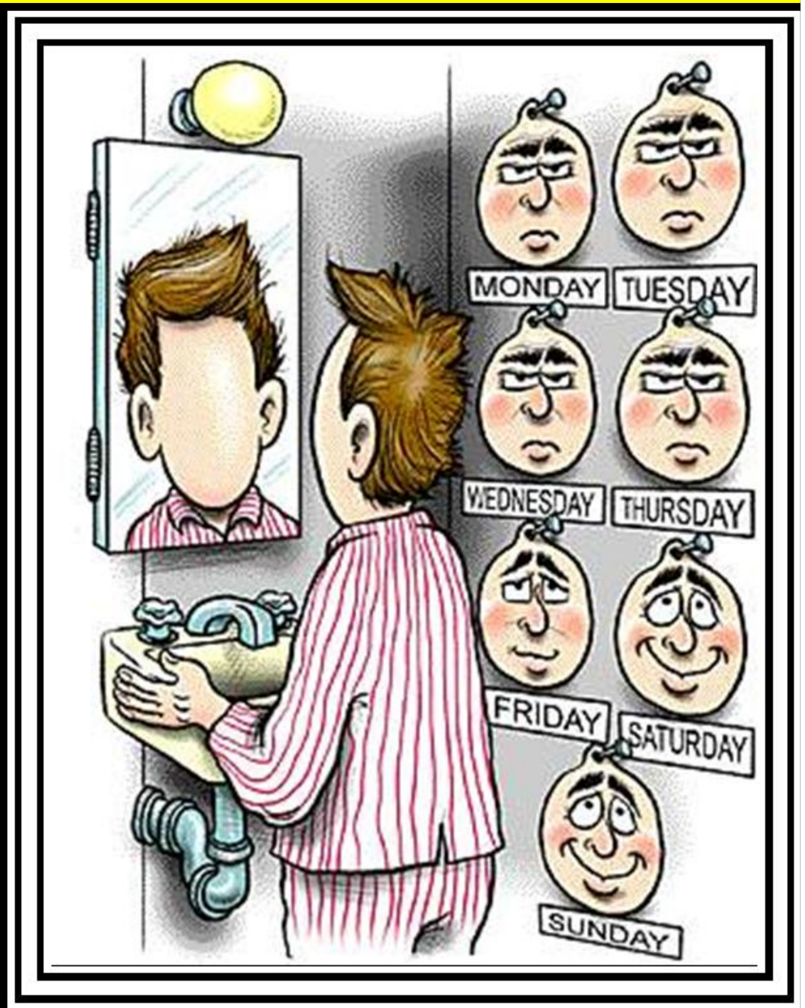
Computer Science
Department

Technion-Israel Institute of Technology

# Matching Isometric Manifolds by Flat Embedding

Ron Kimmel




Geometric Image Processing Lab

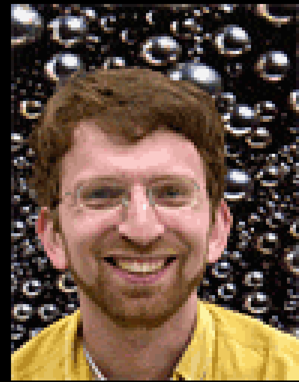
# *Students*

Michal Aharon
CS Technion

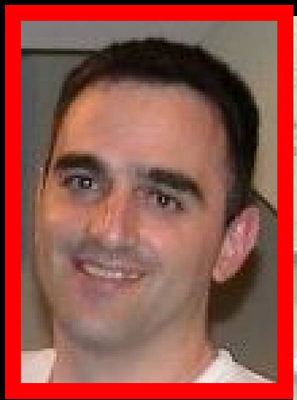
Ilya Blayvas
CS Technion

Alexander Bronstein
EE-CS Technion

Michael Bronstein
CS Technion

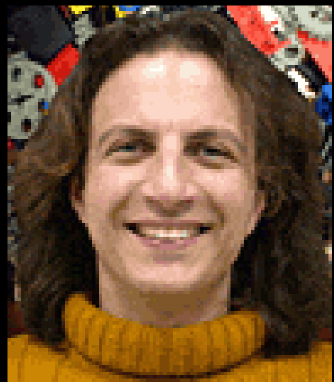


Alexander Brook
Math Technion



Asi Elad
IO-Image

Roman Goldenberg
CS Technion



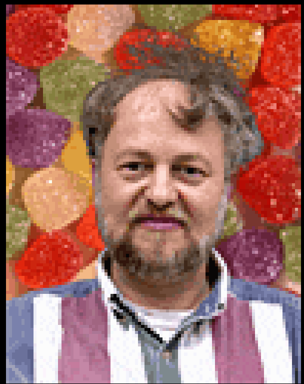
Eyal Gordon
CS Technion

Michal Holzman Gazit
EE-CS Technion

Alon Spira
CS Technion

Gil Zigelman
3DV Systems

# *Colleagues*


Alfred M. Bruckstein
CS Technion


Laurent Cohen
Paris Dauphine


Vicent Caselles
Math. Barcelona


Nahum Kiryati
Tel Aviv Univ.

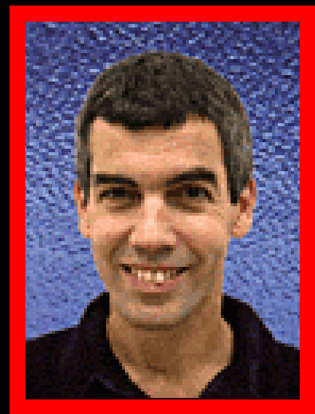

Ehud Rivlin
CS Technion


Michael Rudszky
CS Technion

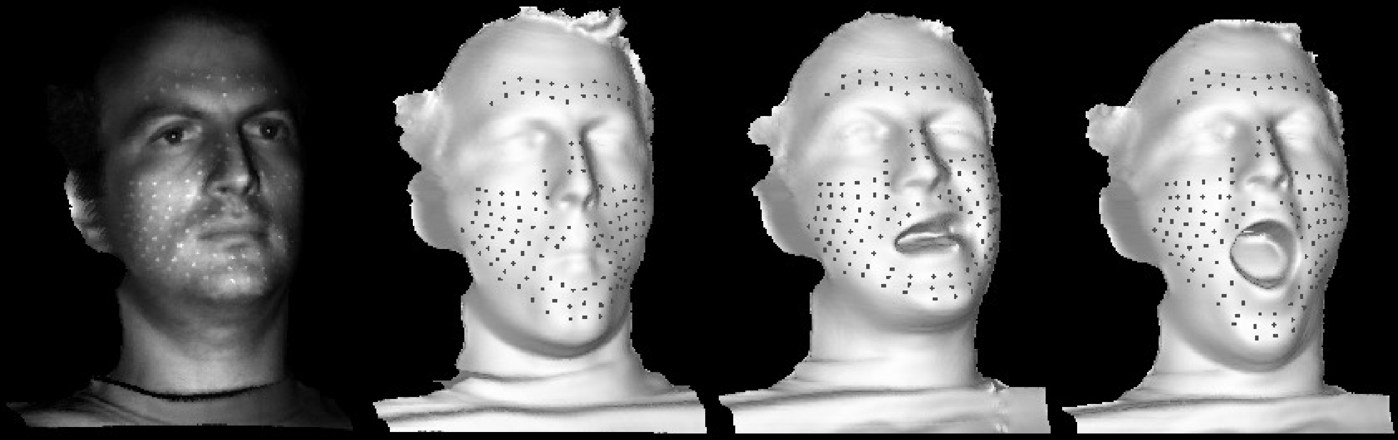

Guillermo Sapiro
EE Minnesota

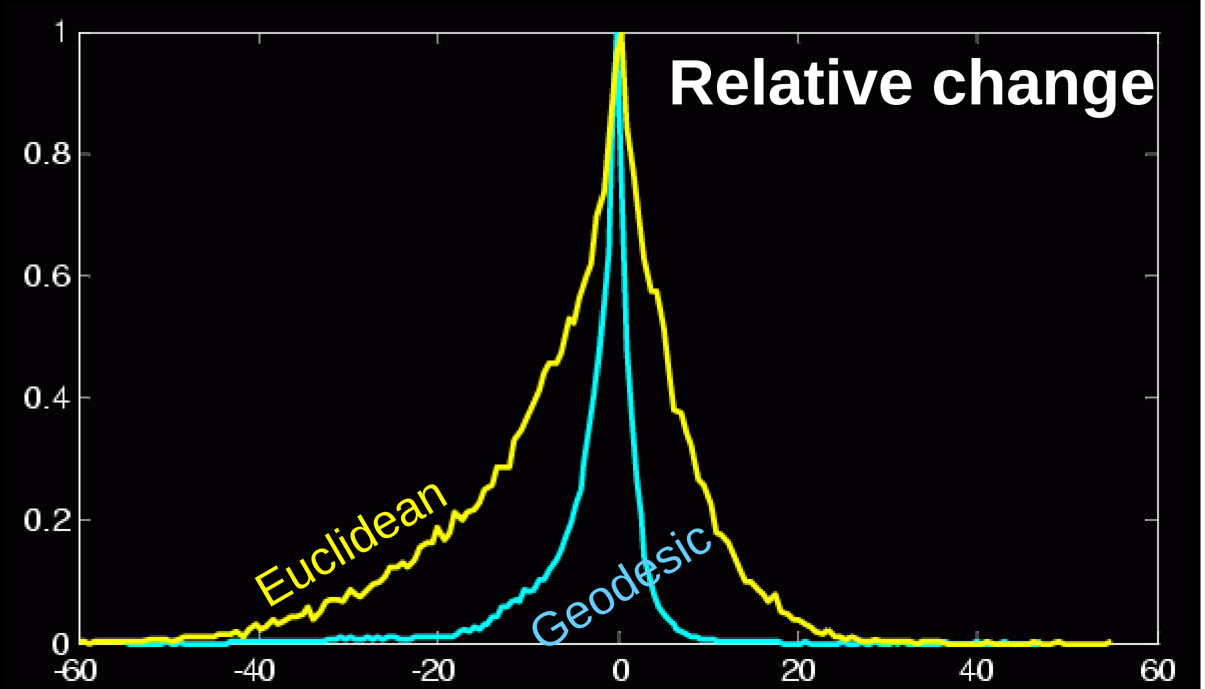

James Sethian
UC Berkeley


Nir Sochen
Tel Aviv Univ.


Irad Yavneh
CS Technion

# Expressions are ~isometries






Bronstein Bronstein Kimmel *2005*

# Open mouth & isometric expressions

On MDS,
Lip Reading,
Texture Mapping,
and
Isometric Signatures

# How it all started… (my story)

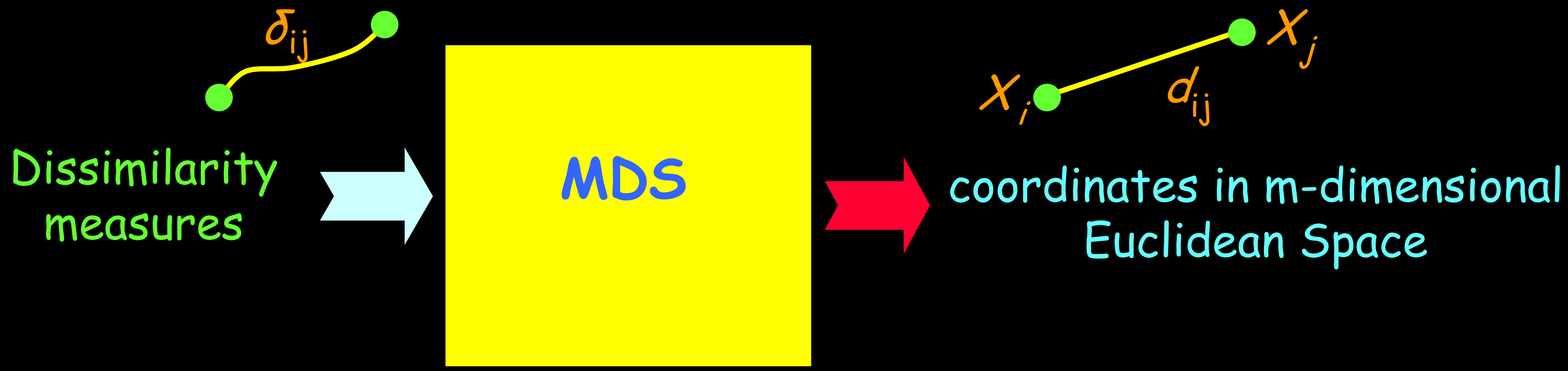
| | | | |
|---|---|---|---|
| 1989 | Schwartz Shaw Wolfson | TPAMI | 2D flattening |
| 1995 | Kimmel Amir Bruckstein | TPAMI | geodesic distances |
| 1998 | Kimmel Sethian | PONAS | geodesic distances |
| 2000 | Zigelman Kimmel Kiryati | TVCG | texture mapping |
| 2001 | A. Elad Kimmel | TPAMI | isometric signatures |
| 2003 | Bronstein$^2$ Kimmel | AVBPA | 3D face recognition |
| 2003 | Spira Kimmel | scale space | geodesic distances |
| 2004 | Bronstein$^2$ Spira Kimmel | ECCV | 3D recognition without 3D |
| 2005 | Bronstein$^2$ Kimmel | scale space | S3 embedding |
| 2005 | Bronstein$^2$ Kimmel | IJCV | 3D recognition |
| 2005 | Bronstein$^2$ Kimmel | 2Bsub. | Isometry/S2 embedding |

# Multidimensional scaling

# Multidimensional scaling

- MDS is a family of methods that map similarity measurements among objects, to points in a small dimensional Euclidean space.

- Enables to explore the geometric structure of the data.




$$Stress = \frac{\Sigma w_{ij} (\delta_{ij} - d_{ij}(X))^2}{\Sigma w_{ij} \delta_{ij}^2}$$

# A simple example


Rotation

Reflection

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| 12.7 | 0 | 32.8 | 29.4 | 14.6 | 8.8 | 4.5 | 6.4 | 16.4 | 13.8 |
| 4.1 | 13.9 | 13.2 | 14.9 | 9.0 | 8.2 | 16.6 | 19.6 | 25.4 | 0 |

Young et al. 1930,
Torgerson&Gower 1952,1958,1966.

# Classical scaling

Given n points in $R^k$, denote $\mathbf{p}_i = [x_i^1, x_i^2, ..., x_i^k]^{\mathbf{T}}$

Define coordinates vector $\mathbf{P} = [\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_n]^{\mathbf{T}}$

The Euclidean distance between 2 points:

$$d_{ij}^2 = \left|\mathbf{p}_i - \mathbf{p}_j\right|^2 = \left|\mathbf{p}_i\right|^2 - 2\mathbf{p}_i\mathbf{p}_j^{\mathbf{T}} + \left|\mathbf{p}_j\right|^2$$

$$\boxed{\mathbf{Q} = \begin{bmatrix} |\mathbf{p}_1|^2 & |\mathbf{p}_1|^2 & \cdots & |\mathbf{p}_1|^2 \\ |\mathbf{p}_2|^2 & |\mathbf{p}_2|^2 & \cdots & |\mathbf{p}_2|^2 \\ . & . & \cdots & . \\ |\mathbf{p}_n|^2 & |\mathbf{p}_n|^2 & \cdots & |\mathbf{p}_n|^2 \end{bmatrix}}$$

$$\mathbf{D} = \mathbf{Q} - 2\mathbf{P}\mathbf{P}^{\mathbf{T}} + \mathbf{Q}^{\mathbf{T}}$$

$$\mathbf{JDJ} = \mathbf{J}(\mathbf{Q} + \mathbf{Q}^{\mathbf{T}} - 2\mathbf{P}\mathbf{P}^{\mathbf{T}})\mathbf{J} = \mathbf{0} + \mathbf{0} - 2\tilde{\mathbf{P}}\tilde{\mathbf{P}}^{\mathbf{T}}$$

$$\boxed{\mathbf{J} = \mathbf{I} - \tfrac{1}{n}\mathbf{1}\mathbf{1}^T}$$

$$-\tfrac{1}{2}\mathbf{JDJ} = \mathbf{U}\Lambda\mathbf{U}^{\mathbf{T}} \quad \Rightarrow \quad \boxed{\tilde{\mathbf{P}}^{\mathbf{T}} = \Lambda^{\frac{1}{2}}\mathbf{U}^{\mathbf{T}}}$$

# Classical scaling

Matlab code
for 2D flattening

$$J = eye(n) - ones(n)./n;$$

$$B = -0.5 * J * D * J;$$

$$[U, L] = eigs(B, 2, 'LM');$$

$$newy = sqrt(L(1,1)).*U(:,1);$$

$$newx = sqrt(L(2,2)).*U(:,2);$$

Or a single command (Matlab 6.5)

$$Y = cmdscale(D);$$

$$newx = Y(1,:);$$

$$newy = Y(2,:);$$

Zigelman, Kimmel, Kiryati IEEE TVCG 2002

# Analyzing and Synthesizing Lips Movements

Lip-reading is a difficult task.

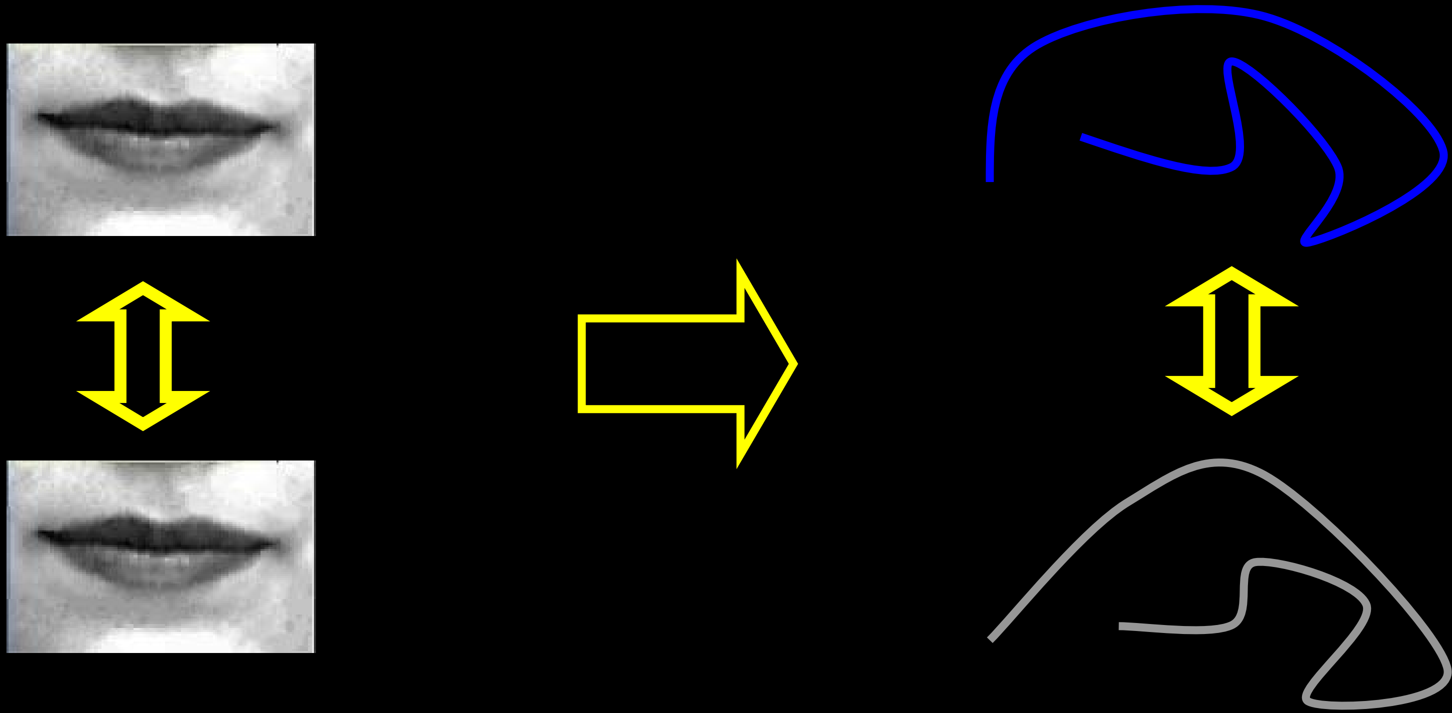
Goal: Recognize a limited set of words.

Previous Lip synthesis:
video-rewrite (Bregler)
eigenfacemasks (Van Gool).

Previous Lip Reading:
HMM (Bregler, 1998) (the bartender problem),
PCA (Li et al. 1997) (eigensequences),
NN and SVM (Bregler, Duchnowski).

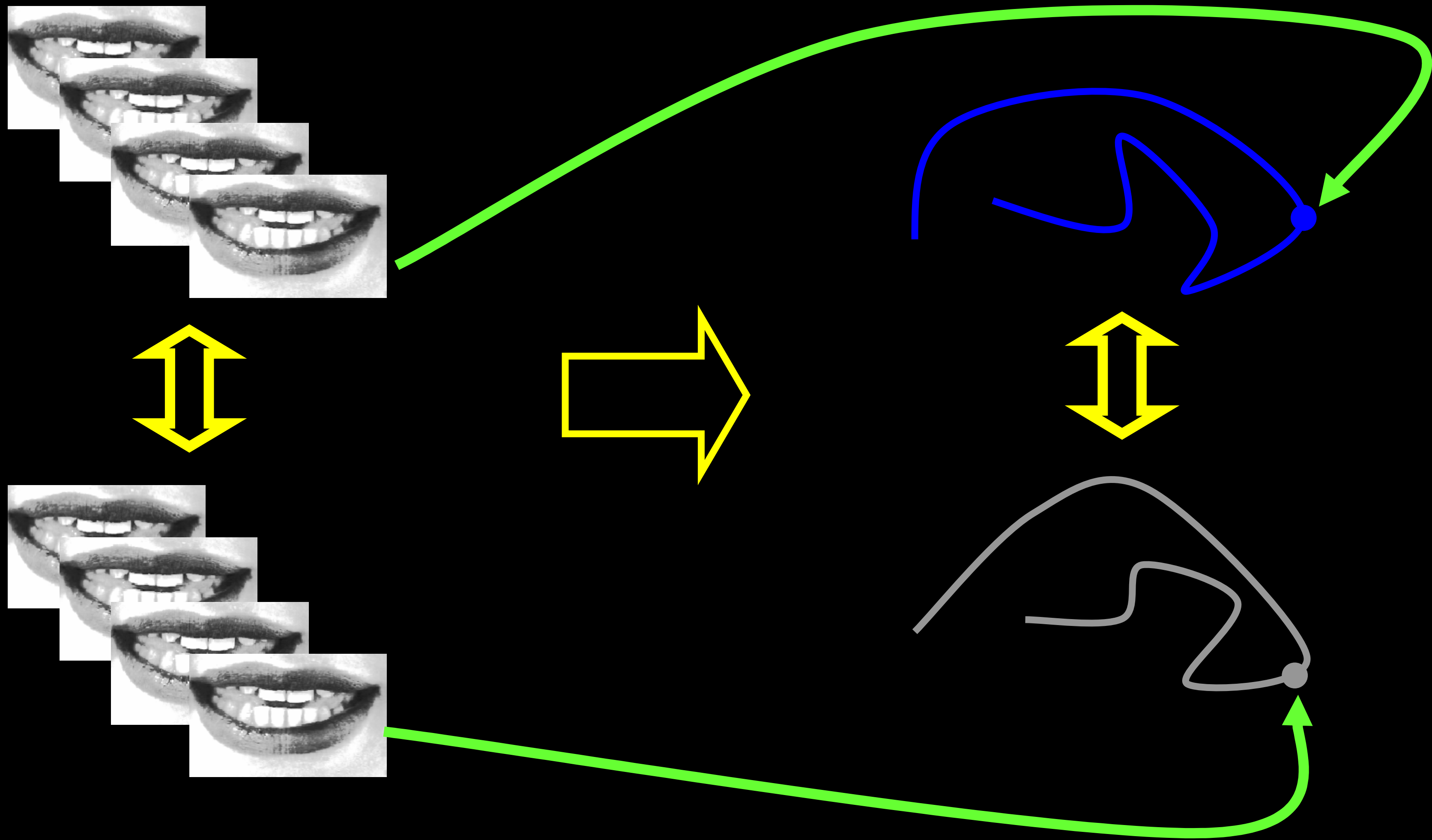
Michal Aharon, Kimmel 2004

# Main Idea



Michal Aharon, Kimmel 2004

# Main Idea



Michal Aharon, Kimmel 2004

# Framework

**Image acquisition (single speaker).**

**Lips area extraction**

**Distances measurement**

**Dim. reduction (flat embedding)**

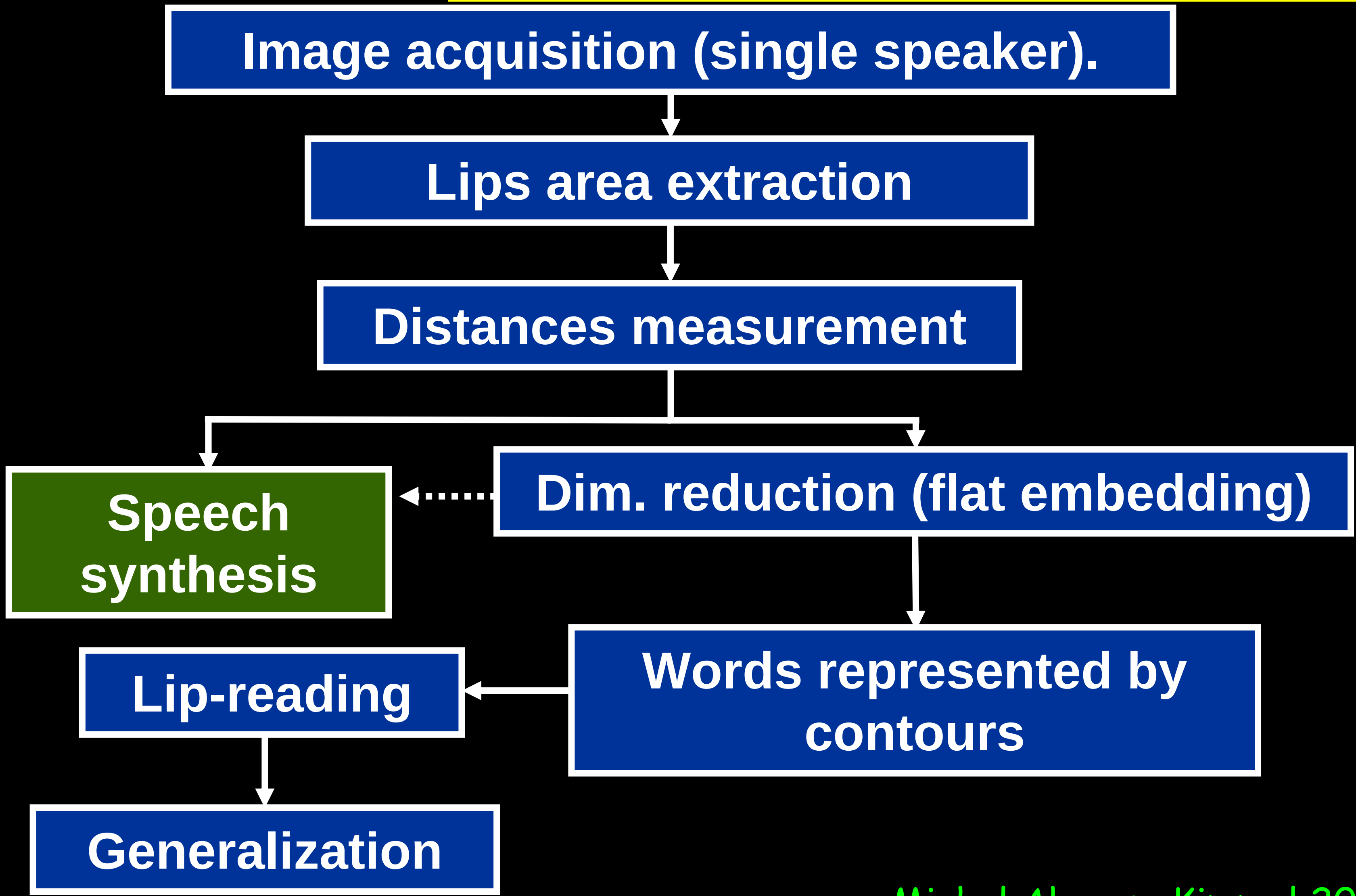
**Words represented by contours**

**Lip-reading**

**Generalization**

Michal Aharon, Kimmel 2004

# Framework

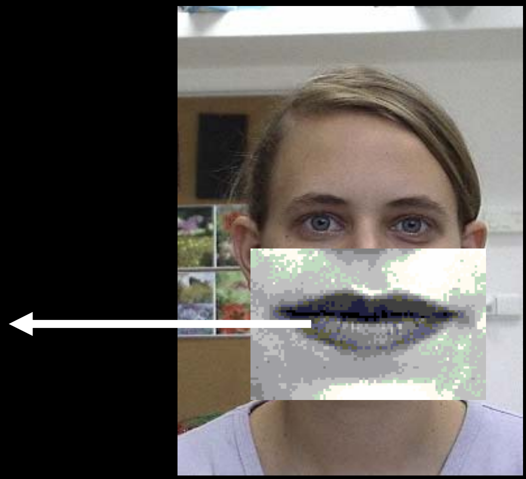



Michal Aharon, Kimmel 2004

# Image acquisition (single speaker).

# Lips Area extraction

- nose used for alignment
- images aligned using affine model
- The mouth area section is extracted as gray level images.



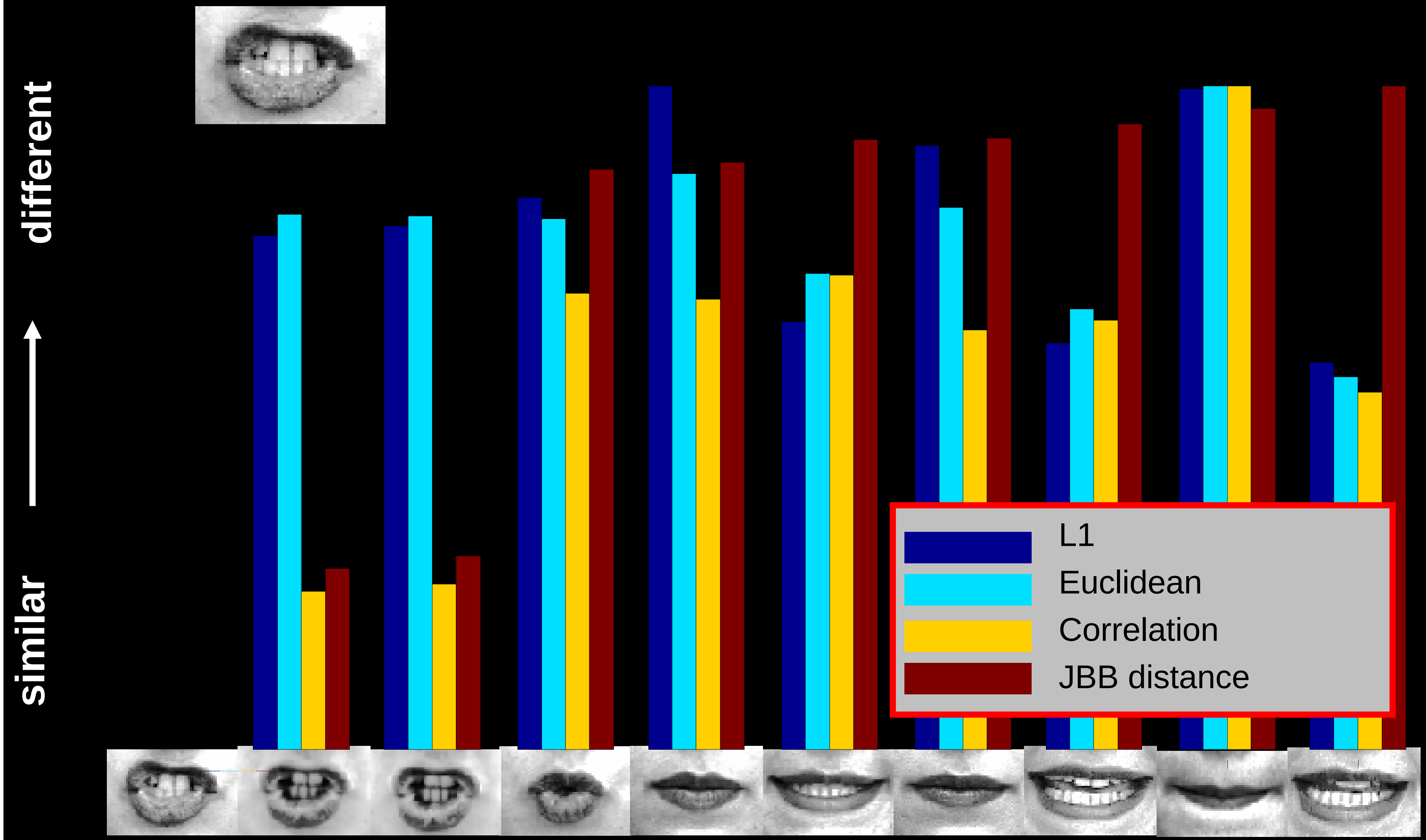
# Distances measurement

- Variation on Jacobs, Belhumeur, Basri (1998)

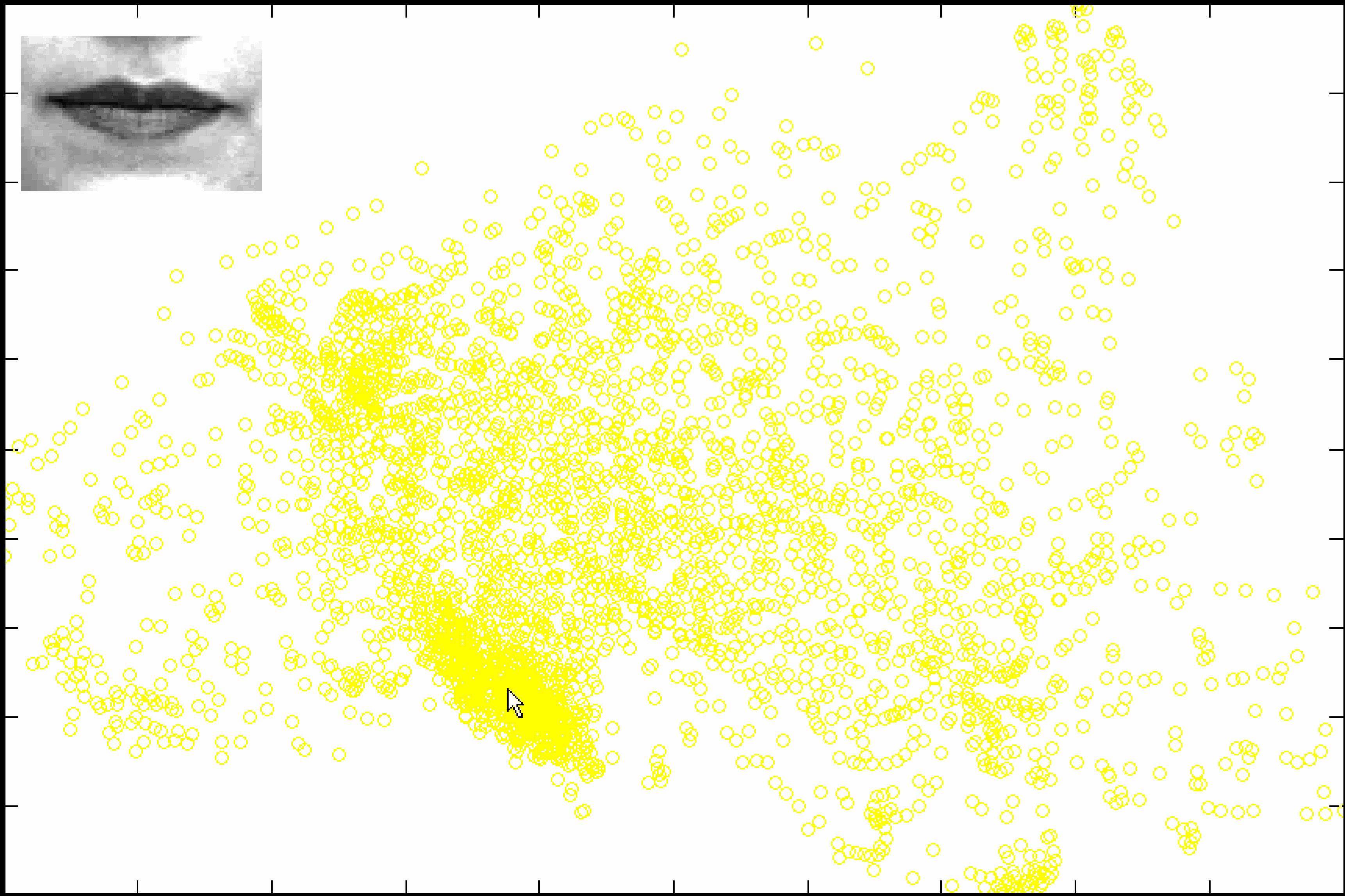
$$E(I,J)=\iint I\cdot J\cdot\left|\nabla\left(\frac{I}{J}\right)\right|\cdot\left|\nabla\left(\frac{J}{I}\right)\right|dxdy$$

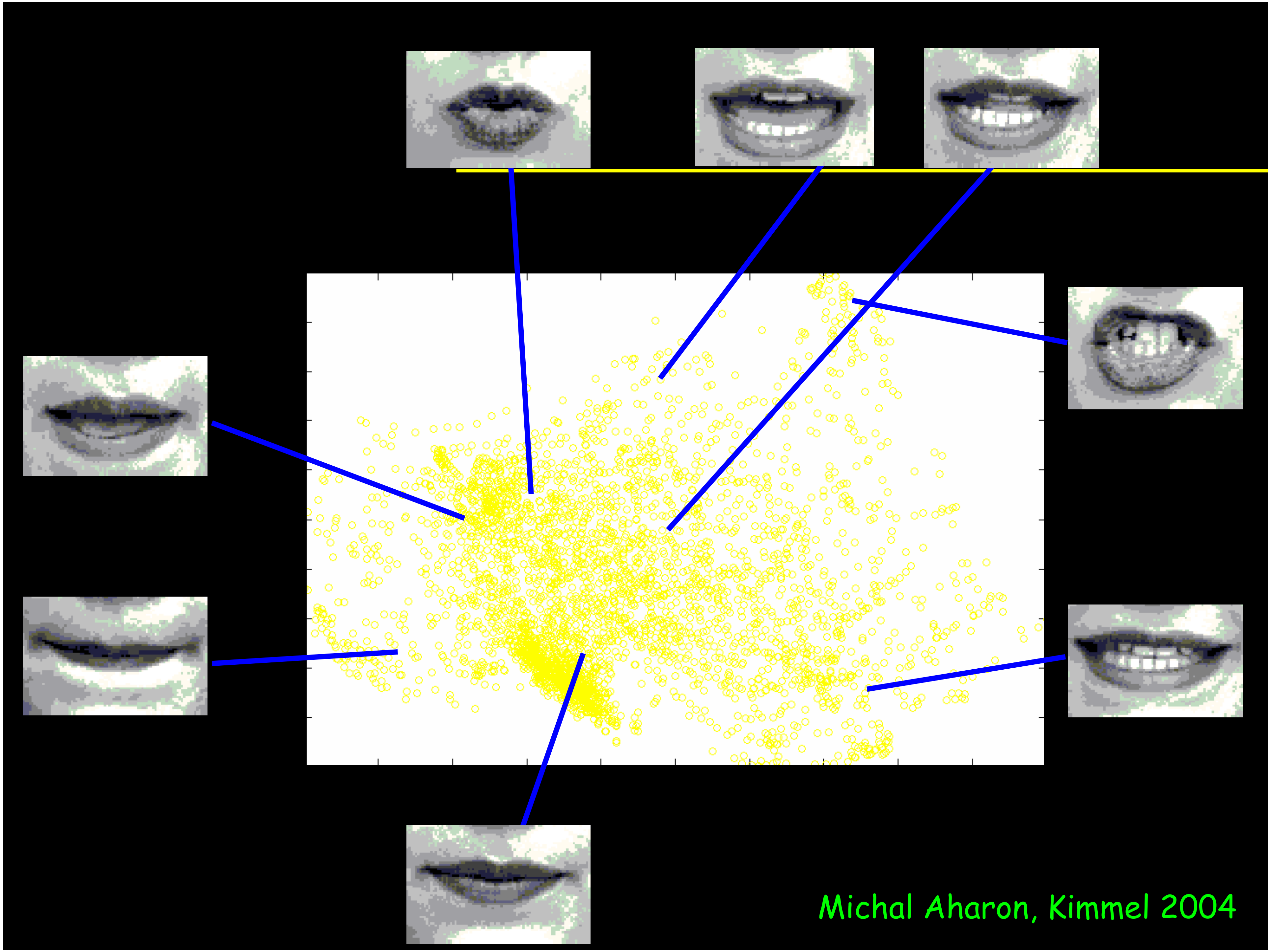
- ~ Invariant to changes in illumination.
- Symmetric consideration.
- Avoid singularities in dark areas.

# Comparing JBB to other measures

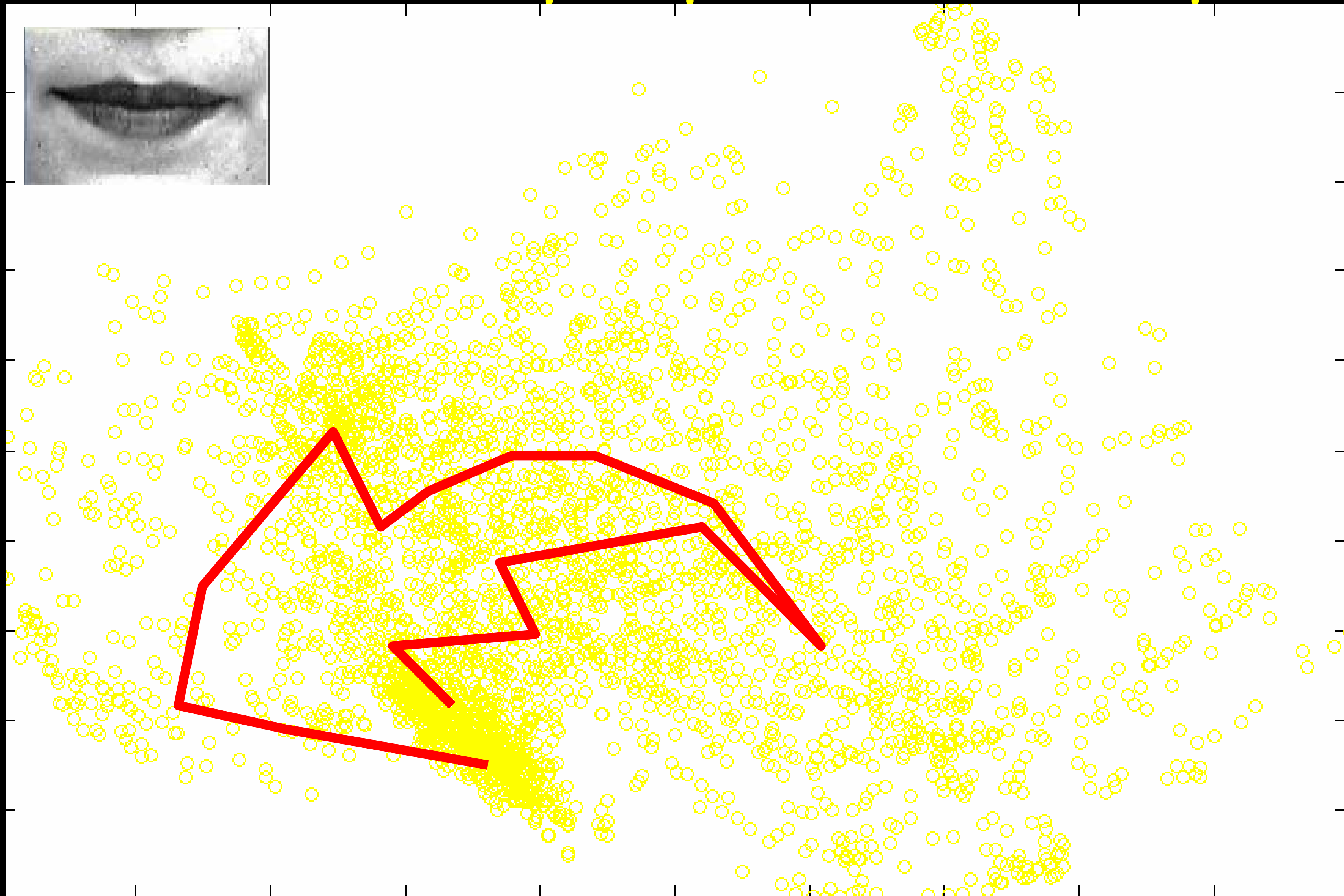


# Lip representation space




Michal Aharon, Kimmel 2004

# Lip representation space



# Speech Synthesis

- Linear interpolation in intensity space is not natural.
- The transition between syllables should be embedded in the lip representation space.

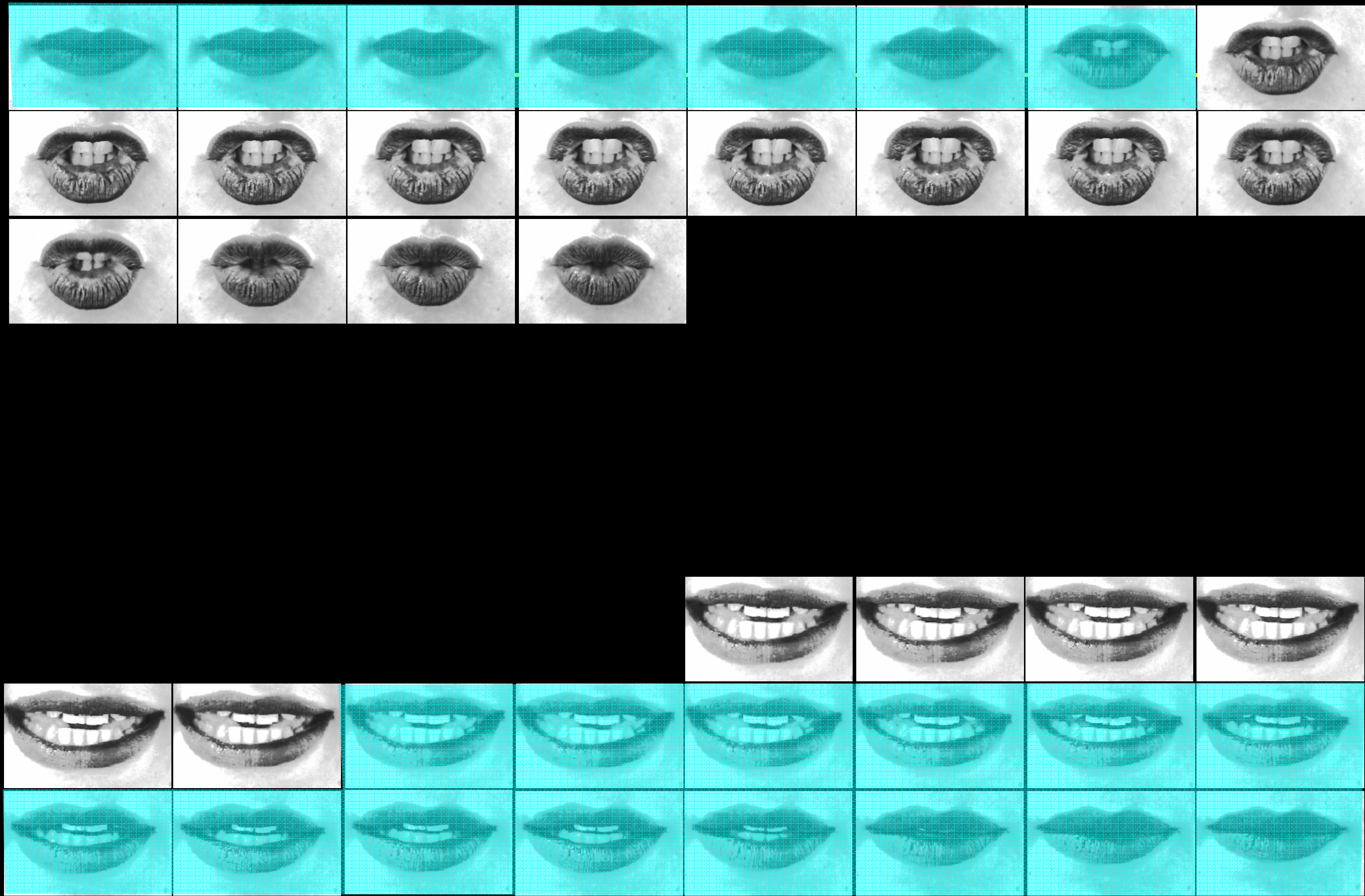
Michal Aharon, Kimmel 2004

# Visual Articulation Signature

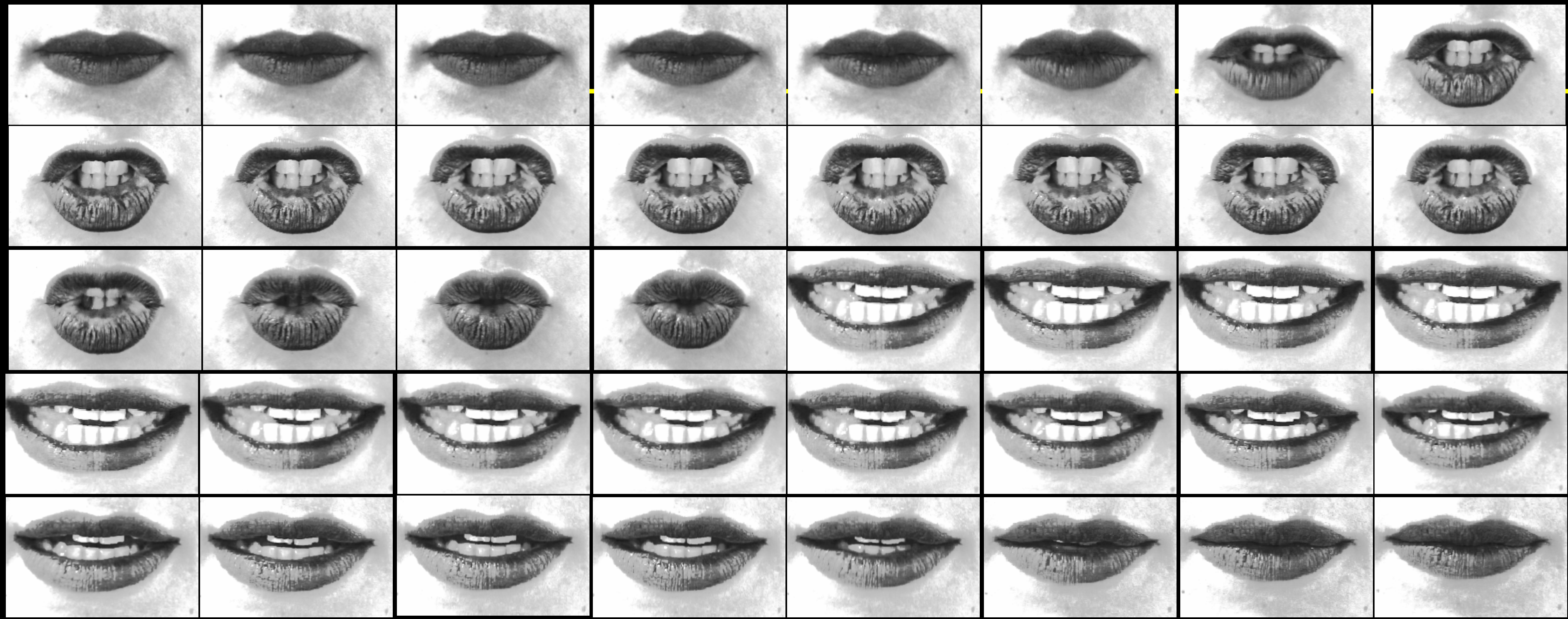
- *Visual articulation signature* (VAS) – the series of mouth configurations that occur in order for a sound to be vocalized.

# Synthesis - SHU+KI

# Synthesis - SHU+KI



# Synthesis - SHU+KI



# Synthesis - Example

- Concatenate the VAS of the syllables that make the word.
- Smooth the transition between each successive VAS.

Shu

Ki

Shuki – simple concatenation

Shuki– smooth transition

# Synthesis - Example

- Concatenate the VAS of the syllables that make the word.
- Smooth the transition between each successive VAS.

Mi

La

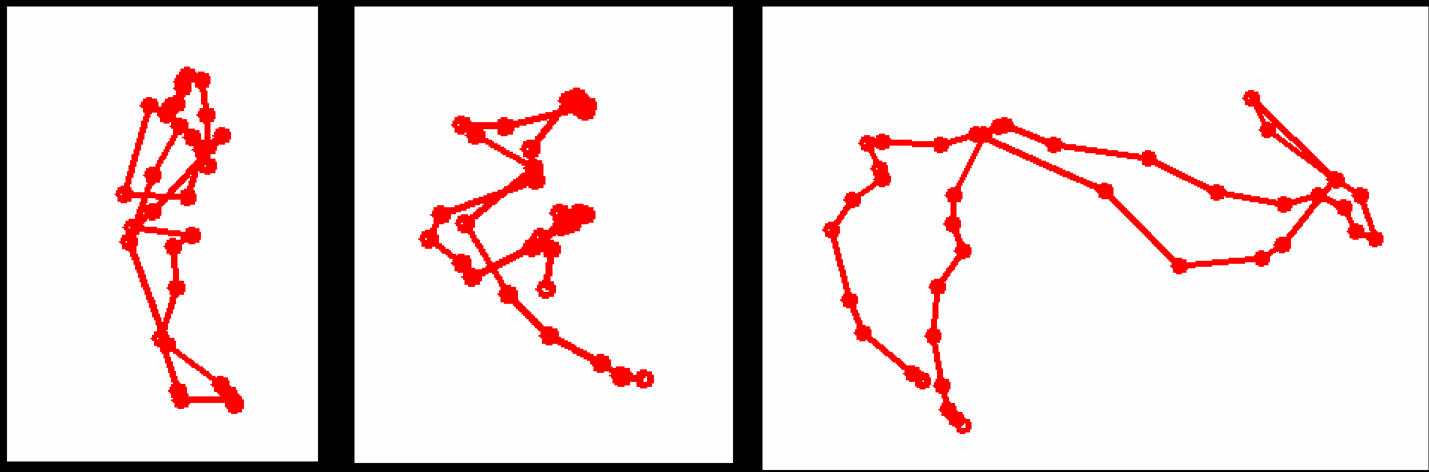
MiLa – simple concatenation

MiLa – smooth transition

# Lip-reading Test

16 different words

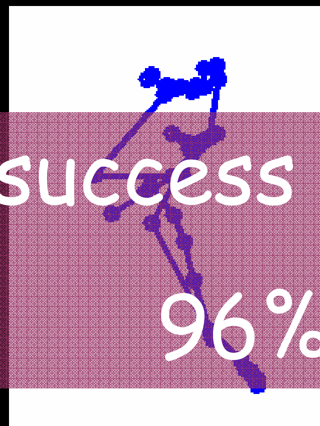
"Coffee" "Cola" "Champagne"



...

The lips area is extracted and a new contour is computed

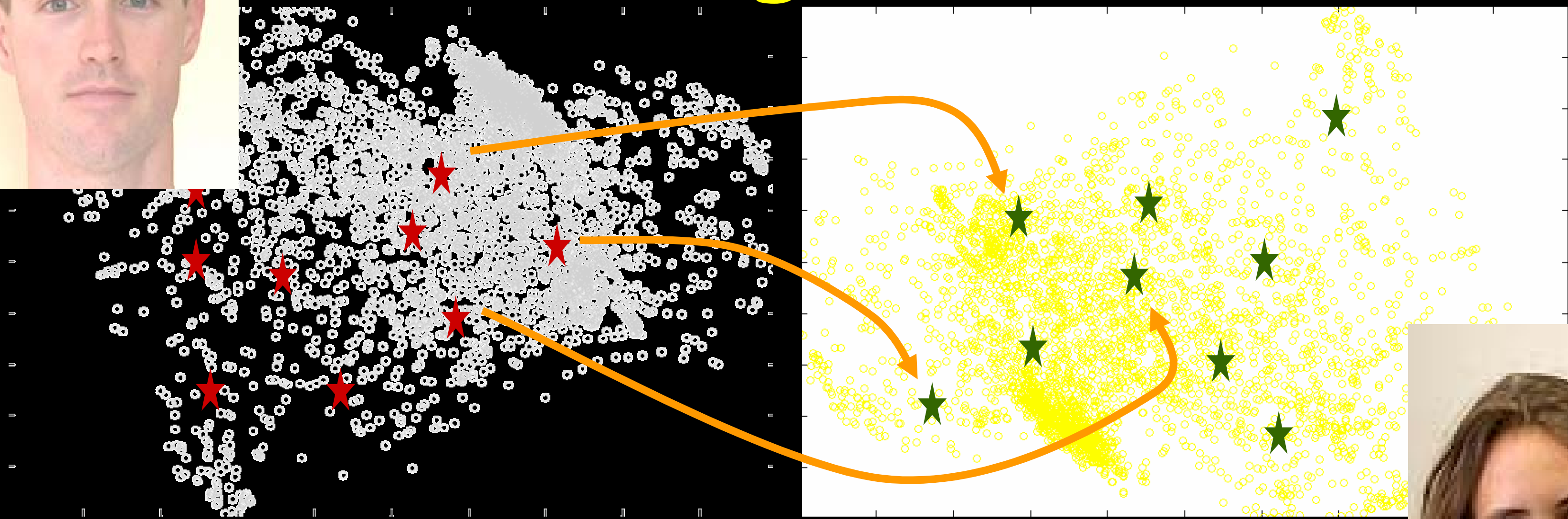
The word contour is then matched to all others.



success rate

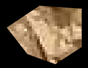
96%

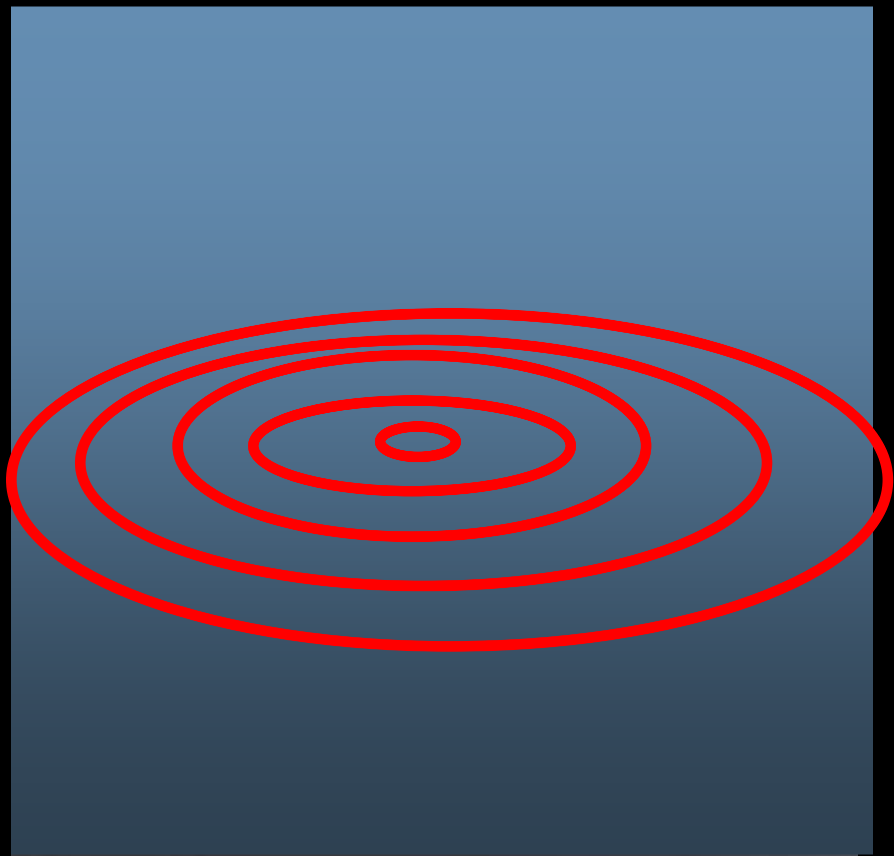
# Generalization: Images as anchors

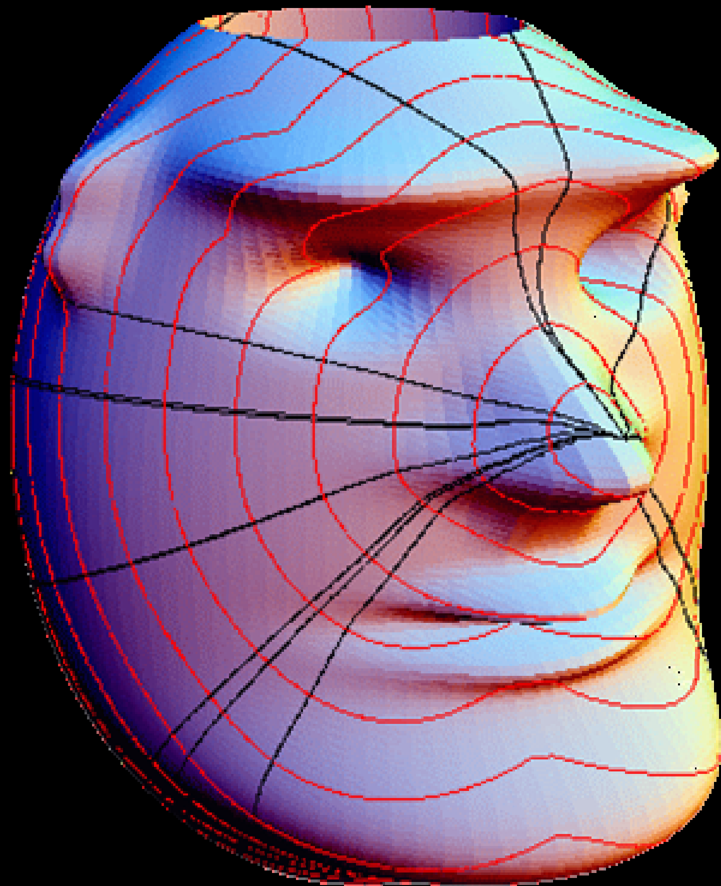


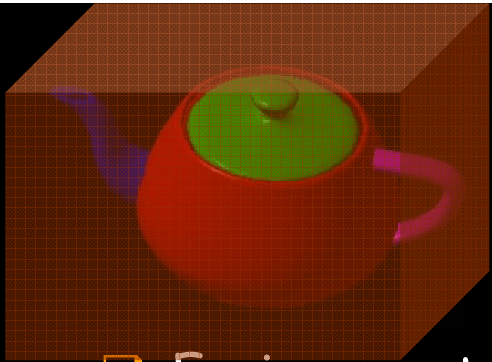

1st hit recognition rate 70%, 2nd hit rate 80%.

# Distance Maps and Minimal Geodesics

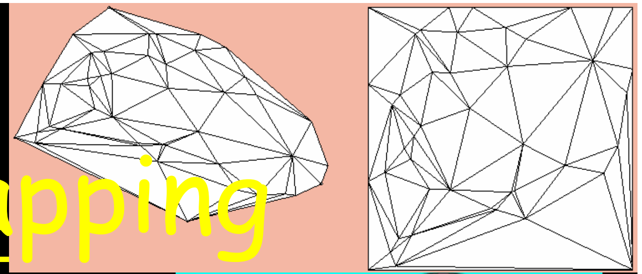


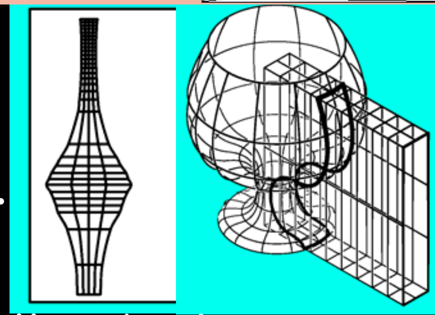






# Texture Mapping





- Environment mapping: Blinn, Newell (76).
- Environment mapping: Greene, Bier and Sloan (86).
- Free-form surfaces: Arad and Elber (97).
- Polyhedral surfaces: Floater (96, 98), Levy and Mallet (98).
- Multi-dimensional scaling: Schwartz, Shaw and Wolfson (89).

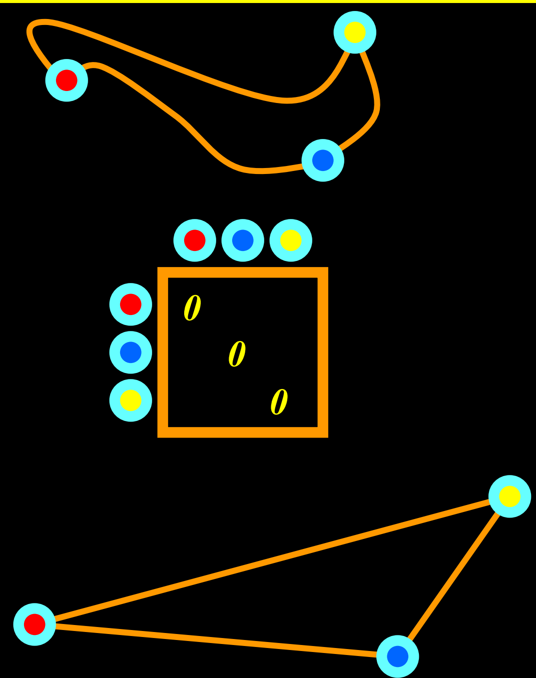
**Difficulties:**

- Need for user intervention.
- Local and global distortions.
- Restrictive boundary conditions.
- High computational complexity.

# Flattening via MDS

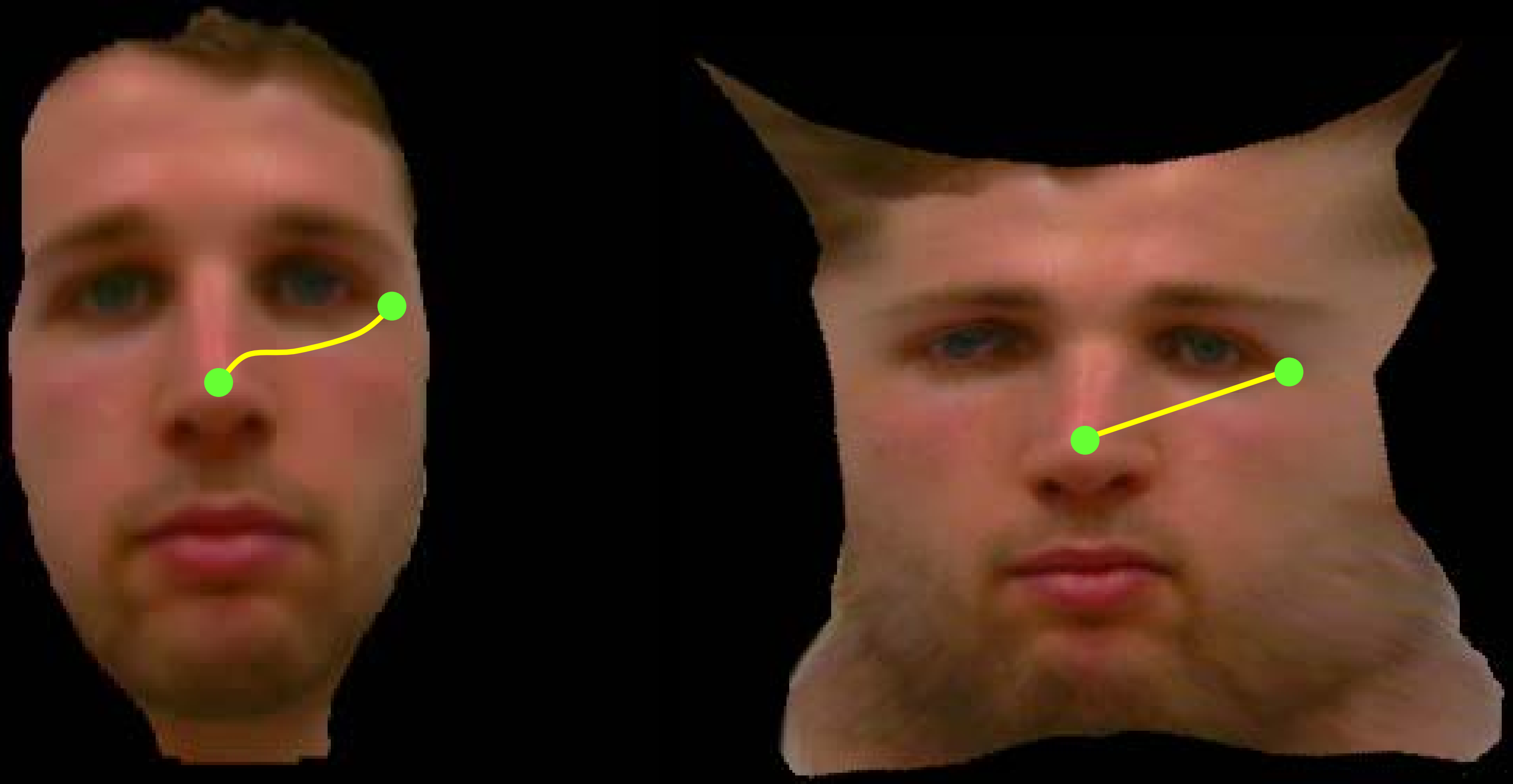
- Compute geodesic distances between pairs of points.
- Construct a square distance matrix of geodesic distances^2.
- Find the coordinates in the plane via multi-dimensional scaling.

  The simplest is `classical scaling'.
- Use the flattened coordinates for texturing the surface, while preserving the texture features.



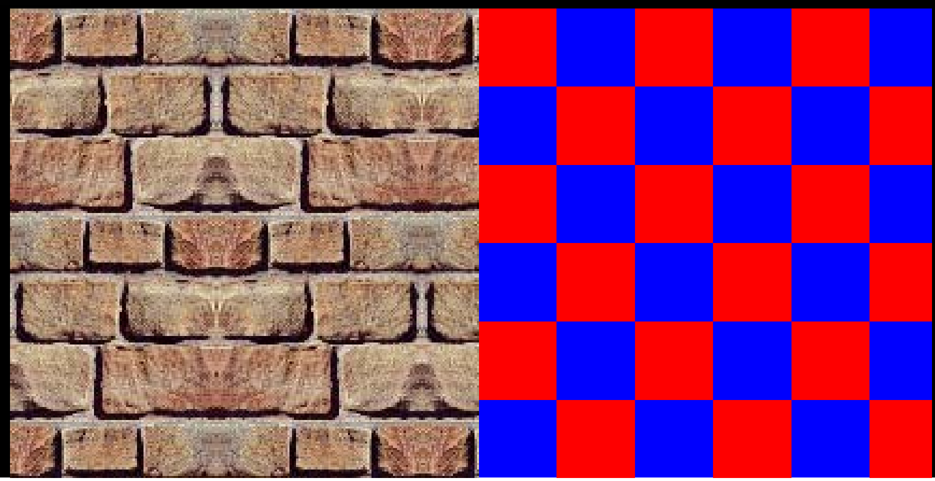

Zigelman, Kimmel, Kiryati, **IEEE T. on Visualization and Computer Graphics** 2002.

# Flattening

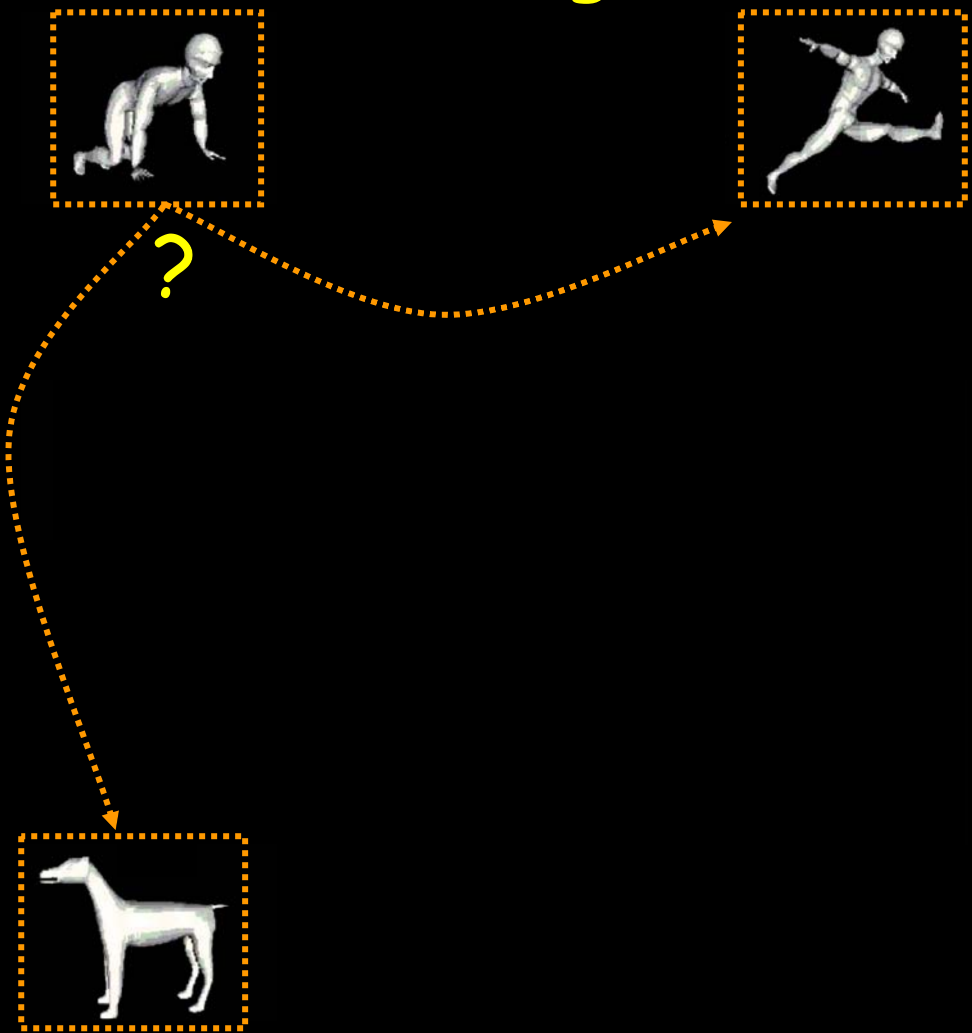


Gil Zigelman

# Flattening

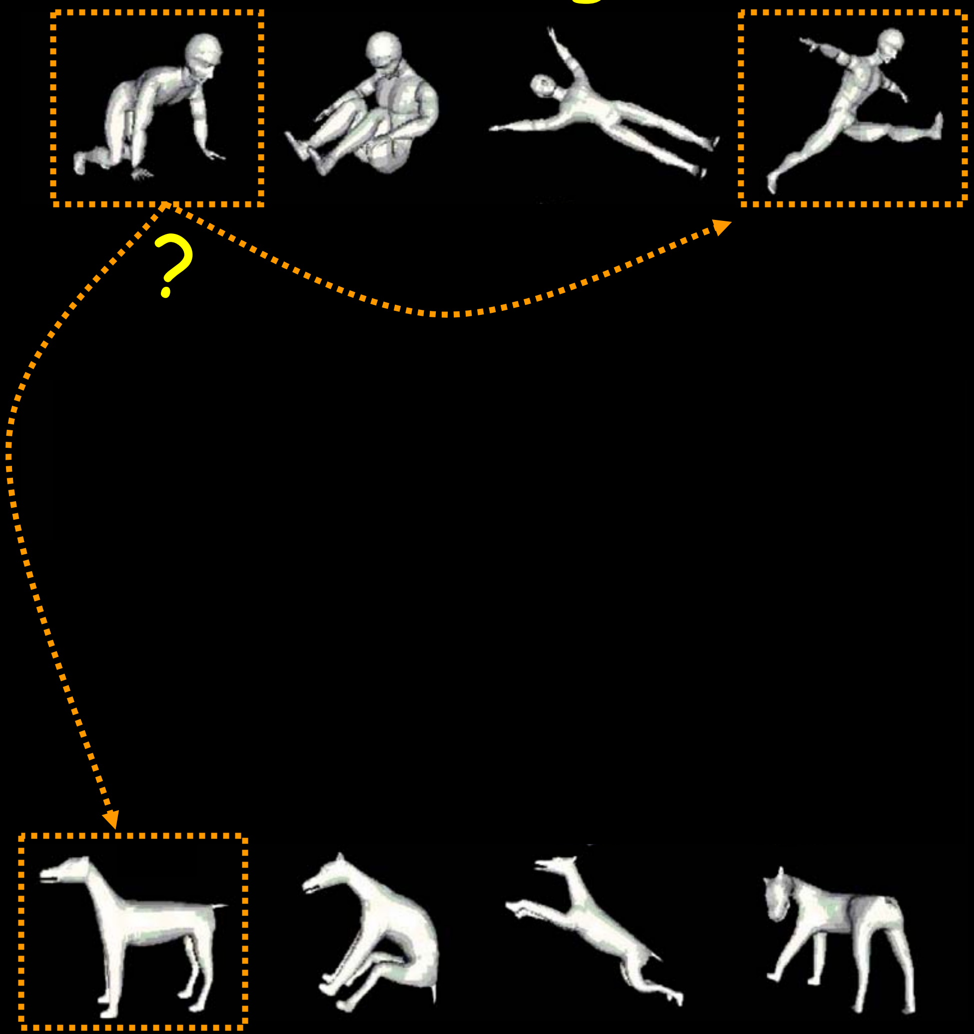


# Bending invariant signatures



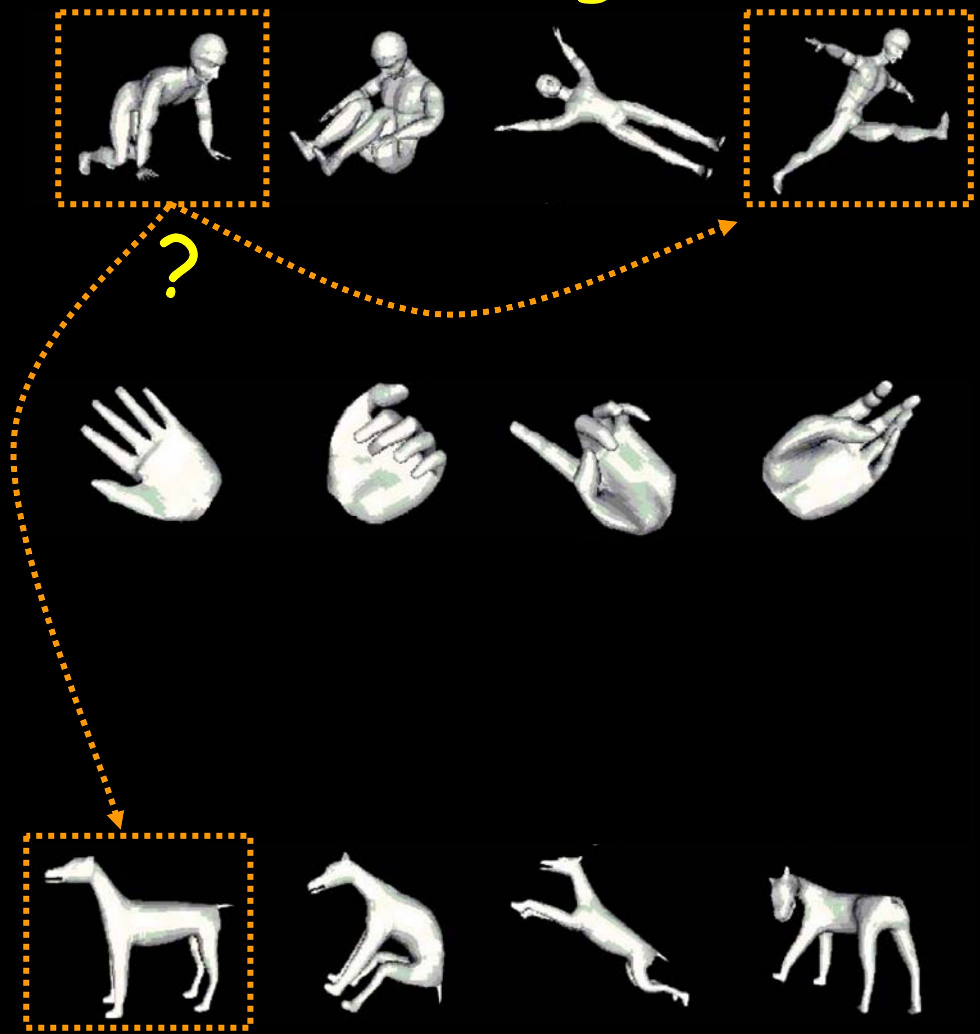

Asi Elad and Kimmel, *CVPR'2001/PAMI'2003*

# Bending invariant signatures




Elad and Kimmel, *CVPR'2001/PAMI'2003*

# Bending invariant signatures

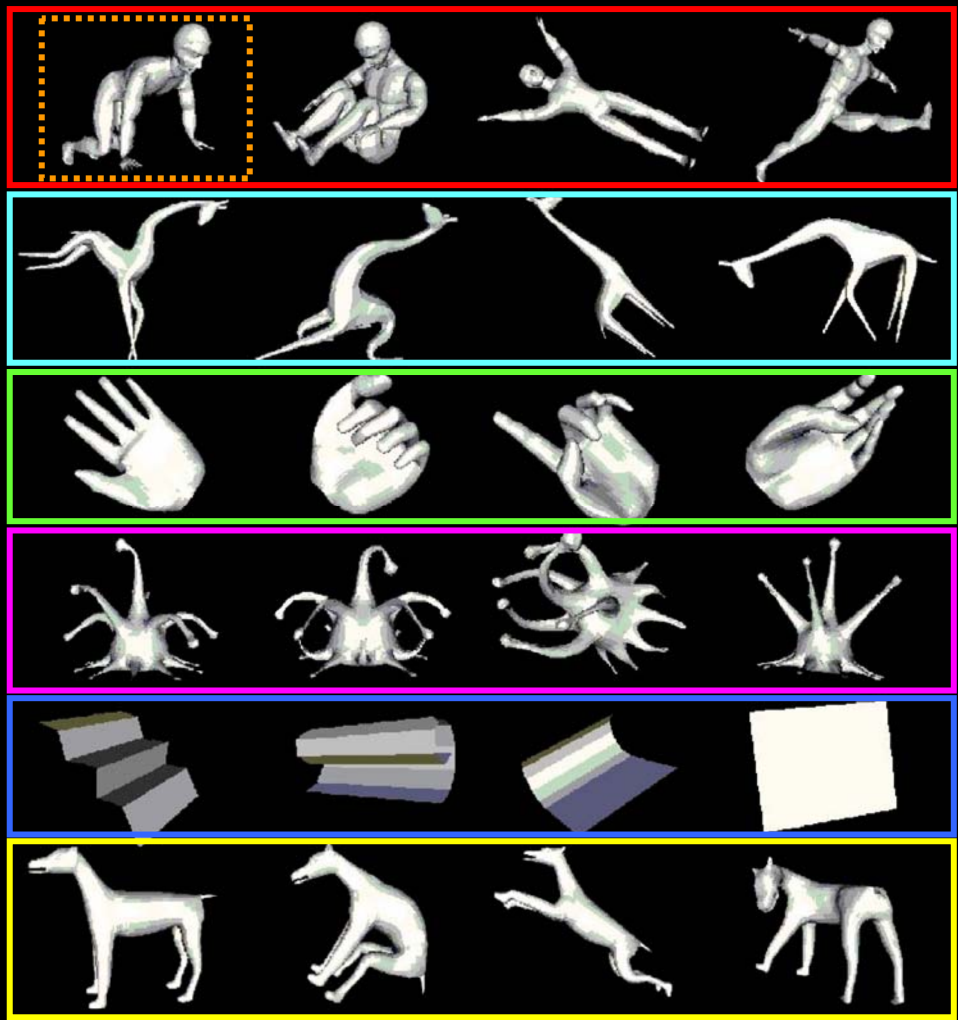


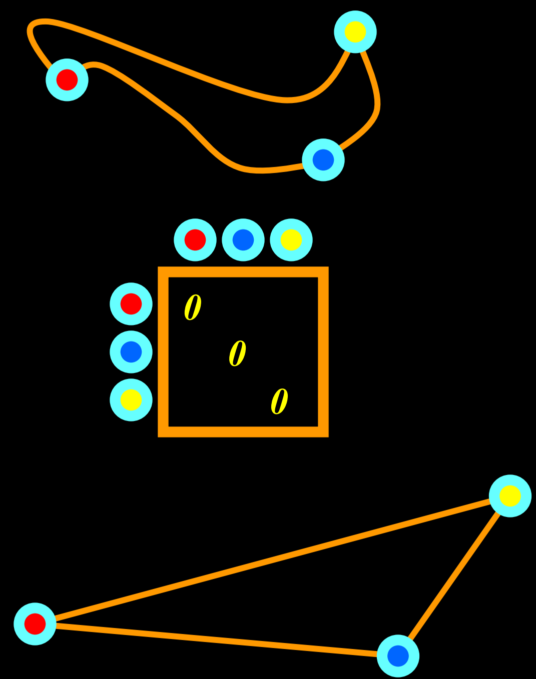

Elad and Kimmel, *CVPR'2001/PAMI'2003*

# Bending Invariant Signatures

Elad and Kimmel, *CVPR'2001/PAMI'2003*

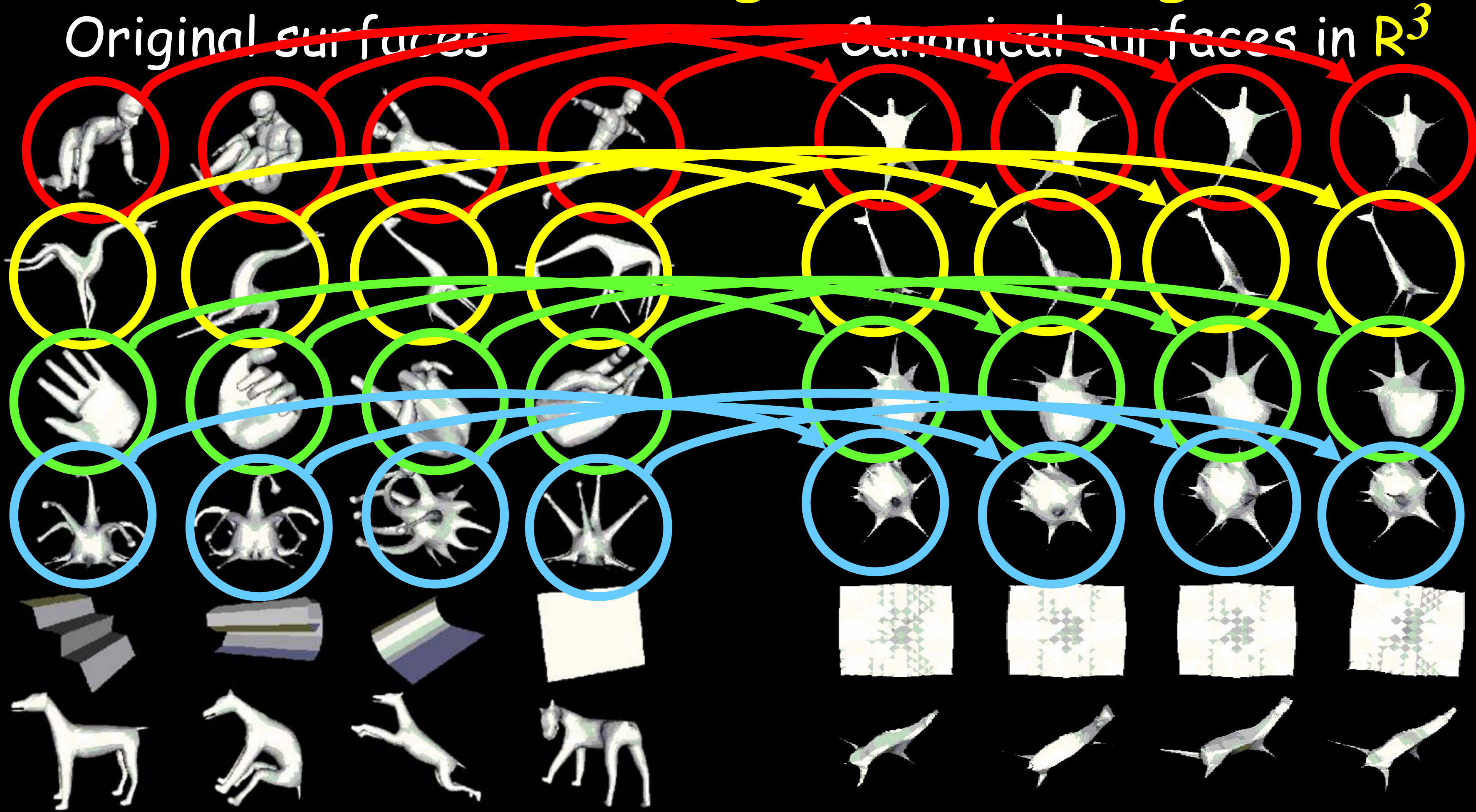
# Bending Invariant Signatures



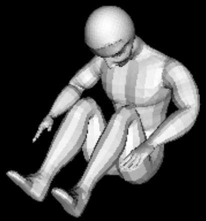


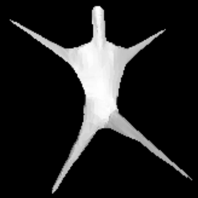

Elad and Kimmel, *CVPR'2001/PAMI'2003*

# Bending invariant signatures



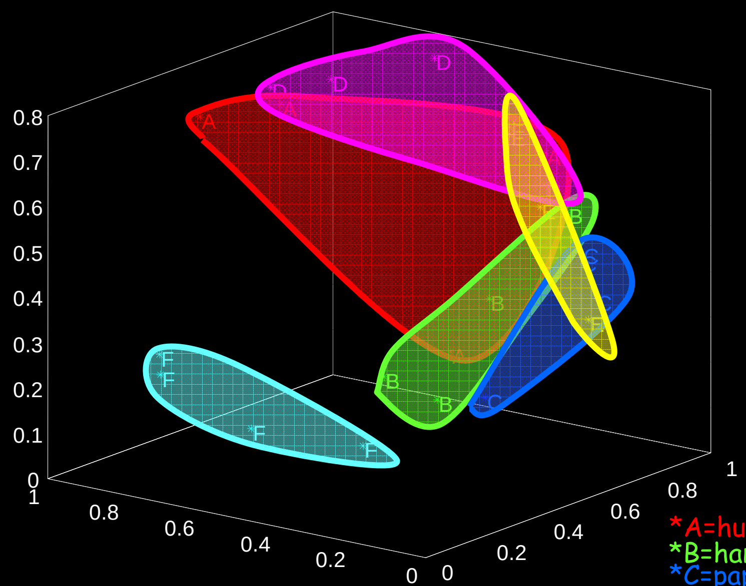

Elad and Kimmel, *CVPR'2001/PAMI'2003*





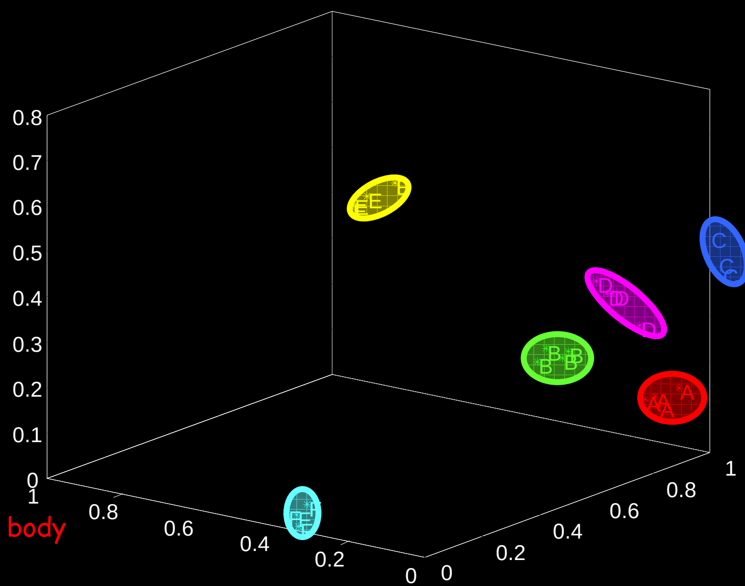
# Classification

Original surfaces



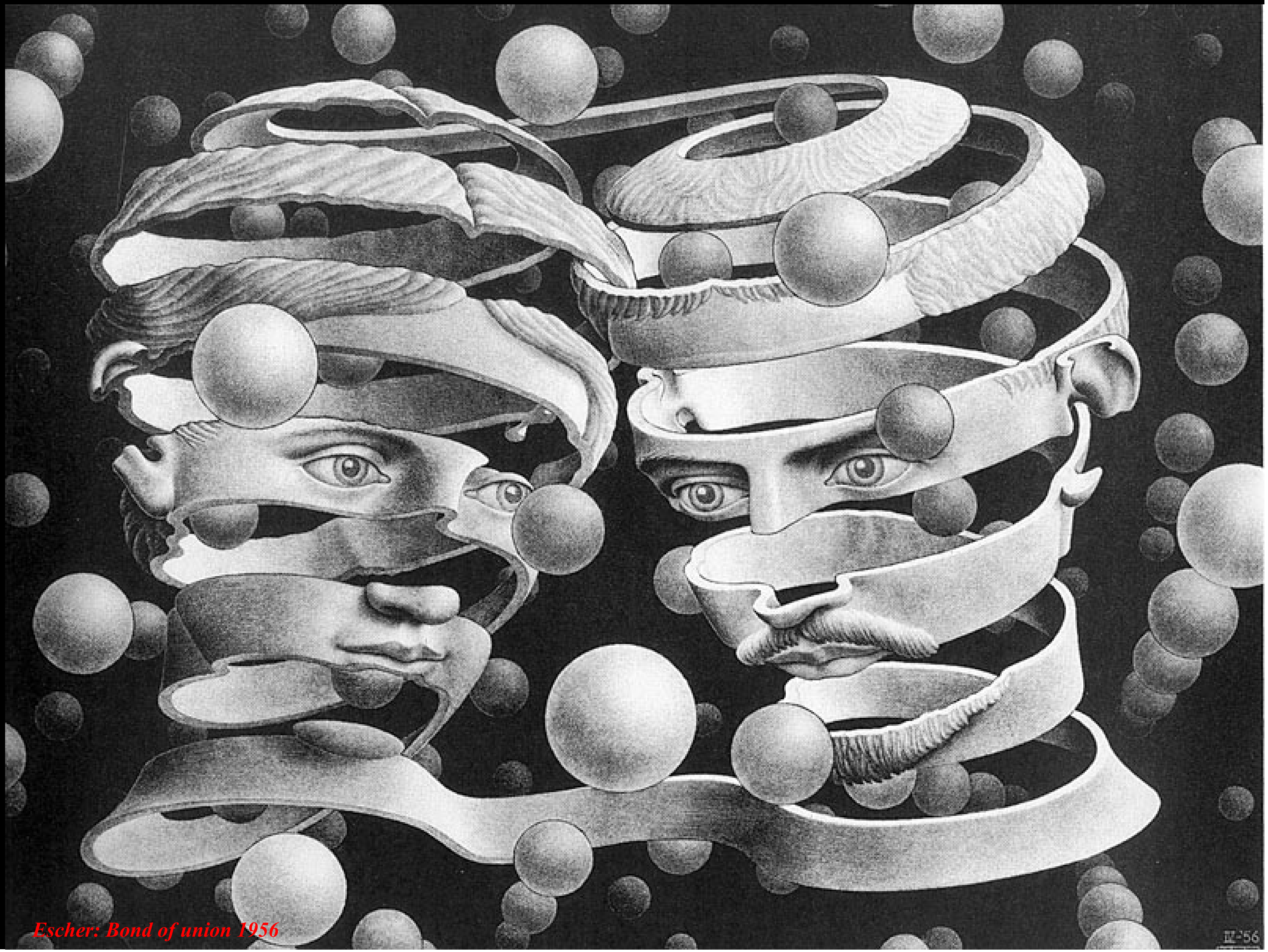


Our signatures



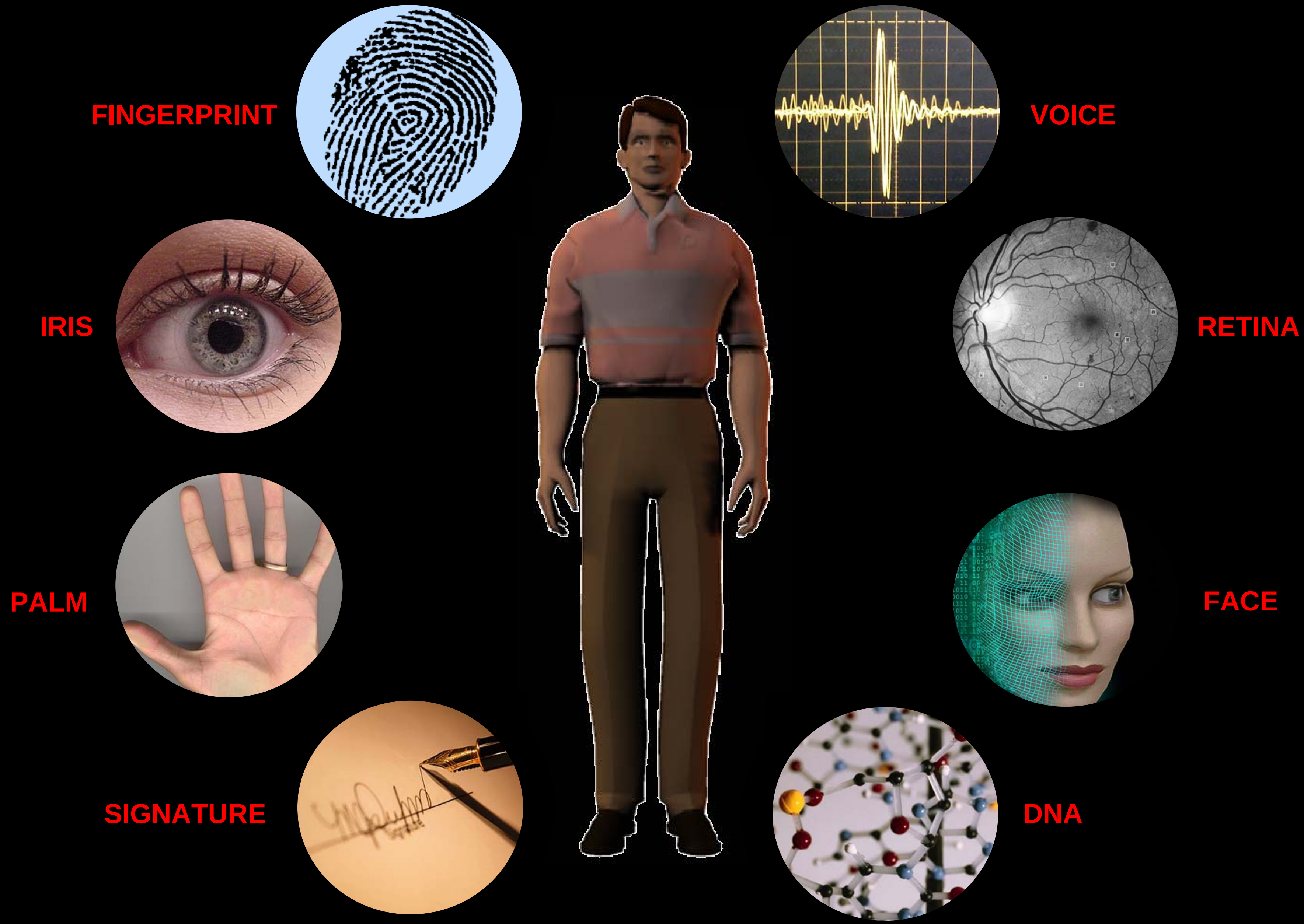
Elad Kimmel *2001*

# Face Recognition



*Escher: Bond of union 1956*

# Some BIOMETRIC Techniques




Weight, Height, Color, Smell, …

# Twins test

who is Michael?

# FACE RECOGNITION: PROBLEMS

**False acceptance:** accept impostors as authenticated persons.

**False rejection:** fail to recognize an authenticated person.

ILLUMINATION and MAKEUP

image = light reflected from the face,
different illuminations yield different images,
thus recognized as different subjects.

FACIAL EXPRESSIONS

Modern face recognition algorithms unable to deal with facial expressions.

**SOLUTION: THREE DIMENSIONAL FACE RECOGNITION**

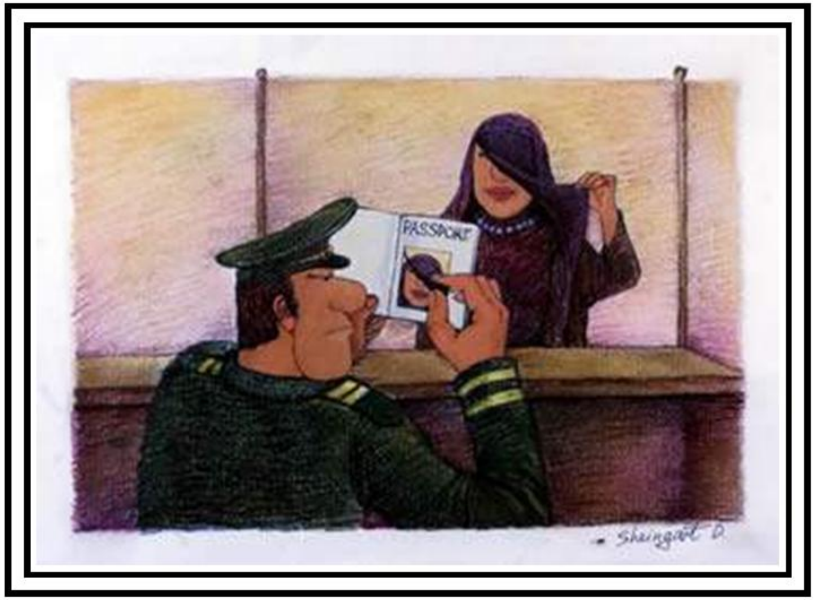
# 2D face recognition

<u>CONNs</u>

- Less accurate than other biometrics
- Sensitive to environment conditions, postures, and expressions.
- Sensitivity to fooling (makeup)

<u>PROs</u>

- No direct contact
- Friendly to users (indirect contact)
- Passive monitoring (surveillance)
- Low cost

# Why 3D?

2D information is sometimes misleading…

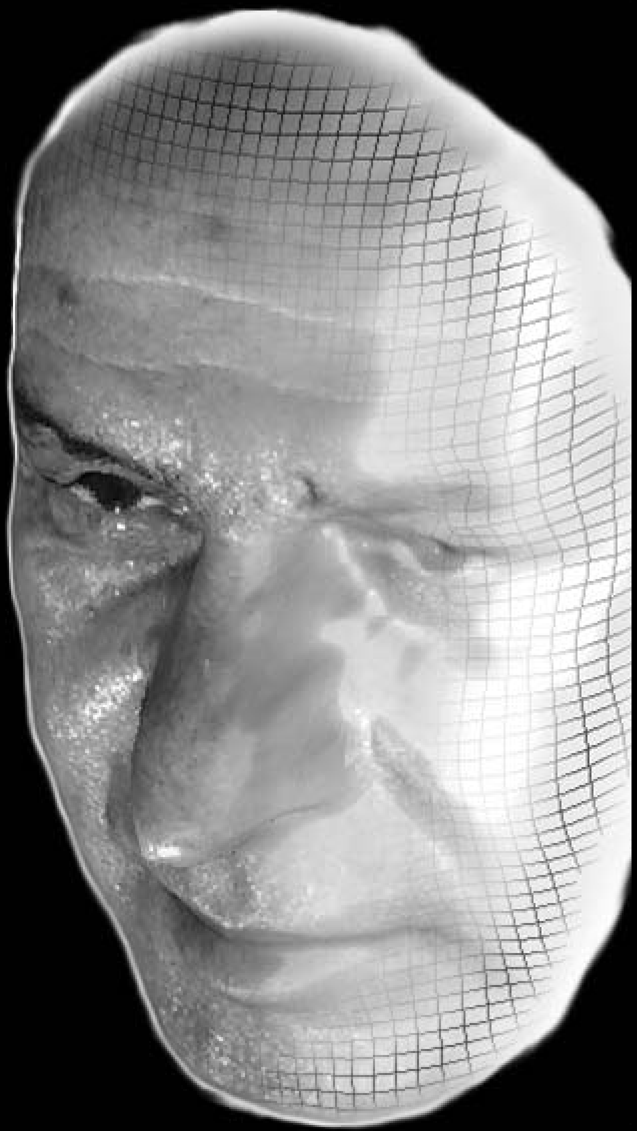
# Why 3D?

...and affected by different factors

Can a 2D system tell that these images are all of the same subject?

# Coded light range video camera

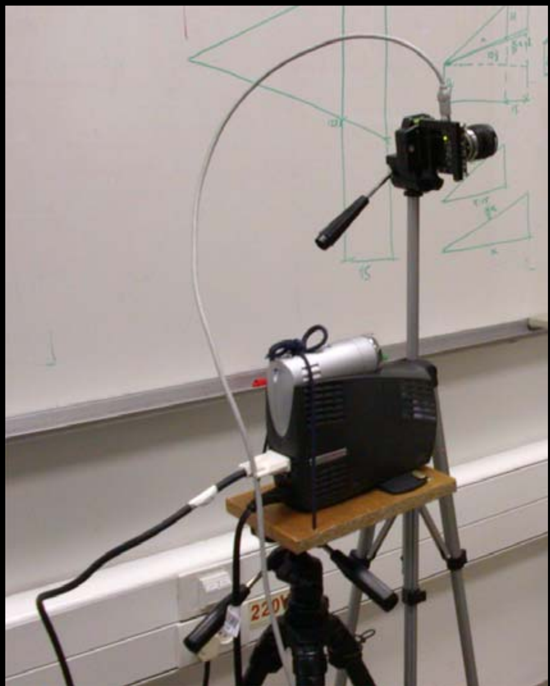
*N* patterns allow angular (~depth) resolution of $2^N$.

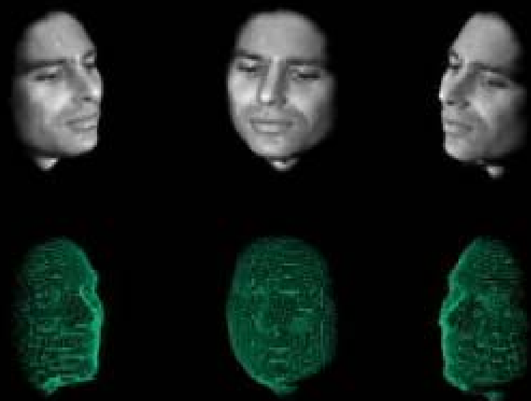


ex-minister of science, Matan Vilnai.

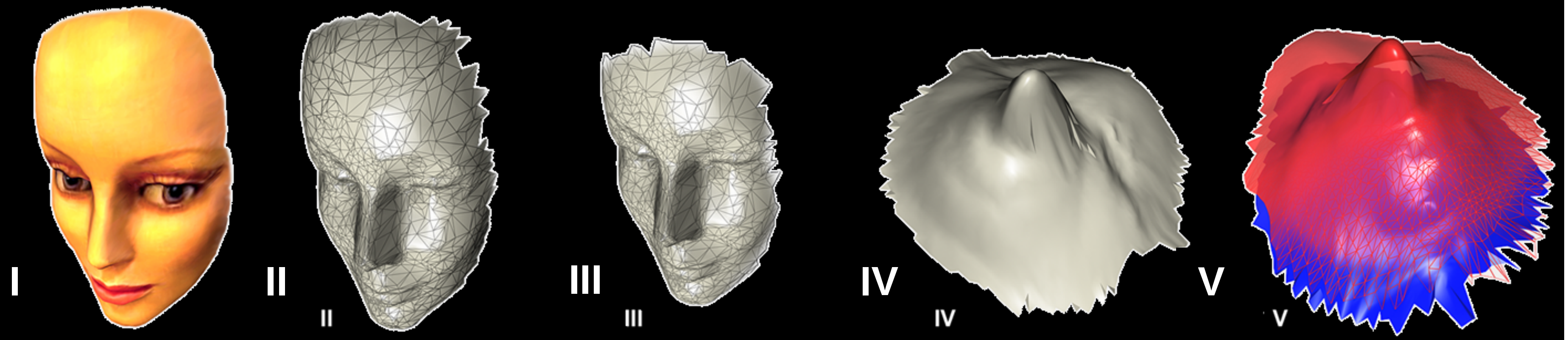
Our new scanner works at less than 150msec,
Components cost ~3k$

# Coded light range video camera





# 3D recognition via geometric invariants

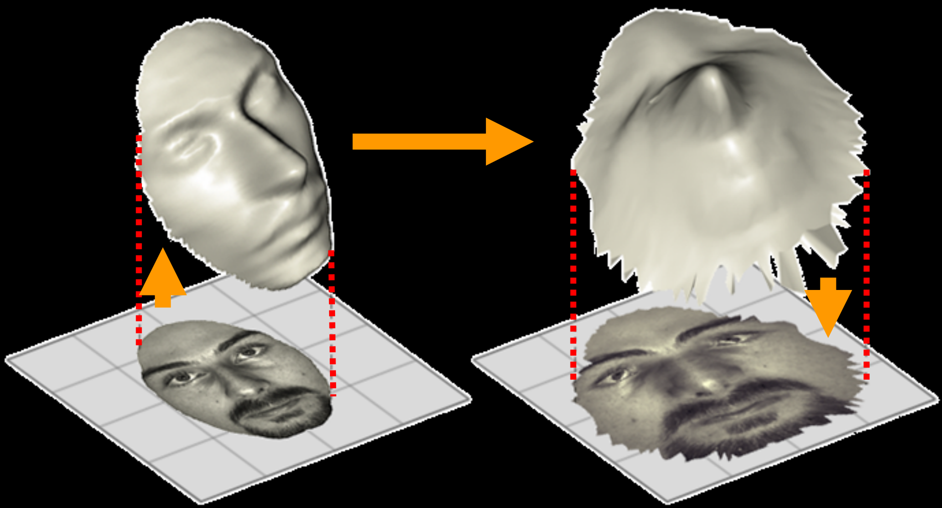



- Range camera acquires facial surface (**I**).
- The surface is smoothed (**II**), subsampled and cropped (**III**).
- Fast marching computes geodesic distances on the surface.
- Facial surface is flattened via MDS (**IV**).
- Rigid surface matching using the canonical surfaces (**V**).

A. Bronstein, M. Bronstein and R. Kimmel, "3D face recognition using geometric invariants", 2003

# Eigenforms

- The training set include: flattened texture + canonical image.
- Applying eigendecomposition to the two sets, we get two eigenspaces.
- The resulting eigenvectors are our eigenforms.

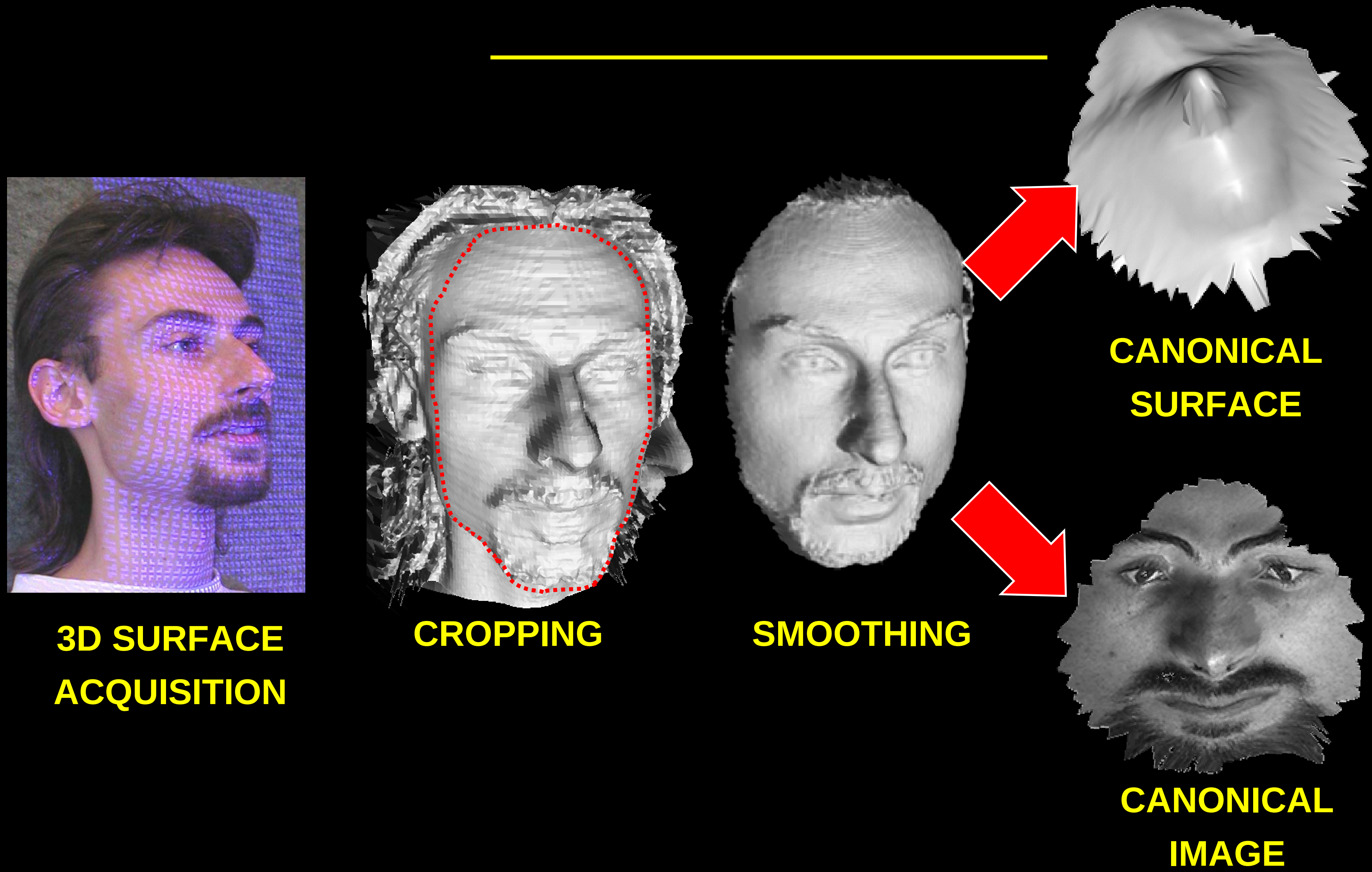


flattened texture

canonical image

A. Bronstein, M. Bronstein and R. Kimmel, “Expression-invariant 3D face recognition“, **AVBPA**, 2003

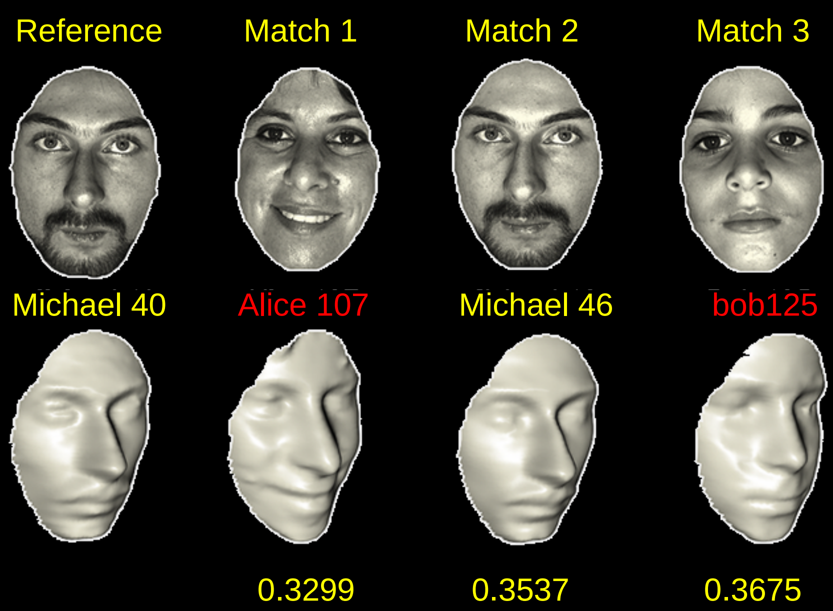
# 3D FACE RECOGNITION STAGES



# TWINS TEST I

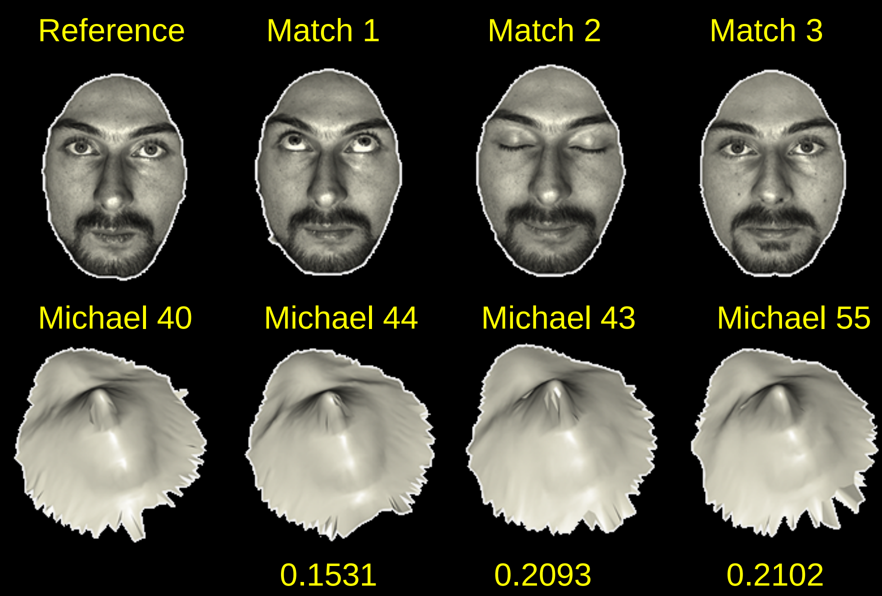
- Recognizing **twins**, a challenging test for face recognition.

SURFACE MATCHING




Facial surfaces as rigid objects -> inaccurate.

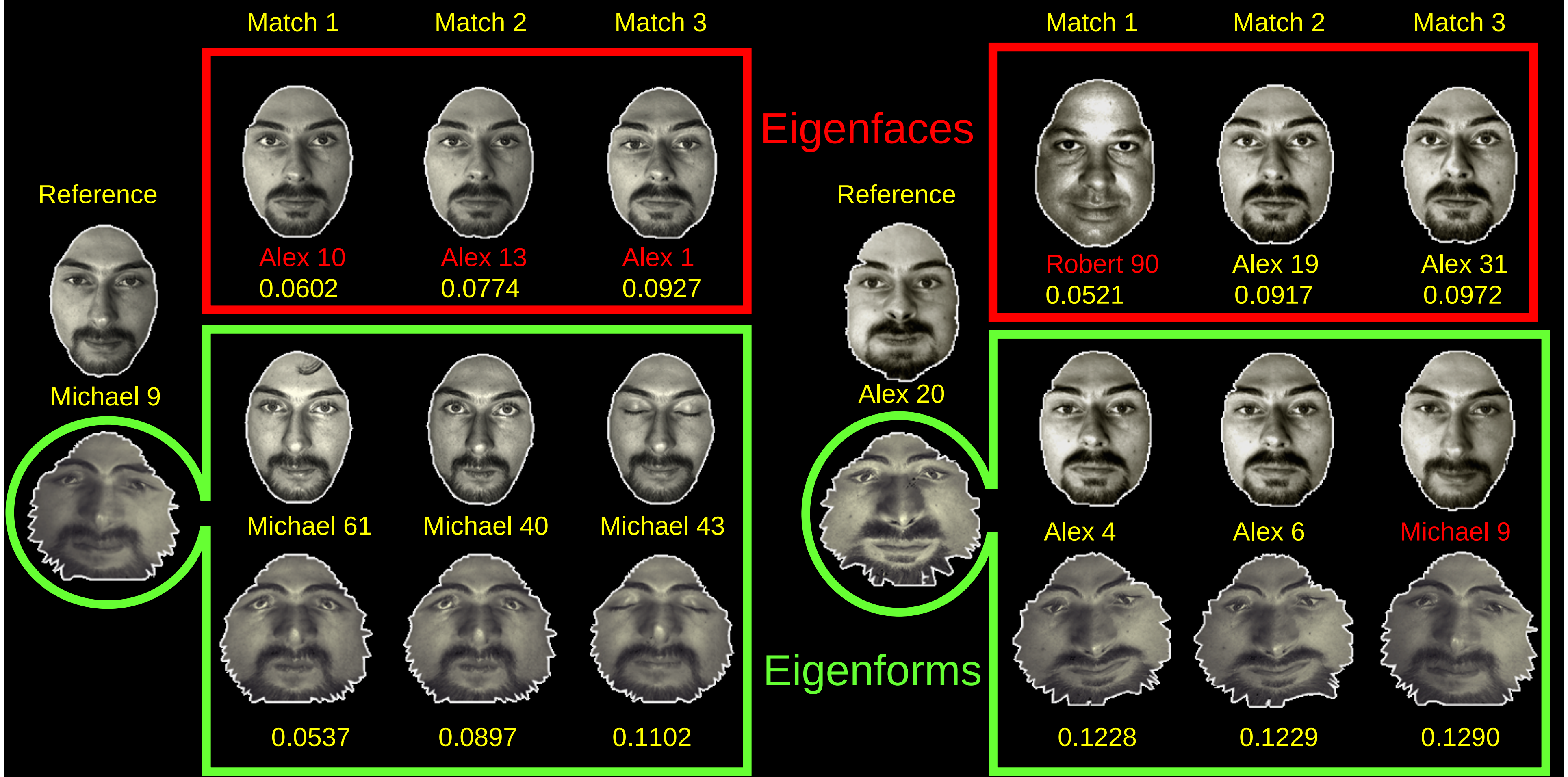
CANONICAL FORM MATCHING




Canonical forms tell apart identical twins.

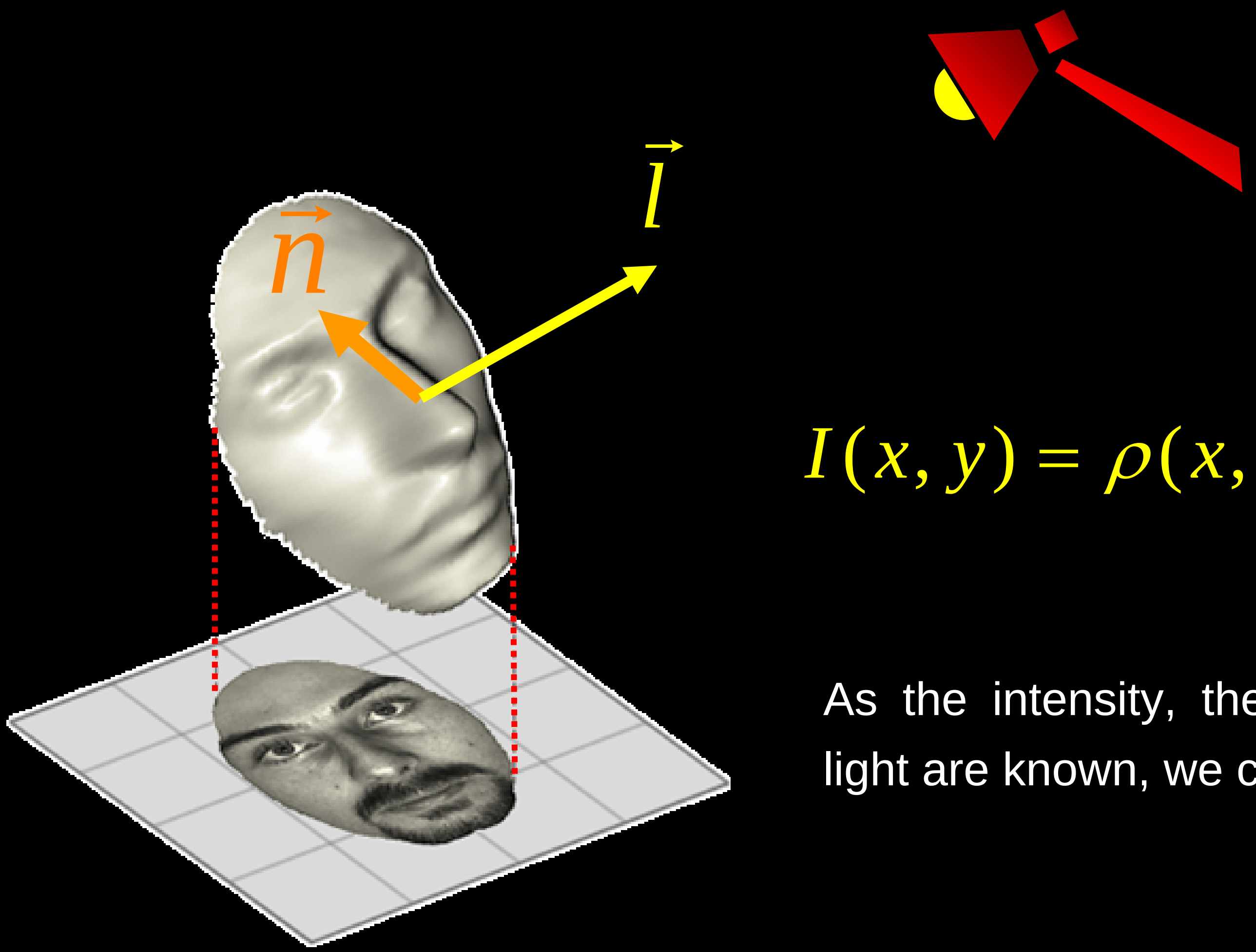
A. Bronstein, M. Bronstein and R. Kimmel, 2003

# Twins test II




A. Bronstein, M. Bronstein and R. Kimmel, "3D face recognition using geometric invariants", 2003

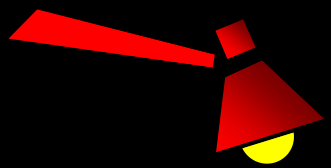
# Invariance to the effects of pose on the intensity



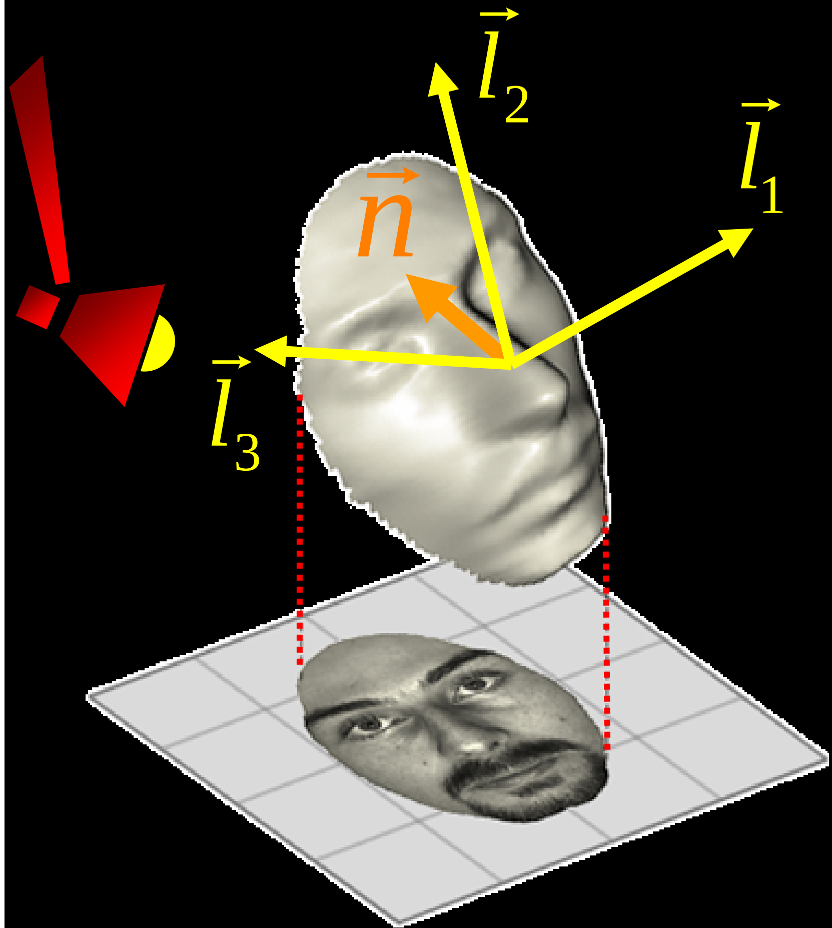

$$I(x,y) = \rho(x,y)\left\langle \vec{l}, \vec{n}(x,y) \right\rangle$$

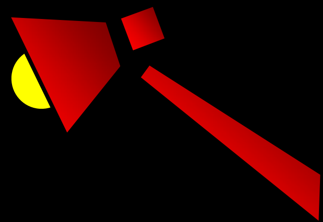
As the intensity, the normal (surface), and light are known, we can solve for the albedo

A. Bronstein, M. Bronstein, E. Gordon, and R. Kimmel, 2003



# Do we need the facial surface?



$$I_1(x,y) = \rho(x,y)\left\langle \vec{l}_1, \vec{n}(x,y) \right\rangle$$

$$I_2(x,y) = \rho(x,y)\left\langle \vec{l}_2, \vec{n}(x,y) \right\rangle$$

$$I_3(x,y) = \rho(x,y)\left\langle \vec{l}_3, \vec{n}(x,y) \right\rangle$$

Photometric Stereo:

- Compute the normal from 3 images: same camera different light sources.
- The normal is enough for all computations.
- No need to integrate the surface using Poisson solvers.

A. Bronstein, M Bronstein, R. Kimmel and A. Spira., ECCV 2004

# In the news

- CNN-news, W-NBC, Rueters, Washington-Post, Channel-2, Haaretz, Maariv, Yahoo-news, SIAM news, and more than 60 other TV and newspapers all over the world…

CNN.com/TECHNOLOGY

SEARCH The Web CNN.com Search ENHANCED BY Google

Home Page
World
U.S.
Weather
Business at CNNMoney
Sports at SI.com
Politics
Law
Technology
Science & Space
Health
Entertainment
Travel
Education
Special Reports

Visit the international edition. Click here

SERVICES
Video
Newswatch
E-Mail Services
CNN To Go

SEARCH
Web CNN.com
Search ENHANCED BY Google

# Twins crack face recognition puzzle

Monday, March 10, 2003 Posted: 9:51 AM EST (1451 GMT)

**HAIFA, Israel (Reuters) -- For a fleeting moment, Mohamed Atta appeared on an airport security camera minutes before he boarded one of the planes which crashed into the World Trade Center on September 11, 2001.**

Was there any way the camera or its operator would have been able to identify Atta as a suspect before he hijacked and flew the first of two planes into the twin towers?

Israelis Michael and Alex Bronstein think they have the answer.

Story Tools
SAVE THIS
EMAIL THIS
PRINT THIS
MOST POPULAR

The computer whiz-kids -- 22-year-old identical twins almost impossible to tell apart -- have applied a new technology to recognizing faces in a way that may yet revolutionize international security.

"I said it to them as a joke: if you succeed in building a system that can distinguish between the two of you, you'll get (a grade of) 100," said the twins' professor, Ron Kimmel of the Technion Institute in Haifa.

"They succeeded and got 100. They are brilliant."

The technology scans and maps the human face as a three-dimensional surface, providing a far more accurate reference for identifying a person than current systems, most of which rely on two-dimensional images, Kimmel said.

The product can potentially meet a wide range of security needs in a world shaken by the September 11 attacks and a series of bombings blamed on Osama Bin Laden's al Qaeda network, of which Atta was a suspected member.

Kimmel and one of his former pupils, Assi Elad, had already developed the algorithms used as building-blocks for the face-recognition system. The

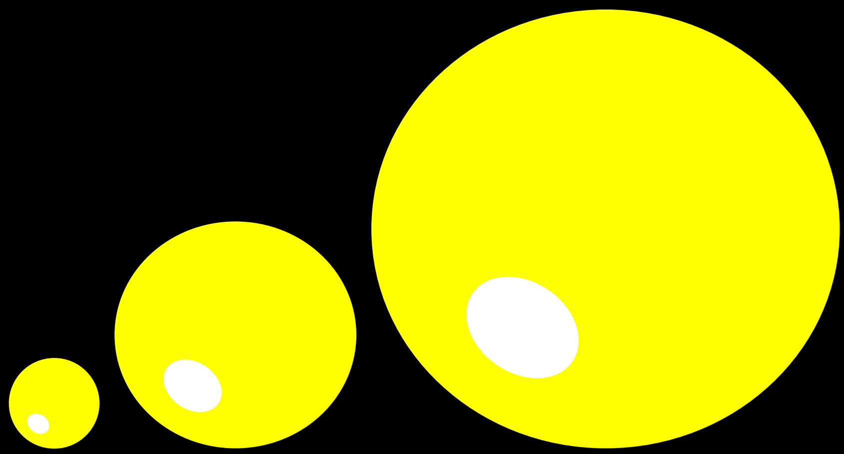


הארץ Online העיתון המודפס

מצאו את ההבדלים

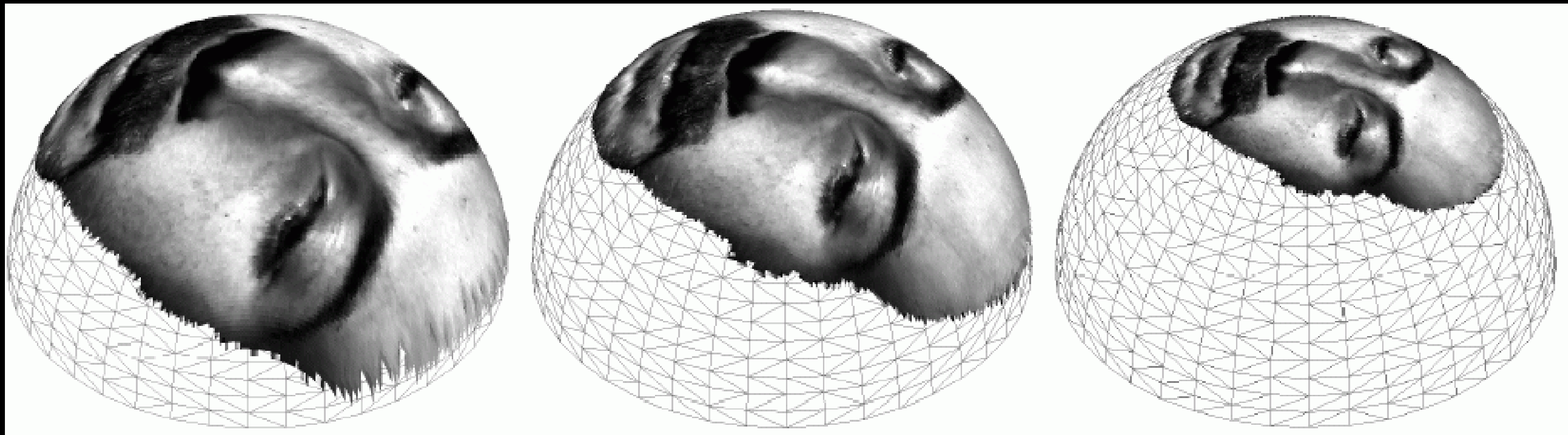
# Why flat embedding?

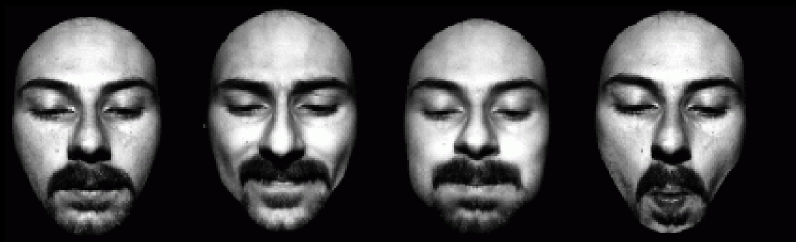
- Is there another space that better captures the face intrinsic geometry yet still provides the convenient comparison property of a flat embedding space?

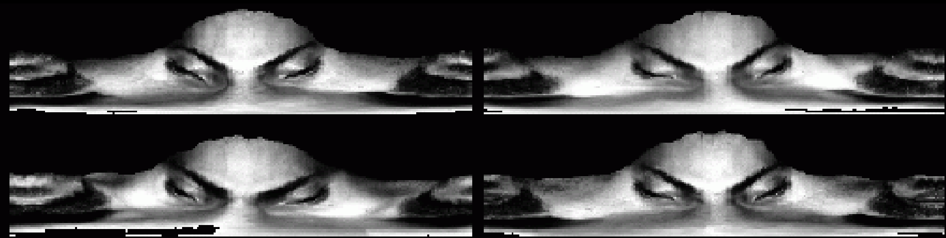


$$S^n$$

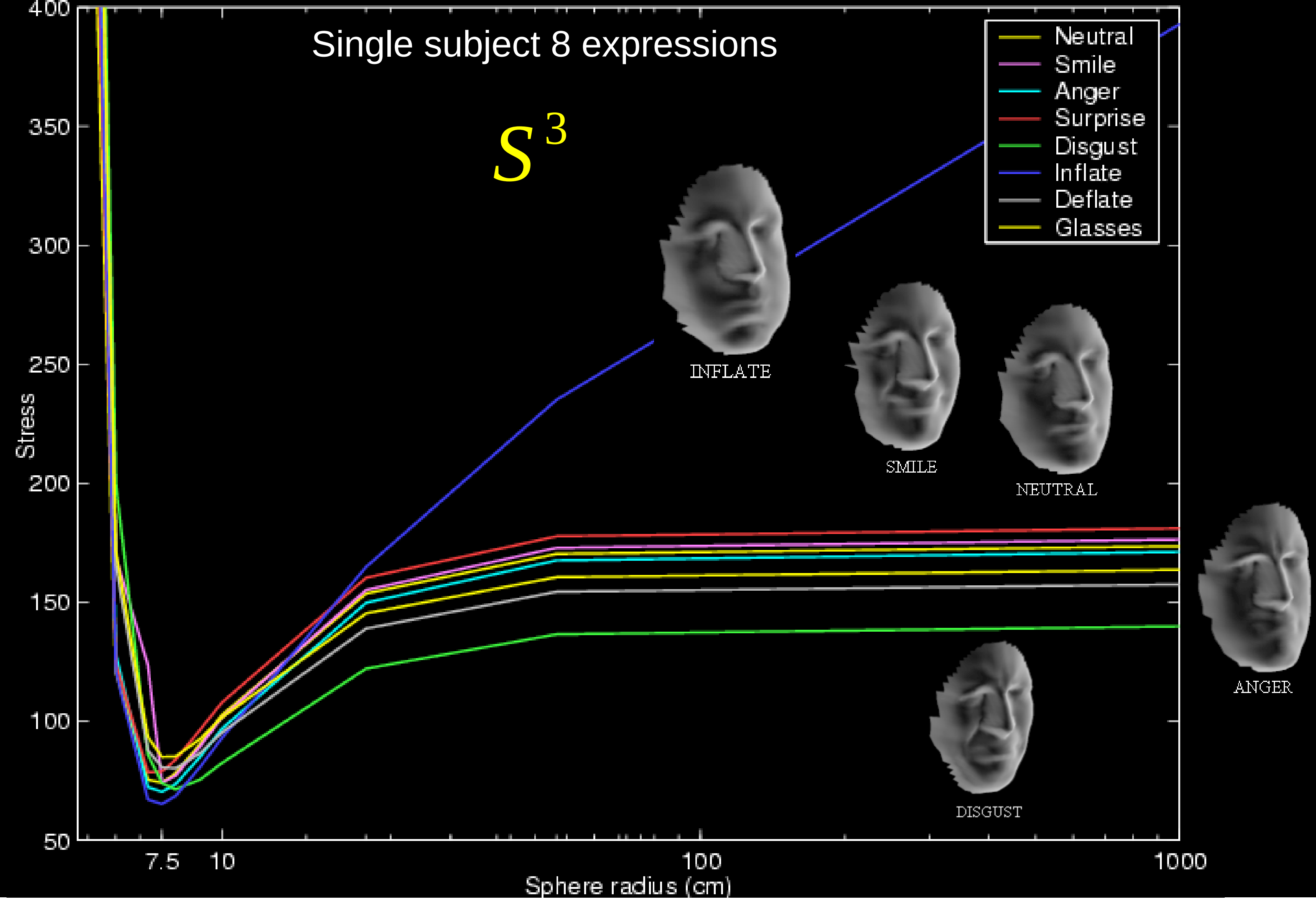
# Embedding in $S^2$



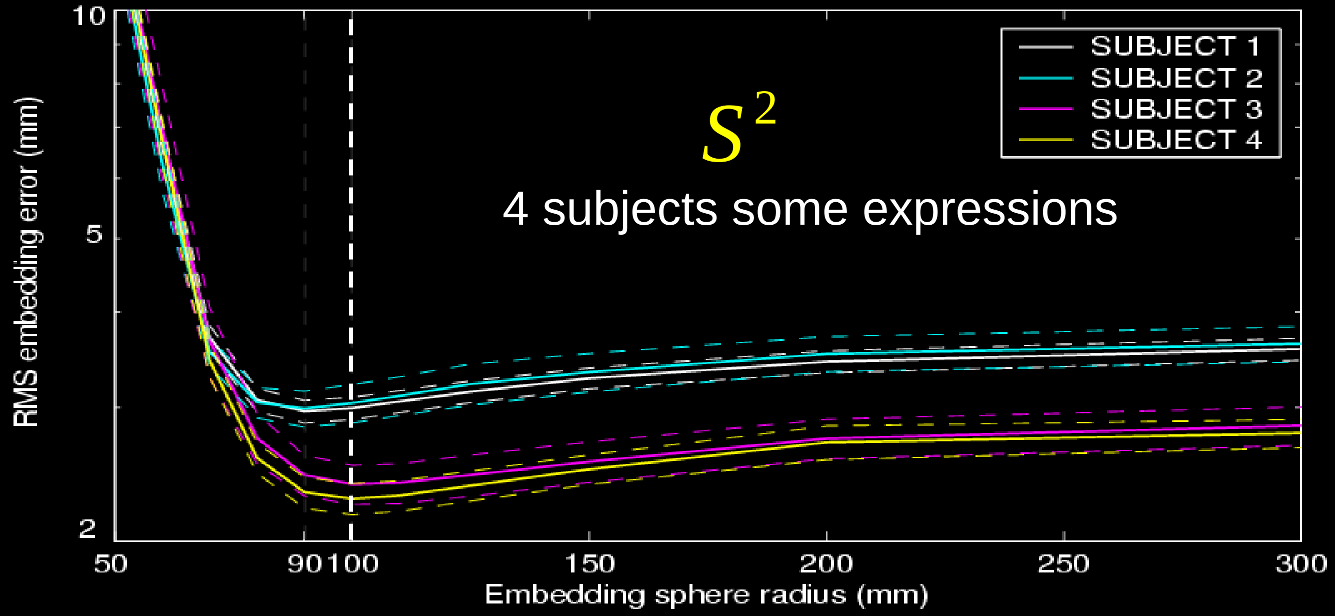


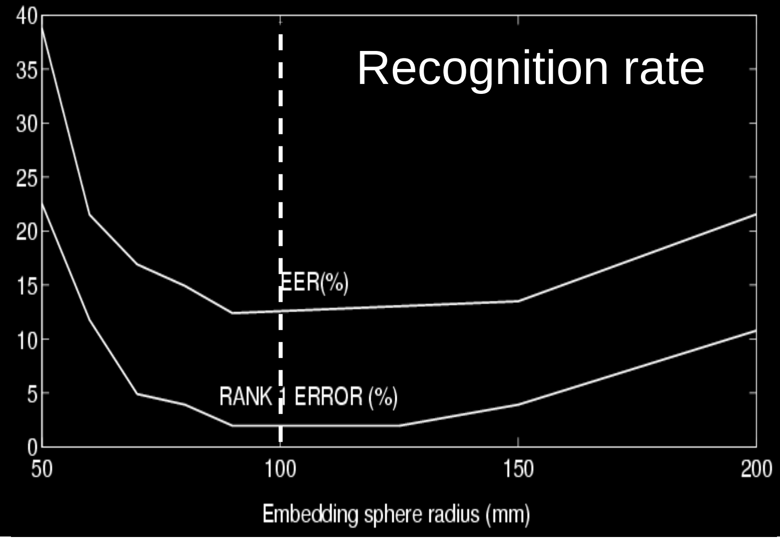


# Spherical embedding error

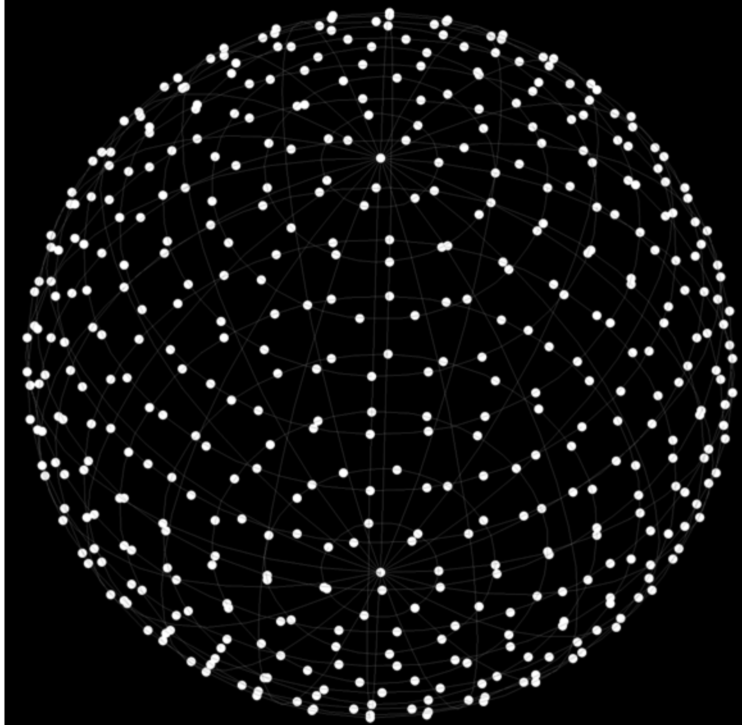


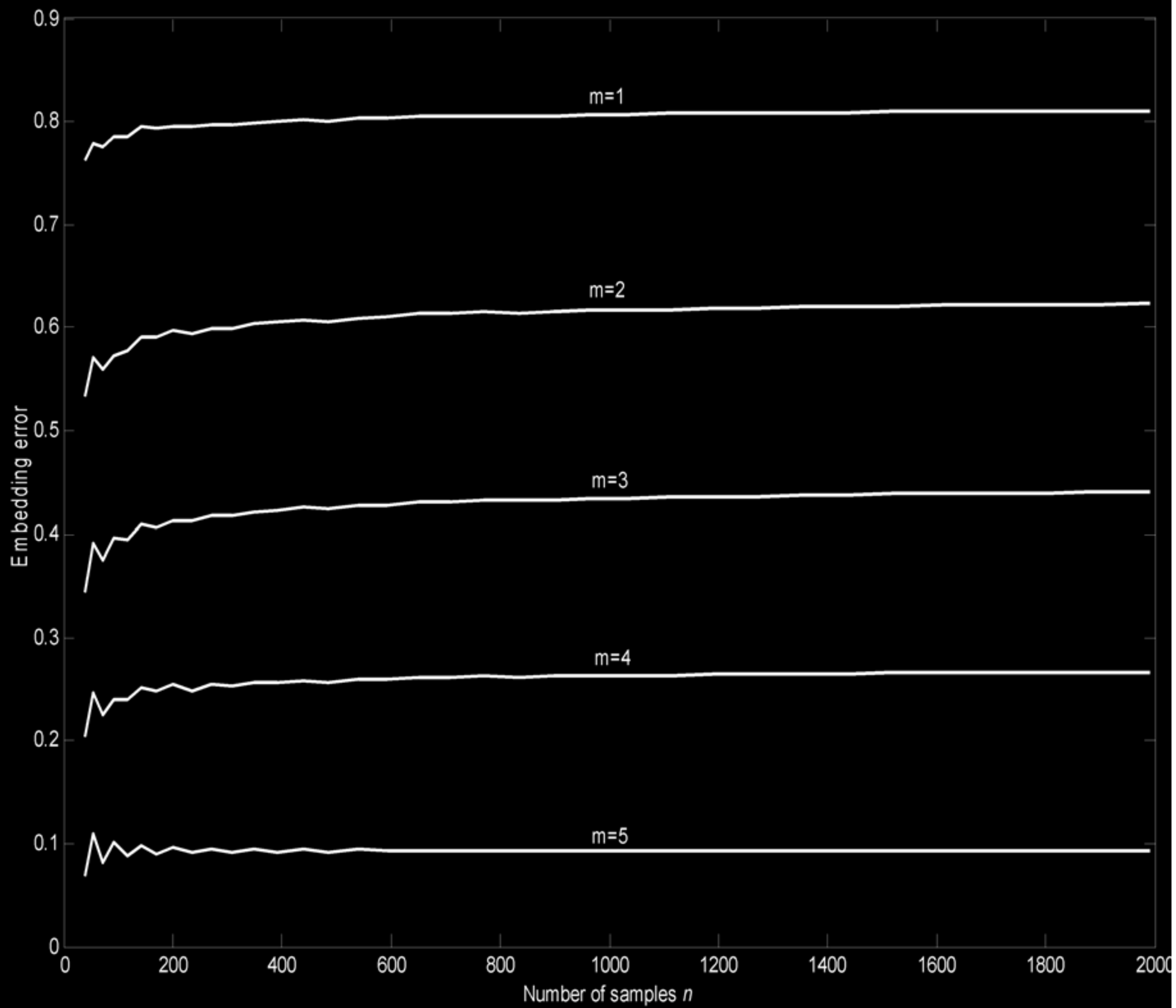
# Spherical embedding error



# FLAT EMBEDDING SAMPLED SURFACES

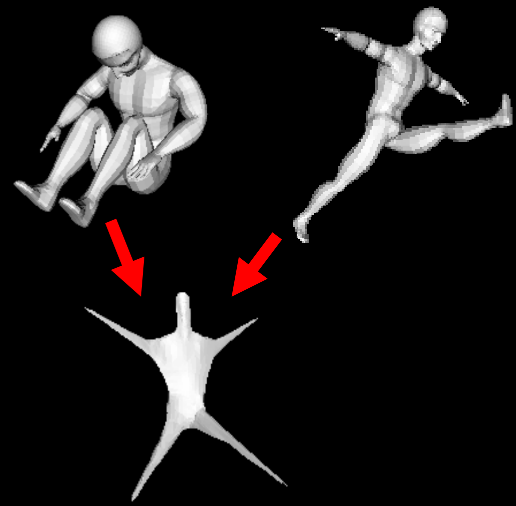


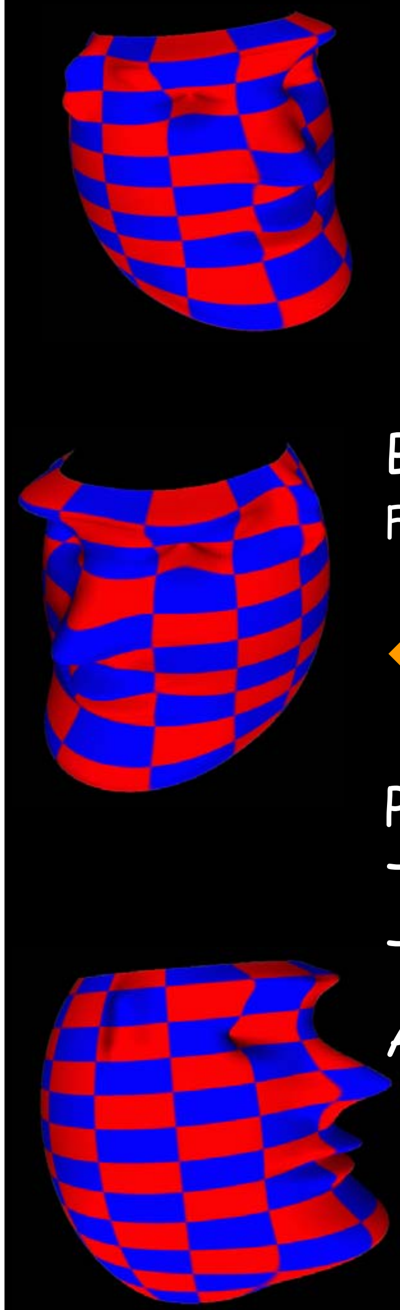



Embedding a sphere in $\mathbf{R}^m$. Asymptotic behavior with number of samples: embedding error decreases as embedding dimension $m$ grows.

A. Bronstein, M. Bronstein and R. Kimmel, “A geometric framework for 3D face recognition“, 2002

# Flat embedding?





Beyond Bourgain theorem (1985):

For n points metric space there is a flat embedding with distortion O(log n)

- Does a smooth surface has a flat embedding with bounded distortion error?

Partial answers:

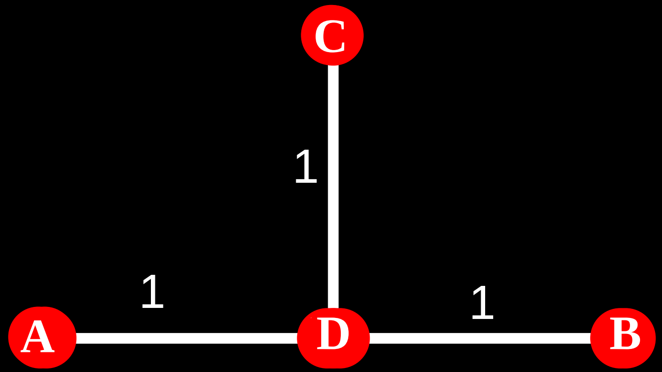
Tight bound for 1D curves (spectral distortion)

Tight numerical bound for spheres (spectral distortion)

And numerical evidence for faces.

Next goal: A sampling theorem for isometric surfaces

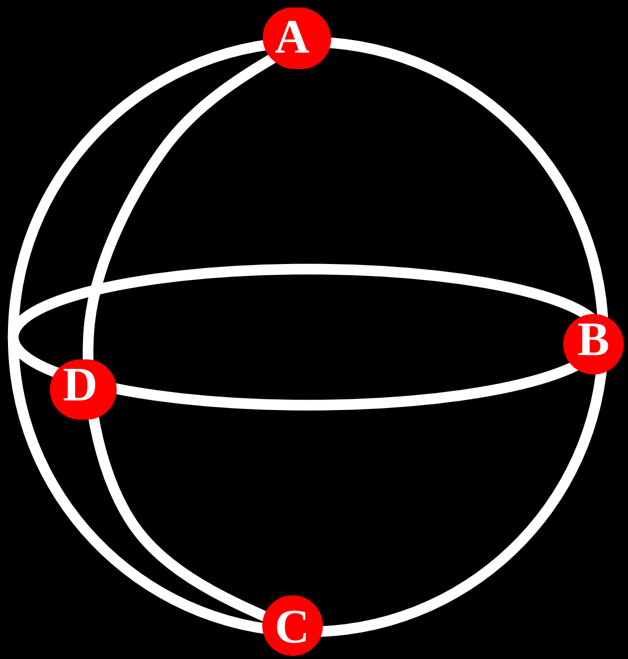
# Embedding arbitrary metric spaces in flat domains

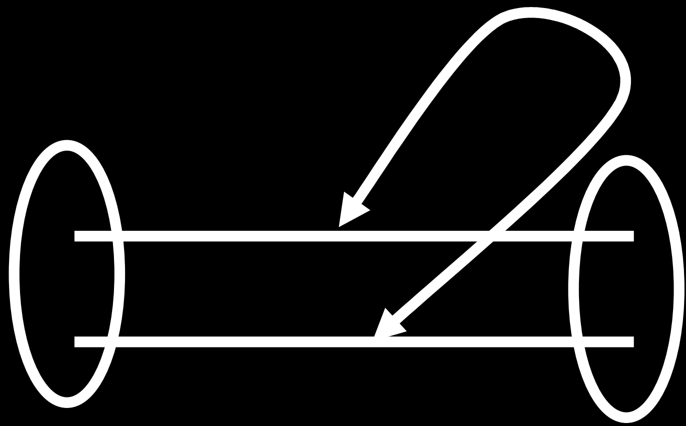



| | A | B | C | D |
|---|---|---|---|---|
| A | 0 | 2 | 2 | 1 |
| B | 2 | 0 | 2 | 1 |
| C | 2 | 2 | 0 | 1 |
| D | 1 | 1 | 1 | 1 |


Personal communication N. Linial

# Embedding arbitrary metric spaces in flat domains



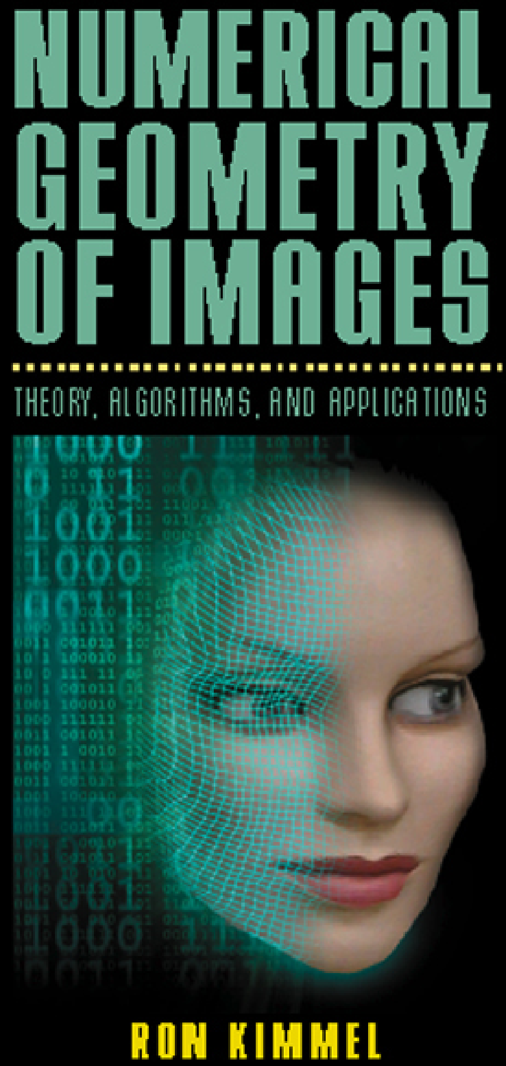
Personal communication N. Linial

Released: Oct. 2003
Publisher: Springer

ISBN: 0387955623

## Future directions:

Better models
Better numerical methods
Better coding
Better hardware
Better analysis of the
large scale problem




Front cover rendered by A&M Bronstein


3DFACE