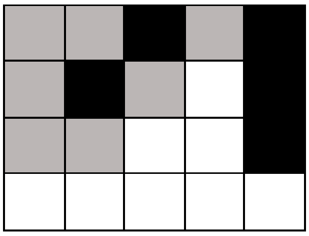
Exercise Multimedia Technology
WS 2003/2004

Sheet 11 (January 23th, 2004)

Exercise 11.1 Color analysis

The image below shows a magnified image clipping.



- Invent an algorithm for calculating an image's histogram. You can expect the image to be of size N x M using 256 different gray values per pixel.

```
void generateHistogram(unsigned char image[][]) {
  int histogram[256];
  for(int index = 0; index < 256; index++) histogram[index] = 0;

    for(int y = 0; y < N; y++) {
      for(int x = 0; x < M; x++) {
        int index = image[x][y];
        histogram[index]++;
      } // for
    } // for
} // generateHistogram
```

- Calculate the histogram for the image above.

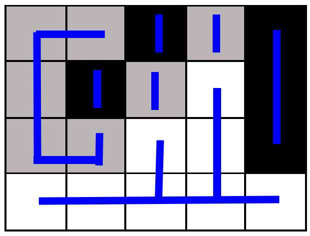
  Black = 5 pixels

  Gray = 7 pixels

  White = 8 Pixels

- Calculate the CCV based on the image above. The neighborhood shall be defined as a 4-pixel neighborhood, which means that only vertical and horizontal pixels are considered but no diagonal ones. How many regions does the above image contain? What is the average size of a region in pixels?

The following figure shows each of the 7 unique regions as a blue (connected) line.



7 regions in an image of about 20 pixels means that the average region size is about 2.86 pixels.

The color coherence vector has a 2D entry for each gray (or color) value which is contained in the image. The first value of a particular entry stands for the percentage of pixels below the average region size, the second entry stands for the percentage of pixels above the average region size (or vice versa, we don't care, but make clear what you mean).

Example for the gray pixels: There is one region of 5 pixels (> 2.86) and here are two regions of 1 pixel (< 2.86). So 2 out of 20 pixels are below the average region and 5 pixels out of 20 are above the average region.

CCV = <(2/20, 3/20), (2/20, 5/20), (0/20, 8/20)> (in the order black, gray, white)

Note: If you provided the absolute number of pixels in the CCV as suggested in the paper, this is also ok.

Note: In case you might want to have further information on color coherence vectors, read section 3 of the following paper. It contains a very compact description of CCV. The threshold *tau* mentioned in the paper corresponds to the average size of the regions calculated above.

http://www.informatik.uni-mannheim.de/informatik/pi4/data/pass96comparing.pdf

Exercise 11.2 Cut detection

(1) How would you compare two histograms H_1 and H_2 ?

Each histogram is a vector with 256 components. So any metrics for vectors would do the job (for example, the euclidean one).

(2) Develop an algorithm which detects cuts based on color histograms. The image should be given

as char image[height][width][frame].

Generate the histogram H(t) for frame t and H(t+1) for frame t+1 as in the first exercise 11.1 and calculate the difference as suggested in exercise 11.2(1). If the distance between two histograms exceeds a predefined threshold, there might be a cut between t and t+1.

(3) Comment on the reliability of histograms in the context of cut-detection and provide solutions to possible problems.

When verifying hypotheses statistically, two kinds of mistakes can occur which are referred to as alpha- and beta-mistakes. Alpha mistakes happen if the hypothesis was rejected even though it was actually true. Beta mistakes mean that the hypothesis was accepted even though is was actually not justified. So in doing cut-detection, we assume that there is a cut between t and t+1 (the hypothesis).

Changing histograms are not always caused by cuts. A sudden change of the image could as well be caused by an explosion in the film or by changing light conditions. If, for example, politicians are surrounded by reporters, simple cut detection algorithms often misinterpret the hails of flash bulbs as being many cuts. This is a typical example of a beta-mistake. To solve it one could try to find patterns of many cuts each occurring rapidly in turn.

On the other hand, filmmakers use a variety of cut techniques other than hard cuts (e. g., the so-called disolves). These effects, which gradually change from one image to another, result in gradually changing histograms as well. Histogram-based cut-detection approaches are not suitable for these cases, since they produce alpha-mistakes.

Exercise 11.3 Gradients

The magnitude of the gradient of an image I (x,y) is derived as follows:

- $|\nabla I(x,y)| = \sqrt{\left(\frac{\partial I(x,y)}{\partial x}\right)^2 + \left(\frac{\partial I(x,y)}{\partial y}\right)^2}$

The partial derivations can e. g., be approximated as follows:

- $\frac{\partial I(x,y)}{\partial x} \approx \frac{I(x+1,y) - I(x-1,y)}{2}$
- $\frac{\partial I(x,y)}{\partial y} \approx \frac{I(x,y+1) - I(x,y-1)}{2}$

The following table gives you the gray values of an image.

100	100	200	100	200	100
100	100	200	100	200	100
100	100	200	200	200	100
100	100	200	200	200	100
100	100	200	100	200	100
100	100	200	100	200	100

Calculate the gradients for the second row and the gradient magnitudes for the third column. If you need additional values at the borders, set them to zero.

<(100-0)/2, (100-100)/2>, <(200-100)/2, (100-100)/2>, <0, 0>, <0, 50>, <0, 0>, <-100, 0>

100, 0, 50, 50, 0, 100

To what kind of visual effect does the gradient yield? Explain why.

The gradient image (or to be even more precise, the gradient magnitude image) surrounds regions of equal gray values with an outline, especially if they come along with a significantly differing background. Highly textured backgrounds remain highly textured because every change in gray values results in two gradient pixels, one for the change from the background to the foreground and one for the second change back to the background.

Exercise 11.4 Panoramic images

(1) In the lecture you have learned how to produce a tube-like reconstruction of several images into a panoramic one. Extend the approach to a spherical projection. What additional degrees of freedom does the spherical approach give the photographer?

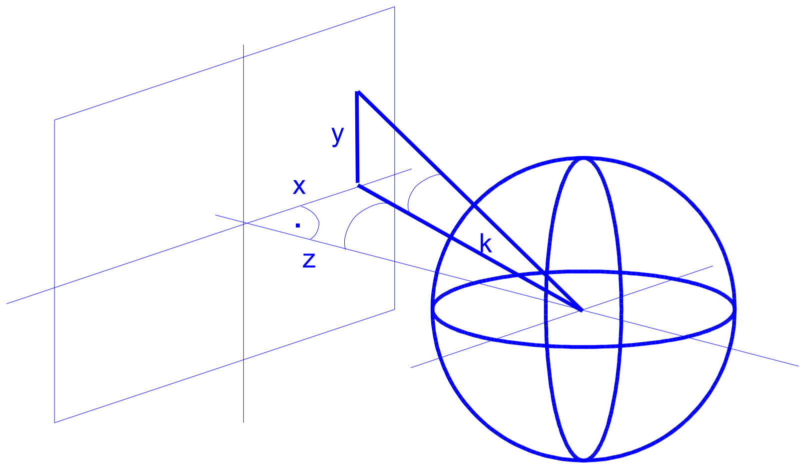
In a spherical projection the photographer is allowed not only to rotate the camera around the vertical axis but also around the horizontal axis.

Consider the slide *Cylindrical Panoramas (1)* in lecture 8.3-15. The cylindrical projection projected a flat image onto a cylinder. Thus the coordinates (x, y) were mapped onto the coordinates (theta, v).

The spherical projection differs from the cylindrical one only by the fact that two angles rather than only one are needed. As we already know the angle theta for the horizontal component, we need another angle for the vertical component – let's call it alpha.

α

θ



theta could be computed since x and z are known and they are perpendicular to one another. Exactly the same holds true for y and k. Thus the angle alpha is computed as follows:

$$k = \sqrt{x^2 + z^2}$$
$$y = \tan(\alpha) k$$
$$\Rightarrow y / k = \tan(\alpha)$$
$$\Rightarrow \alpha = \arctan\left(y / \sqrt{x^2 + z^2}\right)$$

(2) What kinds of transforms have been applied between frame 1 and frame 2 (see next figure)? Find a parametric description for each transform.

Note: Let's assume that the lower left corner of the square in frame 1 is the origin of the coordinate system.

(1) The transform is a shear into
the horizontal direction

$$\begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

(2) The transform is a rotation and
a translation

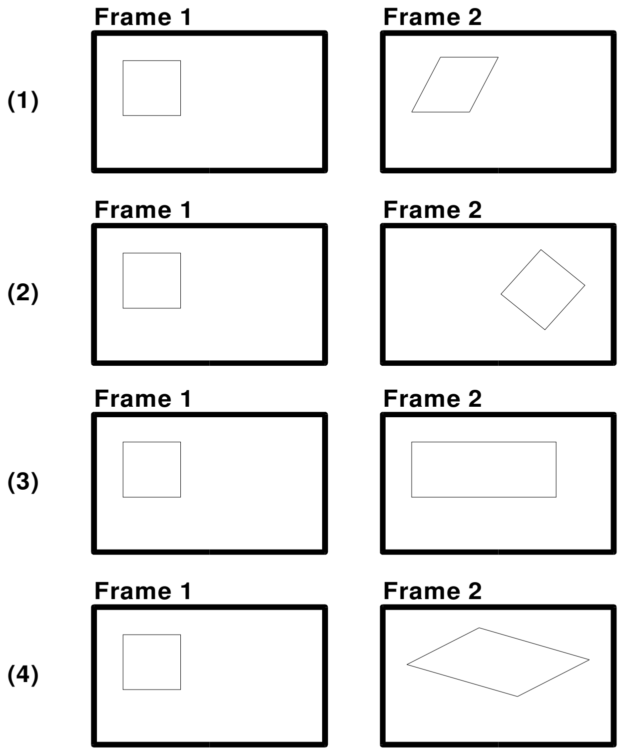
$$\begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix}$$

$$\begin{pmatrix} \cos(\alpha) & -\sin(\alpha) & t_x \\ \sin(\alpha) & \cos(\alpha) & t_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

(3) The transform is a scaling of
the horizontal component

$$\begin{pmatrix} s & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

(4) There is no unique solution to
this transform. It might have
been a shear, then a
horizontal scaling and finally a rotation but other transforms are
possible as well.



(3) Prove that a rotation can be expressed by means of two shears.

Actually, using only two shears only allows rotations of about 0 degrees. So of course you have to use three shears rather than two. The following equation shows that the rotation can be replaced by three shears in 2D.

$$\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix}\begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix} =$$

$$\begin{pmatrix} 1+ab & a \\ b & 1 \end{pmatrix}\begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1+ab & c+abc+a \\ b & bc+1 \end{pmatrix}$$

$$\begin{pmatrix} 1+ab & c(1+ab)+a \\ b & bc+1 \end{pmatrix} \overset{!}{=} \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

(I) $1+ab = \cos\alpha$ (III) in (I) $\Rightarrow 1 + a\sin\alpha = \cos\alpha \Rightarrow a = \frac{\cos\alpha - 1}{\sin\alpha}$

(II) $c(1+ab)+a = -\sin\alpha$

(III) $b = \sin\alpha$

(IV) $bc+1 = \cos\alpha$ (III) in (IV) $\Rightarrow \sin\alpha \cdot c + 1 = \cos\alpha \Rightarrow c = \frac{\cos\alpha - 1}{\sin\alpha}$

a, b, c in (II) $\Rightarrow \frac{\cos\alpha - 1}{\sin\alpha}\left(1 + \frac{\cos\alpha - 1}{\sin\alpha} \cdot \sin\alpha\right) + \frac{\cos\alpha - 1}{\sin\alpha} = -\sin\alpha \quad | \cdot \sin\alpha$

$\Rightarrow (\cos\alpha - 1)(1 + \cos\alpha - 1) + \cos\alpha - 1 = -\sin^2\alpha$

$\Rightarrow \cos^2\alpha - \cos\alpha - 1 + \cos\alpha = -\sin^2\alpha$

$\Rightarrow \sin^2\alpha + \cos^2\alpha = 1$ ✓

Remember: