Geometric Registration for Deformable Shapes

3.4 Probabilistic Techniques

RANSAC · Forward Search · Efficiency Guarantees

31st Annual Conference of the European Association for Computer Graphics
eurographics 2010

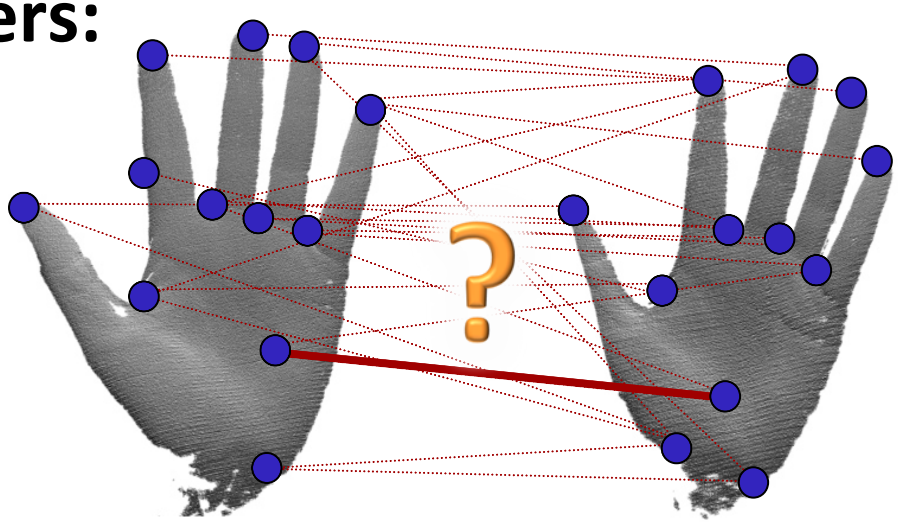
Ransac and Forward Search

The Basic Idea

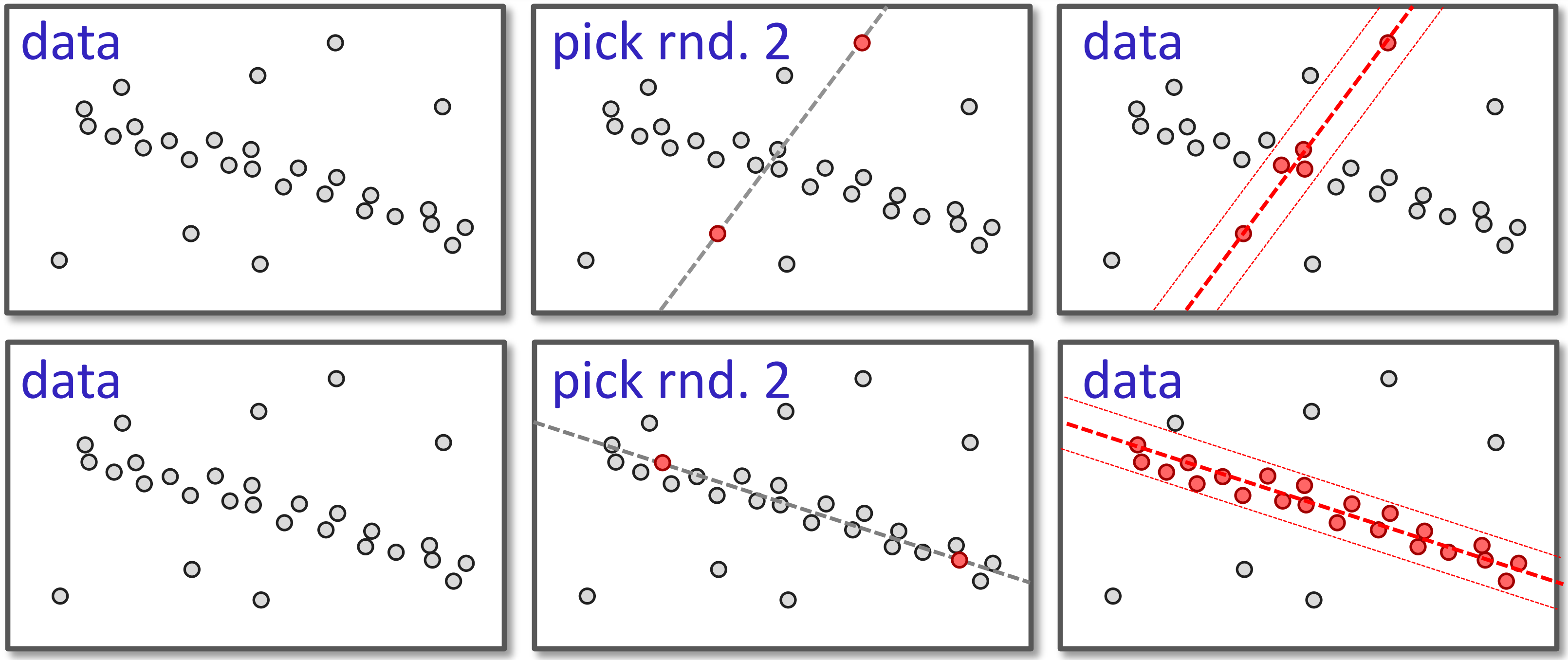
Random Sampling Algorithms

Estimation subject to outliers:

- We have candidate correspondences
- But most of them are bad
- Standard vision problem
- Standard tools: Ransac & forward search



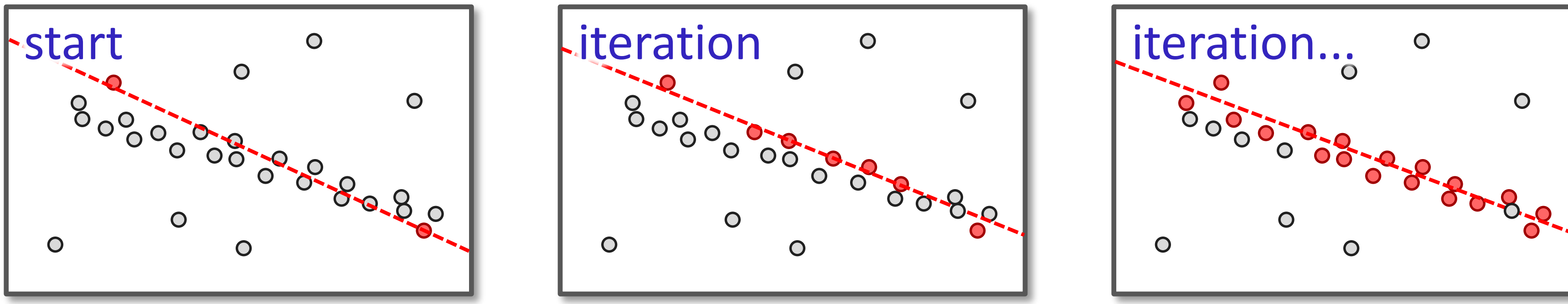
RANSAC

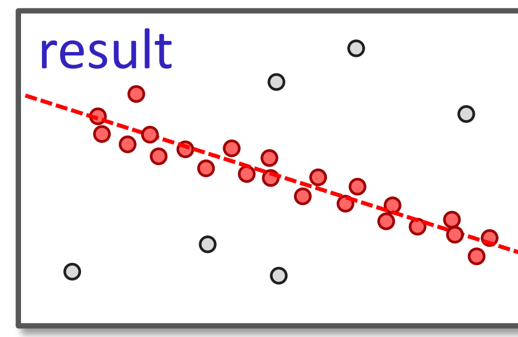


„Standard" RANSAC line fitting example:

- Randomly pick two points
- Verify how many others fit
- Repeat many times and pick the best one (most matches)

Forward Search





Forward Search:

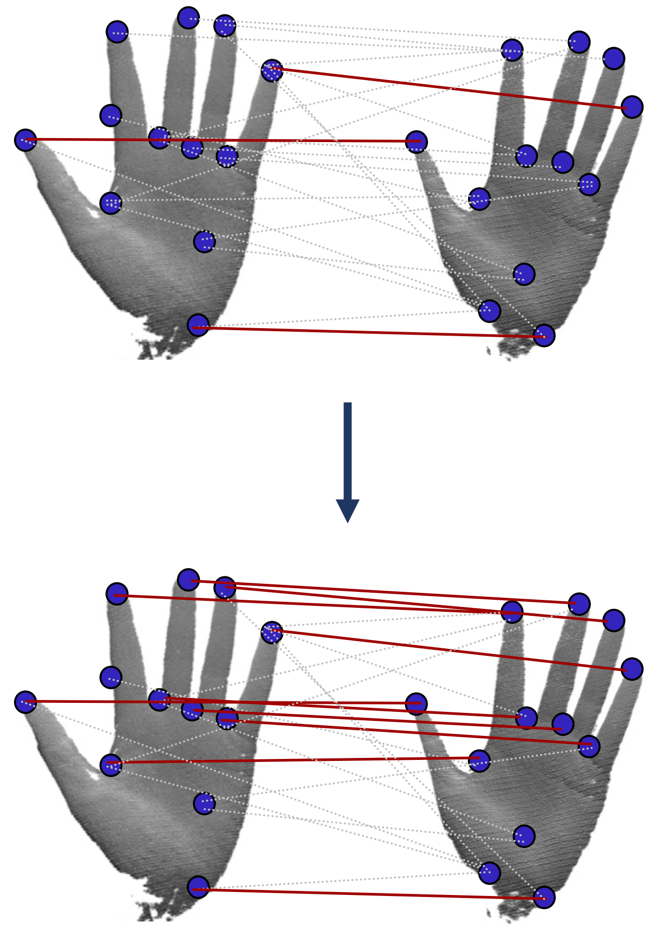
- Ransac variant
- Like ransac,
  but refine model by „growing“
- Pick best match, then recalculate
- Repeat until threshold is reached

Ransac-Based Correspondence Estimation

RANSAC/FWS Algorithm

Idea

- Starting correspondence
- Add more that are consistent
  - Preserve intrinsic distances
- Importance sampling algorithm



Advantages

- Efficient (small initial set)
- General (arbitrary criteria)

Ransac/FWS Details

Algorithm: Simple Idea

- Select correspondences with probability proportional to their plausibility
- First correspondence: Descriptors
- Second: Preserve distance (distribution peaks)
- Third: Preserve distance (even fewer choices)

  ...

- Rapidly becomes deterministic
- Repeat multiple times (typ.: 100x)
  - Choose the largest solution (larges #correspondences)

Ransac/FWS Details

Provably Efficient:

- Theoretically efficient (details later)
- Faster in practice (using descriptors)

Flexible:

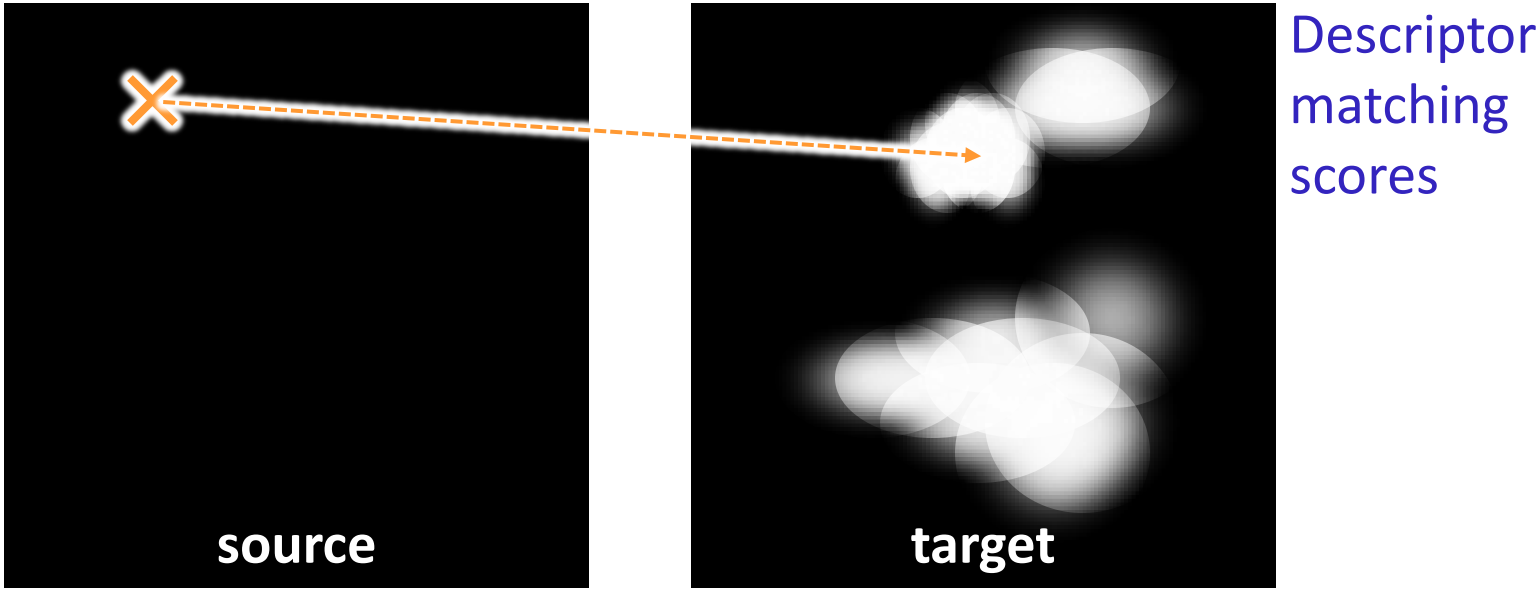
- In later iterations (> 3 correspondences), allow for outlier geodesics
- Can handle topological noise

Foreward Search Algorithm

Forward Search

- Add correspondences incrementally
- Compute match probabilities given the information already decided on
- Iterate until no more matches can found that meet a certain error threshold
- Outer Loop:
  - Iterate the algorithm with random choices
  - Pick the best (i.e., largest) solution

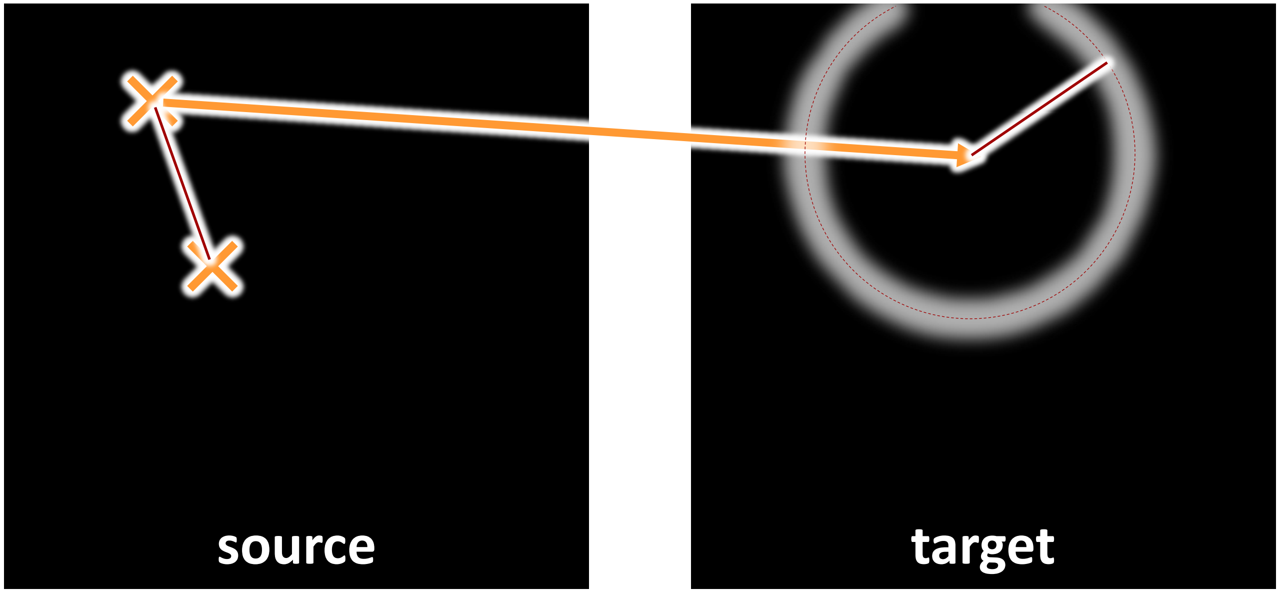
Foreward Search Algorithm



Step 1:

- Start with one correspondence
  - Target side importance sampling:
    prefer good descriptor matches
  - Optional source side imp. sampl: prefer unique descriptors

Foreward Search Algorithm

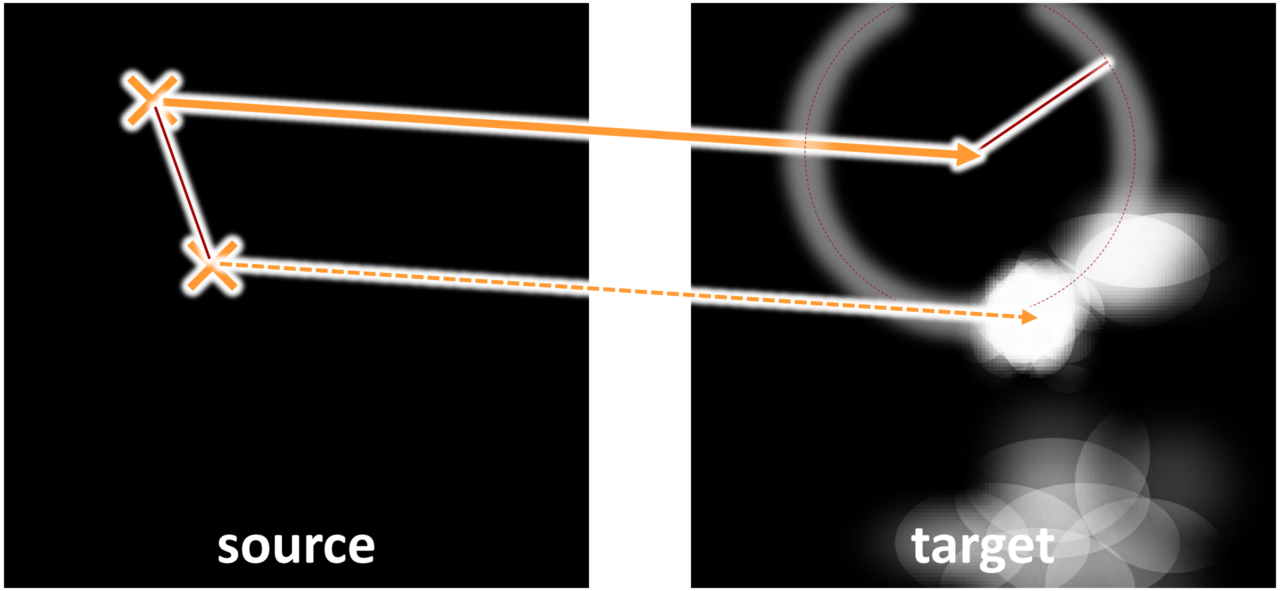


posterior
(distance)

Step 2:

- Compute „posterior“ incorporating geodesic distance
  - Target side importance sampling:
    sample according to descriptor match × distance score
  - Again: optional source side imp. sampl: prefer unique descriptors

Foreward Search Algorithm

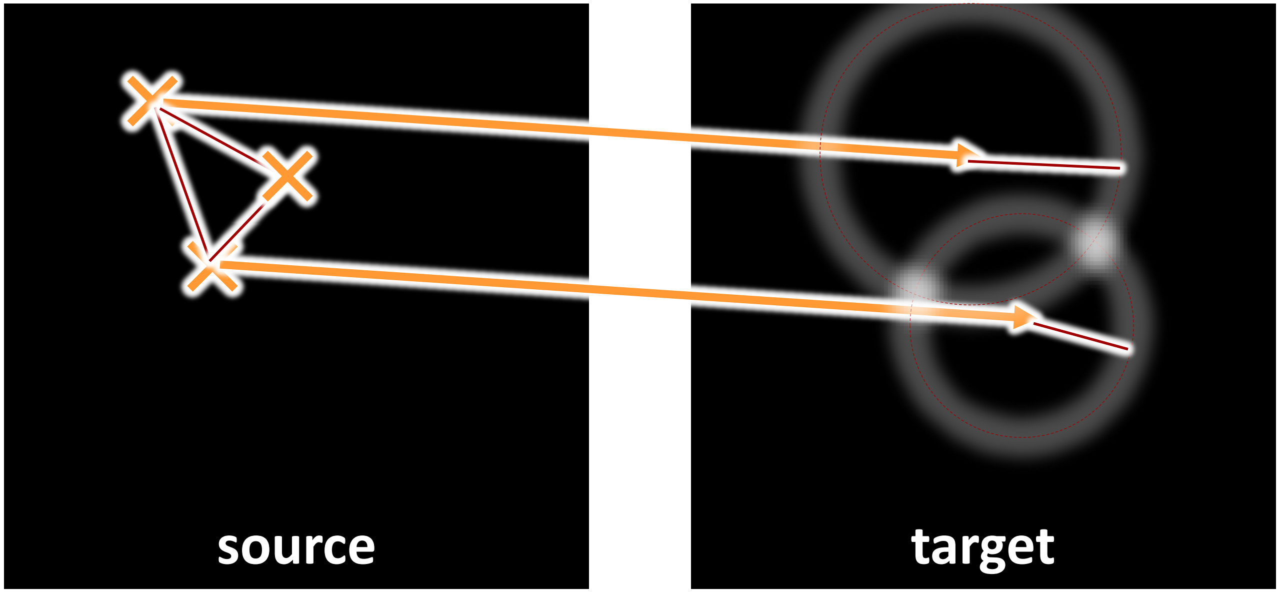


posterior (distance & descriptors)

Step 2:

- Compute „posterior“ incorporating geodesic distance
  - Target side importance sampling:
    sample according to descriptor match × distance score
  - Again: optional source side imp. sampl: prefer unique descriptors

Foreward Search Algorithm

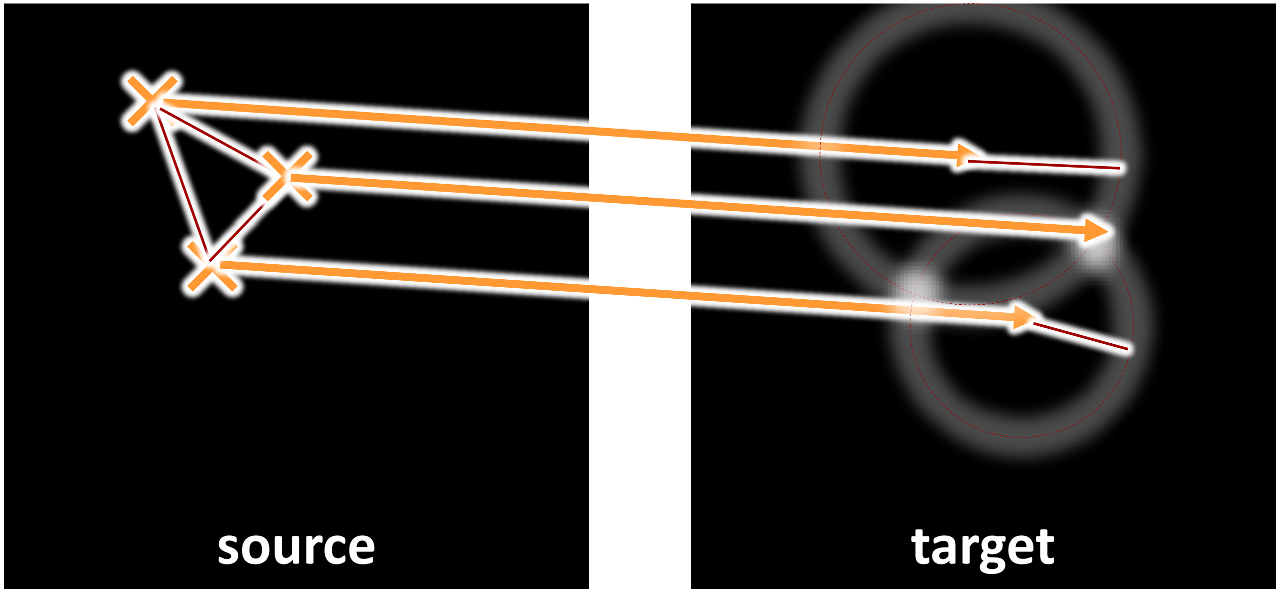


posterior (distance & descriptors)

Step 3:

- Same as step 2, continue sampling...

Foreward Search Algorithm



posterior (distance & descriptors)

Step 3:

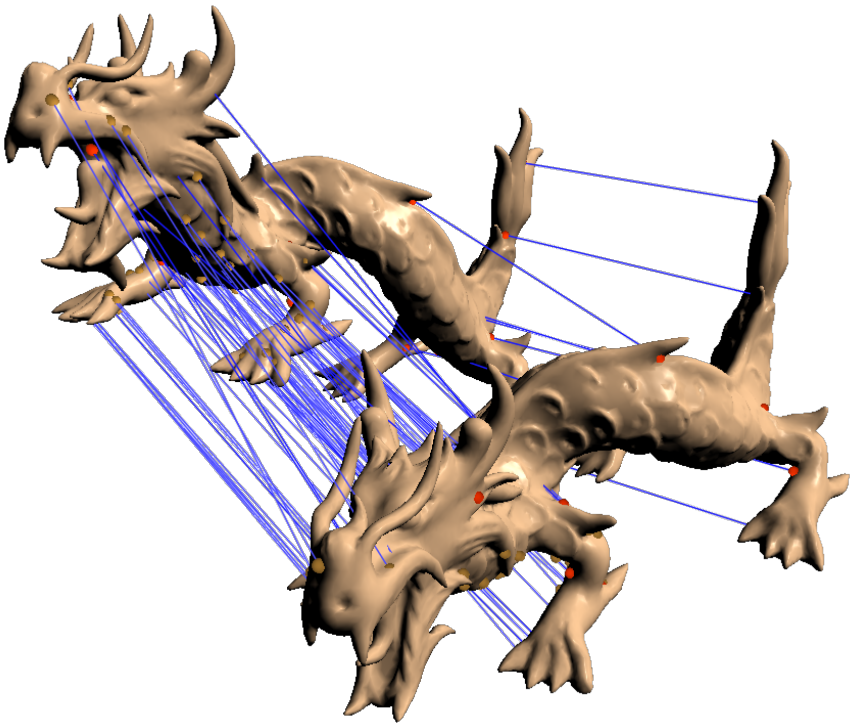
- Same as step 2, continue sampling...

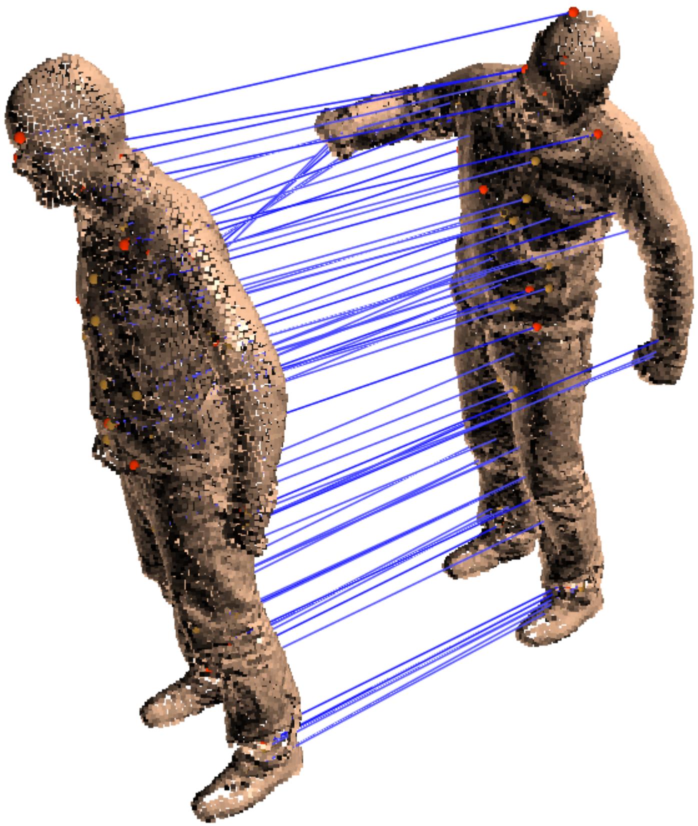
Another View

Landmark Coordinates

- Distance to already established points give a charting of the manifold

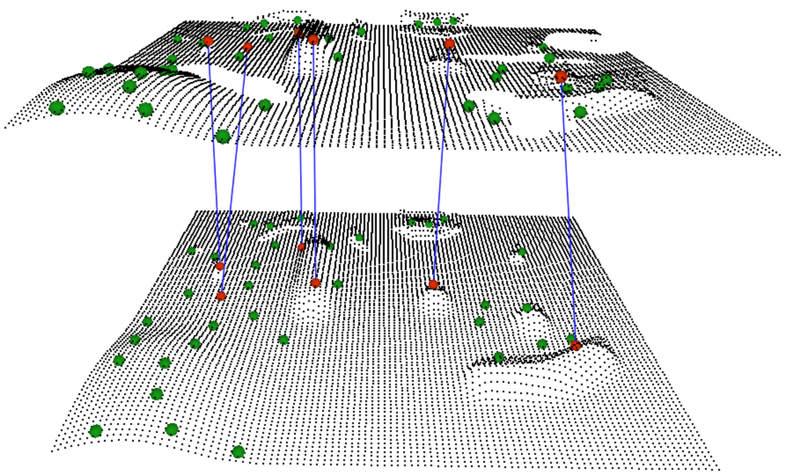
Results



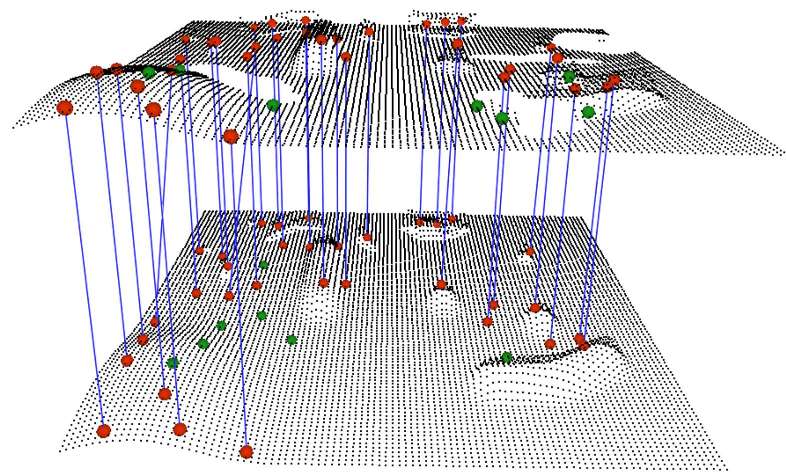


[data sets: Stanford 3D Scanning Repository / Carsten Stoll]

Results: Topological Noise



Spectral Quadratic Assignment
[Leordeanu et al. 05]



Ransac Algorithm
[Tevs et al. 09]

Complexity

How expensive is all of this?

Cost analysis:

- How many rounds of sampling are necessary?

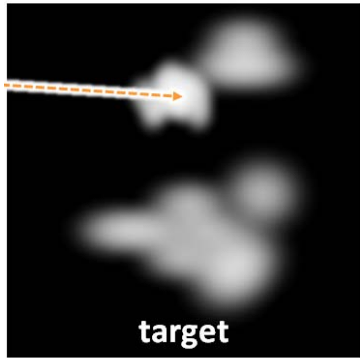
Constraints [Lipman et al. 2009]:

- Assume disc or sphere topology
- An isometric mapping is in particular a conformal mapping
- A conformal mapping is determined by 3 point-to-point correspondences

How expensive is it..?

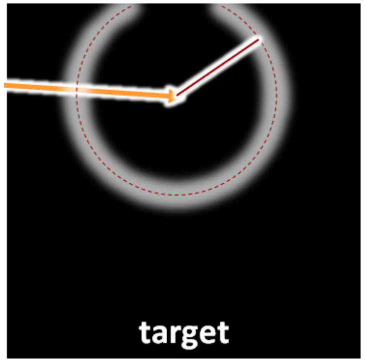
First correspondence:

- Worst case: n trials (n feature points)
- In practice: $k \ll n$ good descriptor matches (typically $k \approx$ 5-20)



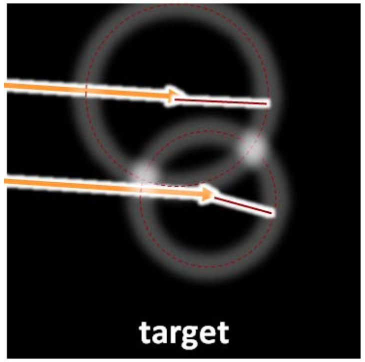
Second correspondence:

- Worst case: n trials, expected: $\sqrt{n}$ trials
- In practice: very few (due to descriptor matching, maybe 1-3)



Last match:

- At most two matches



Costs...

Overall costs:

- Worst case: $O(n^2)$ matches to explore
- Typical: $O(n^{1.5})$ matches to explore

Randomization:

- Exploring m items costs expected $O(m \log m)$ trials
- Worst case bound of $O(n^2 \log n)$ trials
- Asymptotically sharp: $O(c)$-times more trials for shrinking failure probability to $O(\exp(-c^2))$

Costs...

Surface discretization:

- Assume ε-sampling of the manifold (no features): $O(\varepsilon^{-2})$ sample points
- Worst case $O(\varepsilon^{-4} \log \varepsilon^{-1})$ sample correspondences for finding a match with accuracy ε.
- Expected: $O(\varepsilon^{-3} \log \varepsilon^{-1})$.

In practice:

- Importance sampling by descriptors is very effective
- Typically: Good results after 100 iterations

General Case

Numerical errors:

- Noise surfaces, imprecise features: reflected in probability maps (we know how little we might know)

Topological noise:

- Use robust constraint potentials
- For example: account for 5 best matches only

Topologically complex cases:

- No analysis beyond disc/spherical topology
- However: the algorithm will work in the general case (potentially, at additional costs)