How do we (directly) compare two images?

Are these images the same?  Are they similar?
Pixel-wise Comparison

Or, normalized by image area, about 5 grey values per pixel.

\[ 8,140 = \sum_{x}^{x<148} \sum_{y}^{y<161} |A(x,y) - B(x,y)| \]
Consider two vectors/points.

\[
X = \begin{pmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n
\end{pmatrix} \quad Y = \begin{pmatrix}
y_1 \\
y_2 \\
\vdots \\
y_n
\end{pmatrix}
\]

Distance vs. similarity:

\[
S : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R} \\
D : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}
\]

\[
S \propto \frac{1}{D}
\]
Simple Distances (norms)

$L_1$ - City Block Distance

$L_2$ - Euclidean Distance

$L_{\infty}$ - Max Distance

Generalized $L$-norm

$$\sum_{x,y}(|A(x,y) - B(x,y)|)$$

$$\sqrt{\sum_{x,y}(A[x,y] - B[x,y])^2}$$

$$\text{Max}_{x,y}|A[x,y] - B[x,y]|$$

$$\sqrt[l]{\sum_{x,y}(|A(x,y) - B(x,y)|)^l}$$
Curves shown are the set of points that are ‘one unit’ from the origin using different definitions of distance.
Consider the following problem:

*Find the unique point “closest” to k other points.*

For simplicity, do this in $\mathbb{R}$ (a line) with $k = 2$.

See the problem yet?

\[ |2 - 3| + |8 - 3| = 6 \quad |2 - 4| + |8 - 4| = 6 \]
In Comparison, Consider L2

Find the unique point “closest” to k other points.

Using L2,

\[ \sqrt{(2 - 5)^2 + (8 - 5)^2} = \sqrt{18} \quad \sqrt{(2 - 4)^2 + (8 - 4)^2} = \sqrt{20} \]

Best \hspace{2cm} \text{Not as Good}
Motivating Pearson’s Correlation

Let’s play a Game

<table>
<thead>
<tr>
<th>Game 1 X</th>
<th>Y</th>
<th>Y-True</th>
<th>Check</th>
<th>Game 2 X</th>
<th>Y</th>
<th>Y-True</th>
<th>Check</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>3</td>
<td></td>
<td>4</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>4</td>
<td></td>
<td>1</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>2</td>
<td></td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>5</td>
<td></td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5</td>
<td></td>
<td>5</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>2</td>
<td></td>
<td>1</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>4</td>
<td>5</td>
<td></td>
<td>4</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>5</td>
<td>4</td>
<td></td>
<td>5</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>5</td>
<td>2</td>
<td></td>
<td>5</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>4</td>
<td>3</td>
<td></td>
<td>4</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>4</td>
<td></td>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>3</td>
<td></td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>5</td>
<td></td>
<td></td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>3</td>
<td></td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>4</td>
<td></td>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>2</td>
<td></td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>2</td>
<td></td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The results of the Game

- First Game - Random
  - Expected ~20% correct
- Second Game
  - Invert 1-5 to 4-0
  - Add some noise
- Game 2 - features
  - Nearly perfect prediction
  - ... to within one value.
- Punch line
  - Correlation measures predictability!

<table>
<thead>
<tr>
<th>Game 1</th>
<th>Game 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>Y</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td>13</td>
<td>5</td>
</tr>
<tr>
<td>14</td>
<td>3</td>
</tr>
<tr>
<td>15</td>
<td>4</td>
</tr>
<tr>
<td>16</td>
<td>2</td>
</tr>
<tr>
<td>17</td>
<td>1</td>
</tr>
<tr>
<td>18</td>
<td>2</td>
</tr>
<tr>
<td>19</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td>1</td>
</tr>
</tbody>
</table>

Tally

Tally
Pearson’s Correlation

\[
\sum_{x,y} (A(x,y) - \overline{A})(B(x,y) - \overline{B}) \over \sqrt{\sum_{x,y} (A(x,y) - \overline{A})^2} \sqrt{\sum_{x,y} (B(x,y) - \overline{B})^2}
\]

What is the underlying model?
Assumptions of Correlation

\[
\sum_{x,y} (A(x,y) - \bar{A})(B(x,y) - \bar{B}) \\
\sqrt{\sum_{x,y} (A(x,y) - \bar{A})^2} \sqrt{\sum_{x,y} (B(x,y) - \bar{B})^2}
\]

- Two signals vary linearly
  - Constant shift to either signal has no effect.
  - Increased amplitude has no effect.

- This minimizes sensitivity to:
  - changes in (overall) illumination
  - offset or gain.
Special Cases

• Any two linear functions with positive slope have correlation 1.
  – Only the sign of the slope matters.

• Any two linear functions with differently signed slopes have correlation \(-1\).
  – This is called anti-correlation
  – Anti-correlation = correlation for prediction.
  – For matching, it may or may not be as good…

• Correlation undefined for slope = 0 ($\sigma=0$)
Correlation (cont.)

• For Images, correlation is sensitive to:
  – Translation
  – Rotation: in-plane and out-of-plane
  – Scale

• Because it …
  – Assumes pixels align one atop the other.
  – Compares two images pixel by pixel.

• Translation handled by convolution
  – Example, alignment by template matching
Computing Correlation

• Remember adding a constant does not change correlation to any other signal, so
  – Let’s subtract average $A$ from $A()$
  – Let’s subtract average $B$ from $B()$
  – The mean of both signals is now zero
  – Then correlation reduces to:

\[
\frac{A \cdot B}{\sqrt{\sum_{x,y} (A(x,y) - \bar{A})^2} \sqrt{\sum_{x,y} (B(x,y) - \bar{B})^2}}
\]
Computing Correlation (II)

• For zero-mean signals, we can scale them without changing their correlation scores
  – Multiply A by the inverse of its length
  – Multiply B by the inverse of its length
  – Both signals are now unit length
  – Then correlation reduces to:

\[ A \cdot B \]

• Gives rise to ‘Correlation Space’.
Correlation Space

• Why zero-mean & unit-length your images?

• Consider database retrieval
  – Compare new image A …
  – with many images in database.
  – When database images are stored in their zero-mean & unit-length form, then
  – Preprocess A (zero-mean, unit-length)
  – Compute dot products
Correlation Space (II)

• New idea: image as a point in an N-dimensional space
  • N = width \times height

• Zero-mean & unit-length images lie on an N-1 dimensional “correlation space” where the dot product equals correlation.
  – This is a highly non-linear projection.
  – Points lie on an N-1 surface within the original N-dimensional space.

• So consider points in 3-D ....
Correlation Space (II)

• Subtracting mean - translation.
• Length one - project onto sphere.
• Correlation is then:
  – Cosine of angle between vectors (points).
Useful Connection …

• Euclidean distance inverse of correlation in correlation space.

\[
\sqrt{\sum_{x,y} (A[x,y] - B[x,y])^2} = \sqrt{\sum_{x,y} A[x,y]^2 + \sum_{x,y} B[x,y]^2 - 2A[x,y]B[x,y]}
\]

\[
= \sqrt{1 + 1 - 2 \sum_{x,y} A[x,y]B[x,y]}
\]

\[
= \sqrt{2 - 2A \cdot B}
\]

\[
= \sqrt{2 - 2\text{Corr}(A,B)}
\]

Nearest-neighbor classifiers in correlation space maximize correlation
Limitations

- To match images this way, they must be
  - The same width & height
  - In correspondence: coordinates match

- More importantly, objects in the scene must
  - Be in the same location
  - Be at the same scale
  - Be at the same orientation
  - Be seen from the same viewpoint