Recognizing Facial Expressions

Lecture-13

Copyright Mubarak Shah 2003
Homework, Due November 11

- Lecture 9, slide 17, slide 22
- Lecture 12, page 21 and 22 (three problems).
Program II, Due November 16

• Implement Mean shift Algorithm for tracking
  – Assume that the object location is given in the first frame of the seq
  – Demonstrate your program on known test seqs
  – Demonstrate your program on unknown test seqs in the lab
  – Write a short report: method, problems, results, observations.
• Facial expressions reflect the emotional stage of a person.
• Recognizing facial expression from video sequences is a challenging problem.
• Applications
  – Perceptual user interface
  – Video compression (MPEG-4)
  – Synthesis of facial expressions
Facial Expressions

• Joy
  – The eyebrows are relaxed. The mouth is open, and mouth corners pulled back toward ears.

• Sadness
  – The inner eyebrows are bent upward. The eyes are slightly closed. The mouth is relaxed.

• Anger
  – The inner eyebrows are pulled downward and together. The eyes are wide open. The lips are pressed against each other or opened to expose teeth.

Copyright Mubarak Shah 2003
Facial Expressions

• Fear
  – The eyebrows are raised and pulled together. The inner eyebrows are bent upward. The eyes are tense and alert.

• Disgust
  – The eyebrows and eyelids are relaxed. The upper lip is raised and curled, often asymmetrically.

• Surprise
  – The eyebrows are raised. The upper eyelids are wide open, the lower relaxed. The jaw is open.

Copyright Mubarak Shah 2003
FACIAL EXPRESSIONS

RAISE EYE BROWS

SMILE

Copyright  Mubarak Shah 2003
Black and Yacoob Algorithm

• Given the location of the face, eyes, brows, and mouth estimate the rigid motion of the face using pseudo perspective motion model.
• Use the face motion to register images through warping.
• Estimate relative motion of face features (eyes, mouth, brows).
• The estimated feature motions are used to predict locations of features in the next frame, and the process is repeated.
• The estimated motion is used to classify the facial expressions.
Face - Planar

Eyes - Affine

Mouth, Brows - Affine + Curvature

Rigid Face Transformation

Non-rigid Facial Features
Affine

\[ u(x, y) = a_1 x + a_2 y + b_1 \]
\[ v(x, y) = a_3 x + a_4 y + b_2 \]

\[
\begin{bmatrix}
  u(x, y) \\
  v(x, y)
\end{bmatrix}
= \begin{bmatrix}
  x & y & 1 & 0 & 0 & 0 \\
  0 & 0 & 0 & x & y & 1
\end{bmatrix}
\begin{bmatrix}
  a_1 \\
  a_2 \\
  b_1 \\
  a_3 \\
  a_4 \\
  b_2
\end{bmatrix}
\]
Affine

\[ u(x, y) = a_1 x + a_2 y + b_1 \]
\[ v(x, y) = a_3 x + a_4 y + b_2 \]

Expansion or contraction

Rotation around Z

Squashing or stretching

\[ \text{divergence} = u_x + v_y = a_1 + a_4 \]

\[ \text{curl} = -(u_y - v_x) = -(a_2 - a_3) \]

\[ \text{deformation} = (u_x - v_y) = (a_1 - a_4) \]
Pseudo Perspective

\[ u(x, y) = a_1 + a_2 x + a_3 y + a_4 x^2 + a_5 xy \]

\[ v(x, y) = a_6 + a_7 x + a_8 y + a_4 xy + a_5 y^2 \]

\[ a_4 = \text{yaw: rotation around y-axis} \]

\[ a_5 = \text{pitch: rotation around x-axis} \]

\[
\begin{bmatrix}
u(x, y) \\
v(x, y)
\end{bmatrix} = 
\begin{bmatrix}
1 & x & y & x^2 & xy & 0 & 0 & 0 \\
0 & 0 & 0 & xy & y^2 & 1 & x & y
\end{bmatrix}
\begin{bmatrix}
a_1 \\
a_2 \\
a_3 \\
a_4 \\
a_5 \\
a_6 \\
a_7 \\
a_8
\end{bmatrix}
\]
Pseudo Perspective

\[ u(x, y) = a_1 + a_2 x + a_3 y + a_4 x^2 + a_5 xy \]
\[ \nu(x, y) = a_6 + a_7 x + a_8 y + a_4 xy + a_5 y^2 \]

\[ a_4 = \text{yaw} \]
\[ a_5 = \text{pitch} \]
Affine with Curvature

\[ u(x, y) = a_1 x + a_2 y + b_1 \]
\[ v(x, y) = a_3 x + a_4 y + b_2 + cx^2 \]

\[
\begin{bmatrix}
  u(x, y) \\
  v(x, y)
\end{bmatrix} =
\begin{bmatrix}
  x & y & 1 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & x & y & 1 & x^2
\end{bmatrix}
\begin{bmatrix}
  a_1 \\
  a_2 \\
  b_1 \\
  a_3 \\
  a_4 \\
  b_2 \\
  c
\end{bmatrix}
\]
Rules for Classifying Expressions

• Anger
  – B: inward lowering of brows and mouth contraction
  – E: outward raising of brows and mouth expansion

• Disgust
  – B: mouth horizontal expansion and lowering of brows
  – E: mouth contraction and raising of brows

• Happiness
  – B: upward curving of mouth and expansion or horizontal deformation
  – E: downward curving of mouth and contraction or horizontal deformation
Rules for Classifying Expressions

• Surprise
  – B: raising brows and vertical expansion of mouth
  – E: lowering brows and vertical contraction of mouth

• Sadness
  – B: downward curving of mouth and upward-inward motion in the inner parts of brows
  – E: upward curving of mouth and downward-outward motion in inner parts of brows

• Fear
  – B: expansion of mouth and raising-inwards inner parts of brows
  – E: contraction of mouth and lowering inner parts of brows
Smile Expression

Upward-outward motion of mouth corners results in \(-\text{ve}\) curvature.

Horizontal and overall vertical stretching result in \(+\text{ve}\) div & def.

Some upward trans is caused by raising of lower and upper lips due to stretching of the mouth (\(a3\) is \(-\text{ve}\)).
Smile

Figure 8: Smile experiment: facial expression tracking.
Smile Mouth Parameters

Figure 9: Smile mouth parameters. For translation, solid and dashed lines indicate horizontal and vertical motion respectively.
Anger

Figure 10: Anger experiment: facial expression tracking. Features every 15 frames.
Anger Motion Parameters

Figure II: Anger motion parameters; the solid line indicates the right eye or brow while the dashed line indicates the left eye or brow.
Surprise
Surprise Motion Parameters
Blinking

Figure 14: Blinking experiment: facial feature tracking. Features every four frames.
Blinking Motion Parameters for Eyes
Rotation
Rotate Face motion parameters

$P_0 \text{ rot y}$

$P_1 \text{ rot X}$
Rotation Motion Parameters

![Graphs showing rotation motion parameters for Mouth, Brows, Vertical Translation, Curvature, Divergence, Deformation, and Curvature.](#)
Mid-level predicates for Mouth

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Threshold</th>
<th>Derived Predicates</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_0$</td>
<td>$&gt; 0.25$</td>
<td>Mouth rightward</td>
</tr>
<tr>
<td></td>
<td>$&lt; -0.25$</td>
<td>Mouth leftward</td>
</tr>
<tr>
<td>$a_3$</td>
<td>$&lt; -0.1$</td>
<td>Mouth upward</td>
</tr>
<tr>
<td></td>
<td>$&gt; 0.1$</td>
<td>Mouth downward</td>
</tr>
<tr>
<td>$Div$</td>
<td>$&gt; 0.02$</td>
<td>Mouth expansion</td>
</tr>
<tr>
<td></td>
<td>$&lt; -0.02$</td>
<td>Mouth contraction</td>
</tr>
<tr>
<td>$Def$</td>
<td>$&gt; 0.005$</td>
<td>Mouth horizontal deformation</td>
</tr>
<tr>
<td></td>
<td>$&lt; -0.005$</td>
<td>Mouth vertical deformation</td>
</tr>
<tr>
<td>$Curl$</td>
<td>$&gt; 0.005$</td>
<td>Mouth clockwise rotation</td>
</tr>
<tr>
<td></td>
<td>$&lt; -0.005$</td>
<td>Mouth counterclockwise rotation</td>
</tr>
<tr>
<td>$e$</td>
<td>$&lt; -0.0001$</td>
<td>Mouth curving upward (<em>U</em> like)</td>
</tr>
<tr>
<td></td>
<td>$&gt; 0.0001$</td>
<td>Mouth curving downward</td>
</tr>
</tbody>
</table>
Mid-level predicates for Head

Table 4: The mid-level predicates derived from deformation and motion parameter estimates as applied to head motion.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Threshold</th>
<th>Derived Predicates</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_0$</td>
<td>&gt; 0.5</td>
<td>Head rightward</td>
</tr>
<tr>
<td></td>
<td>&lt; −0.5</td>
<td>Head leftward</td>
</tr>
<tr>
<td>$a_3$</td>
<td>&lt; −0.5</td>
<td>Head upward</td>
</tr>
<tr>
<td></td>
<td>&gt; 0.5</td>
<td>Head downward</td>
</tr>
<tr>
<td>$D_{iv}$</td>
<td>&gt; 0.01</td>
<td>Head expansion</td>
</tr>
<tr>
<td></td>
<td>&lt; −0.01</td>
<td>Head contraction</td>
</tr>
<tr>
<td>$D_{ef}$</td>
<td>&gt; 0.01</td>
<td>Head horizontal deformation</td>
</tr>
<tr>
<td></td>
<td>&lt; −0.01</td>
<td>Head vertical deformation</td>
</tr>
<tr>
<td>$C_{url}$</td>
<td>&gt; 0.005</td>
<td>Head clockwise rotation</td>
</tr>
<tr>
<td></td>
<td>&lt; −0.005</td>
<td>Head counterclockwise rotation</td>
</tr>
<tr>
<td>$p_0$</td>
<td>&lt; −0.00005</td>
<td>Head rotating rightward around the neck</td>
</tr>
<tr>
<td></td>
<td>&gt; 0.00005</td>
<td>Head rotating leftward around the neck</td>
</tr>
<tr>
<td>$p_1$</td>
<td>&lt; −0.00005</td>
<td>Head rotating forward</td>
</tr>
<tr>
<td></td>
<td>&gt; 0.00005</td>
<td>Head rotating backward</td>
</tr>
</tbody>
</table>
Parameter values used for classifying expressions

<table>
<thead>
<tr>
<th>Expr.</th>
<th>B/E</th>
<th>Feature</th>
<th>$a_0$</th>
<th>$a_3$</th>
<th>Div</th>
<th>Cowl</th>
<th>Def</th>
<th>c</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anger</td>
<td>B</td>
<td>Mouth</td>
<td></td>
<td></td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R. Brow</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L. Brow</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R. Eye</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L. Eye</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Anger</td>
<td>E</td>
<td>Mouth</td>
<td></td>
<td>+</td>
<td>0</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R. Brow</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L. Brow</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R. Eye</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L. Eye</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Happiness</td>
<td>B</td>
<td>Mouth</td>
<td></td>
<td></td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Happiness</td>
<td>E</td>
<td>Mouth</td>
<td></td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Surprise</td>
<td>B</td>
<td>Mouth</td>
<td></td>
<td>+</td>
<td>0</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R. Brow</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L. Brow</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R. Eye</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L. Eye</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Surprise</td>
<td>E</td>
<td>Mouth</td>
<td></td>
<td>+</td>
<td>0</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R. Brow</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L. Brow</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R. Eye</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L. Eye</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Forty Test Subjects
## Results

<table>
<thead>
<tr>
<th>Expression</th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surprise</td>
<td>91%</td>
</tr>
<tr>
<td>Happiness</td>
<td>95%</td>
</tr>
<tr>
<td>Anger</td>
<td>90%</td>
</tr>
<tr>
<td>Disgust</td>
<td>93%</td>
</tr>
<tr>
<td>Fear</td>
<td>83%</td>
</tr>
<tr>
<td>Sadness</td>
<td>100%</td>
</tr>
</tbody>
</table>
Beginning of Anger Expression
Frames from 10 Video Clips
## Results

<table>
<thead>
<tr>
<th>Expression</th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surprise</td>
<td>86%</td>
</tr>
<tr>
<td>Happiness</td>
<td>95%</td>
</tr>
<tr>
<td>Anger</td>
<td>80%</td>
</tr>
<tr>
<td>Disgust</td>
<td>50%</td>
</tr>
<tr>
<td>Fear</td>
<td>100%</td>
</tr>
<tr>
<td>Sadness</td>
<td>60%</td>
</tr>
</tbody>
</table>