24}} = \frac{d_{12}d_{31}}{d_{32}} = \frac{d_{14}d_{31}}{d_{34}}.
\] (26)

The constraints $d_{ij}d_{kl} = d_{il}d_{kj}$ guarantee equality among all these expressions. Indeed, we have that for example $d_{12}d_{23}d_{31} = d_{21}d_{32}d_{13}$ because

\[
\frac{d_{12}d_{23}d_{31}}{d_{13}d_{21}d_{32}} = \frac{d_{42}d_{43}d_{31}}{d_{43}d_{41}d_{32}} = \frac{d_{42}d_{41}}{d_{43}} = 1.
\] (27)

This shows that if the quadratic constraints hold, we can solve for the diagonal elements in (25), and the resulting matrix will have rank one.

2. Synthetic data experiment

For completeness, we show in Figure 1 the mean value of the mean reprojection and reconstruction errors for the synthetic data used in Section 5.2 of our submission. Recall that the quality of the reconstruction was evaluated in that section by measuring how well it predicts the reprojection of the remaining points in the dataset as well as their 3D reconstruction, once again registered to the ground truth through a homography. We have used the same data as in Section 5.2 to construct the curves shown in Figure 1. They show the mean values of the mean reprojection and reconstruction errors, given respectively in pixel and mm, for 40 random choices of the 7 point correspondences and different values of the standard deviation $\sigma$ (in pixel units) of Gaussian noise added to the image coordinates. As noted in our submission, Figure 1 shows that both the linear trifocal tensor estimation are occasionally thrown completely off course for “bad” choices of the 7 correspondences, without a clear winner in this case.

References


Figure 1. Experiments with synthetic data using the mean of the mean reprojection and reconstruction errors instead of their median. Compare to Figure 8 in our submission.