

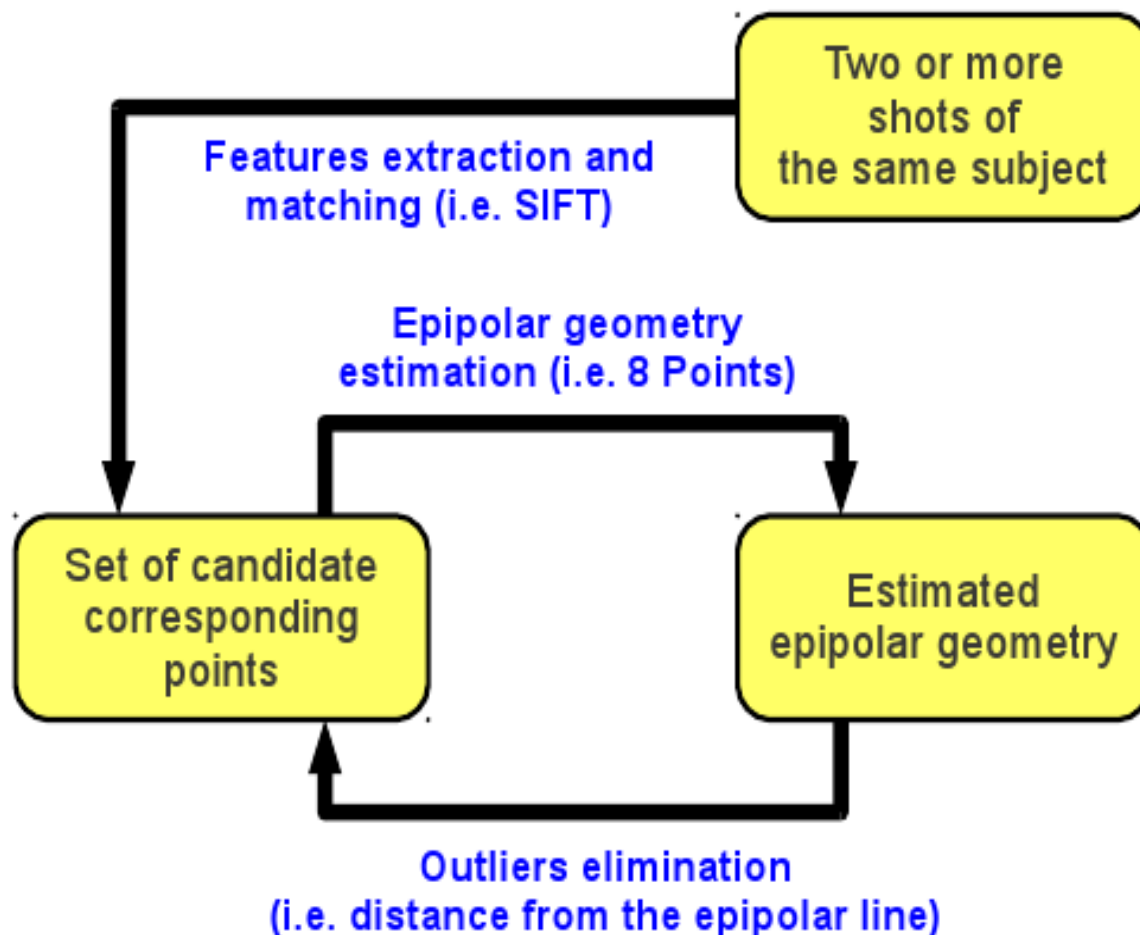


A Game-Theoretic Approach to the Enforcement of Global Consistency in Multi-View Feature Matching

Emanuele Rodolà, Andrea Albarelli, and Andrea Torsello



Inliers and Bundle Adjustment





Typical matching strategies are based on the use of local information such as feature descriptors.

Global coherence checks are only introduced after a first estimation. (filtering)

Filtering approaches are not very robust w.r.t. Outliers (or structured noise)



Game theoretic inlier selection

We use a game-theoretic approach to drive the selection of correspondences that satisfy a global compatibility criterion

- Correspondences receive payoffs from other correspondences according to their fit to a global geometric constraint
- Evolution selects a subset of correspondences that are all geometrically consistent

This happens by modeling each correspondence as a strategy in a non cooperative game and by letting correspondences play a game where the payoff among two strategies is given by their compatibility.

The rationale of this approach is that groups that are coherent w.r.t. a global geometric constraint have large mutual payoffs and thus are more likely to survive.



Game theoretic inlier selection

A large set of matching features (i.e. strategies) is selected with a standard technique (and some outliers are included)

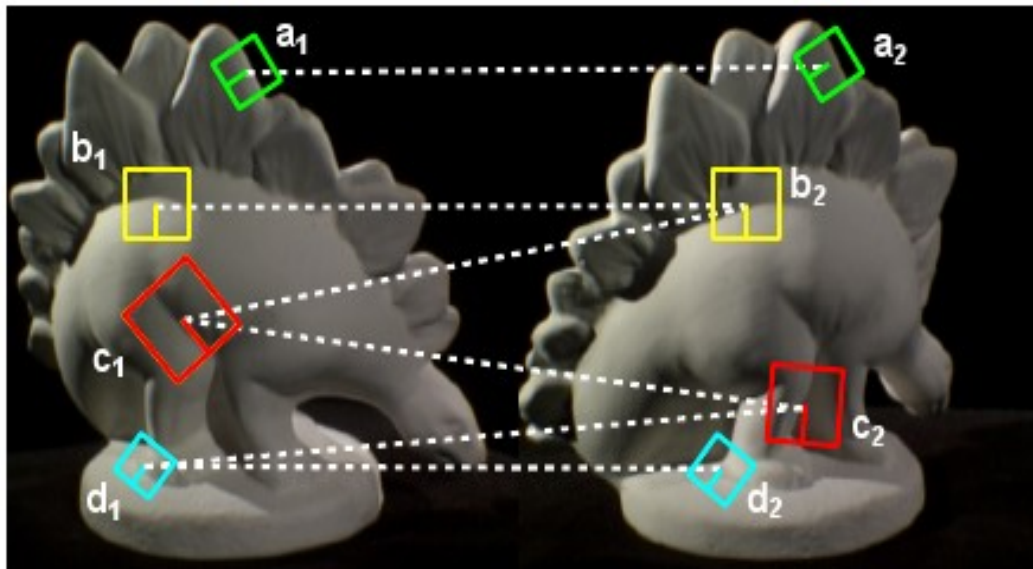
An initial population vector x (of size equal to the number of candidates) is initialized to some point of the standard simplex;

A payoff matrix C is defined for each pair of strategies according to the consistency model

The population is evolved through the replicator equation:

$$\vec{x}_i(t+1) = x_i(t) \frac{(C\vec{x}(t))_i}{\vec{x}(t)^T C \vec{x}(t)}$$

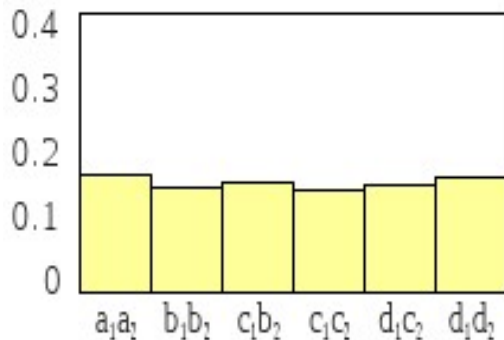
An example of evolution



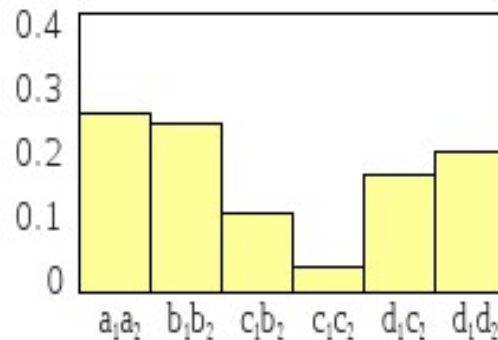
π

	a_1a_2	b_1b_2	c_1b_2	c_1c_2	d_1c_2	d_1d_2
a_1a_2	0	1	0.1	0.1	0.7	0.9
b_1b_2	1	0	0	0.1	0.7	0.9
c_1b_2	0.1	0	0	0	0.6	0.4
c_1c_2	0.1	0.1	0	0	0	0.1
d_1c_2	0.7	0.7	0.6	0	0	0
d_1d_2	0.9	0.9	0.4	0.1	0	0

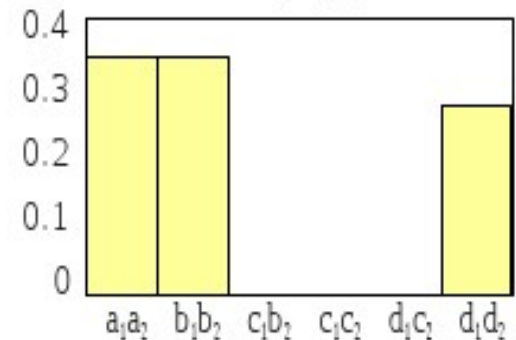
$T=0$



$T=1$



$T=10$





Geometric Consistency Models

We compare **two** geometric consistency approaches based on modern point descriptors (SIFT, SURF)

- A semi-local approach that that extract local patches consistent with a common affine transformation
- A global approach that tries to impose consistency with a common 3D rigid transformation



Semi-local Affine Model

We need to define a compatibility w.r.t. affinity between pairs of candidate correspondences.

We use the orientation and scale information in the feature descriptors to infer an affine transformation between the corresponding features

Correspondence imply transformation

Two correspondences are compatible if they define similar transformations

i.e., compatible correspondences must be subject to the same affine transformation of the local patches



Compatibility in the Affine Model

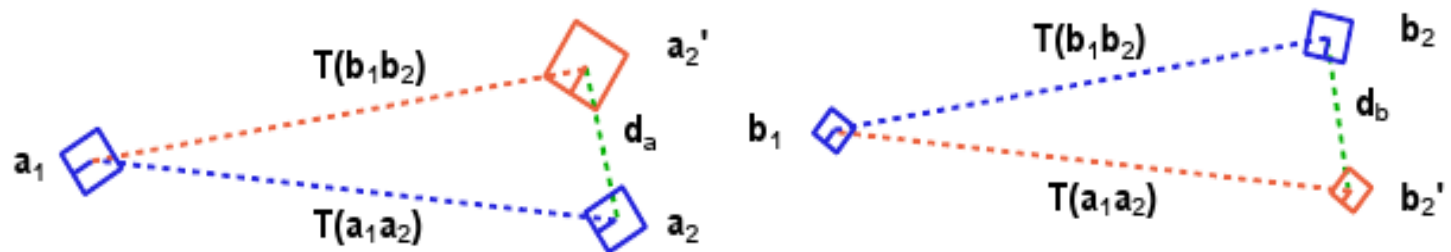
To this extent we create two virtual points:

$$a'_2 = T(b_1, b_2)a_1$$

$$b'_2 = T(a_1, a_2)b_1$$

And we define the affine compatibility as:

$$\Pi((a_1, a_2), (b_1, b_2)) = e^{-\lambda \max(|a_2 - a'_2|, |b_2 - b'_2|)}$$



Penalize re-projection error between the two affine transformations



Limits of the Affine Model

The affine assumption works only locally

The approach extracts only local patches of corresponding points

Repeat the game extracting new sets of correspondences
(image based peel off strategy)



Global Consistency Model

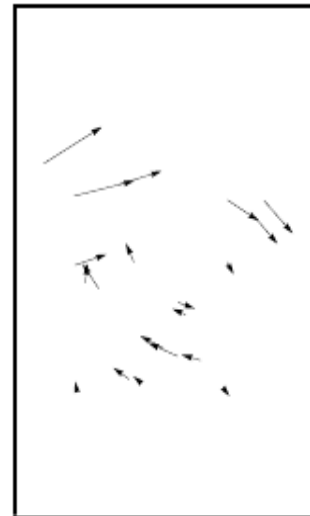
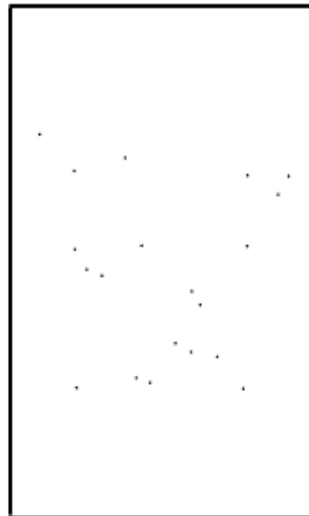
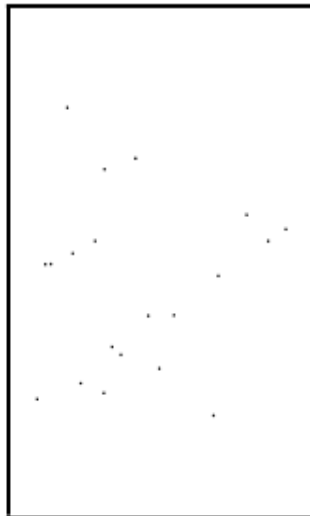
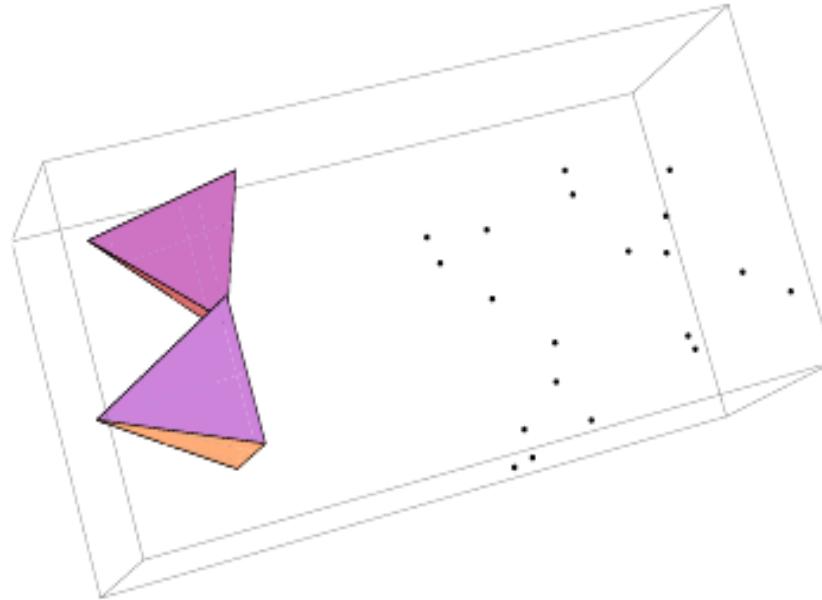
Impose consistency with a common 3D rigid transformation

Impose rigidity by requiring conservation of distances between underlying 3D points

Use local scale and orientation information provided by modern feature descriptors (SIFT, SURF) to constrain the transformations



Uniform 3D motion non-uniform 2D motion

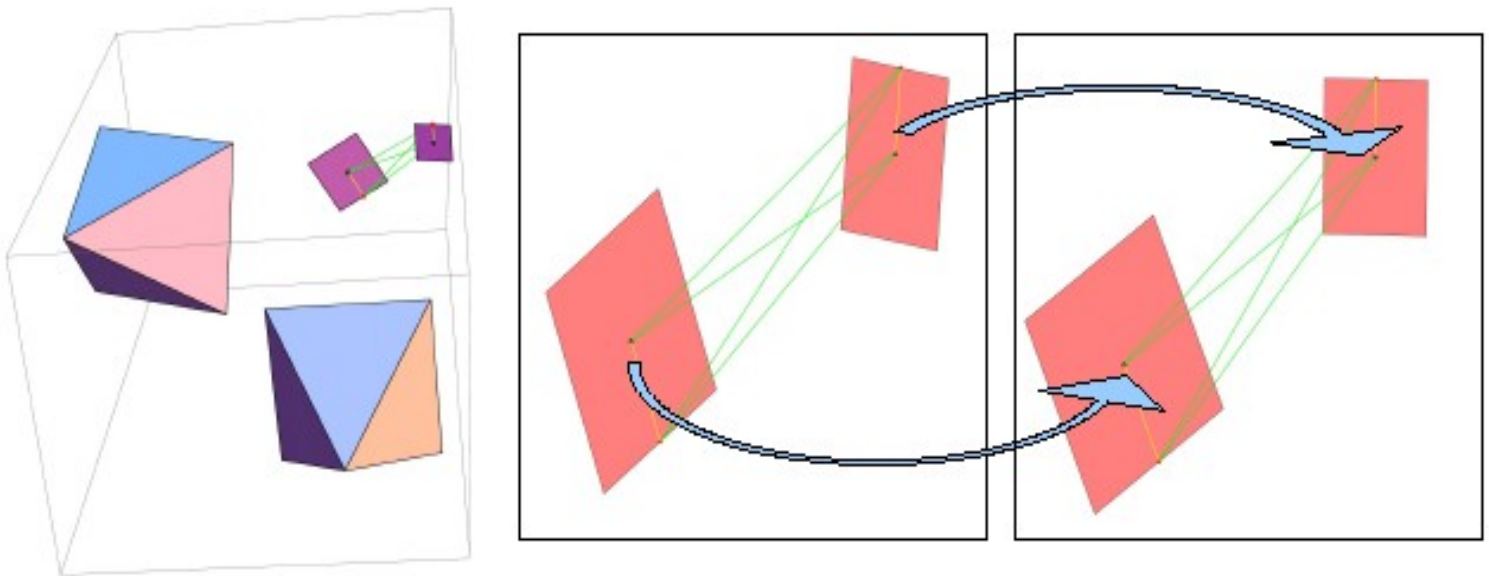




Virtual points

Scale and orientation offer depth information and a second virtual point.

The conservation of the distances in green enforces consistency with a 3D rigid transformation.





Points and Rigid Transformations

4 points are not enough to fully constrain the transformation

Free parameter a : ratio of real scale of 3D patches

Assume features have similar scales and optimize over available variation

In our experiments $[0.5, 2]$



Global 3D Model

Corresponding points

$$p_1^1 = \frac{1}{s_1^1} \begin{pmatrix} u_1^1 \\ v_1^1 \\ f \end{pmatrix}, p_2^1 = \frac{a}{s_2^1} \begin{pmatrix} u_2^1 \\ v_2^1 \\ f \end{pmatrix}, p_1^2 = \frac{1}{s_1^2} \begin{pmatrix} u_1^2 \\ v_1^2 \\ f \end{pmatrix}, p_2^2 = \frac{a}{s_2^2} \begin{pmatrix} u_2^2 \\ v_2^2 \\ f \end{pmatrix}$$

Virtual points

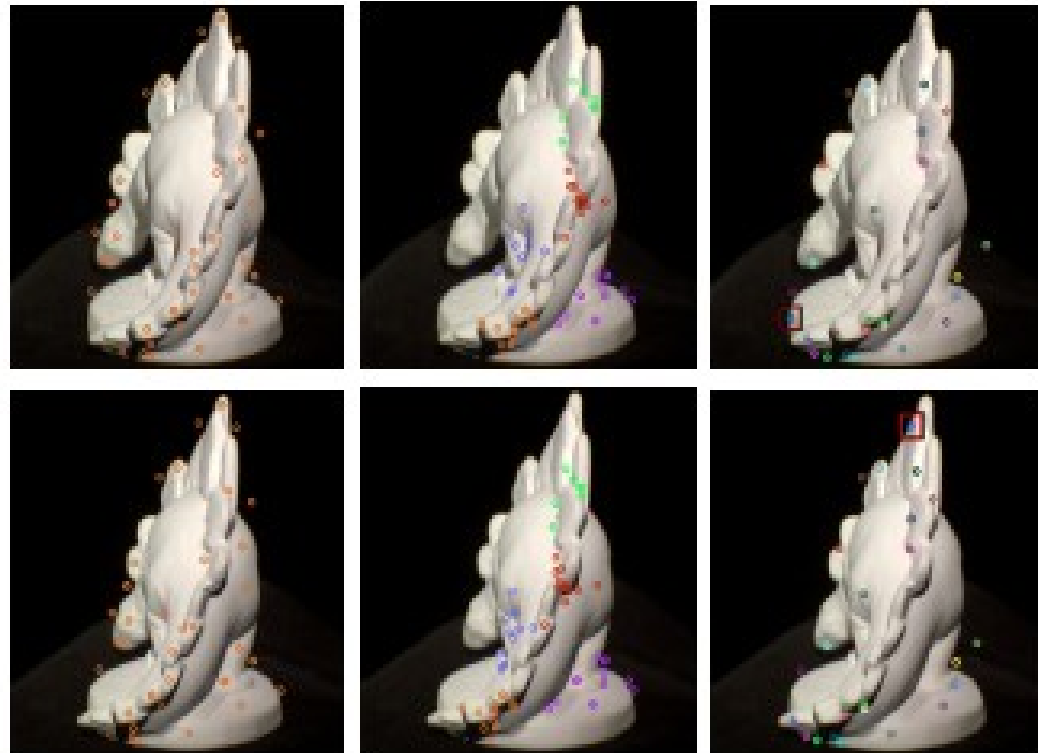
$$q_1^1 = p_1^1 + \begin{pmatrix} \cos \theta_1^1 \\ \sin \theta_1^1 \\ 0 \end{pmatrix}, q_2^1 = p_2^1 + a \begin{pmatrix} \cos \theta_2^1 \\ \sin \theta_2^1 \\ 0 \end{pmatrix}, q_1^2 = p_1^2 + \begin{pmatrix} \cos \theta_1^2 \\ \sin \theta_1^2 \\ 0 \end{pmatrix}, q_2^2 = p_2^2 + a \begin{pmatrix} \cos \theta_2^2 \\ \sin \theta_2^2 \\ 0 \end{pmatrix}$$

Dissimilarity of matches

$$d(m_1, m_2, a) = (\|p_1^1 - p_2^1\|^2 - \|p_1^2 - p_2^2\|^2)^2 + (\|p_1^1 - q_2^1\|^2 - \|p_1^2 - q_2^2\|^2)^2 + \\ (\|q_1^1 - p_2^1\|^2 - \|q_1^2 - p_2^2\|^2)^2 + (\|q_1^1 - q_2^1\|^2 - \|q_1^2 - q_2^2\|^2)^2$$



Experimental Results

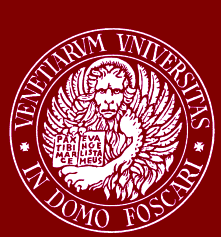


GT-3Drigid

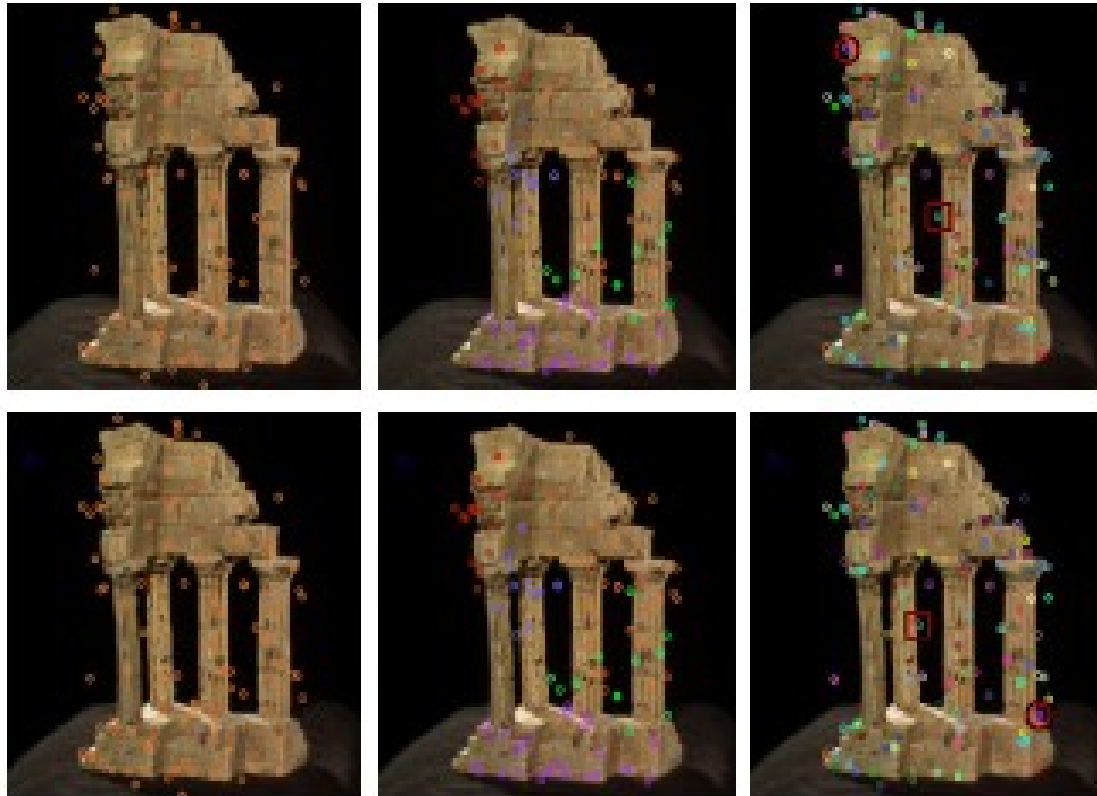
GT-2Daffine

Bundler

Matches	262.5 ± 61.4	271.1 ± 64.2	172.4 ± 79.5
$\Delta\alpha$	0.0668 ± 0.0777	0.0497 ± 0.0810	0.0767 ± 0.1172
$\Delta\gamma$	0.4393 ± 0.4963	0.3184 ± 0.3247	0.6912 ± 0.8793



Experimental Results



GT-3Drigid

GT-2Daffine

Bundler

Mathces 535.7 ± 38.7

564.3 ± 37.2

349.3 ± 36.2

$\Delta\alpha$ 0.1326 ± 0.0399 0.0989 ± 0.0224 0.1414 ± 0.0215

$\Delta\gamma$ 0.0809 ± 0.0144 0.0792 ± 0.0091 0.0850 ± 0.0065



Conclusions

Presented a game theoretic approach to enforce consistency in feature matching

Robustness is achieved by enforcing global geometric consistency in a pairwise setting

Only highly compatible matches are enforced while incompatible correspondences are driven to extinction.

Experimental comparisons show the ability of our approach to obtain very accurate estimates