Joint Feature Distributions for Image Correspondence

Bill Triggs, CNRS-INRIA, Grenoble, France — www.inrialpes.fr/movi/people/Triggs

Contributions

[1] ***Joint Feature Distributions (JFD's)*** — a general probabilistic framework for inter-image feature matching.
- JFD's are *joint probability distributions of positions of corresponding features* across several images.
- *Probabilistic conditioning* on observed feature positions gives conditional distributions for their correspondents in other images
  $\Rightarrow$ tight probabilistic correspondence-search regions.
- JFD's are *probabilistic characterizations of populations of training features*, not rigid geometric constraints — given suitable parametric forms, they can model geometric constraints, non-rigid motion, distortion...

[2] Simplest example: ***Gaussian-like JFD models*** that generalize & probabilize affine & projective multi-image matching constraints.
- Unlike matching constraints, ***Gaussian JFD correspondence models are stable & accurate even for degenerate geometries***
  — no model selection is needed (*c.f.* epipolar *vs.* homography for small translations, near-planar scenes...)
  — Gaussian JFD's can be viewed as algebraic variants of *Bayesian model averaging over geometric matching constraints*.
- Many other parametric forms are possible, *e.g.* for clustered data use mixtures of Gaussian subpopulation JFD's...

Conclusions

[1] The 2 image "epipolar" JFD is especially simple & effective — it should become a standard matching model.
[2] $\geq$ 3 image JFD models with the index structure of matching constraints are also useful. More general ones exist but are less efficient.
[3] For good results near epipoles use statistically-based error weighting — as in matching constraints, algebraic weighting underweights errors near epipoles.

Analogies between JFD's and Matching Tensors

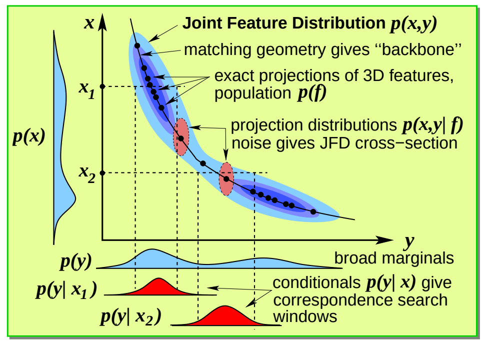
Entity	Matching Constraint Approach	Joint Distribution Approach
3D camera geometry	Camera projection mapping, matrices $\mathsf{P}_i: \mathbf{f} \to x_i = \mathsf{P}_i\,\mathbf{f}$	Conditional feature projection distributions $\mathrm{p}(\mathbf{x}_i \mid \mathbf{f})$
Image signature of camera geometry	Multi-image matching tensors $T_{ij\ldots k}$	Joint Feature Distributions $\mathrm{p}(\boldsymbol{x},\ldots,\boldsymbol{z})$
Inter-image feature transfer	Tensor based feature transfer $x \simeq T_{ij\ldots k}\cdot y\cdot\ldots\cdot z$	Conditional Feature Distributions $\mathrm{p}(\boldsymbol{x}\mid\boldsymbol{y},\ldots,\boldsymbol{z})$
Inter-image feature correspondence	Geometric matching constraints $T_{ij\ldots k}\cdot x\cdot\ldots\cdot z = 0$	Probability that features correspond, $\mathrm{p}(\boldsymbol{x},\ldots,\boldsymbol{z})$, or $\mathrm{p}(\boldsymbol{x}\mid\boldsymbol{y},\ldots,\boldsymbol{z})$
Scene reconstruction	Ray intersection, tensor-based reconstruction	Posterior 3D feature probability $\mathrm{p}(\mathbf{f}\mid\boldsymbol{x},\ldots,\boldsymbol{z})$

Joint Feature Distributions

- Given a population of 3D features $\mathbf{f}$ and probabilistic projection models $\mathrm{p}(\mathbf{x}_i\mid\mathbf{f})$ for their images $\mathbf{x}_1,\ldots,\mathbf{x}_m$, the ***Joint Feature Distribution (JFD)*** of the image features is:

$$\mathrm{p}(\mathbf{x}_1,\ldots,\mathbf{x}_m) \equiv \int \mathrm{p}(\mathbf{x}_1,\ldots,\mathbf{x}_m\mid\mathbf{f})\,\mathrm{p}(\mathbf{f})\,d\mathbf{f} = \int \mathrm{p}(\mathbf{x}_1\mid\mathbf{f})\ldots\mathrm{p}(\mathbf{x}_m\mid\mathbf{f})\,\mathrm{p}(\mathbf{f})\,d\mathbf{f}$$

- Even if the 3D population $\mathrm{p}(\mathbf{f})$ is broad, *the JFD remains highly correlated* — it still encodes precise location information from the feature projections $\mathrm{p}(\mathbf{x}_i\mid\mathbf{f})$.
- ***The JFD can be estimated and used as a matching tool*** — its ***conditional distributions (CFD's)*** $\mathrm{p}(\mathbf{x}_1\mid\mathbf{x}_2,\ldots,\mathbf{x}_m)$ define probabilistic correspondence search regions.



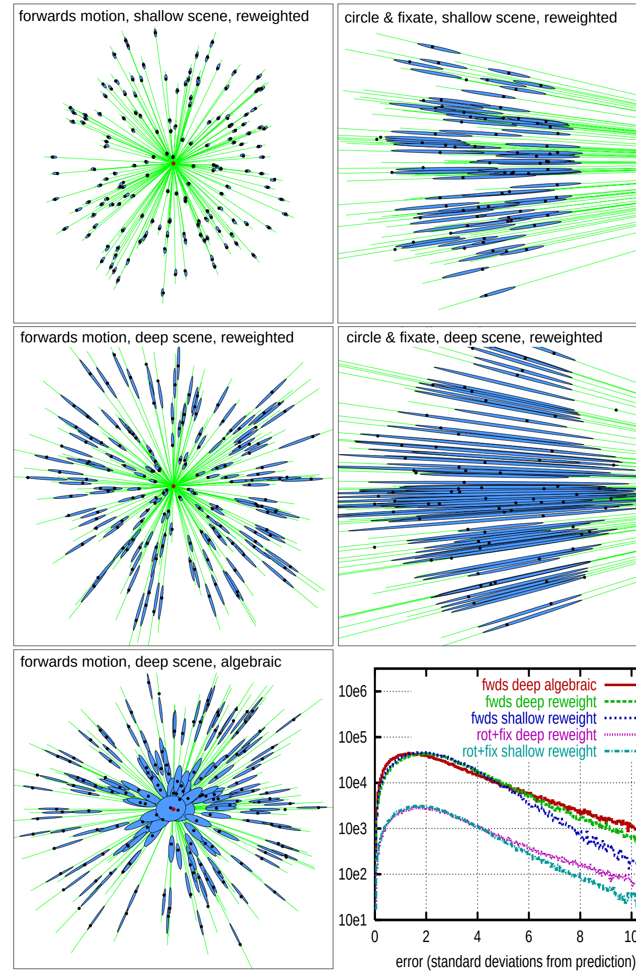
Estimation algorithm — "epipolar" JFD model

- Uses homogeneous representation of Gaussians — see paper for details.

[1] As in linear fundamental matrix estimation, build data matrix M with columns $(1,x,y,x',y',xx',xy',yx',yy')^\top$ — the tensor product of the homogeneous coordinates of the training correspondences.
[2] Regularize and invert 9×9 ***homogeneous scatter matrix*** $\mathsf{M}\,\mathsf{M}^\top$ to get the ***homogeneous information*** (inverse covariance) of the Gaussian JFD model: $W = (\mathsf{M}\,\mathsf{M}^\top + \epsilon)^{-1}$.
[3] To condition on an observed feature $\boldsymbol{x}$, treat W as a tensor $W_{AB\,A'B'}$ and contract against $\boldsymbol{x}\,\boldsymbol{x}^\top$ to get the 3×3 homogeneous information A of the conditional distribution for $\boldsymbol{x}'$, the correspondent of $\boldsymbol{x}$: $\mathsf{A}_{A'B'} \equiv W_{AB\,A'B'}\,\boldsymbol{x}^A\,\boldsymbol{x}^B$
[4] For more accurate results near epipoles, use a more statistically accurate error weighting: $\mathsf{A} \to \mu\mathsf{A}$ with $\mu \sim 1/(\mathsf{A}_{11}+\mathsf{A}_{22})$ (see paper).

Other affine & perspective Gaussian JFD models are similar:
[1] Combine training coordinates into direct sum (affine case) or tensor product (perspective case) "joint image" vectors, and estimate a Gaussian-like scatter model for the vectors.
[2] Condition on given measurements to find Gaussian-like search regions for their correspondents in other images — conditioning uses Schur complement (affine model), tensor contraction (perspective model).
- To probabilize "lines-through-point" matching constraints (homographic, trifocal...), use ***dual scatter matrices*** to represent "uniform distributions of lines" through the given point.

Numerical Experiments — Epipolar JFD search region



Tensor Joint Image Representation

- A new way to view multi-image geometric matching constraints, gives the theoretical foundation for the tensored-Gaussian JFD approach.
- In tensor product representations, matching constraints become linear — represent & image entities by their *Veronese* & *Segré* varieties from algebraic geometry.

Acknowledgements Partly supported by European Union FET-Open project VIBES.