Data-driven 3D shape analysis and synthesis

Evangelos Kalogerakis
UMass Amherst
3D shapes for computer-aided design

Architecture

Interior design
3D shapes for information visualization

Geo-visualization

Scientific visualization
3D shapes for digital entertainment

Video games
Digital representations of 3D shapes

Polygon mesh
Digital representations of 3D shapes
Digitizing our imagination

Professional 3D modeling tools
[Autodesk Maya]
Digitizing our imagination

Computer-Aided Design tools
[Catia]
Digitizing our imagination

General-Purpose Modeling tools
[Google Sketch-up]
3D shape repositories

[Google 3D Warehouse]
Shape understanding

Animal, quadruped, horse, running horse
Why shape understanding?

Shape categorization

Sailing Ship, Galleon

Sailing ship, Yawl

Military ship, Frigate
Why shape understanding?

3D Modeling

Chaudhuri, Kalogerakis, Guibas, Koltun, SIGGRAPH 2011
Why shape understanding?

Shape synthesis
Why shape understanding?

Shape synthesis

Kalogerakis, Chaudhuri, Koller, Koltun, SIGGRAPH 2012
Why shape understanding?

Artistic rendering

Kalogerakis, Nowrouzehahrai, Breslav, Hertzmann, TOG 2012
Why shape understanding?

Texturing

Kalogerakis, Hertzmann, Singh, SIGGRAPH 2010
Why shape understanding?

Character Animation

Simari, Nowrouzezahrai, Kalogerakis, Singh, SGP 2009
How can we perform shape understanding?

It is extremely hard to perform shape understanding with a set of deterministic, manually specified rules!
Key idea: probabilistic models for shapes

Define a probability distribution over high-level shape attributes given geometry (discriminative approach), or both (generative approach).

Learn this distribution by combining training data and expert knowledge.

Efficiently infer unknown attributes given observed evidence.
First part of my talk:
Learning 3D shape segmentation and labeling

Contributions:
Segmentation and labeling of parts with a prob. discriminative model
Major improvements over prior work
Data-driven, learnt from examples

Kalogerakis, Hertzmann, Singh, SIGGRAPH 2010
Second part of my talk:
A generative model of shapes

Contributions:
Learns structural variability in 3D shapes

Automatic shape synthesis in complex domains (airplanes, ships, furniture, game characters)

Kalogerakis, Chaudhuri, Koller, Koltun, SIGGRAPH 2012
Outline

1. Learning 3D shape segmentation and labeling [Kalogerakis et al., SIGGRAPH 2010]

2. A generative model of shapes

3. Other ML applications to graphics and vision

4. Future work
Goal: shape segmentation and labeling

Input Shape → Labeled Shape

Training Meshes:
- Head
- Neck
- Torso
- Leg
- Tail
- Ear
Related work: mesh segmentation

- Shape Diameter [Shapira et al. 2010]
- Randomized Cuts [Golovinskiy and Funkhouser 2008]
- Random Walks [Lai et al. 2008]
- Fitting Primitives [Attene et al. 2006]
Is human-level shape analysis possible without using prior knowledge?

[X. Chen et al. SIGGRAPH 2009]
Must we hand-tune algorithms for each type of shape?

[X. Chen et al. SIGGRAPH 2009]
Related work: image segmentation and labeling

Textonboost
[Shotton et al. ECCV 2006]
Labeling problem statement
Labeling problem statement

\[ C = \{ \text{head, neck, torso, leg, tail, ear} \} \]
Labeling problem statement

\[ c_1, c_2, c_3 \in C \]

\[ C = \{ \text{head, neck, torso, leg, tail, ear} \} \]
Feature vector $\mathbf{x}$

- Surface curvature
- PCA-based descriptors
- Shape diameter
- Average geodesic distances
- 3D contextual features
- Localized descriptors of global shape
Labeling problem statement

\[ c_1, c_2, c_3 \in C \]
\[ C = \{ \text{head, neck, torso, leg, tail, ear} \} \]
Labeling problem statement

\[
\text{model } \ P(c_1, c_2, \ldots, c_n \mid x)
\]
Conditional random field for labeling

\[
P(c_1, c_2, \ldots, c_n | \mathbf{x}) = \frac{1}{Z(\mathbf{x})} \prod_{i=1..n} P(c_i | x_i) \prod_{i,j} P(c_i, c_j | x_{ij})
\]

Unary term
Conditional random field for labeling

\[ P(c_1, c_2, ..., c_n \mid x) = \frac{1}{Z(x)} \prod_{i=1}^{n} P(c_i \mid x_i) \prod_{i,j} P(c_i, c_j \mid x_{ij}) \]
Conditional random field for labeling

\[ P(c_1, c_2, \ldots, c_n \mid x) = \frac{1}{Z(x)} \prod_{i=1}^{n} P(c_i \mid x_i) \prod_{i,j} P(c_i, c_j \mid x_{ij}) \]

**Pairwise term**
Conditional random field for labeling

\[
P(c_1, c_2, ..., c_n \mid x) = \frac{1}{Z(x)} \prod_{i=1..n} P(c_i \mid x_i) \prod_{i,j} P(c_i, c_j \mid x_{ij})
\]
Conditional random field for labeling

\[ P(c_1, c_2, \ldots, c_n \mid x) = \frac{1}{Z(x)} \prod_{i=1..n} P(c_i \mid x_i) \prod_{i,j} P(c_i, c_j \mid x_{ij}) \]

**Unary term**
Unary term

$P(c_i \mid x_i)$
Unary term

Most-likely labels

Unary term

Classifier entropy
Conditional Random Field for labeling

\[ P(c_1, c_2, \ldots, c_n \mid x) = \frac{1}{Z(x)} \prod_{i=1}^{n} P(c_i \mid x_i) \prod_{i,j} P(c_i, c_j \mid x_{ij}) \]

Pairwise term
Pairwise Term

\[ P(c \neq c' | x_{ij}) L(c, c') \]
Maximum A Posteriori assignment

\[
\text{arg max } P(c_1, c_2, \ldots, c_n \mid x) \\
\quad c_1, c_2, \ldots, c_n
\]
Dataset used in experiments

We label 380 meshes (19 categories) from the Princeton Segmentation Benchmark

[Chen et al. 2009]
Quantitative Evaluation

Segmentation

• Our result: 9.5% Rand Index error
• Outperforms all prior work:
  • 15% Randomized Cuts [Golovinskiy and Funkhouser 08]
  • 17% Normalized Cuts [Golovinskiy and Funkhouser 08]
  • 17.5% Shape Diameter [Shapira et al. 08]
  • 21% Core Extraction [Katz et al. 05]
  • 21% Fitting Primitives [Attene et al. 06]
  • 21.5% Random Walks [Lai et al. 08]
  • 21% Intrinsic Symmetry [Solomon et al. 11]
Labeling results
Summary

Use prior knowledge for shape segmentation and labeling
Based on a probabilistic model learned from examples
Significant improvements over the state-of-the-art
Generalization across categories:
Outline

1. Learning 3D shape segmentation and labeling

1. A generative model of shapes
   [Kalogerakis et al., SIGGRAPH 2012]

2. Other ML applications to graphics and vision

3. Future work
Goal: generative model of shape
Related work: generative models of bodies & faces

Works on relatively simple shapes with fixed structure
Based on dense correspondences between input shapes

[Allen et al. SIGGRAPH 2003]
Learning shape structure

We want to model attributes related to shape structure.

Shape styles
Component styles
Number of components
Component geometry
Component placement

model  $P( R, \{S_v\}, \{N_v\}, \{G_v\}, \{T_v\})$
P(R)
P(R)
\[ P(R) \prod_{l \in L} \left[ P(N_l | R) \right] \]
\[ P(R) \prod_{l \in L} \left[ P(N_l \mid R) P(S_l \mid R) \right] \]
\[ P(R) \prod_{l \in L} [P(N_l | R) P(S_l | R)] \]
\[ \Pr(R) \prod_{l \in L} \left[ \Pr(N_l | R) \Pr(S_l | R) \Pr(T_l | N_l) \right] \]
\[
P(R) \prod_{l \in L} \left[ P(N_l | R) P(S_l | R) P(T_l | N_l) P(G_l | S_l) \right]
\]
\[ P(R) \prod_{l \in L} [P(N_l | R) P(S_l | R) P(T_l | N_l) P(G_l | S_l)] \]
\[ \prod_{l \in L} \left[\mathbb{P}(N_l | R) \mathbb{P}(S_l | R) \mathbb{P}(T_l | N_l) \mathbb{P}(G_l | S_l)\right] \]

Latent object style

Latent component style
Learn from training data:
latent styles
lateral edges
parameters of CPDs
Learning

Given observed data $\mathbf{O}$, find structure $\mathbf{G}$ that maximizes:

$$P(\mathbf{G} \mid \mathbf{O}) = \frac{P(\mathbf{O} \mid \mathbf{G}) P(\mathbf{G})}{P(\mathbf{O})}$$

Assuming uniform prior over structures, maximize marginal likelihood:

$$P(\mathbf{O} \mid \mathbf{G}) = \sum_{\mathbf{R}, \mathbf{S}} \int P(\mathbf{O}, \mathbf{R}, \mathbf{S} \mid \Theta, \mathbf{G}) P(\Theta \mid \mathbf{G}) \, d\Theta$$
Learning

Given observed data $O$, find structure $G$ that maximizes:

$$P(G \mid O) = \frac{P(O \mid G)P(G)}{P(O)}$$

Assuming uniform prior over structures, maximize marginal likelihood:

$$P(O \mid G) = \sum_{R,S} \int P(O, R, S \mid \Theta, G)P(\Theta \mid G) \, d\Theta$$
Given observed data $O$, find structure $G$ that maximizes:

$$P(G \mid O) = \frac{P(O \mid G)P(G)}{P(O)}$$

Assuming uniform prior over structures, maximize marginal likelihood:

$$P(O \mid G) = \sum_{R,S} \int P(O, R, S \mid \Theta, G)P(\Theta \mid G) \, d\Theta$$
Learning

Given observed data $\mathbf{O}$, find structure $\mathbf{G}$ that maximizes:

$$P(G | O) = \frac{P(O | G)P(G)}{P(O)}$$

Assuming uniform prior over structures, maximize marginal likelihood:

$$P(O | G) = \sum_{R,S} \int_{\Theta} P(O, R, S | \Theta, G)P(\Theta | G) \ d\Theta$$
Marginal likelihood

\[ P(O \mid G) = \sum_{R,S} \int_{\Theta} P(O, R, S \mid \Theta, G) P(\Theta \mid G) \, d\Theta \]

Summation over all possible assignments to the latent variables
Marginal likelihood

\[ P(O \mid G) = \sum_{R,S} \int_{\Theta} \frac{P(O, R, S \mid \Theta, G) P(\Theta \mid G)}{P(O \mid G)} \, d\Theta \]

need inference for each data instance
Cheeseman-Stutz score

\[ P(O \mid G) \approx P(O^* \mid G) \cdot \frac{P(O \mid G, \tilde{\Theta}_G)}{P(O^* \mid G, \tilde{\Theta}_G)} \]

\( O^* \) is a fictitious dataset composed of training data \( O \) and approximate statistics for latent variables.

\( \tilde{\Theta}_G \) are MAP estimates found by Expectation-Maximization

\[ \tilde{\Theta}_G = \arg \max_{\Theta} P(O \mid G, \Theta) P(\Theta \mid G) \]
Shape synthesis

New shape

Source shapes
(colored parts are selected for the new shape)
Shape synthesis

New shape

Source shapes
(colored parts are selected for the new shape)
Results of alternative models

No latent variables

No lateral edges
User Survey

Training shapes

440

211

414

Synthesized shapes

prefer left

undecided

prefer right
Constrained shape synthesis
Summary

Generative model of shape structure
Learns structural variability from examples
Applicable to a broad range of complex domains
Enables new capabilities for shape processing
Outline

1. Learning 3D shape segmentation and labeling
2. A generative model of shapes
3. Other ML applications to graphics and vision
4. Future work
ML for vision: image sequence geolocation

Want: geo-tags

Kalogerakis, Vesselova, Hays, Efros, Hertzmann, ICCV 2009
How likely are you to travel from one place to another in a given amount of time?

\[ P(L_{t+1} = i | L_t = j, \Delta T_t) \]
Image sequence geolocation

137 photos in a user’s image sequence

Ground truth path

Estimated path
Learning hatching styles

Artist’s hatching illustration

Detected parts of coherent strokes

Learned model of stroke properties and parts

Kalogerakis, Nowrouzehahrai, Breslav, Hertzmann
ACM Transactions on Graphics 2012
Training illustration

Generalization to novel views and objects
Outline

1. Learning 3D shape segmentation and labeling
2. A generative model of shapes
3. Other ML applications to graphics and vision
4. Future work
Shape understanding in the wild

KinectFusion
[Izadi et al., UIST 2011]
Research goals

**Advance shape understanding:**
Joint shape recognition and segmentation
Hierarchical shape categorization
Map NL to shapes and deformation handles
Understand function from shapes, print 3D functional shapes

**Generative models** for:
Variability in symmetries
Architecture
Entire scenes
Images and shapes

**Learning algorithms** for:
Inferring physical/simulation parameters of shapes
Inferring shape deformations
Texturing, placing lights, other artistic rendering styles
Thank you!

My web page (code, data, demos, videos, papers, etc):

http://people.cs.umass.edu/~kalo/