A Tutorial on Image Restoration

Peyman Milanfar*
Electrical Engineering Department
and CfAO

UCSC

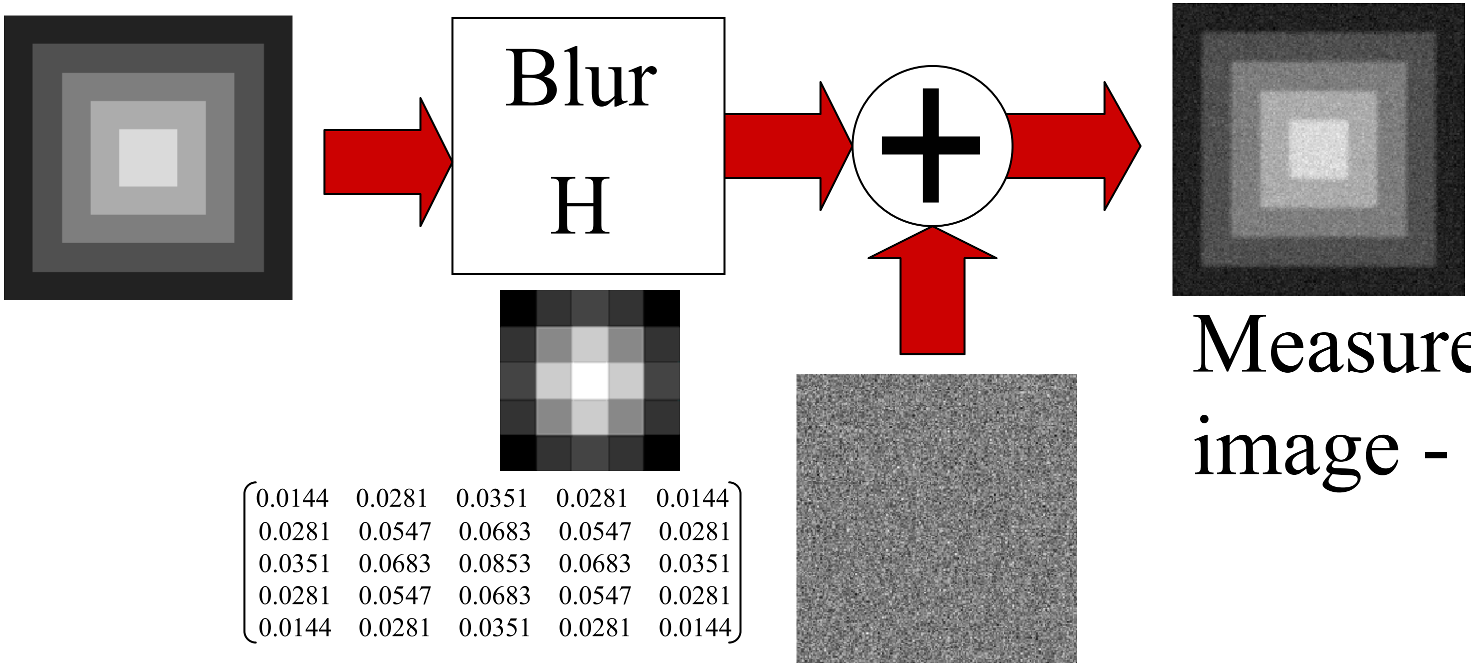
August, 2003

* Based on slides from Michael Elad

Outline

1. The image restoration problem
2. Elementary solutions
3. Statistical estimators
4. Space Varying Restoration
5. Robust Estimation
6. Non-Linear Restoration
7. What Next

The Forward Model



Original image - X

Blur

H

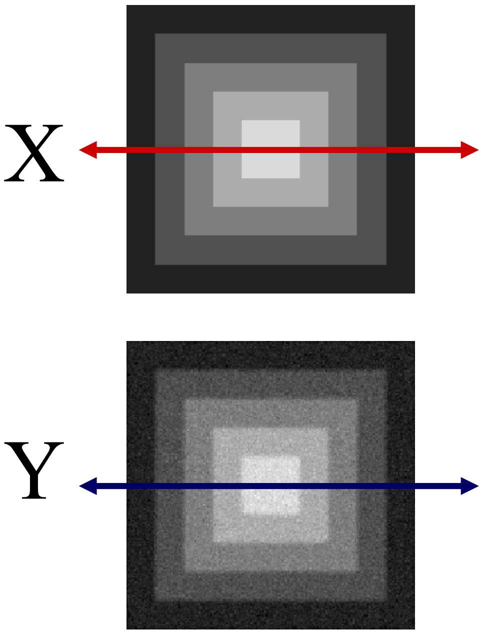
Measured image - Y

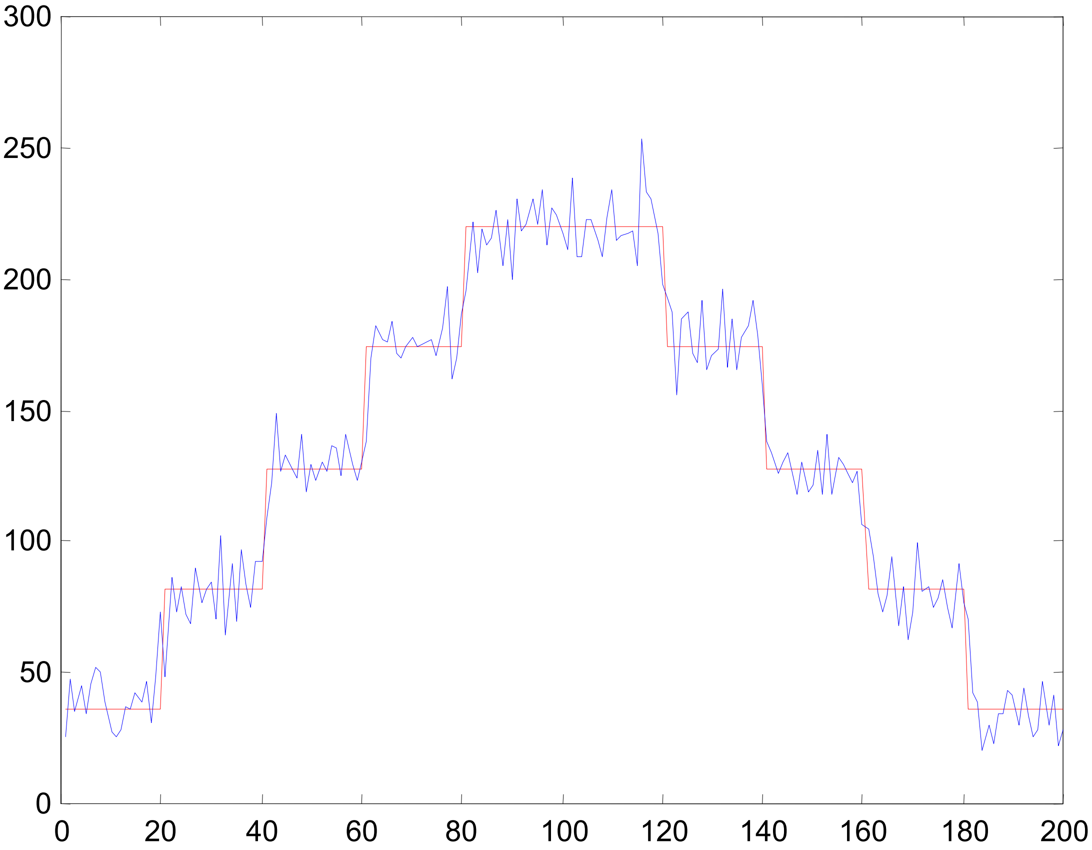
Noise - N

$$\begin{pmatrix} 0.0144 & 0.0281 & 0.0351 & 0.0281 & 0.0144 \\ 0.0281 & 0.0547 & 0.0683 & 0.0547 & 0.0281 \\ 0.0351 & 0.0683 & 0.0853 & 0.0683 & 0.0351 \\ 0.0281 & 0.0547 & 0.0683 & 0.0547 & 0.0281 \\ 0.0144 & 0.0281 & 0.0351 & 0.0281 & 0.0144 \end{pmatrix}$$

X,Y,N of size $M \times M$

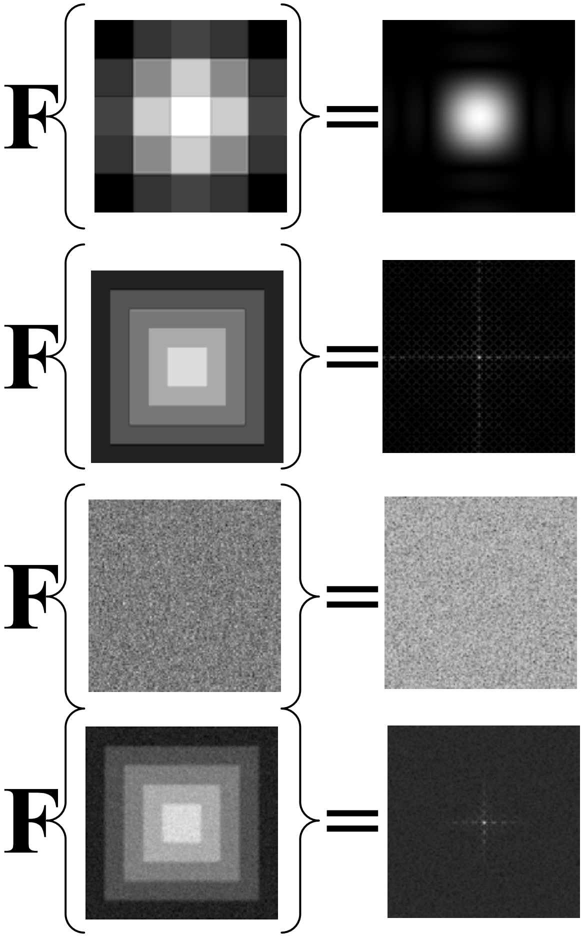
1.2 A Section





1.3 Frequency Domain

Low-pass filter



$= H_f(u,v)$

$\cdot$

$= X_f(u,v)$

$+$

$= N_f(u,v)$

$=$

$= Y_f(u,v)$

$$\boxed{Y_f = H_f X_f + N_f}$$

Point-by-point multiply

1.8 Importance

- The image restoration problem is common in imaging systems.
- Other important problems can be solved using the same tools developed here.
- The problem and its solutions form a very interesting mathematical field, called “Inverse Problems”.

Chapter 2

Elementary Solution

2.1 Basic Idea

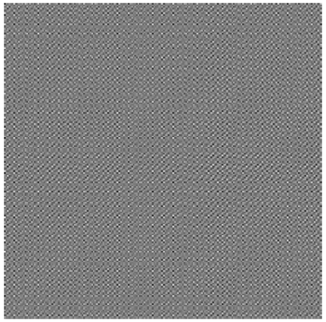
$$Y_f = H_f X_f + N_f$$

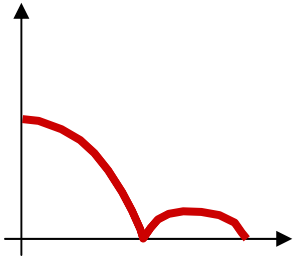
Neglect the noise

$$H_f^{-1} Y_f = X_f + \xcancel{H_f^{-1} N_f}$$

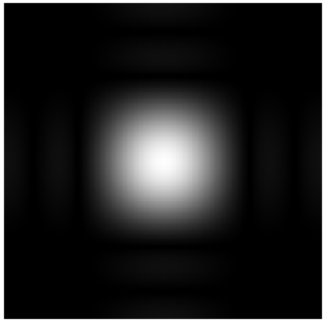
Element-by-Element inverse

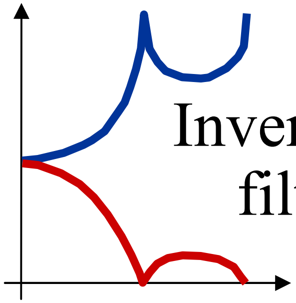
Restored



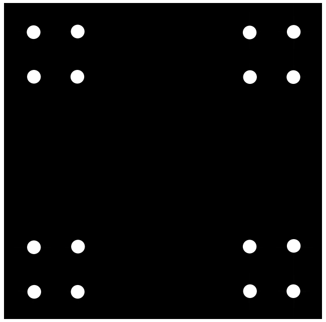


H_f





Inverse filter



2.2 Avoiding Singularities

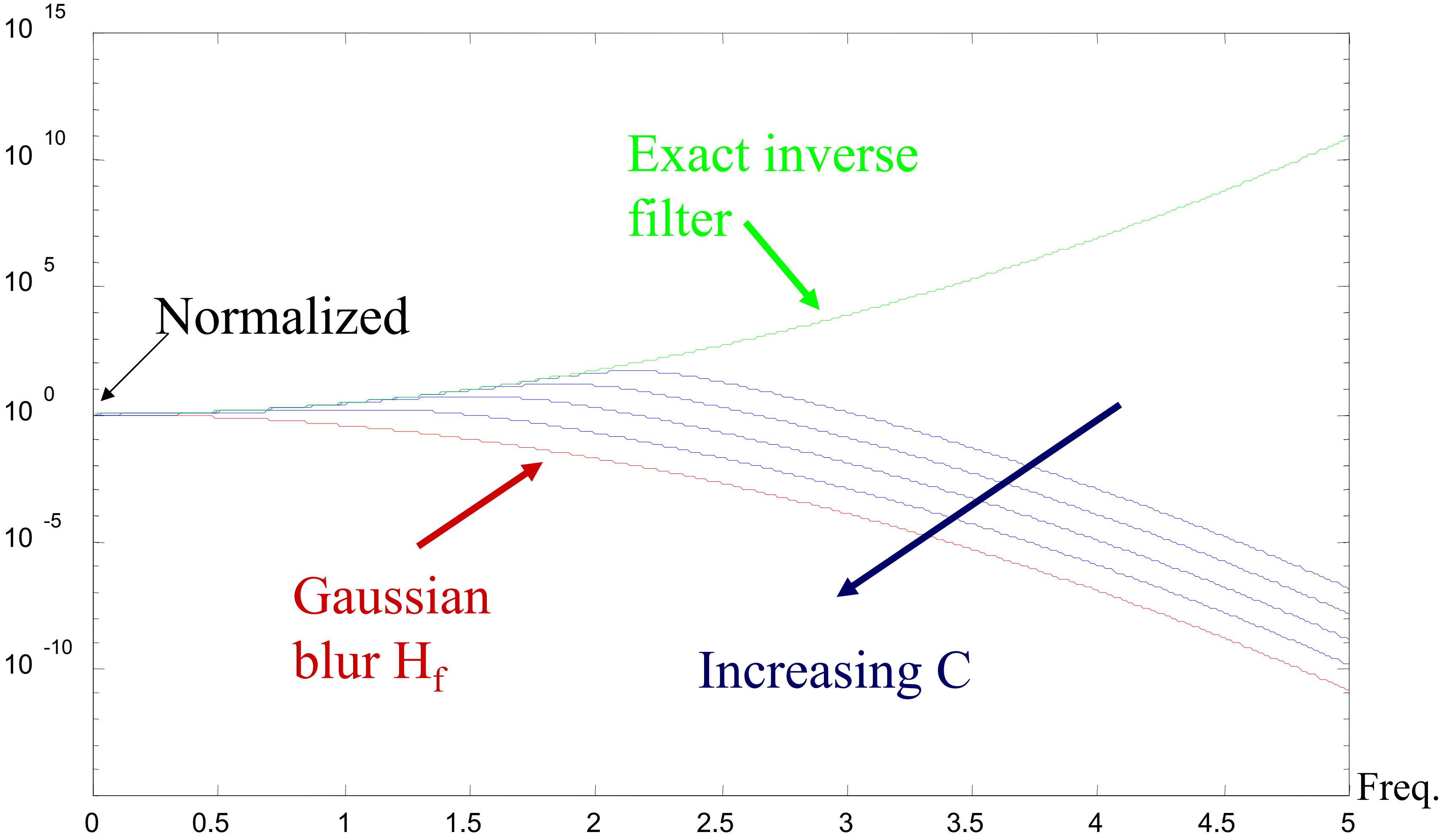
Note: If H_f has zero elements, then at those frequencies, the inverse filter does not exist. Thus use the following inverting equation instead:

$$\frac{H_f^*}{H_f^* H_f + C} \approx \begin{cases} H_f^{-1} & \text{for } \left|H_f\right|^2 >> C \\ \frac{1}{C} H_f^* & \text{for } \left|H_f\right|^2 << C \end{cases}$$

↑

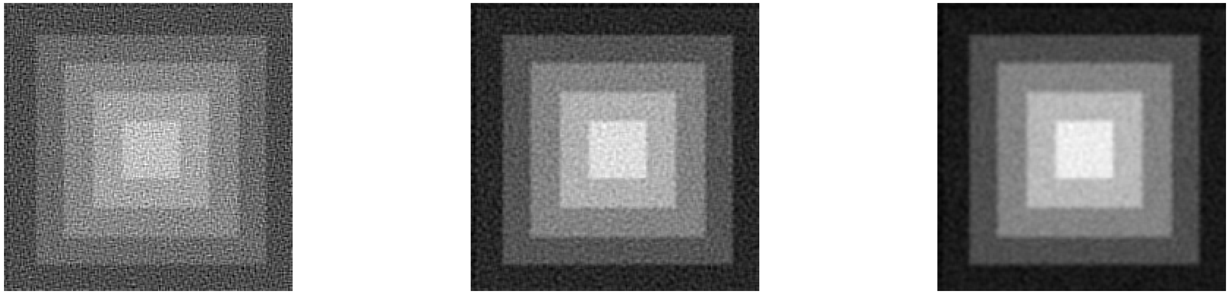
Scales image by 1/C at high frequencies

2.3 Simple Example

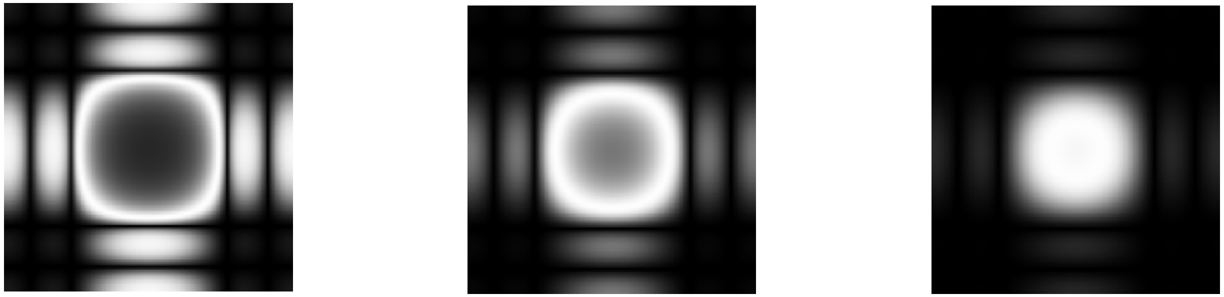


2.4 Restoration Example

Restored X :



Inverse Filter Used:

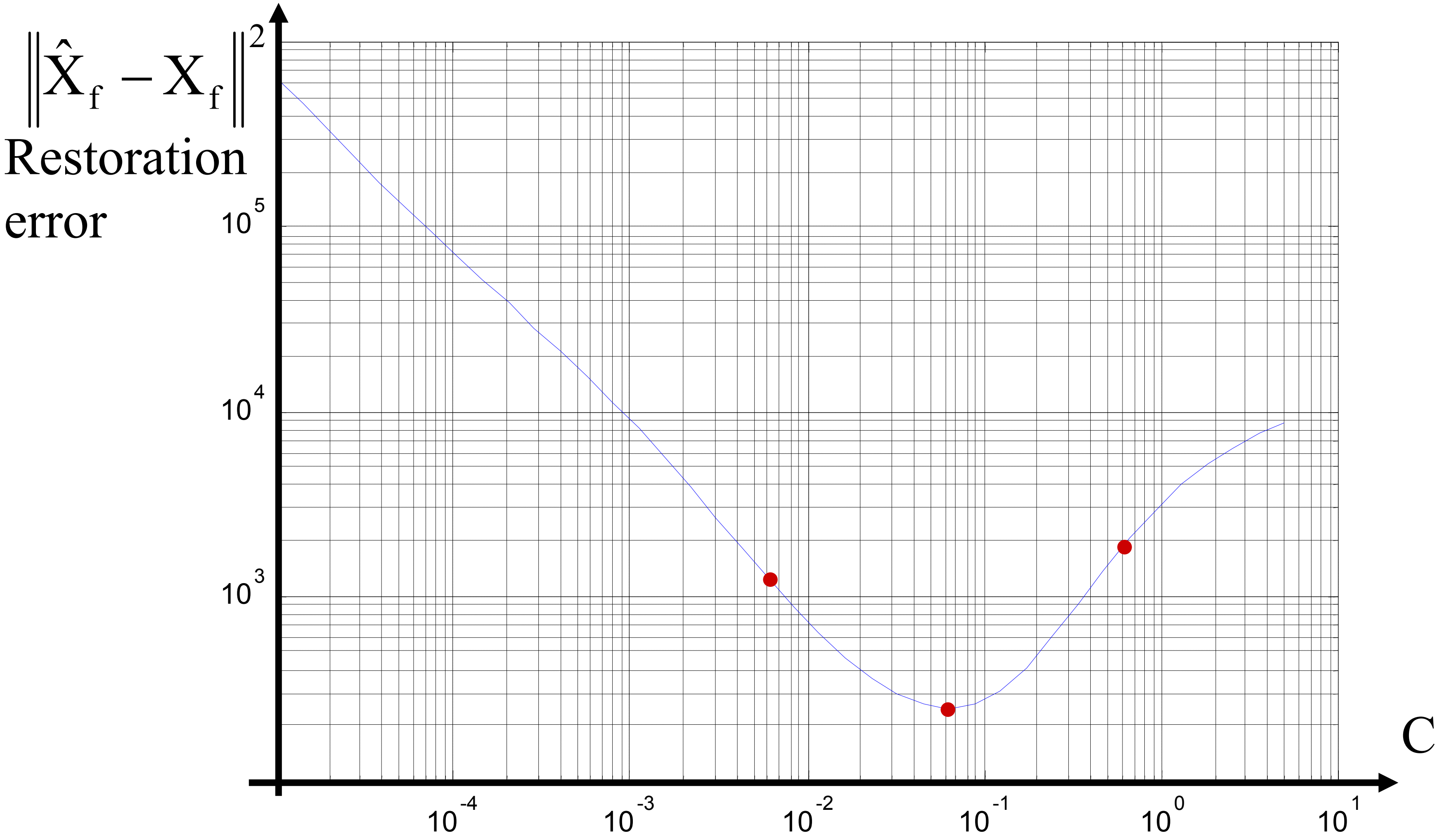


C=0.0063	C=0.063	C=0.63
Error=1264	Error=247	Error=2046

Larger C suppresses high frequency response in the inverse filter

WHAT IS THE BEST CHOICE FOR C?

2.5 The Parameter C



This is not a practical way of finding the best C !!

2.6 Drawbacks

- Finding the best C is not simple.
- In the proposed strategy, C is constant for all frequencies!! Maybe we can gain something by using a frequency dependent value.
- For high frequencies where the noise is dominant (over the signal), the image is amplified by 1/C>>1. This is silly, since we amplify mostly noise!

2.7 Change of Strategy

Previous Strategy - Based on H_f:

Where H_f is high, apply exact inverse, and where low, apply H_f/C

New Startegy - Based on [Signal/Noise]*H_f :

$|H_f|$SNR>>1 - inverse

$|H_f|$SNR<<1 - 0

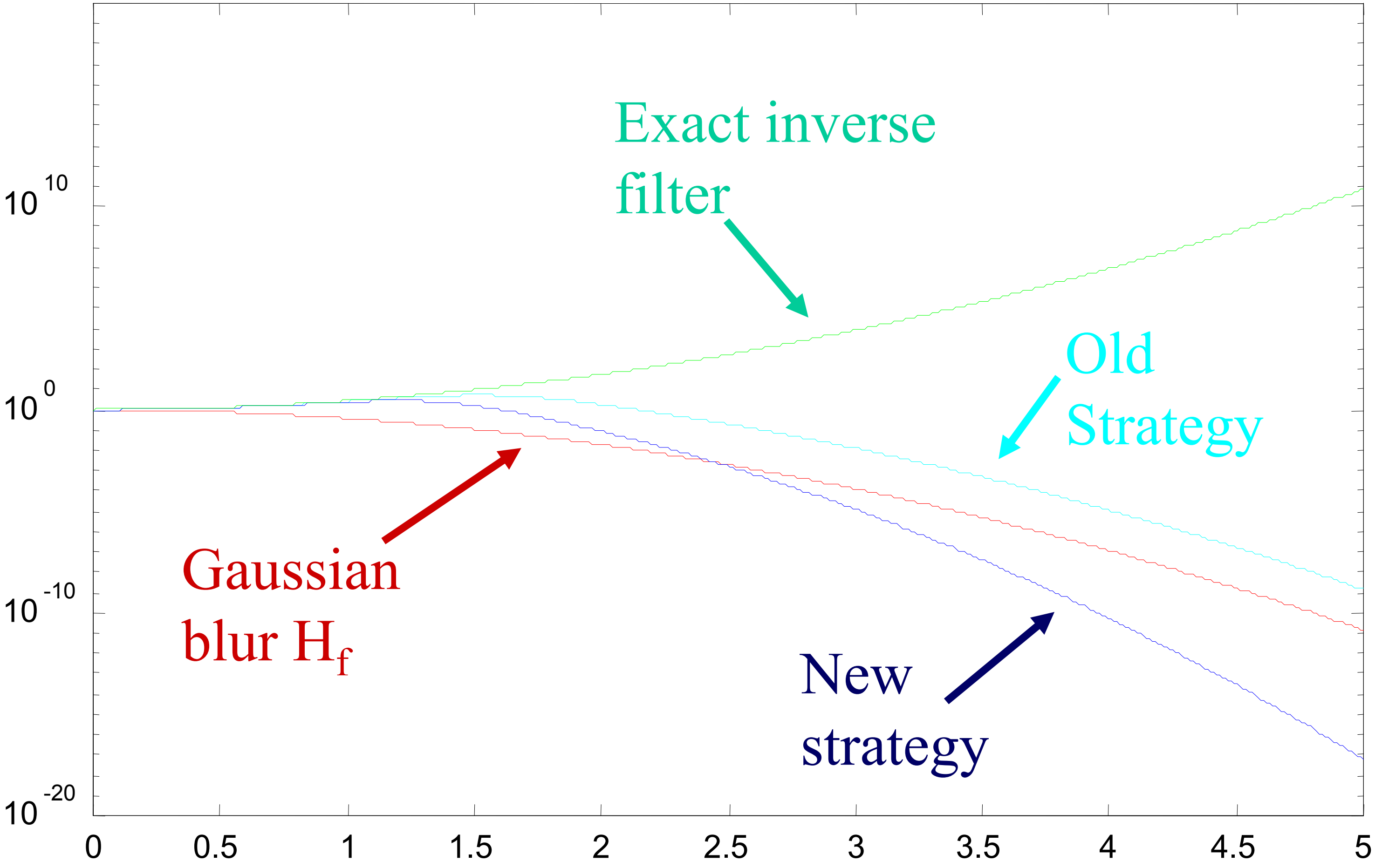
2.8 Wiener Filter

SNR is used in the following manner:

$$\frac{H_f^*}{H_f^* H_f + C \cdot SNR^{-2}} =$$

$$= \begin{cases} H_f^{-1} & |H_f|^2 \cdot SNR^2 >> C \\ \frac{SNR^2}{C} H_f^* & |H_f|^2 \cdot SNR^2 << C \end{cases}$$

2.9 Simple Example

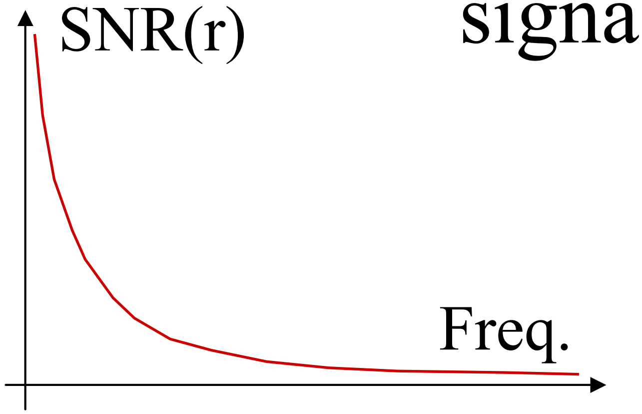


2.10 Choice of SNR

SNR is not known!

⬇

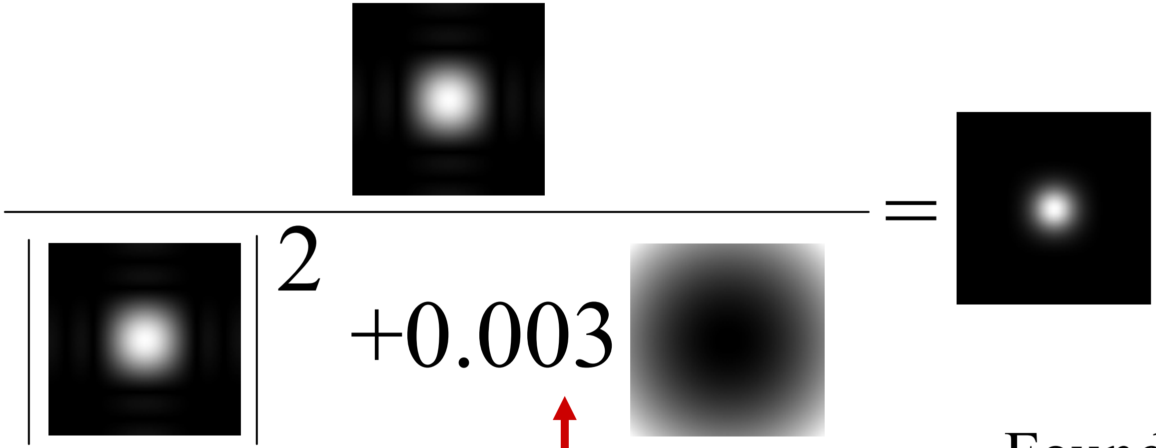
Assume white noise with unit variance & radialy decaying signal:



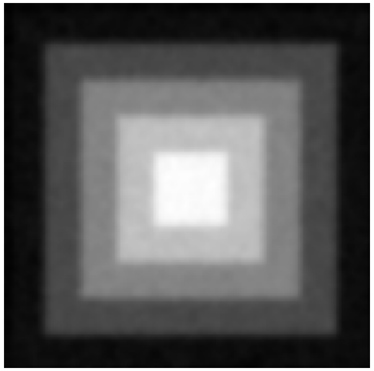
$$N^2(r,\theta)=1$$

$$S^2(r,\theta)=r^{-\rho} \qquad \rho\approx 2$$

2.11 Restoration Example

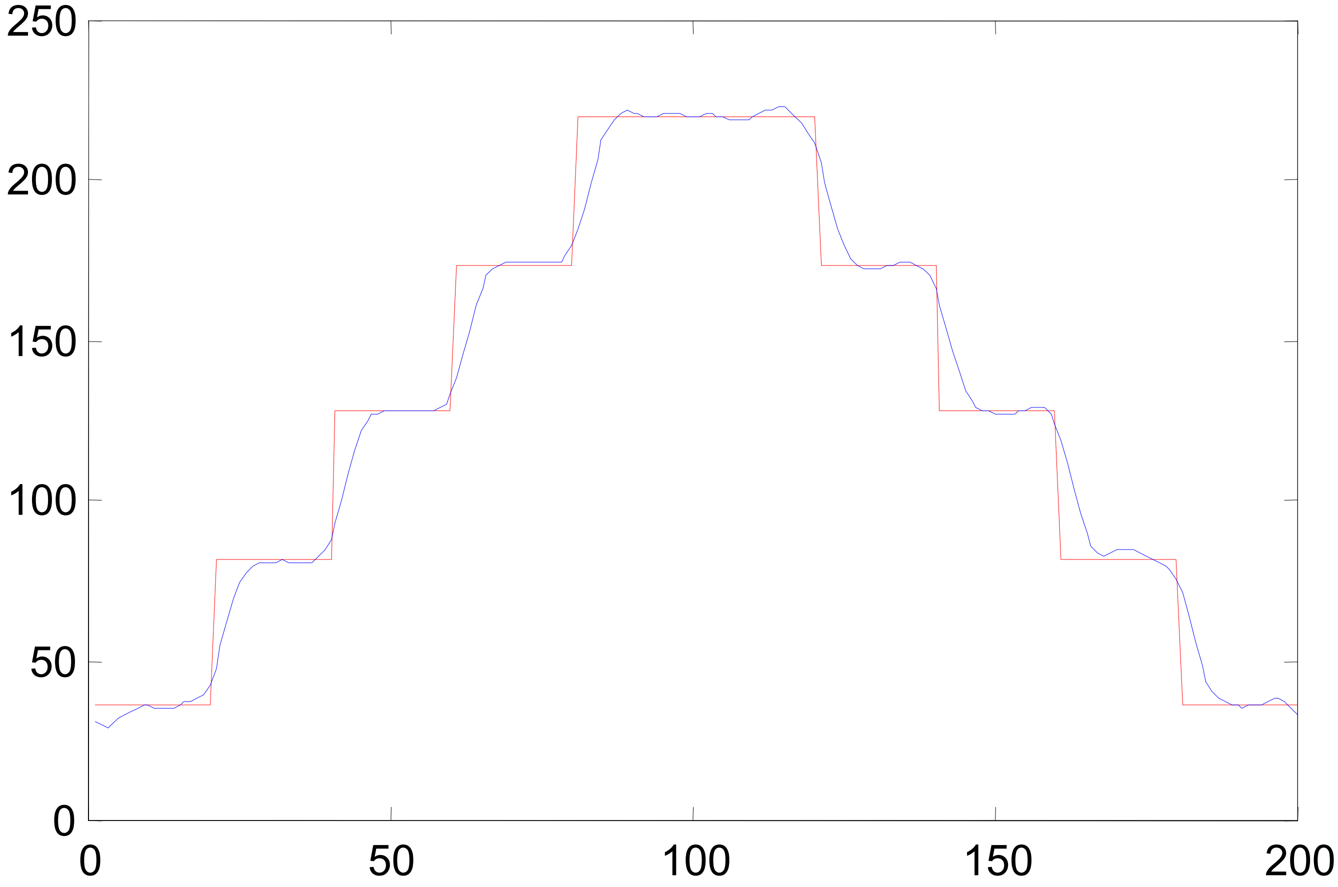


Found empirically



Error=106 (compared to the 247 in the previous approach)

2.12 Is It Enough?



NO !

Chapter 3

Statistical Estimators

1.4 Image Column-Stack

$$\mathbf{CS}\left\{\boxed{\begin{matrix}1 & 2 & 3\\ 4 & 5 & 6\\ 7 & 8 & 9\end{matrix}}\right\} = \begin{pmatrix}1\\ 4\\ 7\\ 2\\ 5\\ 8\\ 3\\ 6\\ 9\end{pmatrix}$$

Using the CS ordering, a linear operator on an image can be represented as a matrix multiplying a vector

1.5 CS for Operators

Image of size $M \times M$: a linear operation is represented as a matrix of size $M^2 \times M^2$

$$\underline{Y} = \mathbf{CS}\left\{X * h\right\} = \mathbf{H} \cdot \mathbf{CS}\left\{X\right\} = \mathbf{H}\underline{X}$$

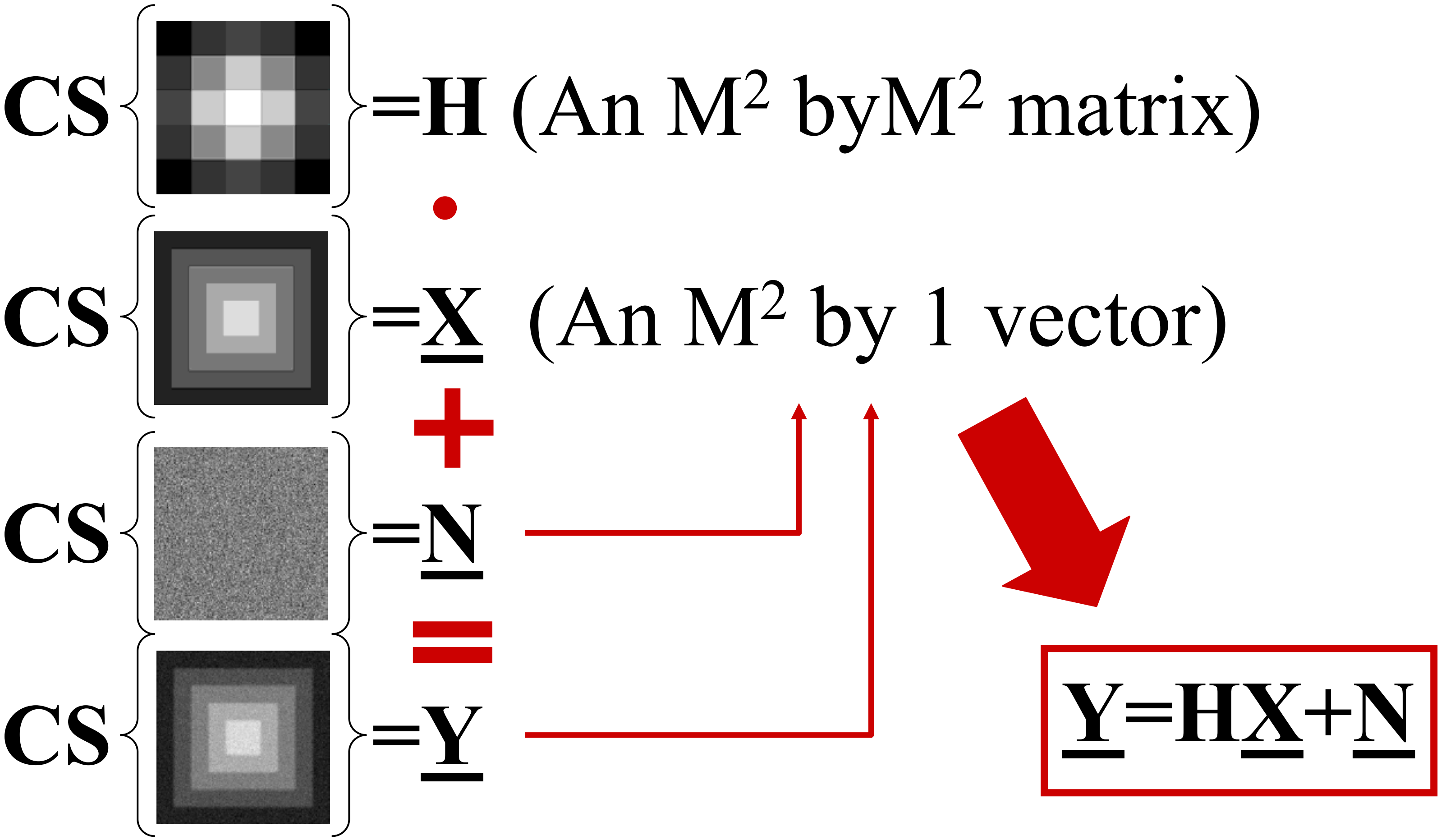
$$\mathbf{CS}\left\{\boxed{\begin{matrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{matrix}} * \boxed{\begin{matrix} 1 & 1 & 1 \end{matrix}}\right\} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 4 \\ 7 \\ 2 \\ 5 \\ 8 \\ 3 \\ 6 \\ 9 \end{pmatrix}$$

1.6 Block Circulant Form

If **H** represents a space invariant operation, it has a block circulant form

$$\begin{pmatrix} A & B & C & D & E & F & G \\ G & A & B & C & D & E & F \\ F & G & A & B & C & D & E \\ E & F & G & A & B & C & D \\ D & E & F & G & A & B & C \\ C & D & E & F & G & A & B \\ B & C & D & E & F & G & A \end{pmatrix}$$

1.7 Gray-Value Domain



CS =H (An M^2 byM^2 matrix)

•

CS =$\underline{X}$ (An M^2 by 1 vector)

+

CS =$\underline{N}$

=

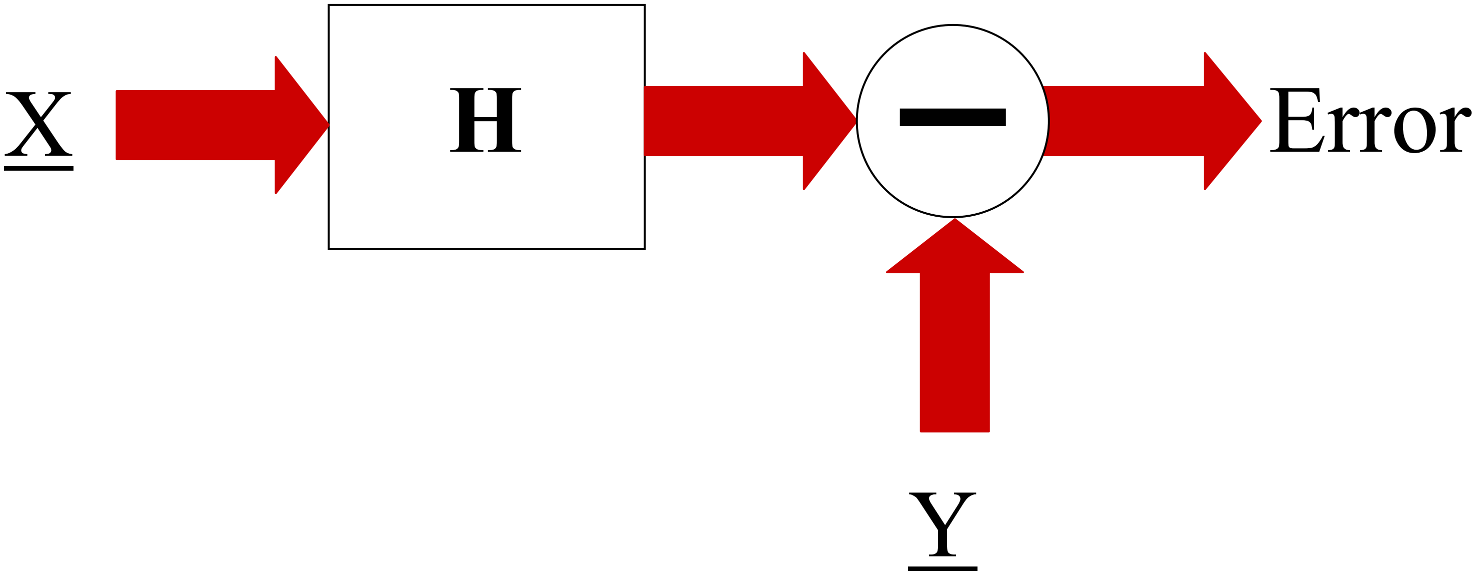
CS =$\underline{Y}$

$$\boxed{\underline{Y}=H\underline{X}+\underline{N}}$$

3.1 Maximum-Likelihood

$$\underline{Y} = \mathbf{H}\underline{X} + \underline{N} \quad \Rightarrow$$

$$\hat{\underline{X}}_{ML} = \underset{\underline{X}}{\mathrm{ArgMin}}\ \varepsilon^{2}_{ML}(\underline{X}) = \underset{\underline{X}}{\mathrm{ArgMin}} \left\| \underline{Y} - \mathbf{H}\underline{X} \right\|^{2}$$



3.2 Solving the ML

$$\underline{\hat{X}}_{ML} = \underset{\underline{X}}{\mathrm{ArgMin}}\, \varepsilon_{ML}^{2}\left(\underline{X}\right)$$

$$\Rightarrow \quad \frac{\partial \varepsilon^{2}\left(\underline{X}\right)}{\partial \underline{X}} = \mathbf{H}^{T}\left[\underline{Y} - \mathbf{H}\underline{X}\right] = 0$$

$$\Rightarrow \quad \underline{\hat{X}}_{ML} = \left[\mathbf{H}^{T}\mathbf{H}\right]^{-1}\mathbf{H}^{T}\underline{Y}$$

H square and non-singular: $\underline{\hat{X}}_{ML} = \mathbf{H}^{-1}\underline{Y}$

3.3 MAP (Regularization)

If **H** is singular, we can apply the following equation instead:

$$\underline{\hat{X}}_{\mathrm{MAP}} = \left[\mathbf{H}^{T}\mathbf{H} + \lambda \cdot \mathbf{I}\right]^{-1} \mathbf{H}^{T} \underline{Y}$$

Small constant

Identity matrix

$$\varepsilon^{2}_{\mathrm{MAP}}\left(\underline{X}\right) = \left\|\underline{Y} - \mathbf{H}\underline{X}\right\|^{2} + \lambda\left\|\underline{X}\right\|^{2}$$

i.e., prefer solutions with smaller norm

3.4 Better MAP Prior

$$\varepsilon^2_{\mathrm{MAP}}\left(\underline{X}\right)=\left\|\underline{Y}-\mathbf{H}\underline{X}\right\|^2+\lambda\left\|\mathbf{D}\underline{X}\right\|^2$$

Among the solutions, find those who are relatively smooth

(**D** - Laplacian derivative $\boxed{\begin{array}{ccc}0 & 1 & 0\\ 1 & -4 & 1\\ 0 & 1 & 0\end{array}}$)

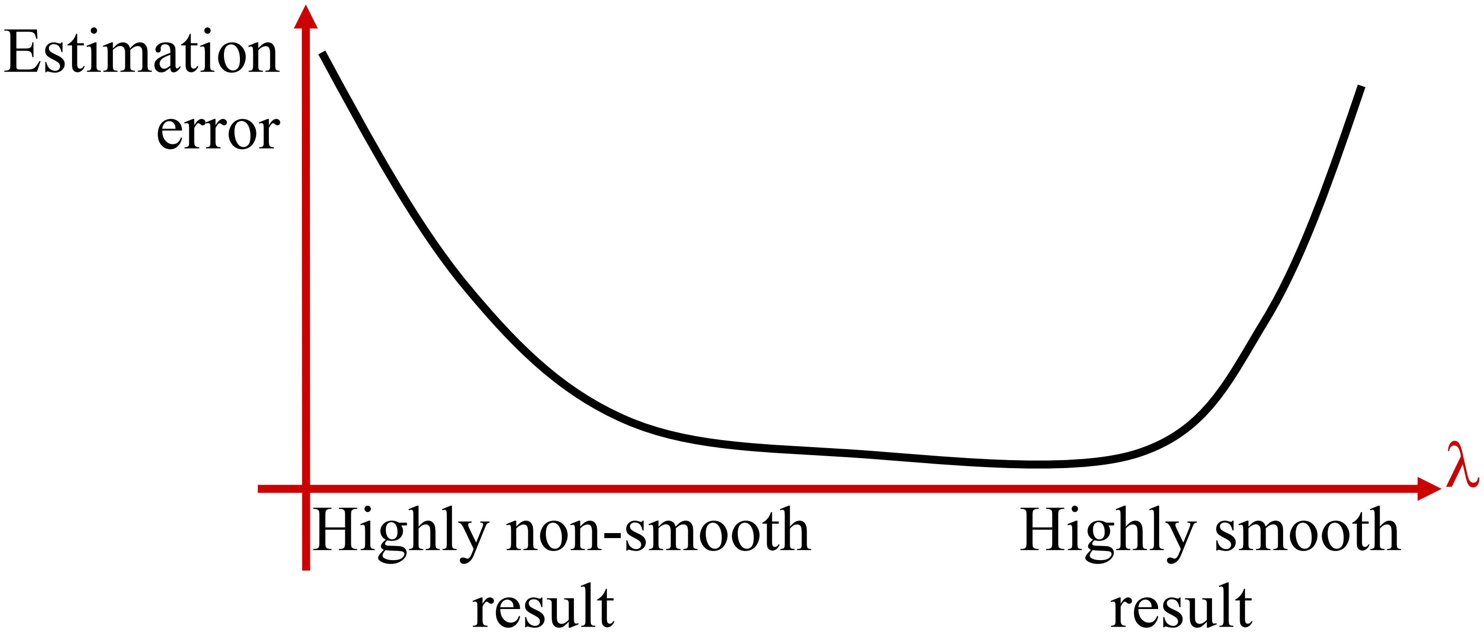
⬇

$$\underline{\hat{X}}_{\mathrm{MAP}}=\left[\mathbf{H}^T\mathbf{H}+\lambda\cdot\mathbf{D}^T\mathbf{D}\right]^{-1}\mathbf{H}^T\underline{Y}$$

3.5 Choice of Lambda

$$\varepsilon^2_{\mathrm{MAP}}\left(\underline{\mathrm{X}}\right)=\left\|\underline{\mathrm{Y}}-\mathbf{H}\underline{\mathrm{X}}\right\|^2+\lambda\left\|\mathbf{D}\underline{\mathrm{X}}\right\|^2$$

Measurement Error   Smoothness Error



3.6 Interpretations?

$$\underline{\hat{X}}_{\mathrm{ML}} = \mathbf{H}^{-1}\underline{Y}$$

→ Inverse filter

⇓

$$\underline{\hat{X}}_{\mathrm{MAP}} = \left[\mathbf{H}^{T}\mathbf{H} + \lambda \cdot \mathbf{I}\right]^{-1}\mathbf{H}^{T}\underline{Y}$$

→ Heuristic approach

⇓

$$\underline{\hat{X}}_{\mathrm{MAP}} = \left[\mathbf{H}^{T}\mathbf{H} + \lambda \cdot \mathbf{D}^{T}\mathbf{D}\right]^{-1}\mathbf{H}^{T}\underline{Y}$$

→ ≈Wiener filter

3.7 Generalizations?

$$\underline{\hat{X}}_{\mathrm{MAP}} = \boxed{\left[\mathbf{H}^{T}\mathbf{H} + \lambda \cdot \mathbf{D}^{T}\mathbf{D}\right]^{-1}\mathbf{H}^{T}}\underline{Y} = \mathbf{R}\underline{Y}$$

R represents a Linear Space Invariant filter (having a block Circulant structure)

Option 1

Extend it to become space varying

Option 2

Extend it to become non-linear

Chapter 4

Space Varying Restoration

4.1 The Basic Idea

$$\varepsilon^2_{\mathrm{MAP}}\left(\underline{X}\right)=\left\|\underline{Y}-\mathbf{H}\underline{X}\right\|^2+\lambda\left\|\mathbf{D}\underline{X}\right\|^2$$

- The expression $\left\|\mathbf{D}\underline{X}\right\|^2$ forces smoothness all over the image, including on edges!
- The basic idea is to force smoothness in a manner that is spatially adaptive, so that only truly smooth regions are forced to become smooth

4.2 Weighted Smoothness

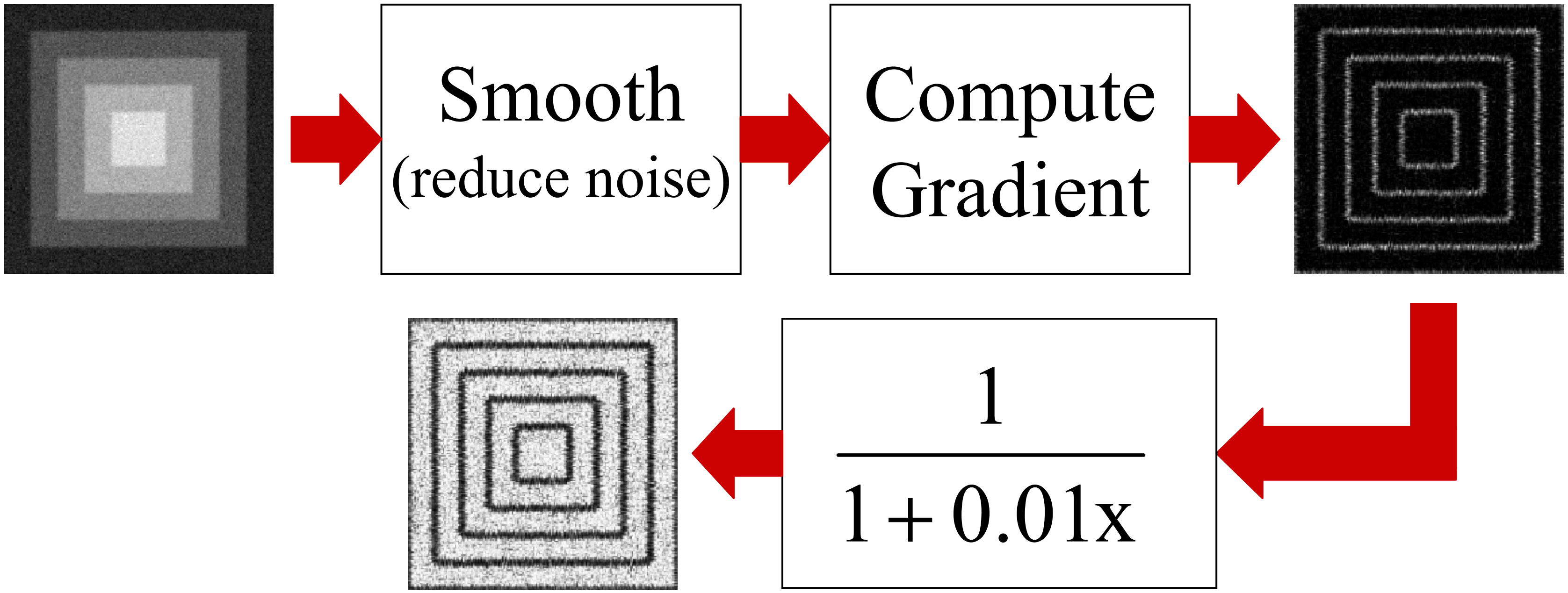
$$\varepsilon^2_{\mathrm{MAP}}(\underline{\mathrm{X}}) = \left\|\underline{\mathrm{Y}} - \mathbf{H}\underline{\mathrm{X}}\right\|^2 + \lambda(\mathbf{D}\underline{\mathrm{X}})^{\mathrm{T}}\mathbf{W}(\mathbf{D}\underline{\mathrm{X}})$$

Where W is a diagonal weight matrix, forcing smoothness with different strengths at different locations in the image

⬇

$$\underline{\hat{\mathrm{X}}}_{\mathrm{MAP}} = \left[\mathbf{H}^{\mathrm{T}}\mathbf{H} + \lambda\cdot\mathbf{D}^{\mathrm{T}}\mathbf{W}\mathbf{D}\right]^{-1}\mathbf{H}^{\mathrm{T}}\underline{\mathrm{Y}}$$

4.3 Ex.- How to Build W



Smooth
(reduce noise)

Compute
Gradient

$$\frac{1}{1+0.01x}$$

In the smooth areas, w(i,j) is close to 1

On the edge areas, w(i,j) is close to zero

4.4 Choice of Lambda

$$\varepsilon^2_{\mathrm{MAP}}\left(\underline{\mathrm{X}}\right)=\left\|\underline{\mathrm{Y}}-\mathbf{H}\underline{\mathrm{X}}\right\|^2+\lambda\left(\mathbf{D}\underline{\mathrm{X}}\right)^{\mathrm{T}}\mathbf{W}\left(\mathbf{D}\underline{\mathrm{X}}\right)$$

If w(i,j) in the interval [0,1], we can use higher value of λ :

Smooth regions - stronger smoothing

Non-Smooth regions - no smoothing

4.5 Numerical Solution

Minimizing the following function

$$\varepsilon^2_{\mathrm{MAP}}\left(\underline{X}\right)=\left\|\underline{Y}-\mathbf{H}\underline{X}\right\|^2+\lambda\left(\mathbf{D}\underline{X}\right)^{\mathrm{T}}\mathbf{W}\left(\mathbf{D}\underline{X}\right)$$

Results in

$$\hat{\underline{X}}_{\mathrm{MAP}}=\boxed{\left[\mathbf{H}^{\mathrm{T}}\mathbf{H}+\lambda\cdot\mathbf{D}^{\mathrm{T}}\mathbf{W}\mathbf{D}\right]}^{-1}\mathbf{H}^{\mathrm{T}}\underline{Y}$$

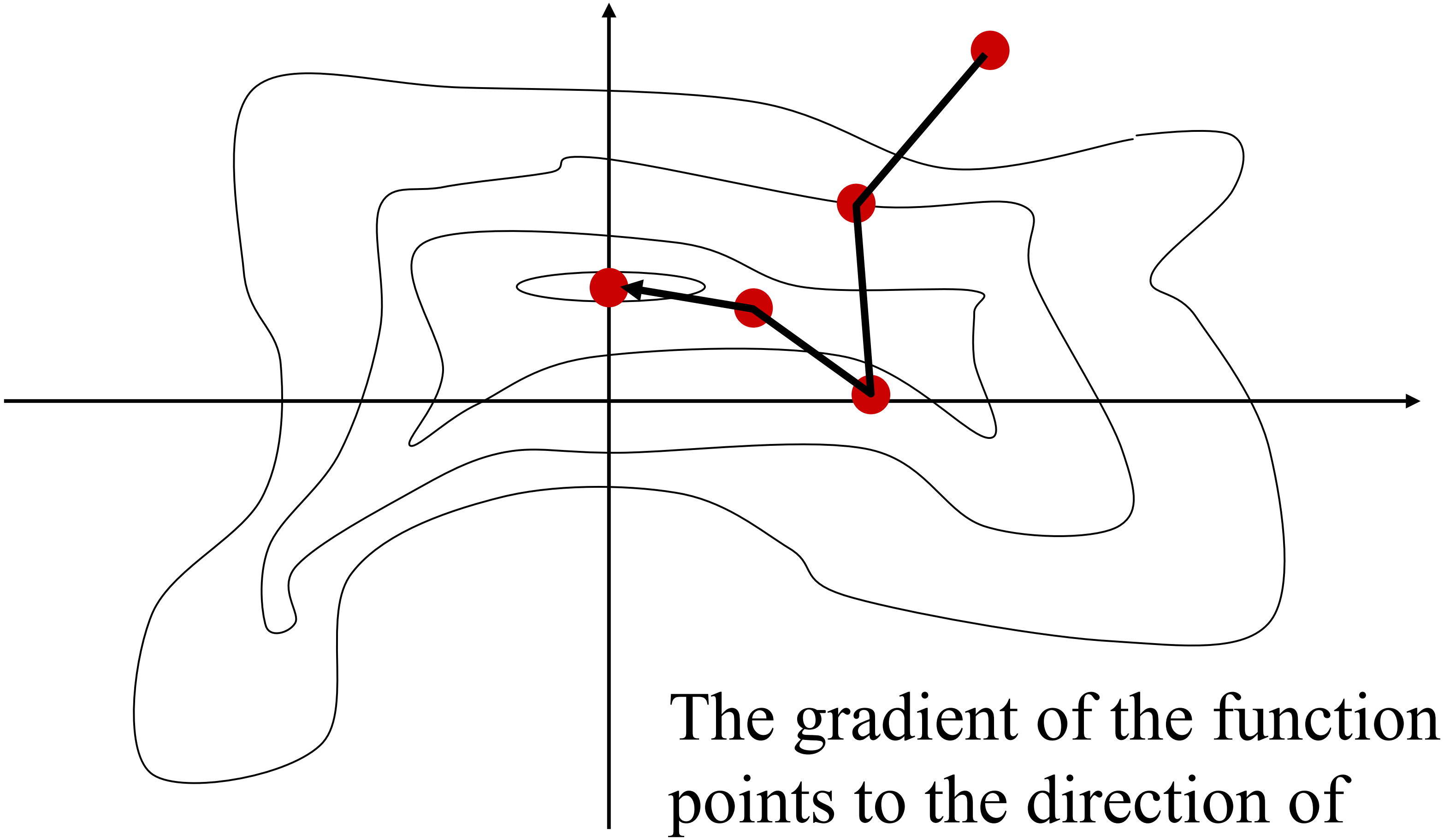
This matrix is huge! For a 1000 by 1000 image X, its size is 1e6 by 1e6 entries

4.6 SD Algorithm

The Steepest Descent algorithm - Minimization of f($\underline{X}$) is done by the following iterative procedure:

- Initialize by some $\underline{X}_0$ (the closer it is to the solution the faster the algorithm converges)
- Iterate by $\underline{X}_{k+1} = \underline{X}_k - \mu \left. \frac{\partial f(\underline{X})}{\partial \underline{X}} \right|_{\underline{X}_k}$, where μ should be small for convergence.

4.7 Basic Idea



The gradient of the function points to the direction of maximal slope

4.8 MAP Solution

$$\varepsilon^2_{\mathrm{MAP}}\left(\underline{\mathrm{X}}\right)=\left\|\underline{\mathrm{Y}}-\mathbf{H}\underline{\mathrm{X}}\right\|^2+\lambda\left(\mathbf{D}\underline{\mathrm{X}}\right)^{\mathrm{T}}\mathbf{W}\left(\mathbf{D}\underline{\mathrm{X}}\right)$$

$$\Downarrow$$

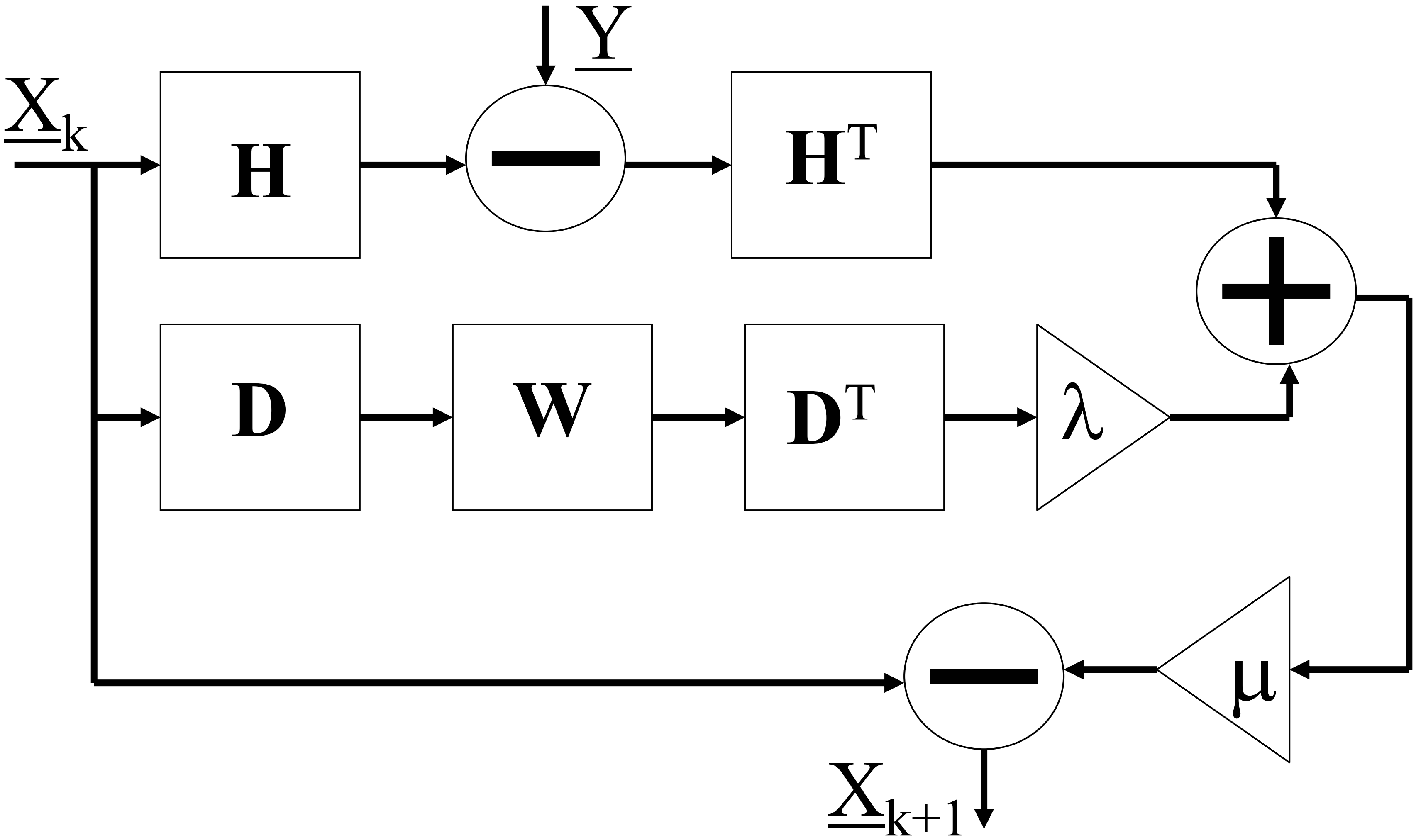
$$\frac{\partial\varepsilon^2_{\mathrm{MAP}}\left(\underline{\mathrm{X}}\right)}{\partial\underline{\mathrm{X}}}=2\mathbf{H}^{\mathrm{T}}\left[\mathbf{H}\underline{\mathrm{X}}-\underline{\mathrm{Y}}\right]+2\lambda\mathbf{D}^{\mathrm{T}}\mathbf{W}\mathbf{D}\underline{\mathrm{X}}$$

$$\Downarrow$$

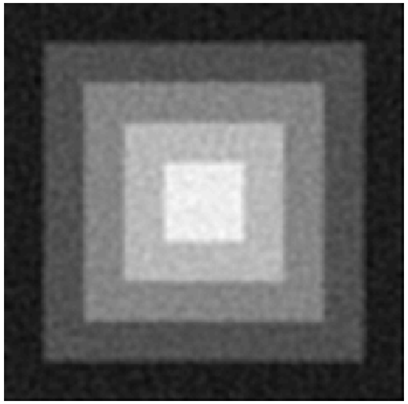
$$\underline{\mathrm{X}}_{\mathrm{k+1}}=\underline{\mathrm{X}}_{\mathrm{k}}-\mu\left\{\mathbf{H}^{\mathrm{T}}\left[\mathbf{H}\underline{\mathrm{X}}_{\mathrm{k}}-\underline{\mathrm{Y}}\right]-\lambda\mathbf{D}^{\mathrm{T}}\mathbf{W}\mathbf{D}\underline{\mathrm{X}}_{\mathrm{k}}\right\}$$

$\underline{\mathrm{X}}_0$ can be chosen to be $\underline{\mathrm{Y}}$

4.9 Restoration System

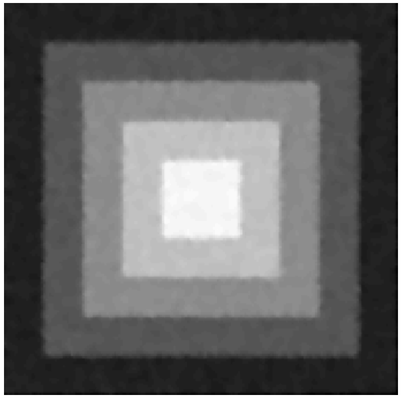


4.10 Example



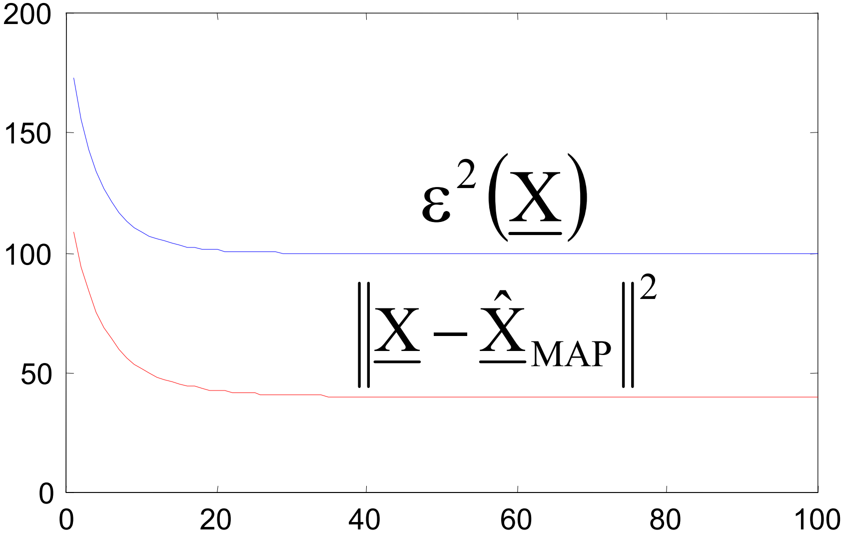
No weight

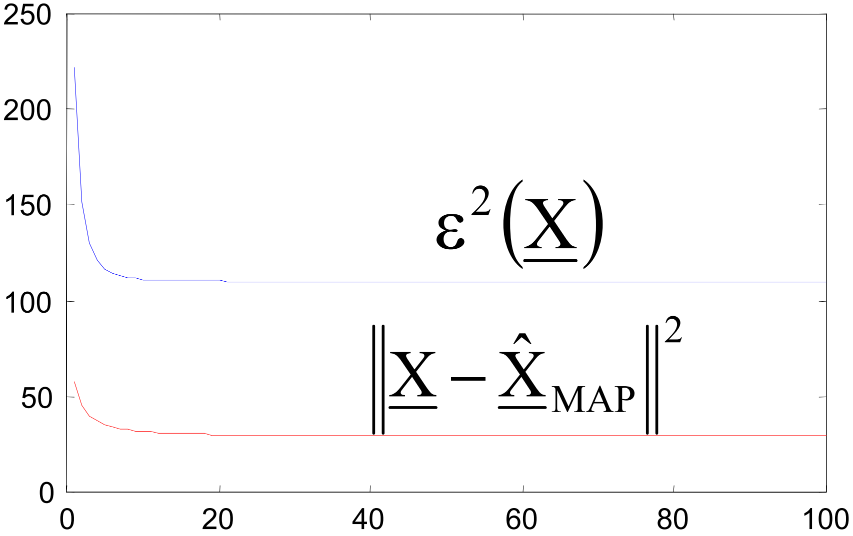
Error=40.7



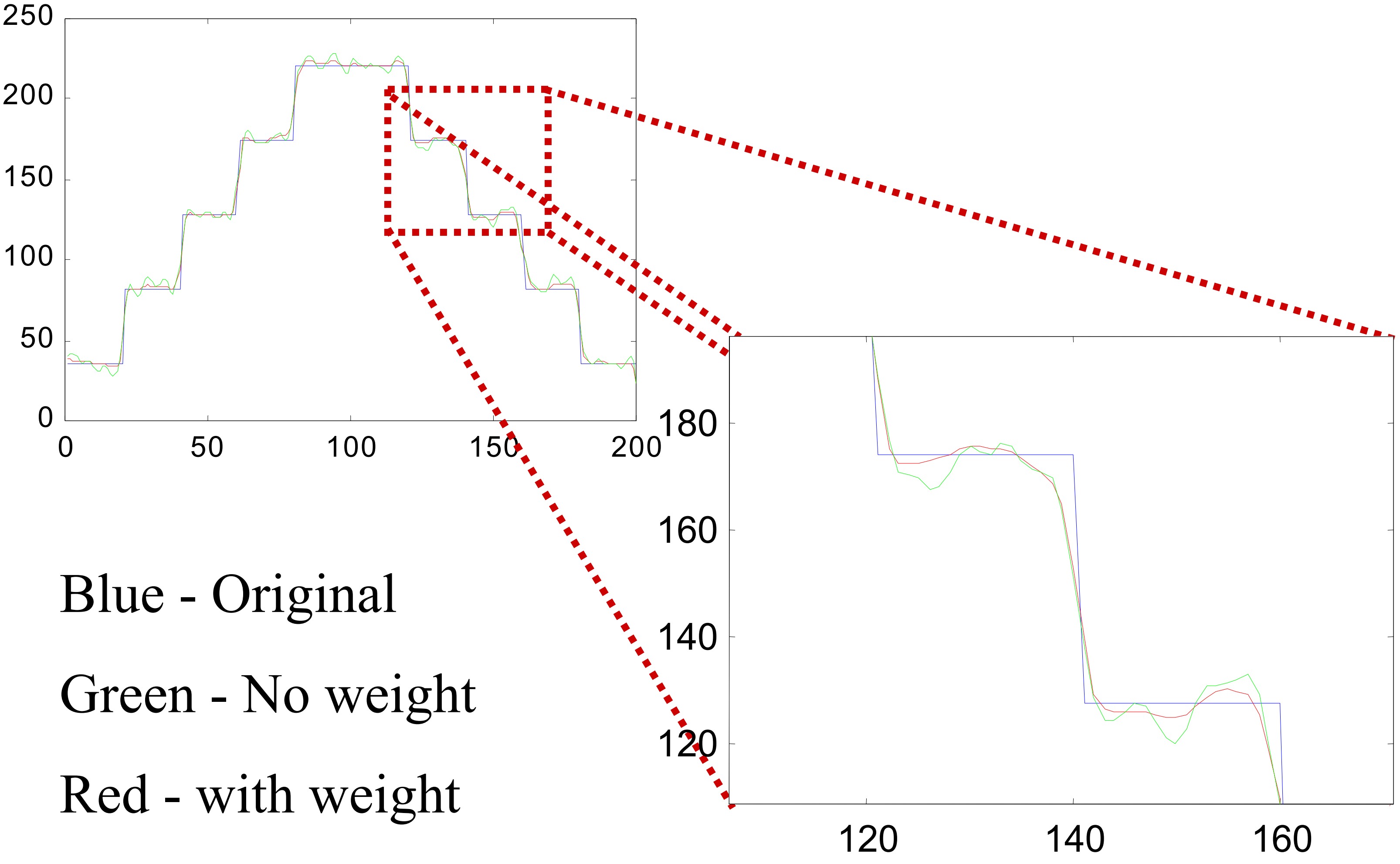
With weight

Error=29.2





4.11 Result Behavior



Blue - Original

Green - No weight

Red - with weight

Chapter 5

Robust Estimation

5.1 Mean Estimation

Assume that a physical value X is unknown. L measurements of it are given:

$$\left\{x_k\right\}_{k=1}^{L}$$

In order to estimate X we can propose the Least Squares approach:

$$\hat{X} = \underset{X}{\mathrm{ArgMin}} \sum_{k=1}^{L} \left(X - x_k\right)^2 = \frac{1}{L} \sum_{k=1}^{L} x_k$$

5.2 Outliers

Lets assume that one measurement (the first) was extremely bad and gave

$$x_1 = X + 1e10$$

In this case, the estimation will be biased by 1e10/L

The estimation approach is too sensitive to outliers

5.3 Robust Estimation

Instead $$\hat{X} = \underset{X}{\text{ArgMin}} \sum_{k=1}^{L} \left(X - x_k \right)^2$$

Use $$\hat{X} = \underset{X}{\text{ArgMin}} \sum_{k=1}^{L} \rho\left\{ X - x_k \right\}$$

- Where ρ is non-negative, symmetric, and monotonically increasing function.
- For $\rho(x)=x^2$, we have the original least-squares equation

5.4 Example 1

For ρ(x)=|x| we get

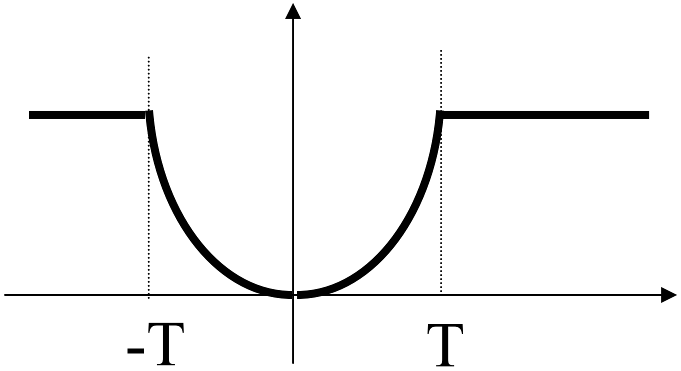
$$\hat{X} = \underset{X}{\text{ArgMin}} \sum_{k=1}^{L} \left| X - x_k \right| \Rightarrow \sum_{k=1}^{L} \text{sign}\left\{X - x_k\right\} = 0$$

$$\Rightarrow \hat{X} = \text{Median}\left\{x_1, x_2, \ldots, x_L\right\}$$

In this case, if one of the measurements is "going wild", there is almost no impact on the result

5.4 Example 2

$$\rho(x)=\begin{cases} x^2 & |x|\le T \\ T^2 & |x|>T \end{cases}$$



- For measurements falling in [-T,T], the behavior is as in the classic $\rho(x)=x^2$ case
- Outliers ($|x|>T$), do not impact the estimation at all!

5.5 Summary

- In the robust estimation paradigm, we are not restricting ourselves to $\rho(x)=x^2$, but rather using other $\rho(x)$ functions
- By using less-convex functions, we get better treatment for outliers
- The problem with the resulting approach is that it might yield non-convex optimization problems

Chapter 6

Robust Smoothness

6.1 The Basic Idea

Instead of using

$$\varepsilon_{\mathrm{MAP}}^{2}\left(\underline{\mathrm{X}}\right)=\left\|\underline{\mathrm{Y}}-\mathbf{H}\underline{\mathrm{X}}\right\|^{2}+\lambda\left\|\mathbf{D}\underline{\mathrm{X}}\right\|^{2}$$

Let us use

$$\varepsilon_{\mathrm{MAP}}^{2}\left(\underline{\mathrm{X}}\right)=\left\|\underline{\mathrm{Y}}-\mathbf{H}\underline{\mathrm{X}}\right\|^{2}+\lambda\rho\left\{\mathbf{D}\underline{\mathrm{X}}\right\}$$

where $\rho(\mathbf{D}\underline{\mathrm{X}})$ stands for applying $\rho(x)$ per each entry and summing all the results

6.2 A simple 1D Example

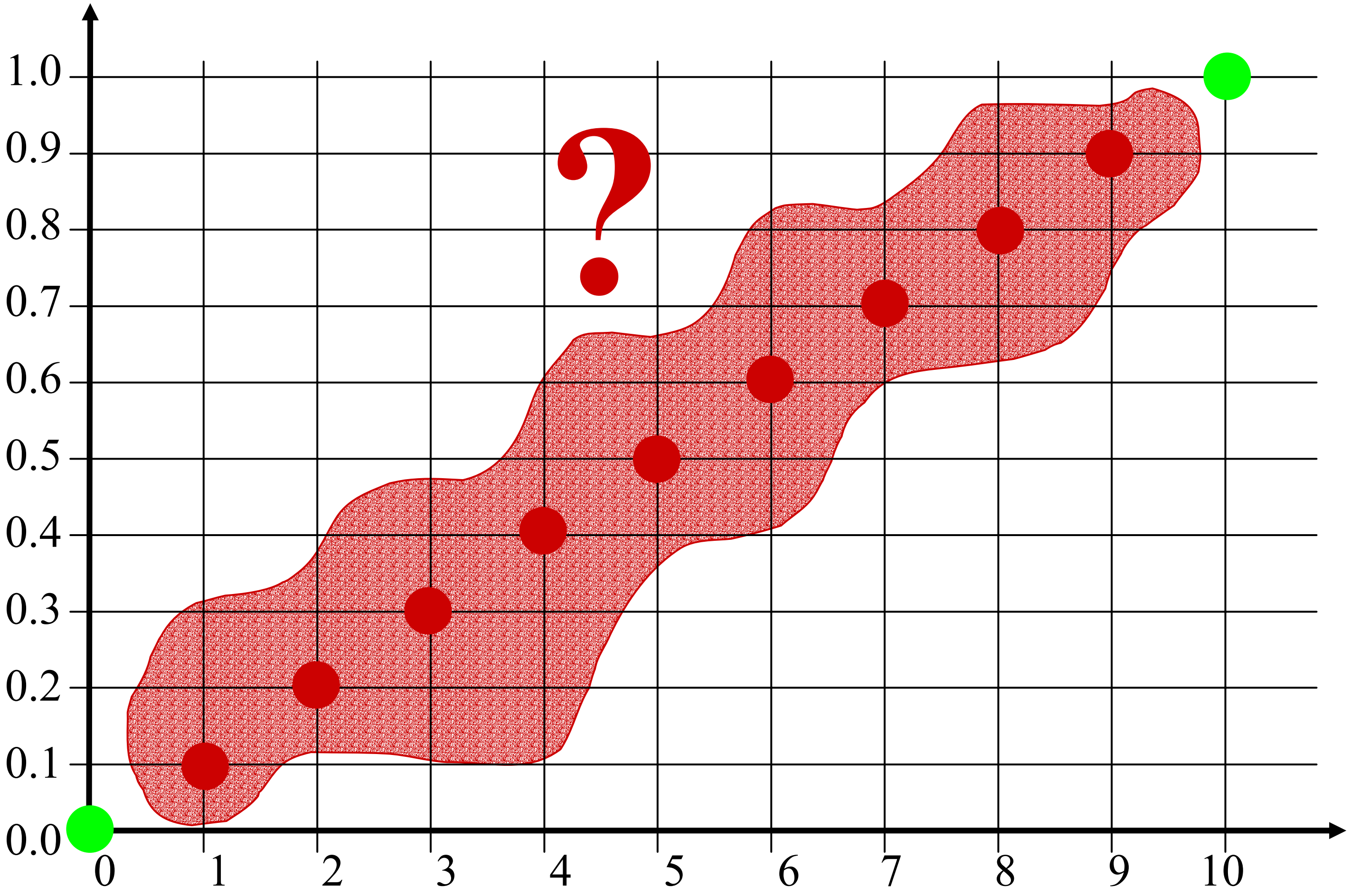
Assume that for a signal x we know that

$$x_0 = 0, \ \ x_{10} = 1$$

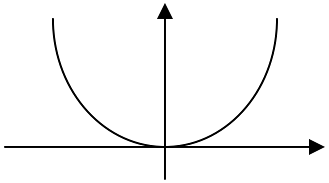
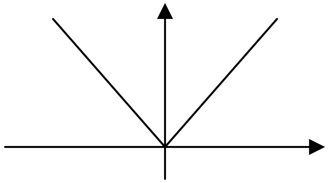
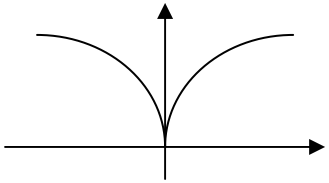
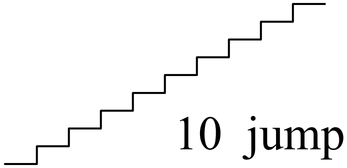
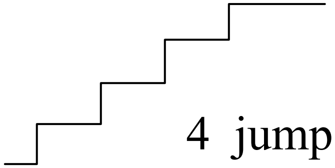
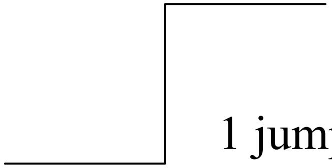
and we would like to find x_2, x_3, … , x_9, such that the following penalty is minimized:

$$\varepsilon^2(\underline{X}) = \sum_{k=1}^{10} \rho\left\{x_k - x_{k-1}\right\}$$

6.3 The Problem



6.4 The Example's Results

			
 10 jumps of 0.1	0.10	1.00	3.16
 4 jumps of 0.25	0.25	1.00	2.00
 1 jump of 1	1.00	1.00	1.00

6.5 Robust Smoothness

$$\varepsilon^2_{\mathrm{MAP}}\left(\underline{X}\right)=\left\|\underline{Y}-\mathbf{H}\underline{X}\right\|^2+\lambda\rho\left\{\mathbf{D}\underline{X}\right\}$$

Looking at the image D$\underline{X}$, small values correspond to smooth regions, and thus should be treated as before. High values correspond to edges.

Instead of suppressing edges, we reduce the cost for these points by using an appropriate $\rho(x)$ function

6.6 Iterative Solution

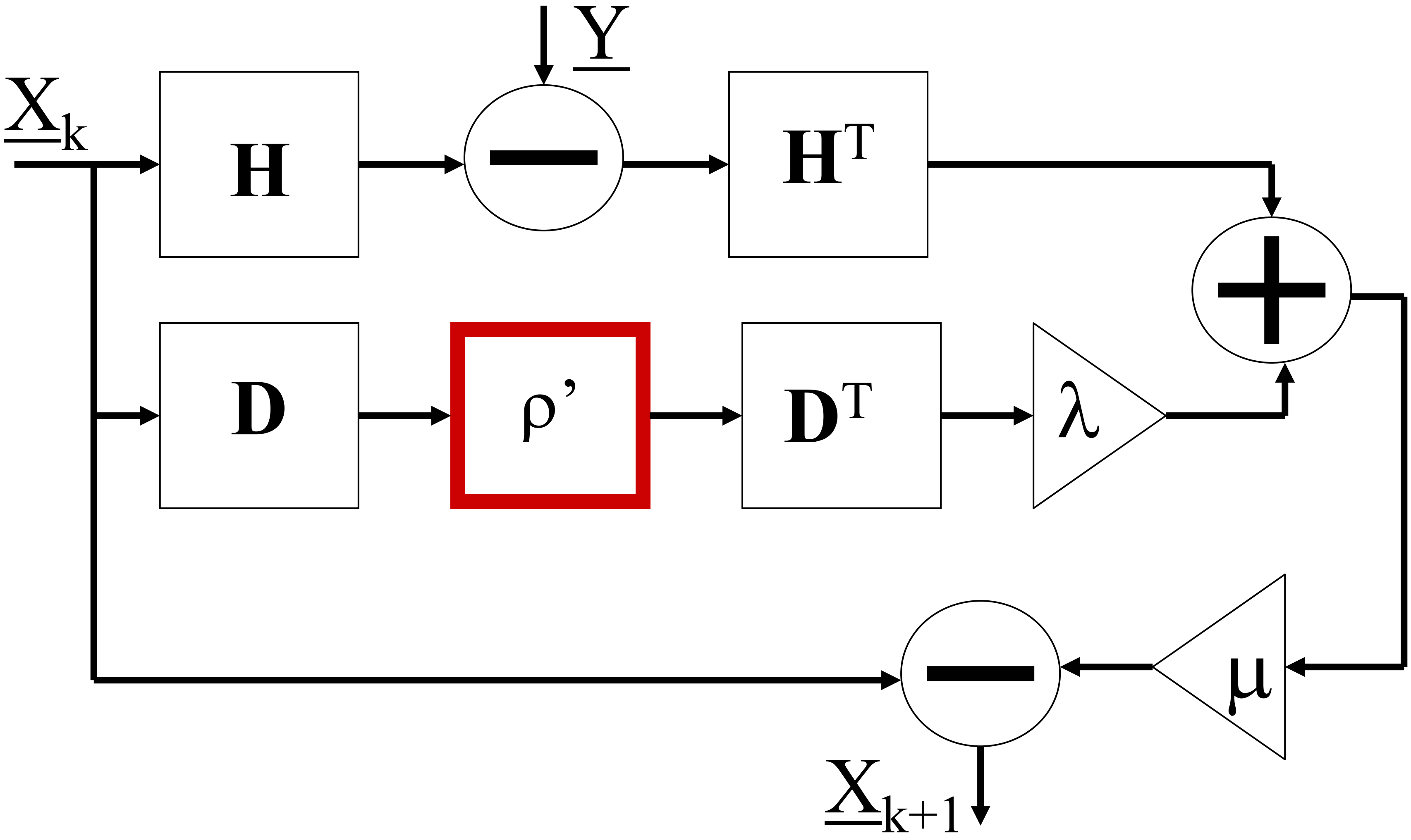
$$\varepsilon^2_{\mathrm{MAP}}(\underline{X}) = \left\|\underline{Y} - \mathbf{H}\underline{X}\right\|^2 + \lambda\rho\{\mathbf{D}\underline{X}\}$$

$$\frac{\partial \varepsilon^2_{\mathrm{MAP}}(\underline{X})}{\partial \underline{X}} = 2\mathbf{H}^{T}[\mathbf{H}\underline{X} - \underline{Y}] + 2\lambda\mathbf{D}^{T}\rho'\{\mathbf{D}\underline{X}\}$$

$$\underline{X}_{k+1} = \underline{X}_{k} - \mu\left\{\mathbf{H}^{T}[\mathbf{H}\underline{X}_{k} - \underline{Y}] - \lambda\mathbf{D}^{T}\rho'\{\mathbf{D}\underline{X}_{k}\}\right\}$$

$\underline{X}_0$ can be chosen to be $\underline{Y}$

6.7 Iterative System



6.8 Effective Weight

Weight approach $\underline{X}_{k+1} = \ldots.. - \lambda \mathbf{D}^{T} \mathbf{W}\mathbf{D}\, \underline{X}_{k}$

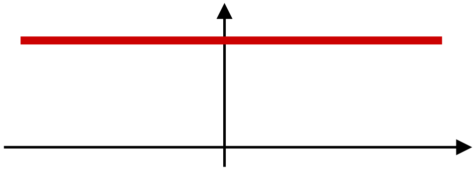
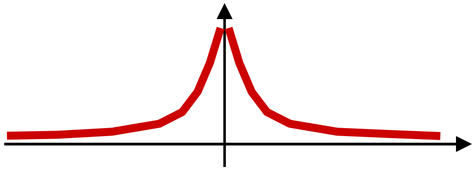
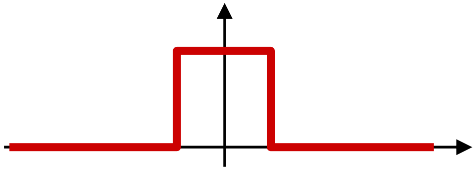
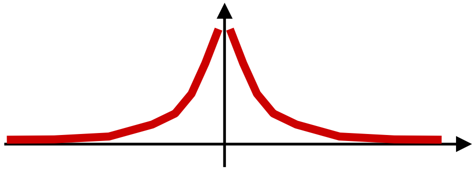
Robust approach $\underline{X}_{k+1} = \ldots.. - \lambda \mathbf{D}^{T} \rho'\left\{\mathbf{D}\underline{X}_{k}\right\}$

In order to get that they are equivalent we should require

$$\rho'\left\{\mathbf{D}\underline{X}_{k}\right\} = \mathbf{W}\mathbf{D}\underline{X}_{k} \quad \Rightarrow \quad w(x) = \frac{\rho'(x)}{x}$$

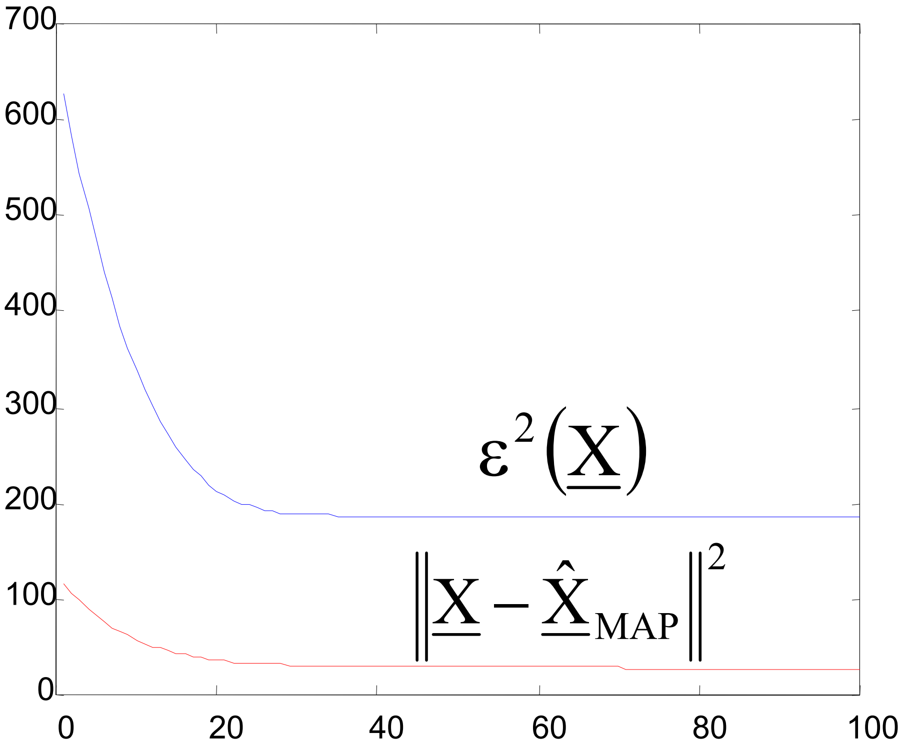
which means that we have to re-compute the weights at every iteration

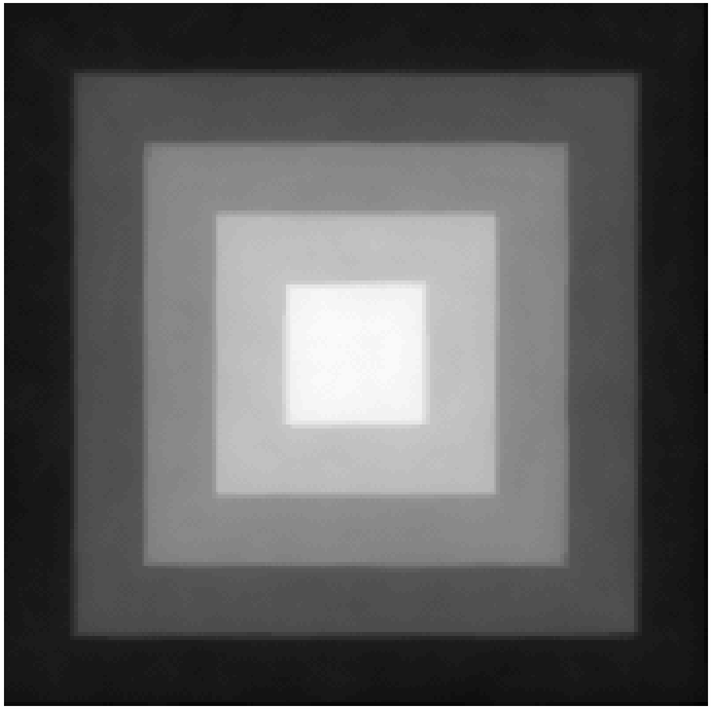
6.9 Robust Functions

$\rho(x)$	$\rho'(x)$	$\rho'(x)/x$
$0.5x^2$	x	
$\lvert x \rvert$	$\mathrm{sign}(x)$	
$\begin{cases} 0.5x^2 & \lvert x \rvert \le T \\ 0.5T^2 & \text{otherwise} \end{cases}$	$\begin{cases} x & \lvert x \rvert \le T \\ 0 & \text{otherwise} \end{cases}$	
$\sqrt{\lvert x \rvert}$	$\dfrac{\mathrm{sign}(x)}{\sqrt{\lvert x \rvert}}$	

6.11 Example

Using $\rho(x)=|x|$ and $\underline{X}_0=\underline{Y}$, Error=27.6





6.10 Local Minimum

For NON-CONVEX ρ(x)

$$\varepsilon^{2}_{\mathrm{MAP}}\left(\underline{\mathrm{X}}\right)=\left\|\underline{\mathrm{Y}}-\mathbf{H}\underline{\mathrm{X}}\right\|^{2}+\lambda\rho\left\{\mathbf{D}\underline{\mathrm{X}}\right\}$$

is not convex as well!

The solution may be different for different initializations

Chapter 7

What Next ?

7.1 The Chosen Prior

Definition: A ‘prior’ is a function f(X) that gives very low values for ‘good’ images and high values for ‘bad’ images

We have seen so far:

1. $f(\underline{X}) = \left\| \mathbf{D}\underline{X} \right\|^2$
2. $f(\underline{X}) = \left[\mathbf{D}\underline{X}\right]^T \mathbf{W} \left[\mathbf{D}\underline{X}\right]$
3. $f(\underline{X}) = \rho\left\{\mathbf{D}\underline{X}\right\}$

7.2 Better Priors

Priors should exploit properties of images in each application: e.g.

1. Presence of point sources,
2. Multi-resolution behavior,
3. Edges directionality,
4. Relation between colors
5. If possible, higher level description

7.3 Learning Priors

Instead of GUESSING how images should look, why not let them tell their story.

⬇

Take a large set of images and learn from them what is the best prior (or at least, what are the parameters for a pre-specified parametric prior)

Very well suited to astronomical applications.

7.4 More General Problems

1. Image Scaling problem
2. Super-Resolution problem
3. Image sequence restoration problem
4. Other imaging degradation models for blur and noise
5. Other inverse problems having the same structure

Topic : Multiframe Resolution Enhancement

Resolution Enhancement Idea

- Given multiple low-resolution moving images of a scene (a video), generate a high resolution image (or video).

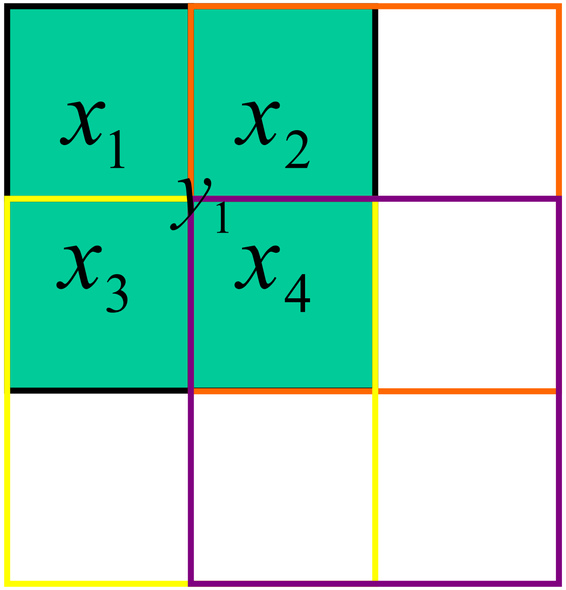
$$\underbrace{\text{frame}_1, \text{frame}_2, \cdots, \text{frame}_{N\text{-}1}, \text{frame}_N}_{\text{High Resolution Frame}}$$

$$\underbrace{\text{frame}_1, \text{frame}_2, \cdots,}_{\text{High Resolution Frame}_1} \; \underbrace{\cdots, \text{frame}_{N\text{-}1}, \text{frame}_N}_{\text{High Resolution Frame}_2}$$

"Trading off time resolution or view diversity to gain spatial resolution"

Resolution Enhancement Model

- A simple model relating the low-resolution blurry image to the high resolution crisper image.



“PSF”

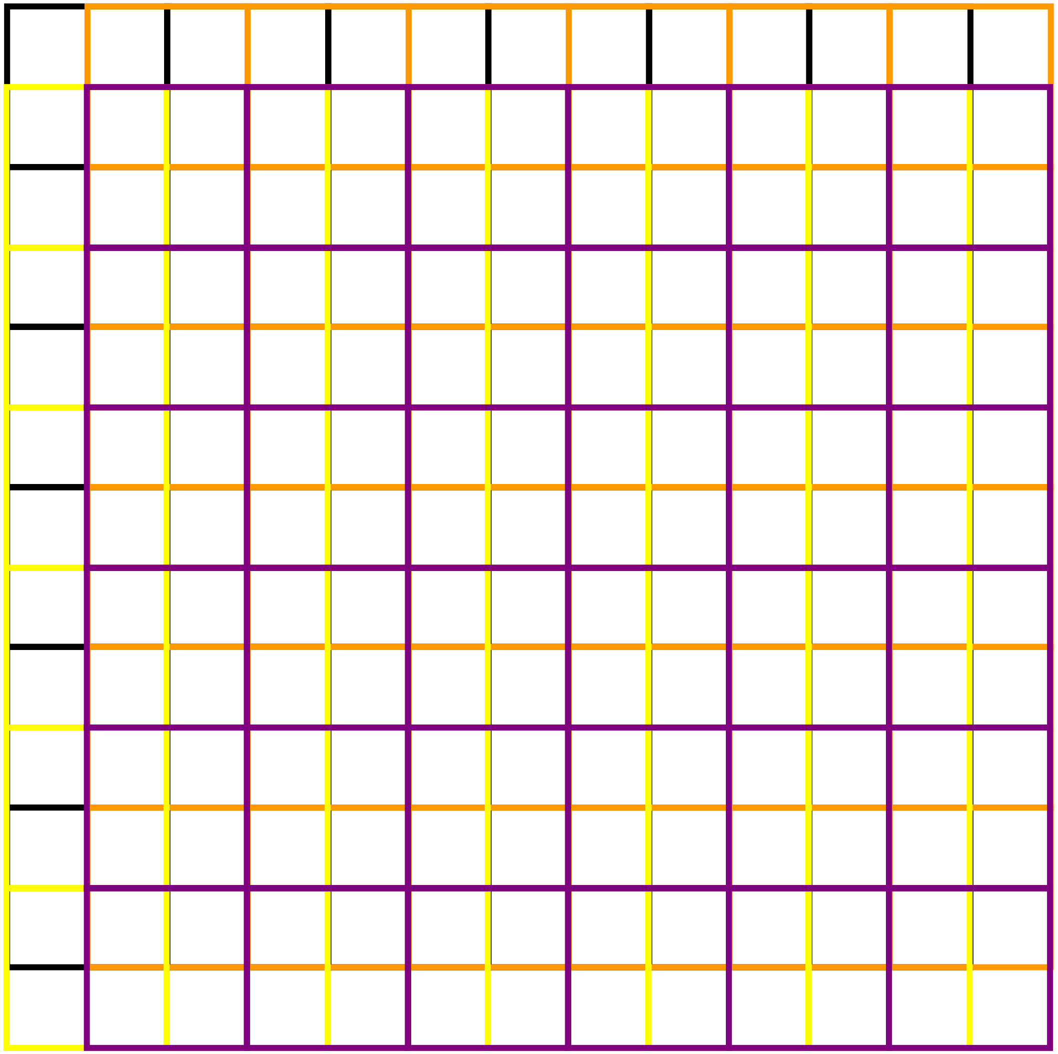
$$y_1 = a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + e_1$$

$$y_2 = 0 \cdot x_1 + a_1x_2 + 0 \cdot x_3 + a_3x_4 + e_2$$

$$y_3 = 0 \cdot x_1 + 0 \cdot x_2 + a_1x_3 + a_2x_4 + e_3$$

$$y_4 = 0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 + a_1x_4 + e_4$$

Low vs High Res Pixels



x2 enhancement

The Mathematical Model

$$\text{k-th frame} \longrightarrow \underline{y}_k = A_k \underline{X} + \underline{e}_k \qquad for \qquad 1 \le k \le N$$

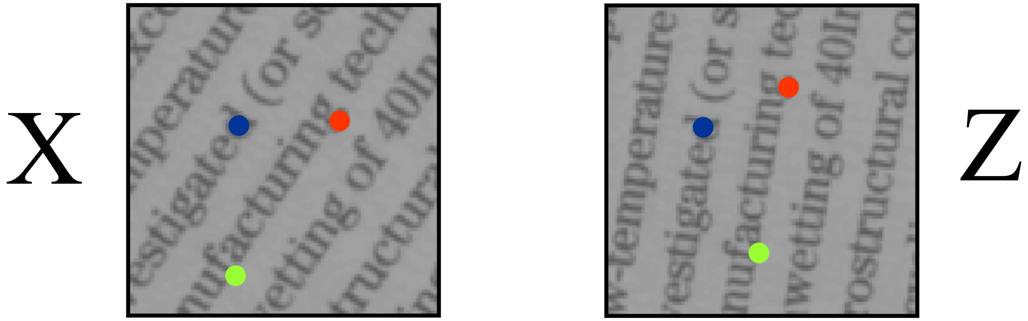
$$A_k = \underbrace{D_k}_{\text{Downsampling}} \underbrace{H_k}_{\text{Blurring}} \underbrace{F_k}_{\text{Warping}}$$

$$A_k = [T_{k,1} \quad T_{k,2} \quad \cdots \quad \underbrace{T_{k,N^2}}_{\text{Upper-banded, “nearly” Toeplitz}}]$$

$$\text{BTTB system*} \longrightarrow \underline{Y} = A\underline{X} + \underline{e}$$

*PCG-based methods developed by N. Nguyen, 2000

The Warping Operation as a Matrix (F)



Per every point in X find a matching point in Z

$$\begin{bmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} 1 & & \mathbf{0} \\ & 1 & \\ \mathbf{0} & & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_j \\ \vdots \\ z_N \end{bmatrix}$$

F[j,i]=1

Additional Intricacies

- Sampling is not a point operation – there is a blur
- Motion (warping) must be estimated!
- Motion may include perspective warp, local motion, etc.
- The system is typically underdetermined and ill-conditioned.
- 10’s or 100’s of thousands of unknown variables and data.

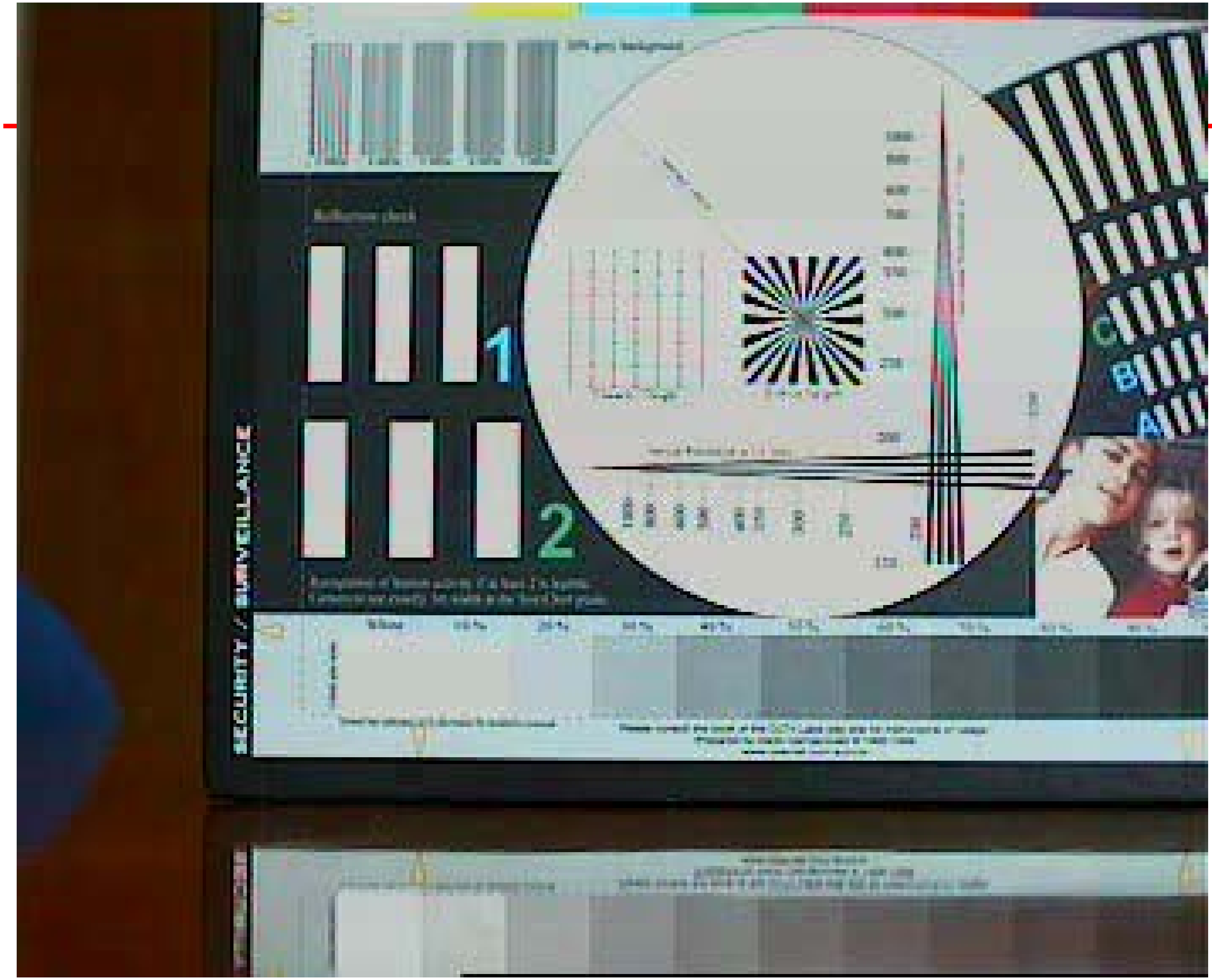
Some Examples: License Plate Reading

Digital Video Camera from 2nd story window

More Examples

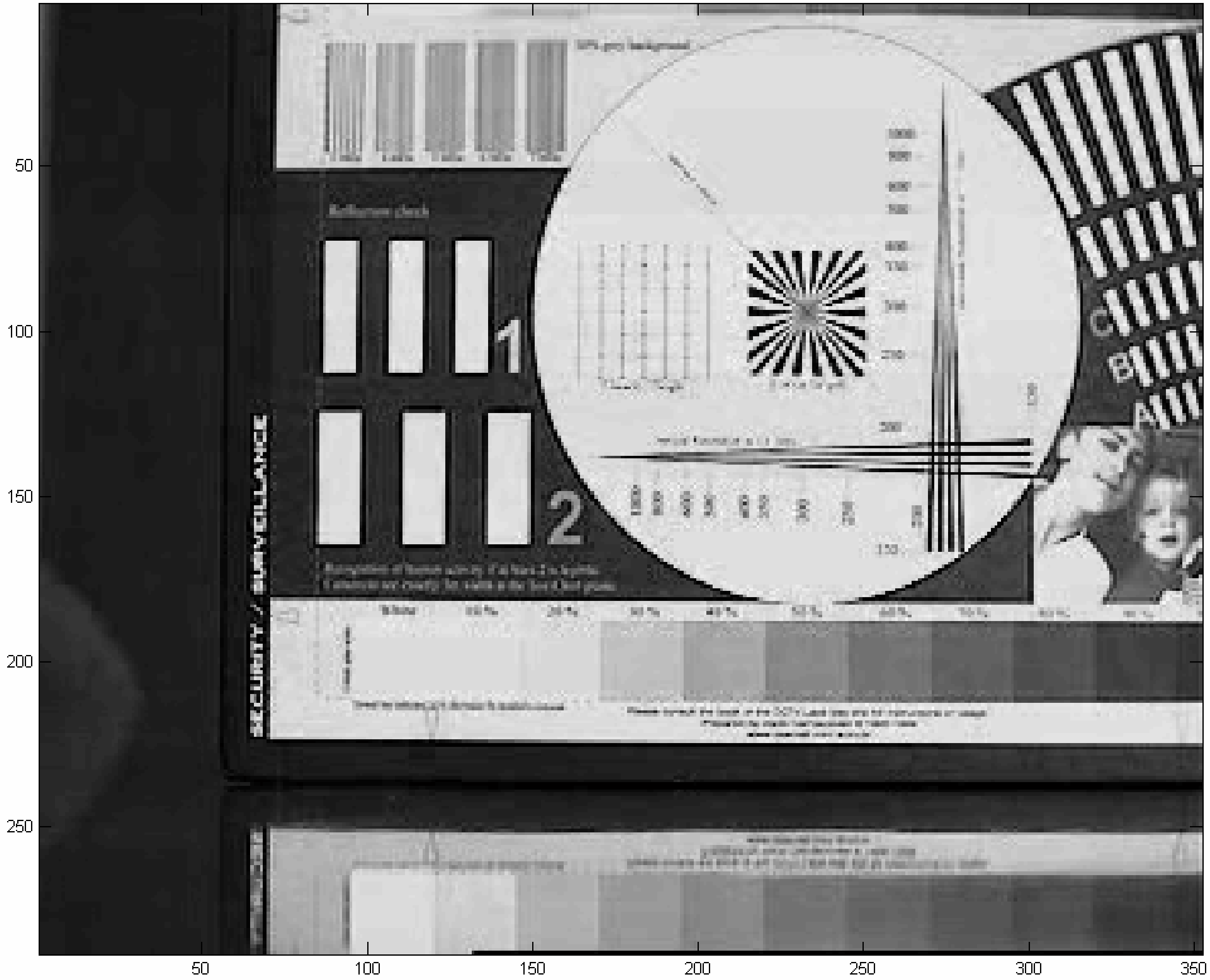
Infrared Camera (Night Vision)

Data Courtesy Wright Labs USAF



Data Courtesy Vigilant Technology

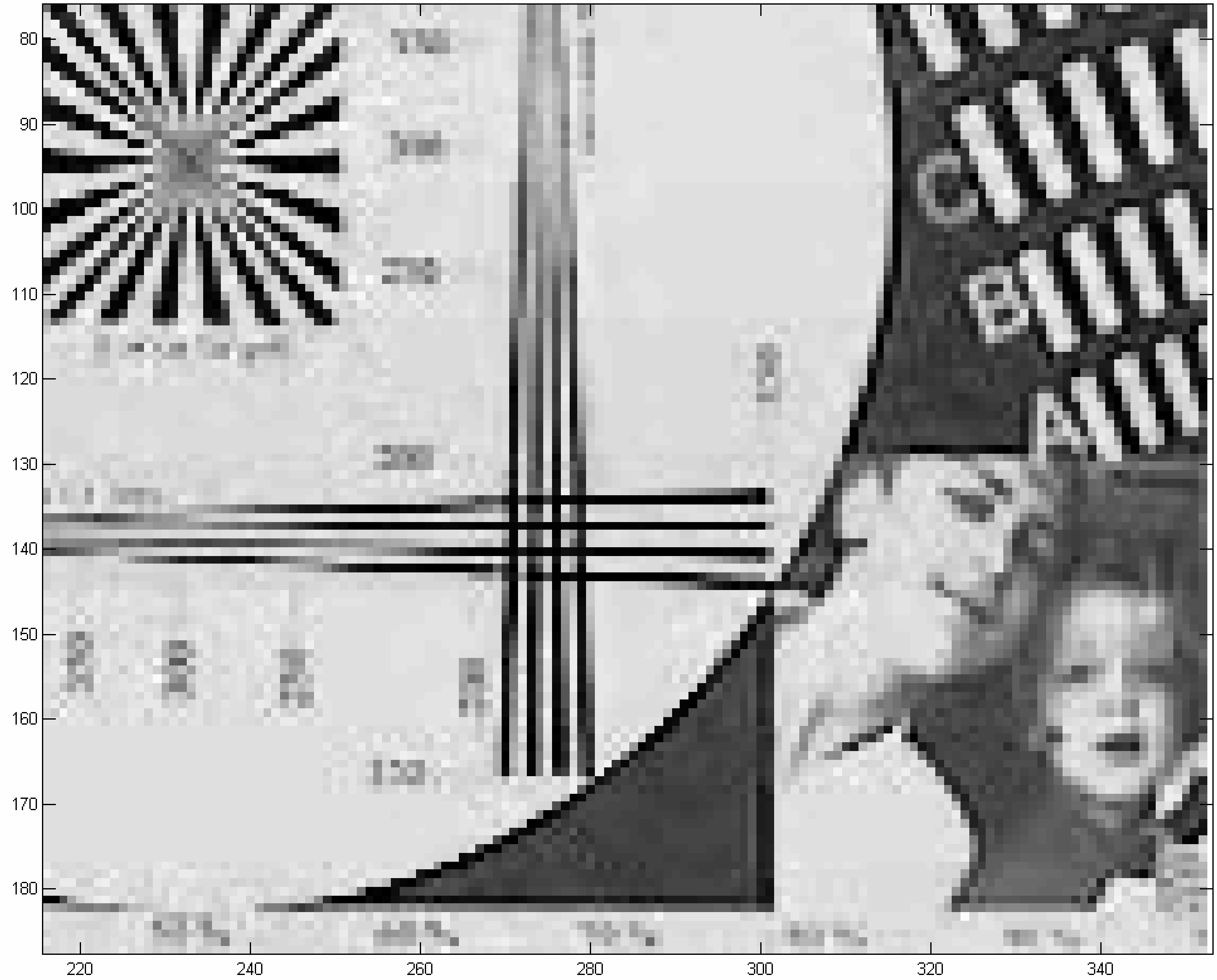
Before


SECURITY / SURVEILLANCE
1
2
50
100
150
200
250
50
100
150
200
250
300
350

After

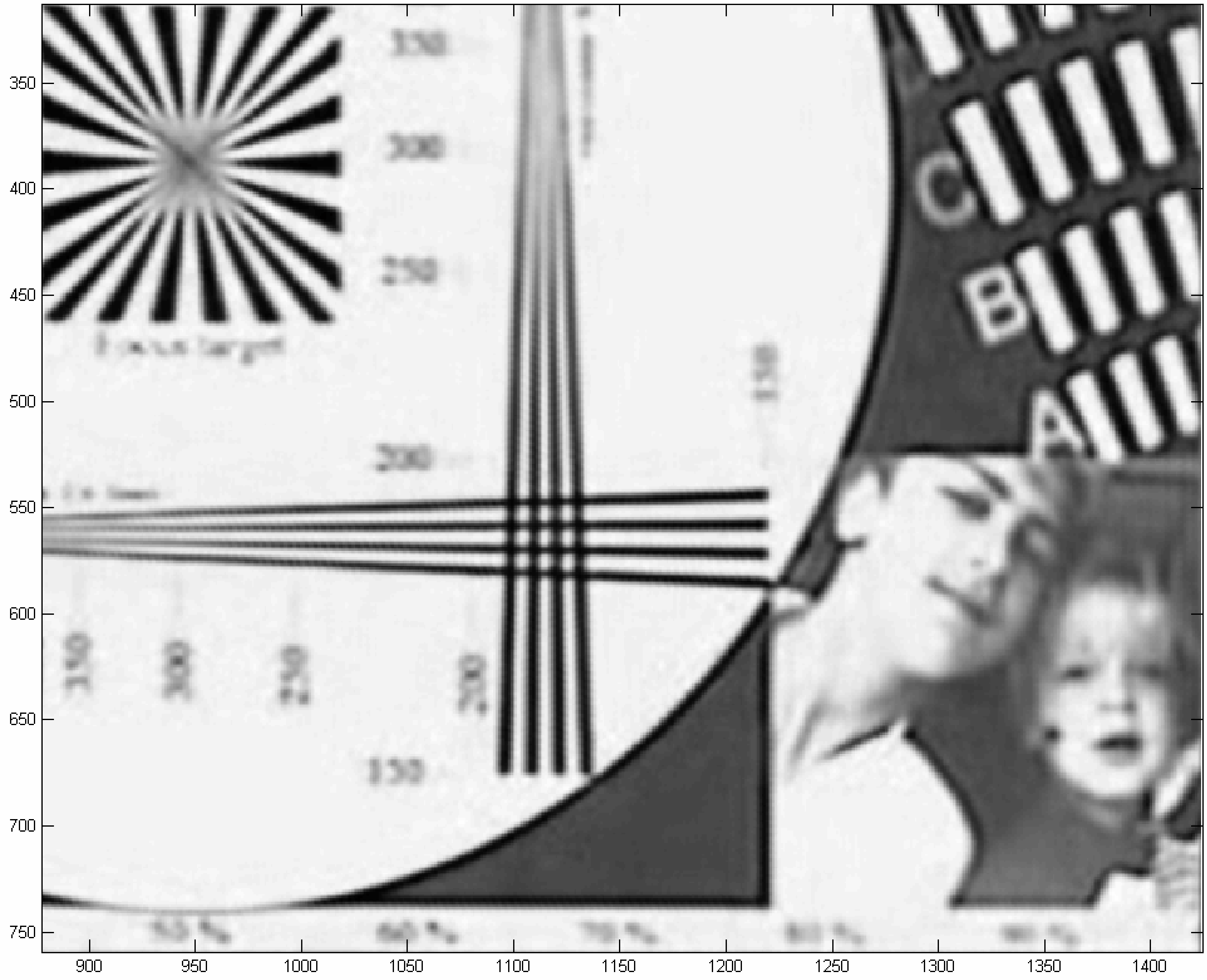

100
200
300
400
500
600
700
800
900
1000
1100
200
400
600
800
1000
1200
1400
Reflection check
1
2
SECURITY / SURVEILLANCE
Focus target
C
B
A
1000
800
600
500
400
350
300
250
200
150
White
10 %
20 %
30 %
40 %

Detail Before



Data Courtesy Vigilant Technology

Detail After



Data Courtesy Vigilant Technology

Conclusions

- The field of image restoration has been around for more than 40 years, with more than 10,000 papers, and numerous active researchers in engineering and elsewhere.

- There is much more to say because
  - New applications continue to pop up,
  - The results are still not satisfactory,
  - As computers improve, complicated solutions become practical