Multiscale 3D Feature Extraction and Matching

Hadi Fadaifard and George Wolberg

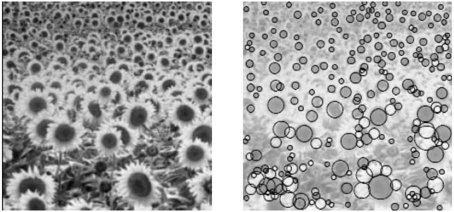
Graduate Center, City University of New York

May 18, 2011

Overview

1. We present a scale-space based surface representation useful for shape matching
2. The representation is
   - insensitive to noise
   - computationally efficient
   - capable of automatic scale selection
3. We refer to our representation as **Curvature Scale Space 3D (CS3)**.

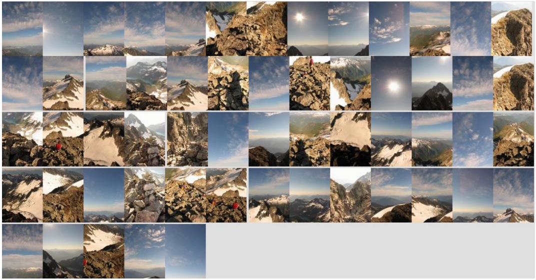
Motivation:



(a) (b)

Figure: Blob detection with automatic scale selection (Lindeberg, '94).

Motivation—SIFT → AutoStitch, Photosynth, etc



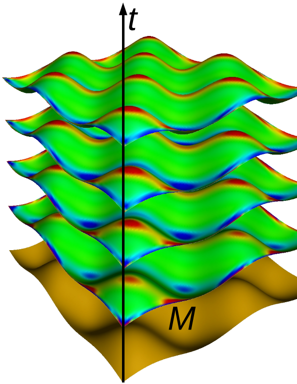
↓

Scale Space Representation of Surface Signals (Continuous)

- Scale-space representation, $F : \mathcal{M} \times \mathbb{R} \rightarrow \mathbb{R}^n$, of signal $f : \mathcal{M} \rightarrow \mathbb{R}^n$ on surface $\mathcal{M}$ is defined as the solution to the heat (diffusion) equation

$$\frac{\partial}{\partial t} F(\mathbf{x}; t) = \Delta_{\mathcal{M}} F(\mathbf{x}; t) \ , \tag{1}$$

  with the initial condition $F(\mathbf{x}; 0) = f(\mathbf{x})$.
  - $\Delta_{\mathcal{M}}$ is the Laplace-Beltrami operator



Scale Space Representation of Surface Signals (Discrete)

- Surface is represented by polygonal mesh $\mathcal{M} = (\mathcal{V}, \mathcal{E})$, where
  - $\mathcal{V} = \{v_1, \ldots, v_N\}$ is the vertex set
  - $\mathcal{E} = \{e_{ij} | v_i \text{ is connected to } v_j\}$ is the edge set
- Let $f : \mathcal{V} \to \mathbb{R}^n$ denote the initial signal on $\mathcal{M}$.
- Let

$$\mathbf{f} = \begin{pmatrix} f(v_1) & \ldots & f(v_N) \end{pmatrix}^\top . \tag{2}$$

- The solution to the heat equation can be estimated using the *backward Euler method*

$$(\mathbf{I} - \lambda \mathbf{L})\mathbf{f}^\lambda = \mathbf{f} , \tag{3}$$

  where
  - λ is a time step,
  - vector $\mathbf{f}^\lambda$ contains the signal values at time λ,
  - $\mathbf{I}$ is the $N \times N$ identity matrix, and $\mathbf{L} = (w_{ij})_{N \times N}$ is the Laplacian matrix with elements

$$w_{ij} = \begin{cases} -1 & \text{for } i = j , \\ \frac{1}{|\mathcal{N}(i)|} & \text{for } j \in \mathcal{N}(i) , \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

Scale Space Representation of Surface Signals (3)

Advantages:

- Equation
$$(\mathbf{I} - \lambda \mathbf{L})\mathbf{f}^{\lambda} = \mathbf{f}^{0} \ , \tag{5}$$
is sparse and can be solved efficiently using the Conjugate Gradient Method.

Disadvantages:

- What is the relation between λ and the time parameter t in the original heat equation?
- The resulting transfer function of Eq. (5) is
$$h(\omega) = (1 + \lambda\omega^2)^{-1} \ , \tag{6}$$
where ω denotes surface signal frequency (Desbrun, '99).
- However, to be consistent with the scale-space representation of signals in $\mathbb{R}^n$, we desire the transfer function to be a Gaussian.

Scale Space Representation of Surface Signals (4)

- The scale-space representation of the surface signal $\mathbf{f}$ is then given by the sequence $(\mathbf{F}^0, \ldots, \mathbf{F}^{L-1})$, which is obtained iteratively using

$$\mathbf{F}^l = \begin{cases} (\mathbf{I} - \lambda_{l-1}\mathbf{L})^{-1}\mathbf{F}^{l-1} & \text{if } l > 0 \\ \mathbf{f} & \text{if } l = 0 \end{cases}, \tag{7}$$

  where $(\lambda_0, \ldots, \lambda_{L-1})$ denotes a sequence of time steps used at intermediate levels $l = 0, \ldots, L-1$.
- Transfer function of $\mathbf{F}^0 \rightarrow \mathbf{F}^L$ is

$$h_L(\omega) = \prod_{l=0}^{L-1} (1 + \lambda_l \omega^2)^{-1}, \tag{8}$$

  which approaches a Gaussian as L grows (central limit theorem).

Scale Space Representation of Surface Signals (5)

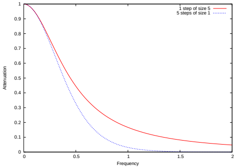


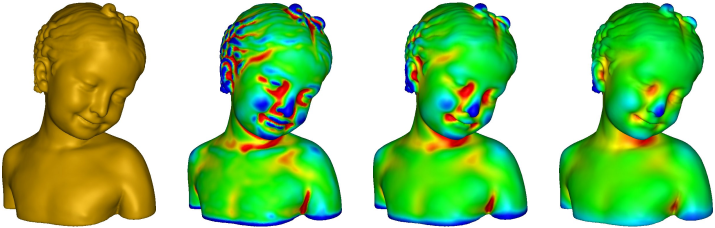
Figure: $h_L(\omega)$ at $L = 1$ (red) and $L = 5$ (blue).

- We define the *scale parameter* at level L as the *inverse of variance of the transfer function* at L:

$$t_L = \frac{\int_{-\infty}^{\infty} \left(\prod_{l=0}^{L-1} (1 + \lambda_l \omega^2)^{-1} \right) d\omega}{\int_{-\infty}^{\infty} \left(\omega^2 \prod_{l=0}^{L-1} (1 + \lambda_l \omega^2)^{-1} \right) d\omega} \tag{9}$$

Curvature Scale Space 3D (CS3)

- **Curvature Scale Space 3D (CS3)** of surface $\mathcal{M}$ is defined as the scale-space representation of the *surface (mean) curvatures*.



(a) (b) (c) (d)

Figure: The CS3 representation of the Bimba model shown in (a), at scales (b) $t = 3.0$, (c) $t = 7.5$, (d) $t = 13.8$.

- We introduce the *scale-invariant* Laplacian of Curvatures (si-LoC) as

$$\Delta^{si}\mathbf{F}^{l} = \frac{\Delta\mathbf{F}^{l} - \bar{\mathbf{F}}^{l}}{\sigma_{l}}\ , \tag{10}$$

where

$$\Delta\mathbf{F}^{l} = \frac{2(\mathbf{F}^{l+1} - \mathbf{F}^{l})}{t_{l+1} - t_{l}}\ . \tag{11}$$

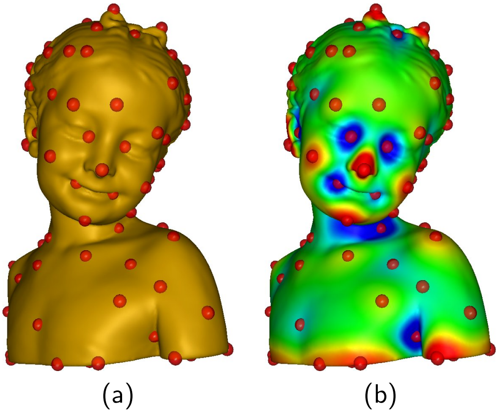
and

$$\bar{\mathbf{F}}^{l} = \frac{1}{N}\mathbf{1}^{\top}\Delta\mathbf{F}^{l}\mathbf{1}, \quad \sigma_{l} = \frac{1}{\sqrt{N}}\|\Delta\mathbf{F}^{l} - \bar{\mathbf{F}}^{l}\|\ , \tag{12}$$

denote the vector-form mean, and standard deviation of the LoC values at level l, respectively;

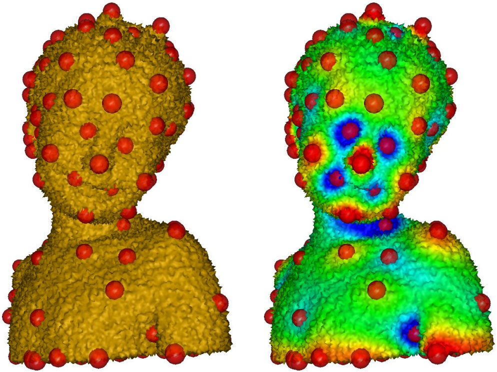
- N is the total number of vertices in $\mathcal{M}$, and $\mathbf{1}$ is an N-dimensional vector of all 1's.

Feature Extraction using CS3 (2)



Extracted features on the Bimba model at $t = 21.7$; the false-colors in (b) reflect the response of the Δ^{si} at each vertex on the model.

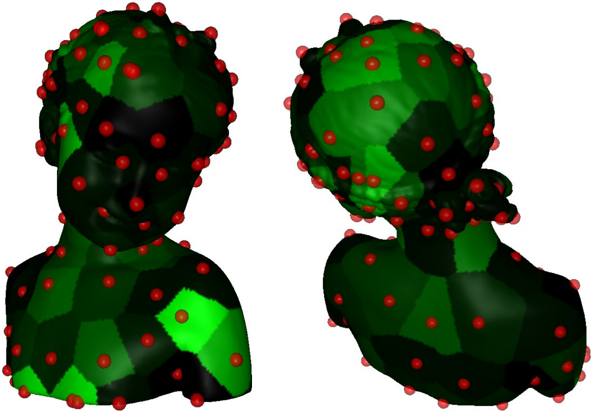
Feature Extraction using CS3 (3)



(a) (b)

Extracted features on a noisy (80% Gaussian) version of the model at $t = 21.7$

Feature Extraction using CS3 (4)


(a) (b)

Slippage of keypoints due to noise.

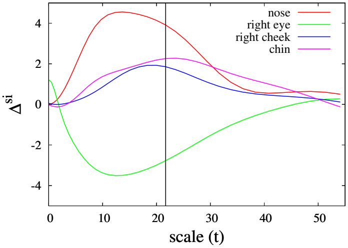
Feature Extraction using CS3 (5)

Displacement statistics:

min	max	avg.	std.	avg. edge len	avg. disp. (nbrsz)
0.000	0.127	0.023	0.022	0.0057	4.044

159 keypoints on noiseless mesh
166 keypoints on noisy mesh

Feature Extraction using CS3 (6)

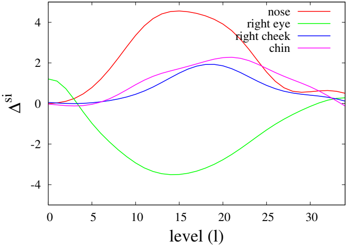


(a) original

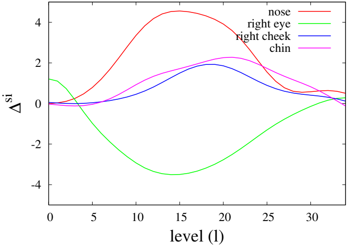
(b) noisy

Figure: Plots of the *scale-invariant* LoC values of a few vertices on the Bimba model

Feature Extraction using CS3 (7)



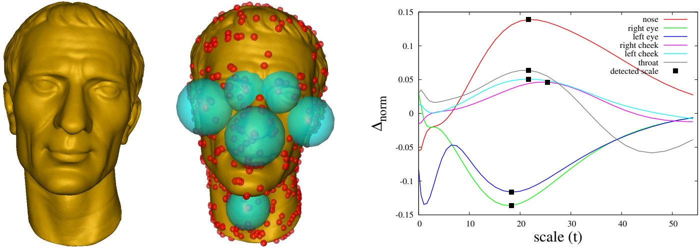
(a) original



(b) spatial scaling: $\times 100$

Figure: Comparison of scale-invariant LoC curves of the Bimba model with different spatial scales

Feature Extraction with Auto. Scale Selection

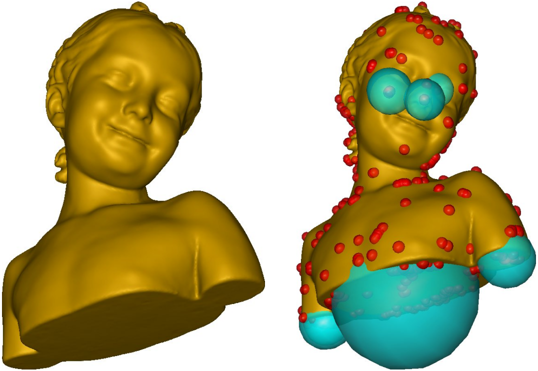


(a) (b)

Figure: Automatic scale selection on the Caesar model. (a) Original model. (b) Estimated scales at a few locations on the model. (c) Plots of the scale-normalized Laplacian of the surface mean curvatures at the selected vertices as functions of scale; the locations of the filled squares on the scale-axis indicate the detected scale for the keypoints.

The radii of the blue spheres are drawn proportional to the detected scales at the extracted keypoints.

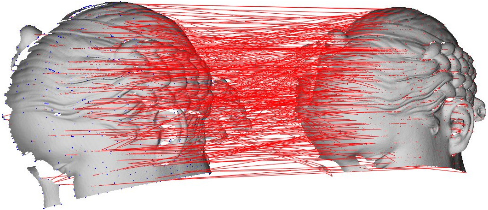
Feature Extraction with Auto. Scale Selection (2)



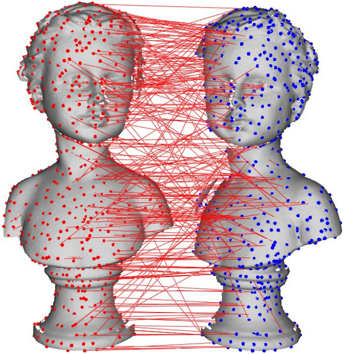
(a) (b)

Figure: Automatic scale selection on the Bimba model. (a) Original model; (b) estimated scales at a few locations on the model.

Application: Surface Registration

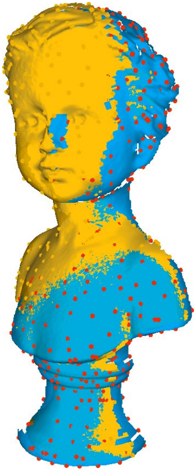


(a)

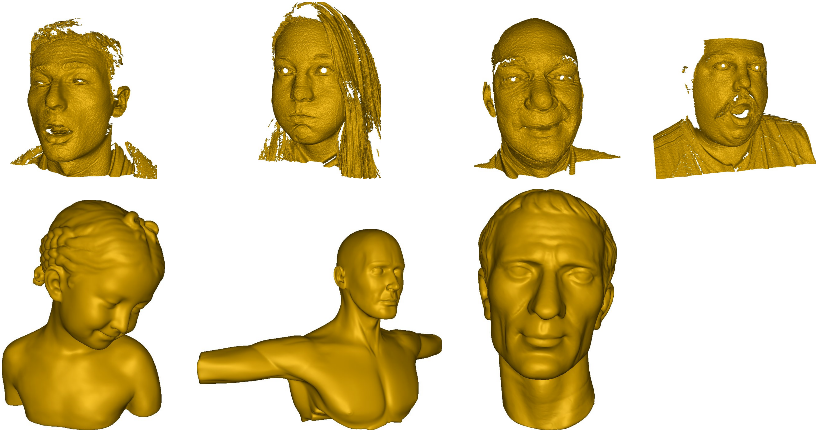


(b)

Application: Surface Registration Results

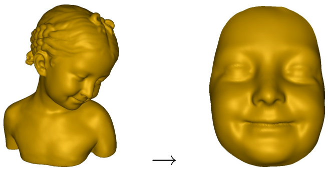


Application: 3D Face Extraction/Recognition



Sample inputs to our face extraction/recognition system

Application: 3D Face Extraction



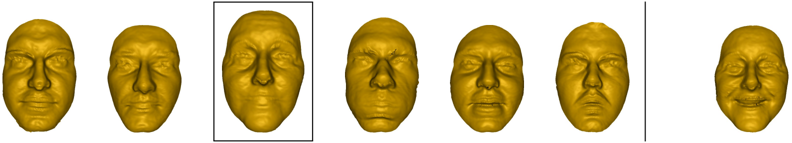
input $\longrightarrow$ output

- Performance of our extraction system was tested on a set of 1068 models.
- The correct extraction rate of the system was 92.13%.
- The extraction time on a model with 113.4K vertices was 180secs on a 2.4 GHz CPU.

3D Face Recognition

Face Recognition:

- Input: database of 3D faces with known classes + a query model
- Output: determine to which class the query belongs



(a) training set (gallery) (b) query (probe)

Examples of 3D faces used in our 3D face recognition system

GavabDB Dataset

- Contains 3D face scans of 61 individuals (used 7 scans per individual—total of 427 scans)
- Contains the following poses/expressions:
  - 1 scan looking up,
  - 1 scan looking down,
  - 2 frontal scans,
  - 1 scan with random gesture,
  - 1 scan with laughter,
  - 1 scan with smile.

Comparison of Results with Other Works (GavabDB)

Pose	This Work	Mahoor '09	Berretti '07
Frontal	95.08%	95.0%	94%
Smile	93.44%	83.6%	85%
Laughter	80.33%	68.9%	81%
Random gesture	78.69%	63.4%	77%
Looking down	88.52%	85.3%	80%
Looking up	85.25%	88.6%	79%
Overall	86.89%	82.83%	84.29%

Approach	Controlled	Non-controlled
This work	98.36%	96.02%
Moreno (PCA)	82.00%	76.20%
Moreno (SVM)	90.16%	77.90%

Future Work

- Improve the performance of the face extraction system (92.13% accuracy).
- Test the performance of the CS3 representation in a 3D face recognition system (more extensive tests).