***Lecture 6***

***Fourth stage***

***Medical Physical Department***

***Medical Image Analysis***

**Noise Reduction**

***By***

***Asst. Prof. Dr. Mehdi Ebady Manaa***

1. **Noise Reduction**

***Noise Reduction*** Noise is usually modeled stationary, additive, and with zero mean. A noisy image g is related to the unknown noise-free image f through g = f + n, where n is the zero mean noise. Noise removal through linear filtering consists of estimating the expected value E(g). Since E(n) = 0, we have

E(g) = E(f ) + E(n) = E(f ) = f, because the deterministic function f has E(f ) = f Hence, noise reduction schemes try to reconstruct E(g) from various assumptions about f as follows.

* **Linear filtering** assumes f to be locally constant. E is then estimated by averaging over a local neighborhood.
* **Median filtering** assumes that noise is normally distributed, f is locally constant, *except for the edges, the signal at the edges is higher than the noise, and the edges are locally straight*.
* **Diffusion filtering** assumes that f is locally constant, *except at the edges*, and that the properties of the edges and noise can be differentiated by amplitude or frequency.
* **Bayesian image** restoration requires f to be locally smooth, *except for the edges*. It further requires that in some local neighborhood edge pixels are not the majority of all pixels in that neighborhood. Since most of the assumptions are not true everywhere in the image, filtering results in the various filter-specific artefacts
1. **Noise Reduction by Linear Filtering**

If f is constant in some neighborhood around (i,j), E(i,j) can be estimated by *averaging over this neighborhood*. The operation can be carried out in the spatial domain as a convolution with a mean (sometimes called boxcar filter) of size s:



where “\*” stands for *the convolution operation*. The convolution kernel cboxcar,s is a square matrix of size s × s with s being odd and the filter being centered.



The estimate of E(g(i,j)) improves with the size of s, but the likelihood that f is constant in this neighborhood for all locations (i,j) decreases, leading to *increased blurring at the edges.* Question, is this filter HPF, or LPF?

 

Fig. 6.1 neighborhood averaging Filter

1. **Edge-Preserving Smoothing: Median Filtering**

A median filter is a nonlinear rank order filter that selects the result at some location (i,j) *from an ordered list of values*. The values are taken from pixels in the neighborhood of (i,j). A filter is a median filter if the median rank is selected from the list. A 3 ×3 median filter would thus sort nine pixel values—the pixels *in an 8- neighborhood of (i,j) plus (i,j) itself*—and select the one ranked at position **five**. The neighborhood is usually square-shaped covering an odd number of pixels (e.g., 3 × 3, 5 × 5, or 7 × 7 neighborhoods). *The result of a median filter approaches تقترب the expected value with increasing filter size if Gaussian noise can be assumed.* Under such conditions, the median filter has similar noise reduction capabilities as the linear filters presented previous section. However, there is more to the median filter. If the neighborhood region of a median filter contains an edge, it preserves the edge under the following conditions (compare results in Fig. 6.2).

* The edge is straight within the neighborhood region.
* The signal difference of the two regions incident تحدثto the edge exceeds the noise amplitude.
* The signal is locally constant within each of the two regions

 

Fig. 6.2 Comparison between boxcar filter (a) and median filter (b) of the same size. It can be seen in the enlarged part that the median filter preserves sharp edges while the boxcar filter produces blur

1. **Edge-Preserving Smoothing: Diffusion Filtering**

The median filter has two disadvantages:

1. It does not remove noise at edges even if the edge follows the implicit edge model, and ,
2. it may alter edges in a random fashion that does not follow the edge model.

***Diffusion filtering*** is an alternative that enables smoothing at edges. It may also accommodate a wide range of edge models. Diffusion filtering uses the diffusion of a liquid or gaseous material as a model for noise reduction. Using homogeneous and inhomogeneous diffusion for data filtering was first proposed by Perona and Malik (1990) and has since become very popular for edge-preserving smoothing. The popularity is probably due to the relative simplicity of the concepts, the intuitive behavior of a diffusion process, and a fairly simple implementation. For applying a diffusion process as a filter, *image intensity is taken as material density*. Noise is taken as density variation and diffusion is carried out iteratively. After an infinite number of iterations, homogeneous diffusion levels any density inhomogeneity resulting in a noise-free image without edges. Diffusion across edges should be inhibited for edge enhancement. Since the boundaries are unknown (otherwise edge-preserving smoothing would be trivial), the edge response from edge enhancement is used to indicate the potential boundary locations. Inhomogeneous diffusion treats such boundary locations as a semi-permeable material. ***The process will still level densities (i.e., image intensities), but the process will be slower at potential edge locations.*** It will not be prohibited, however, because the gradient response may also be caused by noise. Hence, noise removal should stop after a number of iterations to prevent leveling the intensity difference between objects. The stopping criterion depends on the image characteristics and on the parameters of the diffusion equation. Preserving boundaries may be improved if, instead of restricting any kind of diffusion at edges, it is only restricted across the edges. This is called anisotropic diffusionالانتشار متباين الخواص Allowing diffusion parallel to an edge enables noise removal by smoothing while inhibiting diffusion across an edge. *Gradient direction is used as a discriminative feature between noise and the edges. The gradients between adjacent edge pixels tend to have similar directions while this is not true for the gradients of adjacent noise pixels*. Diffusion in regions with noise pixels will be in random directions supporting homogenization while it will be directed in boundary regions supporting contrast enhancement.

 

Fig. 6.3 (a) Homogeneous diffusion is only dependent on the density gradient, (b) inhomogeneous diffusion decrease at edges, (c) anisotropic diffusion decrease at edge in edge normal direction

 

Fig. 6.4 Comparison of the different types of diffusion: (a) homogeneous diffusion, (b) inhomogeneous diffusion, (c) anisotropic inhomogeneous diffusion

1. **Edge-Preserving Smoothing: Bayesian Image Restoration**

Linear filtering assumes existing prior knowledge about the noise, but nothing about the image. Restoration in a probabilistic framework is able to incorporate smoothness constraints into the method. For this purpose, the image characteristics are assumed to be representable by a Markov random field (MRF). Determining an unknown image function f from some observation g—which does not necessarily have to be an image, but must be associated with the image in some known fashion—given some probabilistic knowledge about the nature of the mapping from f to g is a powerful tool based on a simple concept. Its application fields in image processing are image reconstruction—e.g., in the reconstruction of nuclear images -noise removal, segmentation, and classification. For simplifying the notation, we assume that the pixels of an image are ordered as a vector. A vector **f** shall be restored from an observed vector **g**. An individual pixel in the vector will be represented by *fi* and *gi*, respectively. Noise reduction in a Bayesian framework searches for an image **f** that maximizes the conditional probability of observing a noisy image g, given that the true image is f. In Bayesian notation, we have

 P(f|g) ∝ P(g|f) · P(f).

The term P(g|f) is the data term since it describes the dependency of the observation on the unknown undistorted data f. The term P(f) comprises domain knowledge about f independent of an observation (e.g., that in most images neighboring scene elements have similar values.