Imperial College of Science Technology and Medicine

Department of Electrical and Electronic Engineering

Digital Image Processing

PART 3

IMAGE RESTORATION

Academic responsible

Dr. Tania STATHAKI

Room 812

Ext. 46229

Email: t.stathaki@ic.ac.uk

What is image restoration?

Image Restoration refers to a class of methods that aim to remove or reduce the degradations that have occurred while the digital image was being obtained.

All natural images when displayed have gone through some sort of degradation:

- during display mode
- acquisition mode, or
- processing mode

The degradations may be due to

- sensor noise
- blur due to camera misfocus
- relative object-camera motion
- random atmospheric turbulence
- others

In most of the existing image restoration methods we assume that the degradation process can be described using a mathematical model.

How well can we do?

Depends on how much we know about

- the original image
- the degradations

(how accurate our models are)

Image restoration and image enhancement-differences:

- Image restoration differs from image enhancement in that the latter is concerned more with accentuation or extraction of image features rather than restoration of degradations.
- Image restoration problems can be quantified precisely, whereas enhancement criteria are difficult to represent mathematically.

Image observation models

Typical parts of an imaging system: image formation system, a detector and a recorder. A general model for such a system could be:

$$y(i,j) = r[w(i,j)] + n(i,j)$$

$$w(i,j) = H[f(i,j)] = \int\int_{-\infty}^{\infty} h(i,j,i',j') f(i',j') di' dj'$$

$$n(i,j) = g[r[w(i,j)]]n_1(i,j) + n_2(i,j)$$

where $y(i,j)$ is the degraded image, $f(i,j)$ is the original image and $h(i,j,i',j')$ is an operator that represents the degradation process, for example a blurring process.

Functions $g(\cdot)$ and $r(\cdot)$ are generally nonlinear, and represent the characteristics of detector/recording mechanisms. $n(i,j)$ is the additive noise, which has an image-dependent random component $g[r[H[f(i,j)]]]n_1(i,j)$ and an image-independent random component $n_2(i,j)$.

Detector and recorder models

The response of image detectors and recorders in general is nonlinear.

An example is the response of image scanners

$$r(i,j) = \alpha w(i,j)^{\beta}$$

where α and β are device-dependent constants and $w(i,j)$ is the input blurred image.

For photofilms

$$r(i,j) = \gamma \log_{10} w(i,j) - r_0$$

where γ is called the gamma of the film, $w(i,j)$ is the incident light intensity and $r(i,j)$ is called the optical density. A film is called positive if it has negative γ.

Noise models

The general noise model

$$n(i,j) = g[r[w(i,j)]]n_1(i,j) + n_2(i,j)$$

is applicable in many situations. Example, in photoelectronic systems we may have $g(x) = \sqrt{x}$. Therefore

$$n(i,j) = \sqrt{\alpha w(i,j)^{\beta}}\, n_1(i,j) + n_2(i,j)$$

where n_1 and n_2 are zero-mean, mutually independent, Gaussian white noise fields.

The term $n_2(i,j)$ may be referred as thermal noise.

In the case of films there is no thermal noise and the noise model is

$$n(i,j) = \sqrt{\gamma \log_{10} w(i,j) - r_o}\, n_1(i,j)$$

Because of the signal-dependent term in the noise model, restoration algorithms are quite difficult. Often $w(i,j)$ is replaced by its spatial average, μ_w, giving

$$n(i,j) = g\left[r\left[\mu_w\right]\right]n_1(i,j) + n_2(i,j)$$

which makes $n(i,j)$ a Gaussian white noise random field.

A lineal observation model for photoelectronic devices is

$$y(i,j) = w(i,j) + \sqrt{\mu_w}\,n_1(i,j) + n_2(i,j)$$

For photographic films with $\gamma = -1$

$$y(i,j) = -\log_{10} w(i,j) - r_0 + a n_1(x,y)$$

where r_0, a are constants and r_0 can be ignored.

The light intensity associated with the observed optical density $y(i,j)$ is

$$I(i,j) = 10^{-y(i,j)} = w(i,j)10^{-an_1(i,j)} = w(i,j)n(i,j)$$

where $n(i,j) \hat{=} 10^{-an_1(i,j)}$ now appears as multiplicative noise having a log-normal distribution.

Keep in mind that we are just referring to the most popular image observation models. In the literature you can find a quite large number of different image observation models!

Image restoration algorithms are based on (derived from) the above image formation models!

A general model of a simplified digital image degradation process

A simplified version for the image restoration process model is

$$y(i,j) = H\left[f(i,j)\right] + n(i,j)$$

where

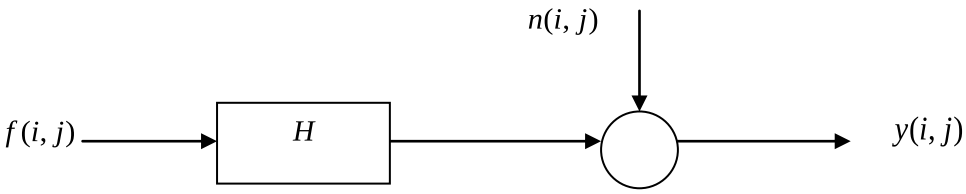
$y(i,j)$ the degraded image

$f(i,j)$ the original image

H an operator that represents the degradation process

$n(i,j)$ the external noise which is assumed to be image-independent

We see in the figure below a schematic diagram for a generic degradation process described by the above simplified model



Possible classification of restoration methods

Restoration methods could be classified as follows:

- **deterministic:** we work with sample by sample processing of the observed (degraded) image
- **stochastic:** we work with the statistics of the images involved in the process

- **non-blind:** the degradation process H is known
- **blind:** the degradation process H is unknown
- the degradation process H could be considered **partly known**

From the viewpoint of implementation:

- **direct**
- **iterative**
- **recursive**

Definitions

We again consider the general degradation model

$$y(i, j) = H[f(i, j)] + n(i, j)$$

If we ignore the presence of the external noise $n(i, j)$ we get

$$y(i, j) = H[f(i, j)]$$

H is ***linear*** if

$$H[k_1 f_1(i, j) + k_2 f_2(i, j)] = k_1 H[f_1(i, j)] + k_2 H[f_2(i, j)]$$

H is ***position*** (or ***space***) ***invariant*** if

$$H[f(i - a, j - b)] = y(i - a, j - b)$$

From now on we will deal with linear, space invariant type of degradations.

In a real life problem many types of degradations can be approximated by linear, position invariant processes!

Advantage: Extensive tools of linear system theory become available.

Disadvantage: In some real life problems ***nonlinear*** and ***space variant*** models would be more appropriate for the description of the degradation phenomenon.

Typical linear position invariant degradation models

- Motion blur. It occurs when there is relative motion between the object and the camera during exposure.

$$h(i) = \begin{cases} \dfrac{1}{L+1}, & \text{if } -\dfrac{L}{2} \le i \le \dfrac{L}{2} \\ 0, & \text{otherwise} \end{cases}$$

- Atmospheric turbulence. It is due to random variations in the reflective index of the medium between the object and the imaging system and it occurs in the imaging of astronomical objects.

$$h(i, j) = K \exp\left(-\frac{i^2 + j^2}{2\sigma^2} \right)$$

- Uniform out of focus blur

$$h(i, j) = \begin{cases} \dfrac{1}{\pi R^2}, & \text{if } \sqrt{i^2 + j^2} \le R \\ 0, & \text{otherwise} \end{cases}$$

- Uniform 2-D blur

$$h(i, j) = \begin{cases} \dfrac{1}{(L+1)^2}, & \text{if } -\dfrac{L}{2} \le i, j \le \dfrac{L}{2} \\ 0, & \text{otherwise} \end{cases}$$

- …

Some characteristic metrics for degradation models

- **Blurred Signal-to-Noise Ratio (BSNR)**: a metric that describes the degradation model.

$$\text{BSNR} = 10\log_{10}\left\{ \frac{\frac{1}{MN}\sum_i \sum_j \left[g(i, j) - \bar{g}(i, j) \right]^2}{\sigma_n^2} \right\}$$

$g(i, j) = y(i, j) - n(i, j)$

$\bar{g}(i,j) = E\{g(i,j)\}$

σ_n^2 : variance of additive noise

- **Improvement in SNR (ISNR)**: validates the performance of the image restoration algorithm.

$$\text{ISNR} = 10\log_{10}\left\{\frac{\sum_i\sum_j\left[f(i,j) - y(i,j)\right]^2}{\sum_i\sum_j\left[f(i,j) - \hat{f}(i,j)\right]^2}\right\}$$

where $\hat{f}(i,j)$ is the restored image.

Both BSNR and ISNR can only be used for simulation with artificial data.

One dimensional discrete degradation model

Suppose we have a one-dimensional discrete signal $f(i)$ of size A samples $f(0), f(1), \ldots, f(A-1)$, which is due to a degradation process.

The degradation can be modeled by a one-dimensional discrete impulse response $h(i)$ of size B samples. If we assume that the degradation is a causal function we have the samples $h(0), h(1), \ldots, h(B-1)$.

We form the extended versions of $f(i)$ and $h(i)$, both of size $M \geq A+B-1$ and periodic with period M. These can be denoted as $f_e(i)$ and $h_e(i)$.

For a time invariant degradation process we obtain the discrete convolution formulation as follows

$$y_e(i) = \sum_{m=0}^{M-1} f_e(m)h_e(i-m) + n_e(i)$$

Using matrix notation we can write the following form

$$\mathbf{y} = \mathbf{Hf} + \mathbf{n}$$

$$\mathbf{f} = \begin{bmatrix} f_e(0) \\ f_e(1) \\ \vdots \\ f_e(M-1) \end{bmatrix},$$

$$\underset{(M\times M)}{\mathbf{H}} = \begin{bmatrix} h_e(0) & h_e(-1) & \ldots & h_e(-M+1) \\ h_e(1) & h_e(0) & \ldots & h_e(-M+2) \\ \vdots & \vdots & \ddots & \vdots \\ h_e(M-1) & h_e(M-2) & \ldots & h_e(0) \end{bmatrix}$$

At the moment we decide to ignore the external noise $\mathbf{n}$.

Because h is periodic with period M we have that

$$\underset{(M\times M)}{\mathbf{H}} = \begin{bmatrix} h_e(0) & h_e(M-1) & \dots & h_e(1) \\ h_e(1) & h_e(0) & \dots & h_e(2) \\ \vdots & \vdots & \ddots & \vdots \\ h_e(M-1) & h_e(M-2) & \dots & h_e(0) \end{bmatrix}$$

We define $\lambda(k)$ to be

$$\lambda(k) = h_e(0) + h_e(M-1)\exp(j\frac{2\pi}{M}k) + h_e(M-2)\exp(j\frac{2\pi}{M}2k) + \dots$$

$$\dots + h_e(1)\exp[j\frac{2\pi}{M}(M-1)k], \quad k = 0,1,\dots,M-1$$

Because $\exp[j\frac{2\pi}{M}(M-i)k] = \exp(-j\frac{2\pi}{M}ik)$ we have that

$$\lambda(k) = MH(k)$$

$H(k)$ is the discrete Fourier transform of $h_e(i)$.

I define $\mathbf{w}(k)$ to be

$$\mathbf{w}(k) = \begin{bmatrix} 1 \\ \exp(j\frac{2\pi}{M}k) \\ \vdots \\ \exp[j\frac{2\pi}{M}(M-1)k] \end{bmatrix}$$

It can be seen that

$$\mathbf{H}\mathbf{w}(k) = \lambda(k)\mathbf{w}(k)$$

This implies that $\lambda(k)$ is an eigenvalue of the matrix $\mathbf{H}$ and $\mathbf{w}(k)$ is its corresponding eigenvector.

We form a matrix $\mathbf{w}$ whose columns are the eigenvectors of the matrix $\mathbf{H}$, that is to say

$$\mathbf{W} = \begin{bmatrix} \mathbf{w}(0) & \mathbf{w}(1) & \dots & \mathbf{w}(M-1) \end{bmatrix}$$

$$w(k,i) = \exp\left[j\frac{2\pi}{M}ki\right] \text{ and } w^{-1}(k,i) = \frac{1}{M}\exp\left[-j\frac{2\pi}{M}ki\right]$$

We can then diagonalize the matrix $\mathbf{H}$ as follows

$$\mathbf{H} = \mathbf{W}\mathbf{D}\mathbf{W}^{-1} \Rightarrow \mathbf{D} = \mathbf{W}^{-1}\mathbf{H}\mathbf{W}$$

where

$$\mathbf{D} = \begin{bmatrix} \lambda(0) & & & \mathbf{0} \\ & \lambda(1) & & \\ & & \ddots & \\ \mathbf{0} & & & \lambda(M-1) \end{bmatrix}$$

Obviously $\mathbf{D}$ is a diagonal matrix and

$$D(k,k) = \lambda(k) = MH(k)$$

If we go back to the degradation model we can write

$$\mathbf{y} = \mathbf{H}\mathbf{f} \Rightarrow \mathbf{y} = \mathbf{W}\mathbf{D}\mathbf{W}^{-1}\mathbf{f} \Rightarrow \mathbf{W}^{-1}\mathbf{y} = \mathbf{D}\mathbf{W}^{-1}\mathbf{f} \Rightarrow$$

$$Y(k) = MH(k)F(k), k = 0,1,\ldots, M-1$$

$Y(k), H(k), F(k), k = 0,1,\ldots, M-1$ are the $M-$sample discrete Fourier transforms of $y(i), h(i), f(i)$, respectively.

So by choosing $\lambda(k)$ and $\mathbf{w}(k)$ as above and assuming that $h_e(i)$ is periodic, we start with a matrix problem and end up with M scalar problems.

Two dimensional discrete degradation model

Suppose we have a two-dimensional discrete signal $f(i, j)$ of size $A \times B$ samples which is due to a degradation process.

The degradation can now be modeled by a two dimensional discrete impulse response $h(i, j)$ of size $C \times D$ samples.

We form the extended versions of $f(i, j)$ and $h(i, j)$, both of size $M \times N$, where $M \geq A + C - 1$ and $N \geq B + D - 1$, and periodic with period $M \times N$. These can be denoted as $f_e(i, j)$ and $h_e(i, j)$.

For a space invariant degradation process we obtain

$$y_e(i, j) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} f_e(m,n) h_e(i-m, j-n) + n_e(i, j)$$

Using matrix notation we can write the following form

$$\mathbf{y} = \mathbf{Hf} + \mathbf{n}$$

where $\mathbf{f}$ and $\mathbf{y}$ are $MN-$dimensional column vectors that represent the lexicographic ordering of images $f_e(i, j)$ and $h_e(i, j)$ respectively.

$$\mathbf{H} = \begin{bmatrix} \mathbf{H_0} & \mathbf{H_{M-1}} & \cdots & \mathbf{H_1} \\ \mathbf{H_1} & \mathbf{H_0} & \cdots & \mathbf{H_2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{H_{M-1}} & \mathbf{H_{M-2}} & \cdots & \mathbf{H_0} \end{bmatrix}$$

$$\mathbf{H}_j = \begin{bmatrix} h_e(j,0) & h_e(j,N-1) & \cdots & h_e(j,1) \\ h_e(j,1) & h_e(j,0) & \cdots & h_e(j,2) \\ \vdots & \vdots & \ddots & \vdots \\ h_e(j,N-1) & h_e(j,N-2) & \cdots & h_e(j,0) \end{bmatrix}$$

The analysis of the diagonalisation of $\mathbf{H}$ is a straightforward extension of the one-dimensional case.

In that case we end up with the following set of $M \times N$ scalar problems.

$$Y(u,v) = MNH(u,v)F(u,v)(+N(u,v))$$

$$u = 0,1,\ldots, M-1, v = 0,1,\ldots, N-1$$

In the general case we may have two functions $f(i), A \le i \le B$ and $h(i), C \le i \le D$, where A, C can be also negative (in that case the functions are **non-causal**). For the periodic convolution we have to extend the functions from both sides knowing that the convolution is $g(i) = h(i) * f(i), A + C \le i \le B + D$.

Deterministic approaches to restoration

DIRECT METHODS

1. Inverse filtering

The objective is to minimize

$$J(\mathbf{f}) = \|\mathbf{n}(\mathbf{f})\|^2 = \|\mathbf{y} - \mathbf{Hf}\|^2$$

We set the first derivative of the cost function equal to zero

$$\frac{\partial J(\mathbf{f})}{\partial \mathbf{f}} = 0 \Rightarrow -2\mathbf{H}^{\mathbf{T}}(\mathbf{y} - \mathbf{Hf}) = \mathbf{0}$$

$$\mathbf{H}^{\mathbf{T}}\mathbf{Hf} = \mathbf{H}^{\mathbf{T}}\mathbf{y}$$

If $M = N$ and $\mathbf{H}^{-1}$ exists then

$$\mathbf{f} = \mathbf{H}^{-1}\mathbf{y}$$

According to the previous analysis if $\mathbf{H}$ (and therefore $\mathbf{H}^{-1}$) is block circulant the above problem can be solved as a set of $M \times N$ scalar problems as follows

$$F(u,v) = \frac{H^*(u,v)Y(u,v)}{|H(u,v)|^2} \Rightarrow f(i,j) = \mathfrak{I}^{-1}\left[\frac{H^*(u,v)Y(u,v)}{|H(u,v)|^2}\right] = \frac{Y(u,v)}{H(u,v)}$$

Computational issues concerning inverse filtering

(I)

Suppose first that the additive noise $n(i,j)$ is negligible. A problem arises if $H(u,v)$ becomes very small or zero for some point (u,v) or for a whole region in the (u,v) plane. In that region inverse filtering cannot be applied.

- Note that in most real applications $H(u,v)$ drops off rapidly as a function of distance from the origin !

Solution: if these points are known they can be neglected in the computation of $F(u,v)$.

(II)

In the presence of external noise we have that

$$\hat{F}(u,v) = \frac{H^*(u,v)\big(Y(u,v) - N(u,v)\big)}{|H(u,v)|^2} =$$

$$\frac{H^*(u,v)Y(u,v)}{|H(u,v)|^2} - \frac{H^*(u,v)N(u,v)}{|H(u,v)|^2} \Rightarrow$$

$$\hat{F}(u,v) = F(u,v) - \frac{N(u,v)}{H(u,v)}$$

If $H(u,v)$ becomes very small, the term $N(u,v)$ dominates the result.

Solution: again to carry out the restoration process in a limited neighborhood about the origin where $H(u,v)$ is not very small.

This procedure is called **pseudoinverse filtering**.

In that case we set

$$\hat{F}(u,v) = \begin{cases} \dfrac{H^*(u,v)Y(u,v)}{|H(u,v)|^2} & H(u,v) \neq 0 \\ 0 & H(u,v) = 0 \end{cases}$$

or

$$\hat{F}(u,v) = \begin{cases} \dfrac{H^*(u,v)Y(u,v)}{|H(u,v)|^2} & |H(u,v)| \geq \varepsilon \\ 0 & \text{otherwise} \end{cases}$$

- In general, the noise may very well possess large components at high frequencies (u,v), while $H(u,v)$ and $Y(u,v)$ normally will be dominated by low frequency components.
- ε is a small number chosen by the user.

2. Constrained least squares (CLS) restoration

It refers to a very large number of restoration algorithms.

The problem can be formulated as follows.

minimize

$$J(\mathbf{f}) = \|\mathbf{n}(\mathbf{f})\|^2 = \|\mathbf{y} - \mathbf{H}\mathbf{f}\|^2$$

subject to

$$\|\mathbf{C}\mathbf{f}\|^2 < \varepsilon$$

where

$\mathbf{C}\mathbf{f}$ is a high pass filtered version of the image.

The idea behind the above constraint is that the highpass version of the image contains a considerably large amount of noise!

Algorithms of the above type can be handled using optimization techniques.

Constrained least squares (CLS) restoration can be formulated by choosing an $\mathbf{f}$ to minimize the Lagrangian

$$\min\left(\|\mathbf{y} - \mathbf{H}\mathbf{f}\|^2 + \alpha\|\mathbf{C}\mathbf{f}\|^2\right)$$

Typical choice for $\mathbf{C}$ is the 2-D Laplacian operator given by

$$\mathbf{C} = \begin{bmatrix} 0.00 & -0.25 & 0.00 \\ -0.25 & 1.00 & -0.25 \\ 0.00 & -0.25 & 0.00 \end{bmatrix}$$

α represents either a Lagrange multiplier or a fixed parameter known as **regularisation parameter**.

α controls the relative contribution between the term $\|\mathbf{y} - \mathbf{H}\mathbf{f}\|^2$ and the term $\|\mathbf{C}\mathbf{f}\|^2$.

The minimization of the above leads to the following estimate for the original image

$$\mathbf{f} = \left(\mathbf{H}^\mathbf{T}\mathbf{H} + \alpha\mathbf{C}^\mathbf{T}\mathbf{C}\right)^{-1}\mathbf{H}^\mathbf{T}\mathbf{y}$$

Computational issues concerning the CLS method

(I) Choice of α

The problem of the choice of α has been attempted in a large number of studies and different techniques have been proposed.

One possible choice is based on a **set theoretic approach**: a restored image is approximated by an image which lies in the intersection of the two ellipsoids defined by

$$Q_{\mathbf{f}|\mathbf{y}} = \{\mathbf{f} \mid \|\mathbf{y} - \mathbf{H}\mathbf{f}\|^2 \leq E^2\} \text{ and}$$

$$Q_\mathbf{f} = \{\mathbf{f} \mid \|\mathbf{C}\mathbf{f}\|^2 \leq \varepsilon^2\}$$

The center of one of the ellipsoids which bounds the intersection of $Q_{\mathbf{f}|\mathbf{y}}$ and $Q_\mathbf{f}$, is given by the equation

$$\mathbf{f} = \left(\mathbf{H}^{\mathbf{T}}\mathbf{H} + \alpha\mathbf{C}^{\mathbf{T}}\mathbf{C}\right)^{-1}\mathbf{H}^{\mathbf{T}}\mathbf{y}$$

with $\alpha = (E/\varepsilon)^2$.

Problem: choice of E^2 and ε^2. One choice could be

$$\alpha = \frac{1}{\text{BSNR}}$$

Comments

With larger values of α, and thus more regularisation, the restored image tends to have more ***ringing***.

With smaller values of α, the restored image tends to have more ***amplified noise effects***.

The variance and bias of the error image in frequency domain are

$$Var(\alpha) = \sigma_n^2 \sum_{u=0}^{M}\sum_{v=0}^{N} \frac{|H(u,v)|^2}{\left(|H(u,v)|^2 + \alpha|C(u,v)|^2\right)^2}$$

$$Bias(\alpha) = \sigma_n^2 \sum_{u=0}^{M-1}\sum_{v=0}^{N-1} \frac{|F(u,v)|^2 \alpha^2 |C(u,v)|^4}{\left(|H(u,v)|^2 + \alpha|C(u,v)|^2\right)^2}$$

The minimum MSE is encountered close to the intersection of the above functions.
A good choice of α is one that gives the best compromise between the variance and bias of the error image.

ITERATIVE METHODS

They refer to a large class of methods that have been investigated extensively over the last decades.

Advantages

- There is no need to explicitly implement the inverse of an operator. The restoration process is monitored as it progresses. Termination of the algorithm may take place before convergence.
- The effects of noise can be controlled in each iteration.
- The algorithms used can be spatially adaptive.
- The problem specifications are very flexible with respect to the type of degradation. Iterative techniques can be applied in cases of spatially varying or nonlinear degradations or in cases where the type of degradation is completely unknown (blind restoration).

A general formulation

In general, iterative restoration refers to any technique that attempts to minimize a function of the form

$$\Phi(\mathbf{f})$$

using an updating rule for the partially restored image.

A widely used iterative restoration method is the method of **successive approximations** where the initial estimate and the updating rule for obtaining the restored image are given by

$$\mathbf{f_0} = \mathbf{0}$$

$$\begin{aligned}\mathbf{f_{k+1}} &= \mathbf{f_k} + \beta\Phi(\mathbf{f_k}) \\ &= \Psi(\mathbf{f_k})\end{aligned}$$

Next we present possible forms of the above iterative procedure.

3. Basic iteration

$$\Phi(\mathbf{f}) = \mathbf{y} - \mathbf{Hf}$$

$$\mathbf{f_0} = \mathbf{0}$$

$$\mathbf{f_{k+1}} = \mathbf{f_k} + \beta(\mathbf{y} - \mathbf{Hf_k}) = \beta\mathbf{y} + (\mathbf{I} - \beta\mathbf{H})\mathbf{f_k}$$

4. Least squares iteration

In that case we seek for a solution that minimizes the function

$$M(\mathbf{f}) = \|\mathbf{y} - \mathbf{Hf}\|^2$$

A necessary condition for $M(\mathbf{f})$ to have a minimum is that its gradient with respect to $\mathbf{f}$ is equal to zero, which results in the normal equations

$$\mathbf{H^T H f} = \mathbf{H^T y}$$

and

$$\Phi(\mathbf{f}) = \mathbf{H^T}(\mathbf{y} - \mathbf{Hf})$$

$$\mathbf{f_0} = \beta\mathbf{H^T y}$$

$$\begin{aligned}\mathbf{f_{k+1}} &= \mathbf{f_k} + \beta\mathbf{H^T}(\mathbf{y} - \mathbf{Hf_k}) \\ &= \beta\mathbf{H^T y} + (\mathbf{I} - \beta\mathbf{H^T H})\mathbf{f_k}\end{aligned}$$

5. Constrained least squares iteration

In this method we attempt to solve the problem of constrained restoration iteratively.

As already mentioned the following functional is minimized

$$M(\mathbf{f},\alpha)=\|\mathbf{y}-\mathbf{Hf}\|^2+\alpha\|\mathbf{Cf}\|^2$$

The necessary condition for a minimum is that the gradient of $M(\mathbf{f},\alpha)$ is equal to zero. That is

$$\Phi(\mathbf{f})=\nabla_{\mathbf{f}}M(\mathbf{f},\alpha)=(\mathbf{H^TH}+\alpha\mathbf{C^TC})\mathbf{f}-\mathbf{H^Ty}$$

The initial estimate and the updating rule for obtaining the restored image are now given by

$$\mathbf{f_0}=\beta\mathbf{H^Ty}$$

$$\mathbf{f_{k+1}}=\mathbf{f_k}+\beta[\mathbf{H^Ty}-(\mathbf{H^TH}+\alpha\mathbf{C^TC})\mathbf{f_k}]$$

It can be proved that the above iteration (known as **Iterative CLS** or **Tikhonov-Miller Method**) converges if

$$0<\beta<\frac{2}{|\lambda_{\max}|}$$

where $\lambda_{\max}$ is the maximum eigenvalue of the matrix

$$(\mathbf{H^TH}+\alpha\mathbf{C^TC})$$

If the matrices $\mathbf{H}$ and $\mathbf{C}$ are block-circulant the iteration can be implemented in the frequency domain.

6. Projection onto convex sets (POCS)

The set-based approach described previously can be generalized so that any number of prior constraints can be imposed as long as the constraint sets are closed convex.

If the constraint sets have a non-empty intersection, then a solution that belongs to the intersection set can be found by the method of POCS.

Any solution in the intersection set is consistent with the a priori constraints and therefore it is a feasible solution.

Let $Q_1,Q_2,\ldots,Q_m$ be closed convex sets in a finite dimensional vector space, with $P_1,P_2,\ldots,P_m$ their respective projectors.

The iterative procedure

$$\mathbf{f_{k+1}}=P_1P_2,\ldots P_m\mathbf{f_k}$$

converges to a vector that belongs to the intersection of the sets $Q_i, i=1,2,\ldots,m$, for any starting vector $\mathbf{f_0}$.

An iteration of the form $\mathbf{f_{k+1}}=P_1P_2\mathbf{f_k}$ can be applied in the problem described previously, where we seek for an image which lies in the intersection of the two ellipsoids defined by

$$Q_{\mathbf{f|y}} = \{\mathbf{f} \mid \|\mathbf{y} - \mathbf{Hf}\|^2 \le E^2\} \text{ and } Q_{\mathbf{f}} = \{\mathbf{f} \mid \|\mathbf{Cf}\|^2 \le \varepsilon^2\}$$

The respective projections $P_1\mathbf{f}$ and $P_2\mathbf{f}$ are defined by

$$P_1\mathbf{f} = \mathbf{f} + \lambda_1\left(\mathbf{I} + \lambda_1\mathbf{H}^{\mathbf{T}}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathbf{T}}(\mathbf{y} - \mathbf{Hf})$$

$$P_2\mathbf{f} = [\mathbf{I} - \lambda_2\left(\mathbf{I} + \lambda_2\mathbf{C}^{\mathbf{T}}\mathbf{C}\right)^{-1}\mathbf{C}^{\mathbf{T}}\mathbf{C}]\mathbf{f}$$

Brief description of other advanced methods

7. Spatially adaptive iteration

The functional to be minimized takes the form

$$M(\mathbf{f},\alpha) = \|\mathbf{y} - \mathbf{Hf}\|^2_{\mathbf{W_1}} + \alpha\|\mathbf{Cf}\|^2_{\mathbf{W_2}}$$

where

$$\|\mathbf{y} - \mathbf{Hf}\|^2_{\mathbf{W_1}} = (\mathbf{y} - \mathbf{Hf})^{\mathbf{T}}\mathbf{W_1}(\mathbf{y} - \mathbf{Hf})$$

$$\|\mathbf{Cf}\|^2_{\mathbf{W_2}} = (\mathbf{Cf})^{\mathbf{T}}\mathbf{W_2}(\mathbf{Cf})$$

$\mathbf{W_1},\mathbf{W_2}$ are diagonal matrices, the choice of which can be justified in various ways. The entries in both matrices are non-negative values and less than or equal to unity.

In that case

$$\Phi(\mathbf{f}) = \nabla_{\mathbf{f}} M(\mathbf{f},\alpha) = (\mathbf{H}^{\mathbf{T}}\mathbf{W_1^T}\mathbf{W_1}\mathbf{H} + \alpha\mathbf{C}^{\mathbf{T}}\mathbf{W_2^T}\mathbf{W_2}\mathbf{C})\mathbf{f} - \mathbf{H}^{\mathbf{T}}\mathbf{W_1}\mathbf{y}$$

A more specific case is

$$M(\mathbf{f},\alpha) = \|\mathbf{y} - \mathbf{Hf}\|^2 + \alpha\|\mathbf{Cf}\|^2_{\mathbf{W}}$$

where the weighting matrix is incorporated only in the regularization term. This method is known as **weighted regularised image restoration**. The entries in matrix $\mathbf{W}$ will be chosen so that the high-pass filter is only effective in the areas of low activity and a very little smoothing takes place in the edge areas.

8. Robust functionals

Robust functionals allow for the efficient supression of a wide variety of noise processes and permit the reconstruction of sharper edges than their quadratic counterparts. We are seeking to minimize

$$M(\mathbf{f},\alpha) = R_n(\mathbf{y} - \mathbf{Hf}) + \alpha R_x\mathbf{Cf}$$

$R_n(), R_x()$ are referred to as **residual** and **stabilizing** functionals respectively.

Computational issues concerning iterative techniques

(I) Convergence

The **contraction mapping theorem** usually serves as a basis for establishing convergence of iterative algorithms.

According to it iteration

$$\mathbf{f_0} = \mathbf{0}$$

$$\mathbf{f_{k+1}} = \mathbf{f_k} + \beta\Phi(\mathbf{f_k}) = \Psi(\mathbf{f_k})$$

converges to a unique fixed point $\mathbf{f}^*$, that is, a point such that $\Psi(\mathbf{f}^*) = \mathbf{f}^*$, for any initial vector, if the operator or transformation $\Psi(\mathbf{f})$ is a **contraction**.

This means that for any two vectors $\mathbf{f_1}$ and $\mathbf{f_2}$ in the domain of $\Psi(\mathbf{f})$ the following relation holds

$$\|\Psi(\mathbf{f_1}) - \Psi(\mathbf{f_2})\| \le \eta \|\mathbf{f_1} - \mathbf{f_2}\|$$

$\eta \le 1$

$\|\cdot\|$ any norm

The above condition is **norm dependent.**

(II) Rate of convergence

The termination criterion most frequently used compares the normalized change in energy at each iteration to a threshold such as

$$\frac{\|\mathbf{f_{k+1}} - \mathbf{f_k}\|^2}{\|\mathbf{f_k}\|^2} \le 10^{-6}$$

RECURSIVE METHODS

1. Kalman filtering

Kalman is a recursive filter based on an autoregressive (AR) parametrization of the prior statistical knowledge of the image.

A global **state vector** for an image model, at pixel position (i, j) is defined as

$\underline{f}(i, j) = [f(i, j), f(i, j-1), \ldots,$

$f(i-1, j+N), f(i-1, j-M+1), \ldots, f(i-M+1, j-M+1)]^T$

The image model is then defined as

$$\underline{f}(i, j) = A\underline{f}(i, j-1) + \underline{w}(i, j)$$

$$\underline{y}(i, j) = H\,\underline{f}(i, j) + \underline{n}(i, j)$$

- the noise terms $\underline{w}(i,j)$ and $\underline{n}(i,j)$, are assumed to be white, zero-mean, Gaussian processes, with covariance matrices $R_{\underline{w}\underline{w}}$ and $R_{\underline{n}\underline{n}}$
- A is the **state transition matrix**
- H is the so called **measurement matrix**

The Kalman filter algorithm

Prediction

$$\underline{\hat{f}}^{+}(m,n) = A\underline{\hat{f}}(m,n-1)$$

$$P^{+}(m,n) = AP(m,n-1)A^{T} + R_{\underline{w}\underline{w}}$$

Update

$$K(m,n) = P^{+}(m,n)H^{T}[HP^{+}(m,n)H^{T} + R_{nn}]^{-1}$$

$$\underline{\hat{f}}(m,n) = A\underline{\hat{f}}(m,n-1) + K(m,n)[\underline{y}(m,n) - HA\underline{\hat{f}}(m,n-1)]$$

$$P(m,n) = [I - K(m,n)H]P^{+}(m,n)$$

where

$$P^{+}(m,n) = E\left\{\left(\underline{f}(m,n) - \underline{\hat{f}}^{+}(m,n)\right)\left(\underline{f}(m,n) - \underline{\hat{f}}^{+}(m,n)\right)^{T}\right\}$$

$$P(m,n) = E\left\{\left(\underline{f}(m,n) - \underline{\hat{f}}(m,n)\right)\left(\underline{f}(m,n) - \underline{\hat{f}}(m,n)\right)^{T}\right\}$$

2. Variations of the Kalman filtering

2.1 Reduced update Kalman filter (RUKF)

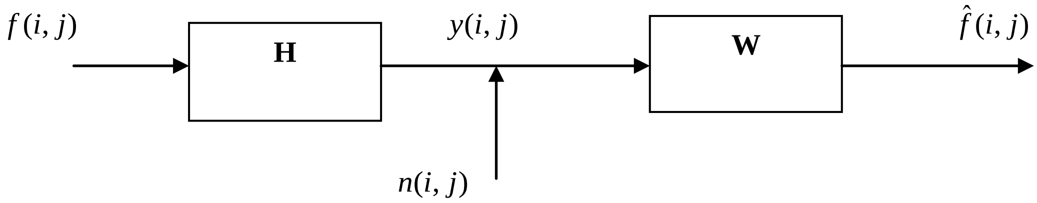
2.2 Reduced order model Kalman filter (ROMKF)

Stochastic approaches to restoration

DIRECT METHODS

1. Wiener estimator (stochastic regularisation)

The image restoration problem can be viewed as a system identification problem as follows:



The objective is to minimize the following function

$$E\{(\mathbf{f}-\hat{\mathbf{f}})^{\mathbf{T}}(\mathbf{f}-\hat{\mathbf{f}})\}$$

To do so the following conditions should hold:

(i) $E\{\hat{\mathbf{f}}\}=E\{\mathbf{f}\}\Rightarrow E\{\mathbf{f}\}=\mathbf{W}E\{\mathbf{y}\}$

(ii) the error must be orthogonal to the observation about the mean

$E\{(\hat{\mathbf{f}}-\mathbf{f})(\mathbf{y}-E\{\mathbf{y}\})^{\mathbf{T}}\}=0$

From (i) and (ii) we have that

$$E\{(\mathbf{Wy}-\mathbf{f})(\mathbf{y}-E\{\mathbf{y}\})^{\mathbf{T}}\}=0\Rightarrow E\{(\mathbf{Wy}+E\{\mathbf{f}\}-\mathbf{W}E\{\mathbf{y}\}-\mathbf{f})(\mathbf{y}-E\{\mathbf{y}\})^{\mathbf{T}}\}=0\Rightarrow$$
$$E\{[\mathbf{W}(\mathbf{y}-E\{\mathbf{y}\})-(\mathbf{f}-E\{\mathbf{f}\})](\mathbf{y}-E\{\mathbf{y}\})^{\mathbf{T}}\}=0$$

If $\tilde{\mathbf{y}}=\mathbf{y}-E\{\mathbf{y}\}$ and $\tilde{\mathbf{f}}=\mathbf{f}-E\{\mathbf{f}\}$ then

$$E\{(\mathbf{W}\tilde{\mathbf{y}}-\tilde{\mathbf{f}})\tilde{\mathbf{y}}^{\mathbf{T}}\}=0\Rightarrow E\{\mathbf{W}\tilde{\mathbf{y}}\tilde{\mathbf{y}}^{\mathbf{T}}\}=E\{\tilde{\mathbf{f}}\tilde{\mathbf{y}}^{\mathbf{T}}\}\Rightarrow \mathbf{W}E\{\tilde{\mathbf{y}}\tilde{\mathbf{y}}^{\mathbf{T}}\}=E\{\tilde{\mathbf{f}}\tilde{\mathbf{y}}^{\mathbf{T}}\}\Rightarrow \mathbf{W}\mathbf{R}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}=\mathbf{R}_{\tilde{\mathbf{f}}\tilde{\mathbf{y}}}$$

If the original and the degraded image are both zero mean then

$\mathbf{R}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}=\mathbf{R}_{\mathbf{yy}}$ and $\mathbf{R}_{\tilde{\mathbf{f}}\tilde{\mathbf{y}}}=\mathbf{R}_{\mathbf{fy}}$.

In that case we have that $\mathbf{W}\mathbf{R}_{\mathbf{yy}}=\mathbf{R}_{\mathbf{fy}}$.

If we go back to the degradation model and find the autocorrelation matrix of the degraded image then we get that

$$\mathbf{y}=\mathbf{Hf}+\mathbf{n}\Rightarrow \mathbf{y}^{\mathbf{T}}=\mathbf{f}^{\mathbf{T}}\mathbf{H}^{\mathbf{T}}+\mathbf{n}^{\mathbf{T}}$$

$E\{\mathbf{y}\mathbf{y}^{\mathbf{T}}\}=\mathbf{H}\mathbf{R}_{\mathbf{ff}}\mathbf{H}^{\mathbf{T}}+\mathbf{R}_{\mathbf{nn}}=\mathbf{R}_{\mathbf{yy}}$

$E\{\mathbf{f}\mathbf{y}^{\mathbf{T}}\}=\mathbf{R}_{\mathbf{ff}}\mathbf{H}^{\mathbf{T}}=\mathbf{R}_{\mathbf{fy}}$

From the above we get the following result

$$\mathbf{W}=\mathbf{R}_{\mathbf{fy}}\mathbf{R}_{\mathbf{yy}}^{-1}=\mathbf{R}_{\mathbf{ff}}\mathbf{H}^{\mathbf{T}}(\mathbf{H}\mathbf{R}_{\mathbf{ff}}\mathbf{H}^{\mathbf{T}}+\mathbf{R}_{\mathbf{nn}})^{-1}$$

and the estimate for the original image is

$$\hat{\mathbf{f}}=\mathbf{R}_{\mathbf{ff}}\mathbf{H}^{\mathbf{T}}(\mathbf{H}\mathbf{R}_{\mathbf{ff}}\mathbf{H}^{\mathbf{T}}+\mathbf{R}_{\mathbf{nn}})^{-1}\mathbf{y}$$

Note that knowledge of $\mathbf{R}_{\mathbf{ff}}$ and $\mathbf{R}_{\mathbf{nn}}$ is assumed.

In frequency domain

$$W(u,v)=\frac{S_{ff}(u,v)H^*(u,v)}{S_{ff}(u,v)\left|H(u,v)\right|^2+S_{nn}(u,v)}$$

$$\hat{F}(u,v)=\frac{S_{ff}(u,v)H^*(u,v)}{S_{ff}(u,v)\left|H(u,v)\right|^2+S_{nn}(u,v)}Y(u,v)$$

Computational issues

The noise variance has to be known, otherwise it is estimated from a flat region of the observed image.

In practical cases where a single copy of the degraded image is available, it is quite common to use $S_{yy}(u,v)$ as an estimate of $S_{ff}(u,v)$. **This is very often a poor estimate !**

1.1 Wiener smoothing filter

In the absence of any blur, $H(u,v)=1$ and

$$W(u,v)=\frac{S_{ff}(u,v)}{S_{ff}(u,v)+S_{nn}(u,v)}=\frac{(SNR)}{(SNR)+1}$$

(a) $(SNR)>>1\Rightarrow W(u,v)\cong 1$

(b) $(SNR)<<1\Rightarrow W(u,v)\cong(SNR)$

(SNR) is high in low spatial frequencies and low in high spatial frequencies so $W(u,v)$ can be implemented with a lowpass (smoothing) filter.

1.2 Relation with inverse filtering

If $S_{nn}(u,v)=0\Rightarrow W(u,v)=\dfrac{1}{H(u,v)}$ which is the inverse filter

If $S_{nn}(u,v)\to 0$

$$\lim_{S_{nn}\to 0}W(u,v)=\begin{cases}\dfrac{1}{H(u,v)} & H(u,v)\neq 0\\[2ex] 0 & H(u,v)=0\end{cases}$$

which is the pseudoinverse filter.

ITERATIVE METHODS

2. Iterative Wiener filters

They refer to a class of iterative procedures, that successively use the Wiener filtered signal as an improved prototype to update the covariance estimates of the original image.

Brief description of the algorithm

Step 0: Initial estimate of $\mathbf{R_{ff}}$

$$\mathbf{R_{ff}}(0) = \mathbf{R_{yy}} = E\{\mathbf{y}\mathbf{y}^{\mathbf{T}}\}$$

Step 1: Construct the i^{th} restoration filter

$$\mathbf{W}(i+1) = \mathbf{R_{ff}}(i)\mathbf{H}^{\mathbf{T}}(\mathbf{H}\mathbf{R_{ff}}(i)\mathbf{H}^{\mathbf{T}} + \mathbf{R_{nn}})^{-\mathbf{1}}$$

Step 2: Obtain the $(i+1)^{\text{th}}$ estimate of the restored image

$$\hat{\mathbf{f}}(i+1) = \mathbf{W}(i+1)\mathbf{y}$$

Step 3: Use $\hat{\mathbf{f}}(i+1)$ to compute an improved estimate of $\mathbf{R_{ff}}$ given by

$$\mathbf{R_{ff}}(i+1) = E\{\hat{\mathbf{f}}(i+1)\hat{\mathbf{f}}^{\mathbf{T}}(i+1)\}$$

Step 4: Increase i and repeat steps 1,2,3,4.

References

[1] ***Digital Image Processing*** by R. C. Gonzales and R. E. Woods, Addison-Wesley Publishing Company, 1992.

[2] ***Two-Dimensional Signal and Image Processing*** by J. S. Lim, Prentice Hall, 1990.

[3] 'Digital Image Restoration', by M.R. Banham and A.K. Katsaggelos**,** ***IEEE Signal Processing Magazine***, pp. 27-41, March 1997.