2. Foundations of scale-space

*"There are many paths to the top of the mountain,
but the view is always the same"* -Chinese proverb.

2.1 Constraints for an uncommitted front-end

To compute any type of representation from the image data, information must be extracted using certain *operators* interacting with the data. Basic questions then are: Which operators to apply? Where to apply them? How should they look like? How large should they be?

Suppose such an operator is the derivative operator. This is a difference operator, comparing two neighboring values at a distance close to each other. In mathematics this distance can indeed become infinitesimally small by taking the limit of the separation distance to zero, but in physics this reduces to the sampling distance as the smallest distance possible. Therefore we may foresee serious problems when we deal with such notions as mathematical differentiation on discrete data (especially for high order), and sub-pixel accuracy.

From this moment on we consider the aperture function as an operator: we will search for *constraints* to pin down the exact specification of this operator. We will find an important result: for an unconstrained front-end there is a *unique* solution for the operator. This is the Gaussian kernel $g(x; \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$, with σ the *width* of the kernel. It is the same bell-shaped kernel we know from probability theory as the probability density function of the *normal distribution*, where σ is the *standard deviation* of the distribution.

Interestingly, there have been many derivations of the front-end kernel, all leading to the unique Gaussian kernel.

This approach was pioneered by Iijima (figure 2.2) in Japan in the sixties [Iijima1962], but was unnoticed for decades because the work was in Japanese and therefore inaccessible for Western researchers.

Independently Koenderink in the Netherlands developed in the early eighties a rather complete multi-scale theory [Koenderink1984a], including the derivation of the Gaussian kernel and the linear diffusion equation.

```
<< FrontEndVision`FEV`;
```

$\sigma = 1$; Plot[$\frac{1}{\sqrt{2\pi\sigma^2}} E^{-\frac{x^2}{2\sigma^2}}$, {x, -4, 4}, ImageSize -> 200];

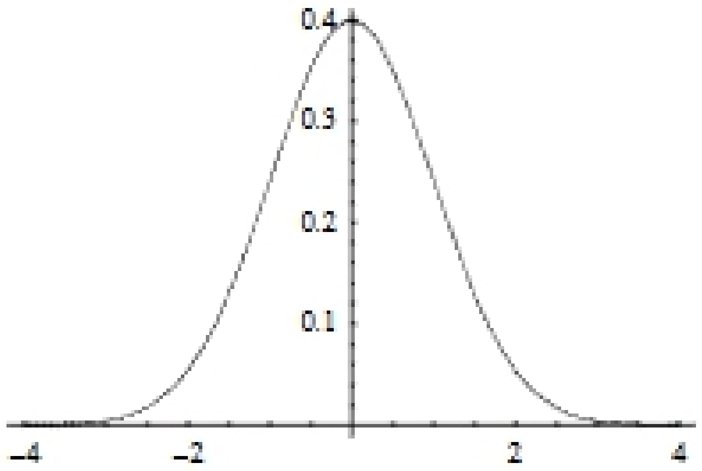

Figure 2.1 The Gaussian kernel with unit standard deviation in 1D.

Koenderink was the first to point out the important relation to the receptive field families in the visual system, as we will discuss in forthcoming chapters. Koenderink's work turned out to be monumental for the development of scale-space theory. Lindeberg pioneered the field with a tutorial book [Lindeberg1994a]. The papers by Weickert, Ishikawa and Imija (who together discovered this Japanese connection) present a very nice review on these early developments [Weickert1997a, Weickert1999a].

```
Show[Import["Iijima.gif"], ImageSize -> 150];
```

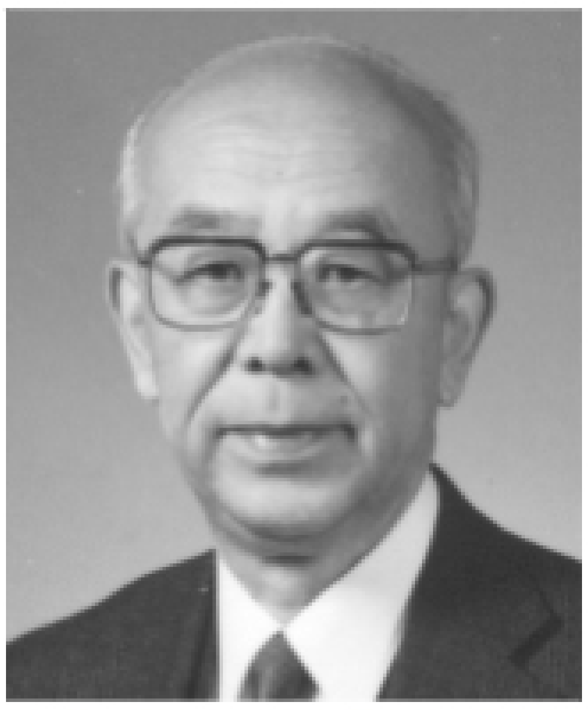

Fig. 2.2 Prof. Taizo Iijima, emeritus prof. of Tokyo Technical University, Japan, was the first to publish the axiomatic derivation of 'the fundamental equation of figure'.

We will select and discuss two fundamentally different example approaches to come to the Gaussian kernel in this book:

1. An axiomatic approach based on dimensional analysis and the notion of having 'no preferences' (section 2.2);
2. An approach based on the maximization of local entropy in the observation (section 2.5);

2.2 Axioms of a visual front-end

The line of reasoning presented here is due to Florack et al. [Florack1992a]. The requirements can be stated as *axioms*, or postulates for an uncommitted visual front-end. In essence it is the mathematical formulation for being uncommitted: "we know nothing", or "we have no preference whatsoever".

• **linearity**: we do not allow any nonlinearities at this stage, because they involve knowledge of some kind. So: no knowledge, no model, no memory;

• **spatial shift invariance**: no preferred location. Any location should be measured in the same fashion, with the same aperture function;

• **isotropy**: no preferred orientation. Structures with a particular orientation, like vertical trees or a horizontal horizon, should have no preference, any orientation is just as likely. This necessitates an aperture function with a circular integration area.

• **scale invariance**: no preferred size, or scale of the aperture. Any size of structure, object, texture etc. to be measured is at this stage just as likely. We have no reason to look only through the finest of apertures. The visual world consists of structures at any size, and they should be measured at any size.

In order to use these constraints in a theory that sets up the reasoning to come to the aperture profile formula, we need to introduce the concept of dimensional analysis.

2.2.1 Dimensional analysis

Every physical unit has a *physical dimension*.

It is this that mostly discriminates physics from mathematics. It was Baron Jean-Baptiste Fourier who already in 1822 established the concept of dimensional analysis [Fourier1955]. This is indeed the same mathematician so famous for his Fourier transformation.

```
Show[Import["Fourier.jpg"], ImageSize → 140];
```

Figure 2.3 Jean-Baptiste Fourier, 1792-1842.