Coincise Computer Vision Notes

(Chapter-1)

Contents

- → Image in the Spatial Domain
- → Image in the Frequency Domain
  - → Colour & Colour Images

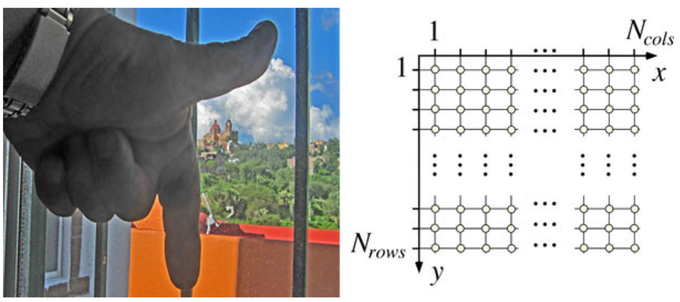
Image in Spatial Domain

A (digital) *image* is defined by *integrating* and *sampling* continuous (analog) data in a spatial domain. It consists of a rectangular array of *pixels* (x, y, u), each combining a location $(x, y) \in \mathbb{Z}^2$ and a value u, the *sample* at location (x, y). $\mathbb{Z}$ is the set of all integers. Points $(x, y) \in \mathbb{Z}^2$ form a *regular grid*. In a more formal way, an image I is defined on a rectangular set, the *carrier*

$$\Omega = \left\{(x, y) : 1 \le x \le N_{cols} \wedge 1 \le y \le N_{rows}\right\} \subset \mathbb{Z}^2 \qquad (1.1)$$

of I containing the *grid points* or *pixel locations* for $N_{cols} \ge 1$ and $N_{rows} \ge 1$.

Image Data is represented in a left hand coordinate system as the diagram below



Grid points for row y

$\{(1, y), (2, y) \ldots (3, y) \ldots\ldots\ldots (N_{col}, y)\}$

where $1 \le y \le N_{rows}$. and grid points for column x $\{(x, 1), (x, 2), \ldots\ldots, (x, N_{row})\}$ where $1 \le x \le N_{col}$

Insert 1.1 (Origin of the Term "Pixel") *The term* pixel *is short for* picture element. *It was introduced in the late* 1960*s by a group at the Jet Propulsion Laboratory in Pasadena, California, that was processing images taken by space vehicles. See* [R.B. Leighton, N.H. Horowitz, A.G. Herriman, A.T. Young, B.A. Smith, M.E. Davies, and C.B. Leovy, Mariner 6 television pictures: First report, Sci.

Pixel → Atomic element of an image.

In grid cell mode → Pixel are adjacent if they are different and their tiny shaded square share at least one point (i.e corner or edge)

Image Window

Image Windows A *window* $W_p^{m,n}(I)$ is a subimage of image I of size $m \times n$ positioned with respect to a *reference point* p (i.e., a pixel location). The default is that $m = n$ is an odd number, and p is the centre location in the window. Figure 1.3 shows the window $W_{(453,134)}^{73,77}(\texttt{SanMiguel})$.

Usually we can simplify the notation to W_p because the image and the size of the window are known by the given context.

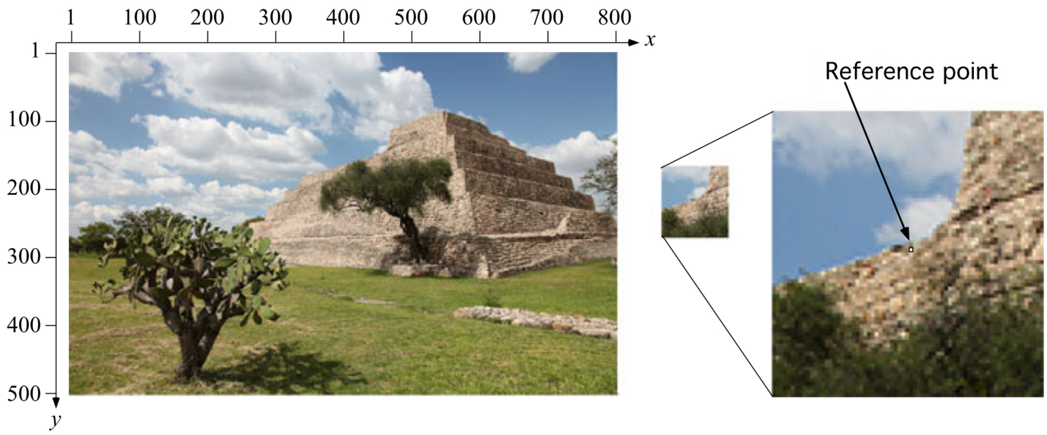


Fig. 1.3 A 73×77 window in the image `SanMiguel`. The marked reference pixel location is at $p = (453, 134)$ in the image that shows the main pyramid at Cañada de la Virgin, Mexico

Scalar Image / Gray Scale Image

Integer value b/w 0 to $2^a - 1$
where 0 = Black $2^a - 1$ = white
and all other gray levels are linearly interpolated b/w black and white

$a = 8$ (common value)

$a = 16$ (Recent Practice)

$G_{max} = 2^a - 1$

Vector-Value of RGB Images

Vector-Valued and RGB Images A *vector-valued image* has more than one *channel* or *band*, as it is the case for scalar images. Image values $(u_1, \ldots, u_{N_{channels}})$ are vectors of length $N_{channels}$. For example, colour images in the common RGB colour model have three channels, one for the red component, one for the green, and one for the blue component. The values u_i in each channel are in the set $\{0, 1, \ldots, G_{\max}\}$; each channel is just a grey-level image. See Fig. 1.4.

Mean

Mean Assume an $N_{cols} \times N_{rows}$ scalar image I. Following basic statistics, we define the *mean* (i.e., the "average grey level") of image I as

$$\mu_I = \frac{1}{N_{cols} \cdot N_{rows}} \sum_{x=1}^{N_{cols}} \sum_{y=1}^{N_{rows}} I(x, y)$$

$$= \frac{1}{|\Omega|} \sum_{(x,y) \in \Omega} I(x, y) \tag{1.2}$$

where $|\Omega| = N_{cols} \cdot N_{rows}$ is the cardinality of the carrier Ω of all pixel locations. We prefer the second way. We use I rather than u in this formula; I is a unique

Variance and Standard Deviation The *variance* of image I is defined as

$$\sigma_I^2 = \frac{1}{|\Omega|} \sum_{(x,y)\in\Omega} \left[I(x,y) - \mu_I\right]^2 \quad ($$

Its root σ_I is the *standard deviation* of image I.

Some well-known formulae from statistics can be applied, such as

$$\sigma_I^2 = \left[\frac{1}{|\Omega|} \sum_{(x,y)\in\Omega} I(x,y)^2\right] - \mu_I^2 \quad ($$

Proof for the formula

$$\sigma_I^2 = \left[\frac{1}{|\Omega|} \sum_{(x,y)\in\Omega} I(x,y)^2\right] - \mu_I^2$$

is given in next page

Proof

$$\mu_I = \frac{1}{N_{col} \cdot N_{rows}} \sum_{x=1}^{N_{cols}} \sum_{y=1}^{N_{rows}} I(x,y)$$

$$= \frac{1}{|\Omega|} \sum_{(x,y)\in\Omega} I(x,y)$$

$$\sigma_I^2 = \frac{1}{|\Omega|} \sum_{(x,y)\in\Omega} \left[I(x,y) - \mu_I \right]^2$$

$$= \frac{1}{|\Omega|} \sum_{(x,y)\in\Omega} \left[I(x,y)^2 + \mu_I^2 - 2 I(x,y) \mu_I \right]$$

$$= \frac{1}{|\Omega|} \sum_{(x,y)\in\Omega} I(x,y) + \frac{1}{|\Omega|} \sum_{(x,y)\in\Omega} \mu_I^2$$

$$- 2 \frac{1}{|\Omega|} \sum_{(x,y)\in\Omega} I(x,y)\,\mu$$

$$= \frac{1}{|\Omega|} \sum_{(x,y)\in\Omega} I(x,y) + \frac{1}{\cancel{|\Omega|}} \cancel{|\Omega|} \mu_I^2 - 2\mu_I^2$$

$$= \frac{1}{|\Omega|} \sum_{(x,y)\in\Omega} I(x,y)^2 - \mu_I^2 \Rightarrow eq\ 1.4$$

Histogram: Tabulated frequencies represented in the form of bar plots.

- Histogram of an image used to represent the frequencies of one band or channel of the image

Assume a scalar image I with pixels (i, j, u), where $0 \le u \le G_{\max}$. We define *absolute frequencies* by the count of appearances of a value u in the carrier Ω of all pixel locations, formally defined by

$$H_I(u) = \left| \left\{ (x, y) \in \Omega : I(x, y) = u \right\} \right| \tag{1.5}$$

Probability density function (PDF)

$$h_I(u) = \frac{H_I(u)}{|\Omega|}$$

$H_I(0)$, $H_I(1)$, $H_I(2)$ . . . . $H_I(G_{max})$

are the grey level histogram of scaled Image I

Mean and Variance from relative frequencies

Mean

$$\mu_I = \frac{1}{|\Omega|} \sum_{(x,y)\in\Omega} I(x,y)$$

$$= \frac{1}{|\Omega|} \sum_{u=0}^{G_{max}} u\, H(u)$$

$$= \sum_{u=0}^{G_{max}} u \frac{H(u)}{|\Omega|}$$

$$\boxed{= \sum_{u=0}^{G_{MAX}} u\, h_I(u)}$$

Imp

Variance

$$\sigma_I = \frac{1}{|\Omega|} \sum_{x,y \in \Omega} |I(x,y) - \mu_I|^2$$

$$= \frac{1}{|\Omega|} \sum_{u=0}^{u_{max}} |u - \mu_I|^2 |H(u)|$$

$$= \frac{1}{|\Omega|} \sum_{u=0}^{u_{max}} (u - \mu_I)^2 H(u)$$

$$\boxed{= \sum_{u=0}^{u_{max}} (u - \mu_I)^2 h_I(u)}$$

Absolute and relative *cumulative frequencies* are defined as follows, respectively:

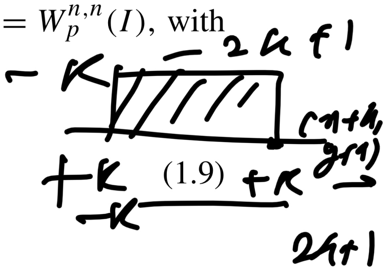
$$C_I(u) = \sum_{v=0}^{u} H_I(v) \quad \text{and} \quad c_I(u) = \sum_{v=0}^{u} h_I(v) \tag{1.8}$$

Those values are shown in *cumulative histograms*. Relative frequencies a rable to the *probability function* $\Pr[I(p) \leq u]$ of discrete random numbe

Window stats

Value Statistics in a Window Assume a (default) window $W = W_p^{n,n}(I)$, with $n = 2k + 1$ and $p = (x, y)$. Then we have in window coordinates

$$\boxed{\mu_W = \frac{1}{n^2} \sum_{i=-k}^{+k} \sum_{j=-k}^{+k} I(x+i, y+j)} \tag{1.9}$$



See Fig. 1.6. Formulas for the variance, and so forth, can be adapted analogously.

Contrast

Contrast $C(I)$ of an image is the absolute difference between the pixel value and the mean at the adjacent pixel

$$\boxed{C(I) = \frac{1}{|\Omega|} \sum_{(x,y)\in\Omega} |I(x,y) - \mu_A(x,y)|}$$

where $\mu_A(x,y)$ is the mean value of pixel location adjacent to pixel location (x,y)

Normalization of Two Functions Let μ_f and σ_f be the mean and standard deviation of a function f. Given are two real-valued functions f and g with the same discrete domain, say defined on arguments $1, 2, \ldots, T$, and non-zero variances. Let

$$\boxed{\alpha = \frac{\sigma_g}{\sigma_f}\mu_f - \mu_g \quad \text{and} \quad \beta = \frac{\sigma_f}{\sigma_g}} \tag{1.11}$$

$$\boxed{g_{new}(x) = \beta\big(g(x) + \alpha\big)} \quad \leftarrow S \tag{1.12}$$

As a result, the function g_{new} has the same mean and variance as the function f.

Distance Between Two Functions Now we define the distance between two real-valued functions defined on the same discrete domain, say $1, 2, \ldots, T$:

$$d_1(f,g) = \frac{1}{T}\sum_{x=1}^{T}\big|f(x) - g(x)\big| \tag{1.13}$$

$$d_2(f,g) = \frac{1}{T}\sqrt{\sum_{x=1}^{T}\big(f(x) - g(x)\big)^2} \tag{1.14}$$

Both distances are *metrics* thus satisfying the following axioms of a metric:
1. $f = g$ iff $d(f,g) = 0$,
2. $d(f,g) = d(g,f)$ (symmetry), and
3. $d(f,g) \le d(f,h) + d(h,g)$ for a third function h (triangular inequality).

Structural Similarity of Data Measures Assume two different spatial or temporal data measures $\mathscr{F}$ and $\mathscr{G}$ on the same domain $1, 2, \ldots, T$. We first map $\mathscr{G}$ into $\mathscr{G}_{new}$ such that both measures have now identical mean and variance and then calculate the distance between $\mathscr{F}$ and $\mathscr{G}_{new}$ using either the L_1- or L_2-metric.

Two measures $\mathscr{F}$ and $\mathscr{G}$ are *structurally similar* iff the resulting distance between $\mathscr{F}$ and $\mathscr{G}_{new}$ is close to zero. Structurally similar measures take their local maxima or minima at about the same arguments.

Edges

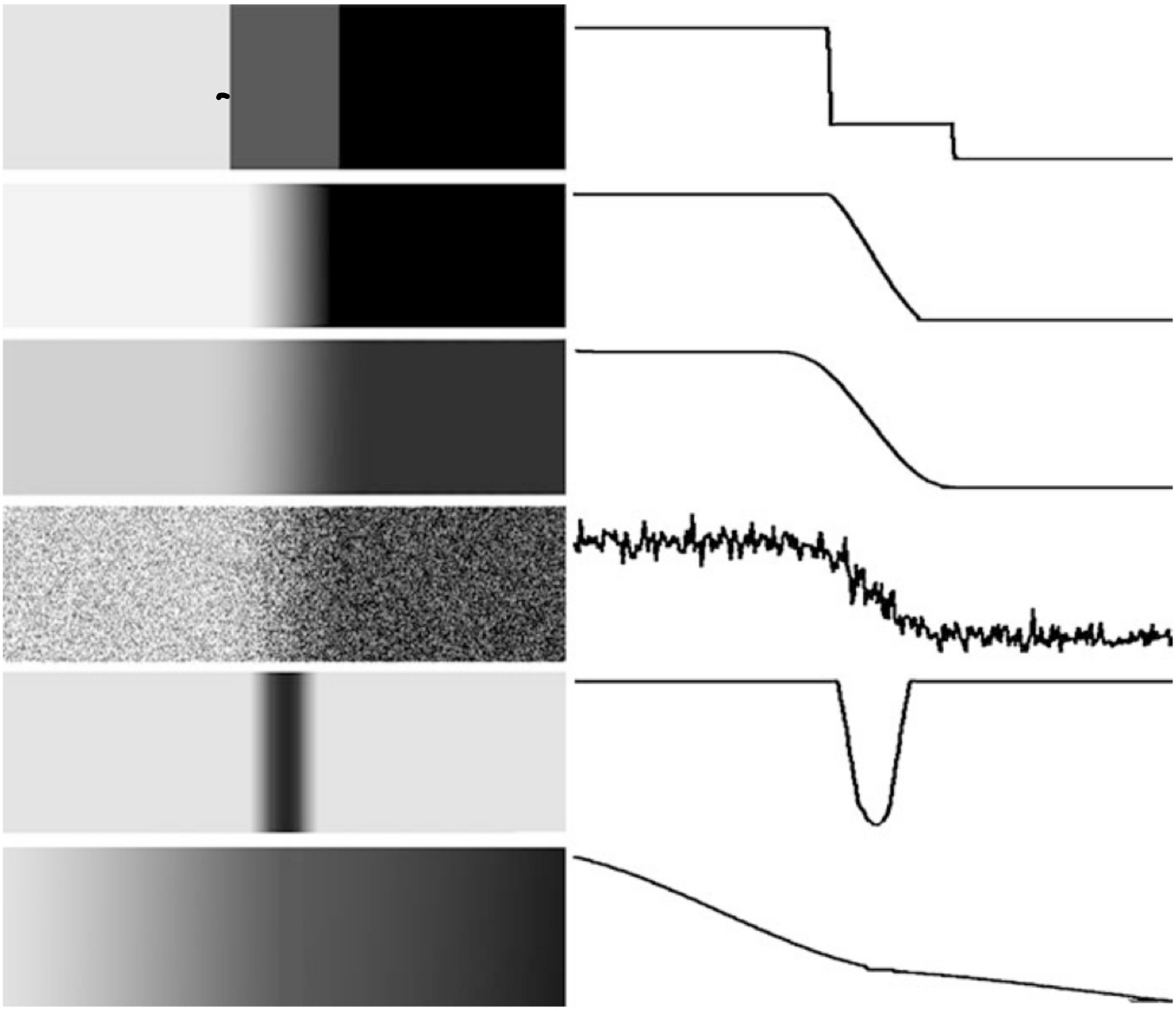
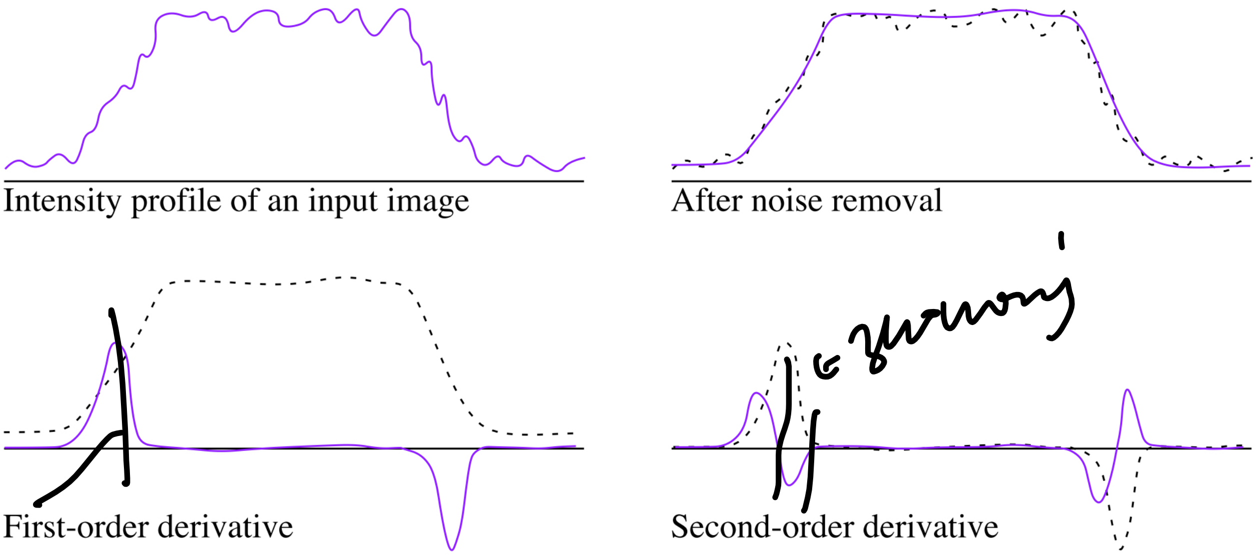


Fig. 1.11 Illustration for the step-edge model. *Left*: Synthetic input images. *Right*: Intensity profiles for the corresponding images on the left. *Top to bottom*: Ideal step-edges, linear edge, smooth edge, noisy edge, thin line, and a discontinuity in shaded region

According to step edge Model Edges are change in local derivatives



edges

zero crossing

Fig. 1.12 Illustration of an input signal, signal after noise removal, first derivative, and second derivative

Image as a continuous Surface

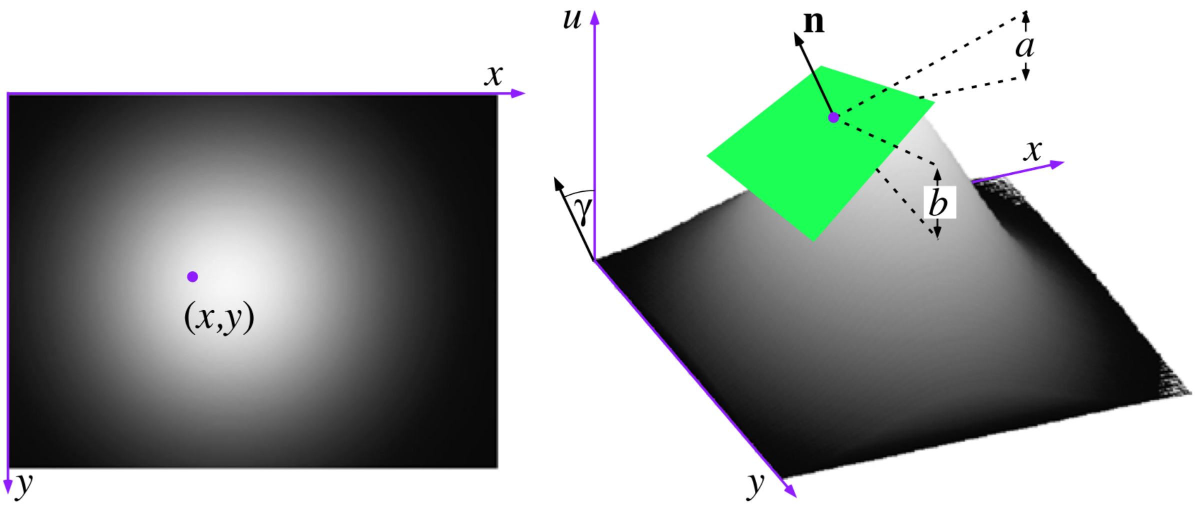


Fig. 1.13 *Left*: Synthetic input image with pixel location (x, y). *Right*: Illustration of tangential plane (in *green*) at pixel $(x, y, I(x, y))$, normal $\mathbf{n} = [a, b, 1]^\top$, which is orthogonal to this plane, and partial derivatives a (in x-direction) and b (in y-direction) in the left-hand Cartesian coordinate system defined by image coordinates x and y and the image-value axis u

First order Derivative

First-Order Derivatives The *normal* $\mathbf{n}$ is orthogonal to the *tangential plane* at a pixel $(x, y, I(x, y))$; the tangential plane follows the surface defined by image values $I(x, y)$ on the xy-plane. The normal has an angle γ with the image-value axis.

The *gradient*

$$\nabla I = \mathbf{grad}\, I = \left[\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}\right]^\top \tag{1.15}$$

combines both partial derivatives at a given point $p = (x, y)$. Read ∇I as "nabla I". To be precise, we should write $[\mathbf{grad}\, f](p)$ and so forth, but we leave pixel location p out for easier reading of the formulae.

The *normal*

$$\mathbf{n} = \left[\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}, +1\right]^\top \tag{1.16}$$

can point either into the positive or negative direction of the u-axis; we decide here for the positive direction and thus $+1$ in the formal definition. The *slope angle*

$$\gamma = \arccos \frac{1}{\|\mathbf{n}\|_2} \tag{1.17}$$

is defined between the u-axis and normal $\mathbf{n}$. The first-order derivatives allow us to calculate the *length* (or *magnitude*) of gradient and normal:

$$\|\mathbf{grad}\, I\|_2 = \sqrt{\left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2} \quad \text{and} \quad \|\mathbf{n}\|_2 = \sqrt{\left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2 + 1} \tag{1.18}$$

Following Fig. 1.12 and the related discussion, we conclude that:

<u>note</u>

Edges are located at the location where $\|grad\, I\|_2$ is or $\|n\|_2$

is local maximum.

Second-Order Derivative

Second-Order Derivatives Second-order derivatives are combined into either the *Laplacian* of I, given by

$$\triangle I = \nabla^2 I = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} \tag{1.19}$$

or the *quadratic variation* of I, given by[2]

$$\left(\frac{\partial^2 I}{\partial x^2}\right)^2 + 2\left(\frac{\partial^2 I}{\partial x \partial y}\right)^2 + \left(\frac{\partial^2 I}{\partial y^2}\right)^2 \tag{1.20}$$

---

[2]To be precise, a function I satisfies the second-order differentiability condition iff $(\frac{\partial^2 I}{\partial x \partial y}) = (\frac{\partial^2 I}{\partial y \partial x})$. We simply assumed in (1.20) that I satisfies this condition.

→ Laplacian is scalar not vector like & normal.

→ Edges are located at the location where laplacian ΔI or quadratic variation defines zero crossing.

Frequency Domain Analysis

The 2D Fourier transform maps the image from spatial domain to the frequency domain.

> **Insert 1.2** (Fourier and Integral Transforms) *J.B.J. Fourier* (1768–1830) *was a French mathematician. He analysed series and integrals of functions that are today known by his name.*
>
> *The Fourier transform is a prominent example of an* integral transform. *It is related to the computationally simpler* cosine transform, *which is used in the baseline JPEG image encoding algorithm.*

Fourier Filtering :- Change of data into frequency domain is called fourier filtering operation and then inverting it back fourier transformed image to modified image in spatial domain using 2D Inverse Discrete transform. The whole process is called fourier filtering. It helps in contrast enhancement, noise removal and smoothening of image

→ In context of FT, we assume that image coordinates are from $0 \le x \le N_{row} - 1$ and $0 \le y \le N_{col} - 1$ otherwise use $x-1$ and $y-1$ in the formula

2D Fourier Transform Formally, the 2D DFT is defined as follows:

$$\mathbf{I}(u, v) = \frac{1}{N_{cols} \cdot N_{rows}} \sum_{x=0}^{N_{cols}-1} \sum_{y=0}^{N_{rows}-1} I(x, y) \cdot \exp\left[-i2\pi \left(\frac{xu}{N_{cols}} + \frac{yv}{N_{rows}}\right)\right] \tag{1.21}$$

for *frequencies* $u = 0, 1, \ldots, N_{cols} - 1$ and $v = 0, 1, \ldots, N_{rows} - 1$. The letter $i = \sqrt{-1}$ denotes (here in the context of Fourier transforms only) the *imaginary unit* of complex numbers.[3] For any real α, the *Eulerian formula*

$$\exp(i\alpha) = e^{i\alpha} = \cos\alpha + i \cdot \sin\alpha \tag{1.22}$$

demonstrates that the Fourier transform is actually a weighted sum of sine and cosine functions, but in the complex plane. If α is outside the interval $[0, 2\pi)$, then it is taken modulo 2π in this formula. The Eulerian number $e = 2.71828\ldots = \exp(1)$.

[3] Physicists or electric engineers use j rather than i, in order to distinguish from the intensity i in electricity.

1.2.2 Inverse Discrete Fourier Transform

The *inverse 2D DFT* transforms a Fourier transform **I** back into the spatial domain:

$$I(x, y) = \sum_{u=0}^{N_{cols}-1} \sum_{v=0}^{N_{rows}-1} \mathbf{I}(u, v) \exp\left[i2\pi \left(\frac{xu}{N_{cols}} + \frac{yv}{N_{rows}} \right) \right] \tag{1.23}$$

Note that the powers of the root of unity are here reversed compared to (1.21) (i.e., the minus sign has been replaced by a plus sign).

Variants of Transform Equations Definitions of DFT and inverse DFT may vary. We can have the plus sign in the DFT and the minus sign in the inverse DFT.

We have the scaling factor $1/N_{cols} \cdot N_{rows}$ in the 2D DFT and the scaling factor 1 in the inverse transform. Important is that the product of both scaling factors in the DFT and in the inverse DFT equals $1/N_{cols} \cdot N_{rows}$. We could have split $1/N_{cols} \cdot N_{rows}$ into two scaling factors, say, for example, $1/\sqrt{N_{cols} \cdot N_{rows}}$ in both transforms.

Complex Plane Properties

$z_1 = a_1 + i b_1 \quad z_2 = a_2 + i b_2$

Property 1

$z_1 + z_2 = (a_1 + a_2) + i(b_1 + b_2)$

Property 2

$z_1 \cdot z_2 = (a_1 a_2 - b_1 b_2) + i(a_1 b_2 + a_2 b_1)$

__Property 3__

Norm $|z| = \sqrt{a^2+b^2}$

Property 4

The conjugate of of complex number $z = a+ib$ is

$z^* = a - ib$

$(z^*)^* = z$

$(z_1 \cdot z_2)^* = z_1^* z_2^*$

$z^{-1} = \|z\|_2^{-2} \cdot z^*$

if $z \neq 0$

Complex Numbers in Polar Coordinates A complex number z can also be written in the form $z = r \cdot e^{i\alpha}$, with $r = \|z\|_2$ and α (the *complex argument of* z) is uniquely defined modulo 2π if $z \neq 0$. This maps complex numbers into polar coordinates (r, α).

A rotation of a vector $[c, d]^\top$ [i.e., starting at the origin $[0, 0]^\top$] about an angle α is the vector $[a, b]^\top$, with

$$a + i \cdot b = e^{i\alpha} \cdot (c + i \cdot d) \tag{1.26}$$

Roots of Unity The complex number $W_M = \exp[i2\pi/M]$ defines the Mth *root of unity*; we have $W_M^M = W_M^{2M} = W_M^{3M} = \cdots = 1$. Assume that M is a multiple of 4. Then we have that $W_M^0 = 1 + i \cdot 0$, $W_M^{M/4} = 0 + i \cdot 1$, $W_M^{M/2} = -1 + i \cdot 0$, and $W_M^{3M/4} = 0 + i \cdot (-1)$.

Insert 1.4 (Fast Fourier Transform) *The properties of* M*th roots of unity,* M *a power of* 2, *supported the design of the original* Fast Fourier Transform (*FFT*), *a time-efficient implementation of the DFT.*

Fourier Transform in terms of Mth roots of unity

$$\mathbf{I}(u, v) = \frac{1}{N_{cols} \cdot N_{rows}} \sum_{x=0}^{N_{cols}-1} \sum_{y=0}^{N_{rows}-1} I(x, y) \cdot W_{N_{cols}}^{-xu} \cdot W_{N_{rows}}^{-yv} \tag{1.27}$$

For any root of unity $W_n = i2\pi/n$, $n \geq 1$, and for any power $m \in \mathbb{Z}$, it follows that

$$\|W_n^m\|_2 = \|e^{i2\pi m/n}\|_2 = \sqrt{\cos(2\pi m/n)^2 + \sin(2\pi m/n)^2} = 1 \tag{1.28}$$

Thus, all those powers are located on the unit circle, as illustrated in Fig. 1.16.

→ In FT low frequency of u and v represent long wavelength of sine and cosine components thus homogeneous additive contribution to the image

→ Large frequencies represent short wavelength thus local discontinuity in I such as edges or intensity outliers

Three Properties of the DFT We consider the 2D Fourier transform of an image I. It consists of two $N_{cols} \times N_{rows}$ arrays representing the real (i.e., the as) and the imaginary part (i.e., the bs) of the obtained complex numbers $a + i \cdot b$. Thus, the $N_{cols} \times N_{rows}$ real data of the input image I are now "doubled". But there is an important *symmetry property*:

$$\mathbf{I}(N_{cols} - u, N_{rows} - v) = \mathbf{I}(-u, -v) = \mathbf{I}(u, v)^* \tag{1.29}$$

(recall: the number on the right is the conjugate complex number). Thus, actually half of the data in both arrays of $\mathbf{I}$ can be directly obtained from the other half. Another property is that

$$\mathbf{I}(0, 0) = \frac{1}{N_{cols} \cdot N_{rows}} \sum_{x=0}^{N_{cols}-1} \sum_{y=0}^{N_{rows}-1} I(x, y) \tag{1.30}$$

which is the mean of I. Because I has only real values, it follows that the imaginary part of $\mathbf{I}(0, 0)$ is always equal to zero. Originating from applications of the Fourier transform in Electrical Engineering, the mean $\mathbf{I}(0, 0)$ of the signal is known as the *DC component* of I, meaning *direct current*. For any other frequency $(u, v) \neq (0, 0)$, $\mathbf{I}(u, v)$ is called an *AC component* of I, meaning *alternating current*.

As a third property, we mention *Parseval's theorem*

$$\frac{1}{|\Omega|} \sum_{\Omega} \left| I(x, y) \right|^2 = \sum_{\Omega} \left| \mathbf{I}(u, v) \right|^2 \tag{1.31}$$

Fourier Relationship

(1) fun (I) $\Longleftrightarrow$ FT in of fun I

$I(x, y) \Longleftrightarrow I(u, v)$

② $\boxed{I * G(x,y) \Leftrightarrow I \circ G(u,v)}$

I convolution of $G(x,y)$ is equal to point by product of I with G in frequency domain.

③ $$\boxed{I(x,y)\,(-1)^{x+y} \Leftrightarrow I\left(u+\frac{N_{col}}{2},\, v+\frac{N_{row}}{2}\right)}$$

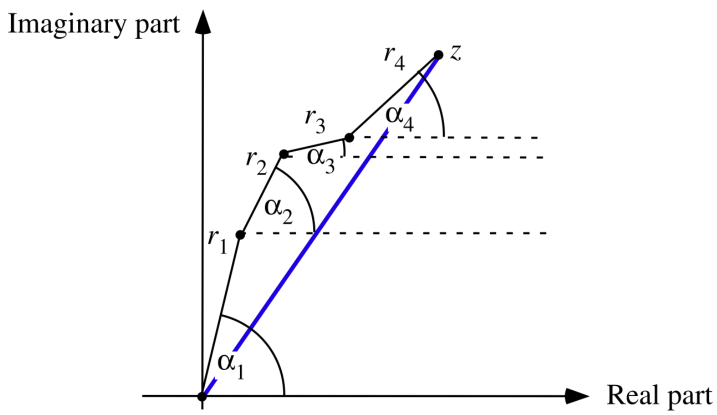
→ Fourier transform is shifted to centred position if input image is multiplied with chessboard pattern of +1 and -1

$$\boxed{\text{(4)}\quad aI(x,y) + bJ(x,y) \Longleftrightarrow aI(u,v) + bJ(u,v)}$$

- Property of linear transformation

Phase Congruency Model

Fig. 1.23 Addition of four complex numbers represented by polar coordinates (r_h, α_h) in the complex plane



Consider a *local Fourier transform*, centred at a pixel location $p = (x, y)$ in image I, using a $(2k+1) \times (2k+1)$ filter kernel of Fourier basis functions:

$$\mathbf{J}(u, v) = \frac{1}{(2k+1)^2} \sum_{i=-k}^{k} \sum_{j=-k}^{k} I(x+i, y+j) \cdot W_{2k+1}^{-iu} \cdot W_{2k+1}^{-jv} \tag{1.32}$$

Ignoring the DC component that has the phase zero (and is of no importance for locating an edge), the resulting Fourier transform $\mathbf{J}$ is composed of $n = (2k+1)^2 - 1$ complex numbers z_h, each defined by the amplitude $r_h = \|z_h\|_2$ and the phase α_h, for $1 \leq h \leq n$.

Figure 1.23 illustrates an addition of four complex numbers represented by the amplitudes and phases, resulting in a complex number z. The four complex numbers (r_h, α_h) are roughly *in phase*, meaning that the phase angles α_h do not differ much (i.e. have a small variance only). Such an approximate identity defines a high *phase congruency*, formally defined by the property measure

$$\mathscr{P}_{ideal\_phase}(p) = \frac{\|z\|_2}{\sum_{h=1}^{n} r_h} \tag{1.33}$$

with z being the sum of all n complex vectors represented by (r_h, α_h). We have that $\mathscr{P}_{ideal\_phase}(p) = 1$ defines perfect congruency, and $\mathscr{P}_{ideal\_phase}(p) = 0$ occurs for perfectly opposing phase angles and amplitudes.

Observation 1.3 *Local phase congruency identifies features in images. Under the phase congruency model, step-edges represent just one narrow class of an infinite range of feature types that can occur. Phase congruency marks lines, corners, "roof edges", and a continuous range of hybrid feature types between lines and steps.*

Colours

RGB Space

The RGB Space Assume that $0 \leq R, G, B \leq G_{\max}$ and consider a multi-channel image I with pixel values $\mathbf{u} = (R, G, B)$. If $G_{\max} = 255$, then we have 16,777,216 different colours, such as $\mathbf{u} = (255, 0, 0)$ for Red, $\mathbf{u} = (255, 255, 0)$ for Yellow, and so forth. The set of all possible RGB values defines the *RGB cube*, a common representation of the RGB colour space. See Fig. 1.30.

The diagonal in this cube, from White at (255, 255, 255) to Black at (0, 0, 0), is the location of all grey levels (u, u, u), which are not colours. In general, a point $\mathbf{q} = (R, G, B)$ in this RGB cube defines either a colour or a grey level, where the

mean

$$M = \frac{R + G + B}{3} \tag{1.43}$$

defines the *intensity* of colour or grey level $\mathbf{q}$.

HSI space

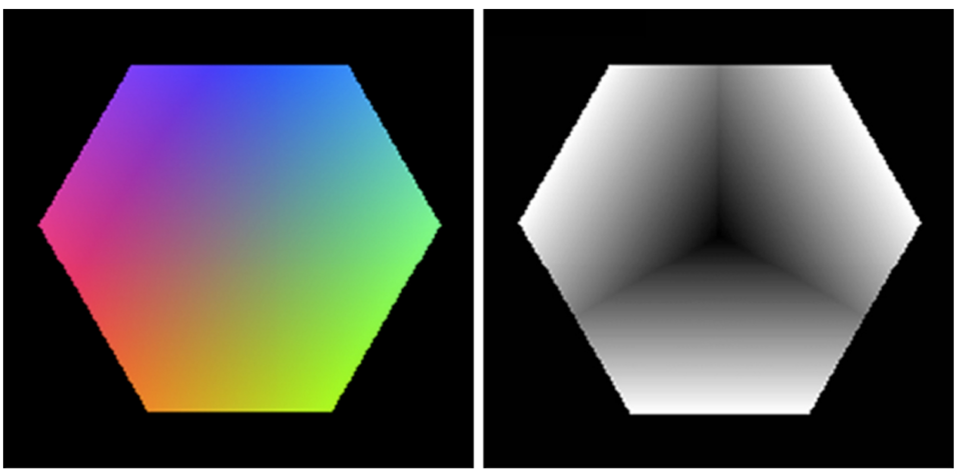


Fig. 1.33 Cuts through the RGB cube at $u = 131$ showing the RGB image I_{131} and saturation values for the same cutting plane

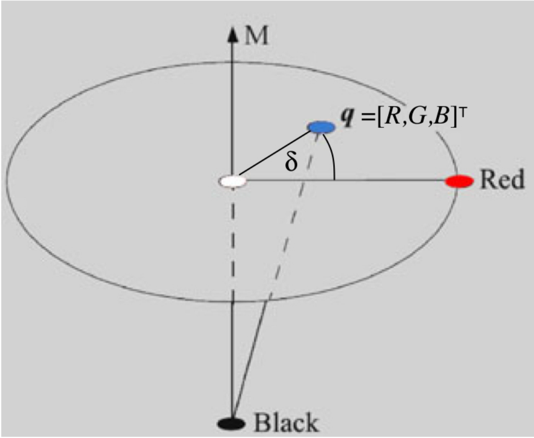


Fig. 1.31 The intensity axis points along the grey-level diagonal in the RGB cube. For the cut with the cube, we identify one colour (here, *Red*) as the reference colour. Now we can describe **q** by the *intensity* (i.e. its mean value), *hue*, which is the angle with respect to the reference colour (*Red* here, and, say, in counter-clockwise order), and *saturation* corresponding to the distance to the intensity axis

The HSI Space Consider a plane that cuts the RGB cube orthogonally to its grey-level diagonal, with $\mathbf{q} = (R, G, B)$ incident with this plane but not on the diagonal (see also Fig. 1.33). In an abstract sense, we represent the resulting cut by a disk, ignoring the fact that cuts of such a plane with the cube are actually simple polygons. See Fig. 1.31.

For the disk, we fix one colour for reference; this is Red in Fig. 1.31. The location of $\mathbf{q}$ in the disk is uniquely defined by an angle H (the *hue*) and a scaled distance S (the *saturation*) from the intersecting grey-level diagonal of the RGB cube. Formally, we have

$$H = \begin{cases} \delta & \text{if } B \leq G \\ 2\pi - \delta & \text{if } B > G \end{cases} \quad \text{with} \tag{1.44}$$

$$\delta = \arccos \frac{(R-G) + (R-B)}{2\sqrt{(R-G)^2 + (R-B)(G-B)}} \quad \text{in } [0, \pi) \tag{1.45}$$

$$S = 1 - 3 \cdot \frac{\min\{R, G, B\}}{R + G + B} \tag{1.46}$$

Altogether, this defines the *HSI colour model*. We represent intensity by M, to avoid confusion with the use of I for images.