Chapter 13

II - Surface features

This chapter focusses on features that characterise surfaces. In particular, a first part discusses how we can characterise surface colors and a second part discusses ways do describe surface textures.

13.1 Surface colours

Already intuitively it is clear that colour can be a very useful feature for image analysis. Being able to observe and measure the colour of materials and objects makes it easier to detect, recognise, or track them.

This section consists of two parts. In the first we indicate how raw colour data can be characterised or represented. To that end, colour spaces are introduced. In the second part, we discuss the extraction of colour features from the raw colour data. These can then be used in the type of applications just mentioned.

13.2 The representation of colour

13.2.1 Generating colours from primaries

In section 3.2 the principles underlying human colour vision were explained. The same principles will be used for the reproduction of colour with sets of three light sources. Applications include the generation of required colours on displays and – as the reverse process – the interpretation of output signals of colour cameras.

Generating colours with primaries

In view of the three-cone theory, the technology of colour displaying is normally based on three different types of sources (e.g. phosphors). These three light sources are the *primary sources* or *primaries*. In general, they will be given rather than being free to choose. First, we need to know their spectral radiant flux $P_j(\lambda), j = 1, 2, 3$. The primary

sources recommended by the CIE are three monochromatic sources with wavelengths $\lambda_1 = 700, \lambda_2 = 546.1, \lambda_3 = 435.8$. According to table 2.1 the corresponding colours are *red, green, and blue.* They are the basis of several standard representations of the colour space, to be discussed later.
For practical applications, such as TV broadcasting, different standard primaries have been defined. These correspond to physically realizable and economically producable sources, that are also sufficiently bright to be useful. Bodies like the EBU [1] and NTSC [2] have defined (different) standard primaries for TV broadcasting. We will refer to these primaries as R_r, G_r, B_r and R_N, G_N, B_N, respectively.

For the following analysis, the precise choice of the primaries is not crucial. Given a source with spectral composition $C(\lambda)$ the colour of which should be matched by mixing the primaries, the question is how much we need from each. The amounts $m_j, j = 1, 2, 3$ should be such that $\sum_{j=1}^{3} m_j P_j(\lambda)$ is perceived as $C(\lambda)$. Therefore, the human cone responses R_i to the colour should equal the response to the mixture of primaries. Hence,

$$R_i(C) = \begin{cases} \int H_i(\lambda) C(\lambda) d\lambda & \text{for the source} \\ \int \left[\sum_{j=1}^{3} m_j P_j(\lambda)\right] H_i(\lambda) d\lambda = \sum_{j=1}^{3} m_j \int H_i(\lambda) P_j(\lambda) d\lambda & \text{for the primaries} \end{cases}$$

where the H_is are the cone sensitivities of human vision introduced in chapter ??. Using shorthand notation

$$l_{i,j} = \int H_i(\lambda) P_j(\lambda) d\lambda \ ,$$

we see that these integrals can be determined "off-line" so to speak. They are fixed once the primaries have been chosen. For the monochromatic CIE primaries the situation is especially simple since $l_{i,j} = H_i(\lambda_j)$. Knowing the human responses R_i to the source, we obtain a linear system of equations to be solved for the m_j:

$$\sum_{j=1}^{3} m_j l_{i,j} = R_i.$$

Notice that solving this system implies inverting the matrix of the $l_{i,j}$. This will be possible as long as the primaries are independent with respect to human vision, i.e. none of them can be produced as a linear combination of the other two. Also notice that changing the primaries, and therefore the l_{ij}s, leads to a linear transformation between the m_js in the two systems of primaries.

Tristimulus and chromaticity coordinates

Since "white" can be considered a natural reference, one will usually specify relative values with respect to the amounts of the primaries needed for some standard white source, written w_j. These *tristimulus values* for the source $C(\lambda)$ are then $T_j = \frac{m_j}{w_j}$. By definition, the standard white source has tristimulus values 1. Deviation from 1 then expresses there

[1] European Broadcasting Union
[2] National Television Systems Committee (USA)

being more or less of the corresponding primary than in the standard white. Such colour coordinates therefore give a more direct idea of the corresponding colour. The corresponding scaling preserves the linearity of the transformations between the systems of primaries. For the CIE primaries, the corresponding tristimulus values are generally called R,G,B, after the colours they represent (red,green,blue). The CIE primaries have been chosen such that CIE standard white W, defined as having a flat energy spectrum in the visible spectrum, has $w_1 = w_2 = w_3$. In order to achieve this the relative power $P_1 : P_2 : P_3$ of the primaries R, G, B, resp., has to be in the proportion 72.1:1.4:1.0. In that case there is not much of a difference between the m_j and the tristimulus values. The value of the w_i for the CIE choice of primaries and white is obtained as follows. First, CIE white having a flat spectrum, $R_i(W) = \int H_i(\lambda)d\lambda$. Thus,

$$\sum_{j=1}^{3} w_j P_j H_i(\lambda_j) = \int H_i(\lambda)d\lambda.$$

This also explains the relative power values for the primaries. Considering the spectral sensitivity curves (see chapter ??), the area below the curve for blue clearly is much less, hence the lower value for the blue primary, although it is primarily this primary which will have to generate output for these cones. On the other hand, the area below the curve sensitive to the largest wavelengths ("red" cones) is comparable to that of the "green' cones, but shows appreciably lower sensitivities to the green primary. This has to be compensated by the red primary, to which sensitivity is still rather low. Hence, the power for the red primary has to be considerable compared to the others. Solving for the w_j one would find they are equal and therefore:

$$w_j = \frac{\int H_i(\lambda)d\lambda}{\sum_{j=1}^{3} P_j H_i(\lambda_j)}.$$

The *spectral matching curves* $T_j(\lambda')$ give the tristimulus values for monochromatic sources $C_{\lambda'}$ with wavelength λ':

$$R_i(C_{\lambda'}) = H_i(\lambda') = \sum_{j=1}^{3} w_j l_{i,j} T_j(\lambda').$$

These are very useful, since knowing the spectral matching curves, the tristimulus values for an arbitrary source $C(\lambda)$ are found as

$$T_j(C) = \int C(\lambda)T_j(\lambda)d\lambda.$$

The spectral matching curves for the CIE primaries are given in fig. 13.1. Negative values indicate colours that cannot be reproduced with the CIE primaries. The corresponding values indicate how much of the corresponding primary should be added *to the given light source* to produce the combination of the other two primaries with positive values. Not being able to produce some colours is by no means a specific drawback of the CIE choice. Whatever choice for the primaries is made, there will always be colours that cannot be

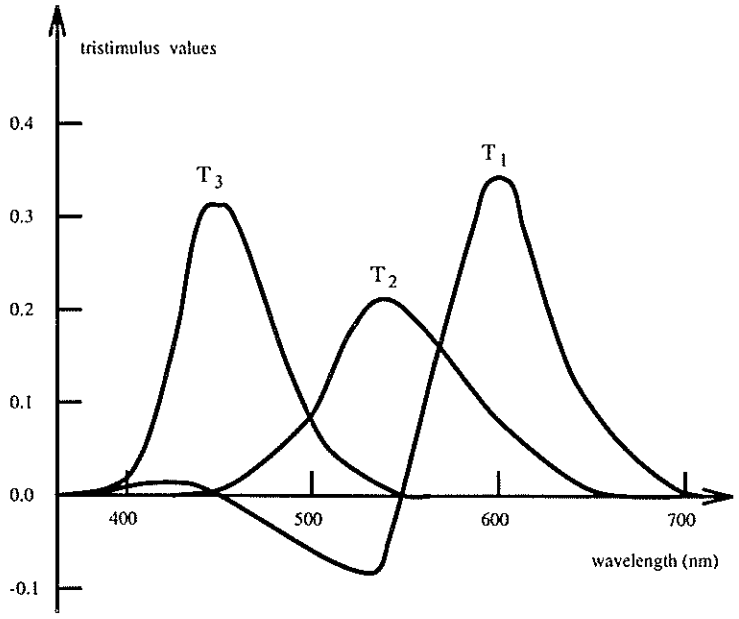


Figure 13.1: *Spectral matching curves for the CIE primaries.*

reproduced if the primaries are to be physically possible. We will come back to this issue later.

The tristumulus representation of colour still contains brightness information. We see that in accordance with Grassman's results, briefly mentioned in section 3.2, multiplying the spectral composition of the colour $C(\lambda)$ by some factor, the tristimulus values will simply scale accordingly. These values therefore still don't capture chrominance information pure. This can be remedied by normalising the tristimulus values, resulting in the following set of *chromaticity coordinates*:

$$t_j = \frac{T_j}{T_1 + T_2 + T_3}.$$

It follows immediately that $t_1 + t_2 + t_3 = 1$. This linear dependence allows the elimination of one coordinate, leading to a two-dimensional colour space. We can think of this colour space as a horizontal section through a squeezed version of the colour solid of fig. 3.2. A pair of chromaticity coordinates specifies saturation and hue, but doesn't contain information on the luminance.

The chromaticity coordinates (r, g) corresponding to the CIE primaries are calculated on the basis of the CIE tristimulus values R, G, B:

$$r = \frac{R}{R+G+B} \qquad g = \frac{G}{R+G+B}.$$

The corresponding colour space is shown in fig. 13.2.

The curved solid line in fig. 13.2 is the *spectrum locus.* It represents the set of monochromatic colours, also referred at as *spectrum colours.* The position of the three monochromatic CIE primaries are indicated and correspond to the points with coordinates $(0, 0)$

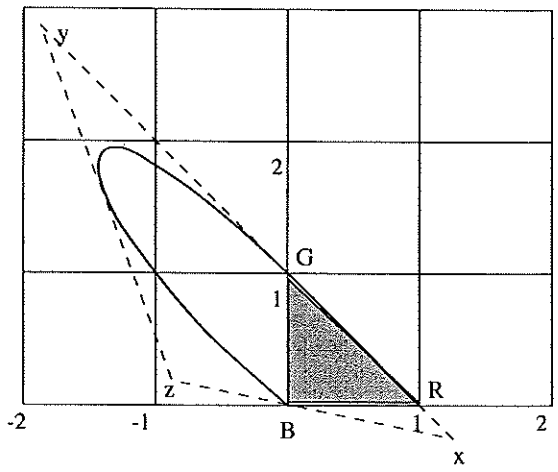


Figure 13.2: *The r-g colour space, with horizontal r and vertical g axis.*

(blue), (1,0) (red), and (0,1) (green). We note that, due to the negative portion of the $r(\lambda)$ curve, we again encounter negative values along the spectral locus. The line segment connecting the extremes of the spectrum locus (close to B and R) is the *line of purples*. These are generated by mixing the blue and red primaries. Colours in the triangle formed by the reference white (coordinates (1/3, 1/3)) and the line of purples are *nonspectral colours*. The only colours that are visible to the human eye are those in the region delineated by the spectral locus and the line of purples. Colours inside this region but outside the triangle of nonspectral colours are called *spectral colours*. Notice the confusing terminology: spectrum colours and spectral colours are not the same! The colours that can be produced as a mixture of the primaries are confined to the triangle $R - G - B$. Other colours have at least one negative chromaticity coordinate. This triangle has been highlighted in the figure.

Transition to other coordinate systems

In order to get rid of the negative values with the CIE primaries, a virtual or normalized tristimulus colour system X, Y, Z of purely positive coordinates has been defined. The corresponding primaries are purely hypothetical, they can in fact not be physically realized. Nevertheless, it is a convenient coordinate system since all colours have corresponding chromaticity coordinates between 0 and 1, as shown by the overlay x, y, z coordinate system in fig. 13.2 (dashed lines). Transitions between primary triples are governed by linear transformations, as noted earlier. The linear transformation from R, G, B to X, Y, Z coordinates is given by:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.490 & 0.310 & 0.200 \\ 0.177 & 0.813 & 0.011 \\ 0.000 & 0.010 & 0.990 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

The transformation has been chosen as to make Y represent luminance. One might be amazed that the coefficients 0.177, 0.813, and 0.011 do not reflect the relative luminous

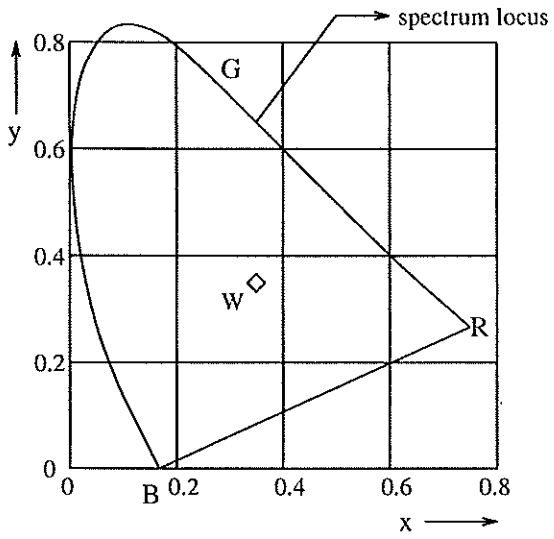


Figure 13.3: *The x-y colour coordinate representation of colour space.*

efficiencies at the wavelengths of R, G, and B, since Y was said to be the luminance. One should not forget, however, that the CIE primaries have been normalized with respect to reference white. For instance, 0.177 is obtained as

$$\frac{v(\lambda_1)72.1}{v(\lambda_1)72.1 + v(\lambda_2)1.4 + v(\lambda_3)1.0}.$$

The coordinates of the reference white remain invariant: $R = G = B = 1$ is mapped to $X = Y = Z = 1$. The chrominance coordinates x, y are found as

$$x = \frac{X}{X+Y+Z} \qquad y = \frac{Y}{X+Y+Z}.$$

The x, y colour space is shown in fig. 13.3. This is the standard CIE *colour triangle* representation usually given in textbooks. The EBU primaries have coordinates

$$\begin{aligned} R_r: &\quad x = 0.64 \quad y = 0.33 \\ G_r: &\quad x = 0.29 \quad y = 0.60 \\ B_r: &\quad x = 0.15 \quad y = 0.06 \end{aligned}$$

and for the NTSC the coordinates are

$$\begin{aligned} R_N: &\quad x = 0.67 \quad y = 0.33 \\ G_N: &\quad x = 0.21 \quad y = 0.71 \\ B_N: &\quad x = 0.14 \quad y = 0.08 \end{aligned}$$

Since colours outside the triangle of the primaries in a chromaticity diagram cannot be produced, the area outside the triangle should be kept small. From fig. 13.2 it seems the CIE primaries are not a particularly good choice in this respect. Minimizing the area

outside the triangle made up of the primaries is not a good criterion for the comparison of primary choices, however. As will be shown shortly, chromaticity coordinates are changed by projective transformations upon changing the primaries. By choosing different primaries, i.e. applying projective transformations, relative areas can be changed almost arbitrarily. In fact, with the CIE R, G, B choice most actual colours can be produced indeed. The fact that some colours near the spectrum locus cannot be produced is all the less important since pure spectrum colours are rather rare in nature.
Also, relative distances between different points in such diagrams do not necessarily indicate relative perceptual differences. A distance in the x-y chromaticity diagram's green region will correspond to a smaller perceptual difference than the same distance in the blue region, for instance. The same holds even more so for the r-g diagram. The CIE therefore has proposed a particular projective transformation that generates a diagram were length more faithully represents perceptual distance. This u-v chromaticity diagram [3] is obtained as follows:

$$u = \frac{4x}{-2x + 12y + 3} \qquad v = \frac{6y}{-2x + 12y + 3}.$$

Practical primaries (e.g. phosphors of a display) are not monochromatic but have a continuous spectral distribution. It is clear that for displaying purposes one should then start from the corresponding primaries, not from the CIE tristimulus coordinates. Physically feasible sources of high luminance that together can produce a wide gamut of colours and can be produced at low cost are preferred. We already mentioned the EBU and NTSC standards for TV. A first thing to note is that the corresponding triangles will fall strictly within the area of perceptable colours. The EBU primaries are shown in fig. 13.4 as R_r, G_r, B_r. As a consequence spectrum colours cannot be produced and at least one of the spectral matching curves will be negative at any wavelength. This is shown in fig. 13.5. For practical purposes the curves are therefore reduced to their positive part. The EBU system also has its standard white C, also indicated in figure 13.4. The x-y chormaticity coordinates of C are $x_C = 0.310$ and $y_C = 0.316$. This reference has been chosen because it corresponds rather well to medium daylight.

Changing the primaries boils down to considering a new triangle in the chromaticity diagram, with the corresponding source colours as its vertices. Two sides of such triangle are to be used as the basis of a new coordinate frame. Moreover, the standard white of one system has to be transformed onto the standard white of the other. This yields four points to be matched and therefore fixes a two-dimensional projective transformation (which has 8 independent degrees of freedom, see also chapter ??).

At this point it is useful to briefly discuss the transformation types relevant for transitions from one set of primaries to another on the level of the different coordinate types. First, on the level of the amounts m_j a linear transformation was shown to take place. For the tristimulus values the reference to white light takes the form of additional scaling

[3] Officially referred to as the CIE 1976 UCS chromaticity diagram. UCS stands for uniform chromaticity scale.

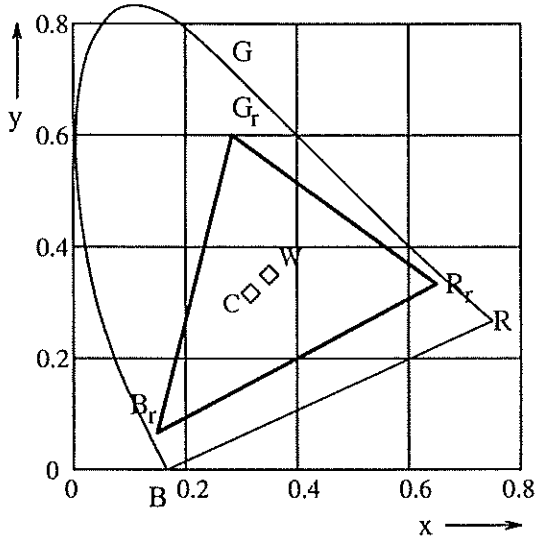


Figure 13.4: *The EBU primaries in the x-y represntation.*

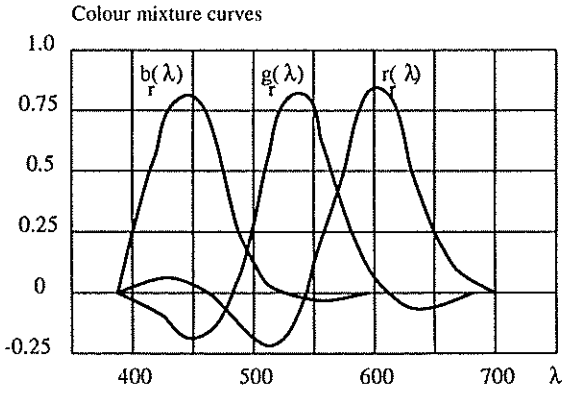


Figure 13.5: *Spectral matching curves for the EBU primaries.*

along the three axes. This operation preserves the linearity of the overall transition:

$$\begin{pmatrix} T_1' \\ T_2' \\ T_3' \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix} \begin{pmatrix} T_1 \\ T_2 \\ T_3 \end{pmatrix}$$

where the c_{ij}s are real constants. Hence, the transformation from R, B, B or X, Y, Z to the EBU and NTSC primaries is a linear transformation. For instance, that from X, Y, Z coordinates to the corresponding NTSC R_N, G_N, B_N coordinates is

$$\begin{pmatrix} R_N \\ G_N \\ B_N \end{pmatrix} = \begin{pmatrix} 1.910 & -0.533 & -0.288 \\ -0.985 & 2.000 & -0.028 \\ 0.058 & -0.118 & 0.896 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} .$$

Note that the tristimulus values are linearly transformed, regardless whether the same standard white is used or not.

The corresponding chromaticity coordinates then can be shown to transform as follows:

$$\begin{aligned} t_1' &= \frac{(c_{11} - c_{13})t_1 + (c_{12} - c_{13})t_2 + c_{13}}{(c_{.1} - c_{.3})t_1 + (c_{.2} - c_{.3})t_2 + c_{.3}} , \\ t_2' &= \frac{(c_{21} - c_{23})t_1 + (c_{22} - c_{23})t_2 + c_{13}}{(c_{.1} - c_{.3})t_1 + (c_{.2} - c_{.3})t_2 + c_{.3}} , \end{aligned}$$

with

$$c_{.1} = c_{11} + c_{21} + c_{31} , \qquad c_{.2} = c_{12} + c_{22} + c_{32} , \qquad c_{.3} = c_{13} + c_{23} + c_{33} .$$

This is a projective transformation in a two-dimensional space. An example is offered by the transition from fig. 13.2 (r, g) to fig. 13.3 (x, y). It might seem that the skewing and rescaling could be modeled by an affine transformation (see chapter ??). Comparing the area of the $R - G - B$ triangle with the area falling outside the triangle, it is clear that the relative area is not preserved, however. This shows the corresponding transformation isn't affine (see chapter ??).

* Identifying a colour on the basis of chromaticity coordinates

Given chromaticity coordinates and the position of the spectrum locus and reference white, it is possible to obtain an idea of how the colour will look like. The procedure is illustrated in fig. 13.6. One draws a line connecting white and the given colour. This line intersects the spectrum locus at the *dominant wavelength* of the colour, i.e. yields its hue. The colour is a mixture of the dominant monochromatic (spectrum) colour and the reference white. A qualitative indication for the saturation of the colour is the proportion A/B. *Complementary chromaticities* are chromaticities containing the reference white on the line that connects them. In appropriate mixtures these colours will produce reference white. Such colours are *complementary colours.*

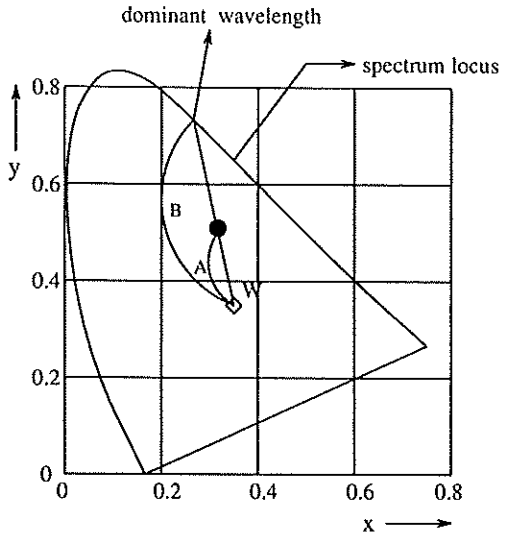


Figure 13.6: *Interpretation of tristimulus values.*

This procedure is very simple and allows to gain insight in the colour without having a coloured chromaticity diagram, which always is easiest of course. Things get more complicated if the chromaticities are given with respect to other primaries and/or white than the ones used for the construction of the given chromaticity diagram. Given the position of the primaries and white used for the generation of the chromaticity coordinates in the available chromaticity diagram, one way to go about is the following. First, consider a simple Cartesian system of coordinates where the primaries A, B, C used for the determination of the given chromaticity coordinates have coordinates (0,0), (1,0), and (0,1). In this system the reference white W will have Cartesian coordinates (0.33,0.33) and the given colour D has the given chromaticity coordinates (d_1, d_2) as its Cartesian coordinates. The situation is depicted in the left half of fig. 13.7. The construction shown in the figure yields four points on both coordinate axes. This allows the calculation of two *cross-ratios.* Cross-ratios are expressions that remain unchanged, i.e. invariant, under projective transformations. Since we know that the chromaticity coordinates are transformed projectively, such expression will be kept unchanged by the transformation. A more detailed discussion on cross-ratios will be given in chapter ??. Let it for the moment suffice to give the definition. If one has four collinear points P_1, P_2, P_3, P_4 with coordinates (x_1, y_1), (x_2, y_2), (x_3, y_3), and (x_4, y_4), then their cross-ratio (P_1, P_2, P_3, P_4) is defined as

$$(P_1, P_2, P_3, P_4) = \frac{\dfrac{x_3 - x_1}{x_2 - x_3}}{\dfrac{x_4 - x_1}{x_2 - x_4}} = \frac{\dfrac{y_3 - y_1}{y_2 - y_3}}{\dfrac{y_4 - y_1}{y_2 - y_4}}.$$

Using the coordinates of the four points along the horizontal axis the corresponding cross-

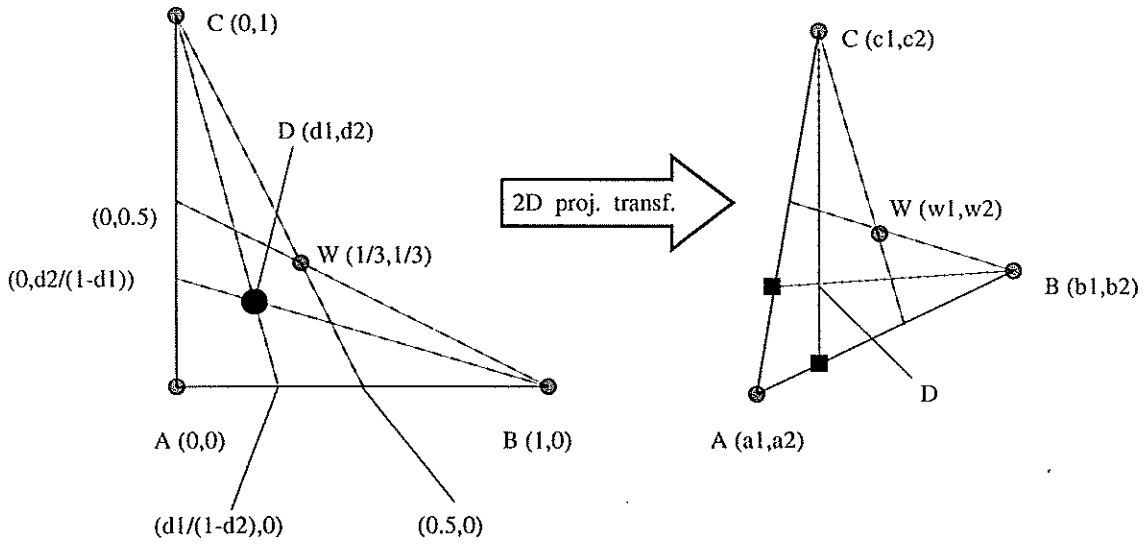


Figure 13.7: *Interpretation of chrominance coordinates in a different frame.*

ratio (numbering the points from left to right) is

$$\frac{d_1 + d_2 - 1}{2d_1 + d_2 - 1}.$$

Similarly, for the vertical axis and numbering the points form bottom to top, the cross-ratio

$$\frac{d_1 + d_2 - 1}{2d_2 + d_1 - 1}.$$

is found. The right half of the figure shows the primaries and white in the available chromaticity diagram. The construction shows the point triples along the lines A-B and A-C that are known: the projection of W on both axes through the opposite primaries C and B, resp., together with two primaries A, B and A, C. The construction is based on the preservation (invariance) of collinearity under projective transformations. Again the reader is referred to chapter ??. The two cross-ratios identify a single point along both axes, indicated by a black square. The construction with the two dotted lines yields D at their intersection. The colour can then be interpreted on the given chromaticity diagram, via the techniques outlined at the beginning of this section or directly on a coloured diagram.

A more direct approach is to determine the projective transformation that maps the four points A, B, C and W (cfr. at the left of fig. 13.7) to their images in the given chromaticity diagram (cfr. at the right of fig. 13.7). The position of the point D is then found by applying the same transformation to the given coordinates (d_1, d_2).

13.2.2 Surface colour as a feature

At first sight, extracting color features seems trivial once a colour image has been taken. Can't we just extract some chromaticity coordinates and use these? Well, unfortunately,

the answer is... in many cases NO! The responses for the three colour bands are not only a function of the reflectance characteristics of the surface, but also of the spectral composition of the illumination. The former is an inherent property of the object's surface and therefore useful for tasks like object recognition or tracking. The latter may change between views and interferes with our attempts to gain knowledge about the surface itself.

Colour constancy

Human colour perception is far more intricate than the collection of the three colour cone responses R_i, $i = 1, 2, 3$ as discussed in chapter ??. This is illustrated by the 'colour constancy' phenomenon. It shows very clearly that the cone responses at the input of the visual system – the retina (see 3.4.1) – undergo further processing along the visual pathway (see section 3.5).

Succinctly put, the colour constancy phenomenon amounts to objects keeping approximately the same perceived colour under changing illumination. Although the spectral reflectance of an object is an inherent characteristic of its surface, changing the illumination's spectral composition will change the spectral composition of the reflected light, and one would therefore expect that the colour would be seen as changing. This is not the case. The human visual system succeeds in transforming the triplet of cone responses into a percept that is almost independent of illumination. As a matter of fact, research indicates that people can consciously identify both the cone responses and the illumination invariant perception. This means that humans notice that the illumination changes and that they can estimate the qualitative aspects of these changes, but that they can simultaneously estimate the colours as they would be under white light. Next, we describe colour constancy in a bit more detail.

A name strongly connected with color constancy is that of Edwin Land. He performed a series of striking experiments with planar, colored patterns coined 'Mondrians', after the Dutch painter Piet Mondriaan. These patterns consist of different patches with a simple, polygonal shape, each with a single, uniform color. Land illuminated his Mondrians with three sources. Each source had a different colour, shining light in a narrow spectral band. He then selected two patches and took two different images. He regulated the sources so that in the second image the second patch reflected the same amounts of light for each of the sources as the first patch did in the first image. Going by what we have said about color perception in chapter ?? one would expect that the perceived colour of the first patch in the first image would match that of the second patch in the second image. But this was not the case. In fact, perceived colour of the same patch in the two images was what roughly stayed constant. Hence, the term 'color constancy'. Conversely, the color of a patch is seen to change when the light reflected from it is kept constant, while that reflected by the surrounding patches changes. The Colour Plate illustrates these points. This illustration shows the same scene twice. The left image shows it under green lighting, the one on the right under normal white light. Now look at the girl's hat in the left image. It looks red. A very similar colour is shown at the bottom right corner of that image. This, however, was copied and overlaid on the left image based on the original colour in the right image,

i.e. both bottom right corners have the same colour! Although the colour of this corner is identical in both, the perceived colour is different.

Also from a computer vision point of view, this capacity to estimate the reflectance characteristics of objects under variable illumination would be extremely useful. Several efforts have been made to model colour constancy. The strategy typically is to estimate the spectral composition of the source whereafter the scene is normalised by simulating how it would look like under a canonical illumination. Illumination independent colour measurements can then be made in the normalized image. On the whole, efforts to achieve true colour constancy have not been very successful so far. Nevertheless, one can extract colour features that remain invariant under changing illumination, even if they do not represent surface reflectance. But before we can discuss strategies for such normalization, we should first analyse the image formation process in some more detail.

Models for observed reflection

We start with the most accurate reflection model. It uses the spectral bidirectional reflectance distribution function f (or spectral BRDF for short)

$$f(x, y, \lambda, \mathbf{n}, \mathbf{e}, \mathbf{v}) = dL_r(x, y, \lambda, \mathbf{n}, \mathbf{v})/dE_i(x, y, \lambda, \mathbf{n}, \mathbf{e})$$

which indicates the radiance L_r of a surface in a direction $\mathbf{v}$ when it is illuminated from a direction $\mathbf{e}$ with irradiance E_i and this for a position (x, y) on the surface with normal $\mathbf{n}$ and for the different wavelengths λ. In order to calculate the reflected radiance one has to integrate the product of the BRDF and the incident radiance over all directions. Then, the spectral reflected radiance under spectral incident radiance L_i within solid angle $d\omega_i$ is

$$\int f(x, y, \lambda, \mathbf{n}, \mathbf{e}, \mathbf{v}) L_i(x, y, \lambda, \mathbf{n}, \mathbf{e}) \cos\theta_i d\omega_i \ ,$$

with θ_i the angle between $\mathbf{n}$ and $\mathbf{e}$. The factor $\cos\theta_i$ has to compensate for the fact that radiance is defined per unit foreshortened area, whereas irradiance is not. This model can deal with complicated lighting situations, including extended light sources, inter-reflections [4], etc. But handling this model is quite complicated. It therefore is interesting to consider a series of simplifications.

A first simplification is to consider a point light source and to discard inter-reflections. Basically, this allows us to drop the integral. Secondly, the above expression yields the radiance, whereas we are interested in the camera response and hence want to know the image irradiance. As we have seen in ch. 4, object radiance and image irradiance are related by the $\cos^4$ law. Rather than manipulate ever more impressive expressions, we simplify the expression for image irradiance as

$$I(x, y, \lambda) = R(x, y, \lambda, \mathbf{e}, \mathbf{n}, \mathbf{v}) E(\lambda, \mathbf{e}, \mathbf{n})$$

with I image irradiance, R a surface reflectivity function, and E the surface irradiance and where (x, y) represents a position in the image and the corresponding (deprojected)

[4] Inter-reflections correspond to incoming light reflected by other surfaces.

position on a surface in the scene (i.e. for ease of reference image coordinates are used as parameterisation of scene surface patches as well), λ is wavelength, $\mathbf{e}$ is the direction of the light falling onto the surface, $\mathbf{n}$ is the surface normal, $\mathbf{v}$ is the viewing direction.

Furthermore, we will assume that light of different wavelengths is reflected in the same proportions, independent of the direction of the normal, the direction of incidence, and the direction of viewing. In other words, the function R can be split into an independent geometric and spectral component:

$$R(x, y, \lambda, \mathbf{e}, \mathbf{n}, \mathbf{v}) = \rho(x, y, \lambda) F(\mathbf{e}, \mathbf{n}, \mathbf{v})$$

ρ is often called the surface *albedo*. It is a function of x and y as the surface properties need not be constant in our model. Of course, also F is dependent on x and y through the vectors $\mathbf{e}, \mathbf{n}.$ and $\mathbf{v}$. So, in this simplified case, a point on the surface keeps its colour, no matter from where it is lit or viewed, as long as the spectral composition of the light source is constant. The brightness can change the function F and different points on the surface may have different colours through ρ.

Intuitively, it is already clear that this assumption does not hold in quite a few practical cases. Imagine looking a a shiny car surface. Where you look at a specular reflection, the car body will look more or less white, i.e. the colour of the illumination. Looking at patches where your direction of view does not correspond with the light source's mirror direction, you will observe the car's true colour. One can stick to the separation of geometric and spectral influences by splitting the function R into a diffuse and a specular component, like

$$R(x, y, \lambda, \mathbf{e}, \mathbf{n}, \mathbf{v}) = \rho_d(x, y, \lambda) F_d(\mathbf{e}, \mathbf{n}, \mathbf{v}) + \rho_s(x, y, \lambda) F_s(\mathbf{e}, \mathbf{n}, \mathbf{v}) \ ,$$

assuming that the independence of these influences holds for the two components separately.

Diffuse reflection For the diffuse reflection component we only consider Lambertian reflection. In that case

$$F_d(\mathbf{e}, \mathbf{n}.\mathbf{v}) E(\lambda, \mathbf{e}, \mathbf{n}) = \bar{\mathbf{e}}(x, y).\mathbf{n}(x, y) E(\lambda)$$

where $\bar{\mathbf{e}}$ is a vector with the direction of the incoming light and a magnitude that specifies the power of that light, and $E(\lambda)$ is a function that specifies the spectral composition of the point light source. The Lambertian assumption implies that the viewing direction has no influence on camera irradiance.

The colour sensor response is specified with the 3-vector $\mathbf{S}$ of colour filter sensitivity functions for the red, geen, and blue bands (be it of polymer filters of a CCD or the cones of our eye). The 3-vector $\mathbf{I}$ of irradiances that finally generate a response are given by

$$\mathbf{I}(x, y) = \bar{\mathbf{e}}(x, y).\mathbf{n}(x, y) \int \rho_d(x, y, \lambda) E(\lambda) \mathbf{S}(\lambda) d\lambda \ .$$

This formulation can again be generalized for multiple sources with the same spectral composition E or for a source with a more complicated geometry in that different vectors

$\bar{\mathbf{e}}$ can be summed and the sum vector can be used just the same. On the other hand, in order to get workable expressions we make a final, further simplification by assuming that the colour filter sensitivities correspond to Dirac impulses at $\lambda_i, i = 1, 2, 3$. This yields

$$I_i(x, y) = \bar{\mathbf{e}}(x, y).\mathbf{n}(x, y)\rho(x, y, \lambda_i)E(\lambda_i)$$

for the three spectral channels $i = 1, 2, 3$ of the sensor. The actual response will typically be a non-linear function of these irradiances. The log function is used to model the response of the human eye; for cameras we have introduced the function I^γ in ch. 4.

Note that a change in colour of the illumination (by switching from $E_1(\lambda_1), E_2(\lambda_2), E_(\lambda_3)$ to $E'_1(\lambda_1), E'_2(\lambda_2), E'_(\lambda_3)$) changes each of the three irradiances by factors E'_i/E_i, that may differ for the channels but are the same at all points. On the other hand, if the spectral signature $E(\lambda)$ of the source is kept the same, but the source position is changed, then the irradiances for the three spectral bands will change by the same factor $\bar{\mathbf{e}}'(x, y).\mathbf{n}(x, y)/\bar{\mathbf{e}}(x, y).\mathbf{n}(x, y)$ which may differ from point to point, however.

Illumination independent colour features

For many computer vision applications, the explicit estimation of the illuminant's spectral characteristics and the extraction of actual spectral surface reflectance is more than what is needed. Rather than really disentangling the influences of surface reflectances and illuminant spectral composition on image irradiance, it often suffices to extract colour features that remain unchanged (or 'invariant') under changing illumination. Of course, such changes can still include changes in the spectral signature or the position of the light source(s).

In the sequel the irradiances for the three spectral bands will be denoted by I_R, I_G, I_B. A change in the spectral signature of the source(s) will change the irradiances by three pixel-independent factors, say α, β, γ: $(I'_R, I'_G, I'_B) = (\alpha I_R, \beta I_G, \gamma I_B)$. It follows that for the irradiances (I_{R1}, I_{G1}, I_{B1}) and (I_{R2}, I_{G2}, I_{B2}) for two different points in the scene, we find that $I'_{R1}/I'_{R2} = (\alpha I_{R1})/(\alpha I_{R2}) = I_{R1}/I_{R2}$ meaning that the ratio of corresponding irradiances for the $R - band$ at two different points is not affected by the change, i.e. is invariant. The same conclusion can be drawn for similar irradiance ratios for the other two spectral bands. Fortunately, the same holds for the camera responses if we discard the dark current, as $(\alpha I_{R1})^\gamma/(\alpha I_{R2})^\gamma = (I_{R1})^\gamma/(I_{R2})^\gamma$. This is not the case if the log response to irradiance has to be used, e.g. to model the response of the human eye or, equivalently, a camera for which the response was tuned to make it similar to that of the eye. In that case, a difference can be taken, as $\log I'_{R1} - \log I'_{R2} = \log(I'_{R1}/I'_{R2}) = \log(I_{R1}/I_{R2}) = \log I_{R1} - \log I_{R2}$ and the same holds for each of the bands. In such case one can also consider to take derivatives of the response over colour edges (more on this later).

In case of a geometric change of the illumination as when e.g. the light comes from a different direction, the three irradiances change with the same factor, say s, but this time the factor changes over the scene: $(I'_R, I'_G, I'_B) = (s(x, y)I_R, s(x, y)I_G, s(x, y)I_B)$. Again, ratios can be used as features that remain unchanged, but here these ratios are formed from different irradiances at the same point: $I'_R/I'_G = I_R/I_G$ and $I'_R/I'_B = I_R/I_B$. Also

here, the responses of a camera would have to be subtracted rather than divided if a log responsivity profile applies.

It might be important to extract features that survive both geometric and spectral changes in illumination. In that case, more complicated ratios can be formed:

$$\frac{I'_{R1}I'_{G2}}{I'_{R2}I'_{G1}} = \frac{\alpha\ s(x_1,y_1)\ \beta\ s(x_2,y_2)}{\alpha\ s(x_2,y_2)\ \beta\ s(x_1,y_1)}\frac{I_{R1}I_{G2}}{I_{R2}I_{G1}} = \frac{I_{R1}I_{G2}}{I_{R2}I_{G1}}\ .$$

Similar expressions can be formed by combining different pairs of spectral bands. Again, a γ response doesn't make a difference, i.e. the same ratio of the responses rather than the irradiances can be used as a feature that remains unchanged. With a log-response expressions like $\log I_{R1} + \log I_{G2} - \log I_{R2} - \log I_{G1}$ should be used.

Changes in the spectral composition and the direction of the illumination can also be dealt with in a simpler way if one focuses on information at the two sides of edges. Here it is assumed that the edge is due to a discontinuity in surface reflectance and not to a discontinuity in surface normal. In that case points on both side of the edge share approximately the same direction of the surface normal. Hence, the factors $s(x_1,y_1) \simeq s(x_2,y_2)$ and one can return to the simpler invariants of the type I_{R1}/I_{R2}. The reason to work along an edge is that otherwise $I_{R1} \simeq I_{R2}$ and the invariant risks to become trivial.

13.3 Surface texture

Texture is difficult to define. It represents aspects of the surface pattern, such as coarseness, directionality, regularity, etc. Fig. 13.8 shows a few textures. Already from this small set, some differences between textures can be observed. Textures (a) and (b) are quite regular. They contain a kind of basic pattern (sometimes called a 'texel') that is repeated in some way (sometimes referred to as the 'placement rule'). Comparing (a) and (b) we wee that (a)'s texel is larger than (b)'s. Texture (a) is said to be coarser. As a matter of fact, the situation is a bit more subtle. The texels in (a) show some subpatterns themselves (sometimes called 'micro-textures'). This texture shows a hierarchical organisation. The textile could be analysed at different scales. We also see differences in the dominant orientations between the four textures. Texture (a) has dominant diagonal orientations, (b) has dominant horizontal and vertical orientations, whereas (c) and (d) do not have orientations that are clearly more present than others. As will be seen, texture analysis techniques typically probe for the forementioned characteristics like regularity, coarseness, and directionality.

Texture features can be used to segment scenes into parts. In remote sensing it is e.g. quite usual to segment scenes into different types of land use or land cover, e.g. which areas are forrest, vineyards, urban, etc. Similarly, texture features can be necessary to inspect parts or to classify samples. Examples are the detection of flaws in fabrics for the textile industry or the classification of skin diseases for dermatologists. In this chapter, we will e.g. see an example of textile inspection.

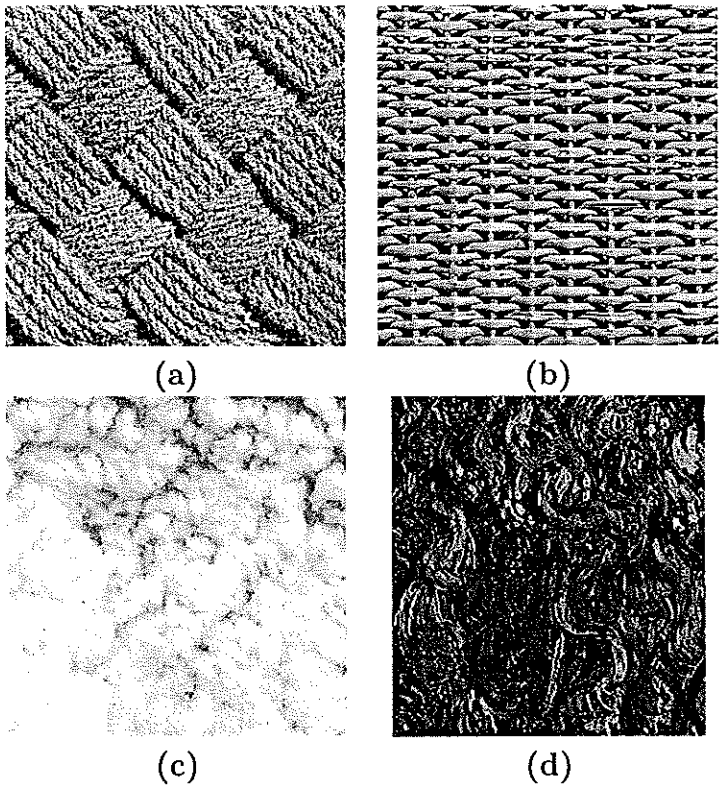


Figure 13.8: *Examples of texture.*

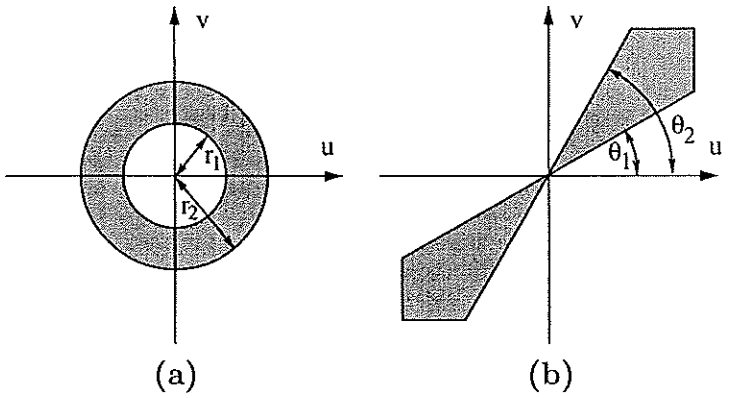


Figure 13.9: *Fourier texture features are based on integrating parts of the power spectrum.: (a) Integration over a ring for coarseness analysis, (b) integration over a sector for the analysis of directionality.*

13.3.1 Fourier features

Frequency related concepts like a texture being periodic or coarse-grained have led vision researchers to experiment with features that are extracted from the Fourier power spectrum of the image. If a texture is highly regular (periodic) it can be expected to have dominant peaks in its power spectrum. The detection of the regularity amounts to the detection of the peaks. One can calculate the entropy of the spectrum to quantify how strongly peaked the spectrum is.

Other Fourier features are used to assess coarseness and directionality. The most popular among the Fourier features are integrated power values, where integration is performed over two different kinds of regions in the Fourier domain, namely rings that that are concentric around the origin (zero frequency) and sectors with this point as their apex. These regions are illustrated in fig. 13.9. Such features correspond with

$$\int_A \int |F(u,v)|^2 du dv$$

with the area of integration A defined as

$$r_1^2 \leq u^2 + v^2 < r_2^2$$

for the rings and

$$\theta_1 \leq \arctan\left(\frac{u}{v}\right) < \theta_2$$

for the sectors.

Power integrated over the rings gives an indication of the coarseness of the texture. A texture is coarse if the power is concentrated in the inner rings. It is fine if its energy is mainly found in the outer rings. Power integrated over the sectors gives an idea of the directionality of the texture, i.e. whether the texture has some dominant orientations in it.

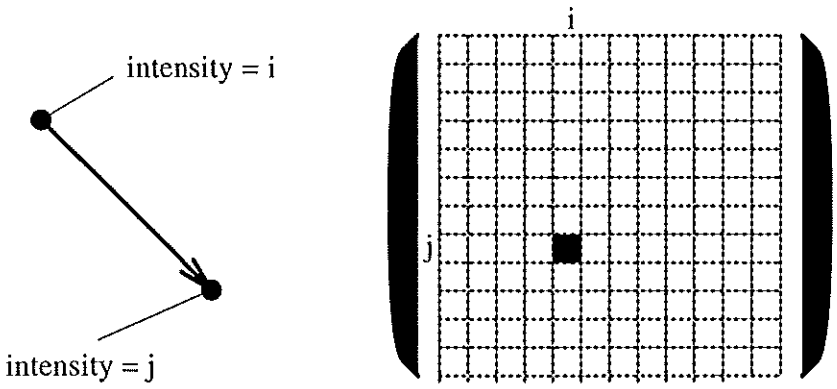


Figure 13.10: *Construction of a cooccurrence matrix: a vector is shifted over the image. For each position, the head and tail of the vector indicate a pixel. The intensity of the tail pixel is used as column coordinate in the cooccurrence matrix. Similarly, the head intensity is used as the row coordinate. At the position defined by this row and column the entry is incremented.*

A disadvantage of these features is that they collect global information about the time, instead of localised information. This is because the power spectrum typically is calculated from a complete image or at least a substantial part. This limits their value for texture-based segmentation. Later in this chapter, a Gabor filter approach is described, that reduces such problems. But even if an image is homogeneous in its texture, comparative tests have shown that Fourier features are less powerful for characterising textures than those described next.

13.3.2 Cooccurrence matrices

In ch. ?? the image histogram was introduced. A *histogram* captures intensity information in a highly compact, but also highly incomplete way. It simply is the discrete counterpart of the continuous probability density distribution $p(i)$ of image intensity i. All information about configurational aspects of the intensity distribution is lost.

A *cooccurrence matrix*, on the other hand, is a related data structure, which preserves information on some aspects of the spatial configurations. Suppose we define a two-dimensional vector, that we shift over the image. For each position of its tail, there is a corresponding pixel at its head. These two pixels (tail and head), define ordered pairs of image points. The cooccurrence matrix C is then simply formed as follows. Suppose we let the column coordinate of the matrix correspond to the intensity of the tail pixel and similarly use the intensity at the head pixel as the row of C. Suppose the intensities of tail and head pixels are i and j, respectively. Then we increment the element $C(i,j)$ each time we find a pixel pair with the corresponding intensities for the tail and head pixels. Dividing all matrix elements by the total number of such pixel pairs then yields a probability for the cooccurrence of the corresponding intensity values (i,j) at points with the specified relative position.

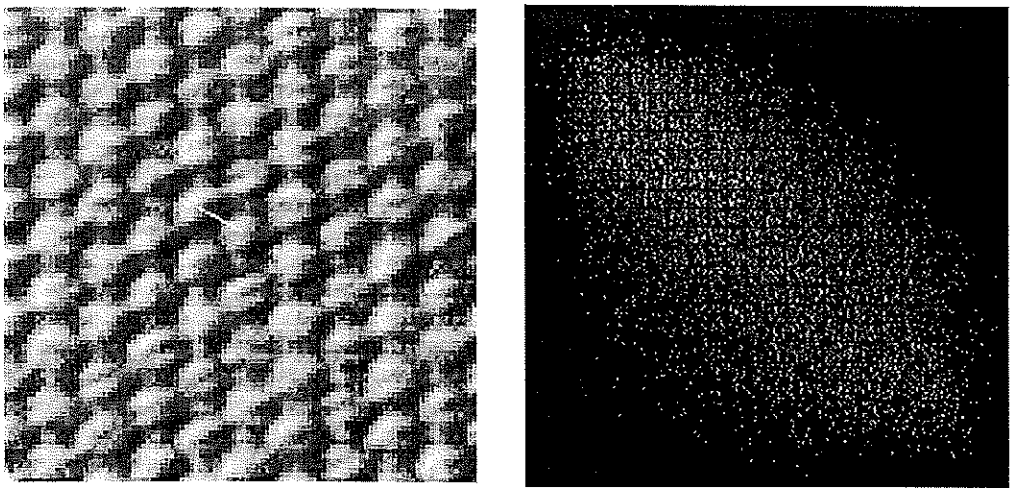


Figure 13.11: *(left) texture with vector superimposed, (right) cooccurrence matrix (256 × 256) with bright spots indicating high entries. The vector corresponds to a period of the texture. High values in the cooccurrence matrix are concentrated near its diagonal.*

A cooccurrence matrix is extracted for a specific vector. Obviously, one can build several such matrices, each for a different choice of vector. One can also use several vectors at once and put the results into a single matrix. If e.g. vectors with the same length but different orientations are combined, the resulting cooccurrence matrix is immune to changes in the orientation of the texture. Of course, the directionality of a texture often is an important feature in its own right that is lost in this way.

Fig. 13.11 and 13.12 show two examples for the same texture. In the texture images (left), the vector used for the construction of the cooccurrence matrices is superimposed. In the image of fig. 13.11 this vector links points at periodic positions. The corresponding cooccurrence matrix is therefore diagonal-like. Indeed, under such conditions head and tail positions will typically have similar intensities. On the other hand, if the vector is chosen such that is doesn't reflect the underlying structure, as in fig. 13.12, then votes will more equally be spread over the cooccurrence matrix, as can be observed there.

Concentration near the diagonal can also give an indication of the texture's coarseness. As long as the length of the vector stays within the grain size of the texture, high entries will typically concentrate near the diagonal. Thus, in order to estimate texture coarseness one can observe from which vector length onwards the entries start to spread out over the entire matrix.

Keeping the whole cooccurrence matrix is not efficient. Typically, a number of features are extracted from this matrix. Table 13.1 gives a number of the so-called 'Haralick features'. Energy and entropy are both measures of how concentrated the entries occur. Energy is high if the cooccurrence matrix has strong peaks. Entropy reaches its highest values if all entries are more or less equal. Maximum probability yields an indication of the strongest of the largest entries. Contrast and homogeneity are more specifically oriented

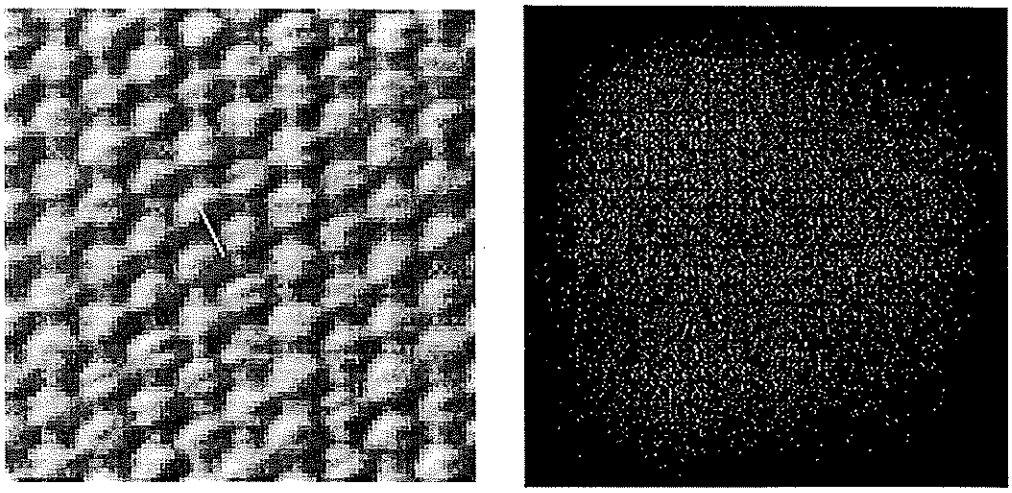


Figure 13.12: *(left) texture with vector superimposed, (right) cooccurrence matrix (256 × 256) with bright spots indicating high entries. The vector does not correspond to a period of the texture. Contributions are smeared out more or less equally over the cooccurrence matrix.*

feature	expression
energy	$\sum_i \sum_j C^2(i,j)$
entropy	$-\sum_i \sum_j C(i,j) \log C(i,j)$
contrast	$\sum_i \sum_j (i-j)^2 C(i,j)$
homogeneity	$\sum_i \sum_j C(i,j)/(1+\lvert i-j \rvert)$
max. probability	$\max_{i,j} C(i,j)$

Table 13.1: *Some of the 'Haralick features' that can be extracted from a cooccurrence matrix.*

feature	1D filter
L3	[1 2 1]
E3	[-1 0 1]
S3	[-1 2 -1]
L5	[1 4 6 4 1]
E5	[-1 -2 0 2 1]
S5	[-1 0 2 0 -1]
W5	[-1 2 0 -2 1]
R5	[1 -4 6 -4 1]
L7	[1 6 15 20 15 6 1]
E7	[-1 -4 -5 0 5 4 1]
S7	[-1 -2 1 4 1 -2 -1]
W7	[-1 0 3 0 -3 0 1]
R7	[1 -2 -1 4 -1 -2 1]
O7	[-1 6 -15 20 -15 6 -1]

Table 13.2: *Filters from which the Laws texture filters are built.*

towards the assessment of whether the entries are concentrated near the diagonal. Then contrast is small and homogeneity is large. These features can be combined into a feature vector, that is fed into a classifier algorithm to distinguish between textures.

13.3.3 Filter banks

A simple but quite effective method for extracting texture information is to convolve the image with a series of convolution filters. The outputs of such 'filter bank' are then combined into a feature vector, used for classification.

Laws filters

A first approach that belongs to this class is the use of a fixed set of filters, called the 'Laws filters'. These are a series of 3×3, 5×5, and 7×7 filters, created by the convolution of row and column filters with coefficients specified in Table 13.2. Usually, the output of these filters is non-linearly transformed and then averaged over larger regions in order to get an 'energy'. This can be the average of the squared outputs or the average of the absolute value. The energies for the different filters are then combined into a feature vector, that can be used for segmentation or classification.

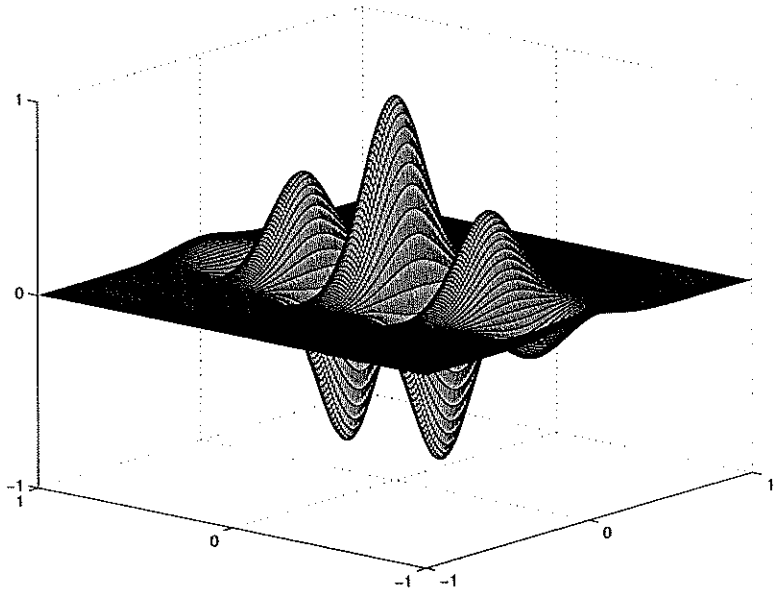


Figure 13.13: *A Gabor filter profile (in the spatial domain): a Gaussian is modulated with a cosine function.*

Gabor filters

The Laws filters were derived more or less *ad hoc*. Gabor filters offer a more principled alternative. These are filters that sample the local spatial frequency content of the image with optimal resolution. They are constructed by modulating a Gaussian with a sine function. For a vertically oriented Gabor filter this yields:

$$g(x,y) = e^{-\frac{x^2+y^2}{4\Delta_{x,y}^2}} \cos(2\pi u^* x + \varphi) \ .$$

The parameter $\Delta_{x,y}$ determines the width of the Gaussian envelope, u^* specifies the frequency of the modulating cosine, while φ specifies its phase. Often $\varphi = 0^o$ or $\varphi = 90^o$. Filters with other orientations can be obtained by applying a rigid rotation of the x, y-plane. Fig. 13.13 shows a typical profile of a Gabor filter.

That a Gabor filter samples the Fourier domain can be seen from its Fourier transform:

$$G(u,v) = \frac{1}{4\pi\Delta_{u,v}^2}\left(e^{-((u-u^*)^2+v^2)/(4\Delta_{u,v}^2)} + e^{-((u+u^*)^2+v^2)/(4\Delta_{u,v}^2)}\right)$$

where $\Delta_{u,v} = 1/(4\pi\Delta_{x,y})$. This real-valued transform consists of two Gaussian lobes in the frequency domain, positioned oppositely at $u = u^*, v = 0$ and $u = -u^*, v = 0$. The Gabor filter being a product of a Gaussian and a cosine, it's Fourier transform is known to be the convolution of a Gaussian and a pair of Dirac impulses. If we rotate the filter, the positions of these lobes will undergo the same rotation in the frequency domain (see chapter 5). An example of a Gabor filter's Fourier transform is given in fig. 13.14. The Gaussian lobes we

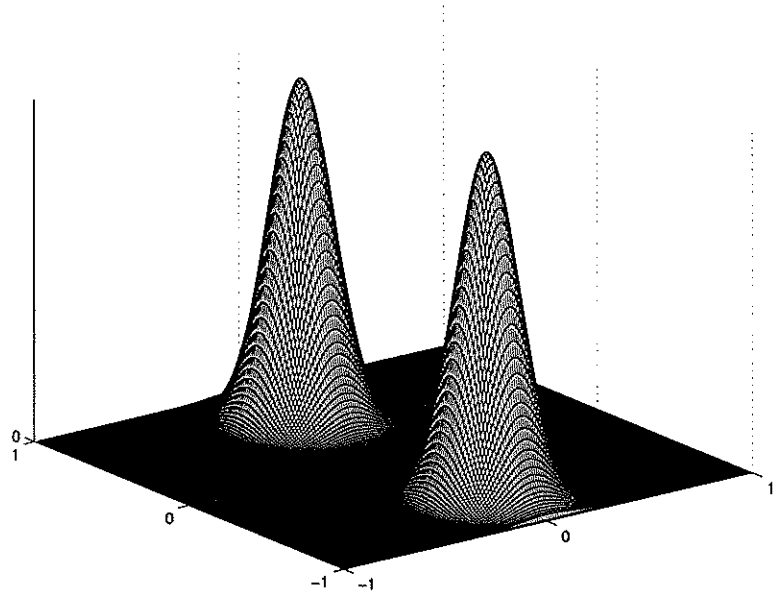


Figure 13.14: *Power spectrum of a Gabor filter.*

used here are circular in shape. They can be made elliptical. This obviously increases the number of parameters needed to specify the Gabor filter.

As features the outputs of a bank of Gabor filters are used, with different scales (Δ) and different orientations. These outputs are again non-linearly transformed and averaged over some neighbourhood. A possible coverage of the frequency space (u, v) with Gabor frequency profiles is shown in fig. 13.15. The example shows how the Gaussian profile can be made elliptical to optimally tessellate the frequency domain.

Fig. 13.16 shows the result for the image in (a) for two Gabor filters, one tuned to detect horizontally oriented structures (b) and one tuned to vertically oriented structures (c).

Note that $\Delta_{x,y}\Delta_{u,v} = 1/4\pi$. This is an important result, as it indicates that the product of the bandwith of the filter in the spatial domain and the spatial frequency domain reaches a minimum. Indeed, define

$$x_{av} = \frac{\int_{-\infty}^{\infty} x f \bar{f} dx}{\int_{-\infty}^{\infty} f \bar{f} dx} \qquad (\Delta x)^2 = \frac{\int_{-\infty}^{\infty} (x - x_{av})^2 f \bar{f} dx}{\int_{-\infty}^{\infty} f \bar{f} dx}$$

$$u_{av} = \frac{\int_{-\infty}^{\infty} u F \bar{F} du}{\int_{-\infty}^{\infty} F \bar{F} du} \qquad (\Delta u)^2 = \frac{\int_{-\infty}^{\infty} (u - u_{av})^2 F \bar{F} du}{\int_{-\infty}^{\infty} F \bar{F} du}$$

with $f = f(x, y)$ the original function and $F = F(u, v)$ its Fourier transform as usual. Bars indicate complex conjugate values. The *uncertainty principle* states that $\Delta x \Delta u \geq 1/4\pi$. A similar relation holds for y and v. Now, it can be shown that Gaussian filters minimize the

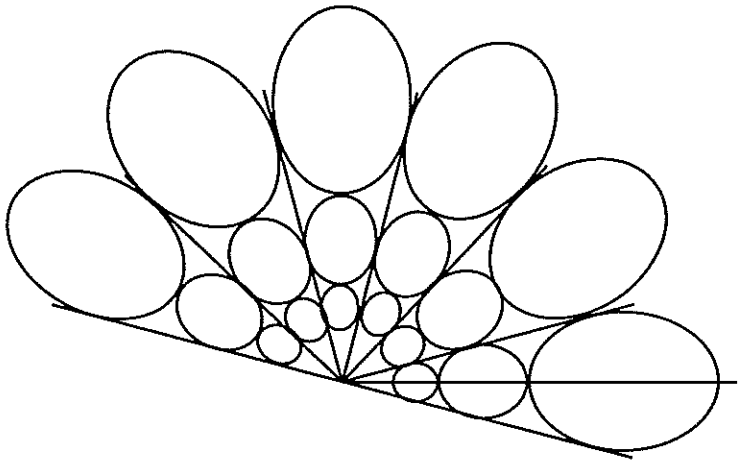
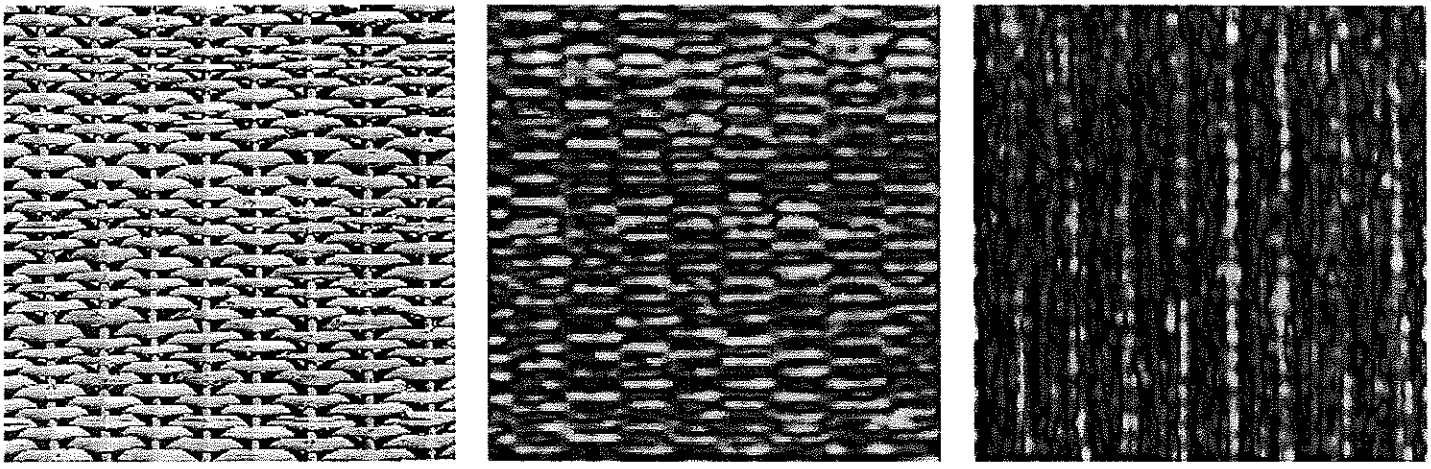


Figure 13.15: *By adapting the parameters of the Gabor filter, one can cover frequency space with a set of them.*



a b c

Figure 13.16: *(a) Input texture, (b) output of a Gabor filter tuned to detect horizontal structures, (c) same for vertical structures.*

uncertainty product, i.e. $\Delta x \Delta u = 1/4\pi$ (as do the more general class of Gabor filters, i.e. cosine modulated Gaussians). These filters therefore realise the best possible simultaneous localisation in the spatial ánd the frequency domain.

Eigenfilters

The Laws and Gabor filter banks have fixed coefficients. If a specific type of texture should be analysed, it can be interesting to finetune the filter set for that type. An example is inspection of a specific textile. One has strong expectations about the structure one will observe and a filter bank can be designed specifically for the task.

Ade proposed to consider small neighbourhoods of points and to extract the structure of such neighbourhoods via PCA (Principal Component Analysis). To fix ideas, suppose we take 3×3 neighbourhood mask. In such neighbourhood there are nine pixels of which the intensities are put in a 9-vector, going in reading order from top-left to bottom-right (we refer to this as 'lexicographical order'). When shifting the mask over the image many samples of this 9-vector are acquired and the corresponding covariance matrix can be calculated. The principal components yield a basis of mutually orthogonal 9-vectors, the eigenvectors. Arranging the 9 components of such eigenvector in the original 3×3 mask structure yields one convolution filter for every eigenvector (principal component), so 9 filters in total. The orthogonality of the filters ensures that they each yield complementary information. The filters coming with the highest eigenvalues reflect the structure of the texture on which the system was trained. In the case of our example this is a flawless piece of textile. Experiments will have to tell which filters perform best in detecting flaws. Filters with small eigenvalues stand a good chance as the variation of their input for correct textures is small and it will be easier to spot deviations from the typical value. Yet, the high eigenvalue filters are very well tuned toward the underlying structure of the material and cannot be discarded.

Before we discuss a real-case example, it is useful to generalise the eigenfilter method to also take the scale of the texture into account. Whereas Ade used filters of a fixed size, one can extract the periodicity of the pattern and adapt the size of the filters accordingly. In textile inspection the periodicity needed in the image will be of the order of 20×20. Using convolution masks of this size is simply not practical. A way out is to consider sparse masks, i.e. masks where most of the elements are set to zero.

Consider the piece of textile shown in fig. 15.9. This textile sample contains a flaw (it runs horizontally over the whole width). In order to detect such flaws, a set of eigenfilters are extracted from a flawless sample, shown in fig. 13.18. The first step is to extract the horizontal and vertical periodicities of the texture. Autocorrelation peaks were used to that end. The horizontal periodicity was found as 17 pixels, the vertical periodicity was found as 13 pixels. Hence, a sparse convolution mask is used with size 13×17. We first try with a simple mask, that only contains 9 non-zero coefficients. at positions as shown in fig. 13.19. The non-zero elements are at a distance of approximately one period or halfway. Hence, the sparse mask very crudely samples the structure of one texture period.

The next step is to calculate the covariance matrix for the underlying intensity data.

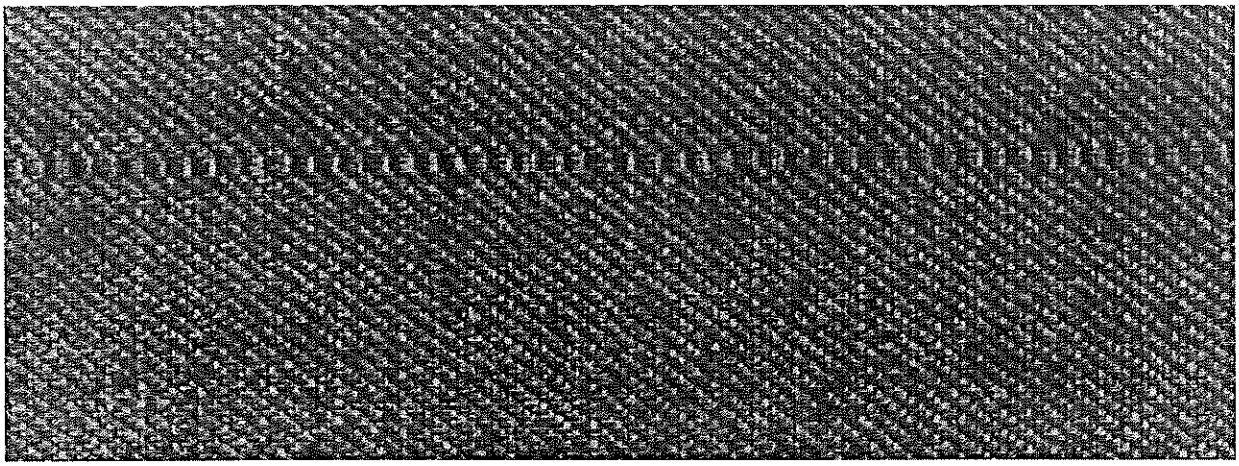


Figure 13.17: *Example of a flaw in a textile sample.*

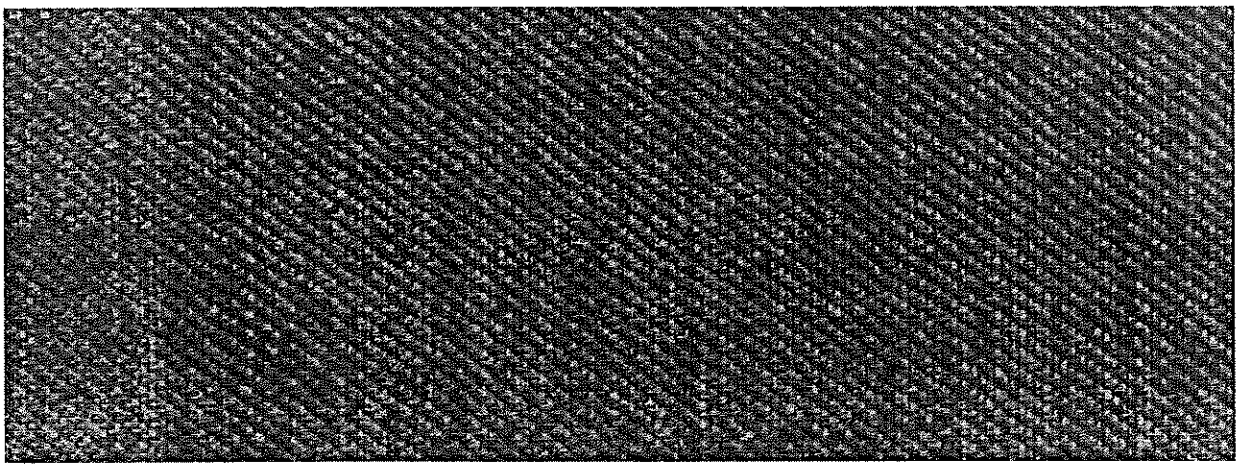


Figure 13.18: *Example of a flawless textile sample, used to extract the eigenfilters.*

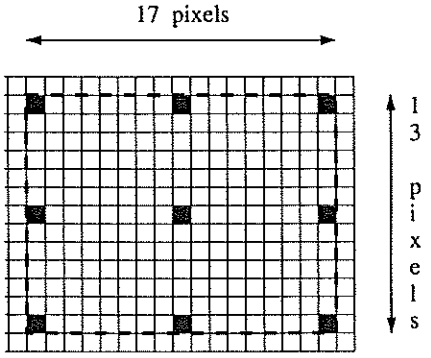


Figure 13.19: *Structure of the eigenfilter convolution masks. The black squares indicate positions in the* 17×13 *convolution mask that have non-zero coefficients.*

The result is

$$\begin{pmatrix} 174.1 & -77.6 & 101.6 & -60.8 & 72.7 & -71.5 & 116.5 & -77.9 & 91.4 \\ -77.6 & 173.9 & -78.2 & 71.5 & -61.7 & 73.1 & -76.4 & 116.4 & -78.4 \\ 101.6 & -78.2 & 173.5 & -70.4 & 71.7 & -62.1 & 95.3 & -77.0 & 116.3 \\ -60.8 & 71.5 & -70.4 & 173.9 & -76.5 & 101.0 & -59.4 & 71.8 & -70.1 \\ 72.7 & -61.7 & 71.7 & -76.5 & 173.7 & -77.1 & 70.6 & -60.3 & 72.1 \\ -71.5 & 73.1 & -62.1 & 101.0 & -77.1 & 173.4 & -69.3 & 70.9 & -60.7 \\ 116.5 & -76.4 & 95.3 & -59.4 & 70.6 & -69.3 & 173.4 & -75.3 & 99.8 \\ -77.9 & 116.4 & -77.0 & 71.8 & -60.3 & 70.9 & -75.3 & 173.2 & -75.9 \\ 91.4 & -78.4 & 116.3 & -70.1 & 72.1 & -60.7 & 99.8 & -75.9 & 172.8 \end{pmatrix}$$

From this matrix the eigenvectors are calculated. These are

$$\begin{pmatrix} 0.35 \\ -0.33 \\ 0.35 \\ -0.30 \\ 0.30 \\ -0.30 \\ 0.35 \\ -0.33 \\ 0.35 \end{pmatrix} \begin{pmatrix} 0.31 \\ 0.12 \\ 0.31 \\ 0.51 \\ -0.20 \\ 0.49 \\ 0.34 \\ 0.12 \\ 0.31 \end{pmatrix} \begin{pmatrix} 0.10 \\ 0.58 \\ 0.10 \\ -0.19 \\ 0.41 \\ -0.19 \\ 0.11 \\ 0.59 \\ 0.09 \end{pmatrix} \begin{pmatrix} 0.46 \\ -0.00 \\ -0.43 \\ 0.30 \\ 0.03 \\ -0.28 \\ 0.41 \\ -0.02 \\ -0.49 \end{pmatrix} \begin{pmatrix} -0.1 \\ -0.15 \\ -0.09 \\ 0.33 \\ 0.83 \\ 0.33 \\ -0.13 \\ -0.14 \\ -0.06 \end{pmatrix}$$

$$\begin{pmatrix} 0.27 \\ 0.11 \\ -0.11 \\ -0.62 \\ 0.00 \\ 0.63 \\ 0.11 \\ -0.12 \\ -0.27 \end{pmatrix} \begin{pmatrix} -0.43 \\ -0.06 \\ -0.54 \\ -0.10 \\ 0.00 \\ 0.09 \\ 0.55 \\ 0.05 \\ 0.4 \end{pmatrix} \begin{pmatrix} 0.08 \\ -0.69 \\ 0.02 \\ -0.09 \\ -0.00 \\ 0.10 \\ -0.02 \\ 0.69 \\ -0.09 \end{pmatrix} \begin{pmatrix} -0.5 \\ -0.00 \\ 0.50 \\ 0.00 \\ 0.02 \\ 0.00 \\ 0.47 \\ -0.01 \\ -0.50 \end{pmatrix}$$

Each of these vectors yields an eigenfilter, with its components as the coefficients. They are given in order of descending eigenvalues, which are from left to right and top to bottom: 805.26, 173.78, 144.82, 102.11, 98.96, 66.76, 64.36, 56.77, and 49.57.

This type of textile can now be inspected on-loom, by convolving newly woven parts by the above filters. We apply them to the example sample of fig. 15.9. Fig. 15.10 shows the results. The flaw shows up in the output of several of the filters.

The next step is to calculate the features' local 'energy'. This step consists of calculating the sum of absolute or squared values over neighbourhoods around each pixel. HIER JUISTE PARAMETERS VOOR HET TEXTUURVOORBEELD The results are shown in fig. 15.11. The goal of this step is to generate images with more homogeneous intensities in the hope of getting at least one or two where the resulting intensity for the

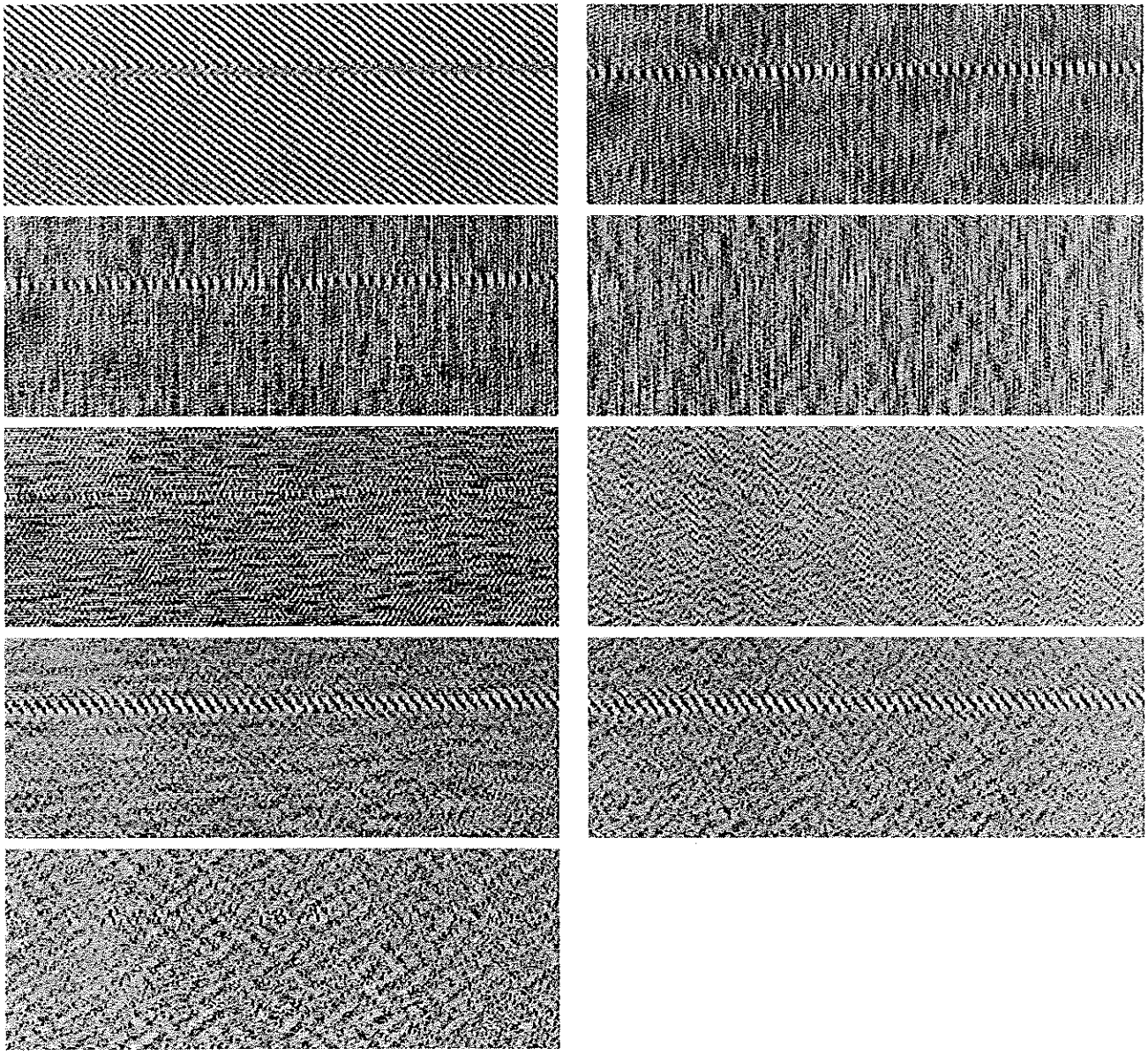


Figure 13.20: *Results of the 9 eigenfilters on the image of fig. 15.9 in lexicographic order.*

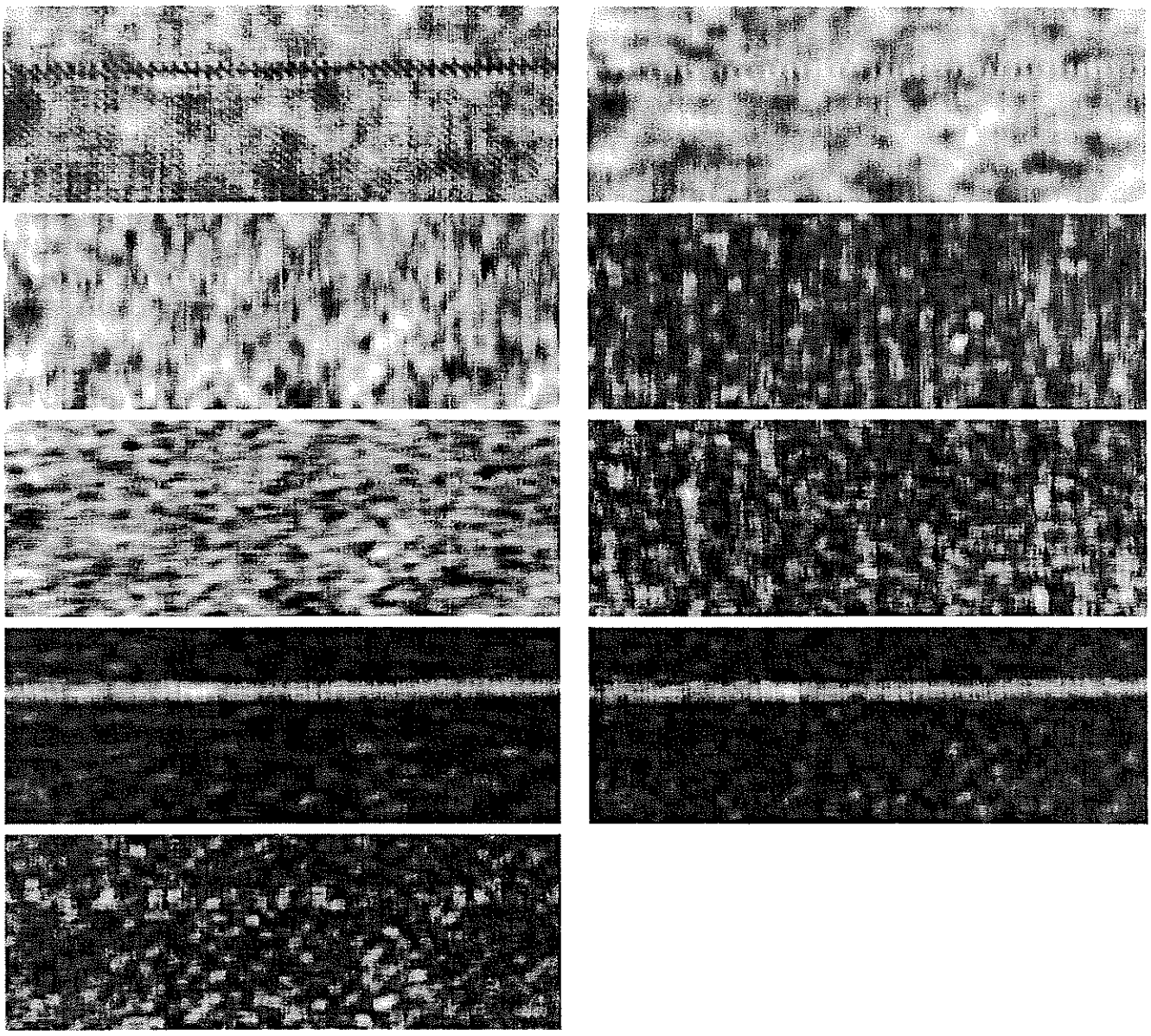


Figure 13.21: *Energies for the outputs of the 9 eigenfilters in fig. 15.10 in lexicographic order.*

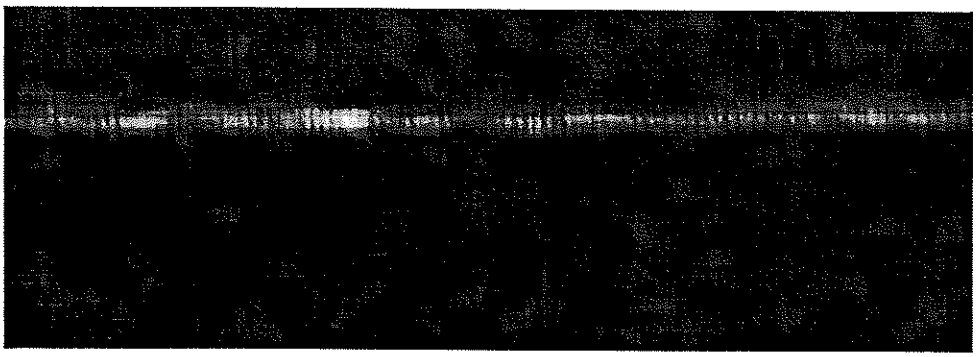


Figure 13.22: *Mahalanobis distance for the image in fig. 15.9.*

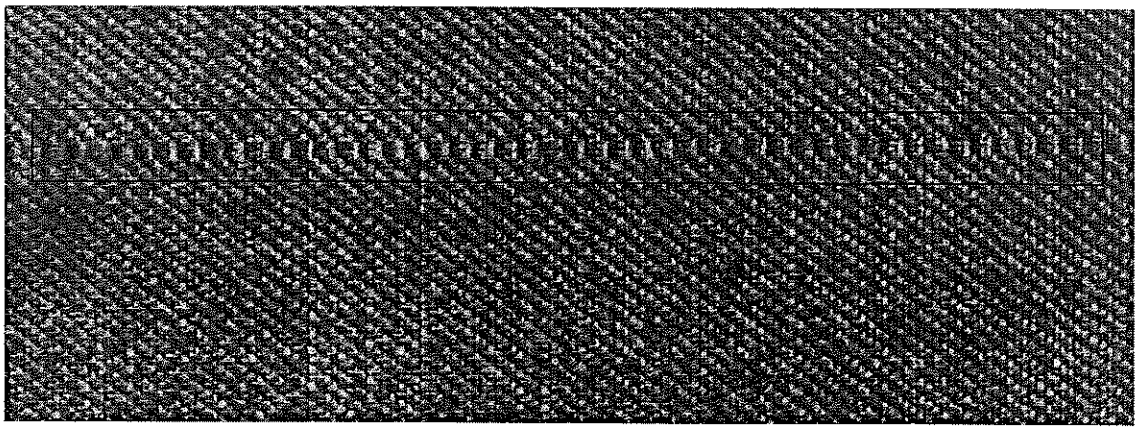


Figure 13.23: *Flaw detected on the basis of the Mahalanobis distance.*

flawed region is substantially different from the resulting intensity for good quality fabric. In the example this is e.g. the case for the energy of the first, seventh and eighth eigenfilter's output.

In a last step, the energy values are combined into a single distance to the typical energies. The Mahalanobis distance was used. This distance sums the differences with the average values for good fabric for the different eigenfilter energies, each time weighted by the inverse of the energy variance for flawless texture. The Mahalanobis distance for our example is shown in fig. 15.12. Based on this distance, the region in fig. 15.13 is classified as a flaw.