Today, first...  

- Color  
  - Electromagnetic Radiation  
  - Illumination - Reflection - Detection  
  - The Human Eye  
  - Digital Cameras  
  - Colour spaces  
  - Pseudo colouring  
  - Multispectral data


Anders Brun  
Centre for Image Analysis  
Swedish University of Agricultural Sciences  
Uppsala University

and then...  

- Compression  
  - Information and Data  
  - Redundancy  
  - Image Quality  
  - Coding  
  - Compression  
  - File formats

Electromagnetic Radiation

- Every imaging system selects one or several spectral windows.  
- For grayscale image we have a single window, typically covering most of the visible spectrum.

Image Formation

The registered image depends on the spectral properties of:  
- The illumination - light from, e.g., the sun or a lamp  
- The object or scene - light can be reflected, absorbed or transmitted  
- The detector - can be, e.g., a camera or the human eye

Light Properties

- Illumination  
  - Achromatic light - White or uncoloured light that contains all visual wavelengths in a complete mix.  
  - Chromatic light - Coloured light.  
  - Monochromatic light - Light with a single wavelength, e.g., a laser.  
- Reflection  
  - Colours we see are typically a mix of wavelengths.  
  - The dominant wavelength reflected by an object decides the "colour tone" or hue.  
  - If many wavelengths are reflected in equal amounts, an object appears to be grey.
What is brightness?

- Radiance, watts [W]
  - The total amount of energy that flows from the light source.
- Luminance, lumens [lm]
  - A measure of the amount of energy an observer perceives.
- Brightness
  - A subjective descriptor that is practically impossible to measure.

Subtractive Colour Mixing

\[ C + M + Y = \text{black} \]

\[ R + G + B = \text{WHITE} \]

CIE 1931 XYZ Color Space

- Based on a large scale study with volunteers evaluate colour samples.
- Representing all of the colours visible to the average person.
- A 2D projection of the XYZ colour space is the CIE chromaticity diagram (to the right).
- Standard white when X+Y+Z
- The typical monitor gamut with three primary colours can’t produce all colors.
- A colour printer with five pigments produce another gamut

The Human Eye

- Rods for lightness, cones for colour.

The typical monitor gamut and a printer gamut

CE 1931 xy chromaticity diagram with a typical monitor gamut and a printer gamut
**The Human Eye**

- About 5 million cone cells per eye. Most localized in the fovea.

- Three types of cones:
  - S, blue, peak at 420 nm, (2%)
  - M, green, peak at 534 nm, (33%)
  - L, red, peak at 564 nm, (65%)

**The Mantis Shrimp**

- 12 types of colour receptors extending the view into the ultraviolet.
- Compound eyes with three different regions giving each eye trinocular vision.
- They are very colourful not only in our visual spectra.

**Digital Camera as Detector**

- Much like the human eye a digital camera has sensors sensitive to three colours. In a CCD chip they are differentiated by a Bayer filter pattern.
- Values are then interpolated to a full RGB image.

**Colour Spaces**

**RGB/CMY, red green blue / cyan magenta yellow**

- RGB is closer to the physiological side of our vision (the three cone types) rather than the psychological.
- \[ [C M Y] = [1.0 \ 1.0 \ 1.0] - [R G B] \]

**Colour Spaces**

**HSV/HSL**

- The HSV and HSL color spaces has intensity decoupled from colour information.
Colour Spaces
Difference between HSV/HSL

- The difference is how saturation and lightness/value hang together. For some people it is non intuitive to have almost white colours that are fully saturated while others likes the lightness property in HSL.

Colour Spaces
CIE L*a*b* or CIELAB

- The most complete colour space specified by the International Commission on Illumination, CIE in 1976.
- Created to:
  - represent all colours visible to the human eye.
  - serve as a device independent model to be used as a reference.
  - be perceptually uniform - equal distance should have equal perceptual difference.

Noise in colour images

- Gaussian noise in all colour channels
- In a HSL representation the noise is most apparent in the H and S channel.

Filtering in colour

- In general all filters used for grey level images can be used for colour images.
- Filtering can be carried on a per-color-plane basis or example on the intensity only.
- There is no right or wrong, but the result differ.

Look out for the H-channel

- You might end up with color artifacts if the H-channel is filtered.
- Remember that the Hue channel is in degrees 0 to 360°.
Is there a better encoding of H?

- H is a scalar number from 0.0 to 1.0 (float)
- However, H = 0.0 ⇔ H = 1.0
- H is an angle! (0 ... 360 deg)

Is there a better encoding of H that makes it easier to filter?
- You tell me...

More grey level methods on colour images

- Histogram equalization on all channels in HSV can make things strange. Only used on the V channel the expected result is acquired.

Segmentation Based on Hue

- Using an intensity decoupled colour space segmentation based on colour can be relative intuitive.
- Setting an interval for the hue around the hue value for red in HSV space the red part of the fish is segmented.

Choosing Color Space

- A colour space can be close to the hardware or close to the application. RGB is close to the output from a CCD, etc.
- Decoupled intensity can be very useful in image processing making it possible to use many grey-scale methods intuitively.
- Some spaces like HSV has a difficult transformation from, e.g., RGB. Singularities may exist.
- Regardless of which colour space is used RGB is often the colour space for the displaying device.

A Remark on Coding

- Depending on the context, color values are encoded by
  - Floating point numbers: 0.0 ... 1.0
  - 8-bit unsigned integers: 0 ... 255
- Often hexadecimals are used to encode RGB (e.g., in HTML)
  - The hex coded color #FF55FF = (255, 85, 255) as (1.0, 0.3, 1.0)

A Short Matlab Example

- The singularity caused by H=0.0 ⇔ H=1.0
- (0r H=0 ⇔ H = 255 for 8-bit coded images)
Pseudo Colouring
For visualization
- The eye can distinguish between ~30 different grey-levels and about 350,000 different colours.
- A grey-level image can be displayed as a colour image for easier visual inspection. Note that the information content is not changed.
- This is often used in modalities when the samples doesn’t correspond to the visual spectrum.

More than three spectral windows
- Why restrict our imaging system to the borders of human vision?
- Multi spectral data is common in remote sensing, astronomy, thermography, etc.
- Multispectral satellite data with 7 spectral bands of the Uppsala urban area.

Guess the colour of:
a) the tower
b) the colibri
(the colour space is RGB)

Many cameras are sensitive to IR or near-IR. Especially webcams, mobile cameras and other cheap CMOS-sensors. By blocking other than IR wavelengths IR photography becomes almost easy.
Summary

- Every imaging system selects one or several spectral windows. Hence, the spectral dimension is unavoidable.
- There are several ways to represent the colour information in the visible spectra. (Colour spaces)
- The RGB colour space is usually unavoidable since viewing devices (LCD/CRT screens) use it.
- Choose colour space based on the specific application and situation.

Redundancy information and data

- Data is not the same thing as information.
- Data is the means with which information is expressed. The amount of data can be much larger than the amount of information.
- Redundant data doesn’t provide additional information.
- Image coding or compression aims at reducing the amount of data while keeping the information by reducing the amount of redundancy.

Different Types of Redundancy

- Coding Redundancy
  - Some gray levels are more common than others.
- Interpixel redundancy
  - The same gray level may cover a large area.
- Psycho-Visual Redundancy
  - The eye can only resolve about 32 gray levels locally.

Image Compression and Decompression

Image Compression

Image compression can be:

- **Reversible** (lossless), with no loss of information.
  - The image after compression and decompression is identical to the original image. Often necessary in image analysis applications.
  - The compression ratio is typically 2 to 10 times.
- **Non reversible** (lossy), with loss of some information.
  - Lossy compression is often used in image communication, compact cameras, video, www, etc.
  - The compression ratio is typically 10 to 30 times.

Image Coding and Compression

- Image coding
  - How the image data can be represented.
- Image compression
  - Reducing the amount of data required to represent an image.
  - Enabling efficient image storing and transmission.
Objective Measures of Image Quality

- **Error** \( e(x, y) = f(x, y) - f(x, y) \)
- **Total Error** \( \varepsilon_{\text{tot}} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (f(x, y) - f(x, y))^2 \)
- **Root-Mean-Square** \( \varepsilon_{\text{RMS}} = \frac{1}{MN} \sqrt{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (f(x, y) - f(x, y))^2} \)
- **Signal-to-Noise Ratio** \( SNR_{\text{dB}} = \frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y)^2}{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (f(x, y) - f(x, y))^2} \)

Subjective Measures of Image Quality

- **Problem**
  - The objective image quality measures previously shown does not always fit with our perception of image quality.
- **Solution**
  - Let a number of test persons rate the image quality of the images on a scale. This will result in a subjective measure of image quality, or rather fidelity, but it will be based on how we perceive the quality of the images.

Information Measure

- If \( p(E) \) is the probability of an event \( E \), then
  \[ I(E) = -\log(p(E)) \]
  is a measure of the information that the event provides.
- The average information is called entropy.
- Shannon Entropy
  \[ H(z) = - \sum_{k=0}^{L-1} p(r_k) \log_2 p(r_k) \]
  where \( r_k \) is gray level \( k \) and \( L \) is the number of gray levels.

Measure the Amount of Data

- The amount of data in an \( M \times N \) image with \( L \) gray levels is equal to \( MN \log_2 \), where
  \[ L_{\text{avg}} = \sum_{k=0}^{L-1} l(r_k) p(r_k) \]
  \( l(r_k) \) is the number of bits used to represent gray level \( r_k \) and \( p(r_k) \) is the probability of gray level \( r_k \) in the image.

Example 3-bit image

Dealing with coding redundancy

- **Basic idea**: Different gray levels occur with different probability (non-uniform histogram). Use shorter code words for the more common gray levels and longer code words for less common gray levels. This is called **Variable Code Length**.
Huffman Coding

- First
  1. Sort the gray levels by decreasing probability
  2. Sum the two smallest probabilities.
  3. Sort the new value into the list.
  4. Repeat 1 to 3 until only two probabilities remains.

- Second
  1. Give the code 0 to the highest probability, and the code 1 to the lowest probability in the summed pair.
  2. Go backwards through the tree one node and repeat from 1 until all gray levels have a unique code.

Example of Huffman coding

\[
\begin{align*}
\text{node 1} & : p(r_1) = 0.610, p(r_2) = 0.390 \rightarrow 0.301, 0.401 \\
\text{node 2} & : p(r_1) = 0.602, p(r_2) = 0.398 \\
\text{node 3} & : p(r_1) = 0.663, p(r_2) = 0.337 \\
\text{node 4} & : p(r_1) = 0.091, p(r_2) = 0.909 \\
\text{node 5} & : p(r_1) = 0.060, p(r_2) = 0.940 \\
\text{node 6} & : p(r_1) = 0.035, p(r_2) = 0.965 \\
\end{align*}
\]
Example of Huffman coding

Huffman Coding

- First
  1. Sort the gray levels by decreasing probability.
  2. Add the two smallest probabilities.
  3. Sort the new value into the list.
  4. Repeat 1 to 3 until only two probabilities remain.

- Second
  1. Give the code 0 to the highest probability, and the code 1 to the lowest probability in the summed pair.
  2. Go backwards through the tree one node and repeat from 1 until all gray levels have a unique code.

Example of Huffman coding

Assigning codes

Assigning codes

Example of Huffman coding

Assigning codes

Example of Huffman coding

Assigning codes
### Example of Huffman coding

#### Assigning codes

<table>
<thead>
<tr>
<th>$r_k$</th>
<th>$p(r_k)$</th>
<th>node 1</th>
<th>node 2</th>
<th>node 3</th>
<th>node 4</th>
<th>node 5</th>
<th>node 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>0.400</td>
<td>0.400</td>
<td>0.400</td>
<td>0.400</td>
<td>0.400</td>
<td>0.400</td>
<td>0.400</td>
</tr>
<tr>
<td>4</td>
<td>0.082</td>
<td>0.082</td>
<td>0.154</td>
<td>0.194</td>
<td>0.152</td>
<td>0.152</td>
<td>0.152</td>
</tr>
<tr>
<td>3</td>
<td>0.066</td>
<td>0.070</td>
<td>0.082</td>
<td>0.114</td>
<td>0.130</td>
<td>0.152</td>
<td>0.152</td>
</tr>
<tr>
<td>2</td>
<td>0.064</td>
<td>0.066</td>
<td>0.068</td>
<td>0.062</td>
<td>0.118</td>
<td>0.118</td>
<td>0.118</td>
</tr>
<tr>
<td>1</td>
<td>0.063</td>
<td>0.064</td>
<td>0.062</td>
<td>0.070</td>
<td>0.114</td>
<td>0.114</td>
<td>0.114</td>
</tr>
<tr>
<td>0</td>
<td>0.051</td>
<td>0.063</td>
<td>0.064</td>
<td>0.064</td>
<td>0.111</td>
<td>0.111</td>
<td>0.111</td>
</tr>
<tr>
<td>5</td>
<td>0.047</td>
<td>0.051</td>
<td>0.051</td>
<td>0.051</td>
<td>0.111</td>
<td>0.111</td>
<td>0.111</td>
</tr>
<tr>
<td>6</td>
<td>0.043</td>
<td>0.043</td>
<td>0.043</td>
<td>0.043</td>
<td>0.111</td>
<td>0.111</td>
<td>0.111</td>
</tr>
</tbody>
</table>

**Huffman Coding**

- The Huffman code is completely reversible, i.e., lossless.
- The table for the translation has to be stored together with the coded image.
- The resulting code is unambiguous.
  - That is, for the previous example, the encoded string `011011101011` can only be parsed into the code words `0, 110, 1110, 1011` and decoded as `7, 4, 5, 0`.
- The Huffman code does not take correlation between adjacent pixels into consideration.
Interpixel Redundancy

Also called spatial or geometric redundancy

- Adjacent pixels are often correlated, i.e., the value of neighboring pixels of an observed pixel can often be predicted from the value of the observed pixel.

Coding methods:

- Run-length coding
- Difference coding

Run-length coding

- Every code word is made up of a pair \((g, l)\) where \(g\) is the gray level, and \(l\) is the number of pixels with that gray level (length or "run").
- E.g., \(56\ 56\ 56\ 82\ 82\ 82\ 83\ 80\ 80\ 80\ 80\)
  creates the run-length code \(56,3\)(82,3)(83,1)(80,4)(56,5)
- The code is calculated row by row.

Difference Coding

- Definition: \(f(x_i) = \begin{cases} x_i & \text{if } i = 0, \\ x_i - x_{i-1} & \text{if } i > 0. \end{cases} \)
- E.g.,

  
  original: 56 56 56 82 82 82 83 80 80 80 80
  
  code: \(f(x_i): 56\ 0\ 0\ 26\ 0\ 0\ 1\ -3\ 0\ 0\ 0\)

  - The code is calculated row by row.

  - Both run-length and difference coding are reversible and can be combined with, e.g., Huffman coding.

Difference and Huffman Coding in Combination

<table>
<thead>
<tr>
<th>Original Image</th>
<th>Difference Coding</th>
</tr>
</thead>
<tbody>
<tr>
<td>9 8 7 7 7 5 5 5</td>
<td>-1 -1 0 0 -2 0 0 0</td>
</tr>
<tr>
<td>7 7 7 7 4 4 5 5</td>
<td>0 0 0 1 0 -1 0 0</td>
</tr>
<tr>
<td>6 6 6 9 9 9 6 6</td>
<td>-1 0 1 0 0 -3 0 0</td>
</tr>
<tr>
<td>9 9 7 7 7 9 9 9</td>
<td>0 -1 0 0 -2 0 0 3</td>
</tr>
<tr>
<td>3 7 7 8 8 8 3 3</td>
<td>-3 0 1 0 0 -5 0 0</td>
</tr>
<tr>
<td>3 3 3 3 3 3 3 3</td>
<td>0 0 0 0 0 0 0 0</td>
</tr>
<tr>
<td>10 10 11 7 7 7 6 6</td>
<td>7 -1 0 -6 0 0 -1 0</td>
</tr>
<tr>
<td>4 4 5 5 5 2 2 6</td>
<td>-1 0 0 3 0 -4 0</td>
</tr>
</tbody>
</table>

Difference and Huffman Coding in Combination

<table>
<thead>
<tr>
<th>Original Image</th>
<th>Difference Coding</th>
</tr>
</thead>
<tbody>
<tr>
<td>9 8 7 7 7 5 5 5</td>
<td>-1 -1 0 0 -2 0 0 0</td>
</tr>
<tr>
<td>7 7 7 7 4 4 5 5</td>
<td>0 0 0 1 0 -1 0 0</td>
</tr>
<tr>
<td>6 6 6 9 9 9 6 6</td>
<td>-1 0 1 0 0 -3 0 0</td>
</tr>
<tr>
<td>9 9 7 7 7 9 9 9</td>
<td>0 -1 0 0 -2 0 0 3</td>
</tr>
<tr>
<td>3 7 7 8 8 8 3 3</td>
<td>-3 0 1 0 0 -5 0 0</td>
</tr>
<tr>
<td>3 3 3 3 3 3 3 3</td>
<td>0 0 0 0 0 0 0 0</td>
</tr>
<tr>
<td>10 10 11 7 7 7 6 6</td>
<td>7 -1 0 -6 0 0 -1 0</td>
</tr>
<tr>
<td>4 4 5 5 5 2 2 6</td>
<td>-1 0 0 3 0 -4 0</td>
</tr>
</tbody>
</table>
Psycho-Visual Redundancy

- If the image will only be used for visual observation much of the information is usually psycho-visual redundant. It can be removed without changing the visual quality of the image. This kind of compression is usually lossy.

50 kB (uncompressed TIFF) 5 kB (JPEG)

Psycho-Visual Redundancy Reduced by Quantization

- 256 gray levels (8 bit) to 16 gray levels (4 bit)

Transform Coding

1. Divide the image into blocks.
2. Transform each block with discrete cosine transform or discrete Fourier transform.
3. The transformed blocks are truncated to exclude the least important data.
4. Code the resulting data using variable length coding, e.g., Huffman coding in a zigzag scan pattern.
File Formats with Lossy Compression

- **JPEG**, Joint Photographic Experts Group, based on a cosine transform on 8x8 pixel blocks and Run-Length coding. Give arise to ringing and block artifacts. (.jpg, .jpeg)
- **JPEG2000**, created by the Joint Photographic Experts Group in 2000. Based on wavelet transform and is superior to JPEG. Give arise only to ringing artifacts and allows flexible decompress and reading. (.jpg, .jp2)

File Formats with Lossless Compression

- **TIFF**, Tagged Image File Format, flexible format often supporting up to 16 bits/pixel in 4 channels. Can use several different compression methods, e.g., Huffman, LZW.
- **GIF**, Graphics Interchange Format. Supports 8 bits/pixel in one channel, that is only 256 colors. Uses LZW compression. Supports animations.
- **PNG**, Portable Network Graphics, supports up to 16 bits/pixel in 4 channels (RGB + transparency). Uses Deflate compression (LZW and Huffman). Great when interpixel redundancy is present.

Choosing image file format

- **Image analysis**
  - Lossless formats are vital. TIFF supports a wide range of different bit depths and lossless compression methods.
- **Images for use on the web**
  - JPEG for photos (JPEG2000 in the future), PNG for illustrations. GIF for small animations. Web browsers have a increasing support for the vector format SVG.
- **Line art, illustrations, logotypes, etc.**
  - Lossless formats such as PNG etc. (or a vector format)

Suggested Exercises

- 6.2, 6.4, 6.6
- 8.1, 8.3, 8.6