Slide 1

Hadi really wanted to be here for the presentation.

Unfortunately, because of visa issues he was unable to attend the conference.

I will be presenting the work instead *(thanks Weihong)*.

The work is titled *Multiscale 3D Feature Extraction and Matching*.

Slide 2

We present a new scale-space based representation for 3D surfaces, which we refer to as the **Curvature Scale Space 3D (CS3)**.

We show that the proposed representation is

- insensitive to noise,
- computationally efficient, and
- capable of automatic scale selection.

The motivation for the work is the scale-space theory for 2D images, which has been extensively studied over the past three decades.

One of the main applications of the scale-space theory (and representation) for images is in feature extraction and descriptor computation.

For example, in the figures below, we show how the scale-space representation can be used to extract the locations of blob-like features in an image (in a manner which is invariant to noise in the input image).

The blob detector used here is defined in terms of the maxima of the Laplacian of the image intensities.

One of the useful features of the scale-space theory is its "automatic scale selection principle," which enables associating a scale (neighborhood size) to the detected structures in images.

For example, Fig. (b) shows, not only the detected locations of blob-like features in the image, but also their detected <u>sizes</u>.

Automatic scale selection is one of the features that we seek in our scale-space representation for 3D surfaces.

Slide 3

Many applications of the scale-space theory has been shown in computer vision.

Perhaps the most well-known and successful example is the Scale Invariant Feature Transform (SIFT) of David Lowe.

One of the earliest uses of SIFT was shown in automatic stitching of panoramic images (shown here):

Given an unorganized set of panoramic images, the proposed AutoStitch approach of Matthew Brown automatically establishes correspondences between the images, registers, and blends them to create the final panoramic image.

Since then, SIFT has been used in more impressive applications such as Photosynth.

The goal of this work is to obtain a similar scale-space representation for 3D surfaces, which can then be used in similar applications.

However, the main application we consider in this work is for 3D face recognition.

Slide 4

We are not going to review related work on scale-space representations for 3D surfaces. Please see the paper for details.

In order to define our proposed **Curvature Scale Space 3D (CS3)** representation for 3D surfaces, we need to first define the scale-space representation for surface signals.

We first present the formulation in the continuous setting, and then show how it can be implemented efficiently in the discrete setting.

The scale-space representation, F (read "upper-case f"), of a continuous surface signal f (read "lower-case f"), which maps each point on surface $\mathcal{M}$ to a value in $\mathbb{R}^n$ (read R n) is defined as the solution to the heat (diffusion) equation; with the initial condition that F at time $t = 0$ is the same as the initial surface signal.

In the above formulation $\Delta_{\mathcal{M}}$ (delta M) denotes the Laplace-Beltrami operator.

The figure on the right shows an example of the scale-space representation of a surface signal. The brown (M level) shows the domain the signal and the colored levels show the signal values at different times.

Slide 5

In the discrete setting, the input surface is represented by polygonal mesh $\mathcal{M}$, with vertex set V and edge set E. V contains N vertices.

Let function f, which maps each mesh vertex to a value in $\mathbb{R}^n$, denote the initial signal on M.

f can then be represented in the form of an N-dimensional array $\mathbf{f}$ (bold-face f), as shown in Eq. (2).

Slide 6

The discretized scale-space representation of the surface signal is then given by the sequence of vectors $\mathbf{F^0}, ..., \mathbf{F}^{L-1}$ (f-zero to f-L minus 1), which are obtained iteratively using Eq. (3).

The definition requires solving a sparse linear system at each level in the scale-space stack.

The system can be solved efficiently using the Preconditioned Conjugate Gradient Method (PCGM).

$\lambda_0, ...\lambda_{L-1}$ denotes the sequence of time steps used at each level in the scale-space stack.

The resulting transfer function of going from level 0 to level L is given by Eq. (4).

Note that a nice property of the above formulation (Eq. (3)) is that as L grows, the corresponding transfer function approaches a Gaussian.

Slide 7

The figure shows the transfer function of the smoothing process at levels 1 and 5 of the scale-space stack of the surface signal.

We define the *scale* associated with level L as the *inverse of the variance of the transfer function* at L.

The scale parameter is needed in our "automatic scale selection" procedure, which is described in the paper.

Note that a somewhat different formulation for t_L is given in the paper. Both formulations worked well in our experiments.

Slide 8

Finally, we define the **Curvature Scale Space 3D (CS3)** representation of surface M, as the *scale-space representation of its mean curvatures.*

The figure shows example of the CS3 representation of a 3D surface at various scales.

Slide 9

The CS3 representation of a 3D surface can be used for keypoint (feature point) extraction on the surface.

For example, the Laplacian of surface curvatures can be easily estimated using Eq. (6), which corresponds to computing the difference between signal values at two consecutive levels in the CS3 stack.

Additionally, we introduce the *scale-invariant* Laplacian of Curvatures (hereafter, referred to as si-LoC (read as "s-i-l-o-c") as in Eq. (7).

$\bar{\mathbf{F}}^l$ (read f bar), and $\Delta\mathbf{F}^l$ (read delta f), denote the mean-vector and standard deviation of the Laplacian of surface curvatures (LoC) at level l.

See paper for more details.

Slide 10

The false-coloring in this figure shows how the distribution of the si-LoC values looks like over a 3D model (at scale $t = 21.7$).

Additionally, the red spheres identify the locations where the si-LoC values are locally maximum or minimum.

As expected, the locations of these extrema correspond to round, or blob-like structures (for example, nose tip, eyes, etc) on the surface.

Slide 11

This figure shows the same thing as the previous slide, except on a noisy version of the model with 80% Gaussian noise.

As can be seen, the si-LoC values are insensitive to noise and the false-coloring looks almost identical to the noise-free case.

Therefore, it can be said that the feature extraction process is also insensitive to noise.

Slide 12

The figure on the left shows how the si-LoC values at a few locations on the original (noise-free) model change with scale in the CS3 representation.

In the figure on the right, we show the same plots, except on the noisy version of the model.

Note how the plots (except at the first few scales) are similar.

This verifies that our proposed representation is insensitive to noise.

Slide 13

(SKIP THIS SLIDE)

In these figures, we compare the si-LoC plots on the original model and its scaled version; the model was scaled by a factor of 100.

Note that the two plots are identical.

This shows that the si-LoC values are in fact invariant to scale changes between 3D models.

Slide 14

(You will probably need to go over this slide very quickly—No need to go into too much detail here)

Note that in most cases, the si-LoC values at the vertices either first decrease and then increase, or first increase and then decrease.

Therefore, a similar "automatic scale selection" approach used for 2D images can be used here to associate a scale (and consequently a neighborhood size) to each keypoint on the surface.

In our approach, the scale where the si-LoC values at a keypoint become locally maximum or minimum across scales was chosen as the scale associated with that keypoint.

If a keypoint has multiple such extrema, multiple scales are associated with it.

In the figure on the right, the black squares indicate the detected scales at a few keypoints on the model on the left.

The blue spheres in Fig. (b) visualize the detected scales at a few keypoints on the model. Each sphere was drawn proportional to the detected scale at the keypoint.

As can be seen, the size of the spheres correspond to the size of the underlying structures on the surface.

Slide 15

This figure shows example of our proposed automatic scale selection procedure applied to a few keypoints on another 3D model.

Note how the sizes of the blue spheres at the bottom of the model are consistent with the actual size of the regions.

Slide 16

The keypoint extraction process and si-LoC values can be used in various shape matching applications.

For example, in these figures, the extrema of si-LoC values on the surfaces were used as keypoints.

Additionally, correspondences between the keypoints on the surfaces were established using the si-LoC values.

The mismatches can be eliminated using techniques such as RANSAC or branch and bound, and the surfaces can be consequently registered.

Slide 17

This figure shows the automatic surface registration results for one of the pairs in the previous slide.

Slide 18

The main application we had in mind for the proposed representation was for 3D face recognition.

The goal was to design a fully automatic 3D face recognition system capable of handling input 3D faces from various sources, and with arbitrary scale.

This slide shows few examples of possible inputs to our system

Slide 19

A component of our recognition system consists of automatic extraction of the face region on an input 3D model.

The details of the approach, which uses the proposed CS3 representation, are given in the paper.

We tested the performance of our face extraction procedure on a dataset of 1068 models.

The correct extraction rate of the system was approximately 92%.

The average extraction time on a 3D model with approximately 110,000 vertices was 3 minutes on a 2.4GHz CPU.

Slide 20

As mentioned earlier, the main goal is to use the cropped faces in a 3D face recognition system.

The face recognition system consists of a database of 3D faces with known classes (gallery).

Given a query model (probe), the objective is then to determine to which class the probe belongs.

Again, in our face recognition system we used the proposed si-LoC values to form the feature vectors, which were used for the classification task.

I will very quickly go over some preliminary results, which were not included in the paper.

The details of the approach will be published soon in the journal version of the paper.

Slide 21

The system was tested on the Gavab dataset which contains scanned 3D human faces with various poses and expressions.

It is also very noisy.

Slide 22

These tables compare the performance of our system with other methods which used the same dataset in their tests.

In all cases, our system outperformed the other competing systems.

Slide 23

In future work, we will improve the performance of the face extraction system (92.13% accuracy).

We will also test the performance of the CS3 representation in a 3D face recognition system (with more extensive tests).

Thank you!