A Bayesian framework for computer vision

In many computer vision applications the input images are noisy and the data possibly sparse. These applications include, (i) identification of abnormalities in human organs, (ii) the computation of anatomical and physiological properties, both for medical images, e.g. liver tumors from CT images, human heart volume from MRI imagery, (iii) depth information for moving robots or automatic mapping of satellite image pairs, and (iv) inspection of faults on large scale manufactures. The common problems are then to "clean" the images from the noisy and/or sparse data and to segment the image. Particular emphasis is placed on image segmentation, that is, separating objects from background.

On June 5, 1992 Davi Geiger from Siemens Corporate Research presented a Bayesian framework for modeling the problem; this model which originated by S. Geman and D. Geman [1] postulates a Markov field framework in order to derive posterior distribution from prior (or belief) together with some noisy data. He then proposed mean field equations approach to obtain fast computational techniques; this work was carried out in [2], as well as in [3].

18.1 The Markov random field approach

Prototypical problems in Computer Vision include:

(i) Enhancement and segmentation from dense image data (the data are noisy);

(ii) Surface reconstruction from sparse data;

(iii) Matching (binocular stereo) two pictures taken from slightly different angles;

(iv) Data fusion (integration of stereo, motion, color, etc.).

An interesting problem is also the enhancement of image sequences in the presence of motion; here the Kalman filter might be used.
There are several different approaches to model images and their boundary: Geman and Geman [1] introduced a Bayesian scheme to restore image. Blake and Zisserman [4] have also used the functional introduced in [1]; however they applied directly optimization method instead of the simulated annealing method of [1]. Mumford and Shah [5] developed a variational approach to which PDE and geometric measure theory methods can be usefully applied.

The basic assumption in all these approaches is that the domain of vision $R$ can be broken into a finite number of regions $R_1 \cup R_2 \cup \cdots \cup R_n$ such that

(i) the image $g$ varies smoothly or slowly within each $R_i$, and

(ii) the image $g$ varies discontinuously on rapidly across the boundary $\Gamma$ between the different $R_i$.

Here $g$ is the intensity of light on film or retina (if the image is taken by camera or eye).

The functional introduced by Mumford and Shah [5] is

$$E(f, \Gamma) = \frac{1}{2} \iint (f - g)^2 \, dx \, dy + \iint |\nabla f|^2 \, dx \, dy + \nu |\Gamma|$$

where $|\Gamma|$ is the total length of $\Gamma$ and $\mu, \nu$ are positive parameters. The variables $f, \Gamma$ have to be chosen so as to minimize $E$.

Geman and Geman [1] introduced a Markov random field approach as explained below.

Consider the so called, weak membrane problem of membrane reconstruction. We divide the domain $R$ on which the surface is to be reconstructed by uniform orthogonal mesh. To each square $R_{ij}$ we assign a random pair

$$(f_{ij}, \ell_{ij})$$

where $f_{ij}$ represents the approximate level of the membrane surface over $R_{ij}$ and

$$\ell_{ij} = 1 \quad \text{if a boundary passes through } R_{ij},$$

$$\ell_{ij} = 0 \quad \text{otherwise.}$$

We have made a small modification on Geman and Geman's model by considering a single line process, and not the original two line processes, a horizontal and vertical one.

We define a probability $P$ by Gibbs distribution

$$P(f_{ij}, \ell_{ij}) = \frac{1}{Z} e^{-\beta [\mu Q(f)_{ij}(1-\ell_{ij}) + \gamma \ell_{ij} \ell_{ij}]}$$

(18.2)