Designing Descriptors

16-385 Computer Vision
Feature matching is important for tasks like…
Object instance recognition

Schmid and Mohr 1997

Sivic and Zisserman, 2003

Rothganger et al. 2003

Lowe 2002
Image mosaicing
If we know where the good features are, how do we match them?
How do we describe an image patch?

Patches with similar content should have similar descriptors.
Challenges of designing a feature descriptor
Photometric transformations
Geometric transformations
e.g. scale, translation, rotation
Designing a feature descriptor
(the search for image invariants)
What is the best descriptor for an image feature?
Image patch

Just use the pixel values of the patch

Perfectly fine if geometry and appearance is unchanged (a.k.a. template matching)

What are the problems?
Image patch

Just use the pixel values of the patch

Perfectly fine if geometry and appearance is unchanged
(a.k.a. template matching)

What are the problems?
How can you be less sensitive to absolute intensity values?
Image gradients

Use pixel differences

Feature is invariant to absolute intensity values

What are the problems?
Image gradients

Use pixel differences

Feature is invariant to absolute intensity values

What are the problems?
How can you be less sensitive to deformations?
Color histogram

Count the colors in the image using a histogram

Invariant to changes in scale and rotation

What are the problems?
Color histogram

Count the colors in the image using a histogram

Invariant to changes in scale and rotation

What are the problems?
Color histogram

Count the colors in the image using a histogram

Invariant to changes in scale and rotation

What are the problems?
How can you be more sensitive to spatial layout?
Spatial histograms

Compute histograms over spatial ‘cells’

Retains rough spatial layout
Some invariance to deformations

What are the problems?
Spatial histograms

Compute histograms over spatial ‘cells’

Retains rough spatial layout
Some invariance to deformations

What are the problems?
How can you be completely invariant to rotation?
Orientation normalization

Use the dominant image gradient direction to normalize the orientation of the patch

What are the problems?
Feature Descriptors

16-385 Computer Vision
Tiny Images
Just downsample it
Simple
Fast
Robust to small affine transformation

What are the problems?
Multi-Scale Oriented Patches (MOPS)

Multi-Scale Oriented Patches (MOPS)

Given a feature \((x, y, s, \theta)\)

Get 40 x 40 image patch, subsample every 5th pixel
(what’s the purpose of this step?)

Subtract the mean, divide by standard deviation
(what’s the purpose of this step?)

Haar Wavelet Transform
(what’s the purpose of this step?)
Given a feature \((x, y, s, \theta)\)

Get 40 x 40 image patch, subsample every 5th pixel (low frequency filtering, absorbs localization errors)

Subtract the mean, divide by standard deviation
(what’s the purpose of this step?)

Haar Wavelet Transform
(what’s the purpose of this step?)
Given a feature \((x, y, s, \theta)\)

Get 40 x 40 image patch, subsample every 5th pixel (low frequency filtering, absorbs localization errors)

Subtract the mean, divide by standard deviation (removes bias and gain)

Haar Wavelet Transform (what’s the purpose of this step?)
Multi-Scale Oriented Patches (MOPS)


Given a feature \((x, y, s, \theta)\)

Get 40 x 40 image patch, subsample every 5th pixel (low frequency filtering, absorbs localization errors)

Subtract the mean, divide by standard deviation (removes bias and gain)

Haar Wavelet Transform (low frequency projection)
Haar Wavelets
(actually, Haar-like features)

Use responses of a bank of filters as a descriptor
Haar wavelet responses can be computed with filtering.

Haar wavelet responses can be computed efficiently (in constant time) with integral images.
# Integral Image

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>19</td>
</tr>
</tbody>
</table>

\[
A(x, y) = \sum_{x' \leq x, y' \leq y} I(x', y')
\]

Can find the sum of any block using 3 operations

\[
A(x_1, y_1, x_2, y_2) = A(x_2, y_2) - A(x_1, y_2) - A(x_2, y_1) + A(x_1, y_1)
\]
\[ A(x_1, y_1, x_2, y_2) = A(x_2, y_2) - A(x_1, y_2) - A(x_2, y_1) + A(x_1, y_1) \]

\[
\begin{array}{ccc}
1 & 5 & 2 \\
2 & 4 & 1 \\
2 & 1 & 1 \\
\end{array}
\]

image

\[
\begin{array}{ccc}
1 & 6 & 8 \\
3 & 12 & 15 \\
5 & 15 & 19 \\
\end{array}
\]

integral image

\[ A(1, 1, 3, 3) = A(3, 3) - A(1, 3) - A(3, 1) + A(1, 1) \]
\[ = 19 - 8 - 5 + 1 \]
\[ = 7 \]
Given an image patch, compute filter responses

filter bank (20 Haar wavelet filters)

vector of filter responses

Responses are usually computed at specified location as a face patch descriptor

When will this feature descriptor fail?