Chapter 1

Stochastic Analysis of Image Acquisition and Scale-space Smoothing

K. Åström and A. Heyden
Dept of Mathematics
Lund University, P O Box 118, S-221 00 Lund
Sweden

1.1 Introduction

Low level image processing is often used to detect and localise features such as edges and corners. It is also used to correlate or match small parts of one image with parts in another. Methods for doing this have been developed for some time, see [6, 9, 10, 11, 13, 16, 17, 19]. However, the stochastic analysis of these algorithms have often been based upon poorly motivated stochastic models. In particular, the effects of image discretisation, interpolation and scale-space smoothing is often neglected or not analysed in detail.

In this chapter, image acquisition, interpolation and scale-space smoothing are modelled into some detail. *Image acquisition* is viewed as a composition of blurring, ideal sampling and added noise, similar to [18]. The discrete signal is analysed after *interpolation*. This makes it possible to detect features on a sub-pixel level. Averaging or *scale-space smoothing* is used to reduce the effects of noise. To understand feature detection in this framework, one has to analyse the effect of noise on interpolated and smoothed signals. In doing so a theory is obtained that connects the discrete and continuous scale-space theories.

The chapter is organised as follows. Section 1.2 treats the image acquisition model. In Section 1.3 a method is proposed where the discrete scale-space is induced from the continuous scale-space theory. The stochastic properties of the intensity error field are discussed in Section 1.4. A short introduction to stationary random fields is given and some important results that are relevant for our model are demonstrated. The ideas are verified with numerical experiments on real images. The analysis of sub-pixel correlation and edge detection is commented upon briefly in Section 1.5. This is described in more detail in [2, 3, 4].

1.2 Image acquisition

To model the image acquisition process, the intensity distribution that would be caught by an ideal camera is first affected by aberrations in the optics of the real camera, e.g. blurring caused by spherical aberration, coma and astigmatism. Other aberrations deform the image, like Petzval field curvature and distorsion, see [12]. Such distorsion can typically be handled by geometric considerations in mid-level vision and will not be commented upon here. One way to model camera blur is to convolve the ideal intensity distribution with a kernel corresponding to the smoothing caused by the camera optics. This process also removes some amount of the high spatial frequencies.

In a video-camera, the blurred image intensity distribution is typically measured by a CCD array. One can think of each pixel intensity as the weighted mean of the intensity distribution in a window around the ideal pixel position. Taking the weighted mean around a position is equivalent to first convolving with the weighting kernel and then ideal sampling. Finally, due to quantisation and other errors, stochastic errors are introduced.

Led by this discussion we will use the following image acquisition model:

$$W_{\text{ideal}} \xrightarrow{\text{blur}} W \xrightarrow{\text{sampling}} w_0 \xrightarrow{\text{noise}} v_0 \ , \tag{1.1}$$

where *upper case letters*, W, denote signals with *continuous* parameters, whereas *lower case letters*, w, denote *discrete* signals, see Fig. 1.1. Here, and often in the sequel, we use the word signal synonymously with function, and discrete signal synonymously with sequence or function defined on $\mathbb{Z}^n$, for some n. These three steps of blurring, sampling and noise will now be discussed in a little more detail.

Here *blurring* is modelled as an abstract operator h, such that $W = h(W_{\text{ideal}})$. We assume that no aliasing effects are present, when the function W is sampled at integer positions. This makes it possible to reconstruct W from the sampled data.

Assumption 1.2.1. *All energy in the high spatial frequencies is cancelled before sampling. The function W is band-limited, i.e. $W \in \mathcal{B}(\mathbb{R}^n)$, where*

$$\mathcal{B}(\mathbb{R}^n) = \{W \in L_2(\mathbb{R}^n) \mid \operatorname{supp} \mathcal{F}W \subset (-1/2, 1/2)^n\}$$

■

In the definition of the Fourier transform, we use the formula

$$\mathcal{F}W(f) = \int_{\mathbb{R}^n} W(\tau) e^{-i2\pi f \cdot \tau} d\tau \ , \tag{1.2}$$

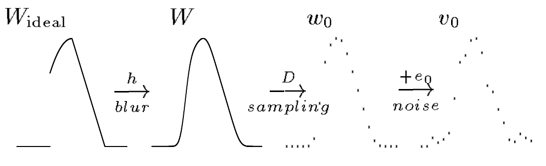

Figure 1.1: Illustration of the image acquisition model.

where $f \cdot \tau$ denotes scalar product.

The *sampling* is assumed to be ideal. Introduce the *sampling* or *discretisation* operator, $D : \mathcal{B} \to l_2$,

$$w(i,j) = (DW)(i,j) = W(i,j) \ . \tag{1.3}$$

Note that the sampling operator maps a continuous signal W onto a discrete signal w.

Finally *noise* is assumed to be an additive stationary random field. Experimentally it is verified that the errors in individual pixel intensities often can be modelled as independent random variables with similar distribution.

These assumptions will serve as an initial model. Further improvements can be made by a more detailed camera acquisition model. Nevertheless, these assumptions will help us to model and analyse the next stage, namely estimating the continuous image intensity distribution from the discrete image. Obviously, it is impossible to reconstruct the original intensity distribution W_{ideal} without prior knowledge. It is, however, reasonable to try to estimate the blurred and distorted intensity distribution W, or to estimate an even more blurred version.

1.3 Interpolation and smoothing

Scale-space theory and its application to computer vision has obvious advantages. In the continuous case, smoothing with the *Gaussian kernel*

$$G_b(x) = \frac{1}{\sqrt{2\pi b^2}} e^{-|x|^2/2b^2} \tag{1.4}$$

is very natural. Under some consistency conditions (symmetry, semi-group property, non-enhancement of local extrema), the Gaussian kernel is the only choice that gives a consistent scale-space theory, cf. [5, 14, 15, 20]. The *smoothing* operator S_b represents convolution with the Gaussian kernel G_b. A signal W is represented at scale b by its smoothed version W_b:

$$W_b = S_b(W) = G_b * W \ . \tag{1.5}$$

The signal W_b is called the *scale-space representation* of W, at scale b. In the sequel subscripts are used to denote different scales.

Scale-space theory in the discrete time case has been investigated in [15]. It turns out that just by sampling a continuous scale-space kernel, one obtains a discrete scale-space kernel. However, in doing so one does not obtain a scale-space theory with all the nice features of the continuous scale-space theory. There are difficulties with fine scales. In particular it is difficult to define higher order derivatives at fine scale levels. For the same reason it is difficult to define local extremum and zero crossings for fine scales. The semi-group property is lost. These questions are discussed in [15].

Interpolation and smoothing

The main idea of our approach is to induce the discrete signal, the scale spaces, etc. from the associated interpolated quantities. By an *interpolation* or *restoration* method we mean an operator that maps a discrete signal, w, to a continuous one, W. The following types of interpolation operators I_F will be used:

$$W(s) = (I_F w)(s) = \sum_i F(s-i)w(i) \ . \tag{1.6}$$

The ideal interpolation operator $I = I_{\text{sinc}}$ is of special interest and we propose to use it with discretisation D as mappings between the continuous and discrete signals to solve the restoration and discrete scale-space problems. In other words we relate the discrete and continuous signals through the operations of discretisation and ideal low-pass interpolation.

Note that if the camera induced blur cancels the high frequency components in W as in Assumption 1.2.1, the deterministic restoration $W_0 = I(w_0)$ is equal to W.

Using these definitions, the discrete and continuous scale-space representations can be defined simultaneously and consistently. We propose the following:

1. If the primary interest is the interpolated continuous signal, then *restore* the scale-space smoothed continuous signal W_b from the discrete signal w_0 first using ideal interpolation and then continuous scale-space smoothing.

2. If the primary interest is a discrete scale-space representation, then use the induced representation from the continuous scale-space representation.

The procedure is illustrated by the diagram:

$$\begin{array}{ccc} W_0 & \xleftarrow{\ I\ } & w_0 \\ {\scriptstyle S_b}\big\downarrow & & \big\downarrow{\scriptstyle s_b} \\ W_b & \xrightarrow{\ D\ } & w_b \end{array} \tag{1.7}$$

Thus, from the discrete signal w_0, the *continuous* scale-space smoothed signal W_b is obtained as $W_b = S_b(I(w))$. The *discrete* scale-space signal $w_b = s_b(w_0)$, is induced from the continuous scale-space signal, i.e.

$$w_b = s_b(w_0) \stackrel{\text{def}}{=} D(S_b(I(w_0))) \ , \tag{1.8}$$

where s_b is introduced as the discrete scale-space smoothing operator. Notice that s_b is a convolution with a kernel g_b,

$$g_b = D(G_b * \mathrm{sinc}) \ . \tag{1.9}$$

The differences between this approach and others, like the sampled Gaussian approach, is very small for large scales but significant for small scales. In fact it can be shown that

$$||\,\mathrm{sinc} * G_b - G_b||_2^2 \leq \frac{1}{b\sqrt{\pi}} \Phi(-\pi b \sqrt{2}) \ , \tag{1.10}$$

where Φ is the normal cumulative distribution function. Notice that the right hand side is small when b is large. The sampled Gaussian approach is also equivalent to using interpolation with the delta distribution followed by Gaussian smoothing. The main motivation for using ideal low-pass interpolation is, however, that the approach is well suited for stochastic analysis as will be shown later. Observe that the interpolated signal W is smooth. Thererfore, there is no difficulty in defining higher order derivatives.

This scale-space theory has several theoretical advantages: It works for all scales. The semi-group property, $s_{\sqrt{a}} s_{\sqrt{b}} = s_{\sqrt{a+b}}$, holds. The coupling to continuous scale-space theory gives a natural way to interpolate in the discrete space. There are no difficulties in defining derivatives at arbitrary scales. It is possible to calculate derivatives at arbitrary interpolated positions. Operators which commute in the continuous theory automatically commute in the discrete theory. The effect of additive stationary noise can easily be modelled. It makes it possible to compare the real intensity distribution with the interpolated distribution.

In practice this scale-space theory is difficult to use for small scale parameters, because of the large tail of the function in (1.9). However, the function $\mathrm{sinc} * G_b$ has a very small tail for larger scales. In practise one may use the approximation $\mathrm{sinc} * G_b \approx G_b$ for large scales, according to (1.10). This simplifies implementation substantially.

1.4 The random field model

The discrete image $v_0 = w_0 + e_0$ is analysed directly or through scale-space smoothing, as illustrated by the diagram:

$$\begin{array}{ccc} W_0 + E_0 & \xleftarrow{\ I\ } & w_0 + e_0 \\ S_b \downarrow & & \downarrow s_b \\ W_b + E_b & \xrightarrow{\ D\ } & w_b + e_b \end{array} \tag{1.11}$$

Note that all operations are linear. The stochastic and deterministic properties can, therefore, be studied separately and the final result is obtained by superposition. Thus with an a priori model on W_{ideal}, for example an ideal edge or corner, it is possible to predict the deterministic parts W_b and w_b. The stochastic properties of the error fields e_0, e_b, E_0 and E_b, will now be studied.

Stationary random fields

The theory of random fields is a simple and powerful way to model noise in signals and images. Stationary or wide sense stationary random fields are particularly easy to use. Denote by $\mathcal{E}$ the expectation value of a random variable.

Definition 1.4.1. A random field $X(t)$ with $t \in \mathbb{R}^n$ is called *stationary* or *wide sense stationary*, if its *mean* $m(t) = m_X(t) = \mathcal{E}[X(t)]$ is constant and if its *covariance function* $r_X(t_1, t_2) = \mathcal{E}[(X(t_1) - m(t_1))(X(t_2) - m(t_2))]$ only depends on the the difference $\tau = t_1 - t_2$. A random field $X(t)$ with $t \in \mathbb{R}^n$ is called *strictly stationary* if for all $(t_1, \ldots, t_n)$ and all τ the stochastic variable $(X(t_1), \ldots, X(t_n))$ has the same probability distribution as $(X(t_1 + \tau), \ldots, X(t_n + \tau))$. ■

For stationary fields we will use $r_X(s,t)$ and $r_X(s-t)$ interchangeably as the *covariance function*. The analogous definition is used for a stationary field in discrete parameters. The notion of *spectral density*

$$R_X(f) = (\mathcal{F} r_X)(f) = \int r_X(\tau) e^{-i2\pi f \cdot \tau} d\tau \tag{1.12}$$

is also important. Again the same definition can be used for random fields with discrete parameters $s \in \mathbb{Z}^n$, but whereas the spectral density for random fields with continuous parameters is defined for all frequencies f, the spectral density of discrete random fields is only defined on an interval $f \in [-1/2, 1/2]^n$. Introductions to the theory of random processes and random fields are given in [1, 7, 8]. In these books you will find that convolution, discretisation and derivation preserves stationarity.

$$\begin{aligned} w = D(W) \quad &\Rightarrow \quad r_w = D(r_W) \\ Y = h * X \quad &\Rightarrow \quad R_Y = R_X |\mathcal{F} h|^2 \\ Y = X' \quad &\Rightarrow \quad r_Y = -r_X'' \end{aligned}$$

It is shown in [2] that ideal interpolation I preserves stationarity as well.

Theorem 1.4.1 (Interpolation of a random field). *Let $e(i_1, \ldots, i_n)$ be a stationary discrete random field with zero mean and covariance function*

$$r_e((i_1, \ldots, i_n), (j_1, \ldots, j_n)) = r_e(i_1 - j_1, \ldots, i_n - j_n) \ ,$$

such that

$$r_e \in l^p, \qquad \text{for some } p < \infty \ .$$

Then the ideal interpolation of the discrete random field,

$$E = I(e) \ , \tag{1.13}$$

is a well defined random field in quadratic mean and E is stationary with covariance function

$$r_E(\tau) = I(r_e)(\tau) \ . \tag{1.14}$$

The corresponding theorem in higher dimensions can be proved in exactly the same manner.

Thus, all operations in the commutative diagram (1.11) preserve stationarity. This simplifies the modelling of errors in scale-space theory. The effects of the operators I, D, S_b and s_b on covariance r and spectral density R are all known by now.

It is often convenient to assume that the discrete noise e_0 can be modelled as white noise, i.e.

$$r_e(k) = \begin{cases} \epsilon^2, & \text{if } k = 0, \\ 0, & \text{if } k \neq 0. \end{cases}$$

It can then be shown that the covariance function of the interpolated and smoothed error field is

$$r_{E_b} = \epsilon^2 \operatorname{sinc} * G_{b\sqrt{2}} \ . \tag{1.15}$$

Remark. The restored image intensity distribution V_b is a sum of a deterministic part W_b and a stationary random field E_b. Notice that the restoration and the residual are *invariant* of the position of the discretisation grid. The effect of discretisation is thus removed. ■

1.5 Sub-pixel correlation and edge detection

Analysis of sub-pixel edge-detection and sub-pixel correlation are two application of our scale-space theory.

In *sub-pixel correlation* the idea is that, at least locally, the images only differ by an unknown translation ρ. Denote by $V = W + E$ and $\bar{V} = \bar{W} + \bar{E}$ the restored intensity fields in two images for a fixed scale b. The deterministic functions are identical except for a translation. For a fixed translation $\rho_0 = (\rho_1, \rho_2)$, we thus have

$$W(t) = \bar{W}(t + \rho_0), \qquad \forall t \ .$$

To determine the translation h with sub-pixel accuracy a least squares integral is minimised,

$$F(\rho) = \int_{t \in \Omega} (V(t) - \bar{V}(t + \rho))^2 dt \ .$$

Furthermore, the residual field $V(t) - \bar{V}(t + \rho)$ can be used to empirically study the stochastic properties of the camera noise e_0. The quality of the estimated sub-pixel translation,

$$\hat{\rho} = \operatorname{argmin} F(\rho) \ ,$$

can be analysed using the statistical model given above. Let $X = \hat{\rho} - \rho_0$ be the error in estimated translation. By linearising the function F it can be shown, see [2, 3], that the probability distribution of X can be approximated with a normal distribution with zero mean and covariance matrix given by

$$C = \mathcal{C}[X] \approx A^{-1} B A^{-T} \ , \tag{1.16}$$

$$A = 2 \int_{t_1 \in \Omega} (V \bar{W} \bar{W}^T)(t_1) dt_1 \ , \tag{1.17}$$

$$B = \int_{t_1 \in \Omega} ((V \bar{W}) * r_{E-\bar{E}})(t_1) \, (V \bar{W})(t_1) dt_1 \ . \tag{1.18}$$

Similarly, *sub-pixel edge detection* can be analysed by modelling the edge W_{ideal}, the blur h, the noise e_0 and analysing its effect on a sub-pixel edge detector. This is studied in [4]. For low levels and low curvature edges it can be shown that the deviations in the normal direction of the detected curve is approximately a stationary random process with respect to the arc-length of the curve.

1.6 Conclusions

In this chapter we have modelled the image acquisition process, taking into account both the deterministic and stochastic aspects. In particular the discretisation process is modeled in detail. This interplay between the continuous signal and its discretisation is very fruitful and the increased knowledge sheds light on scale-space theory, feature detection and stochastic modelling of errors.

The relation between the continuous signal and its discretisation is used to obtain an alternative scale-space theory for discrete signals. It is also used to derive methods of restoring the continuous scale-space representation from the discrete representation. This enables us to calculate derivatives at any position and of any scale.

Furthermore, the stochastic errors in images are modelled and new results are given that show how these errors influence the continuous and discrete scale-space representations and their derivatives. This information is crucial in understanding the stochastic behaviour of scale-space representations as well as fundamental properties of feature detectors. In particular, we have analysed a simple sub-pixel edge detector and a sub-pixel correlator in detail.

In order to validate the theory, experiments and simulations both on real and simulated data have been presented. Good agreement with the theoretical model is achieved.

The work can be extended in several directions. Edges were modelled as straight ideal step edges. It would be interesting to study the effect (the bias) caused by other types of edges and the effect of high curvature edges. The model of image acquisition, interpolation and scale space smoothing can also be used to analyse other feature detectors.

Acknowledgements

The authors thank Georg Lindgren for inspiration and valuable discussions.

Bibliography

[1] A. Adler. *The Geometry of Random Fields.* Wiley, New York, 1985.

[2] K. Åström. *Invariancy Methods for Points, Curves and Surfaces in Computational Vision.* PhD thesis, Dept of Mathematics, Lund University, Sweden, 1996.

[3] K. Åström and A. Heyden. Stochastic analysis of scale-space smoothing. In *Proc. International Conference on Pattern Recognition, Vienna, Austria*, 1996.

[4] K. Åström and A. Heyden. Stochastic analysis of sub-pixel edge detection. In *Proc. International Conference on Pattern Recognition, Vienna, Austria*, 1996.

[5] J. Babaud, A. P. Witkin, M. Baudin, and R. O. Duda. Uniqueness of the Gaussian kernel for scale-space filtering. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 8(1):26–33, 1986.

[6] F. Canny. A computational approach to edge detection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 8(6):676–698, 1986.

[7] H. Cramér and M. R. Leadbetter. *Stationary and Related Stochastic Processes.* Wiley, New York, 1967.

[8] N. A. C. Cressie. *Statistics for Spatial Data.* Wiley, New York, 1991.

[9] E. De Michelli, B. Caprile, P. Ottonello, and V. Torre. Localization and noise in edge detection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 10:1106–1117, 1989.

[10] R. Deriche. Using Canny's criteria to derive an optimal edge detector recursively implemented. *Int. Journal of Computer Vision*, 1:167–187, 1987.

[11] O. Faugeras. *Three-Dimensional Computer Vision.* MIT Press, Cambridge, Mass, 1993.

[12] E. Hecht. *Optics.* Addison-Wesley, Reading, Mass., 1987.

[13] E. Hildreth and D. Marr. Theory of edge detection. *Proceedings of Royal Society of London*, 207:187–217, 1980.

[14] J. J. Koenderink. The structure of images. *Biological Cybernetics*, 50:363–370, 1984.

[15] T. Lindeberg. *Scale-Space Theory in Computer Vision.* Kluwer Academic Publishers, 1994.

[16] D. Marr. *Vision.* W. H. Freeman and Company, 1982.

[17] V. S. Nalwa and T. O. Binford. On detecting edges. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 8:699–714, 1986.

[18] B. Pratt. *Digital Image Processing.* Wiley-Interscience, 1978.

[19] V. Torre and A. Poggio. On edge detection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 8:147–163, 1986.

[20] A. P. Witkin. Scale-space filtering. In *Proc. Eigth International Joint Conference on Artificial Intelligence, Karlsruhe, West Germany*, pages 1019–1022, 1983.