Super Resolution

Laboratoire d'InfoRmatique en Image et Systèmes d'information
LIRIS UMR 5205 CNRS/INSA de Lyon/Université Claude Bernard Lyon 1/Université Lumière Lyon 2/Ecole Centrale de Lyon
Université Claude Bernard Lyon 1, bâtiment Nautibus
43, boulevard du 11 novembre 1918 — F-69622 Villeurbanne cedex
http://liris.cnrs.fr

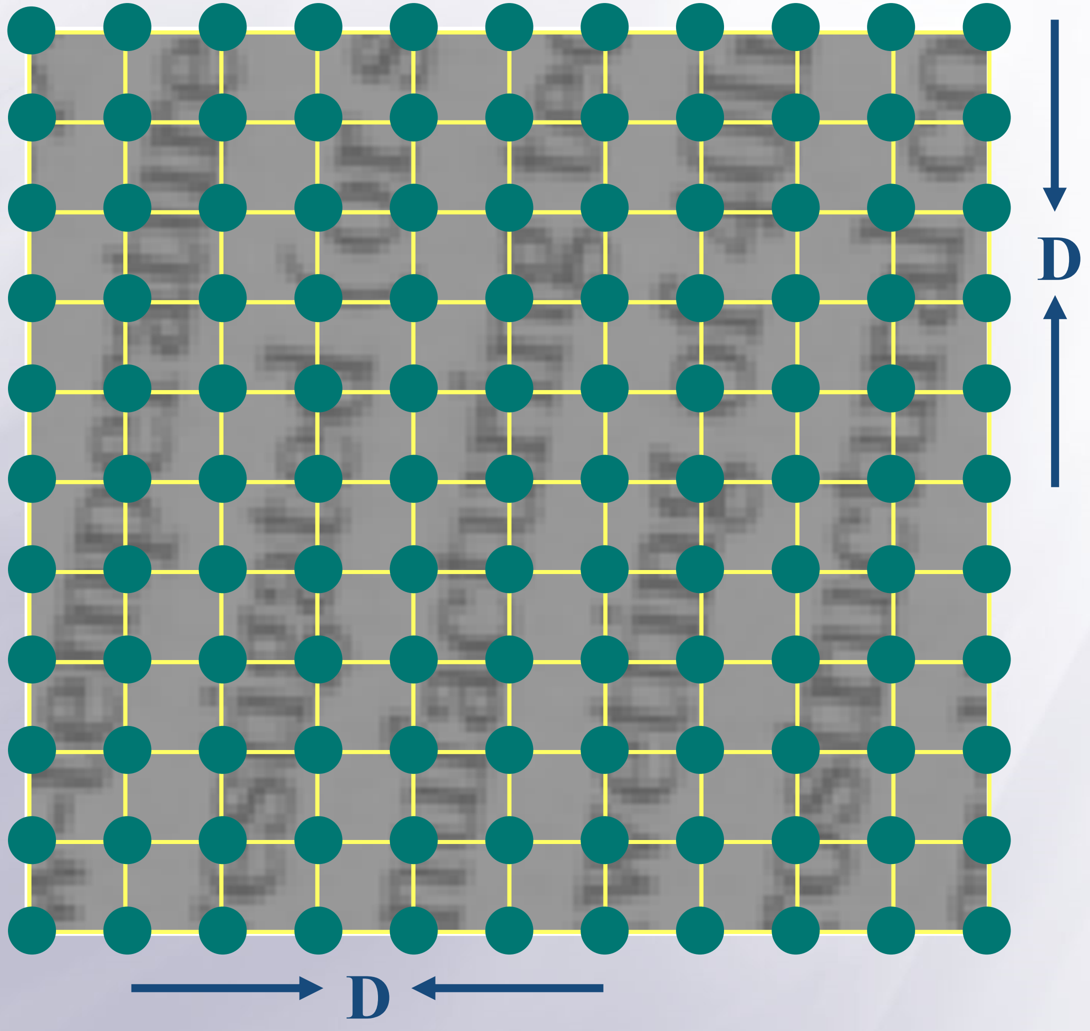
June 11th 2010

Intuition

- Combining multiple incomplete samples of an information, through an adequate process, leads to a more complete and accurate information.

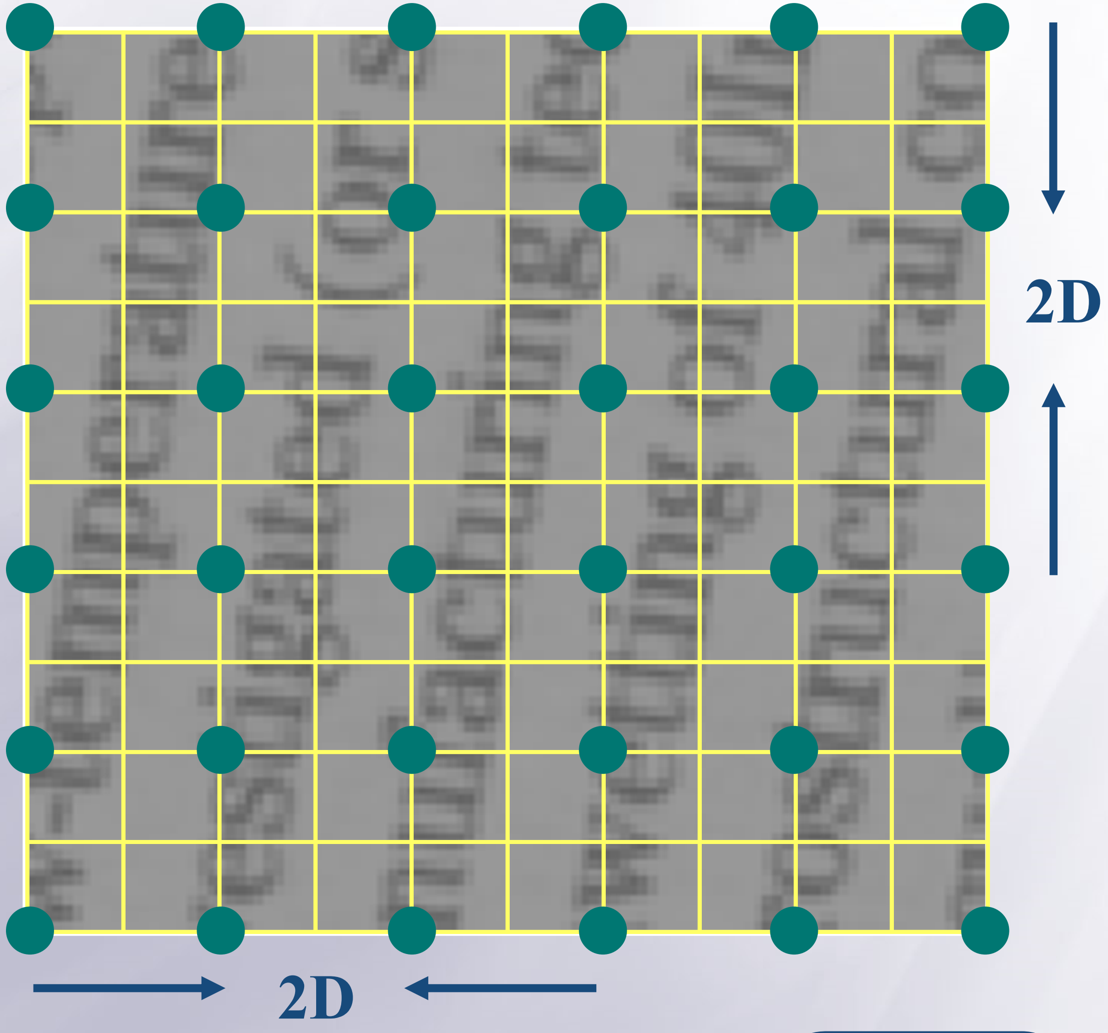
Intuition

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



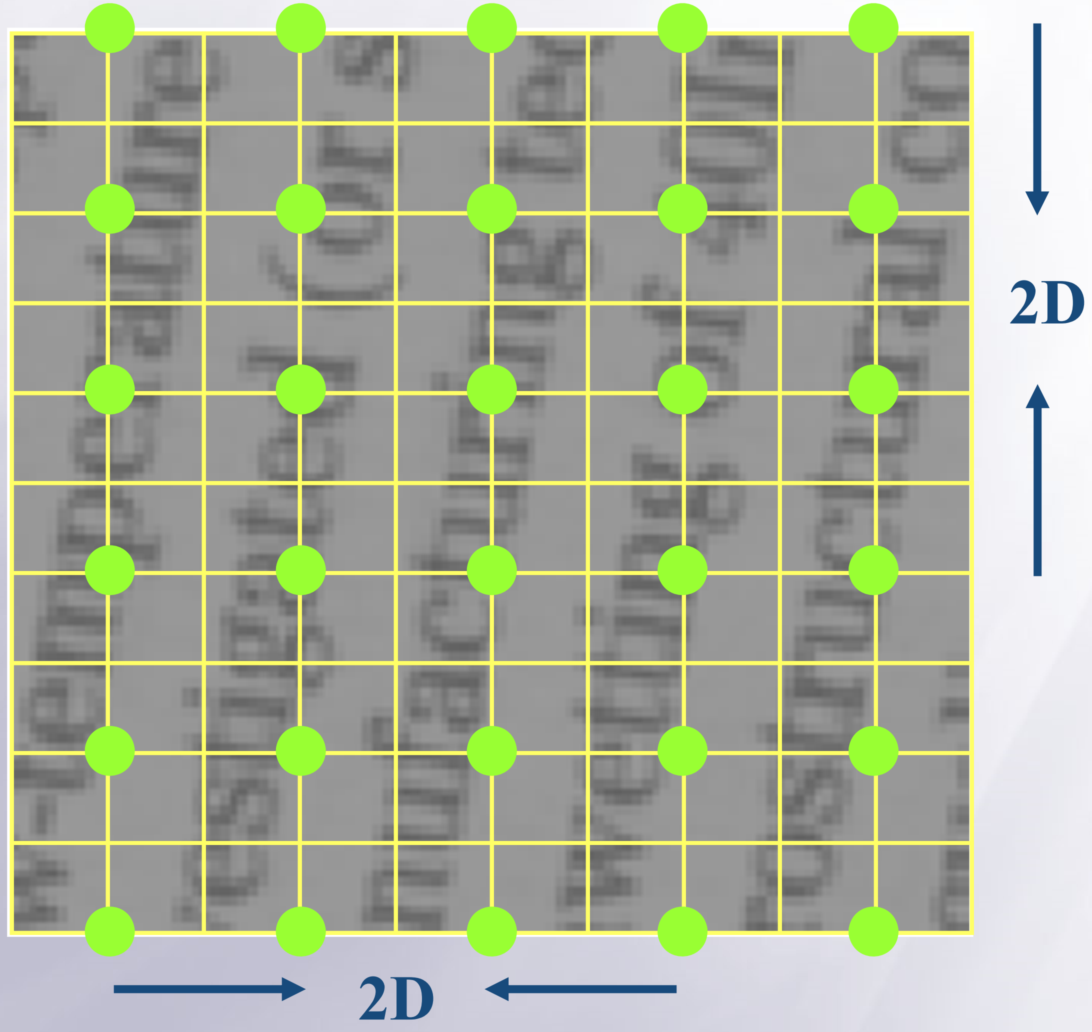
Intuition

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



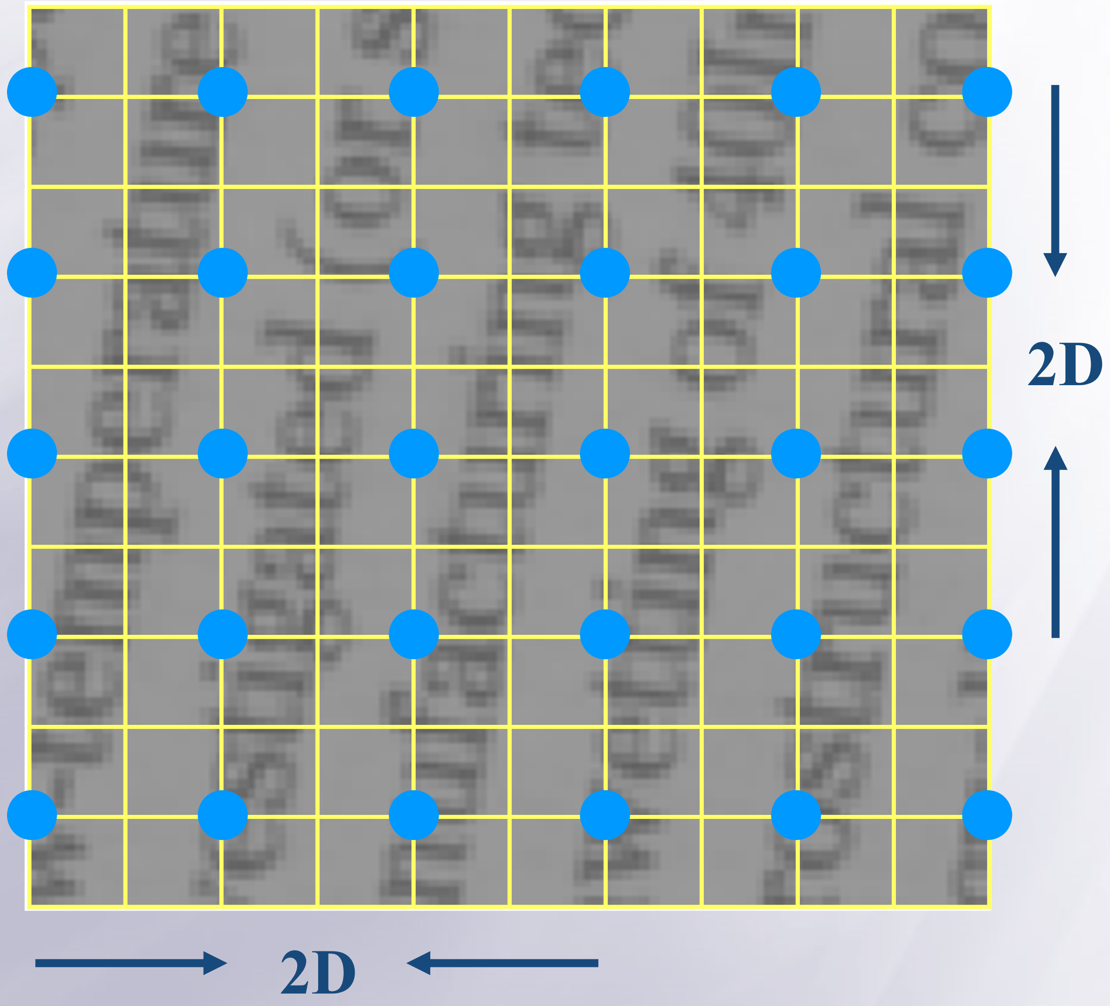
Intuition

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



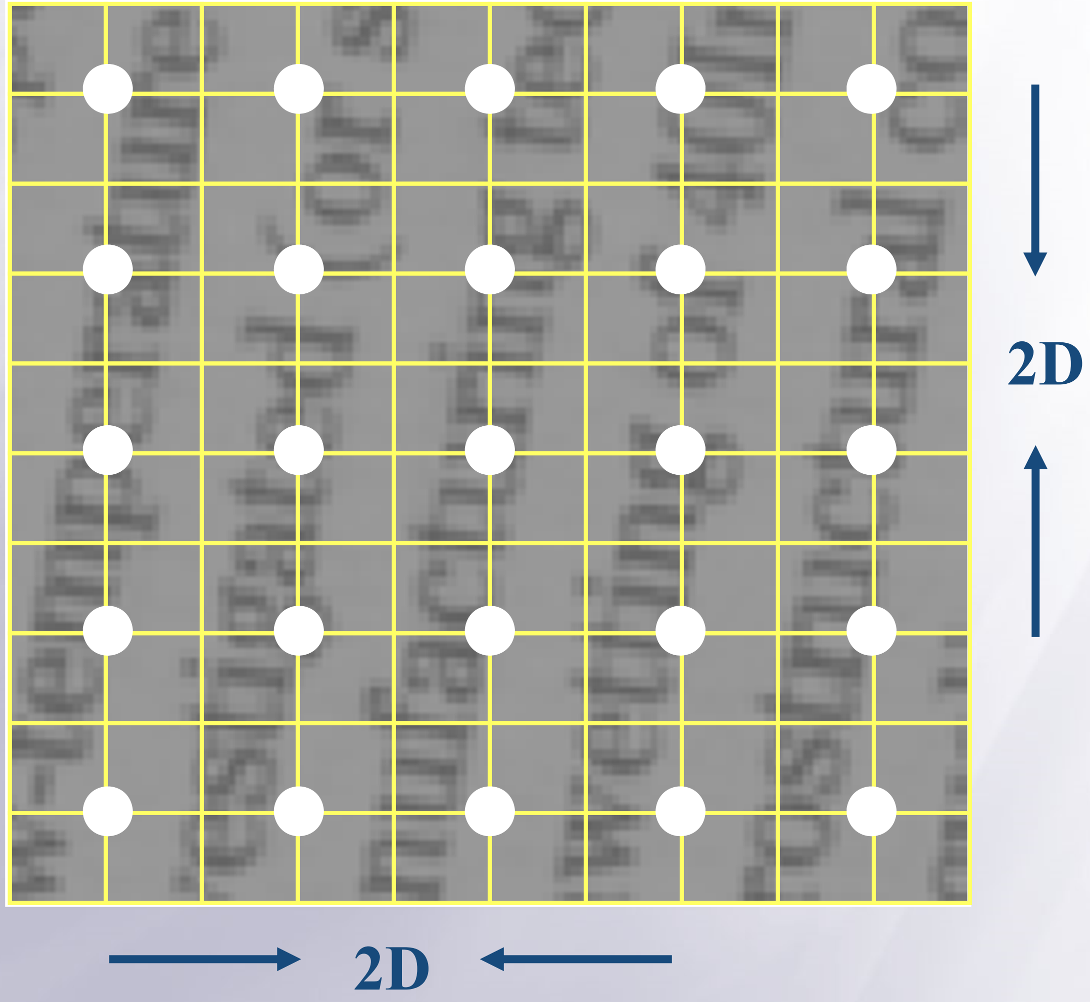
Intuition

Similarly, by shifting down we get a third image:



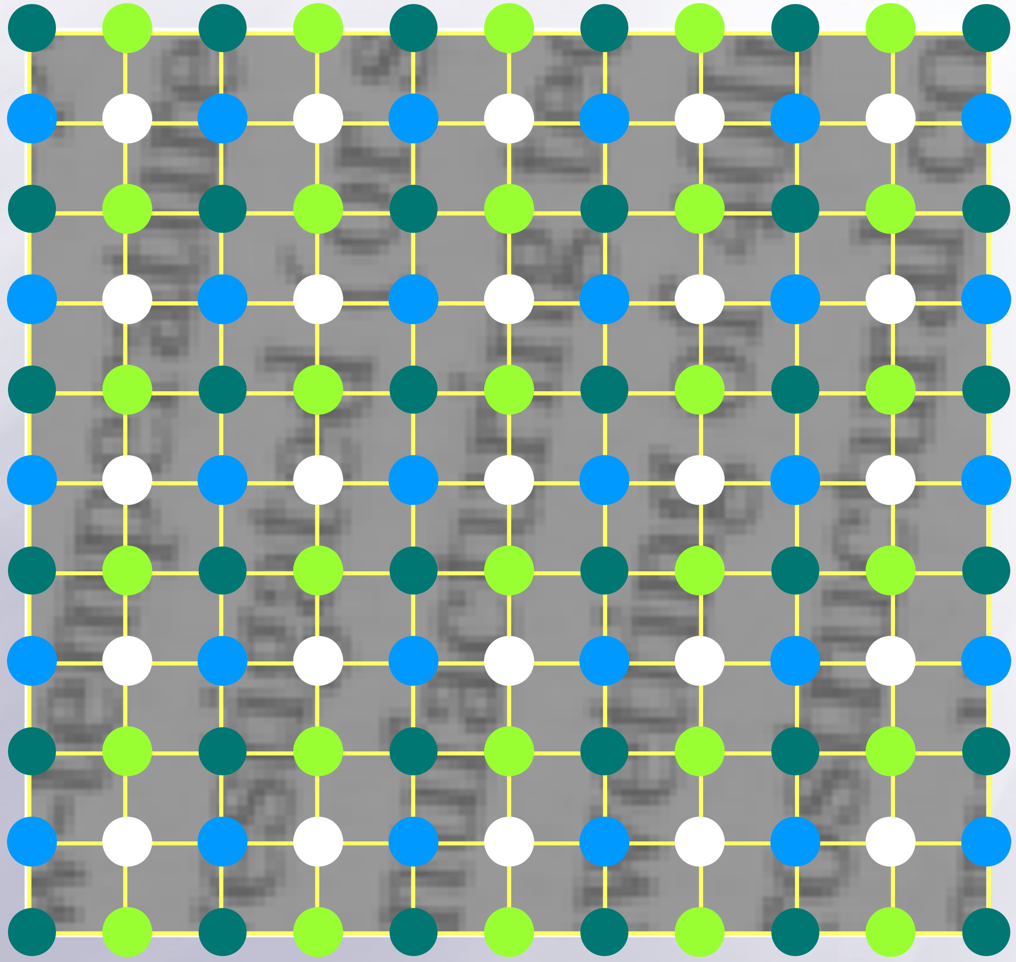
Intuition

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



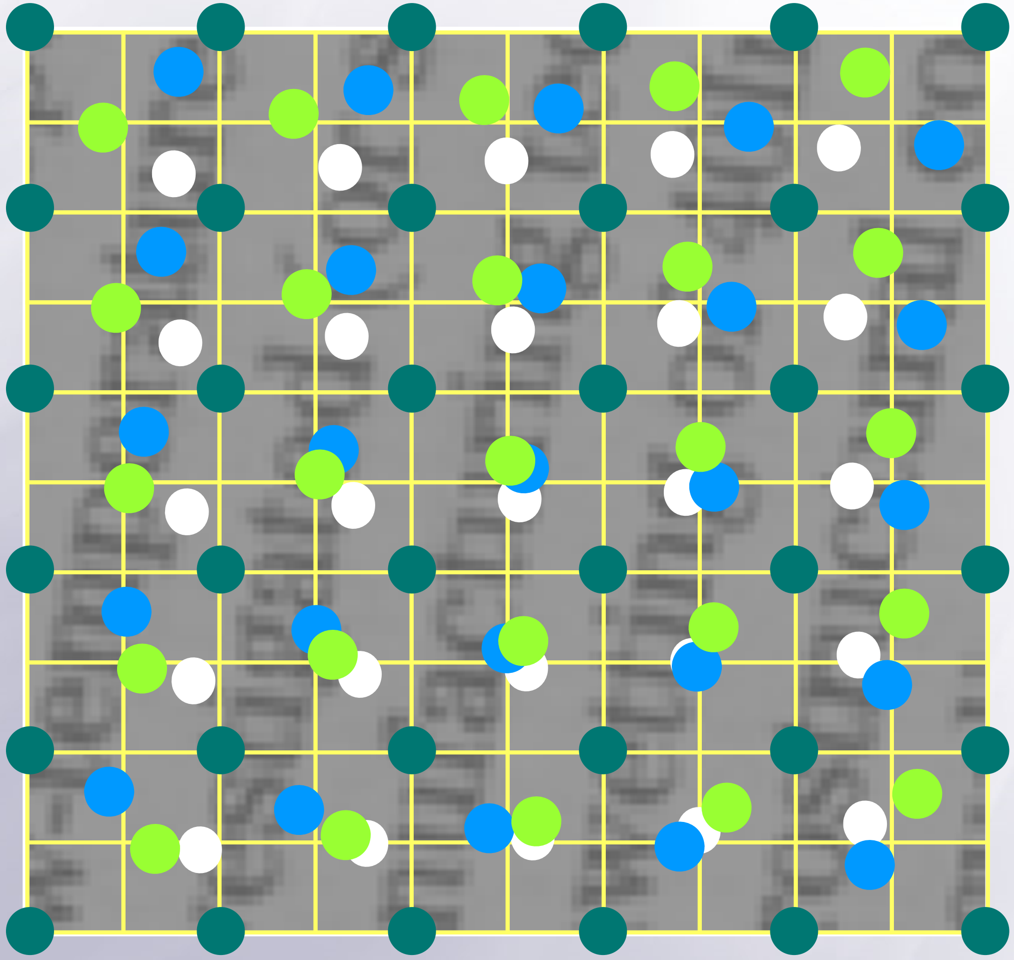
Intuition

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.

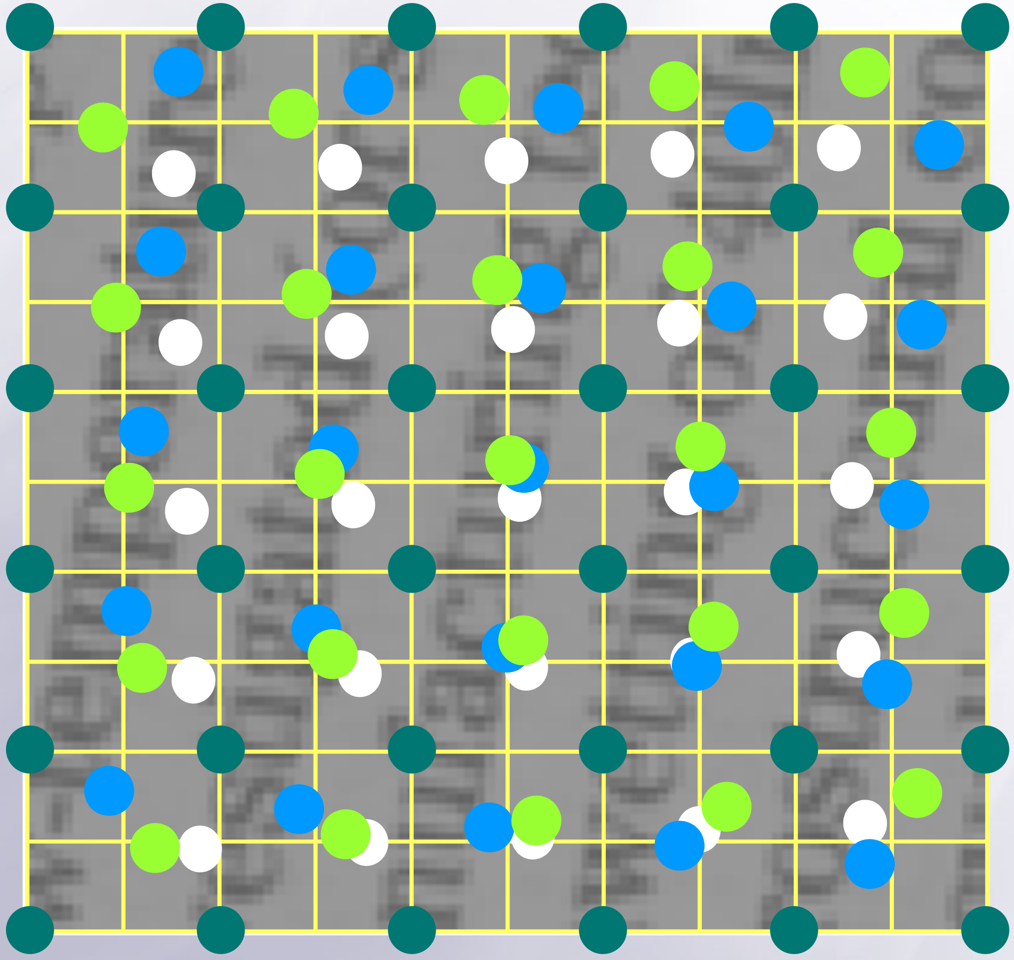
What if the camera displacement is Arbitrary ?
What if the camera rotates? Gets closer to the object (zoom)?



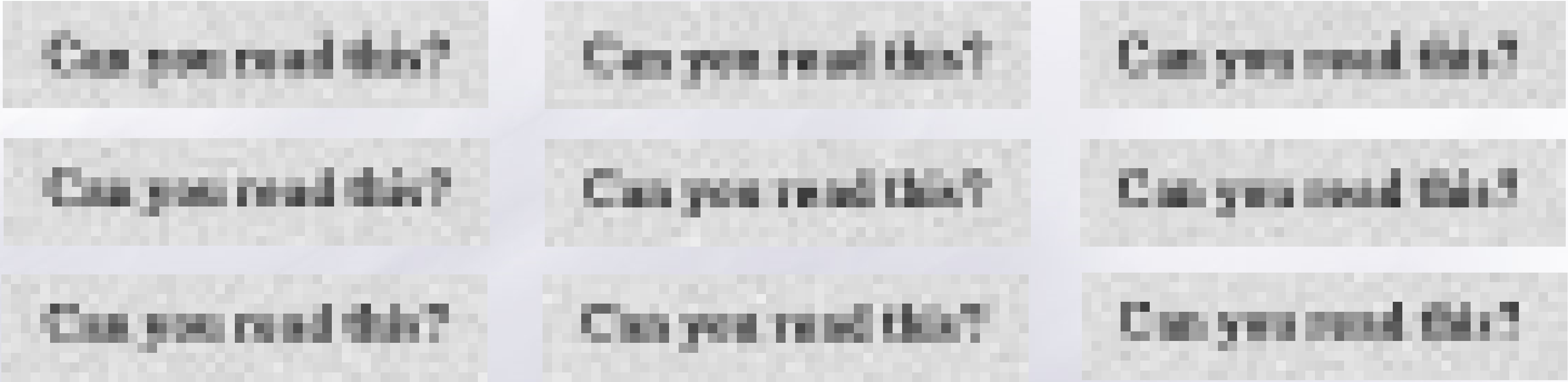
Rotation/Scale/Disp.

There is no sampling theorem covering this case

!

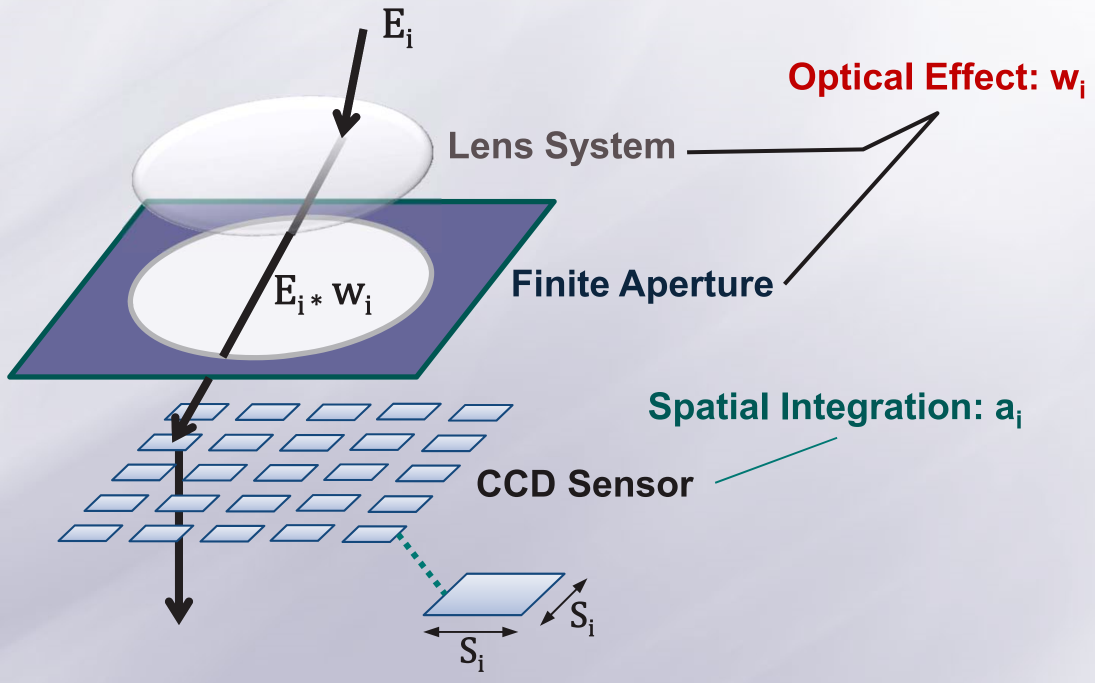


A Small Example



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

Image formation model

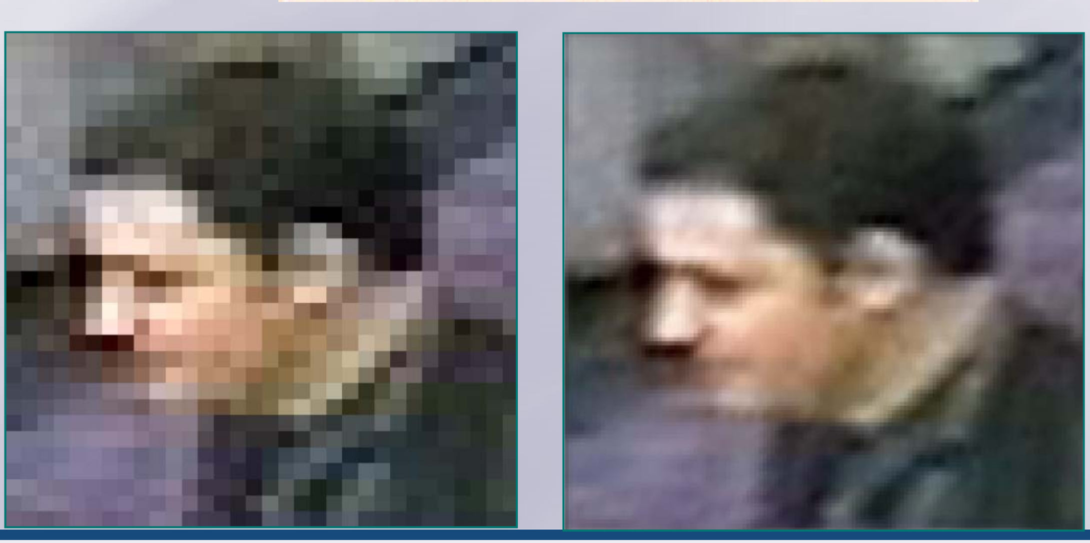


Example - Video

53 images, ratio 1:4

Example – Surveillance

40 images
ratio 1:4



Example – Enhance Mosaics

Example – Enhance Mosaics

Literature

- **Rich literature, numerous algorithms,**

Dating back to the freqeuncy domain approch of Huang an Tsai:

« Multiple image restoration and registration », *Advences in computer Vision and Image Processing*, 1/317-339, 1984.

Over all Process

- Assumptions
  - A. Some small relative motion between camera and the scene
  - B. Or other imaging parameters, such as the amount of defocus blur, variation
- Stages
  - I. Registration : pixels motion estimation from one image to others
  - II. Fusion based on some constraints from the image formation process model’s
- Results
  - Improvement over the input images
  - High frequency of components are usually not recontructed very well

Some practical experiments

- High resolution → random translation → blurred with a guassian → down-sampled
- As many LR pixels in total as the pixels in the HR image
- Providing the exact knowledge of the point spread function (Gaussian) and the sub-pixel translations to SR [Hardie et *al*.]
  - High resolution components are not well recovered
  - A decent reconstruction from input images
  - The performance gets much worse as the magnification increases

inverse problem

- Find the best model such that (at least approximately)

$$d = G(m)$$

- Operator G
  - Describes the explicit relationship between the observed data, d, and the model parameters
  - Also called **forward operator**, **observation operator**, or **observation function**
  - Represents the governing equations that relate the model parameters to the observed data

SR : Position of the problem

- A set of N low images $Lo_i(m)$, $i=1,..,N$
- $m(m,n)$ a vector in Z^2 is the pixel coordinates
- The continuous irradiance light-field: E_i
- The Point Spread Function: $PSF_i(.)$
- $x=(x,y)$ in R^2 the coordin. in the image plane of Lo_i
- Continuous image formation equation

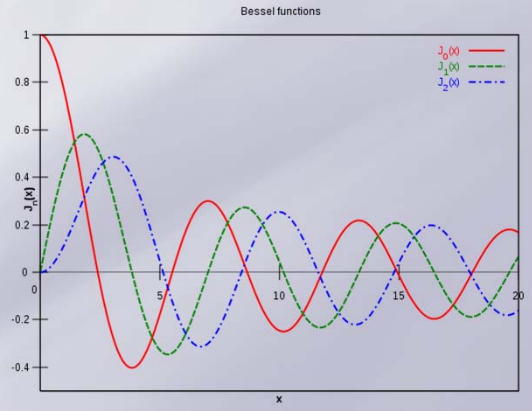
$$Lo_i(\mathbf{m}) = (E_i * PSF_i)(\mathbf{m}) = \int_{Lo_i} E_i(x).PSF(x-\mathbf{m})\,dx$$

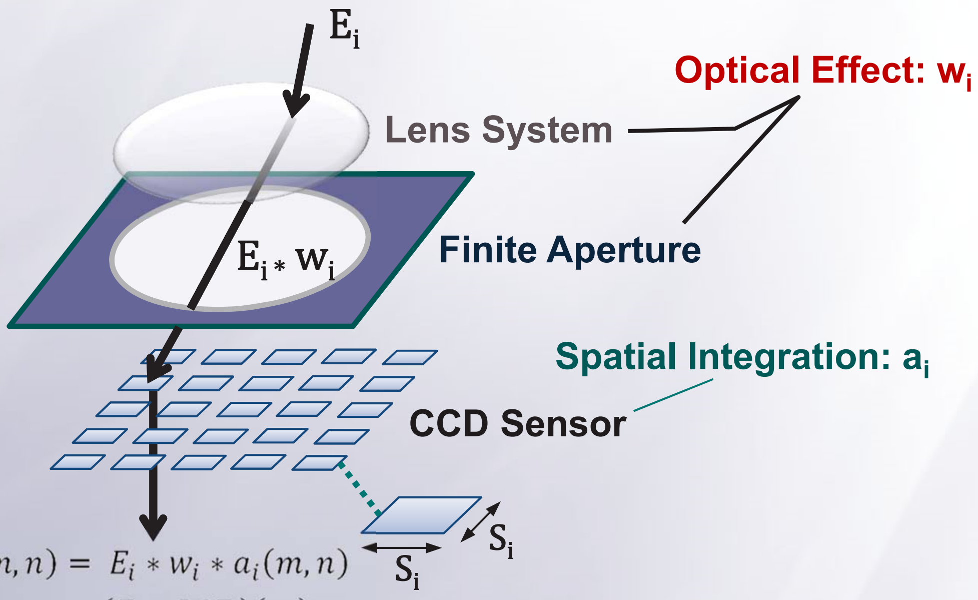
The pixel intensity is the result of convolving the irradiance function with the point-spread function and then sampling it at the discrete pixel locations

Point Spread Function Model

- $w_i(x)$ **models the blurring:**
  - defocus factor (pillbox function – low pass filter)
  - diffraction-limited optical transfer (first-order Bessel function of the first kind)

$$x^2 \frac{d^2 f}{dx^2} + x \frac{dy}{dx} + (x^2 - \alpha^2) f = 0$$





E_i

Optical Effect: w_i

Lens System

Finite Aperture

$E_{i*}w_i$

Spatial Integration: a_i

CCD Sensor

S_i

S_i

$$Lo_i(m,n) = E_i * w_i * a_i(m,n) \\ = (E_i * PSF_i)(\boldsymbol{m})$$

- and $a_i(x)$ **spatisal integration:**

$$a_i(\boldsymbol{x}) = \begin{cases} \frac{1}{S_i^2} & if\ |x| \le \frac{S_i}{2}\ and\ |y| \le \frac{S_i}{2} \\ 0 & otherwise. \end{cases}$$

Super-Resolution

- $Su(\mathbf{p})$ is a superresolved image, where $\mathbf{p}=(p,q)$ in **Z²** according to coordinate frame of $Lo_i(m)$

$$Lo_i(\boldsymbol{m}) = \int_{Su} E_i\big(r_i(\boldsymbol{z})\big)\,.\,PSF_i(r_i(\boldsymbol{z}) - \boldsymbol{m}).\left|\frac{\partial r_i}{\partial \boldsymbol{z}}\right| dz$$

p = m/M, x = ri(z) registration transformation

Determinant of the Jacobian of the registration transformation $r_i(.)$

- Assuming the registration under the pinhole model is correct, the radiance of the scene is transformed, in the same way, by all Lo_i and Su images, we have:

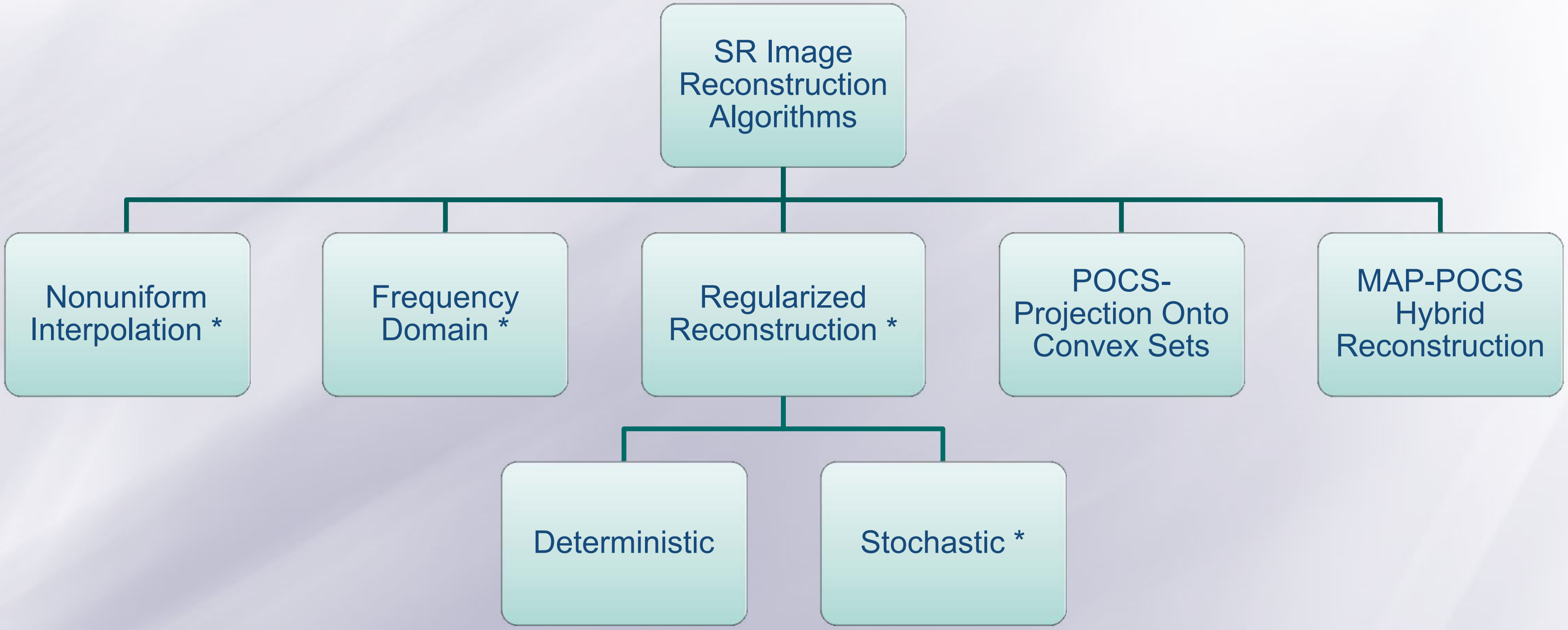
$$Lo_i(\boldsymbol{m}) = \int_{Su} E(\boldsymbol{z})\,.\,PSF_i(r_i(\boldsymbol{z}) - \boldsymbol{m}).\left|\frac{\partial r_i}{\partial \boldsymbol{z}}\right| dz$$

Observation model

Formulation

- Su the high resolution image of size $Ns = L_1\ Nl_1 \ x\ L_2\ Nl_2$ written as a vector $x = [x_1, x_2, \ldots, x_N]^T$
- L_1, L_2 down sampling facors
- Loi the low resoltion image of size $Nl = Nl_1 \ x\ Nl_2$, i in $\{1,..,N\}$ written as a vector $y_i = [y_{i,1},\ldots, y_{i,Nl}]$
- The observation model is $y_i{=}H\ D_i\ W_i\ x + n_i$ where
  - H represente the blur matrix
  - D_i the subsampling matrix
  - W_i the motion matrix
  - n_i the noise matrix
- $y_i{=}M_i\ x + n_i$, i in $\{1,..,N\}$

Categorization of super-resolution approaches



SR Image Reconstruction Algorithms

Nonuniform Interpolation *

Frequency Domain *

Regularized Reconstruction *

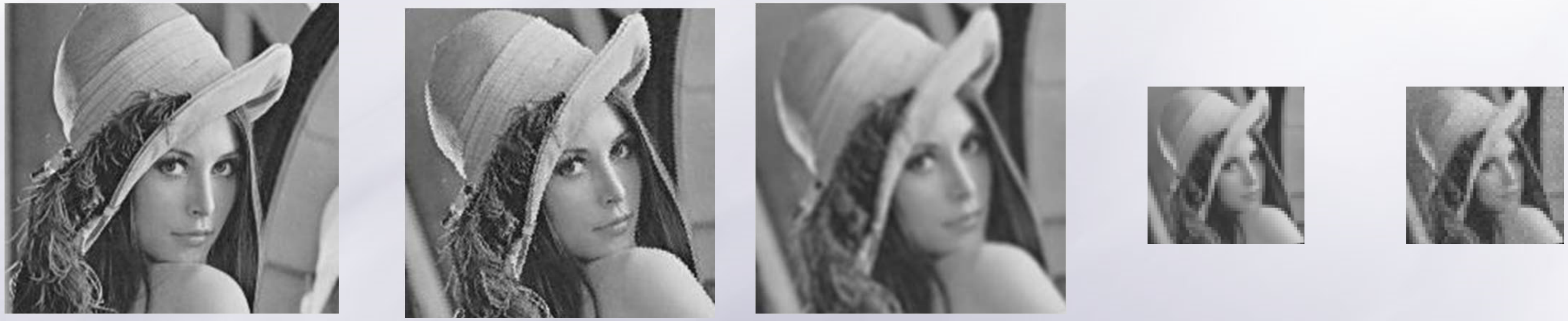
POCS- Projection Onto Convex Sets

MAP-POCS Hybrid Reconstruction

Deterministic

Stochastic *

Image Formation

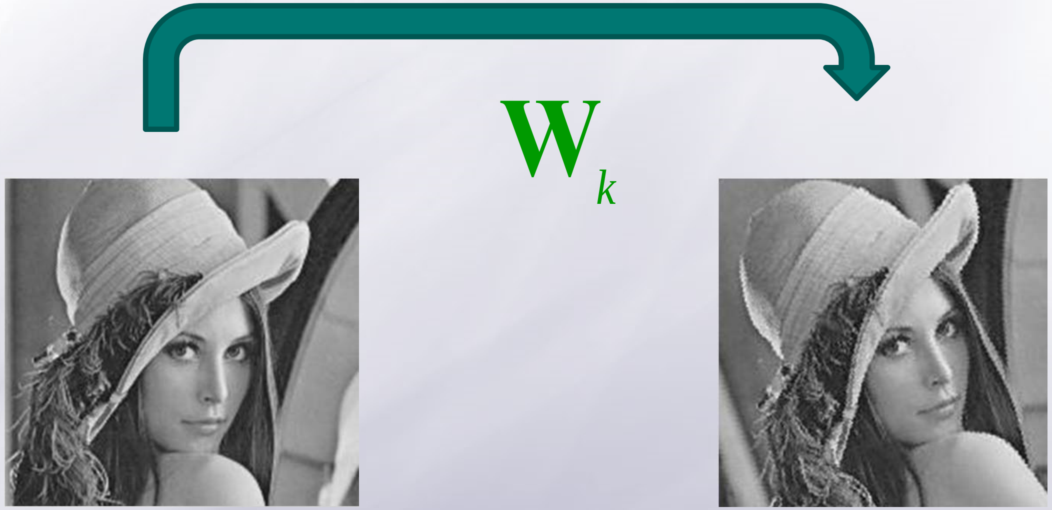


Scene	Geometric transformation	Optical Blur	Sampling	Noise
HR	$\mathbf{W}_k$	$\mathbf{H}_k$	$\mathbf{D}_k$	**LR**

Can we write these steps as linear operators?

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

Geometric Transformation



$\mathbf{W}_k$

Scene

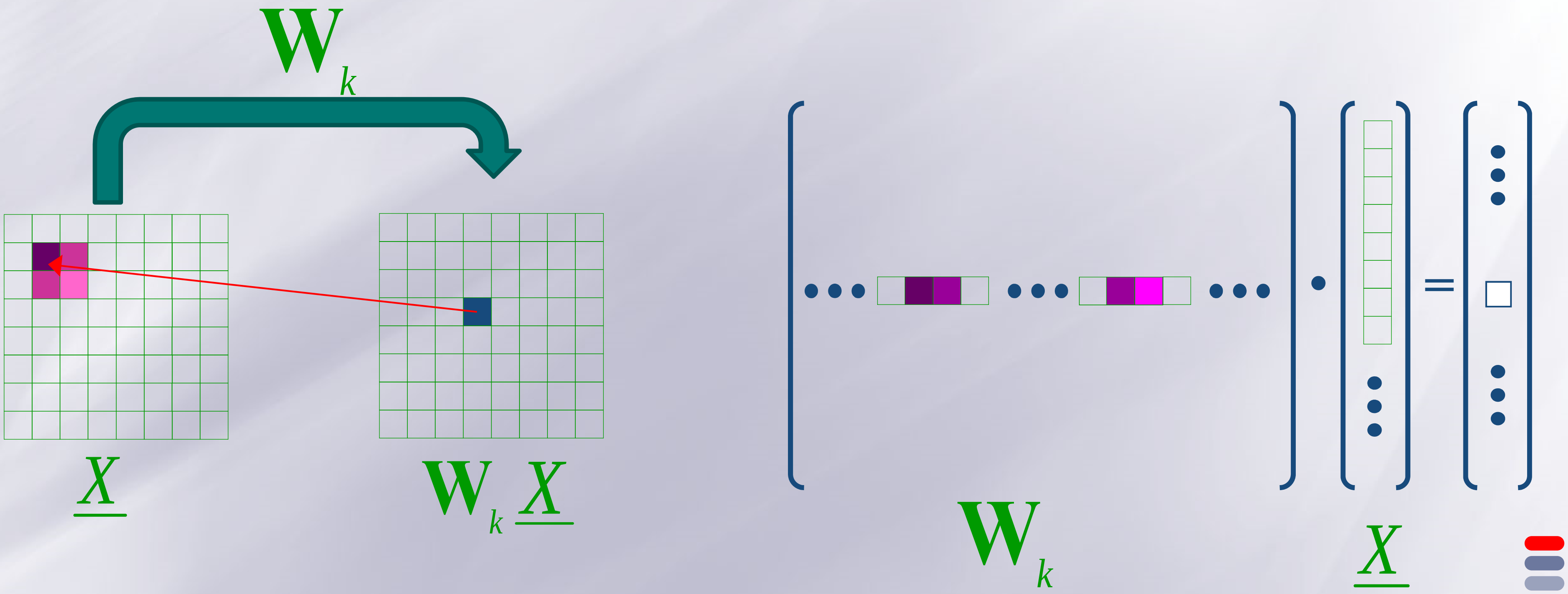
Geometric transformation

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

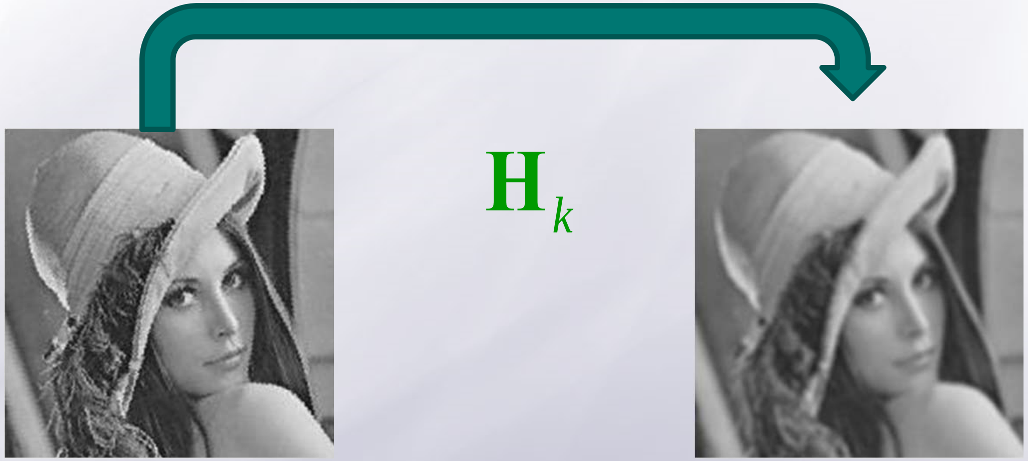
Geometric Transformation

Can be modeled as a linear operation

$$\mathbf{W}_k \underline{X}$$



Optical Blur



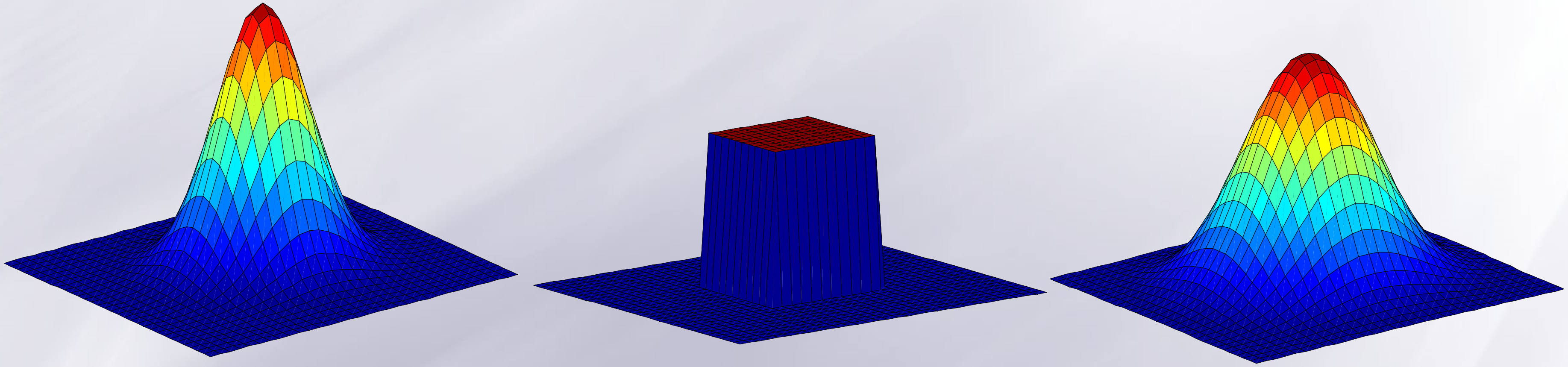
$\mathbf{H}_k$

Geometric transformation

Optical Blur

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

H

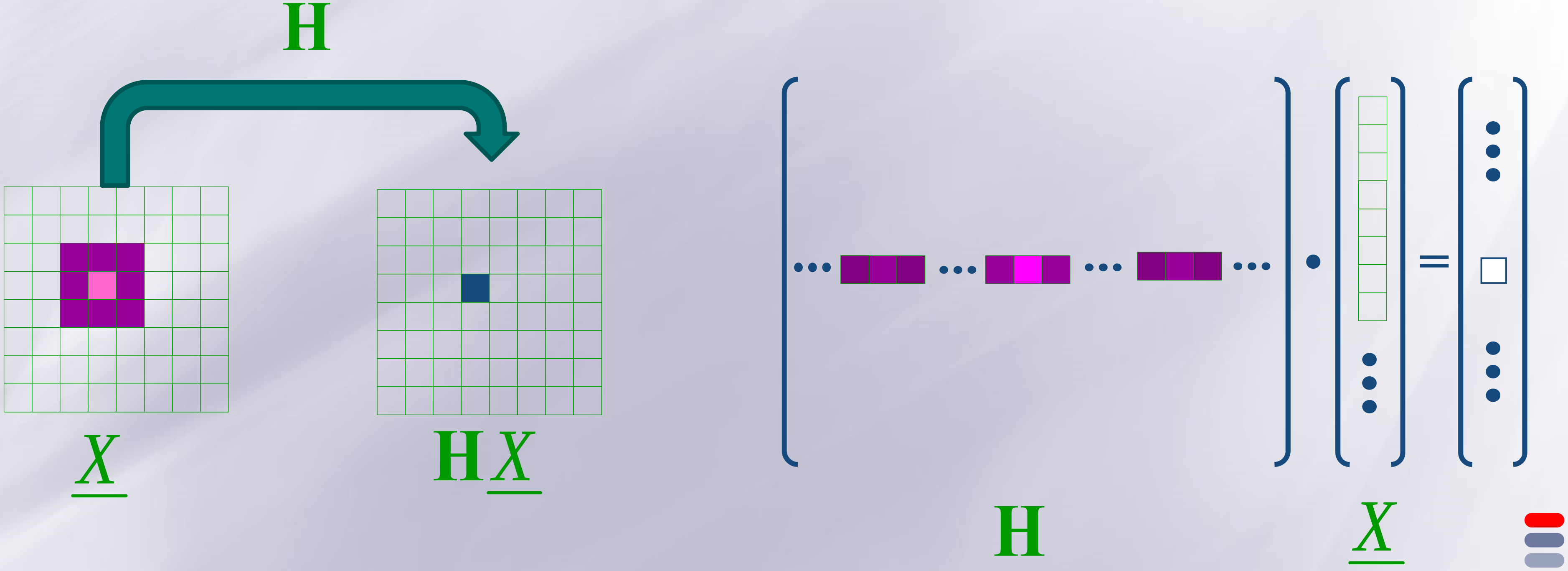


PSF * PIXEL = H

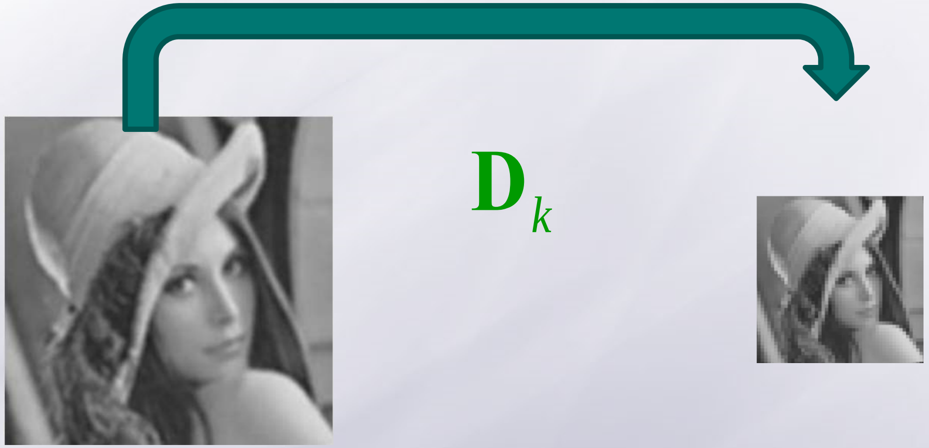
Optical Blur

Can be modeled as a linear operation

$$\mathbf{H}\underline{X}$$



Sampling



$\mathbf{D}_k$

Optical Blur

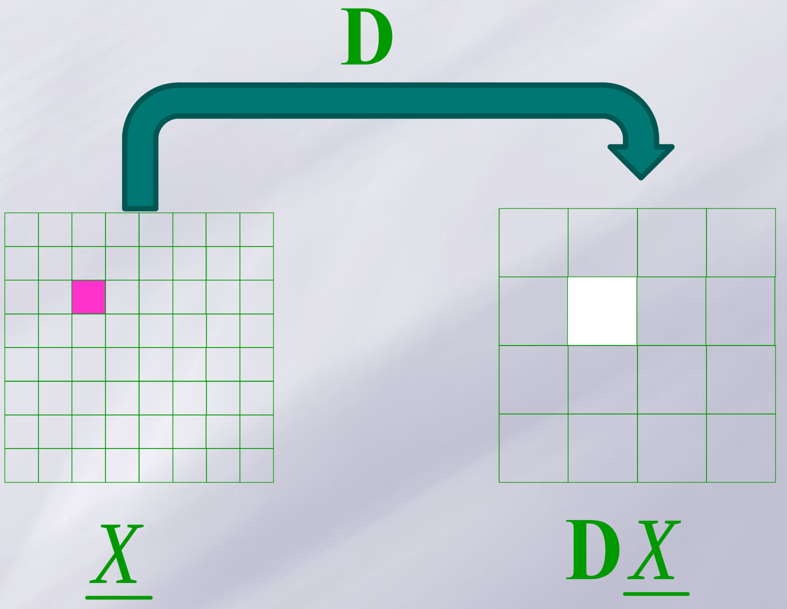
Sampling

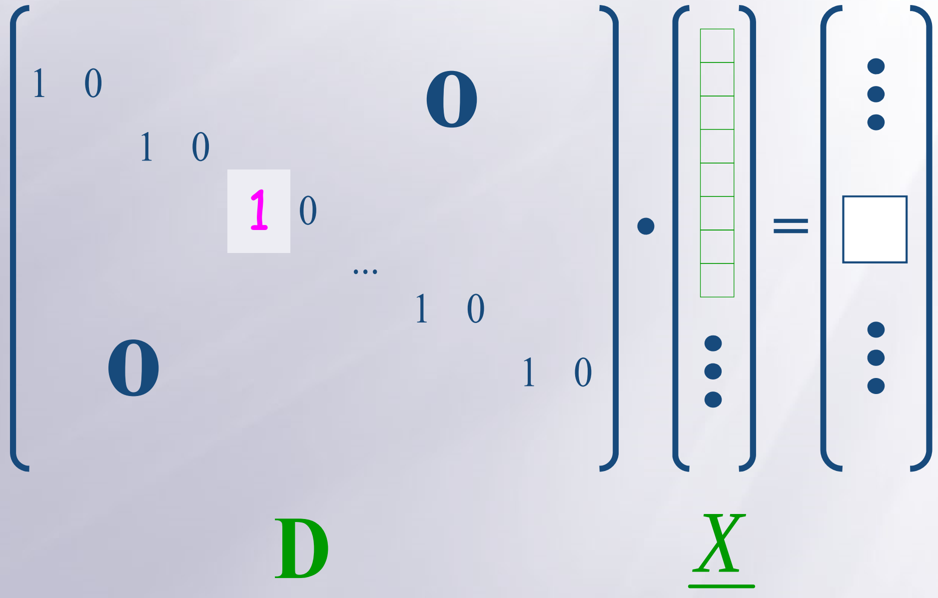
- Pixel operation consists of area integration followed by decimation
- D is the decimation only
- Usually $\mathbf{D}_k = \mathbf{D}$

Decimation

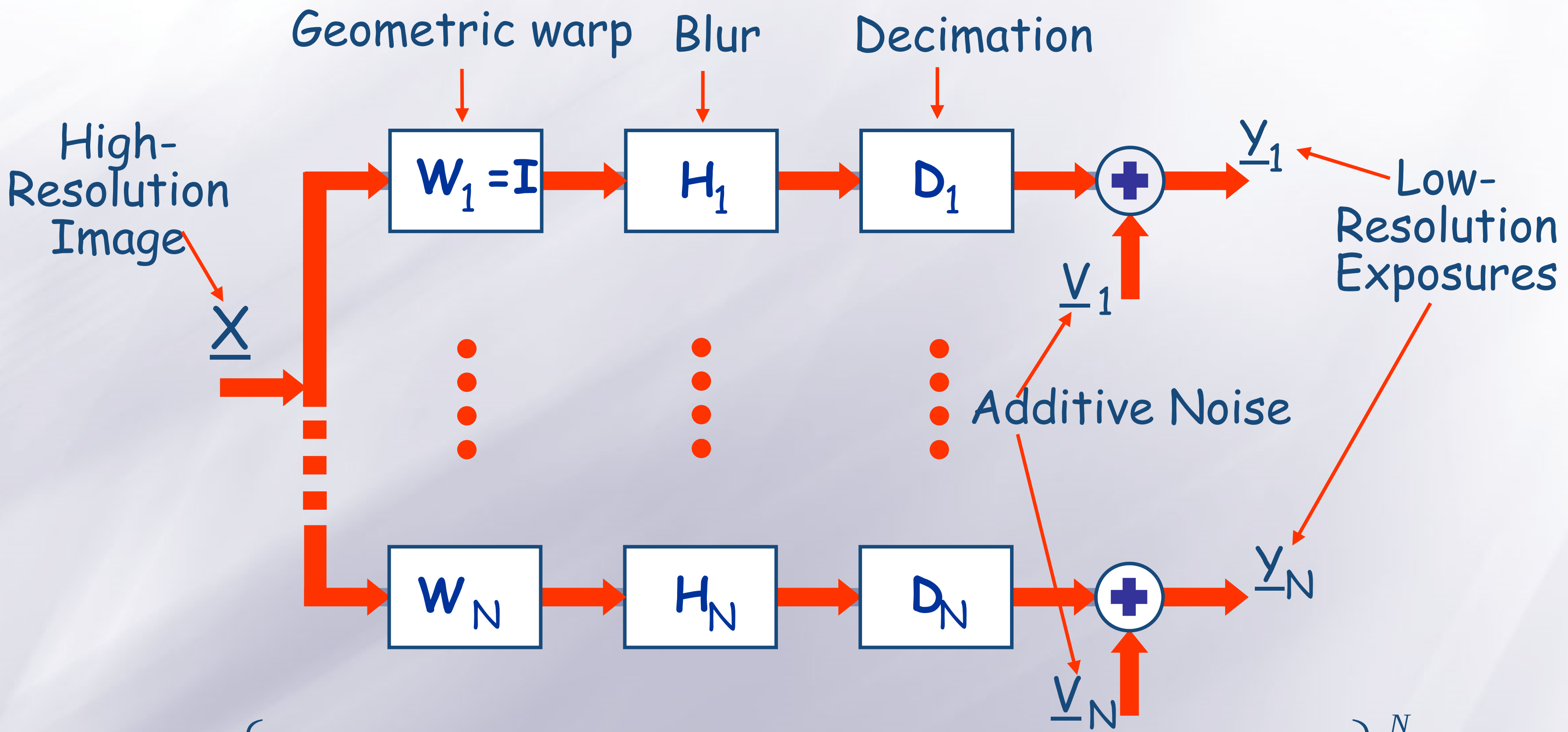
Can be modeled as a linear operation

$\mathbf{D}\underline{X}$



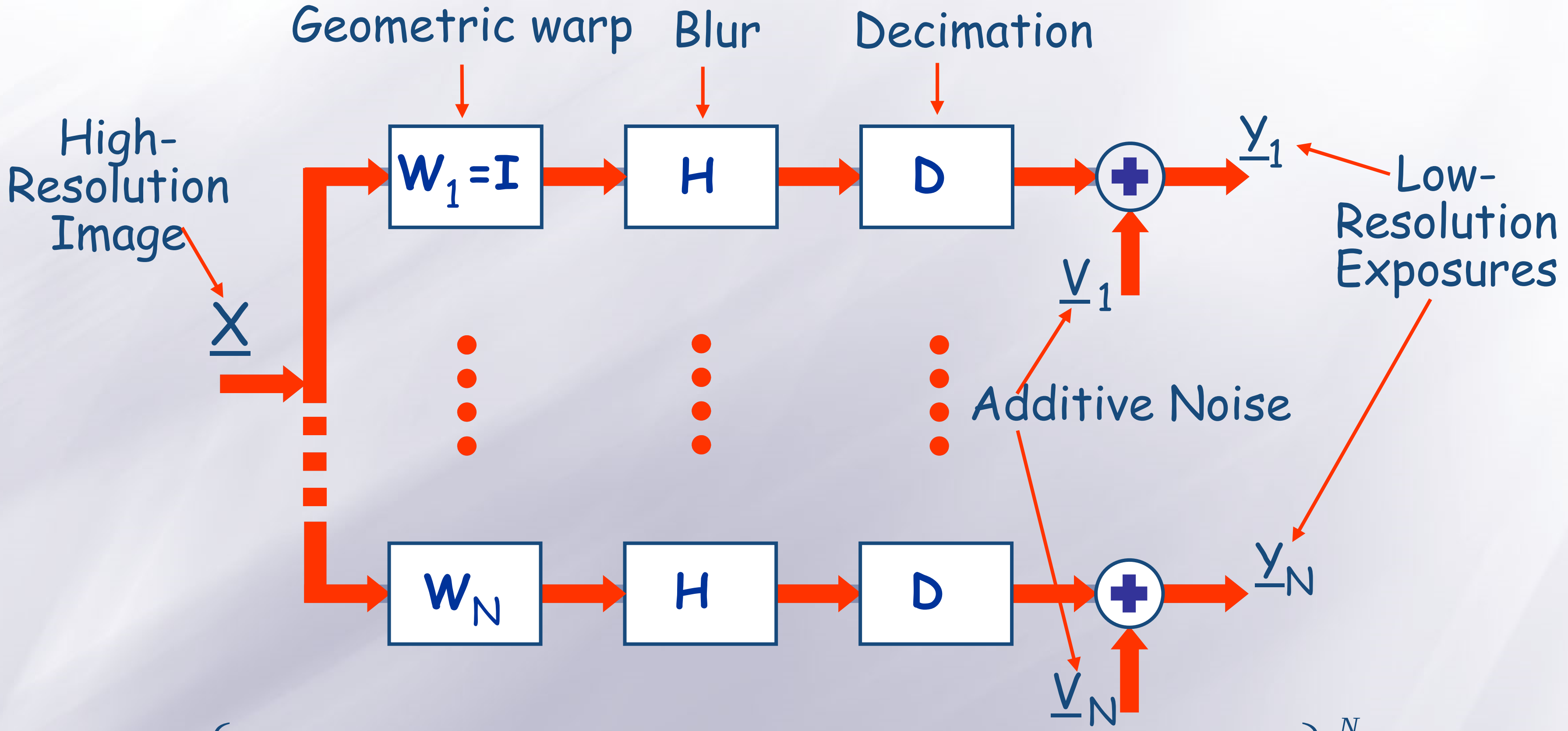


Super-Resolution - Model



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

Simplified Model



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

The Super-Resolution Problem

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

- 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
  - W_k – The warp can be estimated using motion estimation
  - σ_n – The noise can be extracted from the camera / image

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

The Model as One Equation

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

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]

SR - Solutions

- **Maximum Likelihood (ML):**

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

Often ill posed problem!

- **Maximum Aposteriori Probability (MAP)**

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

Smoothness constraint
regularization

ML Reconstruction (LS)

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

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

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

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

Algorithme du gradient

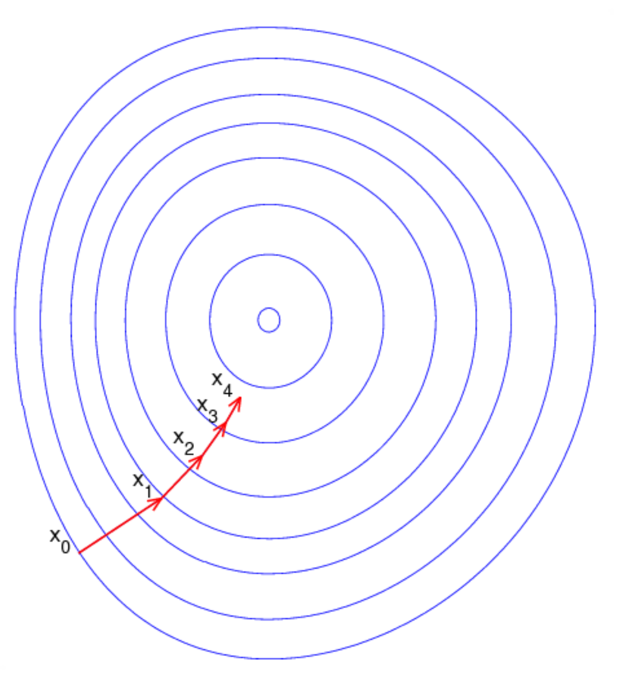
On se donne un point/itéré initial $x_0 \in \mathbb{E}$ et un seuil de tolérance $\varepsilon \geqslant 0$. L'algorithme du gradient définit une suite d'itérés $x_1, x_2, \ldots \in \mathbb{E}$, jusqu'à ce qu'un test d'arrêt soit satisfait. Il passe de x_k à x_{k+1} par les étapes suivantes.

1. Simulation : calcul de $\nabla f(x_k)$.

2. Test d'arrêt : si $\|\nabla f(x_k)\| \leqslant \varepsilon$, arrêt.

3. Calcul du pas $\alpha_k > 0$ par une règle de recherche linéaire sur f en x_k le long de la direction $-\nabla f(x_k)$.

4. Nouvel itéré : $x_{k+1} = x_k - \alpha_k \nabla f(x_k)$.



LS - Iterative Solution

- **Steepest descent**

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

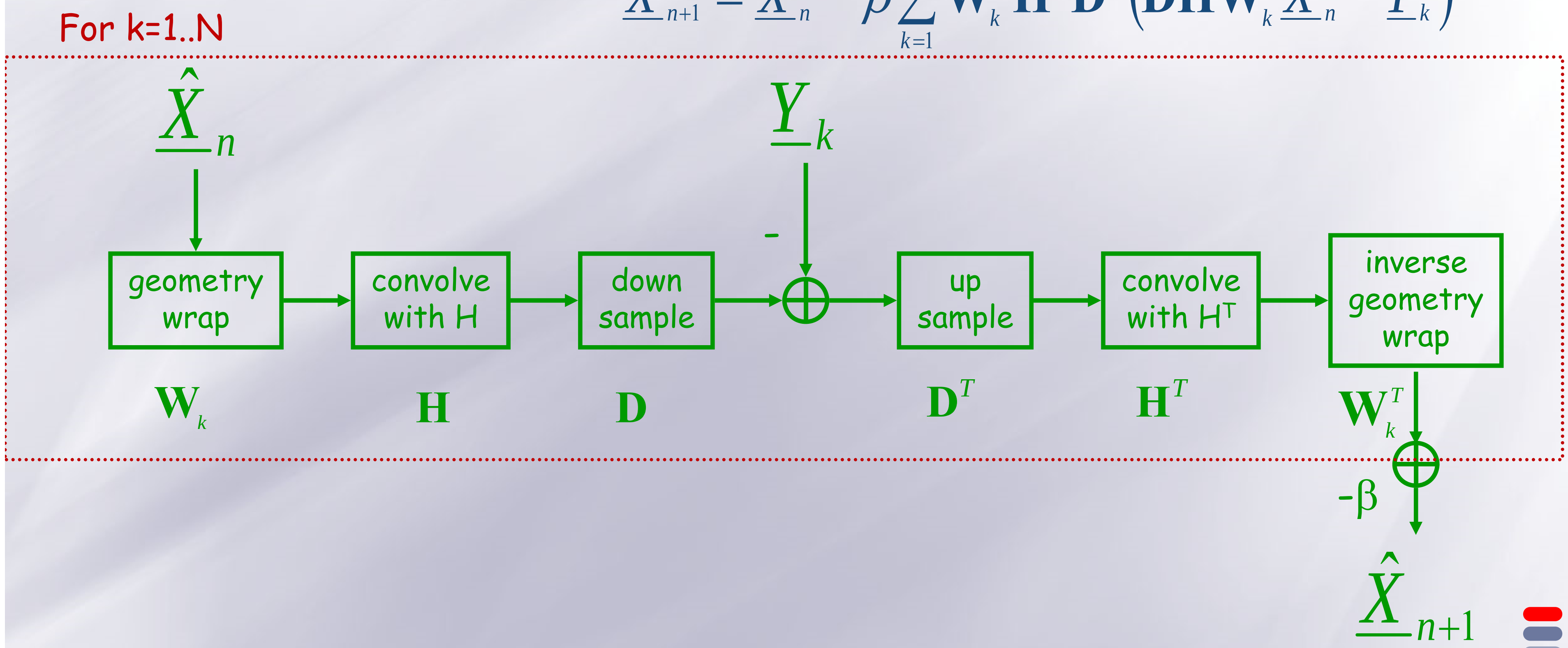
All the above operations can be interpreted as operations performed on images.

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

LS - Iterative Solution

- **Steepest descent**

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



For k=1..N

$\underline{\hat{X}}_n$

$\underline{Y}_k$

-

geometry wrap

convolve with H

down sample

up sample

convolve with H^T

inverse geometry wrap

$\mathbf{W}_k$

$\mathbf{H}$

$\mathbf{D}$

$\mathbf{D}^T$

$\mathbf{H}^T$

$\mathbf{W}_k^T$

$-\beta$

$\underline{\hat{X}}_{n+1}$

Example

HR image

LR + noise
X4

Least squares

Simulated example from Farisu at al.
IEEE trans. On Image Processing, 04

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

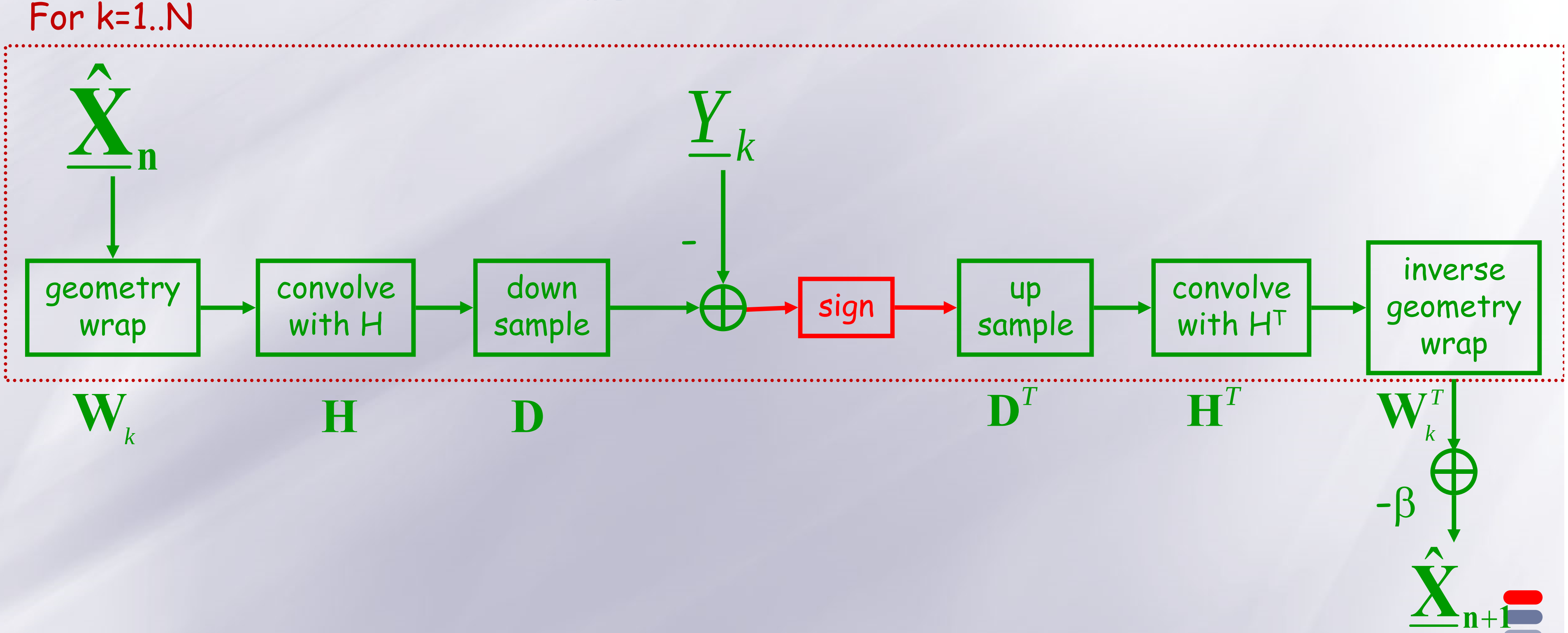
Minimize:
$$\varepsilon^2\left(\underline{X}\right) = \sum_{k=1}^{N} \left\| \mathbf{DHW}_k \underline{X} - \underline{Y}_k \right\|_{\mathbf{1}}$$

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

Robust Reconstruction

- **Steepest descent**

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



Example - Outliers

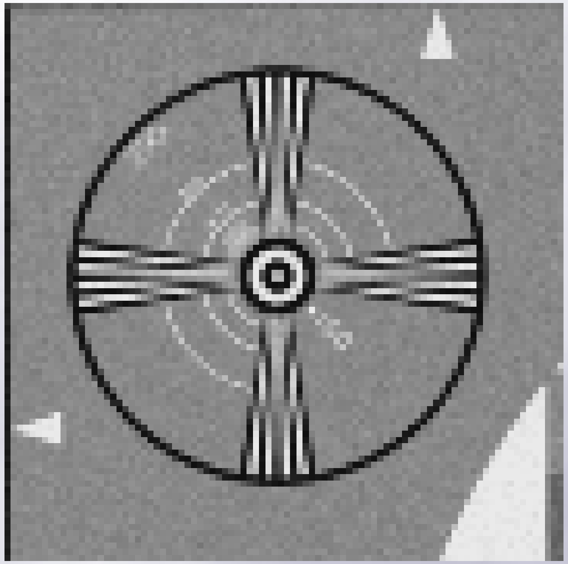
HR image

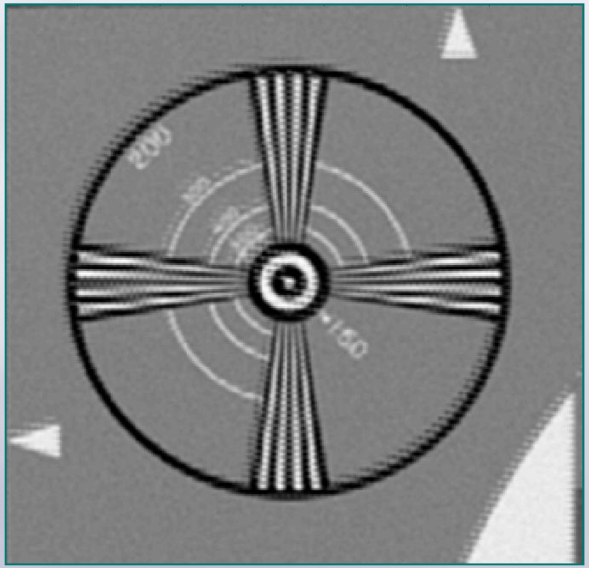
LR + noise
X4

Simulated example from Farisu at al.
IEEE trans. On Image Processing, 04

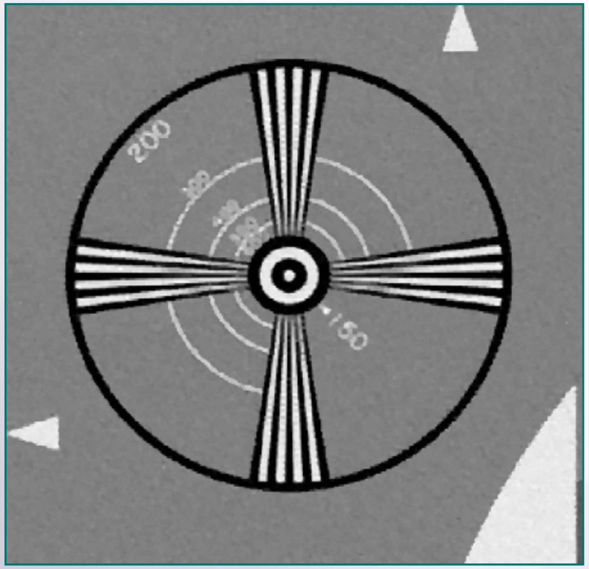
Least squares

Robust Reconstruction





L_2 norm based



L_1 norm based

20 images, ratio 1:4

MAP Reconstruction

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

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

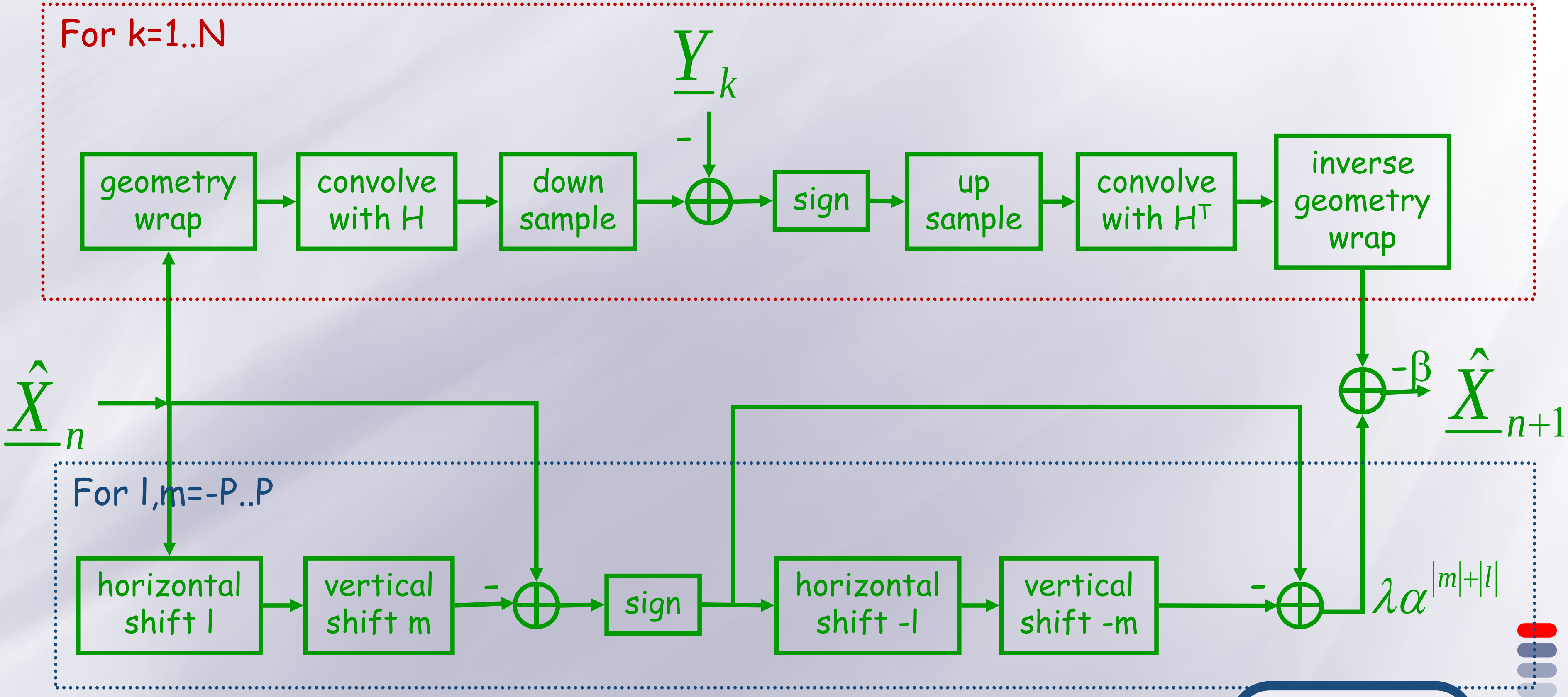
Minimize:

$$\varepsilon^{2}\left(\underline{X}\right)=\sum_{k=1}^{N}\left\| \mathbf{DHW}_{k}\,\underline{X}-\underline{Y}_{k}\right\|_{1}+\lambda\sum_{l=-P}^{P}\sum_{m=-P}^{P}\alpha^{|l|+|m|}\left\|\underline{X}-S_{x}^{l}S_{y}^{m}\,\underline{X}\right\|_{1}$$

$$\begin{aligned}\hat{\underline{X}}_{n+1}=\hat{\underline{X}}_{n}-\beta\Bigg\{&\sum_{k=1}^{N}\mathbf{F}_{k}^{T}\mathbf{H}^{T}\mathbf{W}^{T}\,\mathrm{sign}\left(\mathbf{DHW}_{k}\,\hat{\underline{X}}_{n}-\underline{Y}_{k}\right)\\&+\lambda\sum_{l=-P}^{P}\sum_{m=-P}^{P}\alpha^{|l|+|m|}\left[I-S_{x}^{-l}S_{y}^{-m}\right]\mathrm{sign}\left(\hat{\underline{X}}_{n}-S_{x}^{l}S_{y}^{m}\,\hat{\underline{X}}_{n}\right)\Bigg\}\end{aligned}$$

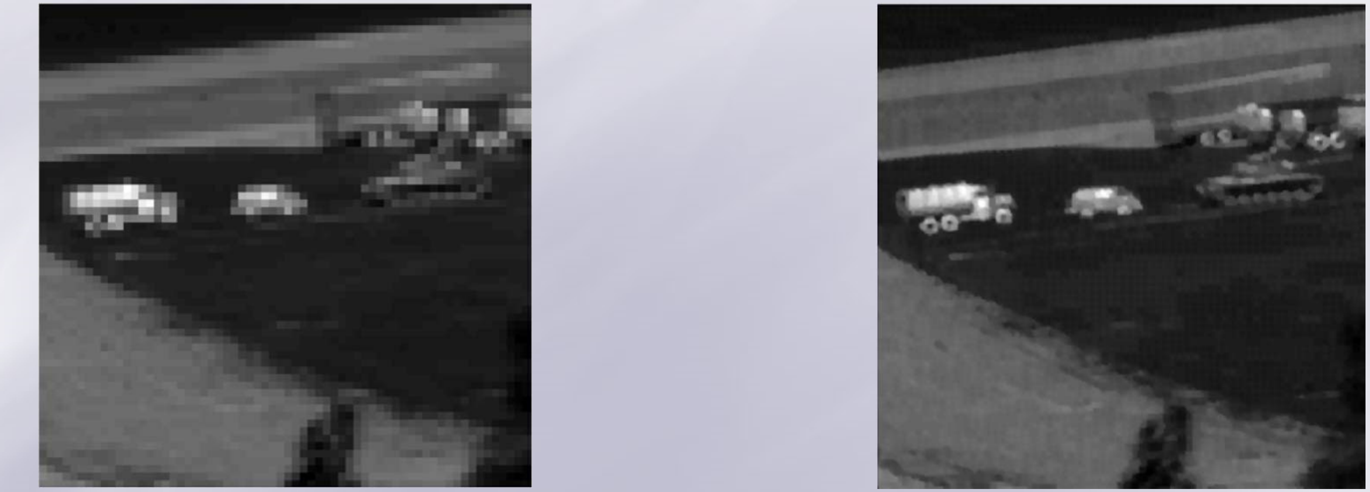
Robust Estimation + Regularization

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



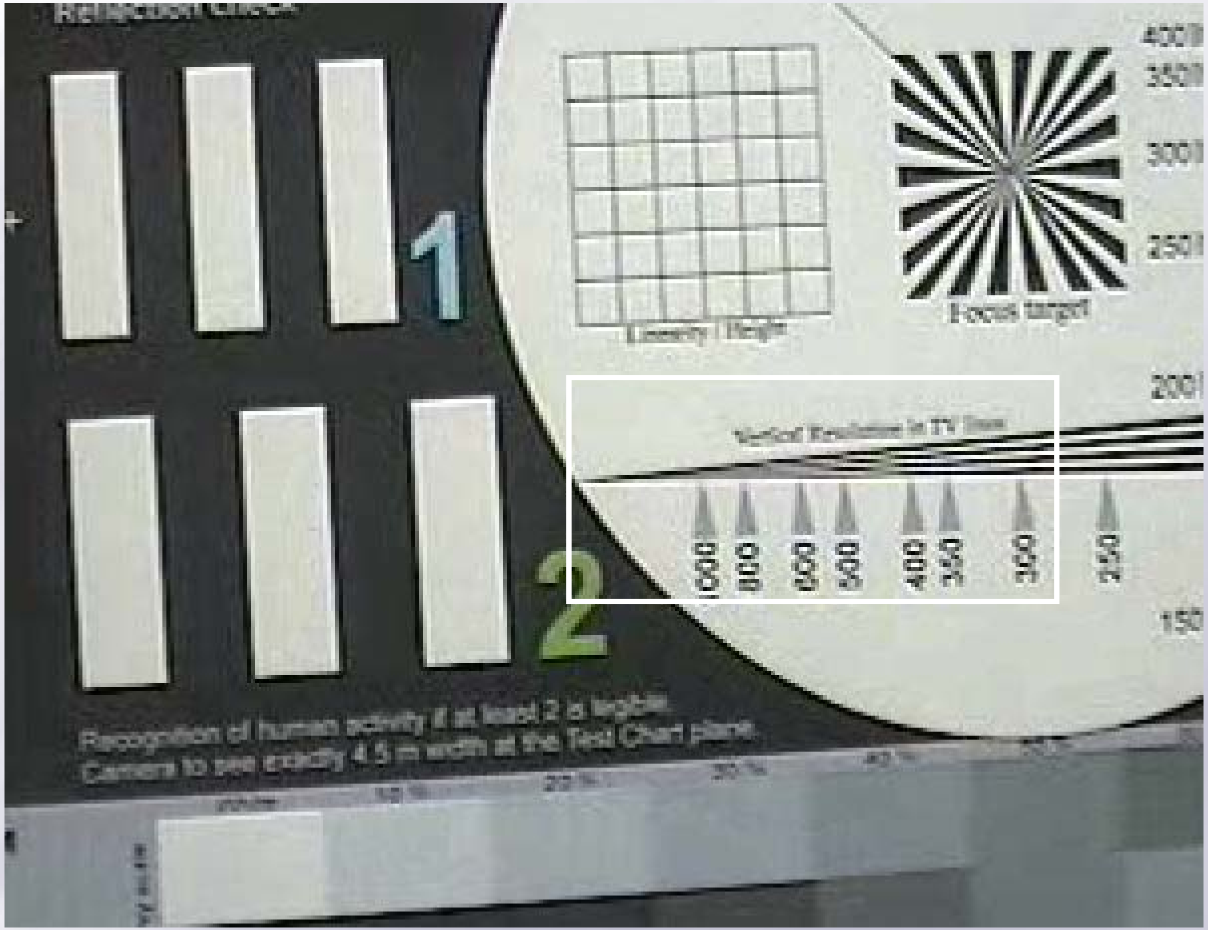
Example

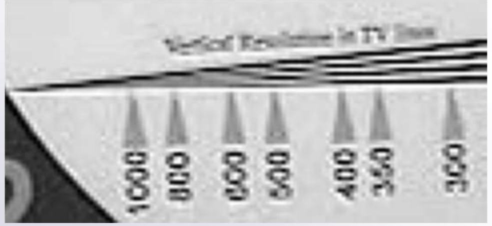
- 8 frames
- Resolution factor of 4

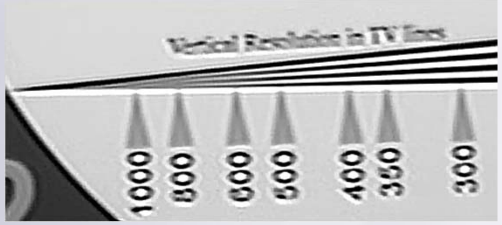


From Farisu at al. IEEE trans. On Image Processing, 04

Example







Images from Vigilant Ltd.

Handling Color in SR

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

Handling color: the classic approach is to convert the measurements to YCbCr, apply the SR on the Y and use trivial interpolation on the Cb and Cr.

Better treatment can be obtained if the statistical dependencies between the color layers are taken into account (i.e. forming a prior for color images).

In case of mosaiced measurements, demosaicing followed by SR is sub-optimal. An algorithm that directly fuse the mosaic information to the SR is better.

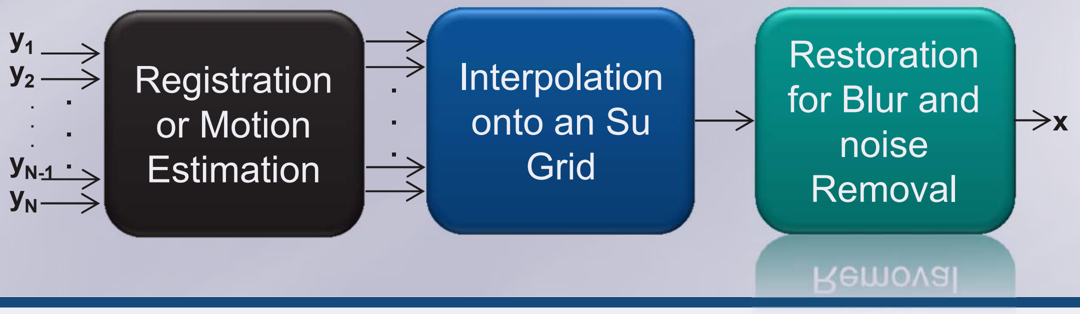
Nonuniform Interpolation Approch

- **Relative motion information estimation**
- **Uniformly spaced sampling Su grid obtained by the single step or iterative method**

J. J. Clark et al., “A transformation method for the reconstruction of functions from nonuniformly spaced samples,” IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-33, pp. 1151-1165, 1985.

J.L. Brown, “Multi-channel sampling of low pass signals,” IEEE Trans. Circuits Syst., vol. CAS-28, pp. 101-106, 1981.

- **Application of a deconvolution method to remove blurring and noise**



y_1

y_2

y_{N-1}

y_N

Registration or Motion Estimation

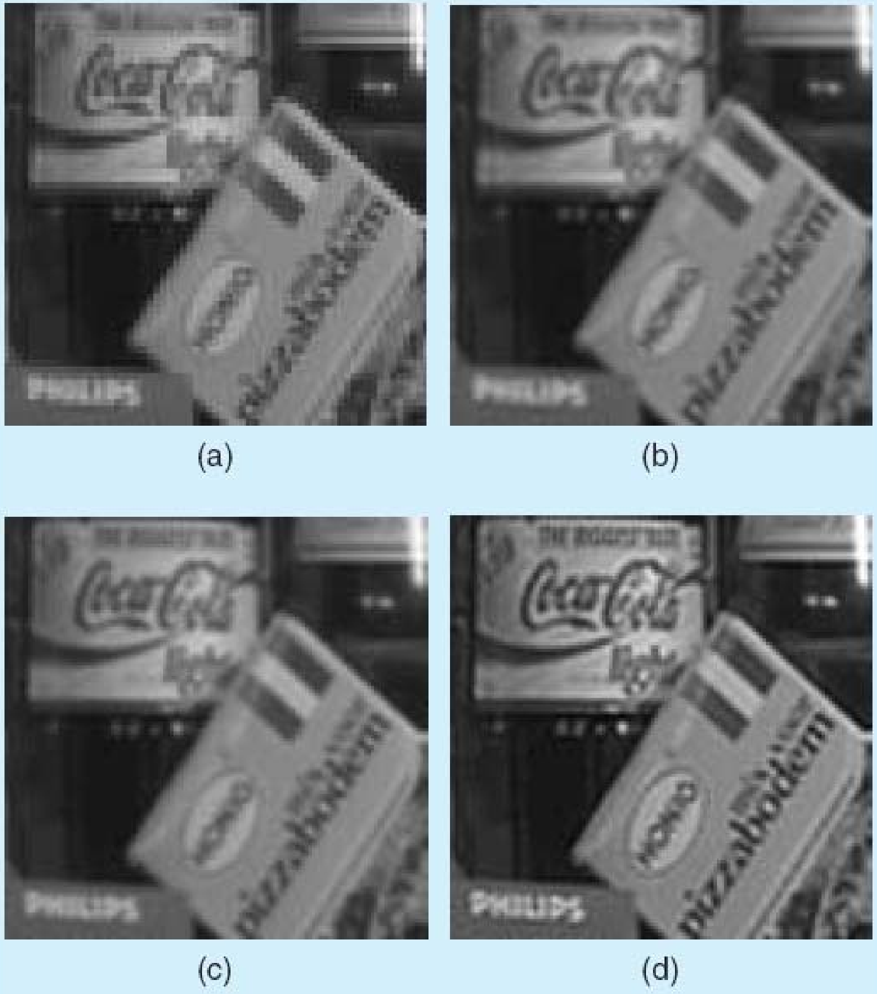
Interpolation onto an Su Grid

Restoration for Blur and noise Removal

x

Nonuniform Interpolation Approach

- (a) nearest neighbor interpolation
- (b) bilinear interpolation
- (c) non uniform interpolation with 4 Lo images
- (d) Deblurring part


(a) (b)
(c) (d)

Frequency Domain Approach

- Based on
  - Shifting property of the Fourier Transform
  - Aliasing relationship between the CFT of an Su image and the DFT of Lo_i images
  - Bandlimited property of the Su image

R.Y. Tsai and T.S. Huang, “Multipleframe image restoration and registration,” in Advances in Computer Vision and Image Processing. Greenwich, CT: JAI Press Inc., 1984, pp. 317-339.

Frequency Domain Approach

- Let a Su image x and its CFT X

- Global translation yield the ith shifted image where the translation vector is known: $x_i = x + t$

- By the shifting property of the CFT, the CFT shifted image Su image can be written: $X_i = f(X)$

- X_i is sampled to generate the observed y_i Lo image

- A system of equations is formulated from the relationship between the CFT of Su and the DFT of the ith observed Lo image

- Finally, the inverse problem is resolved to determine first the DFT of the observed Lo_i images and then CFT coefficient of x

Regularized SR Recons. Approach

- Stochastic Approach based on Bayesian estimation methods
  - PDF (Probability Density Function) of the original image can be established
  - MAP (Maximum A Posteriori) estimator of x maximises the a posteriori PDF with respect to x

$$x = \arg\max P(x|y_1, \ldots, y_n)$$

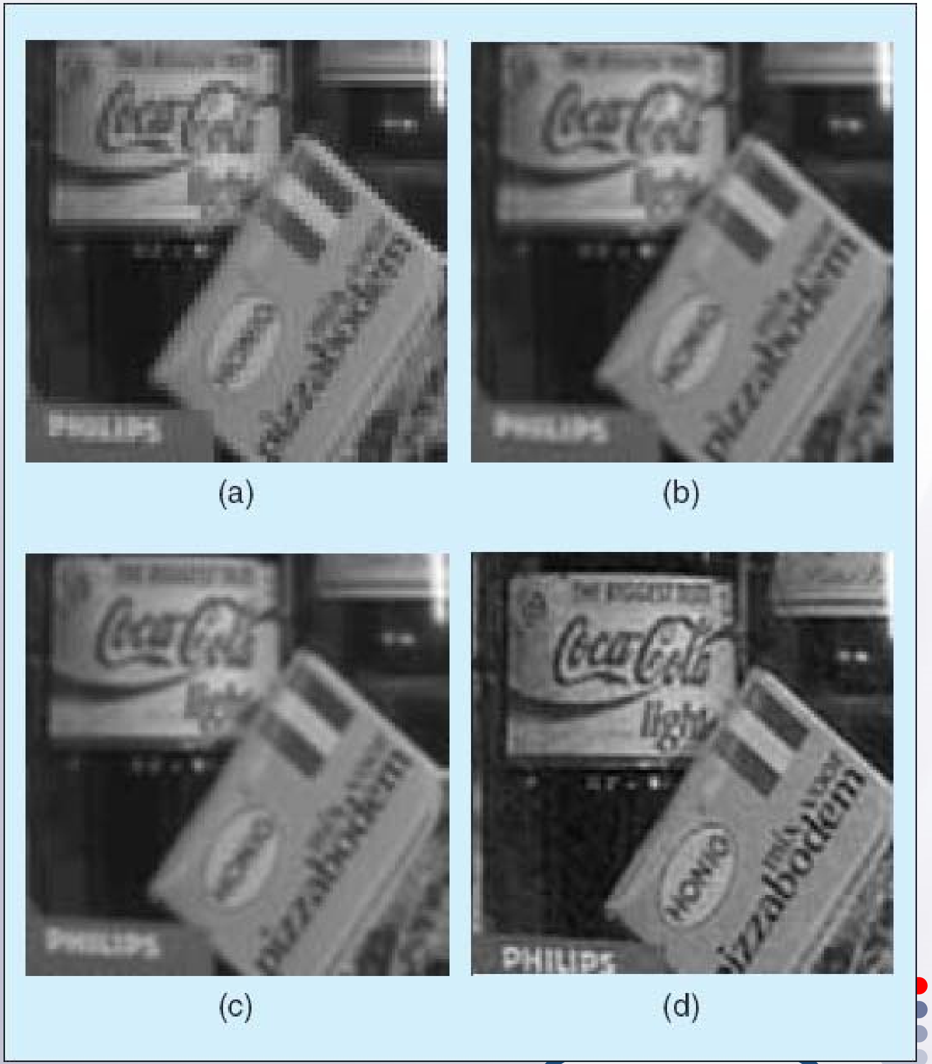
  - Taking the logarithmic function and applying Bayes' theorem to the conditionnal probability, we have

$$x = \arg\max\{\ln P(y_{1,} \ldots y_N |x) + \ln P(x)\}$$

S. Baker and T. Kanade, "Limites on Super-Resoltion and How to Break Them," IEEE Transactions on Pattern Analysis and Machine Intelligence, Sep 2002, Vol. 24(9), pp. 1167 – 1183.

Regularized SR Recons. Approach

- (a) nearest neighbor interpolation
- (b) bilinear interpolation
- (c) non uniform interpolation with 4 Lo images
- (d) MAP with edge-preserving Prior


(a) (b) (c) (d)

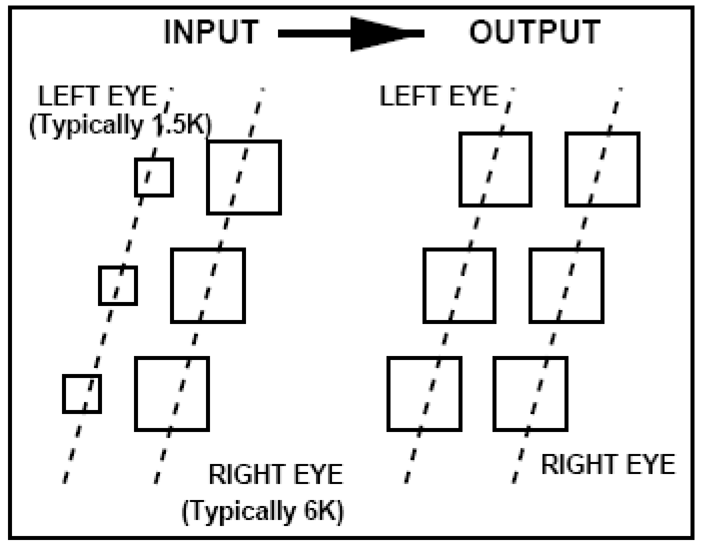
Interesting application of SR

- I. Hybrid Stereo Camera
- II. Super-Resolution of Face Images
- III. Depth Superresolution for ToF 3D Shape Scanning

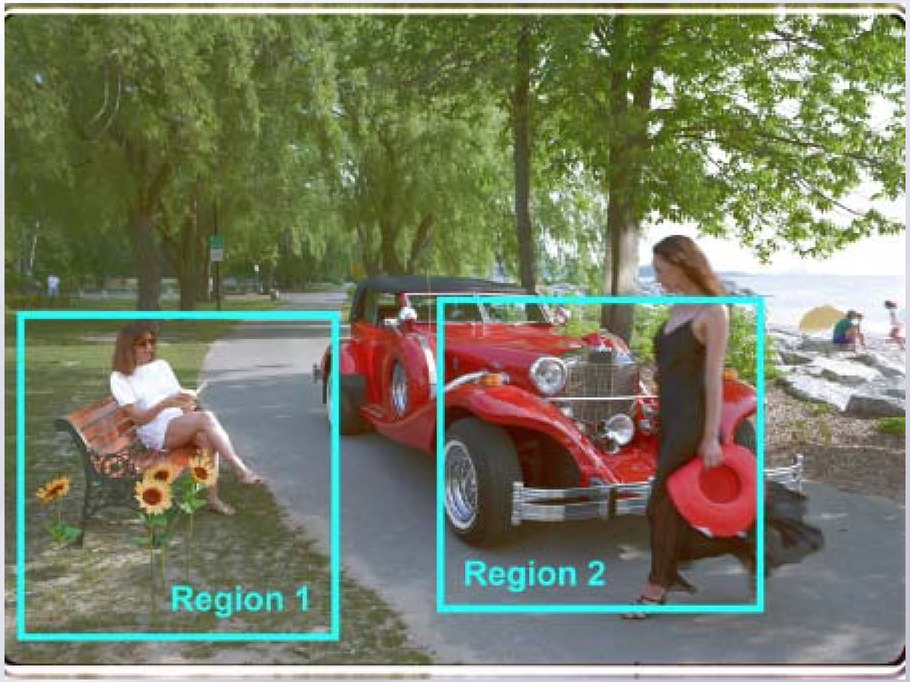
I. Hybrid Stereo Camera

H. S. Sawhney, Y. Guo, K. J. Hanna, R. Kumar, S. Adkins, S. Zhou ’’ Hybrid stereo camera: an IBR approach for synthesis of very high resolution stereoscopic image sequences. ’’ SIGGRAPH 2001: 451-460.

A schematic depicting the hybrid resolution stereo input and the full resolution output.



I. Hybrid Stereo Camera



(A1) Input: Right Original Full-res $(2K \times 2K)$ **(A2) Output**: Left Synthesized Full-res $(2K \times 2K)$

(A3) Input: Left Low-resolution (512×512) **(B1) Input**: Right Low-resolution $(1K \times 1K)$

(B2) Input: Left Original Full-res $(4K \times 4K)$ **(B3) Output**: Right Synthesized Full-res $(4K \times 4K)$

II. Super-Resolution of Face Images

S. Baker and T. Kanade, “Limites on Super-Resoltion and How to Break Them,” IEEE Transactions on Pattern Analysis and Machine Intelligence, Sep 2002, Vol. 24(9), pp. 1167 – 1183.

(a) Input 24 × 32 (b) Hallucinated (c) Hardie *et al.* (d) Original (e) Cubic B-spline

(f) Input 24 × 32 (g) Hallucinated (h) Hardie *et al.* (i) Original (j) Cubic B-spline

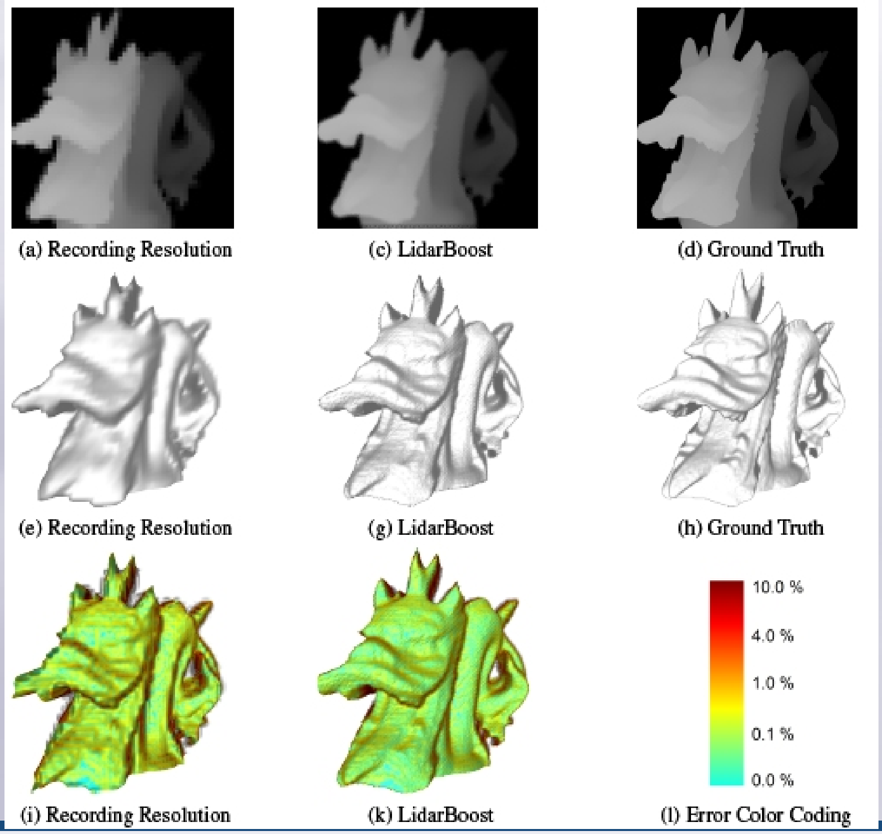
II. Super-Resolution of Face Images


(a) Random Hallucinated (b) Misc. Hallucinated (c) Constant Hallucinated

The results of applying our hallucination algorithm to images not containing faces. As is evident, a face is hallucinated by the proposed algorithm even when none is present, hence the term “hallucination algorithm.”

III. Depth Super-Resolution for ToF 3D Shape Scanning

S. Schuon, C. Theobalt, J. Davis, S. Thrun, "LidarBoost: Depth superresolution for ToF 3D shape scanning," Computer Vision and Pattern Recognition, IEEE Computer Society Conference on, pp. 343-350, 2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009.



(a) Recording Resolution (c) LidarBoost (d) Ground Truth

(e) Recording Resolution (g) LidarBoost (h) Ground Truth

(i) Recording Resolution (k) LidarBoost (l) Error Color Coding

III. Depth Super-Resolution for ToF 3D Shape Scanning

(a) Color Image

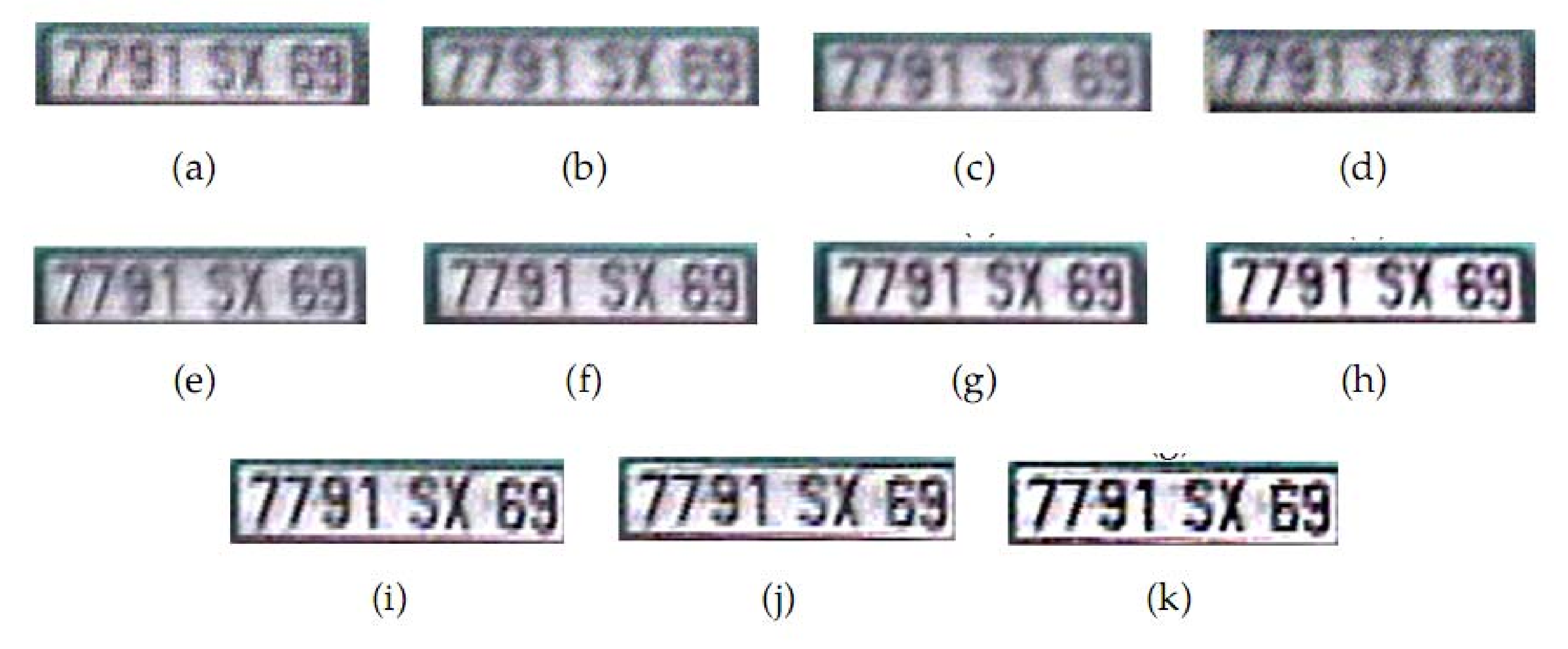
(b) Recording Resolution

(e) LidarBoost

Amélioration de qualité de plaque d'immatriculation

Performance des algorithmes de super-résolution

- Données de vidéosurveillance en contexte réel [Projet UCSD]



- (a) → (d): 4 observations; de (e) à (k): résultats de super-résolution (facteur 2, utilisant 10 observations) de NIL, ML, MAP_GMRF, MAP_HMRF, MAP_DAMRF, MAP_CRE_DAMRF