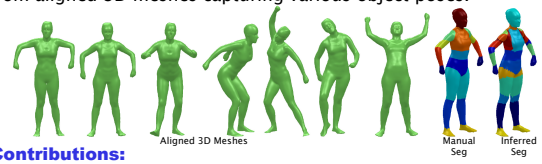
From Deformations to Parts: Motion-based Segmentation of 3D Objects

Soumya Ghosh[1], Erik B. Sudderth[1], Matthew Loper[2] and Michael J. Black[2]

[1] Department of Computer Science, Brown University, [2] Perceiving Systems Department, Max Planck Institute for Intelligent Systems.

Segmentation of Articulated 3D Objects

Goal: To discover distinctly moving parts of an articulated object from aligned 3D meshes capturing various object poses.

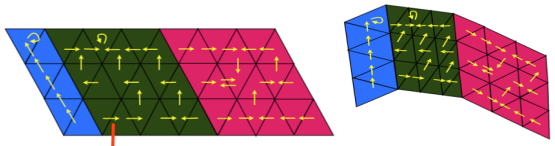


Contributions:

- Modeling: We propose a model that,
  - Discovers a potentially unbounded number of parts,
  - Guarantees spatially connected segmentation,
  - Seamlessly deals with object instances having varying 3D shapes,
  - Is robust to moderate alignment noise.
- Results: State-of-the-art prediction of 3D body deformations.

Distributions over Mesh Partitions

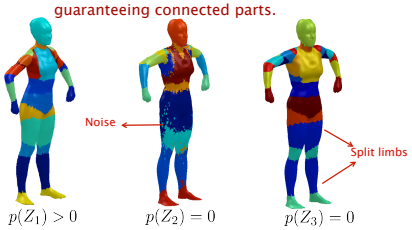
Distance dependent Chinese restaurant process



$$p(c_m = n \mid D, f, \alpha) \propto \begin{cases} f(d_{mn}) & m \neq n, \\ \alpha & m = n. \end{cases}$$

$f(d) = \mathbf{1}[d \leq 1]$ A "window" decay function of width 1 restricts mesh faces to link to immediately adjacent faces, guaranteeing connected parts.

Unlike a standard CRP prior, only partitions with contiguous parts have non-zero prior probability.

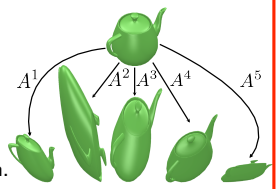


Distributions over Affine Transformations

$$\Sigma \sim \mathcal{IW}(n_0, S_0)$$
$$A \mid \Sigma \sim \mathcal{MN}(M, \Sigma, K)$$

where $A \in \mathbb{R}^{3\times 4}$ is an affine transformation.

- A matrix normal distribution with an inverse Wishart prior over the covariance.
- The mean is set to the identity transformation.
- S_0 is set to match the scale of the translational parameters with that of the mesh coordinates.

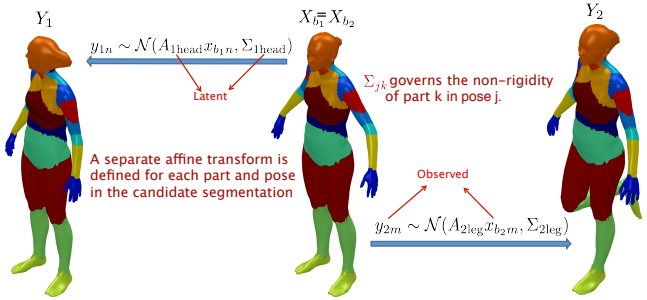


$A^1 \ldots A^5 \sim \mathcal{MNIW}(.)$

Samples from distribution over Affine transformations.

Model and Inference

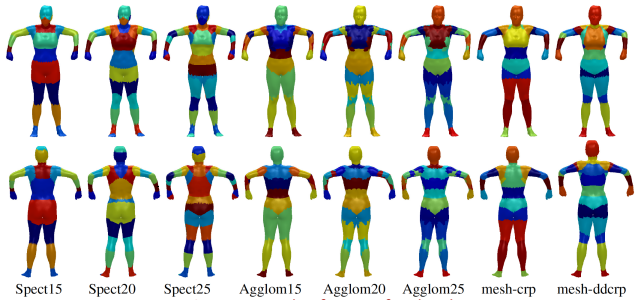
- For each mesh face n, sample an associated link $c_n \sim \mathrm{ddCRP}(\alpha, f, D)$
- For each pose j and each part k, sample an affine transformation A_{jk} and residual noise covariance Σ_{jk} from the MNIW distribution.
- Sample the observed location of each pose triangle relative to its corresponding reference triangle, $y_{jn} \sim \mathcal{N}(A_{j z_n} x_{b_j n}, \Sigma_{j z_n})$



Inference

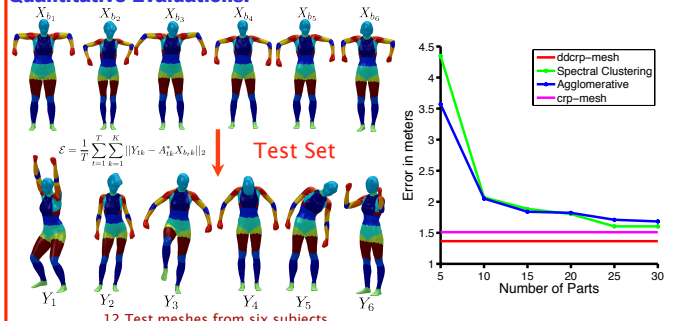
- Collapsed Gibbs sampler that analytically marginalizes latent affine transformations and noise variables.
- Likelihood evaluation independently marginalizes affine and noise variables for each pose: computation scales linearly with number of poses, parallelizable.
- Large moves in the space of candidate partitions through split and merge.

Results: Segmentation of Human Bodies



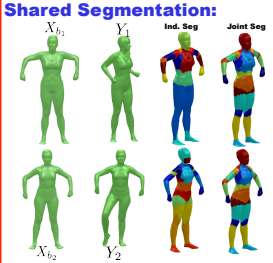
56 Training meshes from two female subjects.

Quantitative Evaluations:



12 Test meshes from six subjects.

Shared Segmentation:



Segmentation of Synthetic Meshes:

http://vosca.cs.technion.ac.il/