

In Defense of Relative Multi-View Geometry

Supplementary Material

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

¹ Inria

² École Normale Supérieure, CNRS, PSL Research University

1 Additional experimental comparisons

We report in this section some additional comparative experiments using the careful implementation of linear trifocal tensor estimation from https://github.com/LauraFJulia/TFT_vs_Fund by Julià and Monasse [1]. Before presenting the results, we want to make it clear that we are using *exactly* the same version of our code here as in the main submission, without *any* improvement or new tricks. Let us start by briefly describing the method implemented in [1]. When the projection matrices are $P_1 = [\text{Id } 0]$, P_2 , and P_3 , the trifocal tensors are expected to be

$$T_j = \mathbf{a}_j \mathbf{b}_4^T - \mathbf{a}_4 \mathbf{b}_j^T \quad (1)$$

where \mathbf{a}_j and \mathbf{b}_j respectively denote the j th column vectors of P_2 and P_3 . Note that \mathbf{a}_4 and \mathbf{b}_4 can respectively be viewed as the epipoles \mathbf{e}_{21} and \mathbf{e}_{31} . The implementation of [1] solves for the projection matrices and a valid trifocal tensor in four steps: (1) precondition the data [2], (2) compute the epipoles \mathbf{e}_{21} and \mathbf{e}_{31} respectively as the intersection of the three lines formed by the left and right null-vectors of T_1 , T_2 and T_3 ; (3) use Eq. (1) to compute the vectors \mathbf{a}_j and \mathbf{b}_j ($j = 1, 2, 3$) and thus complete the reconstruction of the projection matrices using linear least squares; (4) output the valid trifocal tensors defined by (1). See [1] for more details. The results of this method (dubbed 3LTFT) using three views of all 36 points in our dataset have been added to the table of our main submission, together with the corresponding mean reprojection error (Table 1). Except for 3LTFT and the primal and dual versions of our algorithm, all numbers are taken from [3]. The linear trifocal tensor estimation method gives similar performance to two-view projective structure from motion using fundamental matrices. Again, only bundle adjustment has a real edge on our method, but it exploits all six views from the dataset.

We also report below a comparison of our primal algorithm with the linear trifocal estimation method of [1] using their code and their synthetic data. Briefly, one hundred points chosen randomly in a cube centered at the origin with edges 400mm long are projected into three images of size 1800×1200 , corresponding to 35mm cameras with 50mm focal length pointing to the origin. Gaussian noise with zero mean and standard deviation σ between 0 and 2 pixels is added to the image point positions (see [1] for details). Figure 1 compares the results of our primal method using a small number (50) of trials to find good reference points

	Primal	Dual	6ASFM	2PSFM	3LTFT	6PBA
Reproj.	0.8	1.5	2.4	0.8	0.8	0.8
Recons.	0.6%	0.6%	2.8%	1.2%	1.3%	0.2%

Table 1. Quantitative comparison of several methods for projective reconstruction from multiple views on the house dataset: *primal* and *dual* versions of our algorithm and classical approaches to SFM, all using the data preconditioning of [2], namely 6-view affine factorization (*6ASFM*) [4], 2-view projective SFM (*2PSFM*), 3-view linear trifocal tensor estimation (*3LTFT*), and 6-view projective bundle adjustment (*6PBA*). **Top** mean reprojection error (in pixels). **Bottom:** mean relative reconstruction error after projective registration with ground truth,

with the linear method of [1], and demonstrates that is it more robust to noise in this setting.

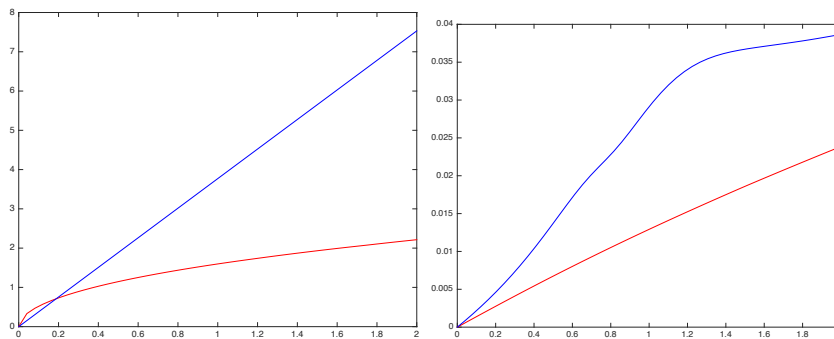


Fig. 1. Comparison of our primal method (in red) with the linear trifocal estimation algorithm of [1] (in blue). The plots show the mean reprojection (**left**) and relative reconstruction (**right**) errors as function of the standard deviation σ of noise added to the image coordinates.

The synthetic data from [1] is not appropriate (as it stands now) for testing the dual version of our algorithm since it only consists of three views of the same scene. If our paper is accepted, we will of course report a much more thorough evaluation of our approach, including its dual version as well as other synthetic and real datasets.

2 Proofs

We next present some technical proofs that were not included in the main part of the paper for lack of space.

2.1 Analytical cameras and duality

The analytical expressions for reduced projections given in the paper in equation (4) can be deduced using the following general result.

Proposition 1. *The unique projective transformation of \mathbb{P}^n that maps $n + 2$ general points $\mathbf{x}_1, \dots, \mathbf{x}_{n+2}$ to the standard basis of \mathbb{P}^n can be described by*

$$\mathbf{y} \mapsto \begin{bmatrix} \frac{|\mathbf{x}_2 \dots \mathbf{x}_{n+1} \mathbf{y}|}{|\mathbf{x}_2 \dots \mathbf{x}_{n+1} \mathbf{x}_{n+2}|} \\ \frac{|\mathbf{x}_1 \mathbf{x}_3 \dots \mathbf{x}_{n+1} \mathbf{y}|}{|\mathbf{x}_1 \mathbf{x}_3 \dots \mathbf{x}_{n+1} \mathbf{x}_{n+2}|} \\ \vdots \\ \frac{|\mathbf{x}_1 \dots \mathbf{x}_n \mathbf{y}|}{|\mathbf{x}_1 \dots \mathbf{x}_n \mathbf{x}_{n+2}|} \end{bmatrix}. \quad (2)$$

Proof. It is clear that (2) describes a projective transformation, since all the expressions are linear in the coordinates of \mathbf{y} . It also follows from elementary properties of determinants that (2) maps $\mathbf{x}_1, \dots, \mathbf{x}_{n+2}$ to the standard basis. \square

The previous result justifies the expression for $A_{\mathbf{c}} \mathbf{x}$ from in Proposition 4 in the paper. We now turn to Cremona transformations.

Proposition 2. *The rational map of \mathbb{P}^3 given by*

$$\hat{\mathbf{y}} = \mathbf{Z} \begin{bmatrix} \frac{1}{|\mathbf{y} \mathbf{x}_2 \mathbf{x}_3 \mathbf{x}_4|} \\ \frac{1}{|\mathbf{z}_1 \mathbf{y} \mathbf{z}_3 \mathbf{z}_4|} \\ \frac{1}{|\mathbf{z}_1 \mathbf{z}_2 \mathbf{y} \mathbf{z}_4|} \\ \frac{1}{|\mathbf{z}_1 \mathbf{z}_2 \mathbf{z}_3 \mathbf{y}|} \end{bmatrix}, \quad (3)$$

where \mathbf{Z} has columns $\mathbf{z}_1, \dots, \mathbf{z}_4$, is a Cremona transformation relative to $\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3, \mathbf{z}_4$ (in the sense of Proposition 2 from the main part of the paper).

Proof. We need to show that $\mathbf{z}_1, \dots, \mathbf{z}_4, \mathbf{y}_1, \mathbf{y}_2$ are in the same projective configuration as $\mathbf{z}_1, \dots, \mathbf{z}_4, \hat{\mathbf{y}}_2, \hat{\mathbf{y}}_1$. Indeed, a projective transformation relating the two sets of points is given by

$$\mathbf{x} \mapsto \mathbf{Z} \begin{bmatrix} \frac{|\mathbf{x} \mathbf{z}_2 \mathbf{z}_3 \mathbf{z}_4|}{|\mathbf{y}_1 \mathbf{z}_2 \mathbf{z}_3 \mathbf{z}_4| |\mathbf{y}_2 \mathbf{z}_2 \mathbf{z}_3 \mathbf{x}_4|} \\ \frac{|\mathbf{z}_1 \mathbf{x} \mathbf{z}_3 \mathbf{z}_4|}{|\mathbf{z}_1 \mathbf{y}_1 \mathbf{z}_3 \mathbf{z}_4| |\mathbf{z}_1 \mathbf{y}_2 \mathbf{z}_3 \mathbf{x}_4|} \\ \frac{|\mathbf{z}_1 \mathbf{z}_2 \mathbf{x} \mathbf{z}_4|}{|\mathbf{z}_1 \mathbf{z}_2 \mathbf{y}_1 \mathbf{z}_4| |\mathbf{z}_1 \mathbf{z}_2 \mathbf{y}_2 \mathbf{x}_4|} \\ \frac{|\mathbf{z}_1 \mathbf{z}_2 \mathbf{z}_3 \mathbf{x}|}{|\mathbf{z}_1 \mathbf{z}_2 \mathbf{z}_3 \mathbf{y}_1| |\mathbf{z}_1 \mathbf{z}_2 \mathbf{z}_3 \mathbf{y}_2|} \end{bmatrix}, \quad (4)$$

where $\mathbf{y}_1, \mathbf{y}_2$ are considered fixed. Note that when $\mathbf{z}_1, \dots, \mathbf{z}_4$ are basis points then (3) yields the standard Cremona transformation

$$\hat{\mathbf{y}} = \left(\frac{1}{y_1}, \frac{1}{y_2}, \frac{1}{y_3}, \frac{1}{y_4} \right)^T, \quad (5)$$

and (4) becomes

$$\mathbf{x} \mapsto \left(\frac{x_1}{y_{11}y_{21}}, \frac{x_2}{y_{12}y_{22}}, \frac{x_3}{y_{13}y_{23}}, \frac{x_4}{y_{14}y_{24}} \right)^T, \quad (6)$$

where $\mathbf{x} = (x_1, x_2, x_3, x_4)^T$, $\mathbf{y}_1 = (y_{11}, y_{12}, y_{13}, y_{14})^T$ and $\mathbf{y}_2 = (y_{21}, y_{22}, y_{23}, y_{24})^T$. \square

We next prove Proposition 6 from the main part of the paper. We recall that the set S in $\mathbb{P}^3 \times \mathbb{P}^3 \times \mathbb{P}^2$ was defined as the set of triples $(\mathbf{c}, \mathbf{x}, \mathbf{u})$ such that $P_{\mathbf{c}}(\mathbf{x}) = \mathbf{u}$ where $P_{\mathbf{c}}$ is a reduced camera.

Proposition 6. *The following properties of the set S hold:*

(1) *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 with Plücker coordinates*

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

(2) *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} .*

Proof. To show (1), we define $\boldsymbol{\xi}_j = \mathbf{c} \vee \mathbf{z}_j$ ($j = 1, \dots, 4$). With our choice of coordinate system, we have

$$\boldsymbol{\xi}_1 = \begin{bmatrix} c_4 \\ 0 \\ 0 \\ 0 \\ c_3 \\ -c_2 \end{bmatrix}, \boldsymbol{\xi}_2 = \begin{bmatrix} 0 \\ c_4 \\ 0 \\ -c_3 \\ 0 \\ c_1 \end{bmatrix}, \boldsymbol{\xi}_3 = \begin{bmatrix} 0 \\ 0 \\ c_4 \\ c_2 \\ -c_1 \\ 0 \end{bmatrix}, \boldsymbol{\xi}_4 = - \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \quad (8)$$

We write $\boldsymbol{\xi} = \rho_1 u_1 \boldsymbol{\xi}_1 + \rho_2 u_2 \boldsymbol{\xi}_2 + \rho_3 u_3 \boldsymbol{\xi}_3$, where the scalars ρ_1, ρ_2, ρ_3 have been chosen so that $\boldsymbol{\xi}_4 = \rho_1 \boldsymbol{\xi}_1 + \rho_2 \boldsymbol{\xi}_2 + \rho_3 \boldsymbol{\xi}_3$. A simple calculation shows that $\rho_1 = -c_1/c_4$, $\rho_2 = -c_2/c_4$, $\rho_3 = -c_3/c_4$, and (7) immediately follows.

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

$$\text{rk} \begin{bmatrix} c_1x_4 - c_4x_1 & u_1c_1 \\ c_2x_4 - c_4x_2 & u_2c_2 \\ c_3x_4 - c_4x_3 & u_3c_3 \end{bmatrix} = 1. \quad (9)$$

This expression follows from

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \sim P_{\mathbf{c}}(\mathbf{x}) = \begin{bmatrix} x_1/c_1 - x_4/c_4 \\ x_2/c_2 - x_4/c_4 \\ x_3/c_3 - x_4/c_4 \end{bmatrix}, \quad (10)$$

after we clear denominators and eliminate a factor of c_4 . We now observe that for fixed \mathbf{x} and \mathbf{u} , the three quadratic equations in \mathbf{c} from the minors of (9) define a twisted cubic curve. Indeed, as shown for example in [5, 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

$$\text{rk} \begin{bmatrix} L_1 & L_2 & L_3 \\ M_1 & M_2 & M_3 \end{bmatrix} = 1, \quad (11)$$

is a twisted cubic if and only if for any (λ, μ) the three linear forms $\lambda L_i + \mu M_i$ ($i = 1, 2, 3$) are independent. It follows that (9) is a twisted cubic in \mathbf{x} if and only if

$$\begin{bmatrix} -\lambda x_4 + \mu u_1 & 0 & 0 & \lambda x_1 \\ 0 & -\lambda x_4 + \mu u_2 & 0 & \lambda x_2 \\ 0 & 0 & -\lambda x_4 + \mu u_3 & \lambda x_3 \end{bmatrix} \quad (12)$$

has rank three for all λ, μ . 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 true under our genericity assumptions. \square

2.2 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 2×3 real matrix \mathbf{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 \mathbf{P} . Note that if \mathbf{P}, \mathbf{P}' are two projections with distinct centers, then every pair of points $(\mathbf{t}, \mathbf{t}')$ in $\mathbb{P}^1 \times \mathbb{P}^1$ will be a ‘‘correspondence’’ for \mathbf{P}, \mathbf{P}' , *i.e.*, there will always exist a point \mathbf{x} in \mathbb{P}^2 such that $\mathbf{P}\mathbf{x} \sim \mathbf{t}$ and $\mathbf{P}'\mathbf{x} \sim \mathbf{t}'$. 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 3. *Given three projections $\mathbf{P}, \mathbf{P}', \mathbf{P}''$ with disjoint centers, there exists a $2 \times 2 \times 2$ ‘‘trifocal tensor’’ \mathcal{T} such that $(\mathbf{t}, \mathbf{t}', \mathbf{t}'')$ in $\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1$ correspond if and only if*

$$\mathcal{T}^{ijk} t_i t'_j t''_k = 0, \quad (13)$$

where we use Einstein notation for summation. Entries for \mathcal{T} are given by

$$\mathcal{T}^{ijk} = (-1)^{ijk} |P_{3-i} P'_{3-j} P''_{3-k}|, \quad i, j, k \in \{1, 2\}, \quad (14)$$

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

1. Any generic $2 \times 2 \times 2$ tensor is a valid trifocal tensor.
2. Given a trifocal tensor \mathcal{T} , there are two projectively distinct sets of parameters for $\mathbf{P}, \mathbf{P}', \mathbf{P}''$.

Proof. These facts are shown in [6], but we give a short proof here for completeness. We first note that for any 2×3 projection matrix \mathbf{P} there is an associated 3×2 “inverse projection” matrix \mathbf{Q} that maps points in \mathbb{P}^1 to the corresponding viewing lines in \mathbb{P}^2 for \mathbf{P} (expressed using dual coordinates). The relation between \mathbf{P} and \mathbf{Q} is simply

$$\mathbf{Q} = \mathbf{P}^T \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}. \quad (15)$$

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

$$|\mathbf{Q}\mathbf{t} \ \mathbf{Q}'\mathbf{t}' \ \mathbf{Q}''\mathbf{t}''| = 0, \quad (16)$$

where $\mathbf{Q}, \mathbf{Q}', \mathbf{Q}''$ denote the inverse projection matrices for $\mathbf{P}, \mathbf{P}', \mathbf{P}''$. The expansion of this determinant, together with (15), immediately yields (13) and (14).

The properties (1) and (2) can be shown computationally. Alternatively, one can explicitly describe a 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 $5 + 5 + 5 - 8 = 7$ which correspond to all $2 \times 2 \times 2$ tensors up to scale.³ \square

We now consider “reduced” projection mappings from \mathbb{P}^2 to \mathbb{P}^1 . Similarly to the 3D case discussed in the paper, a reduced projection is determined by a center together with 3 fixed points in \mathbb{P}^2 in general position. In the following, we will always assume that the three fixed points are basis points for a projective reference frame. This leads to projection matrices of the form

$$\mathbf{P}\mathbf{c} = \begin{bmatrix} 1/c_1 & 0 & -1/c_3 \\ 0 & 1/c_2 & -1/c_3 \end{bmatrix} = \begin{bmatrix} \hat{c}_1 & 0 & -\hat{c}_3 \\ 0 & \hat{c}_2 & -\hat{c}_3 \end{bmatrix}, \quad (17)$$

where $\mathbf{c} = (c_1, c_2, c_3)^T$ is the center of projection, and we write $\hat{c}_i = 1/c_i$ for convenience. Specializing (14) for three reduced cameras $\mathbf{P}\mathbf{c}, \mathbf{P}\mathbf{c}', \mathbf{P}\mathbf{c}''$ with $\mathbf{c} = (1, 1, 1)^T$, $\mathbf{c}' = (c'_1, c'_2, c'_3)^T$, $\mathbf{c}'' = (c''_1, c''_2, c''_3)^T$, yields

$$\begin{aligned} \mathcal{T}^{111} &= 0, & \mathcal{T}^{112} &= \hat{c}'_3 \hat{c}''_1 - \hat{c}'_2 \hat{c}''_1 \\ \mathcal{T}^{121} &= \hat{c}'_1 \hat{c}''_2 - \hat{c}'_1 \hat{c}''_3, & \mathcal{T}^{122} &= \hat{c}'_1 \hat{c}''_3 - \hat{c}'_3 \hat{c}''_1, \\ \mathcal{T}^{211} &= \hat{c}'_2 \hat{c}''_3 - \hat{c}'_3 \hat{c}''_2, & \mathcal{T}^{212} &= \hat{c}'_2 \hat{c}''_1 - \hat{c}'_2 \hat{c}''_3, \\ \mathcal{T}^{221} &= \hat{c}'_3 \hat{c}''_2 - \hat{c}'_1 \hat{c}''_2, & \mathcal{T}^{222} &= 0. \end{aligned} \quad (18)$$

In addition to $\mathcal{T}^{111} = 0$ and $\mathcal{T}^{222} = 0$, we note that the expressions for the remaining six coefficients always sum to zero. Indeed, there are three “synthetic”

³ To 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 $\mathbf{P}, \mathbf{P}', \mathbf{P}''$ simultaneously).

constraints arising from the fact that the standard basis points in each image provide by construction three correspondences. However, these three are the *only* constraints that such a tensor must satisfy. In fact, we can give the following simple algorithm for 2D reduced SFM.

1. Change coordinates in each image \mathbb{P}^1 using three triples of correspondences to restrict to reduced projection matrices as in (17).
2. Estimate the six non-zero coefficients a, b, c, d, e, f of the reduced trifocal tensor (18), under the condition that they sum to zero, using an arbitrary number of correspondences (at least five) in the new image coordinates.
3. Use (18) to recover $\hat{c}'_1, \hat{c}'_2, \hat{c}'_3$ and $\hat{c}''_1, \hat{c}''_2, \hat{c}''_3$, and hence $\mathbf{c}', \mathbf{c}''$ from the six coefficients a, b, c, d, e, f . There are in general *two* solutions, which can be computed as follows. Writing $\rho_{ij} = \hat{c}'_i \hat{c}''_j$, we have that (18) yields the following linear relation

$$\begin{bmatrix} 0 & 0 & -1 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \rho_{12} \\ \rho_{13} \\ \rho_{21} \\ \rho_{23} \\ \rho_{31} \\ \rho_{32} \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix}. \quad (19)$$

The matrix on the left has rank 5, with a null-space generated by $(1, 1, 1, 1, 1, 1)^T$. 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 (19) and t is unknown. However, we may solve for t using the fact that ρ_{ij} must satisfy

$$\rho_{12}\rho_{23}\rho_{31} - \rho_{13}\rho_{21}\rho_{32} = 0. \quad (20)$$

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 ρ_{ij} , it is straightforward to recover \hat{c}'_i and \hat{c}''_j . Indeed, it is sufficient for example to set

$$\begin{aligned} (\hat{c}'_1, \hat{c}'_2, \hat{c}'_3) &= (\rho_{12}/\rho_{32}, \rho_{21}/\rho_{31}, 1), \\ (\hat{c}''_1, \hat{c}''_2, \hat{c}''_3) &= (\rho_{21}/\rho_{23}, \rho_{12}/\rho_{13}, 1). \end{aligned} \quad (21)$$

Note that this procedure is very similar to the “linear” SFM algorithm presented in Section 3.4 of the paper. Indeed, the trilinearities T_1, T_2, T_3, T_4 are closely related to the reduced trifocal tensor, and we will now spell out this relation in detail. We first recall the following fact from [7].

Proposition 4. *Three image points $\mathbf{u}, \mathbf{u}', \mathbf{u}''$ satisfy the trilinearity T_i if and only if the three associated viewing lines $\boldsymbol{\xi}, \boldsymbol{\xi}', \boldsymbol{\xi}''$ admit a common transversal through the basis point \mathbf{z}_i .*

As explained in the paper, the geometric condition from Proposition 4 can be expressed by considering an arbitrary projection $P_{\mathbf{z}_i}$ with center \mathbf{z}_i , and

imposing that the projections $\mathbf{l}, \mathbf{l}', \mathbf{l}''$ of $\boldsymbol{\xi}, \boldsymbol{\xi}', \boldsymbol{\xi}''$ under $P_{\mathbf{z}_i}$ are lines in \mathbb{P}^2 that converge at a point. This is why 2D trifocal tensors come into play.

In the following, we write $\mathbf{c}'_r, \mathbf{c}''_r$ for the points in \mathbb{P}^2 obtained from $\mathbf{c}', \mathbf{c}''$ by excluding the r -th coordinate (e.g., $\mathbf{c}'_2 = (c'_1, c'_3, c'_4)^T$).

Proposition 5. *The trilinearity T_r ($r = 1, 2, 3, 4$) applied to $\mathbf{u}, \mathbf{u}', \mathbf{u}''$ can be written as*

$$\mathcal{T}_r^{ijk}(\mathbf{M}_r \mathbf{u})_i (\mathbf{M}_r \mathbf{u}')_j (\mathbf{M}_r \mathbf{u}'')_k = 0, \quad (22)$$

where \mathcal{T}_r^{ijk} is the reduced 2D trifocal tensor associated with $(1, 1, 1)^T$, $\mathbf{c}'_r, \mathbf{c}''_r$ (see (18) for the coefficients) and

$$\begin{aligned} \mathbf{M}_1 &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \mathbf{M}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & -1 \end{bmatrix}, \\ \mathbf{M}_3 &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \mathbf{M}_4 = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix}. \end{aligned} \quad (23)$$

Proof. This property can be verified computationally with a computer algebra system using the expressions for T_i (given in Proposition 8 of the paper), and for the reduced trifocal tensor (18). A geometric justification based on Proposition 4 is as follows.

Let us focus on T_1 , for notational simplicity. The argument is identical for the other trilinearities. We consider the following simple projection with pinhole \mathbf{z}_1

$$P_{\mathbf{z}_1} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (24)$$

Any triple of image points $\mathbf{u}, \mathbf{u}', \mathbf{u}''$ (in distinct images) determines three lines $\mathbf{l}, \mathbf{l}', \mathbf{l}''$ in the image plane of $P_{\mathbf{z}_1}$, which contain the points $(1, 1, 1)^T$, $\mathbf{c}'_1 = (c'_2, c'_3, c'_4)^T$, $\mathbf{c}''_1 = (c''_2, c''_3, c''_4)^T$. These three points are in fact the images of $\mathbf{c}, \mathbf{c}', \mathbf{c}''$ under $P_{\mathbf{z}_1}$ (i.e., they are three epipoles). We can parameterize the three bundles through $\mathbf{c}_1, \mathbf{c}'_1, \mathbf{c}''_1$ using the reference frame induced by the three points $P_{\mathbf{z}_1}(\mathbf{z}_2), P_{\mathbf{z}_1}(\mathbf{z}_3), P_{\mathbf{z}_1}(\mathbf{z}_4)$. This way, converging triplets $\mathbf{l}, \mathbf{l}', \mathbf{l}''$ are described by a reduced 2D trifocal tensor.

Finally, the coordinates in \mathbb{P}^1 of the lines associated with $\mathbf{u}, \mathbf{u}', \mathbf{u}''$ are given simply by $\mathbf{M}_1 \mathbf{u}, \mathbf{M}_1 \mathbf{u}', \mathbf{M}_1 \mathbf{u}''$. This is because \mathbf{M}_1 is the unique projection matrix that maps $(0, 1, 0)^T, (0, 0, 1)^T, (1, 1, 1)^T$ (which are images of $\mathbf{z}_2, \mathbf{z}_3, \mathbf{z}_4$ for all three cameras) to the standard basis of \mathbb{P}^1 .

In conclusion, a triplet $\mathbf{u}, \mathbf{u}', \mathbf{u}''$ satisfies T_1 if and only if the associated lines $\mathbf{l}, \mathbf{l}', \mathbf{l}''$ in the image plane of $P_{\mathbf{z}_1}$ converge, and this in turn is equivalent to the fact that $\mathbf{M}_1 \mathbf{u}, \mathbf{M}_1 \mathbf{u}', \mathbf{M}_1 \mathbf{u}''$ are three points in \mathbb{P}^1 which satisfy the constraint from the reduced 2D trifocal tensor. \square

Expanding (22) 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 trilinearities:

$$\begin{aligned}
 T_1 &= u_2 u_2' u_3'' (-\check{c}_3' \check{c}_2'' + \check{c}_4' \check{c}_2'') + u_2 u_3' u_2'' (\check{c}_2' \check{c}_3'' - \check{c}_2' \check{c}_4'') + \\
 &\quad u_2 u_3' u_3'' (-\check{c}_4' \check{c}_2'' + \check{c}_2' \check{c}_4'') + u_3 u_2' u_2'' (-\check{c}_4' \check{c}_3'' + \check{c}_3' \check{c}_4'') + \\
 &\quad u_3 u_2' u_3'' (\check{c}_3' \check{c}_2'' - \check{c}_3' \check{c}_4'') + u_3 u_3' u_2'' (-\check{c}_2' \check{c}_3'' + \check{c}_4' \check{c}_3''), \\
 T_2 &= u_1 u_1' u_3'' (-\check{c}_3' \check{c}_1'' + \check{c}_4' \check{c}_1'') + u_1 u_3' u_1'' (\check{c}_1' \check{c}_3'' - \check{c}_1' \check{c}_4'') + \\
 &\quad u_1 u_3' u_3'' (-\check{c}_4' \check{c}_1'' + \check{c}_1' \check{c}_4'') + u_3 u_1' u_1'' (-\check{c}_4' \check{c}_3'' + \check{c}_3' \check{c}_4'') + \\
 &\quad u_3 u_1' u_3'' (\check{c}_3' \check{c}_1'' - \check{c}_3' \check{c}_4'') + u_3 u_3' u_1'' (-\check{c}_1' \check{c}_3'' + \check{c}_4' \check{c}_3''), \\
 T_3 &= u_1 u_1' u_2'' (-\check{c}_2' \check{c}_1'' + \check{c}_4' \check{c}_1'') + u_1 u_2' u_1'' (\check{c}_1' \check{c}_2'' - \check{c}_1' \check{c}_4'') + \\
 &\quad u_1 u_2' u_2'' (-\check{c}_4' \check{c}_1'' + \check{c}_1' \check{c}_4'') + u_2 u_1' u_1'' (-\check{c}_4' \check{c}_2'' + \check{c}_2' \check{c}_4'') + \\
 &\quad u_2 u_1' u_2'' (\check{c}_2' \check{c}_1'' - \check{c}_2' \check{c}_4'') + u_2 u_2' u_1'' (-\check{c}_1' \check{c}_2'' + \check{c}_4' \check{c}_2''), \\
 T_4 &= u_1 u_1' u_2'' (-\check{c}_2' \check{c}_1'' + \check{c}_3' \check{c}_1'') + u_1 u_1' u_3'' (\check{c}_2' \check{c}_1'' - \check{c}_3' \check{c}_1'') + \\
 &\quad u_1 u_2' u_1'' (\check{c}_1' \check{c}_2'' - \check{c}_1' \check{c}_3'') + u_1 u_2' u_2'' (-\check{c}_3' \check{c}_1'' + \check{c}_1' \check{c}_3'') + \\
 &\quad u_1 u_2' u_3'' (\check{c}_3' \check{c}_1'' - \check{c}_1' \check{c}_2'') + u_1 u_3' u_1'' (-\check{c}_1' \check{c}_2'' + \check{c}_1' \check{c}_3'') + \\
 &\quad u_1 u_3' u_2'' (\check{c}_2' \check{c}_1'' - \check{c}_1' \check{c}_3'') + u_1 u_3' u_3'' (-\check{c}_2' \check{c}_1'' + \check{c}_1' \check{c}_2'') + \\
 &\quad u_2 u_1' u_1'' (-\check{c}_3' \check{c}_2'' + \check{c}_2' \check{c}_3'') + u_2 u_1' u_2'' (\check{c}_2' \check{c}_1'' - \check{c}_2' \check{c}_3'') + \\
 &\quad u_2 u_1' u_3'' (-\check{c}_2' \check{c}_1'' + \check{c}_3' \check{c}_2'') + u_2 u_2' u_1'' (-\check{c}_1' \check{c}_2'' + \check{c}_3' \check{c}_2'') + \\
 &\quad u_2 u_2' u_3'' (\check{c}_1' \check{c}_2'' - \check{c}_3' \check{c}_2'') + u_2 u_3' u_1'' (\check{c}_1' \check{c}_2'' - \check{c}_2' \check{c}_3'') + \\
 &\quad u_2 u_3' u_2'' (-\check{c}_2' \check{c}_1'' + \check{c}_2' \check{c}_3'') + u_2 u_3' u_3'' (\check{c}_2' \check{c}_1'' - \check{c}_1' \check{c}_2'') + \\
 &\quad u_3 u_1' u_1'' (\check{c}_3' \check{c}_2'' - \check{c}_2' \check{c}_3'') + u_3 u_1' u_2'' (-\check{c}_3' \check{c}_1'' + \check{c}_2' \check{c}_3'') + \\
 &\quad u_3 u_1' u_3'' (\check{c}_3' \check{c}_1'' - \check{c}_3' \check{c}_2'') + u_3 u_2' u_1'' (-\check{c}_3' \check{c}_2'' + \check{c}_1' \check{c}_3'') + \\
 &\quad u_3 u_2' u_2'' (\check{c}_3' \check{c}_1'' - \check{c}_1' \check{c}_3'') + u_3 u_2' u_3'' (-\check{c}_3' \check{c}_1'' + \check{c}_3' \check{c}_2'') + \\
 &\quad u_3 u_3' u_1'' (-\check{c}_1' \check{c}_3'' + \check{c}_2' \check{c}_3'') + u_3 u_3' u_2'' (\check{c}_1' \check{c}_3'' - \check{c}_2' \check{c}_3'').
 \end{aligned} \tag{25}$$

We can also use Proposition 5 to prove theoretical properties of all four trilinearities. The following statement is equivalent to Proposition 11 in the main part of the paper.

Corollary 1. *The internal constraints of each trilinearity T_r are linear. More precisely, the coefficients in \mathbb{R}^{27} that are entries of T_i for some choice of \mathbf{c}' , \mathbf{c}'' form a linear space, of dimension five. Moreover, the coefficients of T_r characterize \mathbf{c}'_r and \mathbf{c}''_r up to a two-fold ambiguity.*

Proof. It follows from (22) that the coefficients of T_r are images of the coefficients of \mathcal{T}_r^{ijk} under an injective linear map (note that this map is completely described by the four matrices (23)). 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 \mathcal{T}_r^{ijk} , which means that \mathbf{c}'_r and \mathbf{c}''_r are determined up to a two-fold ambiguity. \square

As a final remark, we list the complete set of algebraic constraints considered in Proposition 10 of the paper.

Lemma 1. *A vector $\boldsymbol{\rho} = (\rho_{12}, \dots, \rho_{43})$ in \mathbb{R}^{12} factors as $\rho_{ij} = a_i b_j$ for some vectors $\mathbf{a} = (a_1, a_2, a_3, a_4)^T$ and $\mathbf{b} = (b_1, b_2, b_3, b_4)^T$ if and only if the following 10 polynomials vanish*

$$\begin{aligned}
& \rho_{13}\rho_{42} - \rho_{12}\rho_{43} \\
& \rho_{32}\rho_{41} - \rho_{31}\rho_{42} \\
& \rho_{23}\rho_{41} - \rho_{21}\rho_{43} \\
& \rho_{14}\rho_{32} - \rho_{12}\rho_{34} \\
& \rho_{24}\rho_{31} - \rho_{21}\rho_{34} \\
& \rho_{14}\rho_{23} - \rho_{13}\rho_{24} \\
& \rho_{23}\rho_{34}\rho_{42} - \rho_{24}\rho_{32}\rho_{43} \\
& \rho_{13}\rho_{34}\rho_{41} - \rho_{14}\rho_{31}\rho_{43} \\
& \rho_{12}\rho_{24}\rho_{41} - \rho_{14}\rho_{21}\rho_{42} \\
& \rho_{12}\rho_{23}\rho_{31} - \rho_{13}\rho_{21}\rho_{32}.
\end{aligned} \tag{26}$$

Proof. We consider the ideal I generated by all 2×2 minors of

$$\begin{pmatrix}
\rho_{11} & \rho_{12} & \rho_{13} & \rho_{14} \\
\rho_{21} & \rho_{22} & \rho_{23} & \rho_{24} \\
\rho_{31} & \rho_{32} & \rho_{33} & \rho_{34} \\
\rho_{41} & \rho_{42} & \rho_{43} & \rho_{44}
\end{pmatrix}. \tag{27}$$

The ideal I characterizes variables ρ_{ij} , with $i, j = 1, \dots, 4$ that factor as $\rho_{ij} = a_i b_j$ for some \mathbf{a}, \mathbf{b} . We then use a computer algebra system to eliminate from I the variables $\rho_{11}, \rho_{22}, \rho_{33}, \rho_{44}$ (since our vector $\boldsymbol{\rho}$ does not include them). The result is an ideal that is minimally generated by the polynomials in (26). \square

References

1. Julià, L., Monasse, P.: A critical review of the trifocal tensor estimation. In: PSIVT. (2017)
2. Hartley, R.: In defence of the 8-point algorithm. In: ICCV, Boston, MA (1995) 1064–1070
3. Forsyth, D., Ponce, J.: Computer Vision: A Modern Approach. 2nd edn. Prentice-Hall (2011)
4. Tomasi, C., Kanade, T.: Shape and motion from image streams under orthography: a factorization method. IJCV **9**(2) (1992) 137–154
5. Harris, J.: Algebraic geometry: a first course. Volume 133. Springer Science & Business Media (2013)
6. Quan, L.: Two-way ambiguity in 2d projective reconstruction from three uncalibrated 1d images. PAMI **23**(2) (2001) 212–216
7. Trager, M., Ponce, J., Hebert, M.: Trinocular geometry revisited. IJCV **120**(2) (2016) 134–152