Differential Matching Constraints

Bill Triggs
www.inrialpes.fr/movi/people/Triggs

Contributions

[1] A ***finite difference expansion*** for vision geometry with closely spaced cameras — rotation small or nulled by image alignment, translation $\ll$ scene distance.

[2] Apply formalism to ***projective matching constraints*** and their tensors, for all possible combinations of near coincidence and finite spacing among images.

Main Conclusions

[1] Difference expansion gives ***much simpler results*** than Taylor series approaches.

[2] **Tensor tracking** Propagate matching relations *w.r.t.* base image(s) along an image sequence, matching as you go.

- Difference expansion *linearizes tensor consistency constraints & error model*
- Optimal tensor estimation becomes an easy constrained linear least squares problem.
- Equivalent to *1 iteration of a nonlinear tensor estimator* started at previous tensor.

[3] Tri- & quadrifocal constraints: expansion is possible but it seems simpler to work with raw projections (or equivalently homography-epipole parametrizations).

[4] ***Three cases*** occur during expansion :

[A] Problem is *linear* in expansion variable : first order expansion is an exact but trivial linear change of variables. *E.g.* features occur linearly in matching relations, so *nothing is gained by using displacement/flow instead of image positions*.

[B] Nonlinear problem expanded *w.r.t.* a nonsingular base case : First order expansion is useful and linearizes problem. *E.g.* tensor tracking subject to consistency constraints.

[C] Nonlinear problem, singular base case : Higher order expansion is needed. Constraints may be simpler but are still nonlinear, so it is not clear that much is gained. *E.g.* tensor tracking when too many images coincide and base tensor vanishes.

Finite Matching Tensors & Matching Constraints

- ***Matching tensors*** are *multilinear forms* in 4 projection matrices $\mathbf{P}_i$:

$\boldsymbol{e}_{12} \equiv \boldsymbol{e}(1,1,1,2)$	epipole (vector)	"1" stands for $\mathbf{P}_1$, *etc*
$\boldsymbol{F}_{12} \equiv \boldsymbol{F}(1,1,2,2)$	fundamental matrix	
$\boldsymbol{T}_1{}^{23} \equiv \boldsymbol{T}(1,1,2,3)$	trifocal tensor	
$\boldsymbol{Q}^{1234} \equiv \boldsymbol{Q}(1,2,3,4)$	quadrifocal tensor	

- $\boldsymbol{e}, \boldsymbol{F}, \boldsymbol{T}$ are 'compacted' versions of $\boldsymbol{Q}$ — they are simpler, but contain the same information when the projections are repeated as shown.
- Valid matching tensors obey ***nonlinear consistency constraints*** that guarantee factorization into 3×4 projections — *e.g.* $\boldsymbol{F}_{12}\,\boldsymbol{e}_{12} = \boldsymbol{0}$, $\det(\boldsymbol{F}_{12}) = 0 \ldots$
- Tensors generate multilinear inter-image ***feature matching constraints***, *e.g.*:

$\boldsymbol{x}_1^\top \boldsymbol{F}_{12}\,\boldsymbol{x}_2 = 0$	epipolar constraint
$\boldsymbol{x}_2 \wedge (\boldsymbol{T}_{123} \cdot \boldsymbol{x}_1) \wedge \boldsymbol{x}_3 = \boldsymbol{0}$	trifocal point constraint
$\boldsymbol{l}_2^\top\,(\boldsymbol{T}_1{}^{23} \wedge \boldsymbol{l}_1)\;\boldsymbol{l}_3 = \boldsymbol{0}$	trifocal line constraint
$\boldsymbol{l}_2^\top\,(\boldsymbol{T}_1{}^{23} \wedge \boldsymbol{x}_1)\;\boldsymbol{l}_3 = \boldsymbol{0}$	trifocal point-line constraint

where $\boldsymbol{x}_i$, $\boldsymbol{l}_i$ denote corresponding image points, lines

Projective Finite Difference Expansion

- We measure & use features only at discrete times. ***Finite differences*** $\Delta\boldsymbol{x} = \boldsymbol{x}(\delta t) - \boldsymbol{x}(0)$ are directly observable, derivatives $\frac{d^k\boldsymbol{x}}{dt^k}$ aren't. Taylor expansion $\Delta\boldsymbol{x} = \frac{d\boldsymbol{x}}{dt}\delta t + \frac{1}{2}\frac{d^2\boldsymbol{x}}{dt^2}(\delta t)^2 + \ldots$ is not an appropriate parametrization — infinitely many unobservable unknowns !
- *E.g.* Åström & Heyden's (CVPR'98, IJCV'98) Taylor formalism gives an *infinite series of very complicated differential matching tensors & relations*, whereas finite differences give just a few relatively simple finite expansions.
- For camera projections $\mathbf{P}'_1$ "near" $\mathbf{P}_1$, *etc*, expand matching tensors by powers of $\boxed{\Delta\mathbf{P}_1 \equiv \mathbf{P}'_1 - \mathbf{P}_1}$:

$$
\begin{aligned}
\boldsymbol{e}_{1'}^2 &= \boldsymbol{e}_1^2 + \boldsymbol{e}_{\Delta 1}^2 + \boldsymbol{e}_{\Delta^2 1}^2 + \boldsymbol{e}_{\Delta^3 1}^2 & \qquad \boldsymbol{e}_1^{2'} &= \boldsymbol{e}_1^2 + \boldsymbol{e}_1^{\Delta 2} \\
\boldsymbol{F}_{1'2} &= \boldsymbol{F}_{12} + \boldsymbol{F}_{\Delta 12} + \boldsymbol{F}_{\Delta^2 12} & \boldsymbol{T}_1^{2'3} &= \boldsymbol{T}_1^{23} + \boldsymbol{T}_1^{\Delta 23} \\
\boldsymbol{T}_{1'}^{23} &= \boldsymbol{T}_1^{23} + \boldsymbol{T}_{\Delta 1}^{23} + \boldsymbol{T}_{\Delta^2 1}^{23} & \boldsymbol{Q}^{1'234} &= \boldsymbol{Q}^{1234} + \boldsymbol{Q}^{\Delta 1234}
\end{aligned}
$$

- Here, the **differential matching tensors** are :

$$
\begin{aligned}
\boldsymbol{e}_{\Delta 1}^2 &\equiv 3\,\boldsymbol{e}(\Delta 1,1,1,2) & \boldsymbol{F}_{\Delta 12} &\equiv 2\,\boldsymbol{F}(\Delta 1,1,2,2) & \boldsymbol{T}_1^{\Delta 23} &\equiv \boldsymbol{T}(1,1,\Delta 2,3) \\
\boldsymbol{e}_{\Delta^2 1}^2 &\equiv 3\,\boldsymbol{e}(\Delta 1,\Delta 1,1,2) & \boldsymbol{F}_{\Delta^2 12} &\equiv \boldsymbol{F}(\Delta 1,\Delta 1,2,2) & \boldsymbol{T}_{\Delta 1}^{23} &\equiv 2\,\boldsymbol{T}(\Delta 1,1,2,3) \\
\boldsymbol{e}_1^{\Delta 2} &\equiv \boldsymbol{e}(1,1,1,\Delta 2) & \boldsymbol{Q}^{\Delta 1234} &\equiv \boldsymbol{Q}(\Delta 1,2,3,4) & \boldsymbol{T}_{\Delta^2 1}^{23} &\equiv \boldsymbol{T}(\Delta 1,\Delta 1,2,3)
\end{aligned}
$$

Differential Matching Constraints

- Simply substitute differential tensors into finite matching constraints and expand.
- Quantities like $\Delta\mathbf{P}_1$ depend on homogeneous scales chosen for $\mathbf{P}_1, \mathbf{P}'_1$. Rescaling invariance implies that *differential matching constraints are only defined modulo multiples of the underlying base constraints*.

Differential Epipolar Constraint

$$
0 = \boldsymbol{x}_1^\top \boldsymbol{F}_{12'}\,\boldsymbol{x}_{2'} \approx \boldsymbol{x}_1^\top (\boldsymbol{F}_{12} + \boldsymbol{F}_{1\Delta 2})\,\boldsymbol{x}'_2 \approx \boldsymbol{x}_1^\top \boldsymbol{F}_{12}\,\Delta\boldsymbol{x}_2 + \boldsymbol{x}_1^\top \boldsymbol{F}_{1\Delta 2}\,\boldsymbol{x}_2
$$

$$
\operatorname{trace}\left(\operatorname{cof}(\boldsymbol{F}_{12})\,\boldsymbol{F}_{1\Delta 2}\right) + \det(\boldsymbol{F}_{12}) = 0 \qquad \text{consistency}
$$

$$
\operatorname{trace}\left(\boldsymbol{F}_{12}^\top \boldsymbol{F}_{1\Delta 2}\right) = 0 \qquad \text{normalization}
$$

- 'Seven point' estimation of either $\boldsymbol{F}_{1\Delta 2}$ or $\boldsymbol{F}_{12'}$ from correspondences and $\boldsymbol{F}_{12}$ is a simple constrained linear least squares problem.
- The update amounts to one iteration of a nonlinear estimator for $\boldsymbol{F}_{12'}$ started from $\boldsymbol{F}_{12}$
- ***Tensor tracking*** : this "unrolling of the estimation loop" along the sequence gives a kind of linearized control law for $\boldsymbol{F}(t) = \boldsymbol{F}(\mathbf{P}_1, \mathbf{P}_2(t))$, *c.f.* Soatto & Perona IJCV'97 for $\boldsymbol{E}$-matrix.

Differential Trifocal Constraints

- First order expansion of the 1–2′–3 and 1′–2–3 trifocal point, line and point-line matching constraints modulo the 1–2–3 ones gives:

$$
\begin{aligned}
&\left(\boldsymbol{x}_2 \wedge (\boldsymbol{T}_1^{\Delta 23} \cdot \boldsymbol{x}_1) + \Delta\boldsymbol{x}_2 \wedge (\boldsymbol{T}_1^{23} \cdot \boldsymbol{x}_1)\right) \wedge \boldsymbol{x}_3 \approx 0 && \boldsymbol{x}_2 \wedge \left(\boldsymbol{T}_{\Delta 1}^{23} \cdot \boldsymbol{x}_1 + \boldsymbol{T}_1^{23} \cdot \Delta\boldsymbol{x}_1\right) \wedge \boldsymbol{x}_3 \approx 0 \\
&\left(\boldsymbol{l}_2^\top (\boldsymbol{T}_1^{\Delta 23} \wedge \boldsymbol{l}_1) + \Delta\boldsymbol{l}_2^\top (\boldsymbol{T}_1^{23} \wedge \boldsymbol{l}_1)\right) \boldsymbol{l}_3 \approx 0 && \boldsymbol{l}_2^\top \left(\boldsymbol{T}_{\Delta 1}^{23} \wedge \boldsymbol{l}_1 + \boldsymbol{T}_1^{23} \wedge \Delta\boldsymbol{l}_1\right) \boldsymbol{l}_3 \approx 0 \\
&\left(\boldsymbol{l}_2^\top (\boldsymbol{T}_1^{\Delta 23} \cdot \boldsymbol{x}_1) + \Delta\boldsymbol{l}_2^\top (\boldsymbol{T}_1^{23} \cdot \boldsymbol{x}_1)\right) \boldsymbol{l}_3 \approx 0 && \boldsymbol{l}_2^\top \left(\boldsymbol{T}_{\Delta 1}^{23} \cdot \boldsymbol{x}_1 + \boldsymbol{T}_1^{23} \cdot \Delta\boldsymbol{x}_1\right) \boldsymbol{l}_3 \approx 0
\end{aligned}
$$

- However, it seems easier in practice to use projection matrices . . .

Coincident Images

- Finite tensors & matching constraints take special forms when their base images coincide :

$$
\boldsymbol{T}_1{}^{12} = \boldsymbol{\delta}_1^1 \otimes \boldsymbol{e}_1^2 \qquad \boldsymbol{T}_1{}^{21} = -\boldsymbol{e}_1^2 \otimes \boldsymbol{\delta}_1^1 \qquad \boldsymbol{T}_{A1}{}^{B2\,C2} = \boldsymbol{F}_{A1\,A2}\,\boldsymbol{\epsilon}^{A2\,B2\,C2}
$$

$$
\boldsymbol{x}_2 \wedge \left(\boldsymbol{T}_1{}^{23} \cdot \boldsymbol{x}_1\right) \wedge \boldsymbol{x}_3 = 0 \quad \longrightarrow \quad \left(\boldsymbol{x}_1^\top \boldsymbol{F}_{12}\,\boldsymbol{x}_2\right)\,[\,\boldsymbol{x}_2\,]_\times = 0
$$

- Similarly for the differential tensors :

$$
\begin{aligned}
-\boldsymbol{e}_{\Delta 1}^1 &= \boldsymbol{e}_1^{\Delta 1} = \boldsymbol{e}_1^{1'} & \qquad \boldsymbol{F}_{1\Delta 1} &= \left[\boldsymbol{e}_1^{\Delta 1}\right]_\times = \left[\boldsymbol{e}_1^{1'}\right]_\times \\
\boldsymbol{T}_1^{1\Delta 2} &= \boldsymbol{\delta}_1^1 \otimes \boldsymbol{e}_1^{\Delta 2} & \boldsymbol{T}_{\Delta 1}^{12} &= \boldsymbol{\delta}_1^1 \otimes \boldsymbol{e}_{\Delta 1}^2 - \boldsymbol{T}_1^{\Delta 12}
\end{aligned}
$$

Differential Epipolar Constraint — Coincident Images

- For $2 \to 1$ coincidence, *exact* expansion $\boldsymbol{F}_{12} \longrightarrow \left[\boldsymbol{e}_1^2\right]_\times + \boldsymbol{F}_{1\Delta^2 1}$ gives Viéville & Faugeras' ***first order motion equation*** (ICCV'95, CVIU'96) :

$$
\boldsymbol{x}_1^\top \boldsymbol{F}_{12}^{(s)}\,\boldsymbol{x}_1 + \boldsymbol{x}_1^\top \left[\boldsymbol{e}_1^2\right]_\times \Delta\boldsymbol{x}_1 \approx 0
$$

$$
\boldsymbol{e}_1^{2\,\top} \boldsymbol{F}_{12}^{(s)}\,\boldsymbol{e}_1^2 = 0 \qquad \text{consistency}
$$

$$
\boldsymbol{F}_{12}^{(s)} \equiv \frac{1}{2}\left(\boldsymbol{F}_{12} + \boldsymbol{F}_{12}^\top\right) = \frac{1}{2}\left(\boldsymbol{F}_{1\Delta^2 1} + \boldsymbol{F}_{1\Delta^2 1}^\top\right)
$$

- The consistency constraint remains nonlinear because the base tensor $\boldsymbol{F}_{11}$ vanishes.
- My experiments show *no* advantages (accuracy, simplicity, speed) over standard linear 8 point & nonlinear 7 point estimators. And there is $\mathcal{O}\left(\|\Delta\boldsymbol{F}\|^2\right)$ truncation bias as expected.

Differential Trifocal Constraints — Coincident Images

- The 1-1′-2 or 1′-1-2, 2-1-1′, and 1-1′-1″ differential trifocal constraints are

$$
\begin{aligned}
\boldsymbol{x}_1 \wedge \left(\boldsymbol{T}_1^{\Delta 12} \cdot \boldsymbol{x}_1\right) \wedge \boldsymbol{x}_2 - (\boldsymbol{x}_1 \wedge \Delta\boldsymbol{x}_1)\,\left(\boldsymbol{e}_1^2 \wedge \boldsymbol{x}_2\right)^\top &\approx \boldsymbol{0} \\
\boldsymbol{l}_1^\top \left(\boldsymbol{T}_1^{\Delta 12} \wedge \boldsymbol{l}_1\right)\,\boldsymbol{l}_2 - (\boldsymbol{l}_1 \wedge \Delta\boldsymbol{l}_1)\,\left(\boldsymbol{l}_2^\top \boldsymbol{e}_1^2\right) &\approx \boldsymbol{0} \\
\boldsymbol{l}_1^\top \left(\boldsymbol{T}_1^{\Delta 12} \cdot \boldsymbol{x}_1\right)\,\boldsymbol{l}_2 - \left(\boldsymbol{l}_1^\top \Delta\boldsymbol{x}_1\right)\left(\boldsymbol{l}_2^\top \boldsymbol{e}_1^2\right) &\approx \boldsymbol{0} \\
\boldsymbol{x}_1 \wedge \left(\boldsymbol{T}_2^{\Delta 11} \cdot \boldsymbol{x}_2\right) \wedge \boldsymbol{x}_1 + (\boldsymbol{F}_{12}\,\boldsymbol{x}_2)\,(\boldsymbol{x}_1 \wedge \Delta\boldsymbol{x}_1)^\top + \left(\Delta\boldsymbol{x}_1^\top \boldsymbol{F}_{12}\,\boldsymbol{x}_2\right) \cdot [\,\boldsymbol{x}_1\,]_\times &\approx \boldsymbol{0} \\
\boldsymbol{l}_1^\top \left(\boldsymbol{T}_2^{\Delta 11} \wedge \boldsymbol{l}_2\right)\,\boldsymbol{l}_1 - \left((\boldsymbol{l}_1 \wedge \Delta\boldsymbol{l}_1)^\top \boldsymbol{F}_{12}\right) \wedge \boldsymbol{l}_2 &\approx \boldsymbol{0} \\
\boldsymbol{l}_1^\top \left(\boldsymbol{T}_2^{\Delta 11} \cdot \boldsymbol{x}_2\right)\,\boldsymbol{l}_1 - (\boldsymbol{l}_1 \wedge \Delta\boldsymbol{l}_1)^\top \boldsymbol{F}_{12}\,\boldsymbol{x}_2 &\approx \boldsymbol{0} \\
\boldsymbol{x}_1 \wedge \left(\left(\boldsymbol{T}_1^{23} \cdot \boldsymbol{x}_1\right) + \Delta\boldsymbol{x}_2\,\boldsymbol{e}_1^{3\,\top} - \boldsymbol{e}_1^2\,\Delta\boldsymbol{x}_3^\top\right) \wedge \boldsymbol{x}_1 &\approx \boldsymbol{0} \\
\boldsymbol{l}_1^\top \left(\boldsymbol{T}_1^{(23)} \wedge \boldsymbol{l}_1\right) \boldsymbol{l}_1 + (\boldsymbol{l}_1 \wedge \Delta\boldsymbol{l}_2)\left(\boldsymbol{l}_1^\top \boldsymbol{e}_1^3\right) - \left(\boldsymbol{l}_1^\top \boldsymbol{e}_1^2\right)(\boldsymbol{l}_1 \wedge \Delta\boldsymbol{l}_3) &\approx \boldsymbol{0} \\
\boldsymbol{l}_1^\top \left(\boldsymbol{T}_1^{(23)} \cdot \boldsymbol{x}_1\right) \boldsymbol{l}_1 - \left(\Delta\boldsymbol{l}_2^\top \boldsymbol{x}_1\right)\left(\boldsymbol{l}_1^\top \boldsymbol{e}_1^3\right) + \left(\boldsymbol{l}_1^\top \boldsymbol{e}_1^2\right)\left(\Delta\boldsymbol{l}_3^\top \boldsymbol{x}_1\right) &\approx \boldsymbol{0}
\end{aligned}
$$

- The last equation is Stein & Shashua's CVPR'97 ***tensor brightness constraint***.
- Again, it seems easier in practice to use projection matrices . . .

Acknowledgements Work supported by Esprit LTR project CUMULI.