Computer Vision

18AI742

Dr. rer. nat. D. Antony Louis Piriyakumar,
Dean (R&D),
Cambridge Institute of Technology.

BRIDGE TO THE FUTURE

CAMBRIDGE
INSTITUTE OF TECHNOLOGY

Registration

Dr. D. Antony Louis Piriyakumar,
*Dean (*Research & Development)
Registered Indian patent agent (IN/PA 3041)

Registration

Contents

1) Registering rigid objects
2) Model-based vision: registering rigid objects with projection
3) Registering deformable objects
4) Conclusion
5) Q&A

$$\nabla \cdot \mathbf{E} = \rho / \varepsilon_0$$

$$\nabla \cdot \mathbf{B} = 0$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

$$\nabla \times \mathbf{B} = \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \mu_0 \boldsymbol{j}_c$$

where

$$\nabla = \hat{\mathbf{i}} \frac{\partial}{\partial x} + \hat{\mathbf{j}} \frac{\partial}{\partial y} + \hat{\mathbf{k}} \frac{\partial}{\partial z}$$

Registration

- Registration - finding a transformation that takes one dataset to another
- transformation is rotation, translation, and perhaps scale
- pose—the position and orientation in world coordinates
- MRI image (which is a 3D dataset) of a patient’s interior
- to superimpose on a view of the real patient to help guide a surgeon

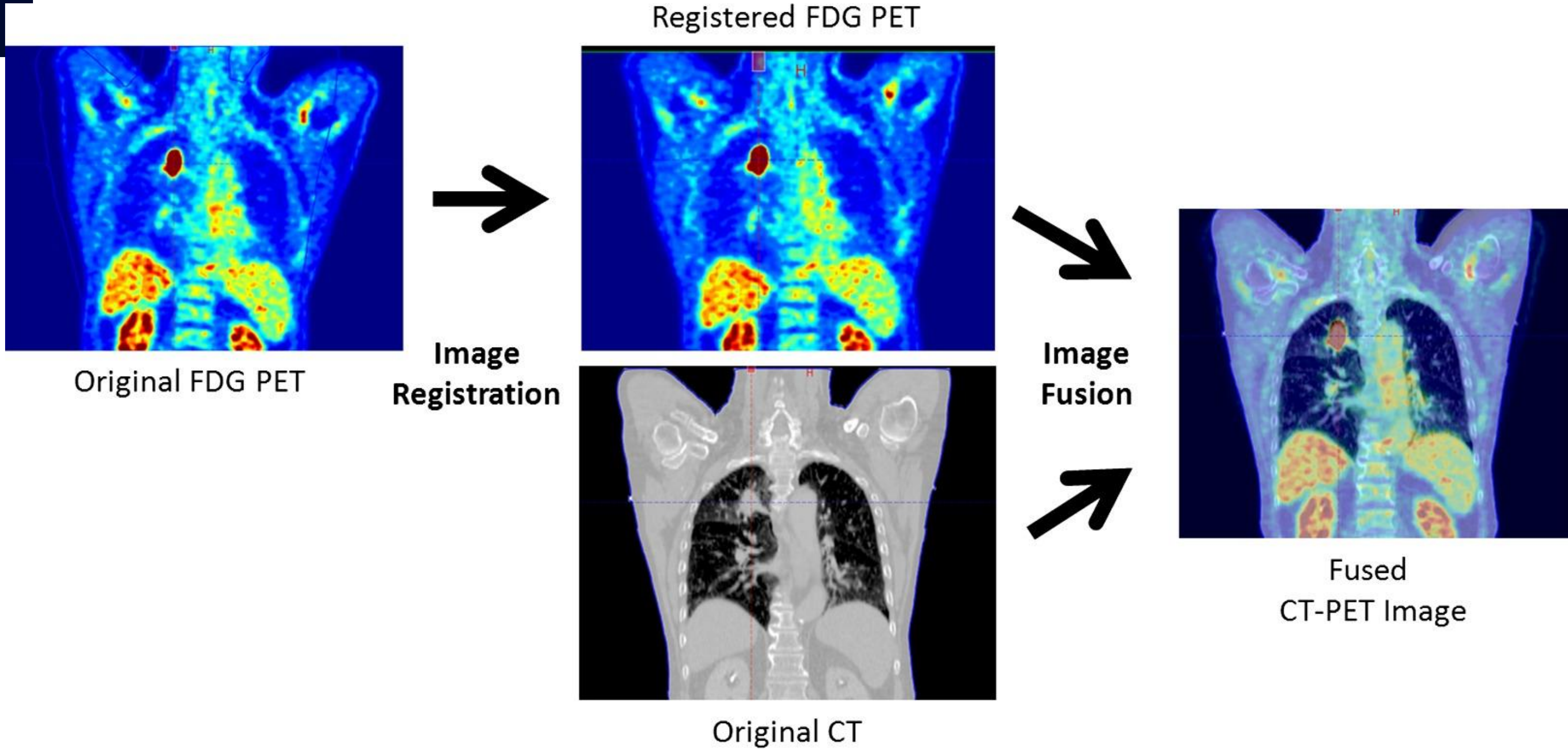
Pose consistency

- different groups of features on rigid object will all report the same pose for the object
- 2D image template of building that we want to find in an overhead aerial image
- use a match quality score to tell whether we have found the right building
- solve these problems using search, exploiting a property
- any search to register rigid objects should be simple

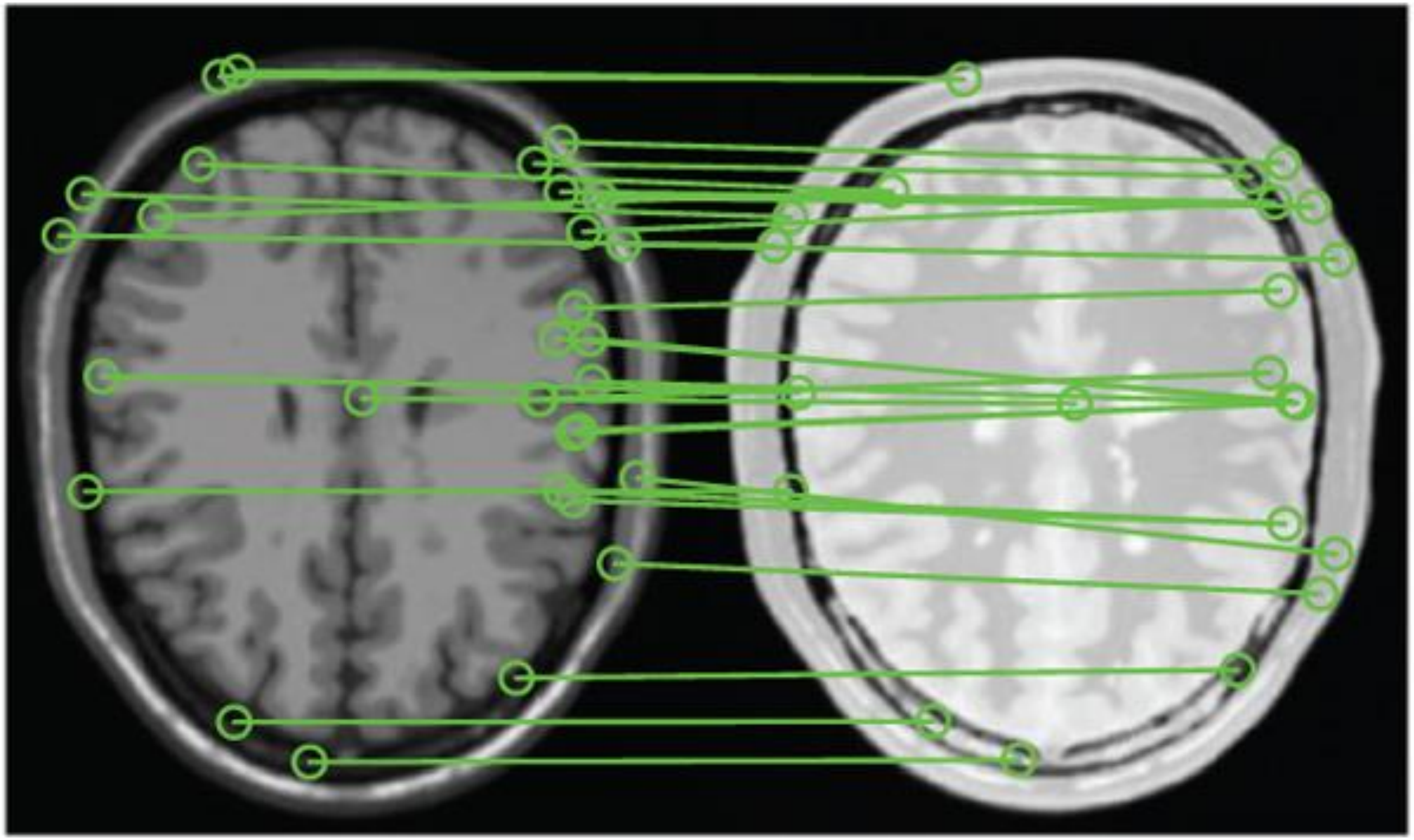
Registration under projection with camera consistency

- image of a 3D object, needs to register the object to the image
- same search algorithms that register datasets of the same dimension?
- Camera consistency - all the features in the image are viewed in the same camera
- need to find only a small set of features to estimate the pose of rigid object

Registration



Registration



Registering rigid objects

- two point sets, S = {xi} a source set and T = {yj}, a target set
- Target set is a rotated, translated, and scaled version of the source set
- c(i) for the index of point in the target set corresponding to the ith source point.
- $$\sum_i \left[(s\mathcal{R}(\theta)x_i + t) - y_{c(i)} \right]^2$$ $$\mathcal{G}(s, \theta, t)\mathcal{S} = \{(s\mathcal{R}(\theta)x_i + t) \mid x_i \in \mathcal{S}\}$$

Iterated closest points algorithm

- For any $y_j \in T$, there is some $z_i \in G$ that is closest.
- Index of the closest such point $c(i, (s, \theta, t))$.
- Estimate of the transformation $(s, \theta, t)(n)$.
- iterating: (a) transforming $p \in S$; (b) for p, finding closest point in T
- (c) re-estimating the transformation with least squares.

Searching for Transformations via Correspondences

- Iterated closest points algo might encounter numerous local minima
- Alternative is to search the space of correspondences
- in the case of rigid objects a relatively small set of correspondences is enough
- tokens, rather than points such as line-segments, corners, or even point-like features blobs

frame-bearing groups for estimating transformations from 2D to 2D

Transformation	Frame-bearing groups
Rigid (Euclidean)	One point and one direction, or two points, or one line and one point
Rigid and scale	Two points, or one line and one point off the line
Affine	Three points, not co-linear

frame-bearing groups for estimating transformations from 3D to 3D

Transformation	Frame-bearing groups
Rigid (Euclidean)	Three points, or one line and one point off the line, or two intersecting lines
Rigid and scale	Three points, or one line and one point off the line, or two intersecting lines and a point off their plane.
Affine	Four points, no two co-planar

CAMBRIDGE
INSTITUTE OF TECHNOLOGY

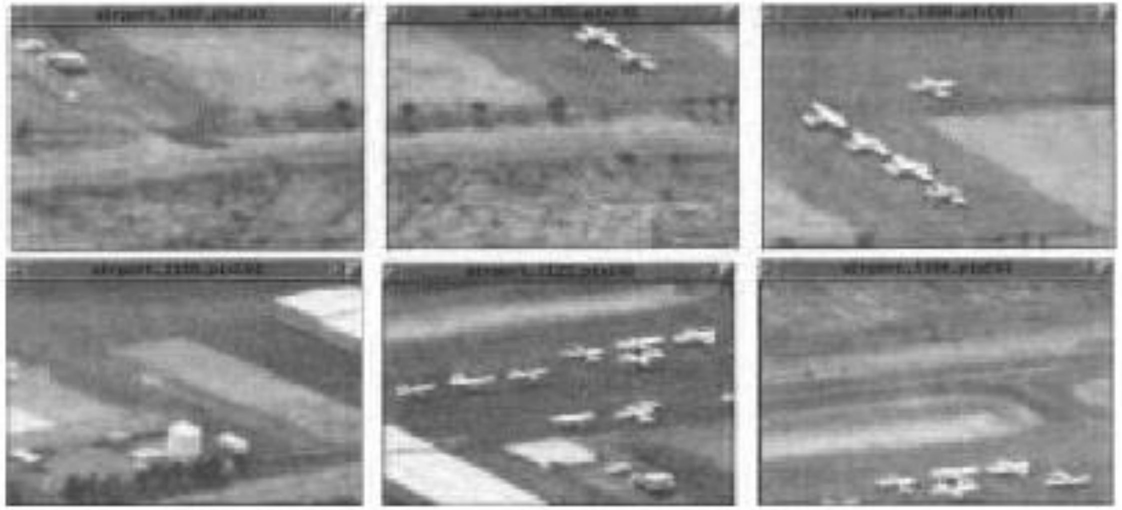
using RANSAC

- Select a frame-bearing group for the target and for the source at random;
- Compute a correspondence between the source and target elements
- from this compute a transformation
- Apply the transformation to the source dataset
- compute a score comparing the transformed source to the target

Application: Building Image Mosaics

- One way to photograph a big, imposing object in detail is
- to take numerous small photographs, then patch them together
- Image mosaics can now be built by registering digital images
- image frames taken by an orthographic camera attached to an aircraft

frames from a video taken by an aircraft overflying an airport for mosaicing



Application: Building Image Mosaics

Cylindrical/Spherical panorama

- Fixed camera collecting video, By registering frames with one another
- we can make estimates of (a) moving objects and (b) background,
- Cylindrical/spherical panorama, a set of pixel samples that mimic the image produced by a
- Cylindrical/spherical camera
- perspective camera were to rotate about its focal point.

Homography

- Nothing about camera is known, the map between the relevant portions of I1 and I2
- plane projective transformation, sometimes known as a homography
- Knowing more about the camera and the circumstances might result in
- a more tightly constrained transformation

Homography

In homogeneous coordinates, the transformation that takes the point $\boldsymbol{x}_1 = (x_1, y_1, 1)$ in $\mathcal{I}_1$ to its corresponding point in $\mathcal{I}_2$, $\boldsymbol{x}_2 = (x_2, y_2, 1)$, has the form of a generic 3×3 matrix with nonzero determinant. Write $\mathcal{H}$ for this matrix. We can estimate its elements using four corresponding points on the plane. Write $\boldsymbol{x}_1^{(i)} = (x_1^{(i)}, y_1^{(i)}, 1)$ for the ith point in $\mathcal{I}_1$, which corresponds to $\boldsymbol{x}_2^{(i)} = (x_2^{(i)}, y_2^{(i)}, 1)$. Now we have

$$\begin{pmatrix} x_2^{(i)} \\ y_2^{(i)} \end{pmatrix} = \begin{pmatrix} \frac{h_{11}x_1^{(i)} + h_{12}y_1^{(i)} + h_{13}}{h_{31}x_1^{(i)} + h_{32}y_1^{(i)} + h_{33}} \\ \frac{h_{21}x_1^{(i)} + h_{22}y_1^{(i)} + h_{23}}{h_{31}x_1^{(i)} + h_{32}y_1^{(i)} + h_{33}} \end{pmatrix},$$

Homography

in the unknown entries of the matrix for each pair of corresponding points, i.e.,

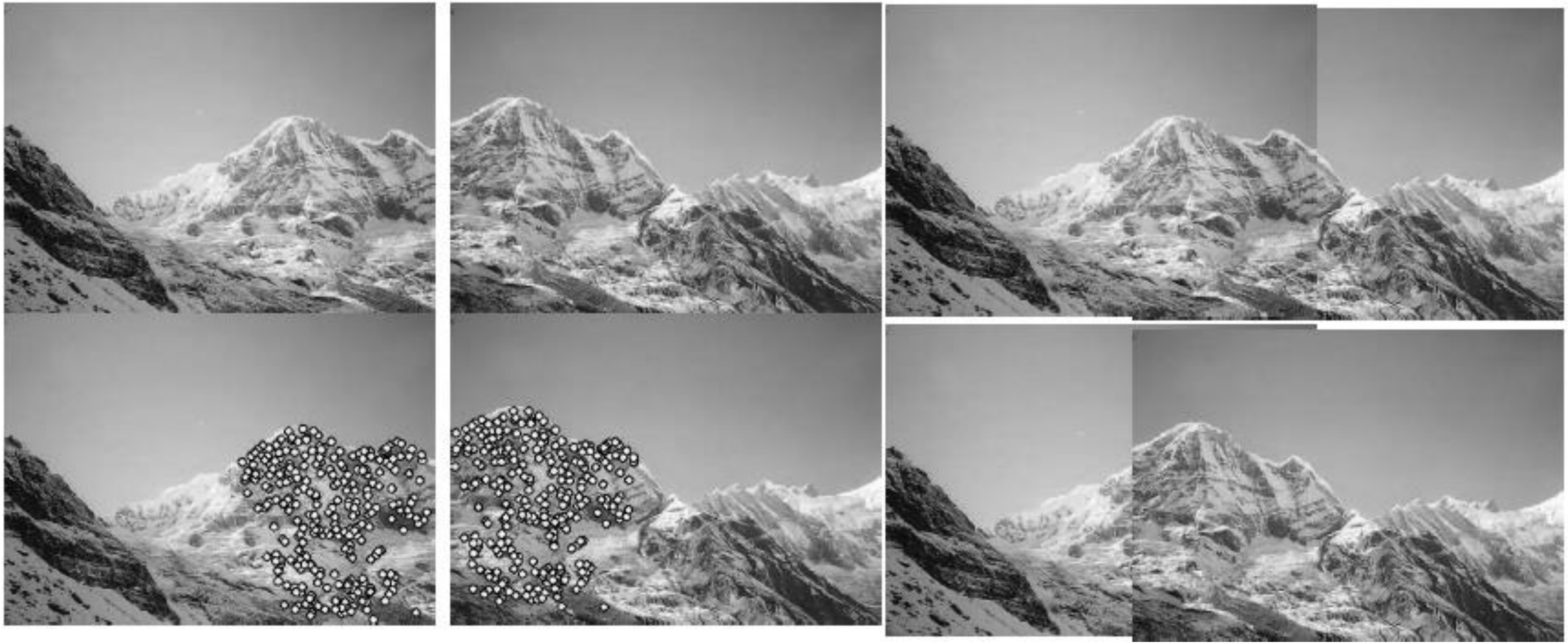
$$
\begin{aligned}
x_2^{(i)}(h_{31}x_1^{(i)} + h_{32}y_1^{(i)} + h_{33}) - (h_{11}x_1^{(i)} + h_{12}y_1^{(i)} + h_{13}) &= 0 \\
y_2^{(i)}(h_{31}x_1^{(i)} + h_{32}y_1^{(i)} + h_{33}) - (h_{21}x_1^{(i)} + h_{22}y_1^{(i)} + h_{23}) &= 0.
\end{aligned}
$$

when we have a large set of corresponding points. In this case, we should minimize as a function of $\mathcal{H}$

$$
\sum_{i \in \text{points}} g\left((x_2^{(i)} - \frac{h_{11}x_1^{(i)} + h_{12}y_1^{(i)} + h_{13}}{h_{31}x_1^{(i)} + h_{32}y_1^{(i)} + h_{33}})^2 + (y_2^{(i)} - \frac{h_{21}x_1^{(i)} + h_{22}y_1^{(i)} + h_{23}}{h_{31}x_1^{(i)} + h_{32}y_1^{(i)} + h_{33}})^2 \right)
$$

where g could be the identity function, which is not a good idea if we have outliers, or is an M-estimator. This function is invariant to the scale of $\mathcal{H}$, so we need

rectified by a translation



rectified with a homography

Error accumulation

- I1, I2, and I3, register I1 to I2, then I2 to I3
- $T_{2\to 1}$ for the estimated transformation that takes image two into image one's frame
- $T_{2\to 1} \circ T_{3\to 2}$ might not be a good estimate of $T_{3\to 1}$
- Resulting in error accumulation

bundle adjustment

- estimate all registrations in one go, using all error terms
- choose a coordinate frame within which to work frame of the first image
- search for a set of maps that take each other image into that frame
- minimize the sum of squared errors between all matching pairs of points.

bundle adjustment

of points. For our example, write $\left(\boldsymbol{x}^{(i)}, \boldsymbol{x}^{(k)}\right)_j$ for the jth tuple consisting of a point $\boldsymbol{x}^{(i)}$ in image i that matches a point $\boldsymbol{x}^{(k)}$ in image k. We would estimate $\mathcal{T}_{2\to1}$ and $\mathcal{T}_{3\to1}$ by minimizing

$$\sum_{j\in 1,\ 2\text{ matches}} g(\| \boldsymbol{x}_j^{(1)} - \mathcal{T}_{2\to1}\boldsymbol{x}_j^{(2)} \|^2)+$$

$$\sum_{j\in 1,\ 3\text{ matches}} g(\| \boldsymbol{x}_j^{(1)} - \mathcal{T}_{3\to1}\boldsymbol{x}_j^{(3)} \|^2)+$$

$$\sum_{j\in 2,\ 3\text{ matches}} g(\| \mathcal{T}_{2\to1}\boldsymbol{x}_j^{(2)} - \mathcal{T}_{3\to1}\boldsymbol{x}_j^{(3)} \|^2)$$

Registration

Contents

1) Registering rigid objects
2) Model-based vision: registering rigid objects with projection
3) Registering deformable objects
4) Conclusion
5) Q&A

$$\nabla \cdot \mathbf{E} = \rho / \varepsilon_0$$

$$\nabla \cdot \mathbf{B} = 0$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

$$\nabla \times \mathbf{B} = \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \mu_0 \boldsymbol{j}_c$$

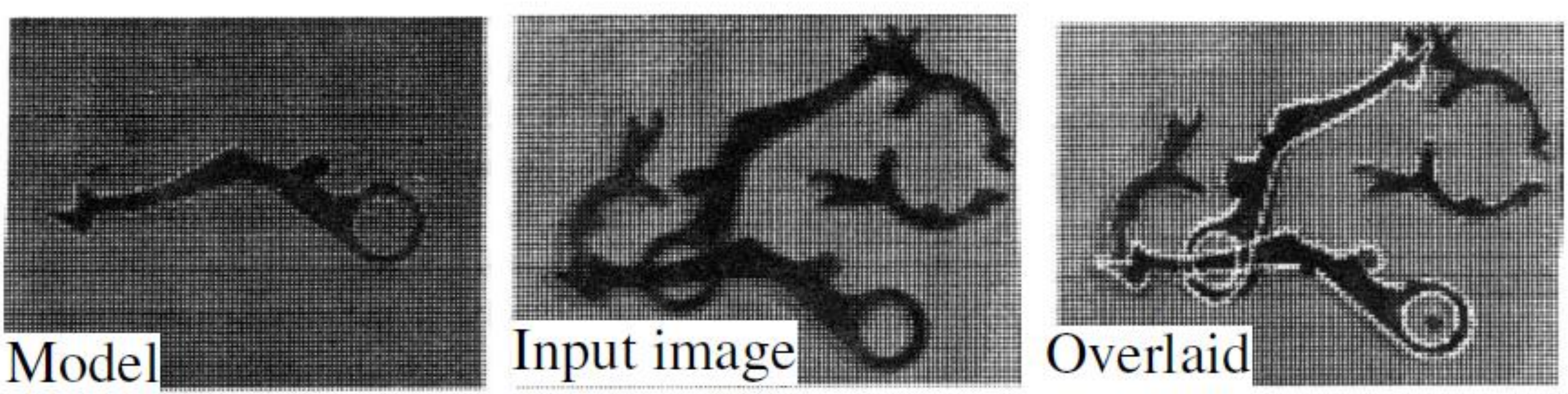
where

$$\nabla = \hat{\mathbf{i}} \frac{\partial}{\partial x} + \hat{\mathbf{j}} \frac{\partial}{\partial y} + \hat{\mathbf{k}} \frac{\partial}{\partial z}$$

orthographic camera

- need to calibrate the camera, and on what camera model we impose.
- cannot determine depth to the 3D object
- correspondences between three points are enough to estimate rotation,
- two observable components of translation, and scale

alignment



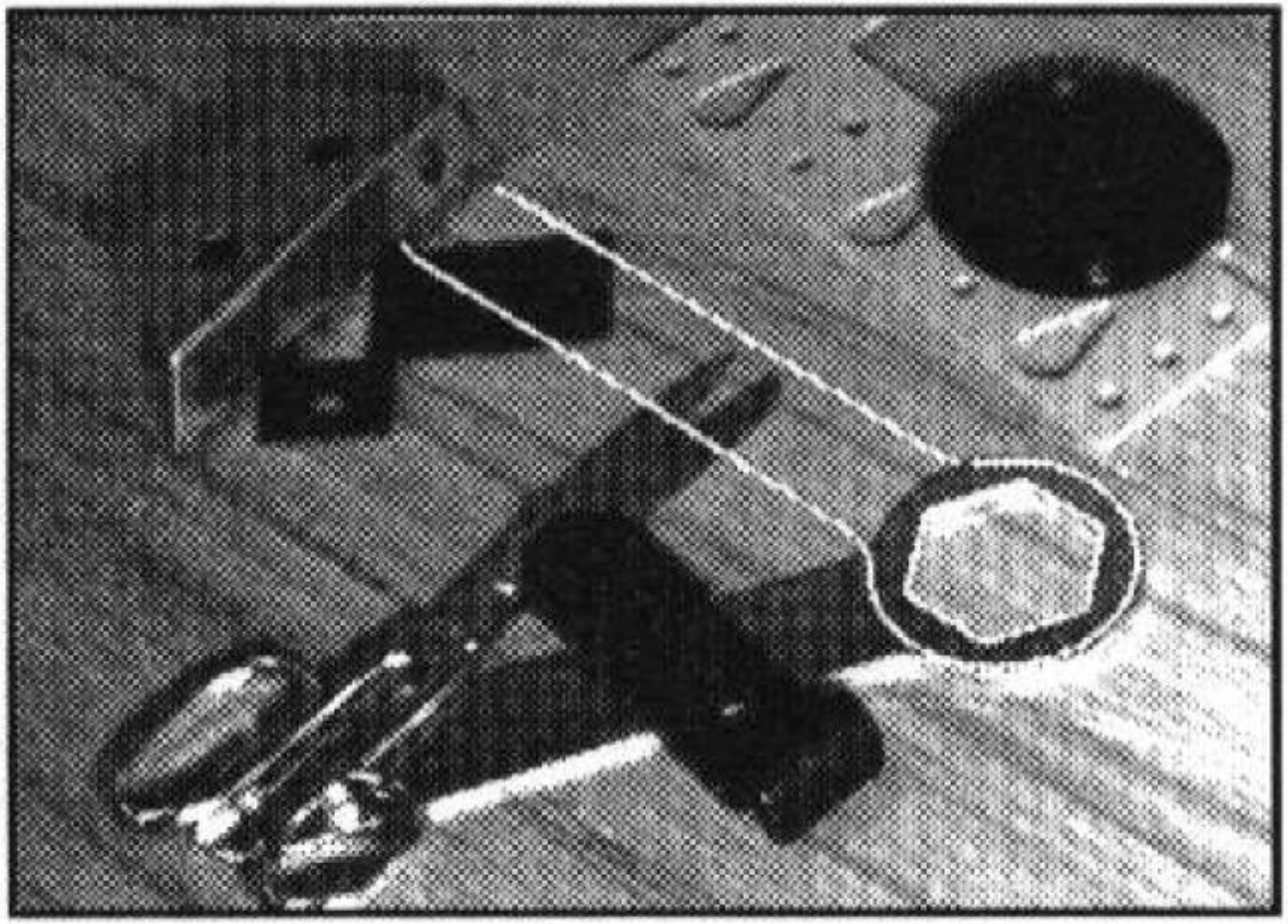
Comparing Transformed and Rendered Source to Target

- render, a general-purpose description for producing an image from models
- encompassing everything from constructing line drawings to
- producing physically accurate shaded images.
- render the transformed object model using our camera model.

Comparing Transformed and Rendered Source to Target

- comparisons should be robust to changes in illumination
- overlay object silhouette edges on image using camera model, and score
- absence of edges lying near these boundaries could well be a quite reliable sign
- affects the contrast sensitivity so that the objects disappear

Edge orientation can be a deceptive cue for verification



Registration

Contents

1) Registering rigid objects
2) Model-based vision: registering rigid objects with projection
3) Registering deformable objects
4) Conclusion
5) Q&A

$$\nabla \cdot \mathbf{E} = \rho/\varepsilon_0$$

$$\nabla \cdot \mathbf{B} = 0$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

$$\nabla \times \mathbf{B} = \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \mu_0 \boldsymbol{j}_c$$

where

$$\nabla = \hat{\mathbf{i}} \frac{\partial}{\partial x} + \hat{\mathbf{j}} \frac{\partial}{\partial y} + \hat{\mathbf{k}} \frac{\partial}{\partial z}$$

Deforming Texture with Active Appearance Models

- texture on the face is an important cue driving the match
- mesh is placed over an image of a face.
- Each point on this triangle has a reference intensity value
- obtain by querying the image at that location on the triangle.

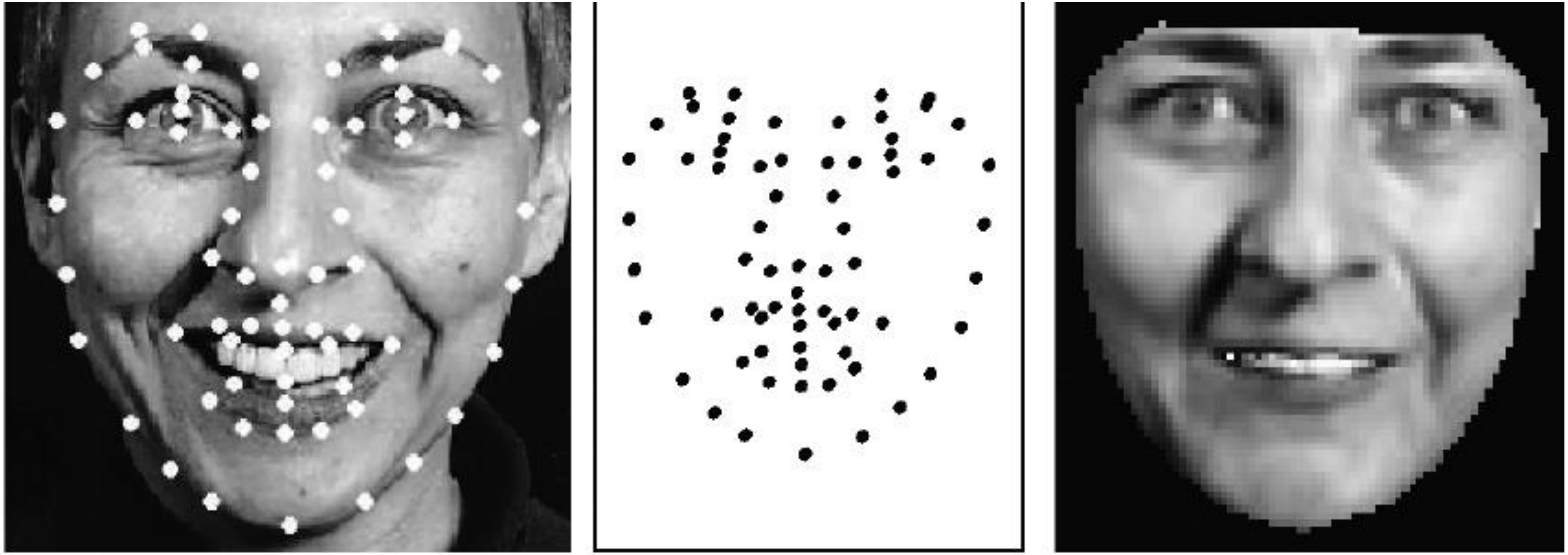
Barycentric coordinates

querying the image at that location on the triangle. Write v_1, v_2, v_3 for the vertices of the triangle. We can represent interior points of the triangle using *barycentric coordinates*; with a point in the reference triangle given by (s, t) such that $0 \le s \le 1$, $0 \le t \le 1$ and $s + t \le 1$, we associate the point

$$p(s, t; v) = sv_1 + tv_2 + (1 - s - t)v_3$$

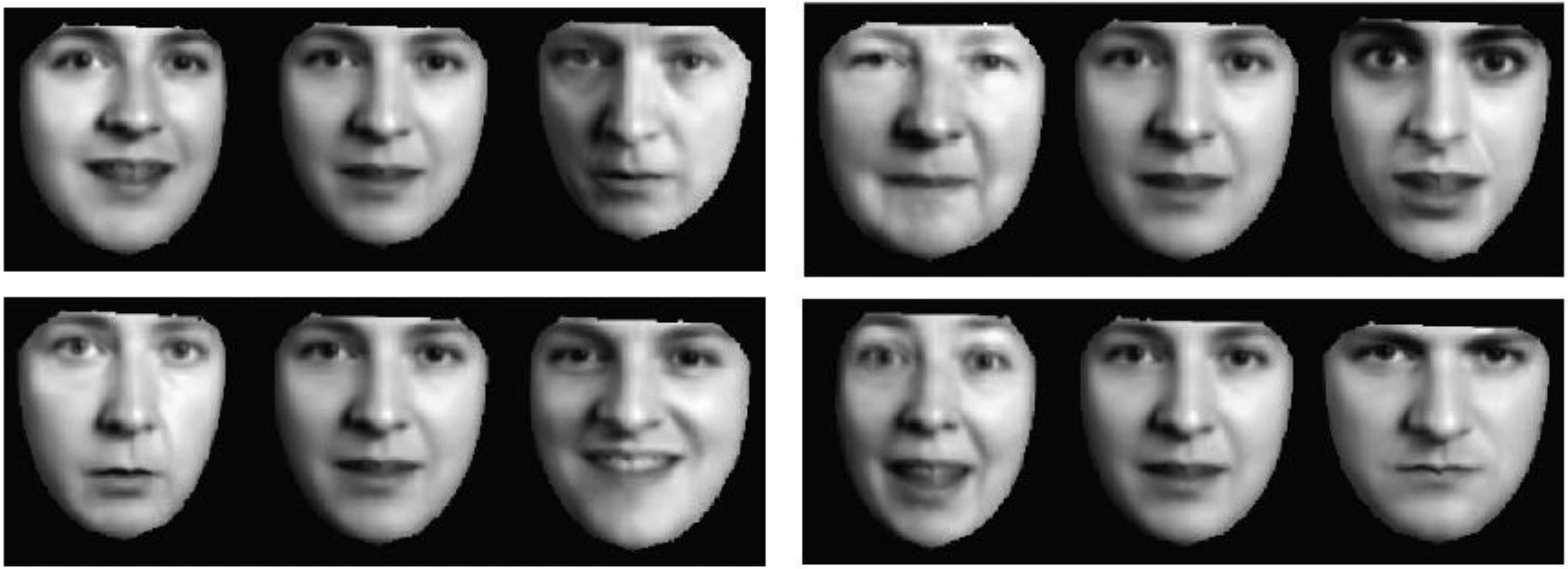
(which lies inside the triangle). The reference intensity value associated with the point (s, t) for the triangle (v_1, v_2, v_3) is $\mathcal{I}_o(p(s, t; v))$.

reasonable triangulation of the original set of points.



Reference points Relaxed points Relaxed intensity

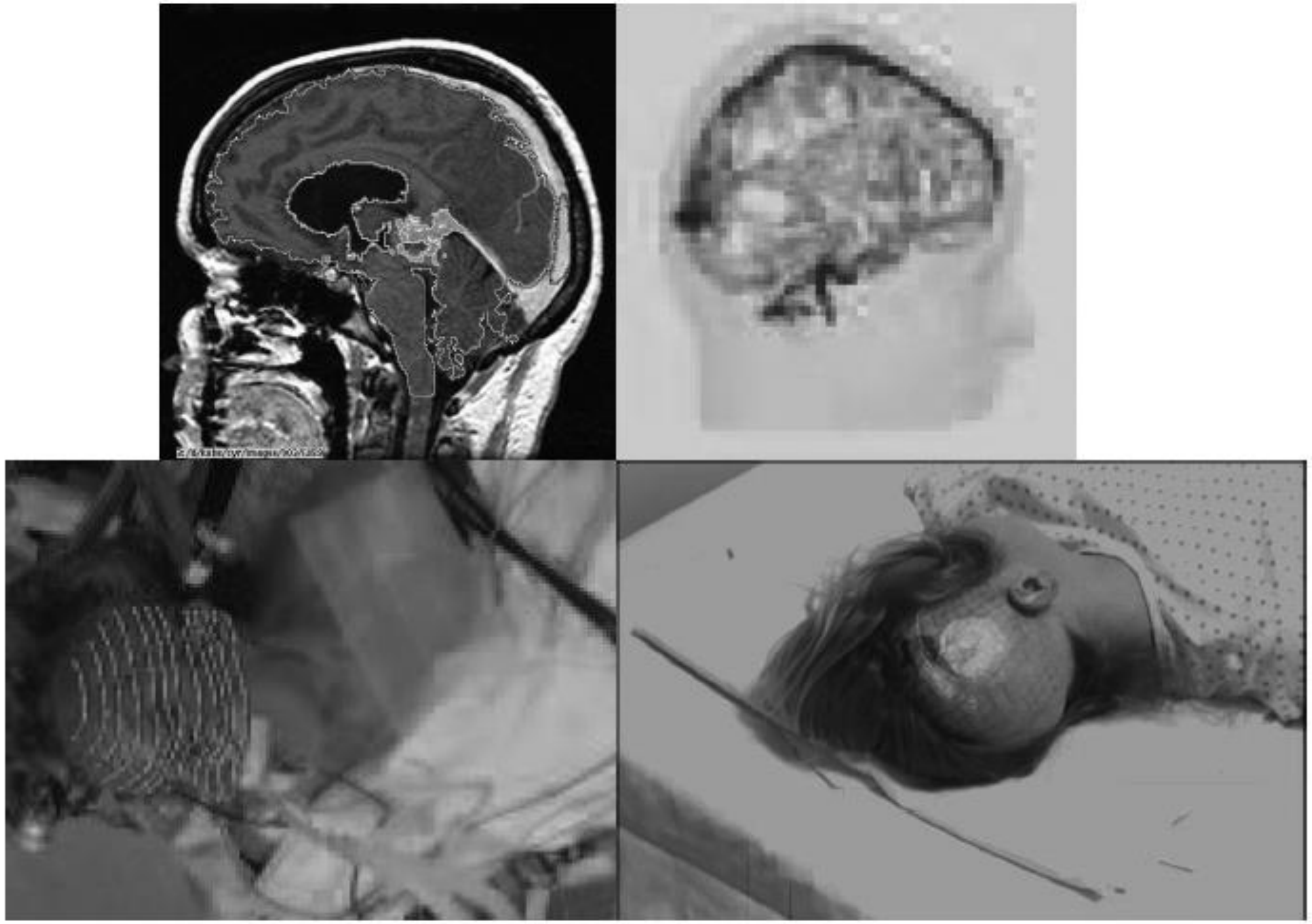
Different face intensity masks generated by moving deformation parameters to different values.



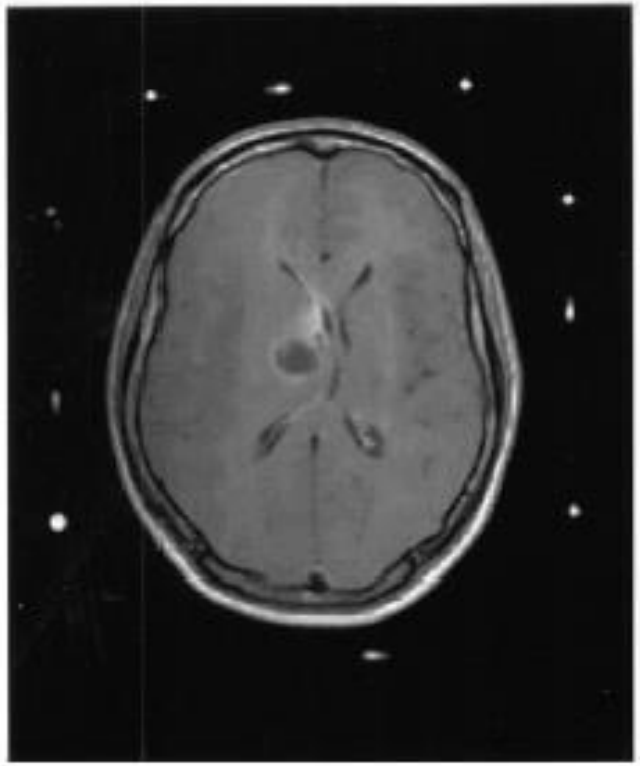
Active appearance models registered to face images

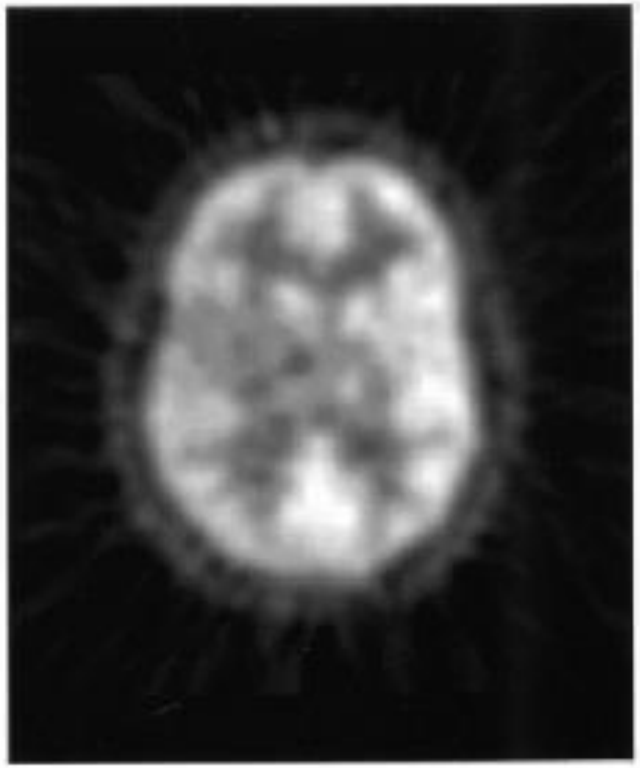
Initial 3 it.s 8 it.s 11 it.s Original Converged

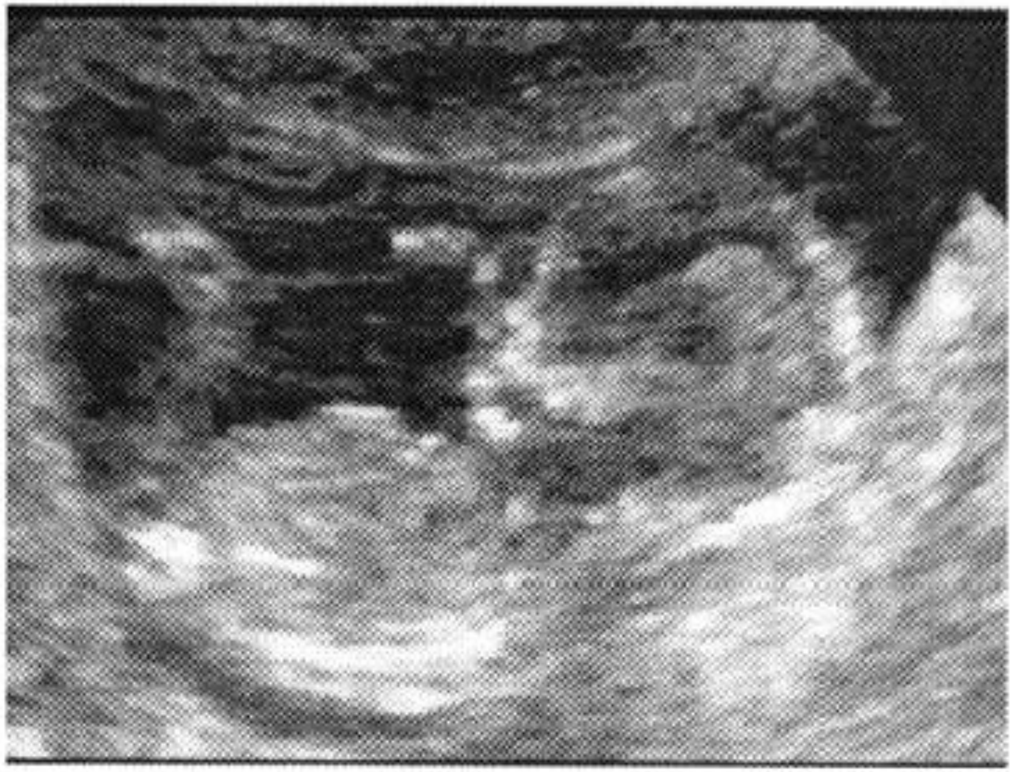
single slice of MRI data with an automatically acquired segmentation overlaid



Images obtained with three different imaging modes







Conclusion

Registration is key technique

- Object detection++
- Ample accuracte algorithms
- Real time applications

Q&A

Contact

- **Prof. D. Antony Louis Piriyakumar**
  Dean (Research and development)
  Cambridge Institute of Technology
- K.R. Puram,
  560036 Bengaluru, India
- Mobile: +91 98459 25132
- E-mail:
  dean_rd@Cambridge.edu.in