Anisotropic diffusion processes in early vision

P. Perona
EECS Dept. – Cory Hall
University of California, Berkeley
Berkeley, CA 94720, USA

Abstract

Images often contain information at a number of different scales of resolution. For instance, the edges of a tree are different at a fine scale than at a coarse scale. At a fine scale we detect the boundary of every single leaf and bark crack, whereas, at a coarse scale just treetop and trunk are outlined. Therefore images have to be analyzed at multiple resolutions. The collection of images from fine to coarse scale that one uses to perform such an analysis is called "scale-space". The definition and generation of a "good" scale-space is a key step in early vision.

The scale-space is traditionally generated by low-pass filtering the image with Gaussian kernels. However, this process is conceptually incorrect, since it results in averaging brightness values belonging to different objects. Moreover, it has the undesirable effects of deforming the shape of the objects in the image and destroying the edge junctions which contain much of the geometrical information.

We can define a scale-space in which object boundaries are respected and smoothing only takes place within these boundaries. Such a scale-space can be generated solving a nonlinear diffusion differential equation forward in time (the scale parameter). The original image is used as the initial condition, and the conduction coefficient $c(x, y, t)$ varies in space and scale as a function of the gradient of the variable of interest (e.g. the image brightness). Such a diffusion process has an intrinsically local and parallel computational structure and may be implemented both on simple analog networks and digital architectures for achieving real-time speed.

Diffusion processes may be used for edge-detection, normalizing, equalizing, and compressing images, and estimating the local noise level in a scale-dependent fashion. Our algorithms are based on comparing the local values of different diffusion processes running in parallel on the same image. All of them can be implemented using passive locally-connected networks composed of capacitors, diodes, and resistors, making it reasonable to propose integrating sensors and early vision processing on the same silicon chip.