Workshop on Computer vision, graphics and Image processing

Optimization

Dr. Uma Mudenagudi

Professor, Department of E and C

BVBCET HUBLI

Class-33, 34 01-07-2010 10.00-1.15pm

Overview

- Motivation
- Optimization methods
- SR problem
- Different solutions to SR problems
- Conclusion

What is Optimization?

- **Optimization** is the mathematical discipline which is concerned with finding the maxima and minima of functions, possibly subject to constraints.

Uses of optimization

- Architecture
- Nutrition
- Electrical circuits
- Economics
- Transportation
- etc.

What do we optimize?

- A real function of n variables

$$f(x_1, x_2, \ldots, x_n)$$

- with or without constrains

Unconstrained optimization

$$\min f(x,y) = x^2 + 2y^2$$

Optimization with constraints

$$\min f(x, y) = x^2 + 2y^2$$
$$x > 0$$

or

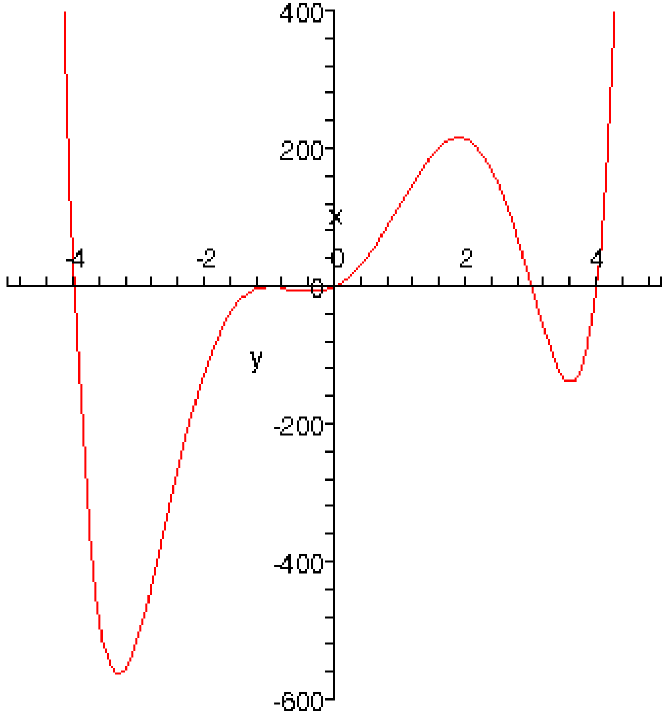
$$\min f(x, y) = x^2 + 2y^2$$
$$-2 < x < 5, y \geq 1$$

or

$$\min f(x, y) = x^2 + 2y^2$$
$$x + y = 2$$

Lets Optimize

- Suppose we want to find the minimum of the function



Review max-min

- What is special about a local max or a

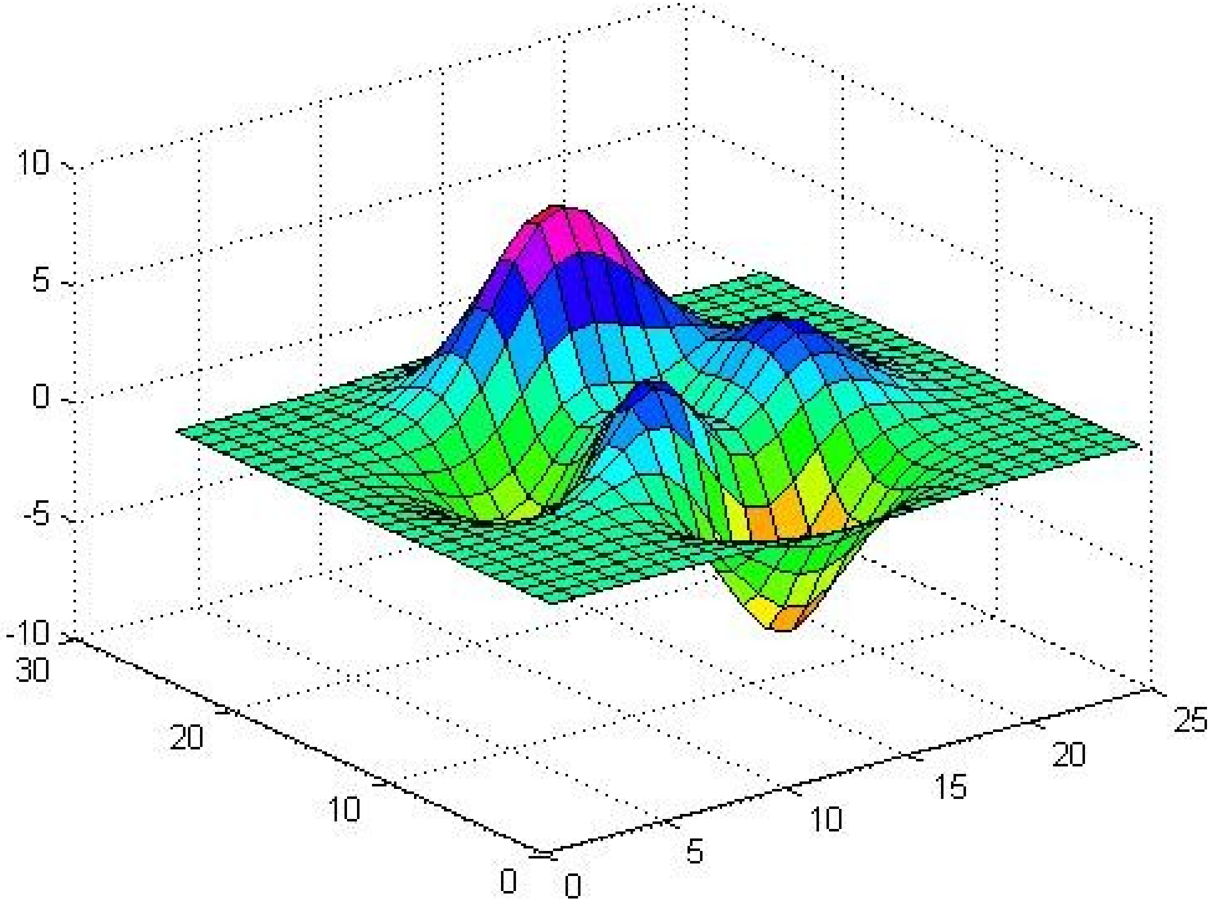
  local min of a function **f** (x)?

  at local max or local min **f**'(x)=0

  f''(x) > 0 if local min

  f''(x) < 0 if local max

Review max-min



Review max-min

- Second Derivative Test
- Local min, local max, saddle point
- Gradient of **f** – vector (d**f/**dx**,** d**f**/dy**,** d**f**/dz)

  direction of fastest increase of **f**
- Global min/max vs. local min/max

Gradient Descent Method, ex

- Minimize function

$$f(x,y) = 0.5(\alpha x^2 + y^2)$$

$$-11 \le x, y \le 11$$

- Minimize function

$$f(x,y) = \cos(x)\cos(y)$$

$$-4 \le x, y \le 4$$

Gradient Descent Method

- Use function gd(alpha,x0)
  - Does gd.m converge to a local min? Is there a difference if $\alpha > 0$ vs. $\alpha < 0$?
  - How many iterations does it take to converge to a local min? How do starting points x0 affect number of iterations?
- Use function gd2(x0)
  - Does gd2.m converge to a local min?
  - How do starting points x0 affect number of iterations and the location of a local minimum?

How good are the optimization methods?

- Starting point
- Convergence to global min/max.
- Classes of nice optimization problems

  Example: $f(x,y) = 0.5(\alpha x^2+y^2)$, $\alpha > 0$

  Every local min is global min.

The Super-Resolution Problem

$$\underline{Y}_k = \mathbf{DHF}_k \underline{X} + \underline{V}_k, \quad \underline{V}_k \sim \mathbf{N}\left\{0, \sigma_n^2\right\}$$

- Given

  $\underline{Y}_k$ – The measured images (noisy, blurry, down-sampled ..)

  H – The blur can be extracted from the camera characteristics

  D – The decimation is dictated by the required resolution ratio

  $\mathbf{F}_k$ – The warp can be estimated using motion estimation

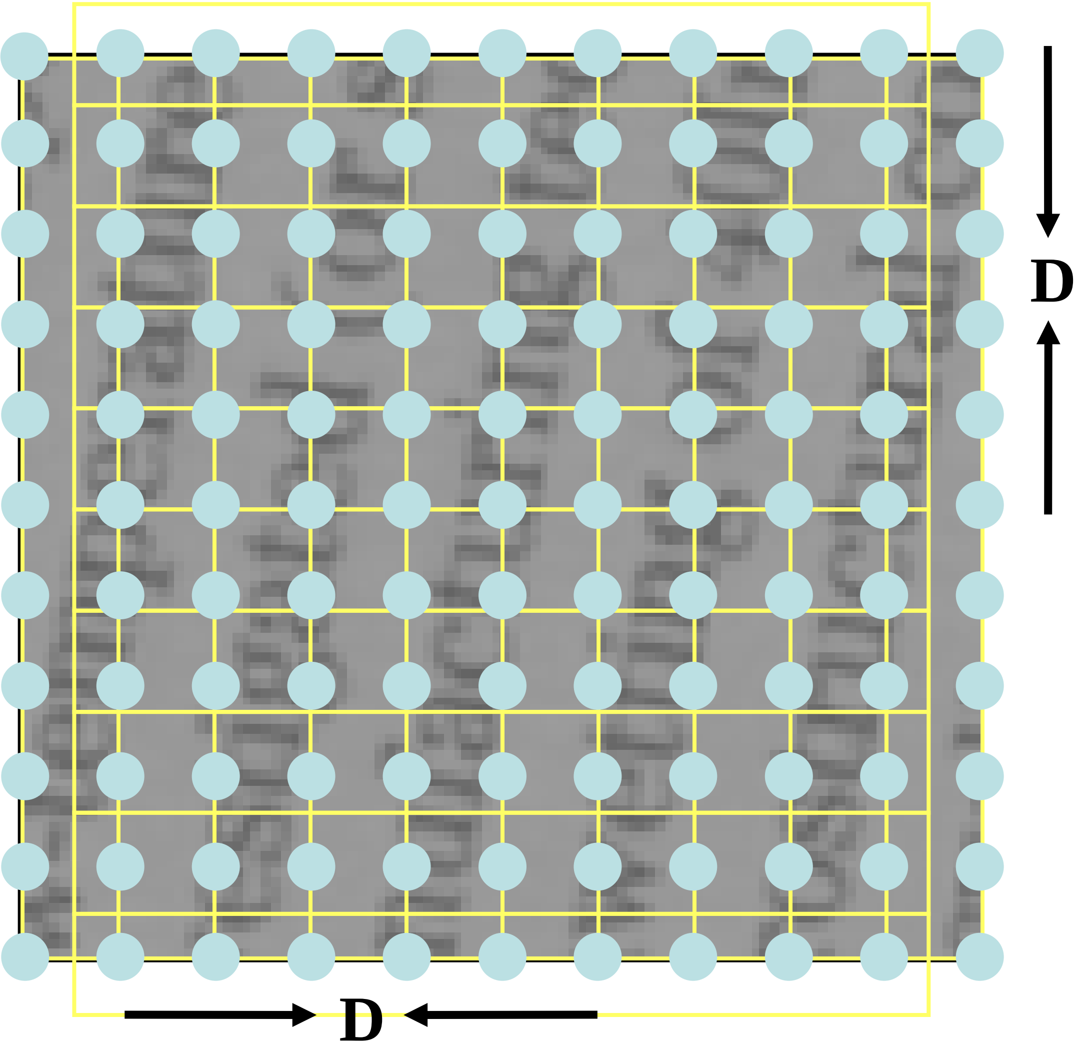
  $\boldsymbol{\sigma}_{\mathbf{n}}$ – The noise can be extracted from the camera / image

- Recover

  $\underline{X}$ – HR image

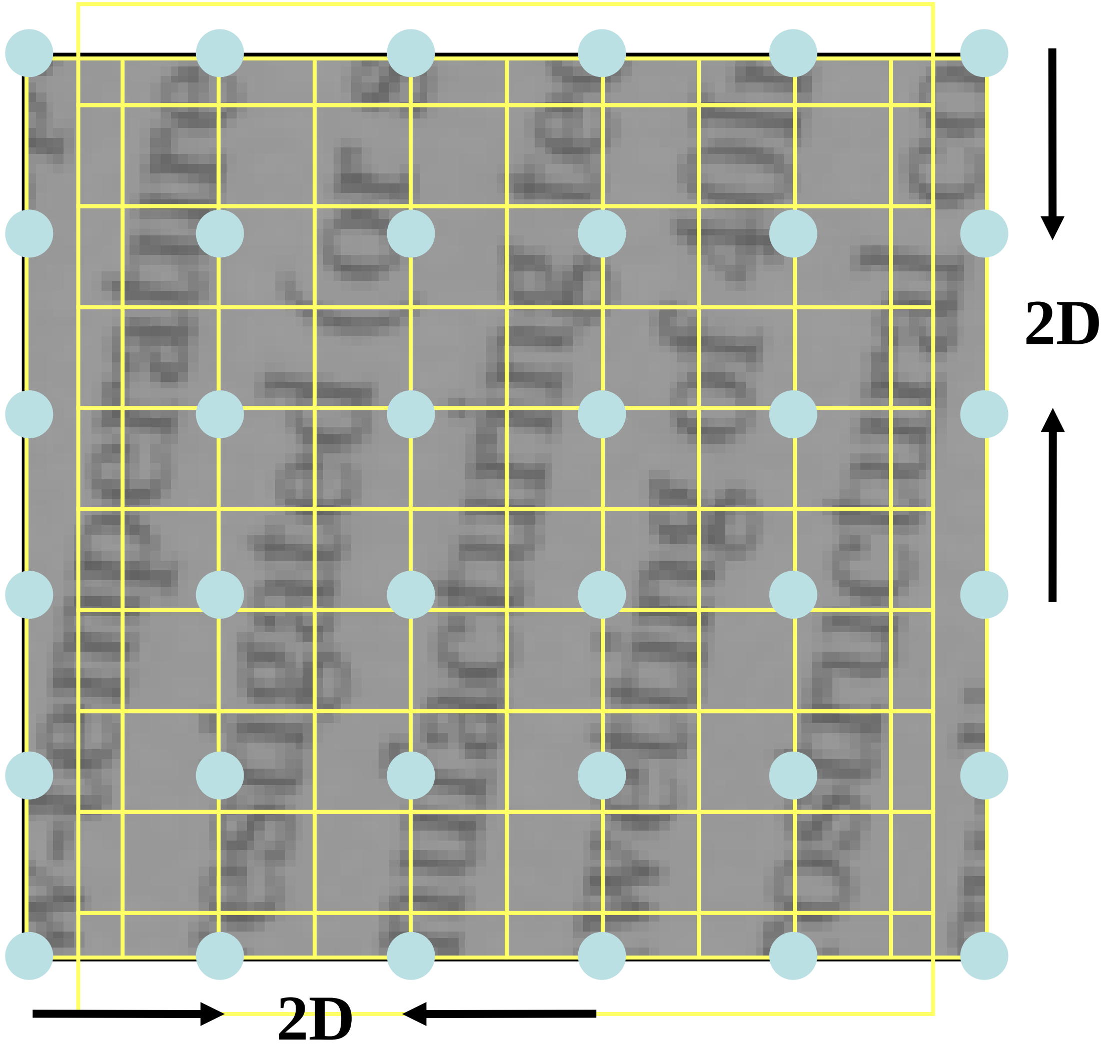
Intuition

For a given band-limited image, the Nyquist sampling theorem states that if a uniform sampling is fine enough (≥**D**), perfect reconstruction is possible.



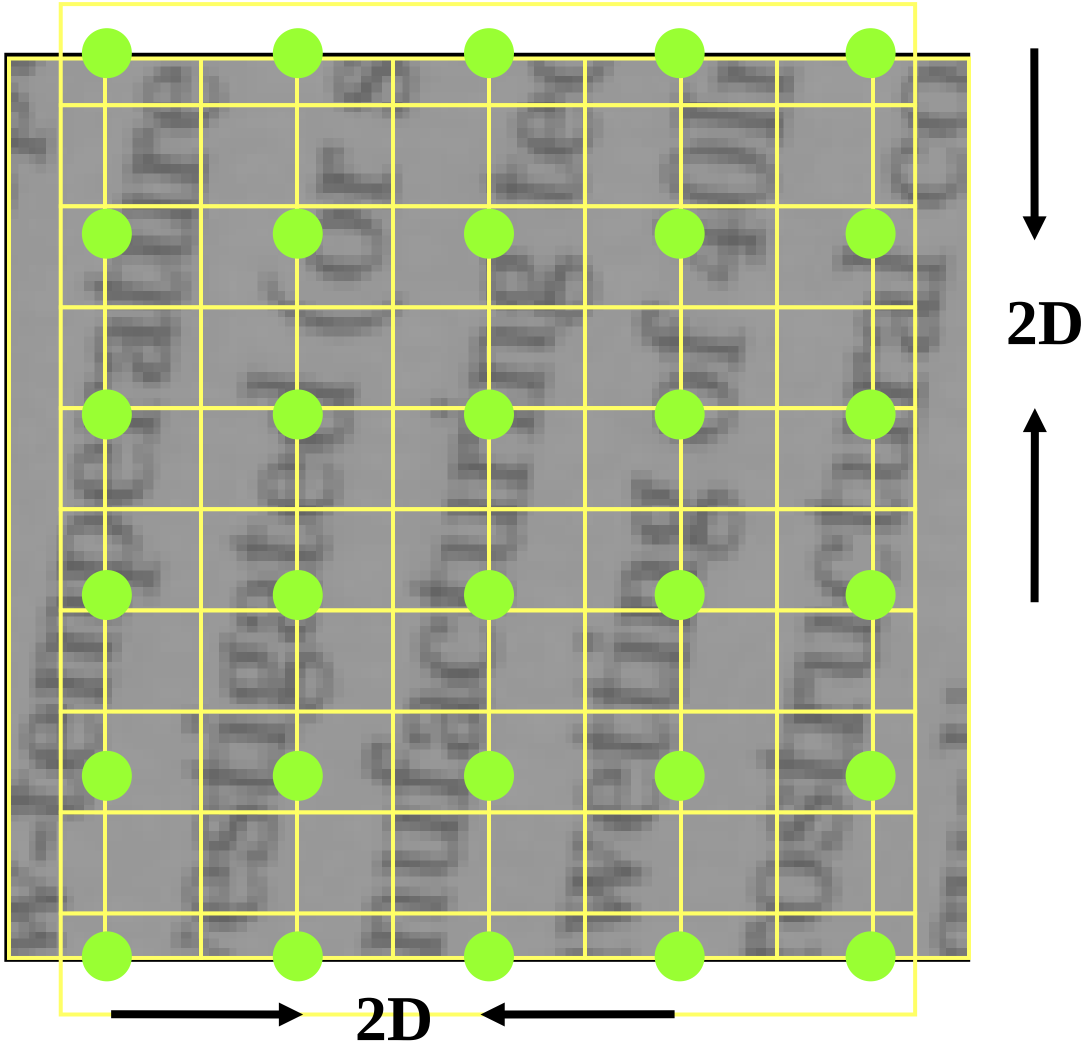
Intuition

Due to our limited camera resolution, we sample using an insufficient 2D grid



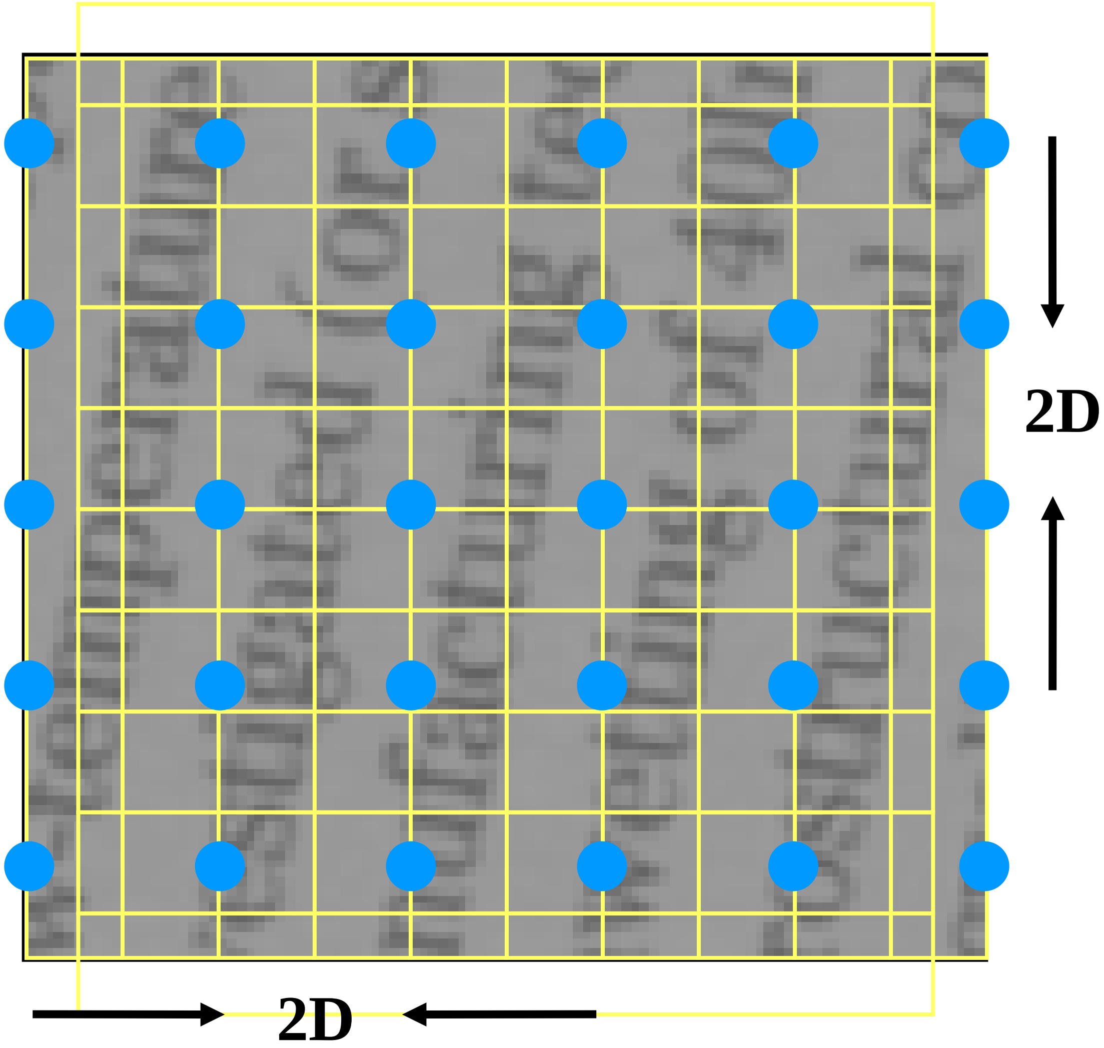
Intuition

However, if we take a second picture, shifting the camera 'slightly to the right' we obtain:



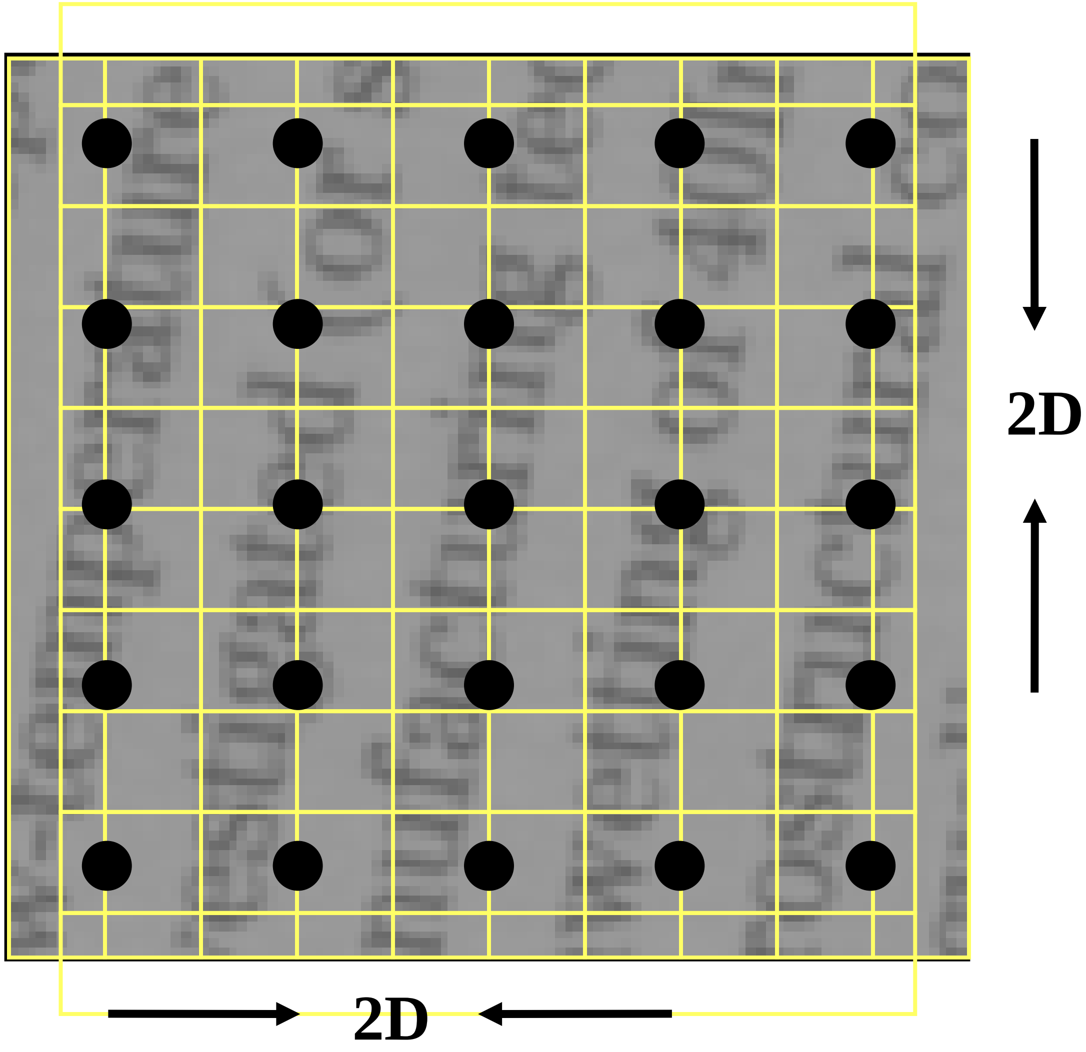
Intuition

Similarly, by shifting down we get a third image:



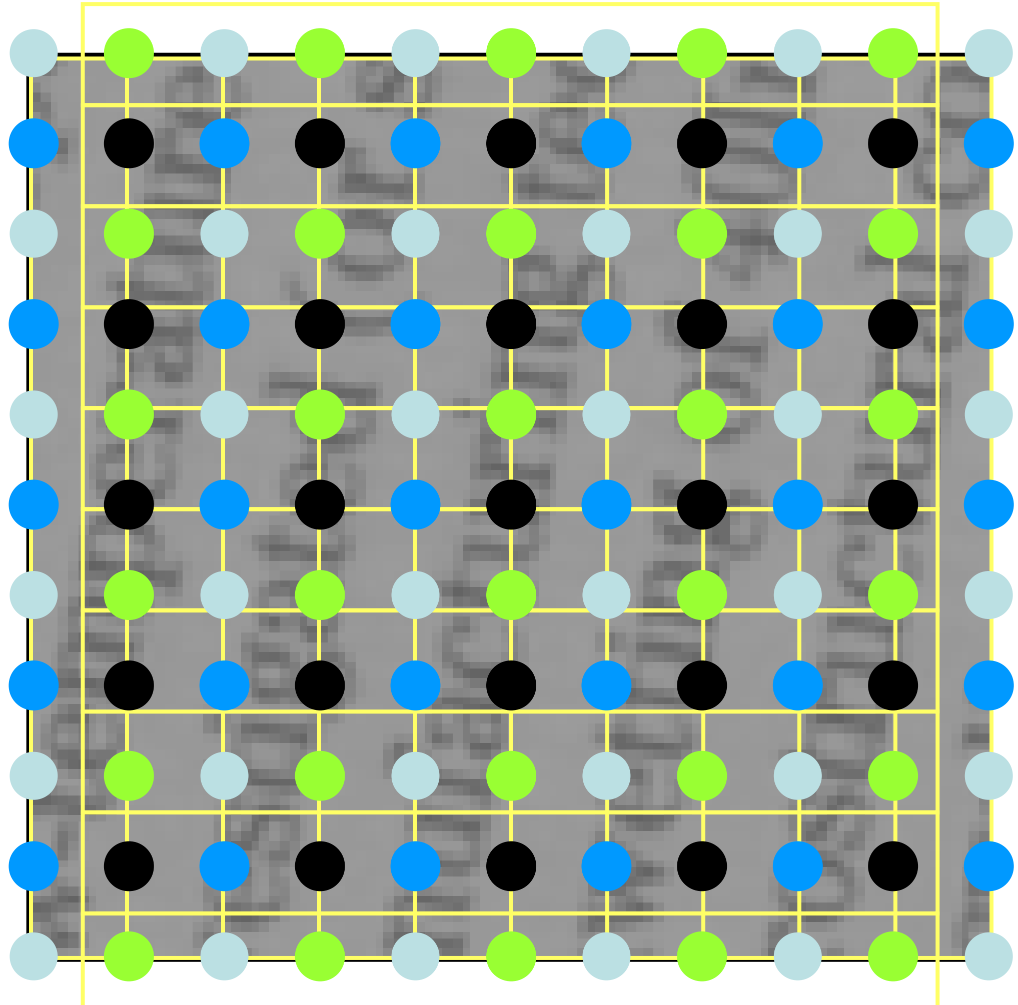
Intuition

And finally, by shifting down and to the right we get the fourth image:



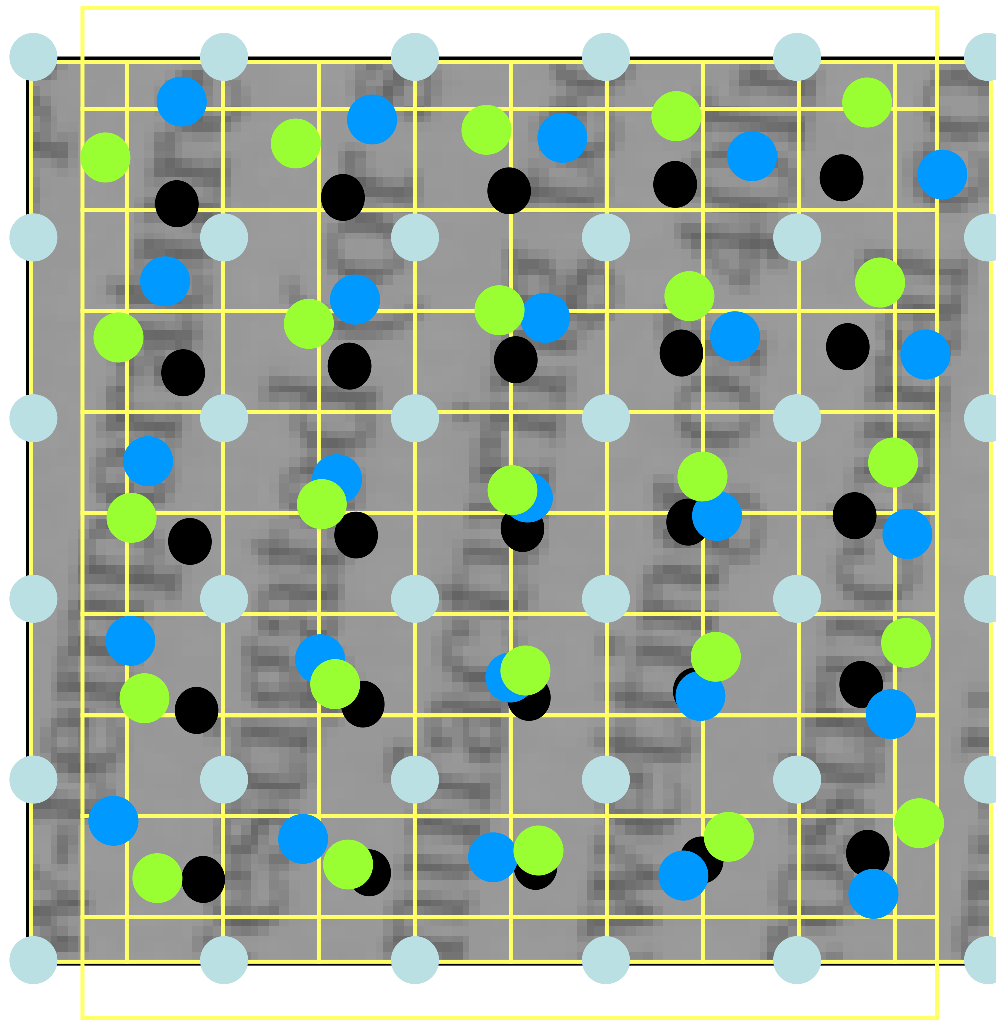
Intuition

It is trivial to see that interlacing the four images, we get that the desired resolution is obtained, and thus perfect reconstruction is guaranteed.



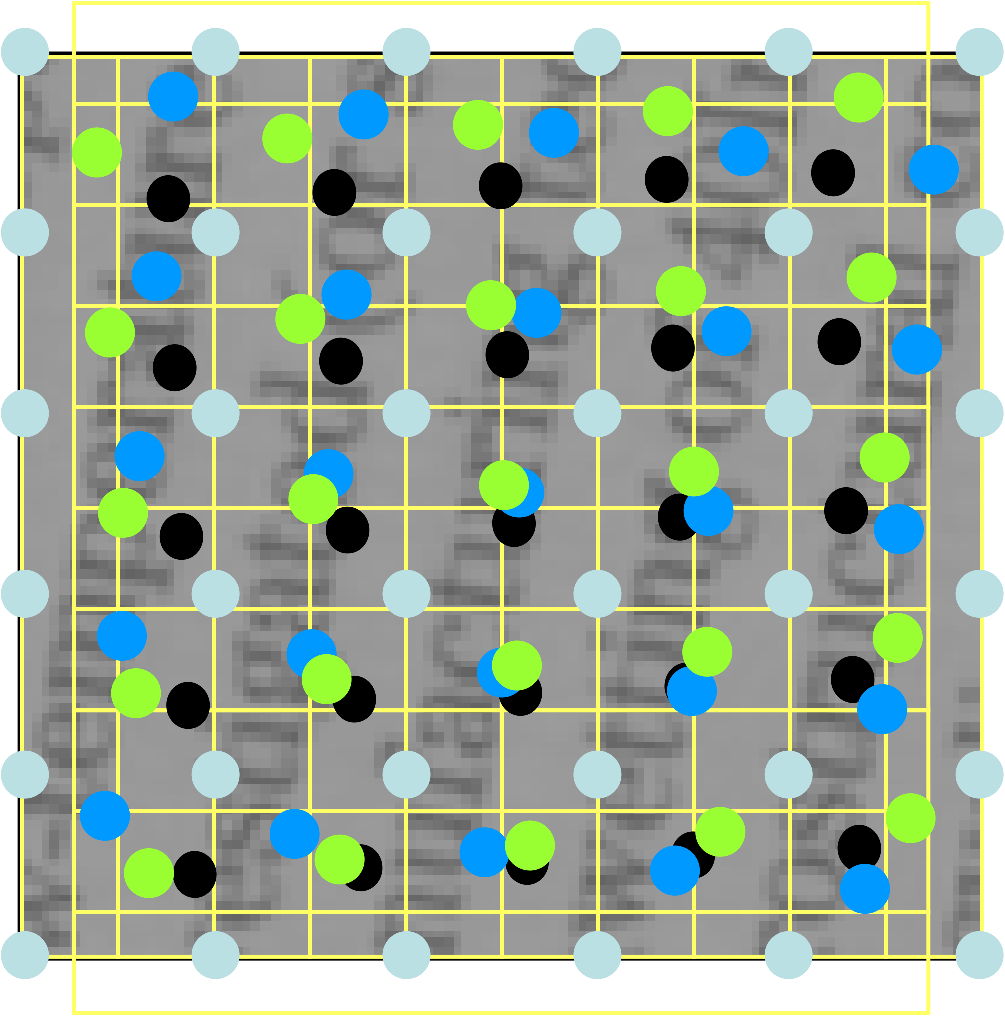
Rotation/Scale/Disp.

What if the camera displacement is Arbitrary ?
What if the camera rotates? Gets closer to the object (zoom)?

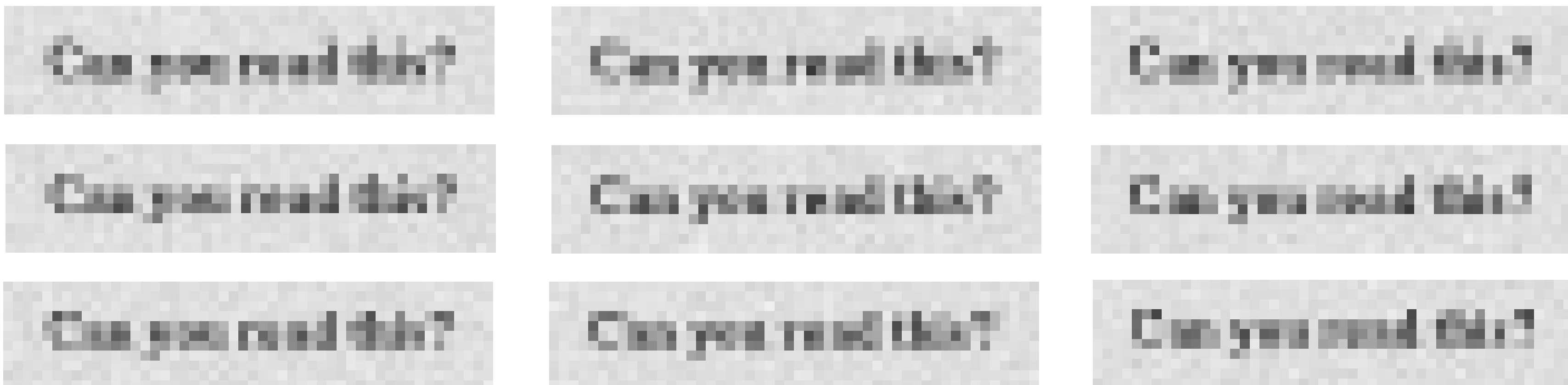


Rotation/Scale/Disp.

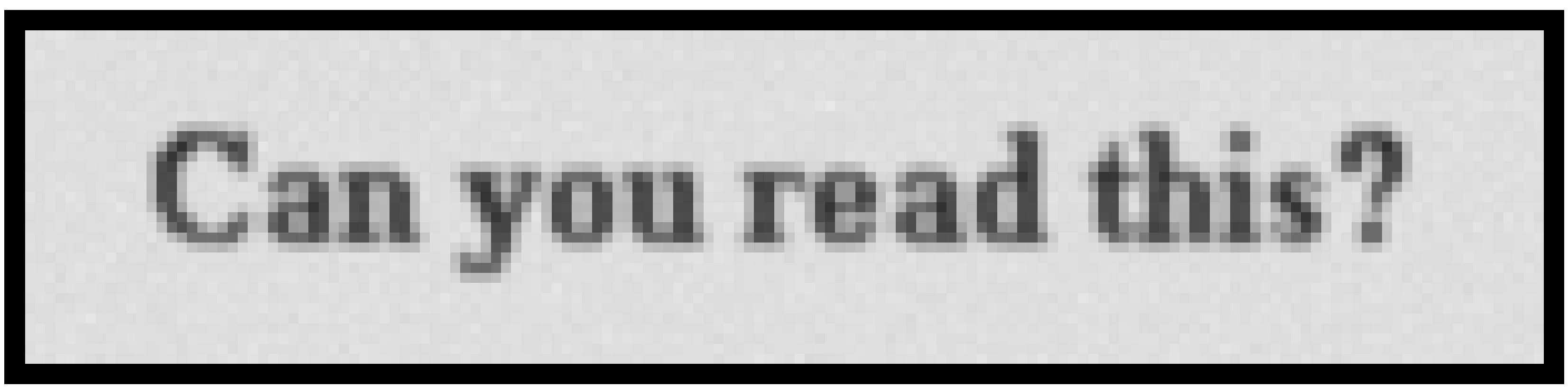
There is no sampling theorem covering this case



A Small Example



3:1 scale-up in each axis using 9 images, with pure global translation between them



Further Complications

- Complicated motion
  - perspective, local motion, …
- Blur
  - sampling is not a point operation
  - Spatially variant blur
  - Temporally variant blur
- Noise
- Changes in the scene

Super-Resolution - Agenda

- The basic idea
- **Image formation process**
- Formulation and solution
- Special cases and related problems
- Limitations of Super-Resolution
- SR in time

Image Formation

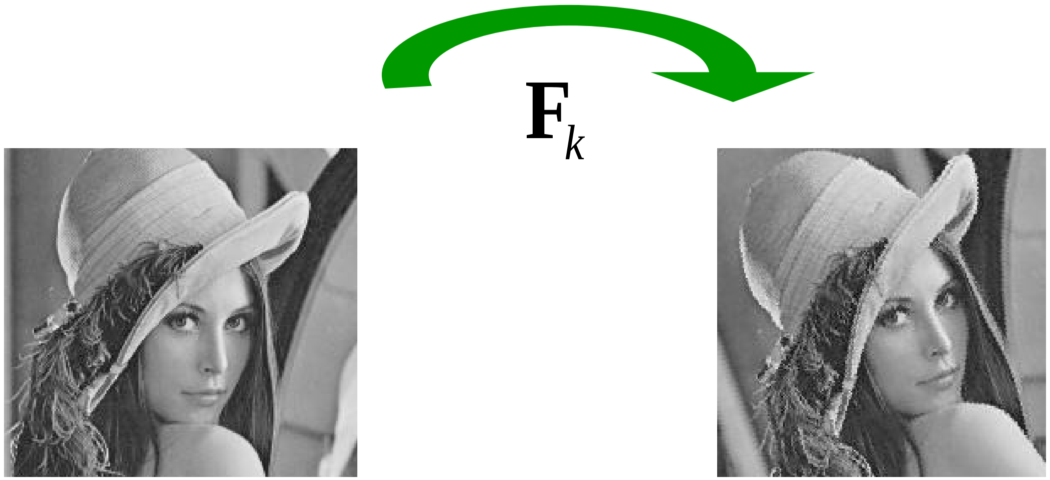
Scene | Geometric transformation | Optical Blur | Sampling | Noise

HR | $\mathbf{F}_k$ | $\mathbf{H}_k$ | $\mathbf{D}_k$ | **LR**

Can we write these steps as linear operators?

$$\mathrm{LR} = \mathbf{D}_k \mathbf{H}_k \mathbf{F}_k \cdot \mathrm{HR}$$

Geometric Transformation



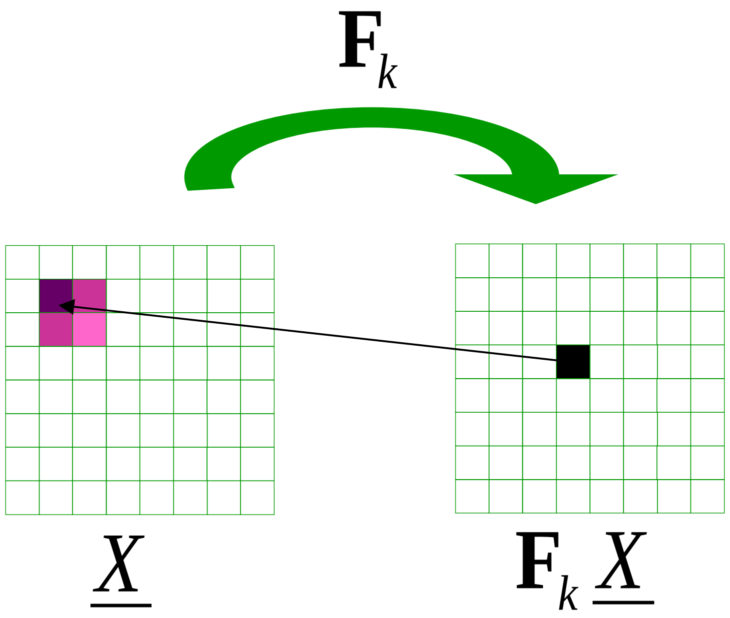
Scene

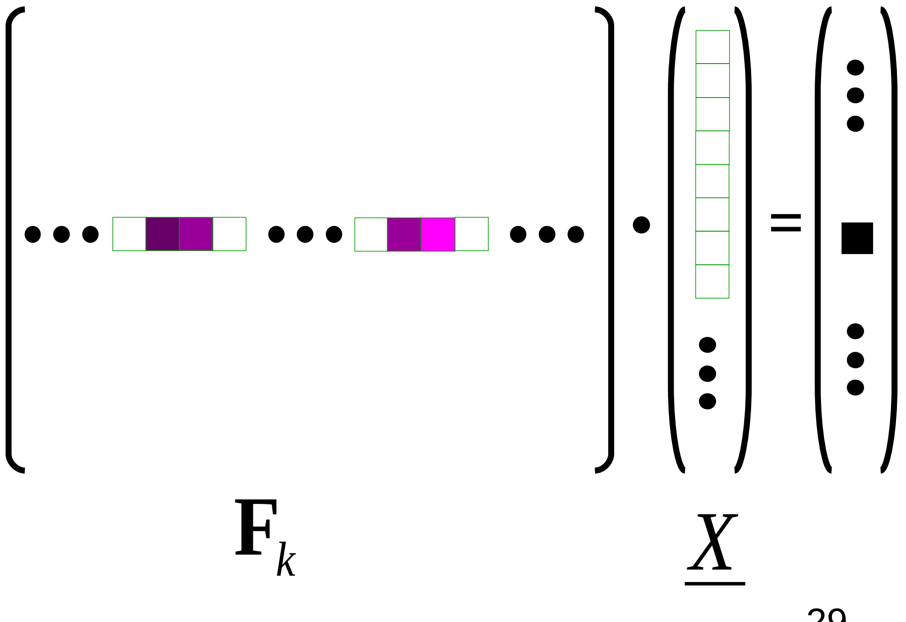
Geometric transformation

- Any appropriate motion model
- Every frame has different transformation
- Usually found by a separate registration algorithm

Geometric Transformation

Can be modeled as a linear operation $\mathbf{F}_k \underline{X}$





Optical Blur

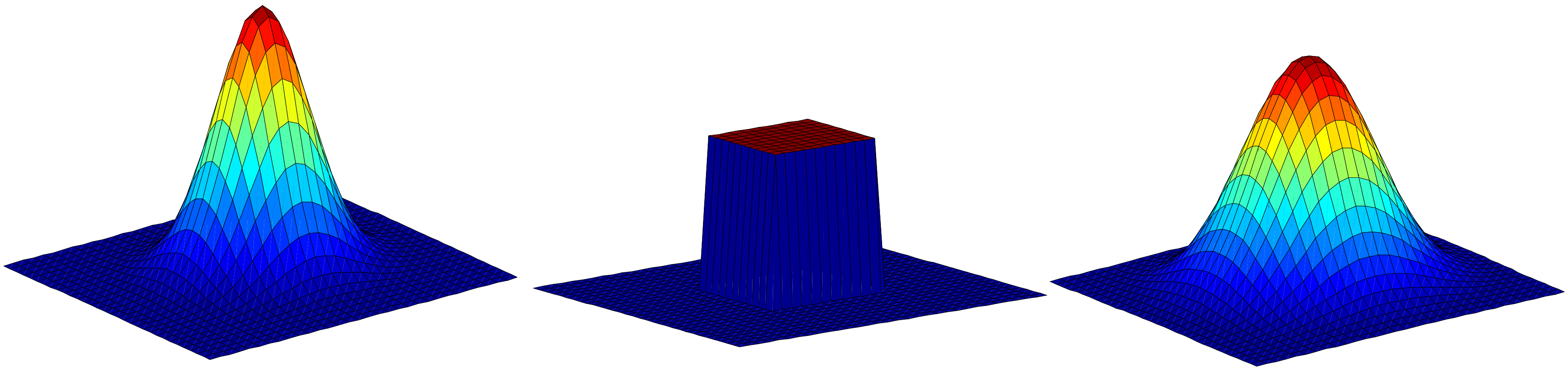
$\mathbf{H}_k$

Geometric transformation

Optical Blur

- Due to the lens PSF and pixel integration
- Usually $\mathbf{H}_k = \mathbf{H}$

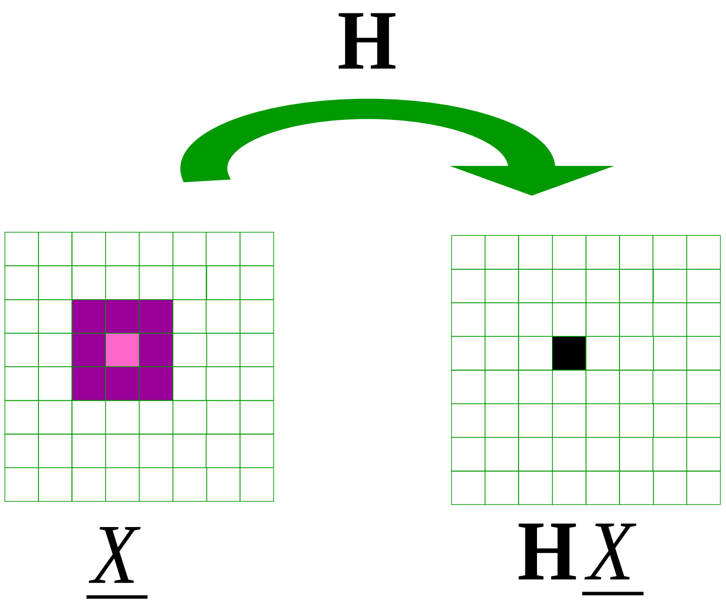
H

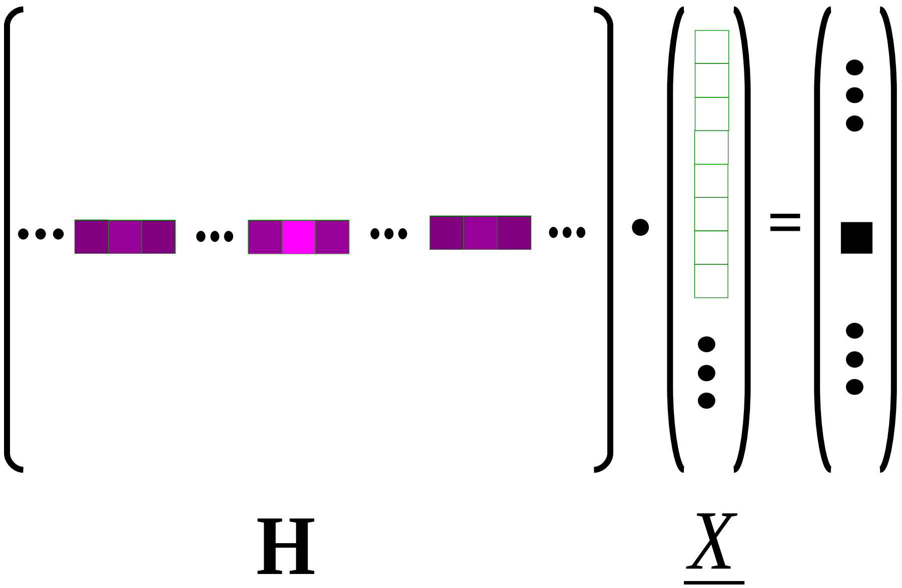


PSF * PIXEL = H

Optical Blur

Can be modeled as a linear operation $\mathbf{H}\underline{X}$





Sampling

$$\mathbf{D}_k$$

Optical Blur

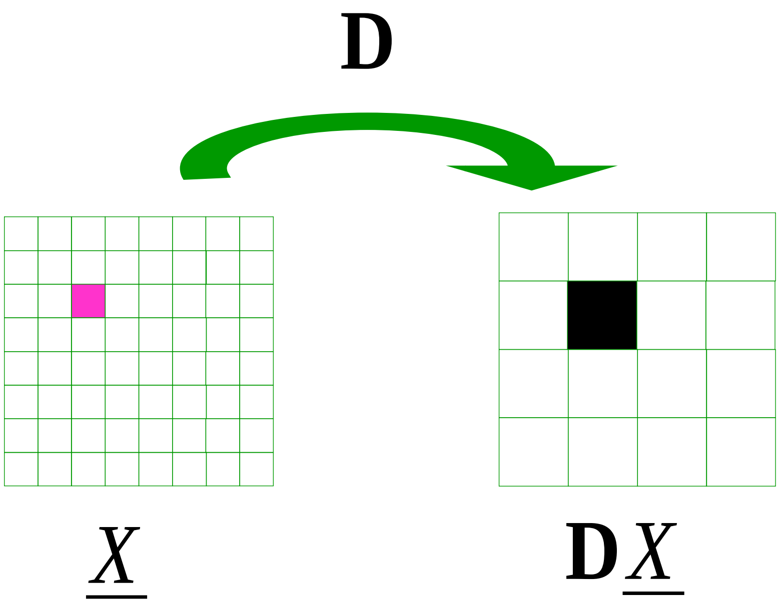
Sampling

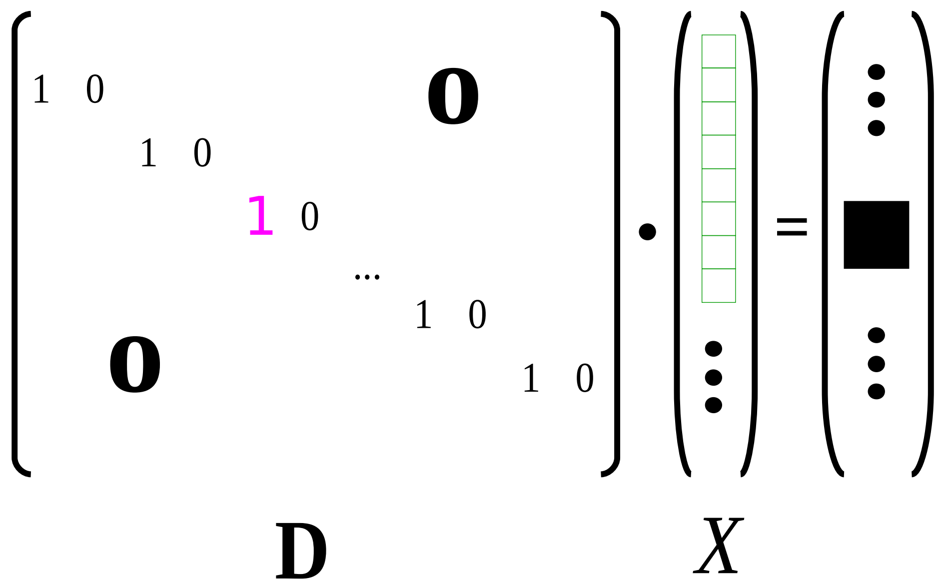
- Pixel operation consists of area integration followed by decimation
- D is the decimation only
- Usually

$$\mathbf{D}_k = \mathbf{D}$$

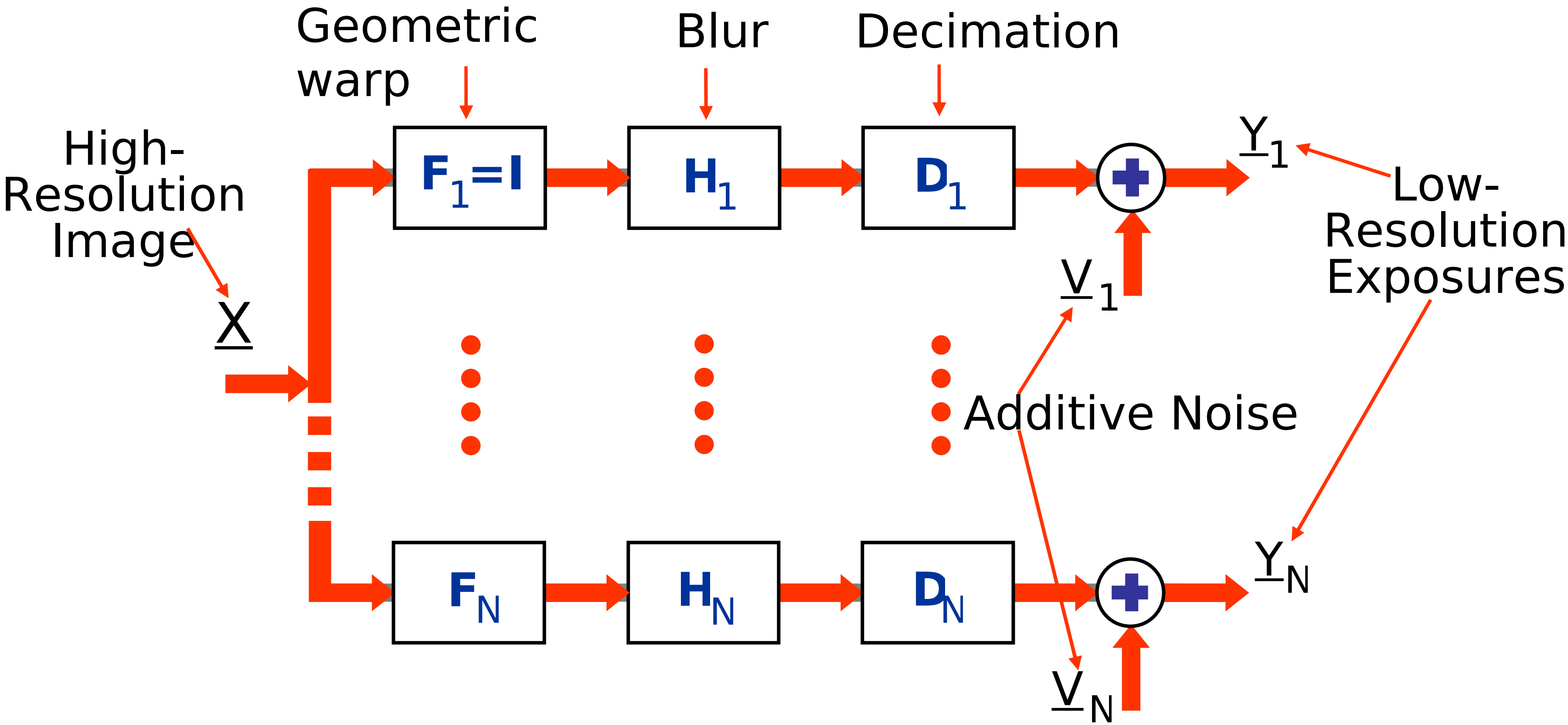
Decimation

Can be modeled as a linear operation $\mathbf{D}\underline{X}$



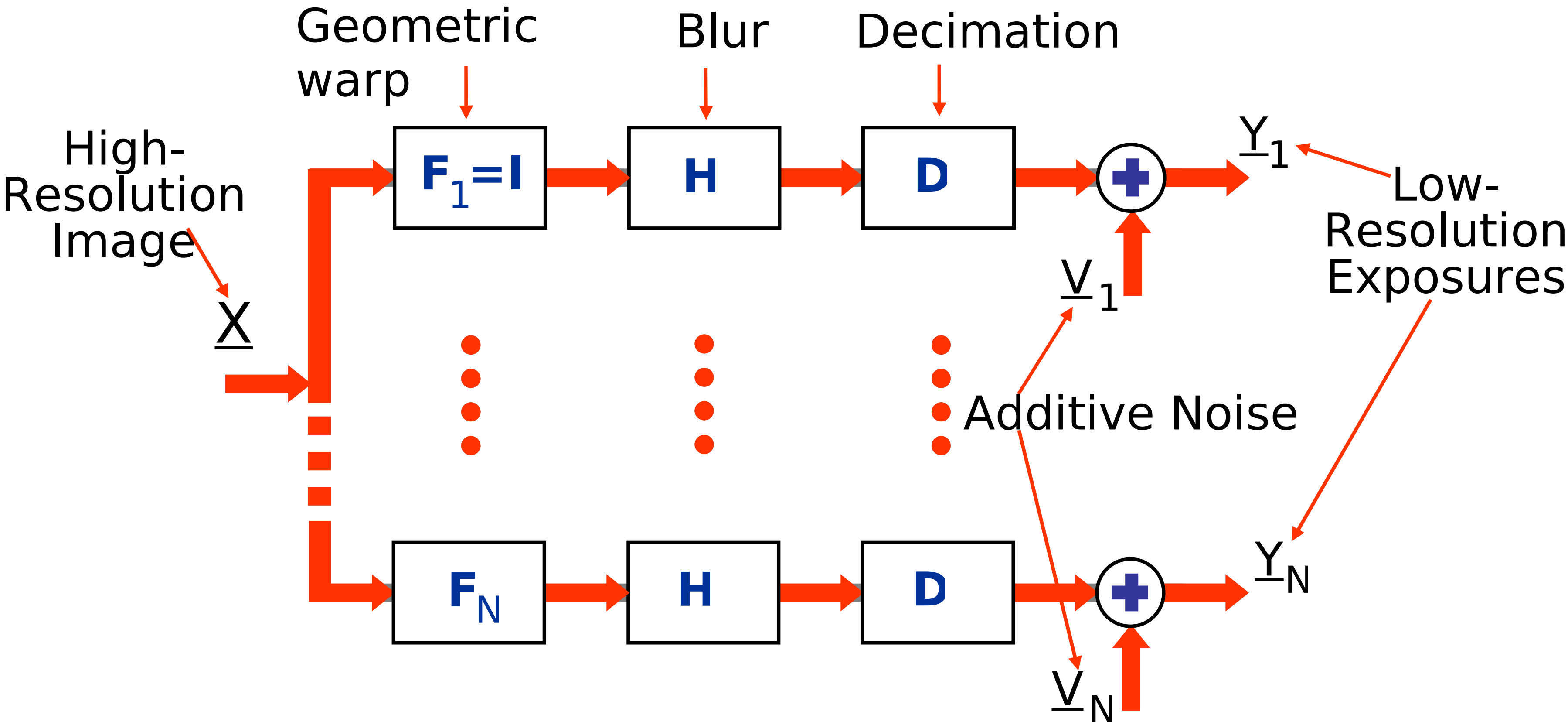


Super-Resolution - Model



$$\left\{ \underline{Y}_k = \mathbf{D}_k \mathbf{H}_k \mathbf{F}_k \underline{X} + \underline{V}_k, \quad \underline{V}_k \sim \mathbf{N}\left\{0, \sigma_n^2\right\} \right\}_{k=1}^{N}$$

Simplified Model



$$\left\{ \underline{Y}_k = \mathbf{DHF}_k \underline{X} + \underline{V}_k , \quad \underline{V}_k \sim \mathbf{N}\left\{ 0, \sigma_n^2 \right\} \right\}_{k=1}^{N}$$

The Super-Resolution Problem

$$\underline{Y}_k = \mathbf{DHF}_k \underline{X} + \underline{V}_k, \quad \underline{V}_k \sim \mathbf{N}\left\{ 0, \sigma_n^2 \right\}$$

- Given
  - $\underline{Y}_k$ – The measured images (noisy, blurry, down-sampled ..)
  - **H** – The blur can be extracted from the camera characteristics
  - **D** – The decimation is dictated by the required resolution ratio
  - $\mathbf{F}_k$ – The warp can be estimated using motion estimation
  - $\boldsymbol{\sigma}_{\mathbf{n}}$ – The noise can be extracted from the camera / image

- Recover
  - $\underline{X}$ – HR image

The Model as One Equation

$$\underline{Y} = \begin{bmatrix} \underline{Y}_1 \\ \underline{Y}_2 \\ \vdots \\ \underline{Y}_N \end{bmatrix} = \begin{bmatrix} \mathbf{D}_1\mathbf{H}_1\mathbf{F}_1 \\ \mathbf{D}_2\mathbf{H}_2\mathbf{F}_2 \\ \vdots \\ \mathbf{D}_N\mathbf{H}_N\mathbf{F}_N \end{bmatrix} \underline{X} + \begin{bmatrix} \underline{V}_1 \\ \underline{V}_2 \\ \vdots \\ \underline{V}_N \end{bmatrix} = \mathbf{G}\underline{X} + \underline{V}$$

r = resolution factor = 4
MXM = size of the frames = 1000X1000
N = number of frames = 10

$\underline{Y}$ of size $\left[NM^2 \times 1\right]$ =[10M×1]

$\mathbf{G}$ of size $\left[NM^2 \times r^2M^2\right]$ =[10M×16M]

$\underline{X}, \underline{V}$ of size $\left[r^2M^2 \times 1\right]$ =[16M×1]

Linear algebra notation is intended only to develop algorithm

SR - Solutions

- Maximum Likelihood (ML):

$$\underline{X} = \arg\min_{\underline{X}} \sum_{k=1}^{N} \| \mathbf{DHF}_k \underline{X} - \underline{Y}_k \|^2$$

Often ill posed problem!

- Maximum Aposteriori Probability (MAP)

$$\underline{X} = \arg\min_{\underline{X}} \sum_{k=1}^{N} \| \mathbf{DHF}_k \underline{X} - \underline{Y}_k \|^2 + \lambda A\{\underline{X}\}$$

Smoothness constraint
regularization

ML Reconstruction (LS)

Minimize:

$$\varepsilon_{ML}^{2}\left(\underline{X}\right)=\sum_{k=1}^{N}\left\| \mathbf{DHF}_k\,\underline{X}-\underline{Y}_k\right\|^2$$

Thus, require:

$$\frac{\partial\,\varepsilon_{ML}^{2}\left(\underline{X}\right)}{\partial\,\underline{X}}=2\sum_{k=1}^{N}\mathbf{F}_k^T\mathbf{H}^T\mathbf{D}^T\left(\mathbf{DHF}_k\,\underline{\hat{X}}-\underline{Y}_k\right)=0$$

$$\underbrace{\sum_{k=1}^{N}\mathbf{F}_k^T\mathbf{H}^T\mathbf{D}^T\mathbf{DHF}_k}_{\mathbf{A}}\cdot\underline{\hat{X}}=\underbrace{\sum_{k=1}^{N}\mathbf{F}_k^T\mathbf{H}^T\mathbf{D}^T\underline{Y}_k}_{\mathbf{B}}$$

$$\mathbf{A}\,\underline{\hat{X}}=\mathbf{B}$$

LS - Iterative Solution

- Steepest descent

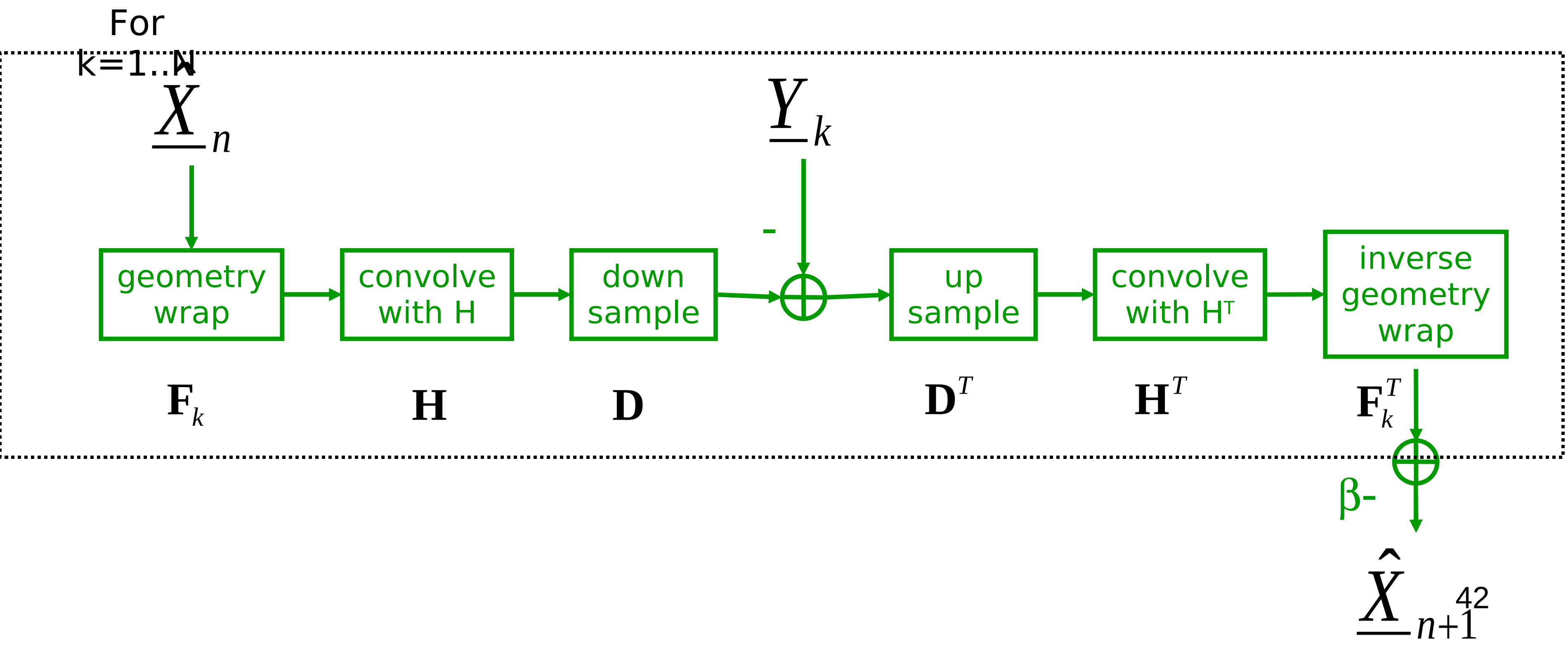
$$\underline{\hat{X}}_{n+1} = \underline{\hat{X}}_n - \beta \sum_{k=1}^{N} \underbrace{\mathbf{F}_k^T \mathbf{H}^T \mathbf{D}^T}_{\text{Back projection}} \underbrace{\left(\mathbf{D}\mathbf{H}\mathbf{F}_k \underline{\hat{X}}_n - \underline{Y}_k \right)}_{\text{Simulated error}}$$

All the above operations can be interpreted as operations performed on images.

There is no actual need to use the Matrix-Vector notations as shown here.

LS - Iterative Solution

- Steepest descent $$\underline{\hat{X}}_{n+1} = \underline{\hat{X}}_n - \beta \sum_{k=1}^{N} \mathbf{F}_k^T \mathbf{H}^T \mathbf{D}^T \left(\mathbf{D}\mathbf{H}\mathbf{F}_k \underline{\hat{X}}_n - \underline{Y}_k \right)$$



Robust Reconstruction

- Cases of measurements outlier:
  - Some of the images are irrelevant
  - Error in motion estimation
  - Error in the blur function
  - General model mismatch

Robust Reconstruction

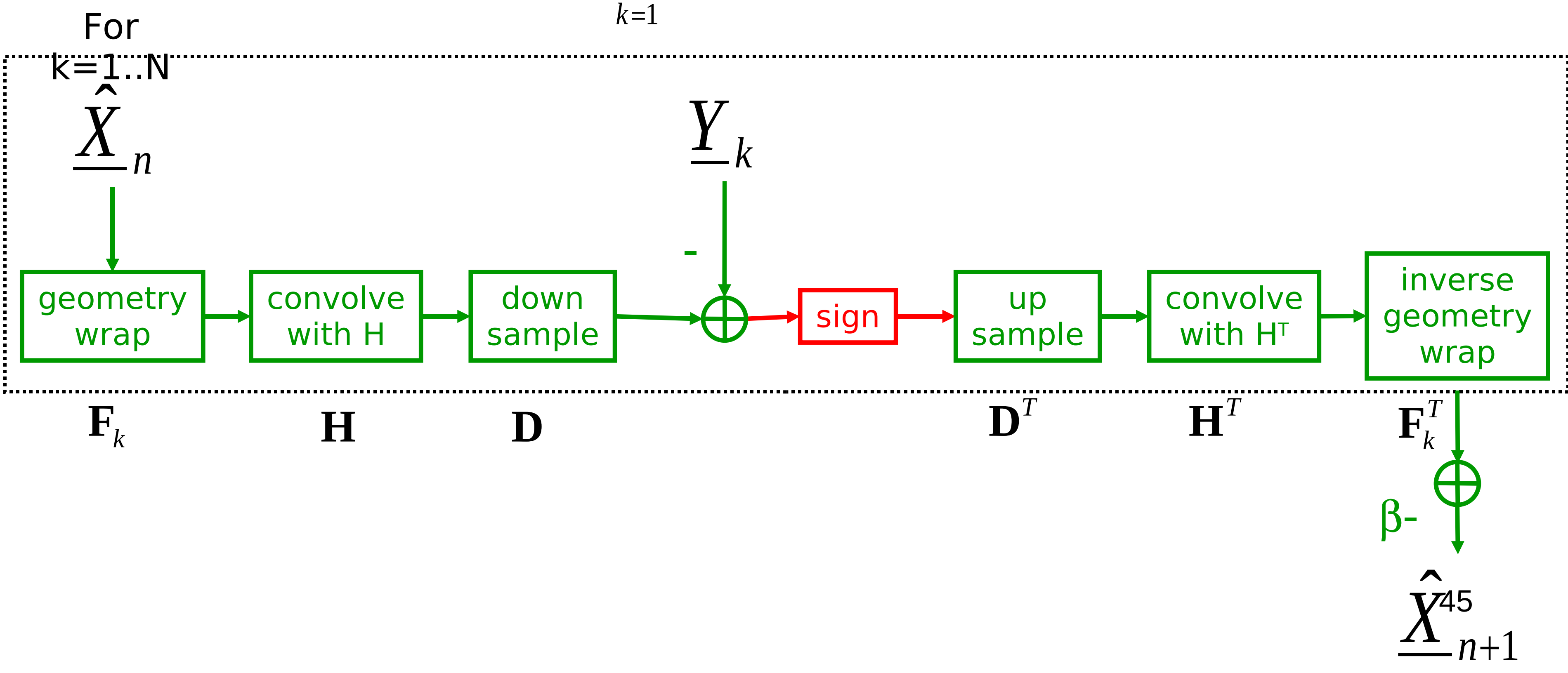
Minimize: $$\varepsilon^2\left(\underline{X}\right)=\sum_{k=1}^{N}\left\|\mathbf{DHF}_k\underline{X}-\underline{Y}_k\right\|_1$$

$$\underline{\hat{X}}_{n+1}=\underline{\hat{X}}_n-\beta\sum_{k=1}^{N}\mathbf{F}_k^T\mathbf{H}^T\mathbf{D}^T\,\mathrm{sign}\left(\mathbf{DHF}_k\underline{\hat{X}}_n-\underline{Y}_k\right)$$

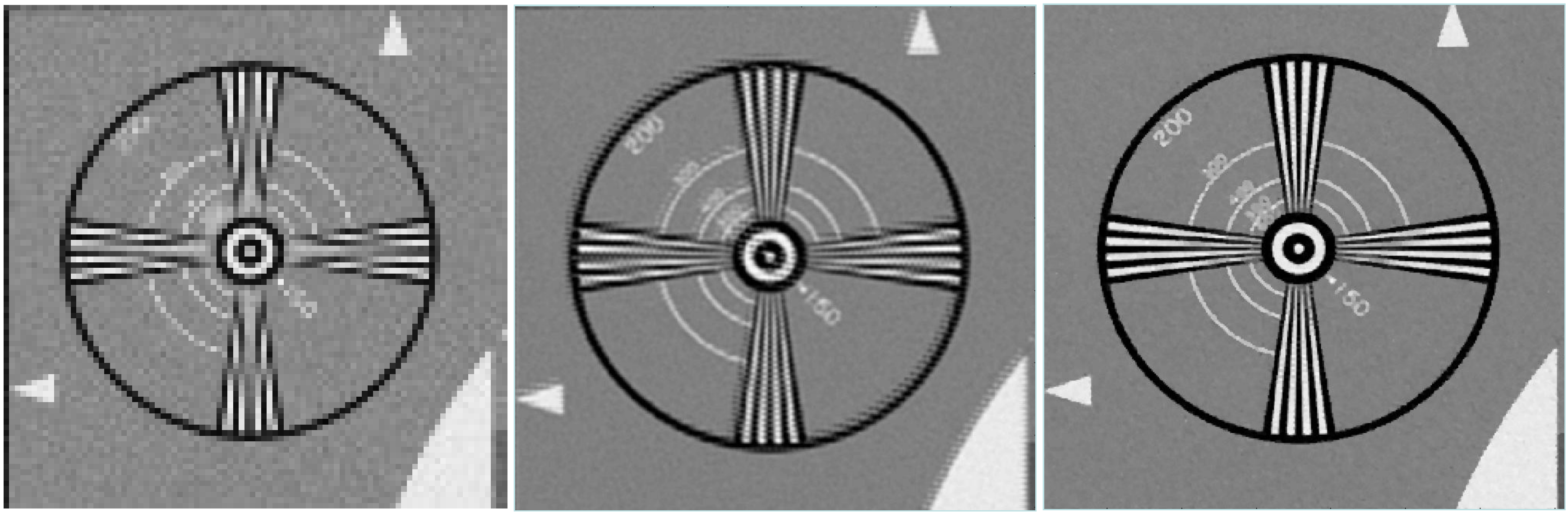
Robust Reconstruction

- Steepest descent

$$\underline{\hat{X}}_{n+1} = \underline{\hat{X}}_n - \beta \sum_{k=1}^{N} \mathbf{F}_k^T \mathbf{H}^T \mathbf{D}^T \operatorname{sign}\left(\mathbf{D}\mathbf{H}\mathbf{F}_k \underline{\hat{X}}_n - \underline{Y}_k \right)$$



Example – Registration Error



L_2 norm based

L_1 norm based

20 images, ratio 1:4

MAP Reconstruction

$$\varepsilon^2_{MAP}\left(\underline{X}\right) = \sum_{k=1}^{N} \left\| \mathbf{DHF}_k \underline{X} - \underline{Y}_k \right\|^2 + \lambda A\left\{\underline{X}\right\}$$

- Regularization term:
  - Tikhonov cost function

$$A_T\left\{\underline{X}\right\} = \left\|\Gamma \underline{X}\right\|^2$$

  - Total variation

$$A_{TV}\left\{\underline{X}\right\} = \left\|\nabla \underline{X}\right\|_1$$

  - Bilateral filter

$$A_B\left\{\underline{X}\right\} = \sum_{l=-P}^{P} \sum_{m=-P}^{P} \alpha^{|l|+|m|} \left\| \underline{X} - S_x^l S_y^m \underline{X} \right\|_1$$

Robust Estimation + Regularization

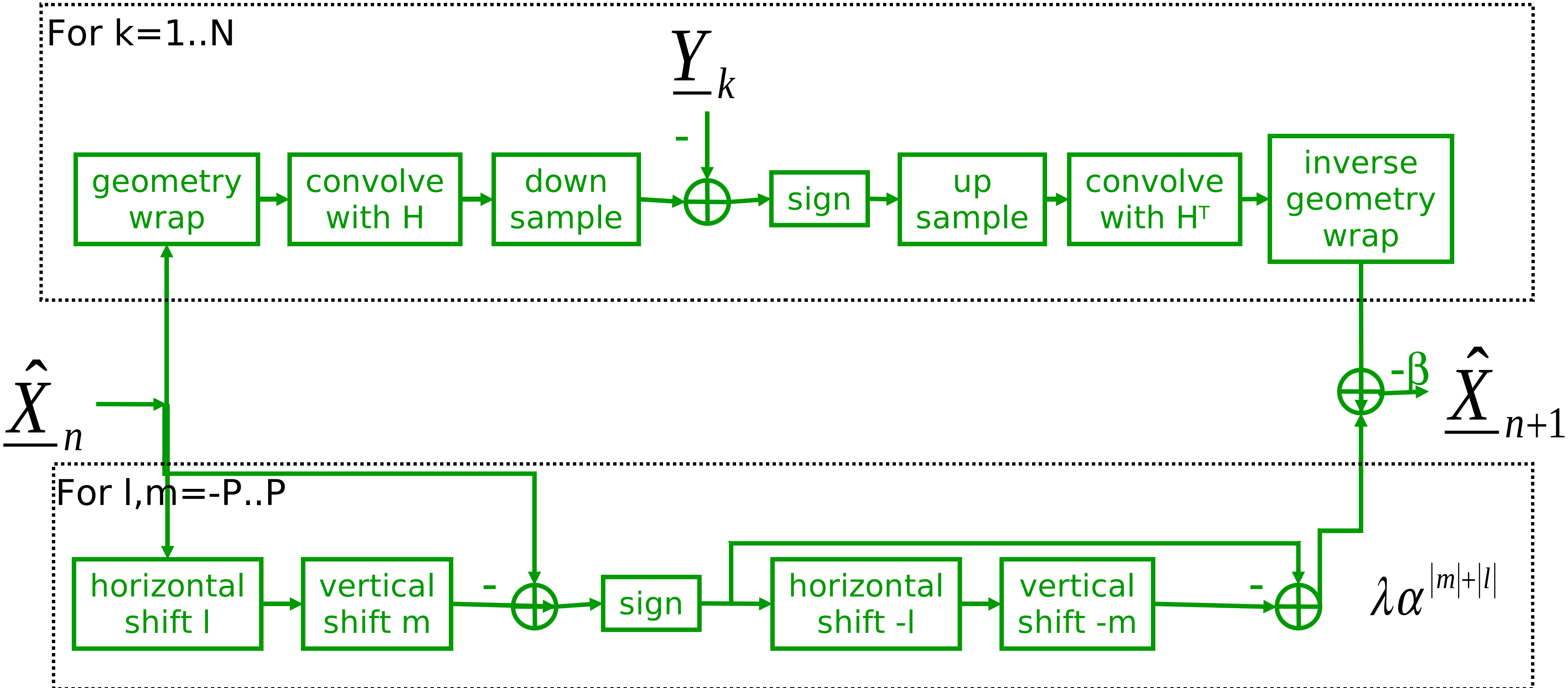
Minimize: $\varepsilon^2\left(\underline{X}\right) = \sum_{k=1}^{N} \left\| \mathbf{DHF}_k \underline{X} - \underline{Y}_k \right\|_1 + \lambda \sum_{l=-P}^{P} \sum_{m=-P}^{P} \alpha^{|l|+|m|} \left\| \underline{X} - S_x^l S_y^m \underline{X} \right\|_1$

$$\underline{\hat{X}}_{n+1} = \underline{\hat{X}}_n - \beta \left\{ \sum_{k=1}^{N} \mathbf{F}_k^T \mathbf{H}^T \mathbf{D}^T \operatorname{sign}\left(\mathbf{DHF}_k \underline{\hat{X}}_n - \underline{Y}_k \right) \right.$$

$$\left. + \lambda \sum_{l=-P}^{P} \sum_{m=-P}^{P} \alpha^{|l|+|m|} \left[I - S_x^{-l} S_y^{-m} \right] \operatorname{sign}\left(\underline{\hat{X}}_n - S_x^l S_y^m \underline{\hat{X}}_n \right) \right\}$$

Robust Estimation + Regularization

$$\underline{\hat{X}}_{n+1} = \underline{\hat{X}}_n - \beta \left\{ \sum_{k=1}^{N} \mathbf{F}_k^T \mathbf{H}^T \mathbf{D}^T \operatorname{sign}\left(\mathbf{DHF}_k \underline{\hat{X}}_n - \underline{Y}_k\right) + \lambda \sum_{l=-P}^{P} \sum_{m=-P}^{P} \alpha^{|l|+|m|} \left[I - S_x^{-l} S_y^{-m}\right] \operatorname{sign}\left(\underline{\hat{X}}_n - S_x^l S_y^m \underline{\hat{X}}_n\right) \right\}$$



From Farisu at al. IEEE trans. On Image Processing, 04

Example

- 8 frames
- Resolution factor of 4

From Farisu at al. IEEE trans. On Image Processing, 04

Conclusions

- Optimization is essential in many problems
- SR using different approaches
- Selection of the algorithm

From Farisu at al. IEEE trans. On Image Processing, 04