General models for rational cameras and the case of two-slit projections

Supplementary material

Matthew Trager[1,2] Bernd Sturmfels[3] John Canny[3] Martial Hebert[4] Jean Ponce[1,2]

[1]École Normale Supérieure, CNRS, PSL Research University [2]Inria [3]UC Berkeley [4]Carnegie Mellon University

This supplementary document contains some technical material not included in the main body of the paper, and presents the algorithms for SfM and self-calibration for two-slit cameras.

1. Calculations with Plücker coordinates

Let $\boldsymbol{\pi}$ be a plane in $\mathbb{P}^3$. We consider a reference frame (π) on $\boldsymbol{\pi}$ described by a 4×3 matrix $\boldsymbol{Y}=[\boldsymbol{y}_1,\boldsymbol{y}_2,\boldsymbol{y}_3]$. The map $\boldsymbol{N}:\mathrm{Gr}(1,3)\dashrightarrow\mathbb{P}^2$ associating any line $\boldsymbol{l}$ not on $\boldsymbol{\pi}$ with the coordinates of the point $\boldsymbol{y}=\boldsymbol{\pi}\wedge\boldsymbol{l}$ for (π) is described by the 3×6 matrix

$$\boldsymbol{N}=\begin{bmatrix}(\boldsymbol{y}_2\vee\boldsymbol{y}_3)^{*T}\\(\boldsymbol{y}_3\vee\boldsymbol{y}_1)^{*T}\\(\boldsymbol{y}_1\vee\boldsymbol{y}_2)^{*T}\end{bmatrix}. \tag{1}$$

Indeed, this is the only linear map $\mathrm{Gr}(1,3)\dashrightarrow\mathbb{P}^2$ such that $\boldsymbol{N}(\boldsymbol{y}_1\vee\boldsymbol{z})=(1,0,0)^T$, $\boldsymbol{N}(\boldsymbol{y}_2\vee\boldsymbol{z})=(0,1,0)^T$, $\boldsymbol{N}(\boldsymbol{y}_3\vee\boldsymbol{z})=(0,0,1)^T$, $\boldsymbol{N}((\boldsymbol{y}_1+\boldsymbol{y}_2+\boldsymbol{y}_3)\vee\boldsymbol{z})=(1,1,1)^T$ for all $\boldsymbol{z}$ not in $\boldsymbol{\pi}$.

Let us now consider two "slits" $\boldsymbol{l}_1,\boldsymbol{l}_2$, that we represent using dual Plücker matrices $\boldsymbol{P}_1^*,\boldsymbol{P}_2^*$. The action of the corresponding essential camera $\boldsymbol{x}\mapsto\lambda_L(\boldsymbol{x})$ can be written as

$$\boldsymbol{x}\mapsto\boldsymbol{l}=(\boldsymbol{l}_1\vee\boldsymbol{x})\wedge(\boldsymbol{l}_2\vee\boldsymbol{x})=\boldsymbol{P}_1^*\boldsymbol{x}\boldsymbol{x}^T\boldsymbol{P}_2^*-\boldsymbol{P}_2^*\boldsymbol{x}\boldsymbol{x}^T\boldsymbol{P}_1^*, \tag{2}$$

where $\boldsymbol{l}=\lambda_L(\boldsymbol{x})$ is given as a dual Plücker matrix. Writing $\boldsymbol{S}_1,\boldsymbol{S}_1^*,\boldsymbol{S}_2,\boldsymbol{S}_2^*,\boldsymbol{S}_3,\boldsymbol{S}_3^*$ for the primal and dual Plücker matrices for $\boldsymbol{y}_2\vee\boldsymbol{y}_3,\boldsymbol{y}_3\vee\boldsymbol{y}_1,\boldsymbol{y}_1\vee\boldsymbol{y}_2$ respectively, and $\boldsymbol{L},\boldsymbol{L}^*$ for the primal and dual Plücker matrices of $\boldsymbol{l}=\lambda_L(\boldsymbol{x})$, we have

$$\begin{aligned}\lambda_L(\boldsymbol{x})\wedge\boldsymbol{\pi}=\boldsymbol{Y}\boldsymbol{N}\boldsymbol{l}=\boldsymbol{Y}\begin{bmatrix}\mathrm{tr}(\boldsymbol{S}_1^*\boldsymbol{L})\\\mathrm{tr}(\boldsymbol{S}_2^*\boldsymbol{L})\\\mathrm{tr}(\boldsymbol{S}_3^*\boldsymbol{L})\end{bmatrix}=\boldsymbol{Y}\begin{bmatrix}\mathrm{tr}(\boldsymbol{S}_1\boldsymbol{L}^*)\\\mathrm{tr}(\boldsymbol{S}_2\boldsymbol{L}^*)\\\mathrm{tr}(\boldsymbol{S}_3\boldsymbol{L}^*)\end{bmatrix}\\=\boldsymbol{Y}\begin{bmatrix}\mathrm{tr}(\boldsymbol{S}_1\boldsymbol{P}_2^*\boldsymbol{x}\boldsymbol{x}^T\boldsymbol{P}_1^*)\\\mathrm{tr}(\boldsymbol{S}_2\boldsymbol{P}_2^*\boldsymbol{x}\boldsymbol{x}^T\boldsymbol{P}_1^*)\\\mathrm{tr}(\boldsymbol{S}_3\boldsymbol{P}_2^*\boldsymbol{x}\boldsymbol{x}^T\boldsymbol{P}_1^*)\end{bmatrix}=\boldsymbol{Y}\begin{bmatrix}\boldsymbol{x}^T\boldsymbol{P}_1^*\boldsymbol{S}_1\boldsymbol{P}_2^*\boldsymbol{x}\\\boldsymbol{x}^T\boldsymbol{P}_1^*\boldsymbol{S}_2\boldsymbol{P}_2^*\boldsymbol{x}\\\boldsymbol{x}^T\boldsymbol{P}_1^*\boldsymbol{S}_3\boldsymbol{P}_2^*\boldsymbol{x}\end{bmatrix},\end{aligned} \tag{3}$$

where equality is written up to scale, and we have used the fact that $\mathrm{tr}(\boldsymbol{A}\boldsymbol{B})=\mathrm{tr}(\boldsymbol{B}\boldsymbol{A})=\mathrm{tr}(\boldsymbol{A}^T\boldsymbol{B}^T)=\mathrm{tr}(\boldsymbol{B}^T\boldsymbol{A}^T)$ for any matrices $\boldsymbol{A},\boldsymbol{B}$. Hence, we recover the expression for a general two-slit camera, already noted in [1, 6]:

$$\boldsymbol{x}\mapsto\boldsymbol{u}=\begin{bmatrix}\boldsymbol{x}^T\boldsymbol{P}_1^*\boldsymbol{S}_1\boldsymbol{P}_2^*\boldsymbol{x}\\\boldsymbol{x}^T\boldsymbol{P}_1^*\boldsymbol{S}_2\boldsymbol{P}_2^*\boldsymbol{x}\\\boldsymbol{x}^T\boldsymbol{P}_1^*\boldsymbol{S}_3\boldsymbol{P}_2^*\boldsymbol{x}\end{bmatrix}. \tag{4}$$

If we choose an "intrinsic" reference frame, so that $\boldsymbol{y}_1=\boldsymbol{l}_2\wedge\pi$ and $\boldsymbol{y}_2=\boldsymbol{l}_1\wedge\pi$, or equivalently $\boldsymbol{P}_1^*\boldsymbol{y}_2=\boldsymbol{P}_2^*\boldsymbol{y}_1=0$, the two-slit projection (4) reduces to

$$\begin{aligned}\boldsymbol{x}\mapsto\boldsymbol{u}&=\begin{bmatrix}\boldsymbol{x}^T\boldsymbol{P}_1^*(\boldsymbol{y}_2\boldsymbol{y}_3^T-\boldsymbol{y}_3\boldsymbol{y}_2^T)\boldsymbol{P}_2^*\boldsymbol{x}\\\boldsymbol{x}^T\boldsymbol{P}_1^*(\boldsymbol{y}_3\boldsymbol{y}_1^T-\boldsymbol{y}_1\boldsymbol{y}_3^T)\boldsymbol{P}_2^*\boldsymbol{x}\\\boldsymbol{x}^T\boldsymbol{P}_1^*(\boldsymbol{y}_1\boldsymbol{y}_2^T-\boldsymbol{y}_2\boldsymbol{y}_1^T)\boldsymbol{P}_2^*\boldsymbol{x}\end{bmatrix}\\&=\begin{bmatrix}-\boldsymbol{x}^T\boldsymbol{P}_1^*\boldsymbol{y}_3\boldsymbol{y}_2^T\boldsymbol{P}_2^*\boldsymbol{x}\\-\boldsymbol{x}^T\boldsymbol{P}_1^*\boldsymbol{y}_1\boldsymbol{y}_3^T\boldsymbol{P}_2^*\boldsymbol{x}\\\boldsymbol{x}^T\boldsymbol{P}_1^*\boldsymbol{y}_1\boldsymbol{y}_2^T\boldsymbol{P}_2^*\boldsymbol{x}\end{bmatrix}\\&=\begin{bmatrix}(\boldsymbol{p}_1^T\boldsymbol{x})\ (\boldsymbol{q}_2^T\boldsymbol{x})\\(\boldsymbol{p}_2^T\boldsymbol{x})\ (\boldsymbol{q}_1^T\boldsymbol{x})\\(\boldsymbol{p}_2^T\boldsymbol{x})\ (\boldsymbol{q}_2^T\boldsymbol{x})\end{bmatrix},\end{aligned} \tag{5}$$

where $\boldsymbol{p}_1=\boldsymbol{P}_1^*\boldsymbol{y}_3=(\boldsymbol{l}_1\vee\boldsymbol{y}_3)$, $\boldsymbol{p}_2=-\boldsymbol{P}_1^*\boldsymbol{y}_1=-(\boldsymbol{l}_1\vee\boldsymbol{y}_1)$, $\boldsymbol{q}_1=\boldsymbol{P}_2^*\boldsymbol{y}_3=(\boldsymbol{l}_2\vee\boldsymbol{y}_3)$, $\boldsymbol{q}_2=-\boldsymbol{P}_2^*\boldsymbol{y}_2=-(\boldsymbol{l}_2\vee\boldsymbol{y}_2)$. Finally, combining (2) and (5), we also obtain an expression for the inverse line projection $\chi:\mathbb{P}^2\dashrightarrow\mathrm{Gr}(1,3)$:

$$\begin{aligned}\boldsymbol{u}\mapsto\lambda_L(\boldsymbol{Y}\boldsymbol{u})&=\boldsymbol{P}_1^*\boldsymbol{Y}\boldsymbol{u}\boldsymbol{u}^T\boldsymbol{Y}^T\boldsymbol{P}_2^*-\boldsymbol{P}_2^*\boldsymbol{Y}\boldsymbol{u}\boldsymbol{u}^T\boldsymbol{Y}^T\boldsymbol{P}_1^*\\&=[-\boldsymbol{p}_2,\boldsymbol{0},\boldsymbol{p}_1]\boldsymbol{u}\boldsymbol{u}^T[\boldsymbol{0},-\boldsymbol{q}_2,\boldsymbol{q}_1]-[\boldsymbol{0},-\boldsymbol{q}_2,\boldsymbol{q}_1]\boldsymbol{u}\boldsymbol{u}^T[-\boldsymbol{p}_2,\boldsymbol{0},\boldsymbol{p}_1]\\&=u_1u_2(\boldsymbol{p}_2\wedge\boldsymbol{q}_2)-u_1u_3(\boldsymbol{p}_2\wedge\boldsymbol{q}_1)-u_2u_3(\boldsymbol{p}_1\wedge\boldsymbol{q}_2)+u_3^2(\boldsymbol{p}_1\wedge\boldsymbol{q}_1).\end{aligned} \tag{6}$$

2. Proofs

Lemma 1. *Let $\boldsymbol{l}_1,\boldsymbol{l}_2$ be two skew lines in $\mathbb{P}^3$. For any point $\boldsymbol{x}$ not on the these lines, we indicate with $\lambda(\boldsymbol{x})$ the unique transversal to $\boldsymbol{l}_1,\boldsymbol{l}_2$ passing through $\boldsymbol{x}$. If $\boldsymbol{\pi}$ and $\boldsymbol{\pi}'$ are two planes intersecting at a line $\boldsymbol{\delta}$ that meets $\boldsymbol{l}_1$ and $\boldsymbol{l}_2$, then the map $f:\boldsymbol{\pi}\dashrightarrow\boldsymbol{\pi}'$ defined, for points $\boldsymbol{y}$ not on $\boldsymbol{\delta}$, as*

$$f(\boldsymbol{y})=\lambda(\boldsymbol{y})\wedge\boldsymbol{\pi}', \tag{7}$$

can be extended to a homography between $\boldsymbol{\pi}$ and $\boldsymbol{\pi}'$.

Proof. Let us fix a coordinate system (π) on $\boldsymbol{\pi}$ given by $\boldsymbol{Y} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{y}_3]$. Up to composing with a projective transformation, we may assume that $\boldsymbol{y}_1 = \boldsymbol{l}_2 \wedge \boldsymbol{\pi}$ and $\boldsymbol{y}_2 = \boldsymbol{l}_1 \wedge \boldsymbol{\pi}$. It is also convenient to define $\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{q}_1, \boldsymbol{q}_2$ as in the previous section, namely $\boldsymbol{p}_1 = (\boldsymbol{l}_1 \vee \boldsymbol{y}_3)$, $\boldsymbol{p}_2 = -(\boldsymbol{l}_1 \vee \boldsymbol{y}_1)$, $\boldsymbol{q}_1 = (\boldsymbol{l}_2 \vee \boldsymbol{y}_3)$, $\boldsymbol{q}_2 = -(\boldsymbol{l}_2 \vee \boldsymbol{y}_2)$. The map $\lambda(\boldsymbol{y})$ can now be written as $\boldsymbol{y} = \boldsymbol{Y}\boldsymbol{u} \mapsto \chi(\boldsymbol{u})$ where χ is given in (6). In particular, since $\boldsymbol{\delta} = \boldsymbol{p}_2 \wedge \boldsymbol{q}_2 = \boldsymbol{y}_1 \vee \boldsymbol{y}_2$ lies on $\boldsymbol{\pi}'$, we can describe $f(\boldsymbol{y})$ as

$$\begin{aligned} &\boldsymbol{y} = \boldsymbol{Y}\boldsymbol{u} \mapsto \chi(\boldsymbol{u}) \wedge \boldsymbol{\pi}' \\ &= -u_1u_3(\boldsymbol{p}_2 \wedge \boldsymbol{q}_1 \wedge \boldsymbol{\pi}') - u_2u_3(\boldsymbol{p}_1 \wedge \boldsymbol{q}_2 \wedge \boldsymbol{\pi}') + u_3^2(\boldsymbol{p}_1 \wedge \boldsymbol{q}_1 \wedge \boldsymbol{\pi}') \\ &= u_1\boldsymbol{y}_1' + u_2\boldsymbol{y}_2' + u_3\boldsymbol{y}_3', \end{aligned} \tag{8}$$

where $\boldsymbol{y}_1' = -(\boldsymbol{p}_2 \wedge \boldsymbol{q}_1 \wedge \boldsymbol{\pi}')$, $\boldsymbol{y}_2' = -(\boldsymbol{p}_1 \wedge \boldsymbol{q}_2 \wedge \boldsymbol{\pi}')$, $\boldsymbol{y}_3' = (\boldsymbol{p}_1 \wedge \boldsymbol{q}_1 \wedge \boldsymbol{\pi}')$. Fixing $\boldsymbol{Y}' = [\boldsymbol{y}_1', \boldsymbol{y}_2', \boldsymbol{y}_3']$ as a reference frame on $\boldsymbol{\pi}'$, the map (7) corresponds to the identity on $\mathbb{P}^2$. Hence, it can be extended to points $\boldsymbol{y}$ on $\boldsymbol{\delta}$ (where $u_3 = 0$), and it is a homography.

We also give a sketch for a more "geometric" argument: we need to show that the a (generic) line $\boldsymbol{m}$ on $\boldsymbol{\pi}$ is mapped by (7) to a line on $\boldsymbol{\pi}'$. If $\boldsymbol{m}$ does not intersect $\boldsymbol{l}_1$ or $\boldsymbol{l}_2$, the union of the common transversals to $\boldsymbol{l}_1, \boldsymbol{l}_2, \boldsymbol{m}$ (that are the lines in $\lambda_L(\boldsymbol{m})$) is a quadric in $\mathbb{P}^3$. The intersection of this quadric with $\boldsymbol{\pi}'$ will have degree two, however it contains the transversal line $\boldsymbol{\delta}$, and hence it is reducible. Since $\boldsymbol{\delta}$ does not belong to the image of (7), we deduce that (the closure of) image of $\boldsymbol{m}$ is a line in $\boldsymbol{\pi}'$. □

Proposition 1. *If $\boldsymbol{A}_1$, $\boldsymbol{A}_2$ describe a parallel two-slit camera, then we can uniquely write*

$$\boldsymbol{A}_1 = \boldsymbol{K}_1 \begin{bmatrix} \boldsymbol{r}_1^T & t_1 \\ \boldsymbol{r}_3^T & t_3 \end{bmatrix}, \ \boldsymbol{A}_2 = \boldsymbol{K}_2 \begin{bmatrix} \boldsymbol{r}_2^T & t_2 \\ \boldsymbol{r}_3^T & t_4 \end{bmatrix}, \tag{9}$$

where $\boldsymbol{K}_1$ and $\boldsymbol{K}_2$ are upper-triangular 2×2 matrices defined up to scale with positive elements along the diagonal, and $\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3$ are unit vectors, with $\boldsymbol{r}_3$ orthogonal to both $\boldsymbol{r}_1, \boldsymbol{r}_2$. Here, $\theta = \arccos(\boldsymbol{r}_1 \cdot \boldsymbol{r}_2)$ is the angle *between the slits, and $|t_4 - t_3|$ is the* distance *between the slits. Moreover, if the matrices $\boldsymbol{K}_1$ and $\boldsymbol{K}_2$ are written as*

$$\boldsymbol{K}_1 = \begin{bmatrix} f_u & u_0 \\ 0 & 1 \end{bmatrix}, \boldsymbol{K}_2 = \begin{bmatrix} 2f_v & v_0 \\ 0 & 1 \end{bmatrix}, \tag{10}$$

then f_u, f_v can be interpreted as "magnifications" in the u and v directions, and (u_0, v_0) as the position of the "principal point".

Proof. The decomposition exists and is unique because of RQ-decomposition of matrices [2, Theorem 5.2.3]. More precisely, if we write $\boldsymbol{A}_1 = [\boldsymbol{M}_1 \,|\, \boldsymbol{t}_1]$, $\boldsymbol{A}_2 = [\boldsymbol{M}_2 \,|\, \boldsymbol{t}_2]$, where $\boldsymbol{M}_1, \boldsymbol{M}_2$ are 2×3, then $\boldsymbol{K}_1$, $\boldsymbol{K}_2$ are the (normalized) upper triangular matrices in the RQ decomposition for $\boldsymbol{M}_1$, $\boldsymbol{M}_2$ respectively.

We next observe that for a pair canonical matrices

$$\boldsymbol{A}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \ \boldsymbol{A}_2 = \begin{bmatrix} 2\cos\theta & 2\sin\theta & 0 & 0 \\ 0 & 0 & 1 & d \end{bmatrix}, \tag{11}$$

the corresponding euclidean orbit is of the form

$$\begin{bmatrix} \boldsymbol{r}_1^T & t_1 \\ \boldsymbol{r}_3^T & t_3 \end{bmatrix}, \ \begin{bmatrix} 2\boldsymbol{r}_2^T & 2t_2 \\ \boldsymbol{r}_3^T & t_3 + d \end{bmatrix}, \tag{12}$$

where $\theta = \arccos(\boldsymbol{r}_1 \cdot \boldsymbol{r}_2)$ (this follows by applying a 4×4 euclidean transformation matrix to (11)). These cameras decompose with $\boldsymbol{K}_1$ being the identity and $\boldsymbol{K}_2 = \mathrm{diag}(2, 1)$.

Finally, if we indicate with $\boldsymbol{p}_1, \boldsymbol{p}_2$ and $2\boldsymbol{q}_1, \boldsymbol{q}_2$ the rows of (12), so that the corresponding camera can be written as $\boldsymbol{x} \mapsto \boldsymbol{u} = (\boldsymbol{p}_1^T\boldsymbol{x}/\boldsymbol{p}_2^T\boldsymbol{x}, 2\boldsymbol{q}_1^T\boldsymbol{x}/\boldsymbol{q}_2^T\boldsymbol{x}, 1)$, then the composition of $\begin{bmatrix} \boldsymbol{p}_1^T \\ \boldsymbol{p}_2^T \end{bmatrix}$, $\begin{bmatrix} \boldsymbol{q}_1^T \\ \boldsymbol{q}_2^T \end{bmatrix}$ with $\boldsymbol{K}_1, \boldsymbol{K}_2$ as in (10) yields the camera

$$\boldsymbol{x} \mapsto \left(f_u \frac{\boldsymbol{p}_1^T\boldsymbol{x}}{\boldsymbol{p}_2^T\boldsymbol{x}} + u_0, f_v \frac{2\boldsymbol{q}_1^T\boldsymbol{x}}{\boldsymbol{q}_2^T\boldsymbol{x}} + v_0, 1 \right)^T. \tag{13}$$

From this we easily deduce the physical interpretations of the entries of $\boldsymbol{K}_1$ and $\boldsymbol{K}_2$. □

We point out that a decomposition with calibration matrices is actually possible for generic finite two-slits (not necessarily "parallel"), if we allow for non triangular matrices $\boldsymbol{K}_1, \boldsymbol{K}_2$. Indeed, the four rows of $\boldsymbol{M}_1, \boldsymbol{M}_2$ will intersect in a linear space of dimension one $\langle \boldsymbol{r} \rangle$, and the second rows of $\boldsymbol{K}_1, \boldsymbol{K}_2$ can describe how to obtain $\boldsymbol{r}$ from $\boldsymbol{M}_1, \boldsymbol{M}_2$. Imposing that the diagonal elements of $\boldsymbol{K}_1, \boldsymbol{K}_2$ are positive, the decomposition is unique, and there are now $6 + 2$ ("analytic" and "3D") intrinsic, and 6 extrinsic parameters, summing up to 14 degrees of freedom of our projective two-slit camera model. On the other hand, the action of general calibration matrices is *not* a linear change of image coordinates, and requires changing retinal plane (in fact, we must switch to a "parallel plane" for the two slits).

Proposition 2. *Let $\boldsymbol{A}_1, \boldsymbol{A}_2$ define a pushbroom camera*

$$\boldsymbol{A}_1 = \begin{bmatrix} \boldsymbol{m}_1^T & t_1 \\ \boldsymbol{0} & 1 \end{bmatrix}, \ \boldsymbol{A}_2 = \begin{bmatrix} \boldsymbol{m}_2^T & t_2 \\ \boldsymbol{m}_3^T & t_3 \end{bmatrix}, \tag{14}$$

such that that $\boldsymbol{m}_1$ and $\boldsymbol{m}_3$ are orthogonal. We can uniquely write

$$\boldsymbol{A}_1 = \boldsymbol{K}_1 \begin{bmatrix} \boldsymbol{r}_1^T & t_1 \\ \boldsymbol{0} & 1 \end{bmatrix}, \boldsymbol{A}_2 = \boldsymbol{K}_2 \begin{bmatrix} \boldsymbol{r}_2^T & t_2 \\ \boldsymbol{r}_3^T & t_3 \end{bmatrix}, \tag{15}$$

where $\boldsymbol{K}_1 = \mathrm{diag}(1/v, 1)$, $\boldsymbol{K}_2 = \begin{bmatrix} f & u \\ 0 & 1 \end{bmatrix}$ (with positive v and f) and $\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3$ are unit vectors, with $\boldsymbol{r}_3$ orthogonal to both $\boldsymbol{r}_1, \boldsymbol{r}_2$. Here, $\theta = \arccos(\boldsymbol{r}_1 \cdot \boldsymbol{r}_2)$ is the angle *between the two slits (or between the direction of motion of the sensor and the parallel scanning planes). Moreover, v can be interpreted as the speed of the sensor, and f and u as the magnification and the principal point of the 1D projection.*

Proof. The proof is similar to that of Proposition 1. The decomposition is unique because of QR-factorization of matrices. The euclidean orbits of "canonical" pushbroom cameras have the form

$$\begin{bmatrix} \boldsymbol{r}_1^T & t_1 \\ \boldsymbol{0} & 1 \end{bmatrix}, \begin{bmatrix} \boldsymbol{r}_2^T & t_2 \\ \boldsymbol{r}_3^T & t_3 \end{bmatrix}, \tag{16}$$

All these cameras decompose with $\boldsymbol{K}_1, \boldsymbol{K}_2$ being the identity. Finally, the physical interpretation of the parameters follows by noting that composing a pushbroom camera (with rows $\boldsymbol{p}_1^T, (0,0,0,1)^T$ and $\boldsymbol{q}_1^T, \boldsymbol{q}_2^T$) with calibration matrices $\boldsymbol{K}_1, \boldsymbol{K}_2$ yields

$$\boldsymbol{x} \mapsto \left(\frac{1}{v} \boldsymbol{p}_1^T \boldsymbol{x}, f \frac{\boldsymbol{q}_1^T \boldsymbol{x}}{\boldsymbol{q}_2^T \boldsymbol{x}} + u, 1 \right)^T. \tag{17}$$

□

Similarly to the case of finite slits, the decomposition based on calibration matrices can be extended to the case of arbitrary pushbroom cameras, by allowing for $\boldsymbol{K}_2$ to be a general 2×2 matrix with positive entries along the diagonal. This gives a total of $4 + 1 + 6 = 11$ free parameters, which agrees with the degrees of freedom of our affine pushbroom model. However, a non-upper triangular matrix $\boldsymbol{K}_2$ does not correspond to a linear change of image coordinates as in (17), but requires changing retinal plane.

Theorem 1. *Let $(\boldsymbol{A}_1, \boldsymbol{A}_2)$, $(\boldsymbol{B}_1, \boldsymbol{B}_2)$ be two general projective two-slit cameras. The set of corresponding image points $\boldsymbol{u}$, $\boldsymbol{u}'$ in $\mathbb{P}^2$ is characterized by the following relation:*

$$\sum_{ijkl} f_{ijkl} \begin{bmatrix} u_1 \\ u_3 \end{bmatrix}_i \begin{bmatrix} u_2 \\ u_3 \end{bmatrix}_j \begin{bmatrix} u'_1 \\ u'_3 \end{bmatrix}_k \begin{bmatrix} u'_2 \\ u'_3 \end{bmatrix}_l = 0, \tag{18}$$

where $\boldsymbol{F} = (f_{ijkl})$ is a $2 \times 2 \times 2 \times 2$ "epipolar tensor". Its entries are

$$f_{ijkl} = (-1)^{i+j+k+l} \cdot \det \left[(\boldsymbol{A}_1)_{3-i}^T \ (\boldsymbol{A}_2)_{3-j}^T \ (\boldsymbol{B}_1)_{3-k}^T \ (\boldsymbol{B}_2)_{3-l}^T \right]. \tag{19}$$

Up to projective transformations of $\mathbb{P}^3$ there are two *configurations $(\boldsymbol{A}_1, \boldsymbol{A}_2), (\boldsymbol{B}_1, \boldsymbol{B}_2)$ compatible with a given epipolar tensor.*

Proof. The inverse line projection (6) can be written as

$$\chi(\boldsymbol{u}) = \sum_{ij} (-1)^{i+j} (\boldsymbol{A}_1)_{3-i} \wedge (\boldsymbol{A}_2)_{3-j} \begin{bmatrix} u_1 \\ u_3 \end{bmatrix}_i \begin{bmatrix} u_2 \\ u_3 \end{bmatrix}_j. \tag{20}$$

The definition of $\boldsymbol{F}$ is simply the condition that $\chi(\boldsymbol{u})$ and $\chi(\boldsymbol{u}')$ as in (20) are concurrent (see also [5]). Up a global scale factor, the elements of $\boldsymbol{F}$ do not depend on the scaling of the 2×4 matrices, and are fixed by projective transformations of $\mathbb{P}^3$. Hence, assuming that the vectors $(\boldsymbol{A}_1)_1, (\boldsymbol{A}_2)_1, (\boldsymbol{B}_1)_1, (\boldsymbol{B}_2)_1$ are independent (which is true generically) we can apply a change of reference frame in $\mathbb{P}^3$ so that the projection matrices have the form

$$\boldsymbol{A}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ c_{11} & c_{12} & c_{13} & c_{14} \end{bmatrix}, \ \boldsymbol{A}_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ c_{21} & c_{22} & c_{23} & c_{24} \end{bmatrix},$$
$$\boldsymbol{B}_1 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ c_{31} & c_{32} & c_{33} & c_{34} \end{bmatrix}. \ \boldsymbol{B}_2 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ c_{41} & c_{42} & c_{43} & c_{44} \end{bmatrix}. \tag{21}$$

The 16 entries of $\boldsymbol{F}$ are now (up to sign) the *principal minors* of the 4×4-matrix $\boldsymbol{C} = (c_{ij})$: more precisely, $f_{ijkl} = (-1)^{i+j+k+l} \det \boldsymbol{C}_{[i-1,j-1,k-1,l-1]}$ where $\boldsymbol{C}_{[i-1,j-1,k-1,l-1]}$ is the submatrix of $\boldsymbol{C}$ where the selected rows and columns correspond to the binary vector $[i-1, j-1, k-1, l-1]$ (for example, $\boldsymbol{C}_{[1,0,0,0]} = (c_{11})$). Determining valid projection matrices $(\boldsymbol{A}_1, \boldsymbol{A}_2)$, $(\boldsymbol{B}_1, \boldsymbol{B}_2)$ given the tensor $\boldsymbol{F}$, is equivalent to finding the entries of the 4×4-matrix $\boldsymbol{C}$ given its principal minors. This problem is studied in [4]. Under generic conditions, the set of all matrices with the same principal minors as $\boldsymbol{C}$ have the form $\boldsymbol{D}^{-1}\boldsymbol{C}\boldsymbol{D}$ or $\boldsymbol{D}^{-1}\boldsymbol{C}^T\boldsymbol{D}$, where $\boldsymbol{D}$ is a diagonal matrix [4]. Each of these two families of matrices is a projective configuration of cameras, and the two configurations are in general distinct (see the discussion in the next section). □

3. Algorithms

3.1. Linear SfM

We assume that we are given pairs of corresponding image points $(\boldsymbol{u}_i, \boldsymbol{u}'_i)$, $i = 1, \ldots, n$, for two unknown two-slit cameras. Each pair yields a linear constraint on the epipolar tensor $\boldsymbol{F}$ in (18). Hence, if $n \geq 15$ correspondences are given, we can compute a linear estimate for $\boldsymbol{F}$. For noisy data, this estimate will not be a valid epipolar tensor, since tensors of the form (18) are not generic. However, it is possible to recover projection matrices from only 13 of the entries of $\boldsymbol{F}$, which avoids the problem of using a valid tensor. A simple scheme for this is as follows:

1. We set out to recover the entries of a 4×4-matrix $\boldsymbol{C}$ given its principal minors. Since we can always replace $\boldsymbol{C}$ with $\boldsymbol{D}^{-1}\boldsymbol{C}\boldsymbol{D}$, where $\boldsymbol{D}$ is a diagonal matrix, we can assume that $c_{12} = c_{13} = c_{13} = c_{14} = 1$ (at least generically). Other similar assignments are possible.

2. Elements on the diagonal and on the first column of $\boldsymbol{C}$ are easily computed given (seven of the entries of) $\boldsymbol{F}$:

   - $c_{11} = -f_{1222}$; $c_{22} = -f_{2122}$; $c_{33} = -f_{2212}$; $c_{44} = -f_{2221}$.
   - $c_{21} = (c_{11}c_{22} - f_{1122})/c_{12}$; $c_{31} = (c_{11}c_{33} - f_{1212})/c_{13}$; $c_{41} = (c_{11}c_{44} - f_{1221})/c_{14}$.

   Here the elements to the right of the equal signs have already been assigned. Hence, we recover 10 entries of $\boldsymbol{C}$ from linear equalities.

3. The remaining six entries of $\boldsymbol{C}$ are pairwise constrained by six elements of $\boldsymbol{F}$. For example, using the minors f_{2112}, f_{1112} (corresponding to rows/columns $2, 3$ and $1, 2, 3$ of $\boldsymbol{C}$) we deduce that c_{32} must satisfy $ac_{32}^2 + bc_{32} + c = 0$ where

$$\begin{aligned} a &= c_{13}c_{21} \\ b &= f_{1112} + c_{11}f_{2112} - c_{13}c_{31}c_{22} - c_{12}c_{21}c_{33} \\ c &= c_{12}c_{31}c_{22}c_{33} - c_{12}c_{31}f_{2112}, \end{aligned} \tag{22}$$

and that $c_{23} = (c_{22}c_{33} - f_{2112})/c_{32}$. Similar relations hold for the pairs c_{24}, c_{42} and c_{34}, c_{43}. This leads to 8 possible matrices $\boldsymbol{C}$, *i.e.*, a finite number of camera configurations. Note however that the entries f_{1111} and f_{2111} of $\boldsymbol{F}$ were never used (which is why we can assume the tensor to be generic): in an ideal setting with no noise, exactly two of the 8 solutions will be consistent with the remaining constraints.

This approach for recovering two-slit projections from the corresponding epipolar tensor relies on some genericity assumptions (*e.g.*, we have often divided by element without verifying that it is not zero), and developing an optimal strategy for this task is outside the scope of our work. Nevertheless, we include as a proof of concept some results.

Experiments. We present a concrete example illustrating some basic properties of the fundamental tensor. We consider the following pairs of projection matrices:

$$\begin{aligned} \boldsymbol{A}_1 &= \begin{bmatrix} -1 & 7 & 4 & 0 \\ 8 & -1 & 13 & 4 \end{bmatrix}, \boldsymbol{A}_2 = \begin{bmatrix} 11 & 6 & -2 & 4 \\ 8 & -1 & 13 & -5 \end{bmatrix} \\ \boldsymbol{B}_1 &= \begin{bmatrix} 14 & 9 & -3 & 8 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \boldsymbol{B}_2 = \begin{bmatrix} -3 & 8 & 10 & 3 \\ 6 & 13 & 5 & 13 \end{bmatrix} \end{aligned} \tag{23}$$

The pair $\boldsymbol{A}_1, \boldsymbol{A}_2$ represents a parallel finite two-slit camera, while $\boldsymbol{B}_1, \boldsymbol{B}_2$ is a pushbroom camera. The associated epipolar tensor (18) is

$$\begin{aligned} \boldsymbol{F} = &\begin{bmatrix} 0 & 0 \\ 21816 & -25650 \end{bmatrix} \begin{bmatrix} 1906 & -2090 \\ -3642 & 5510 \end{bmatrix} \\ &\begin{bmatrix} 880 & 475 \\ 18600 & -11875 \end{bmatrix} \begin{bmatrix} 97 & -380 \\ -1259 & 1425 \end{bmatrix}, \end{aligned} \tag{24}$$

where each 4×4 matrix represents a block $(f_{ijkl})_{kl}$ for fixed i, j. Note that f_{1111} and f_{1112} are zero, since the second rows of $\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{B}_1$ are linearly dependent. Using the approach outlined above, we can use this tensor to recover two matrices $\boldsymbol{C}_1, \boldsymbol{C}_2$ whose principal minors are the entries of $\boldsymbol{F}$ (we must normalize $\boldsymbol{F}$ so that $f_{2222} = 1$) . By construction, $\boldsymbol{C}_1$ and $\boldsymbol{C}_2$ differ only for six elements. We use these matrices to construct two pairs of two-silt cameras, namely

$$\begin{aligned} \boldsymbol{A}_1^1 &= \begin{bmatrix} 1. & 0 & 0 & 0 \\ -3.87 & 1. & 1. & 1. \end{bmatrix} \\ \boldsymbol{A}_2^1 &= \begin{bmatrix} 0. & 1. & 0. & 0. \\ -14.22 & 8.33 & -6.67 & -22.17 \end{bmatrix}, \\ \boldsymbol{B}_1^1 &= \begin{bmatrix} 0. & 0. & 1. & 0. \\ 0.44 & -0.28 & 0.27 & 1.14 \end{bmatrix} \\ \boldsymbol{B}_2^1 &= \begin{bmatrix} 0. & 0. & 0. & 1. \\ -0.86 & 0.26 & 0.15 & 0.88 \end{bmatrix}, \end{aligned} \tag{25}$$

and

$$\begin{aligned} \boldsymbol{A}_1^2 &= \begin{bmatrix} 1. & 0 & 0 & 0 \\ -3.87 & 1. & 1. & 1. \end{bmatrix} \\ \boldsymbol{A}_2^2 &= \begin{bmatrix} 0. & 1. & 0. & 0. \\ -14.22 & 8.33 & 9.25 & 4.24 \end{bmatrix}, \\ \boldsymbol{B}_1^2 &= \begin{bmatrix} 0. & 0. & 1. & 0. \\ 0.44 & 0.20 & 0.27 & -0.07 \end{bmatrix} \\ \boldsymbol{B}_2^2 &= \begin{bmatrix} 0. & 0. & 0. & 1. \\ -0.86 & -1.34 & -2.26 & 0.88 \end{bmatrix}. \end{aligned} \tag{26}$$

Computing the epipolar tensor (18) for both of these pairs yields $\boldsymbol{F}$ as in (24). On the other hand, the two camera configurations are *not* projectively equivalent: indeed, if a projective transformation between the two existed, it would need to be the identity, because five of the eight rows coincide. It is straightforward to verify that it is in fact the second pair that corresponds to the configuration of the original cameras (23).

We now try to recover the same cameras using image correspondences. We consider 70 random points in space, project them using (23), and add some noise to the images. In this case, none of original the eight solutions will be exactly consistent with the last two entries of $\boldsymbol{F}$, however we can consider the two solutions that minimize an "algebraic residual" for these constraints. For image sizes of about 100×100, and noise with a standard deviation of 10^{-5}, we recover the following pairs of cameras (that should be compared with (25)):

$$\begin{aligned} \boldsymbol{A}_1^1 &= \begin{bmatrix} 1. & 0 & 0 & 0 \\ -3.97 & 1. & 1. & 1. \end{bmatrix} \\ \boldsymbol{A}_2^1 &= \begin{bmatrix} 0. & 1. & 0. & 0. \\ -15.26 & 8.44 & -7.60 & -23.18 \end{bmatrix}, \\ \boldsymbol{B}_1^1 &= \begin{bmatrix} 0. & 0. & 1. & 0. \\ 0.42 & -0.25 & 0.27 & 1.17 \end{bmatrix} \\ \boldsymbol{B}_2^1 &= \begin{bmatrix} 0. & 0. & 0. & 1. \\ -0.86 & 0.25 & 0.14 & 0.88 \end{bmatrix}, \end{aligned} \tag{27}$$

and

$$\begin{aligned} \boldsymbol{A}_1^2 &= \begin{bmatrix} 1. & 0 & 0 & 0 \\ -3.97 & 1. & 1. & 1. \end{bmatrix} \\ \boldsymbol{A}_2^2 &= \begin{bmatrix} 0. & 1. & 0. & 0. \\ -15.26 & 8.44 & 9.36 & 4.41 \end{bmatrix} \end{aligned} \tag{28}$$

$$
\boldsymbol{B}_1^2 = \begin{bmatrix} 0. & 0. & 1. & 0. \\ 0.42 & 0.20 & 0.27 & -0.07 \end{bmatrix}
$$
$$
\boldsymbol{B}_2^2 = \begin{bmatrix} 0. & 0. & 0. & 1. \\ -0.86 & -1.30 & -2.42 & 0.88 \end{bmatrix}. \tag{29}
$$

3.2. Minimal SfM

A non-linear "minimal" approach for estimating the epipolar tensor requires 13 corresponding image points. Substituting these correspondences in (18), we obtain an under-determined linear system, which implies that the epipolar tensor is a linear combination $\alpha T_1 + \beta T_2 + \gamma T_3$ for some T_1, T_2, T_3 that generate the corresponding null-space. Since the variety of epipolar tensors has codimension 2 in $\mathbb{P}^{15}$, we expect to find a finite number of feasible tensors in this linear space (up to scale factors). According to [4, Remark 14], the variety of epipolar tensors (that is viewed there as the projective variety for the principal minors of 4×4 matrices) has degree 28. Hence, this minimal approach should lead to 28 complex solutions for $\boldsymbol{F}$, and 56 projective configurations of cameras. Using the computer algebra system `Macaulay2` [3] we have verified (over finite fields) that imposing 13 general linear combinations of the 16 principal minors of the matrix $\boldsymbol{C}$ (so each linear condition can be viewed as a point correspondence), and fixing $c_{12} = c_{13} = c_{14} = 1$, we obtain 56 solutions $\boldsymbol{C}$ in the algebraic closure of the field.

3.3. Self-calibration

We describe a strategy for *self-calibration* for two-slit cameras. We assume that we have recovered a projective reconstruction $\boldsymbol{A}_1^i, \boldsymbol{A}_2^i$ for $i = 1, \ldots, n$ for finite two-slit cameras (that we assume were originally "parallel"). We indicate with $\boldsymbol{Q}$ a "euclidean upgrade", that is, a 4×4-matrix that describes the transition from a euclidean reference frame to the frame corresponding to our projective reconstruction. According to Proposition 1, we may write $\boldsymbol{A}_1^i \boldsymbol{Q} = \boldsymbol{K}_i [\boldsymbol{R}_1^i \,|\, \boldsymbol{t}_1^i]$, $\boldsymbol{A}_2^i \boldsymbol{Q} = \boldsymbol{K}_2^i [\boldsymbol{R}_2^i \,|\, \boldsymbol{t}_2^i]$, where $\boldsymbol{R}_1^i, \boldsymbol{R}_2^i$ are 2×3 matrices with orthonormal rows. In particular, for all $i = 1, \ldots, k$, we have

$$
\begin{aligned}
\boldsymbol{A}_1^i \boldsymbol{Q} \boldsymbol{\Omega}^* \boldsymbol{Q}^T {\boldsymbol{A}_1^i}^T &= \boldsymbol{K}_1^i {\boldsymbol{K}_1^i}^T \\
\boldsymbol{A}_2^i \boldsymbol{Q} \boldsymbol{\Omega}^* \boldsymbol{Q}^T {\boldsymbol{A}_2^i}^T &= \boldsymbol{K}_2^i {\boldsymbol{K}_2^i}^T,
\end{aligned} \tag{30}
$$

where equality is up to scale and $\boldsymbol{\Omega}^* = \mathrm{diag}(1,1,1,0)$. Geometrically, the matrix $\boldsymbol{\Omega}_Q^* = \boldsymbol{Q}\boldsymbol{\Omega}^*\boldsymbol{Q}^T$ represents the *dual of the absolute conic*, in the projective coordinates used in the reconstruction. The equations (30) identify in fact the *dual of the image of the absolute conic* in the two copies of $\mathbb{P}^1$. These are the set of planes containing each slit that are tangent to the absolute conic in $\mathbb{P}^3$.

We now assume that the principal points cameras are at the "origin", so that $\boldsymbol{K}_1^i, \boldsymbol{K}_2^i$ (and hence $\boldsymbol{K}_1^i {\boldsymbol{K}_1^i}^T$ and $\boldsymbol{K}_2^i {\boldsymbol{K}_2^i}^T$) are diagonal. Each row in (30) gives two linear equations in the elements of $\boldsymbol{\Omega}_Q^*$, corresponding to the zeros in the matrices on the right hand side. For example, imposing that the $(1,2)$-entry of $\boldsymbol{K}_1^i {\boldsymbol{K}_1^i}^T$ is zero yields

$$
\begin{aligned}
& a_{11}a_{21}m_{11} + a_{11}a_{22}m_{12} + a_{11}a_{23}m_{13} + a_{11}a_{24}m_{14} \\
& + a_{12}a_{21}m_{21} + a_{12}a_{22}m_{22} + a_{12}a_{23}m_{23} + a_{12}a_{24}m_{24} \\
& + a_{13}a_{21}m_{31} + a_{13}a_{22}m_{32} + a_{13}a_{23}m_{33} + a_{13}a_{24}m_{34} \\
& + a_{14}a_{21}m_{41} + a_{14}a_{22}m_{42} + a_{14}a_{23}m_{43} + a_{14}a_{24}m_{44} = 0,
\end{aligned} \tag{31}
$$

where $\boldsymbol{\Omega}_Q^* = (m_{ij})$, and the elements of $\boldsymbol{A}_1^i = (a_{ij})$ are known. A sufficient number of views allows us to estimate $\boldsymbol{\Omega}_Q^*$ linearly. Finally, from the singular value decomposition of $\boldsymbol{\Omega}_Q^*$, we can compute a matrix $\boldsymbol{Q}'$ such that $\boldsymbol{Q}'\boldsymbol{\Omega}^*\boldsymbol{Q}'^T = \boldsymbol{\Omega}_Q^*$. The matrix $\boldsymbol{Q}'$ is however not uniquely determined, and indeed we can actually only recover a *similarity* upgrade, since any similarity transformation will fix the absolute conic in $\mathbb{P}^3$.

Experiments. To apply our self-calibration scheme, we consider 10 cameras $\boldsymbol{A}_1^i, \boldsymbol{A}_2^i$, $i = 1, \ldots, 10$, of the form $\boldsymbol{A}_1^i = \boldsymbol{K}_1^i [\boldsymbol{R}_1^i \,|\, \boldsymbol{t}_1^i] \boldsymbol{Q}^{-1}$, $\boldsymbol{A}_2^i = \boldsymbol{K}_2^i [\boldsymbol{R}_2^i \,|\, \boldsymbol{t}_2^i] \boldsymbol{Q}^{-1}$, where $\boldsymbol{R}_1^i, \boldsymbol{t}_1^i, \boldsymbol{R}_2^i, \boldsymbol{t}_2^i$ are random parameters for euclidean primitive parallel cameras, $\boldsymbol{K}_1^i, \boldsymbol{K}_2^i$ are random *diagonal* calibration matrices, and $\boldsymbol{Q}$ is a random 4×4 matrix describing a projective change of coordinates. We also add small amounts of noise to the entries of $\boldsymbol{A}_1^i, \boldsymbol{A}_2^i$. The matrices $\boldsymbol{A}_1^i, \boldsymbol{A}_2^i$ represent a *projective* configuration of two-slit cameras. Using (30), we can recover an estimate for $\boldsymbol{\Omega}_Q^* = \boldsymbol{Q}\boldsymbol{\Omega}^*\boldsymbol{Q}^T$ by solving an over-constrained linear system (with 40 equations). From this, we compute a matrix $\boldsymbol{Q}'$ such that $\boldsymbol{Q}'\boldsymbol{\Omega}^*\boldsymbol{Q}'^T \simeq \boldsymbol{\Omega}_Q^*$. For our example, the original data was

$$
\begin{aligned}
\boldsymbol{Q} &= \begin{bmatrix} 1.49 & 0.60 & -0.11 & -1.15 \\ -1.43 & 0.88 & -0.93 & 1.52 \\ -0.38 & -0.21 & 1.83 & -0.55 \\ 0.83 & -0.95 & -0.63 & 0.93 \end{bmatrix}, \\
\boldsymbol{Q}\boldsymbol{\Omega}^*\boldsymbol{Q}^T &= \begin{bmatrix} 1. & -0.58 & -0.34 & 0.28 \\ -0.58 & 1.42 & -0.52 & -0.55 \\ -0.34 & -0.52 & 1.36 & -0.49 \\ 0.28 & -0.55 & -0.49 & 0.77 \end{bmatrix},
\end{aligned} \tag{32}
$$

while our estimates are

$$
\begin{aligned}
\boldsymbol{Q}' &= \begin{bmatrix} -0.43 & 0.21 & 0.35 & 0. \\ 0.67 & 0.26 & 0.08 & 0. \\ -0.04 & -0.69 & 0.03 & 0. \\ -0.34 & 0.26 & -0.28 & 1. \end{bmatrix}, \\
\boldsymbol{Q}'\boldsymbol{\Omega}^*\boldsymbol{Q}'^T &= \begin{bmatrix} 1. & -0.59 & -0.34 & 0.29 \\ -0.59 & 1.44 & -0.51 & -0.56 \\ -0.34 & -0.51 & 1.35 & -0.48 \\ 0.29 & -0.56 & -0.48 & 0.75 \end{bmatrix}.
\end{aligned} \tag{33}
$$

The matrices $\boldsymbol{Q}$, $\boldsymbol{Q}'$ are not close, however one easily verifies that $\boldsymbol{Q}^{-1}\boldsymbol{Q}'$ is (almost) a similarity transformation.

In particular, the cameras $\boldsymbol{A}_1^i\boldsymbol{Q}', \boldsymbol{A}_2^i\boldsymbol{Q}'$, $i = 1, \ldots, 10$ are a "similarity upgrade" of the projective solution. For example, for the first of our 10 original cameras we had $\boldsymbol{K}_1^1 = \text{diag}(4.04, 1)$, $\boldsymbol{K}_2 = \text{diag}(1.37, 1)$, and indeed

$$\begin{aligned} \boldsymbol{A}_1^1\boldsymbol{Q}' &= \begin{bmatrix} -2.07 & -1.29 & 3.23 & 13.25 \\ 0.39 & -0.91 & -0.12 & -0.08 \end{bmatrix}, \\ \boldsymbol{A}_2^1\boldsymbol{Q}' &= \begin{bmatrix} -0.49 & -0.36 & 1.24 & 2.81 \\ 0.38 & -0.91 & -0.12 & 0.53 \end{bmatrix}, \end{aligned} \tag{34}$$

describe a parallel two-slit camera, where the ratios between the norms of the rows (the "magnifications") are respectively 4.05 and 1.38.

References

[1] D. Feldman, T. Pajdla, and D. Weinshall. On the epipolar geometry of the crossed-slits projection. In *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, pages 988–995. IEEE, 2003. 1

[2] G. H. Golub and C. F. Van Loan. *Matrix computations*, volume 3. JHU Press, 2012. 2

[3] D. R. Grayson and M. E. Stillman. Macaulay2, a software system for research in algebraic geometry. Available at http://www.math.uiuc.edu/Macaulay2/. 5

[4] S. Lin and B. Sturmfels. Polynomial relations among principal minors of a 4× 4-matrix. *Journal of Algebra*, 322(11):4121–4131, 2009. 3, 5

[5] J. Ponce, B. Sturmfels, and M. Trager. Congruences and concurrent lines in multi-view geometry. *arXiv preprint arXiv:1608.05924*, 2016. 3

[6] A. Zomet, D. Feldman, S. Peleg, and D. Weinshall. Mosaicing new views: the crossed-slits projection. *PAMI*, 25(6):741–754, 2003. 1