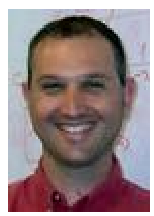
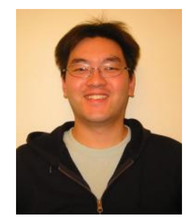
(Super) Resolution: Statistical Definition, Computation, and Fundamental Limits

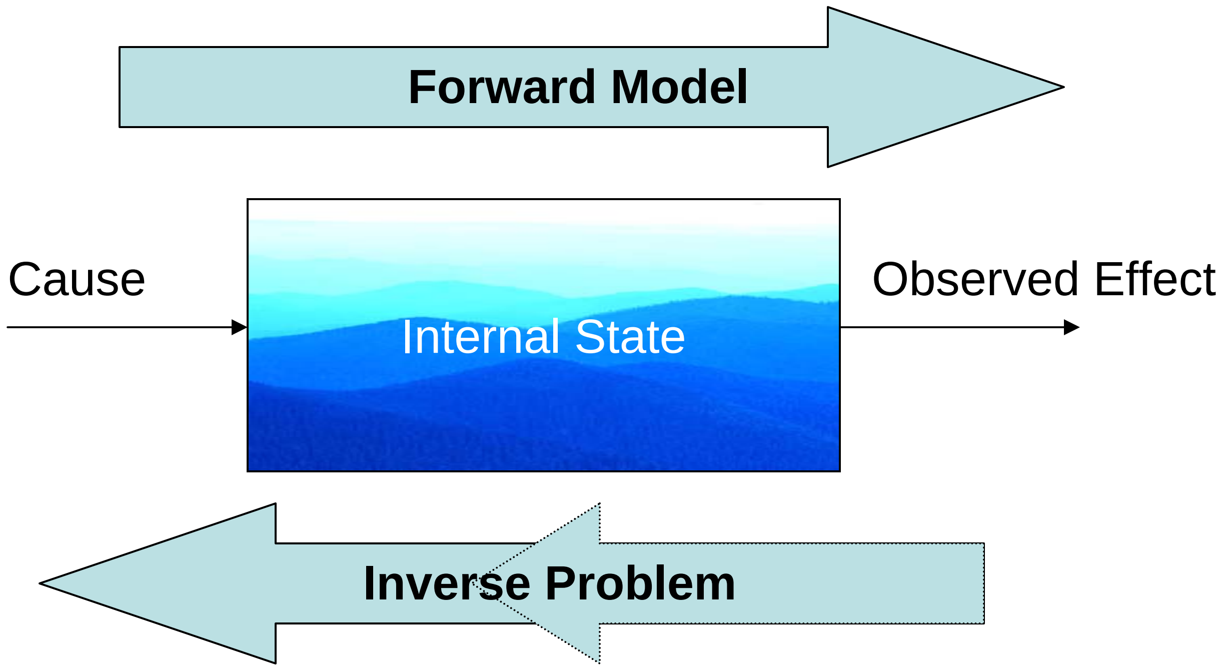
Peyman Milanfar*
EE Department
University of California, Santa Cruz

*Joint work with Sina Farsiu, Miki Elad, Dirk Robinson, Morteza Shahram, Hiro Takeda

International Conference on Super-Resolution Imaging
Hong Kong, August 28-31, 2005

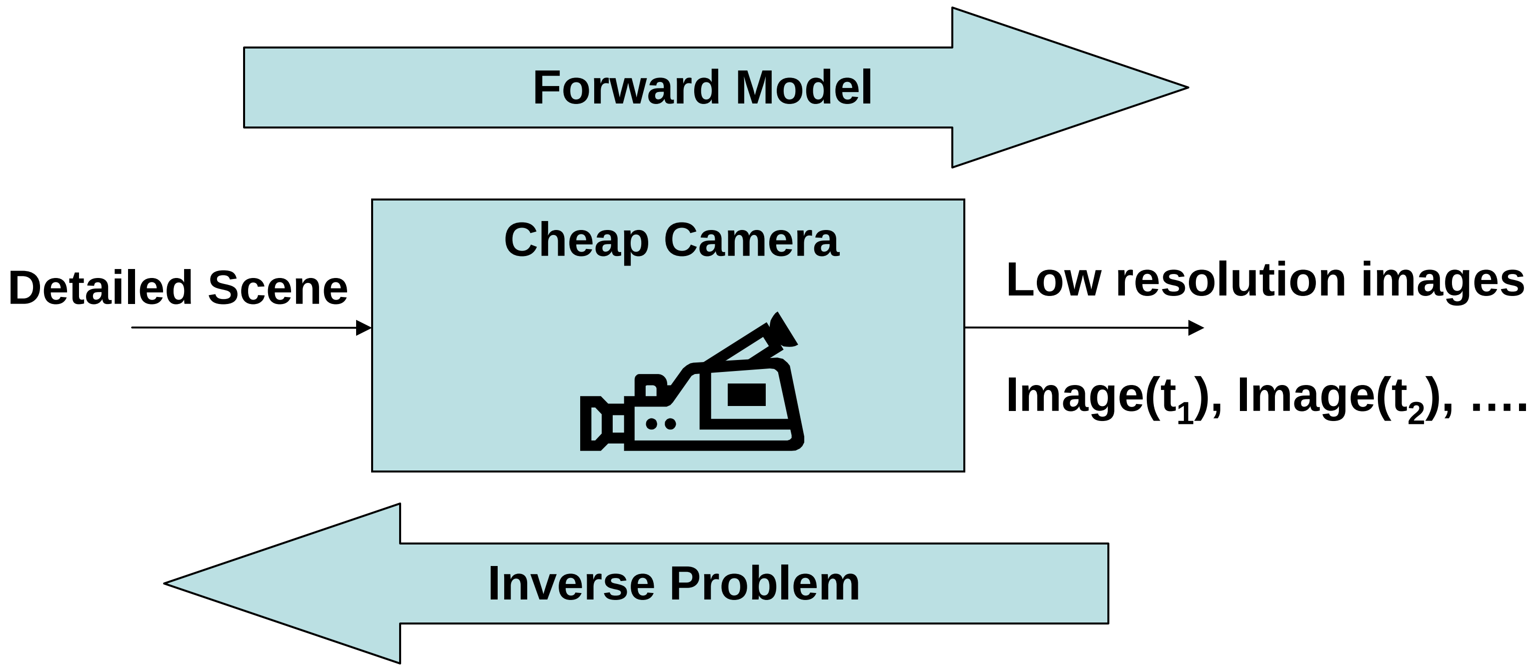
Inverse Problems



- Noise sensitivity
- Non-uniqueness
- Numerical conditioning

Inverse Problem of Interest

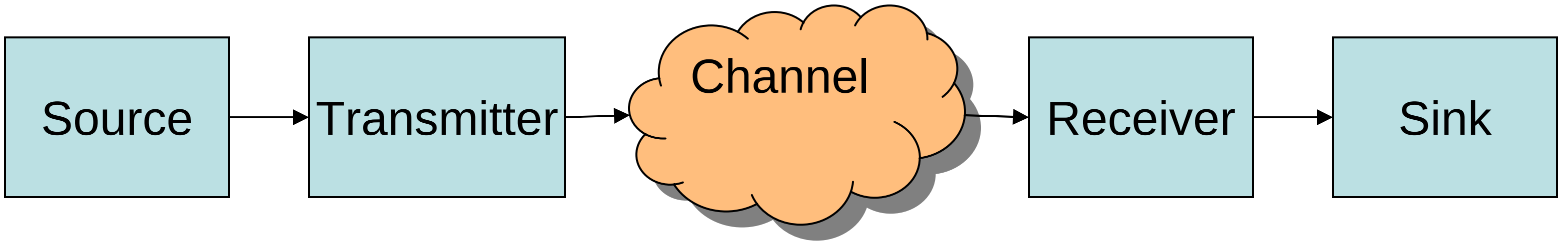
$$LR(x,y,t) = \downarrow_N \left[HR(x,y,t) * PSF(x,y) \right] + noise$$



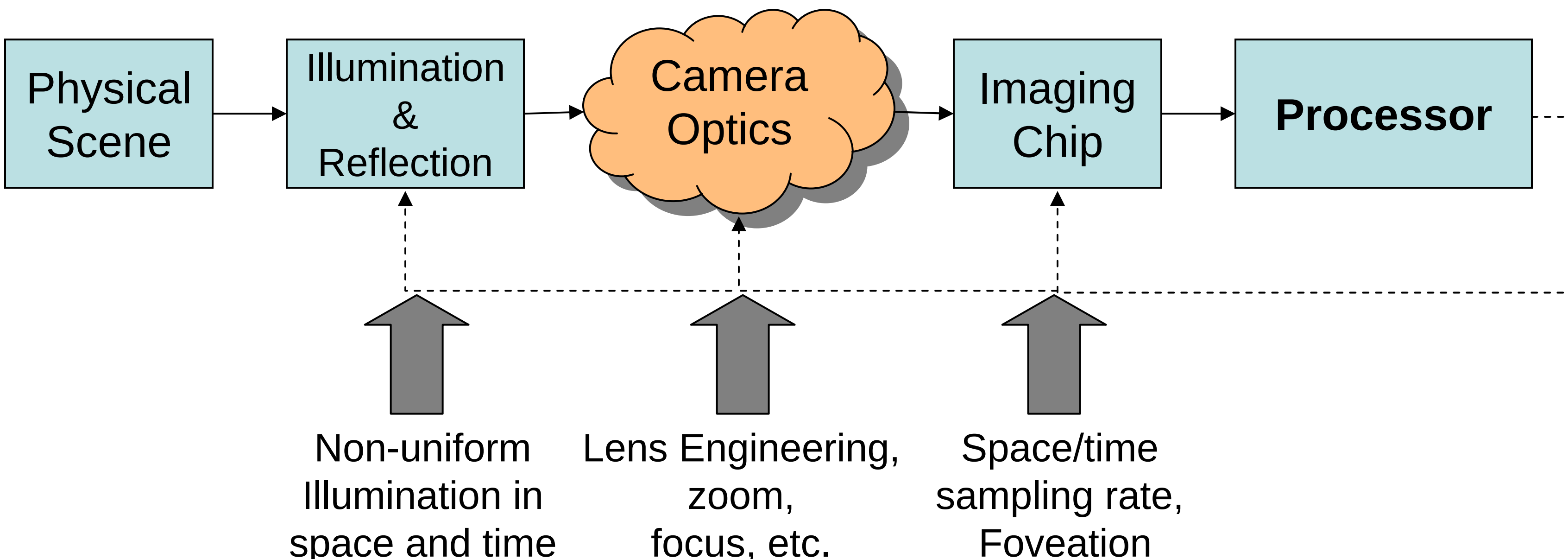
Compute $HR(x,y,t)$

Information in Imaging

Communication System:



Optical Imaging:



Some Informative Analogies

- Imaging (Inv. Probs.)
  - Point-spread function
  - Deconvolution
  - Occlusion
  - “Multi-frame” imaging
  - Image registration
  - Resolution Limits
  - ….
- Communication
  - Channel response
  - Equalization
  - Interference
  - Multi-antenna systems
  - Time-delay estimation
  - “Capacity”
  - ….

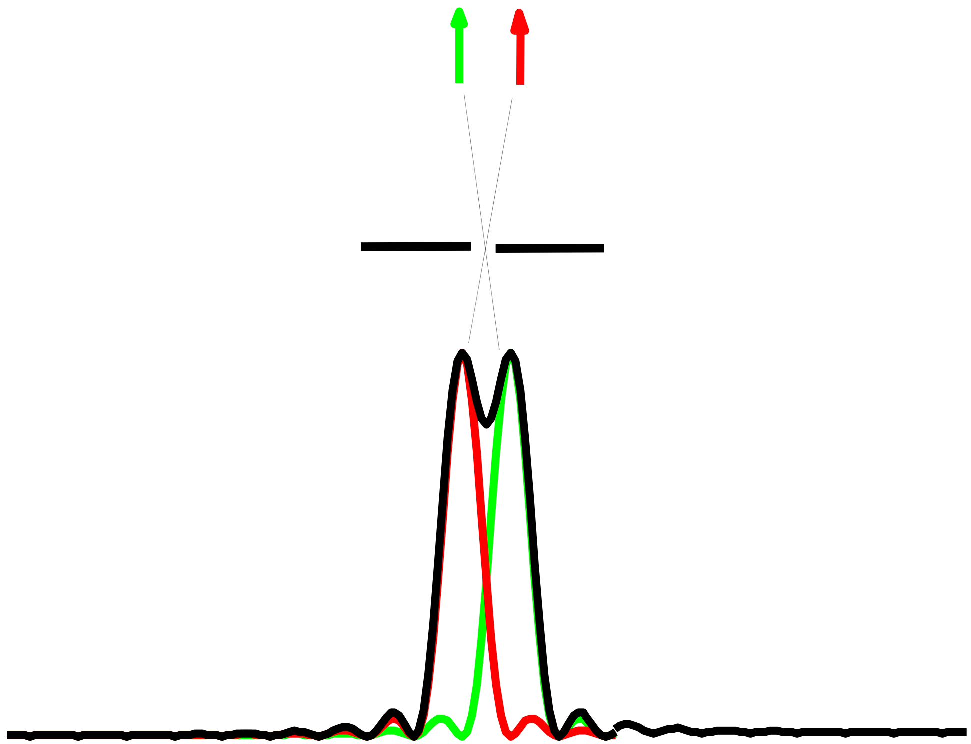
Agenda

- Sensor model and limitations
- Processing algorithms
- SR Performance limits
- Further extensions and directions

Sensing : Resolution Limits of a Canonical Image Sensor

Pinhole Camera

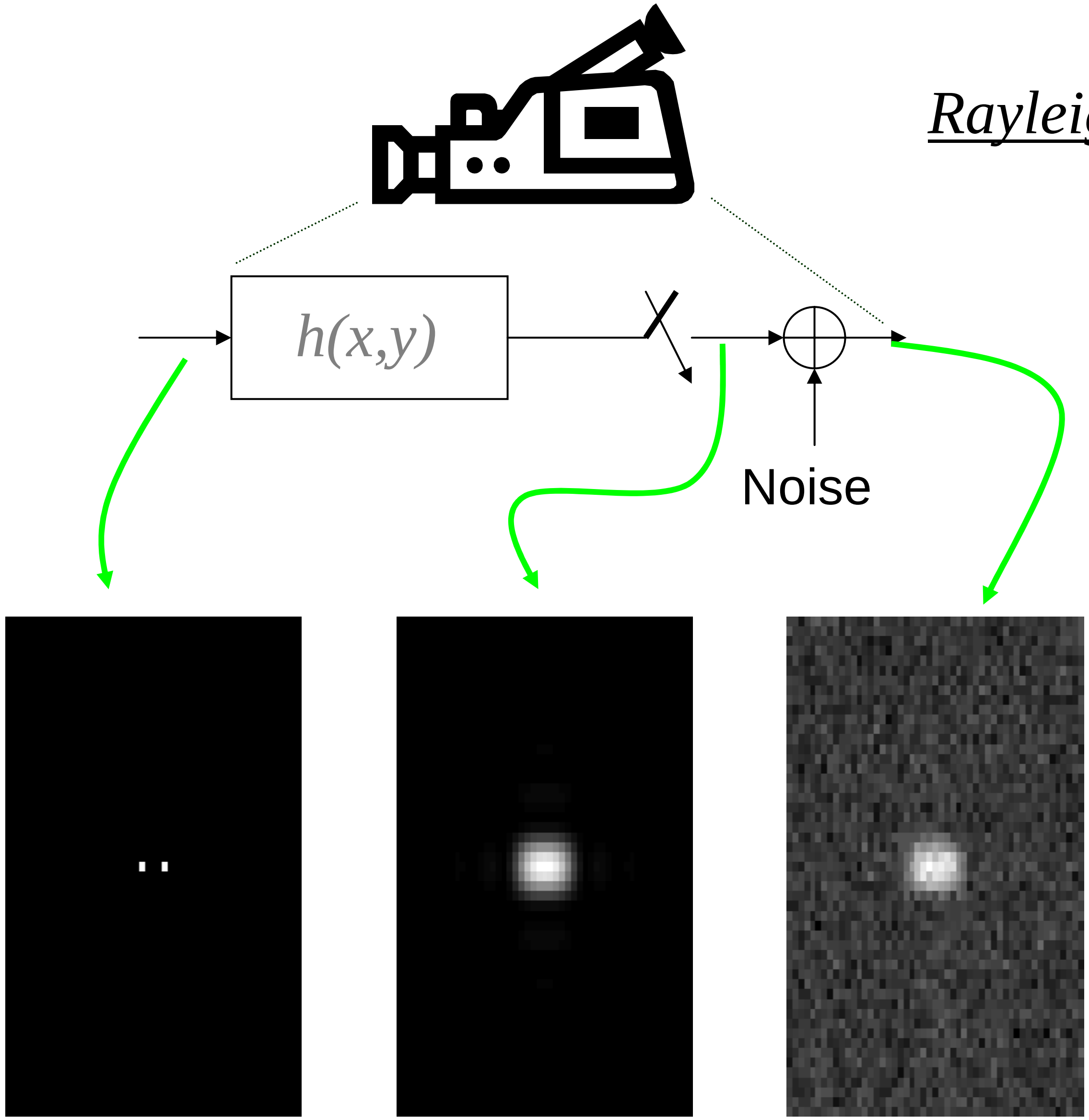
- The image of two sources is the incoherent sum of PSFs, representing the effect of the diffraction



- When the point sources are "too close", according to the Rayleigh criterion these two point sources are not resolvable.

Imaging Closely-Spaced Point Sources

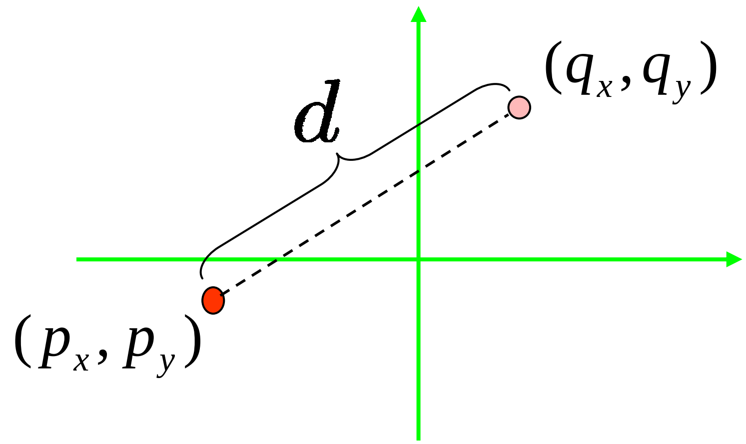
Rayleigh's limit isn't.



Milanfar et al. EE Dept, UCSC

Resolution: Stochastic Problem

- Point sources:



- Measured Signal:

$$\boxed{\alpha}\, h(x_k - p_k, y_l - p_y) + \boxed{\beta}\, h(x_k + q_k, y_l + q_y) + w(x_k, y_l)$$

Resolution as a composite statistical hypothesis test:

Spatial Resolution

$$\begin{cases} \mathcal{H}_0 : d = 0 & \text{One peak is present} \\ \mathcal{H}_1 : d > 0 & \text{Two peaks are present} \end{cases}$$

Optimal Solution and a Scaling Law

- "Information capacity": What is the minimum SNR required to detect the presence of two point sources with high confidence?

Depends on the sensor PSF, the required false alarm and tolerable error rates. (Optimize!)

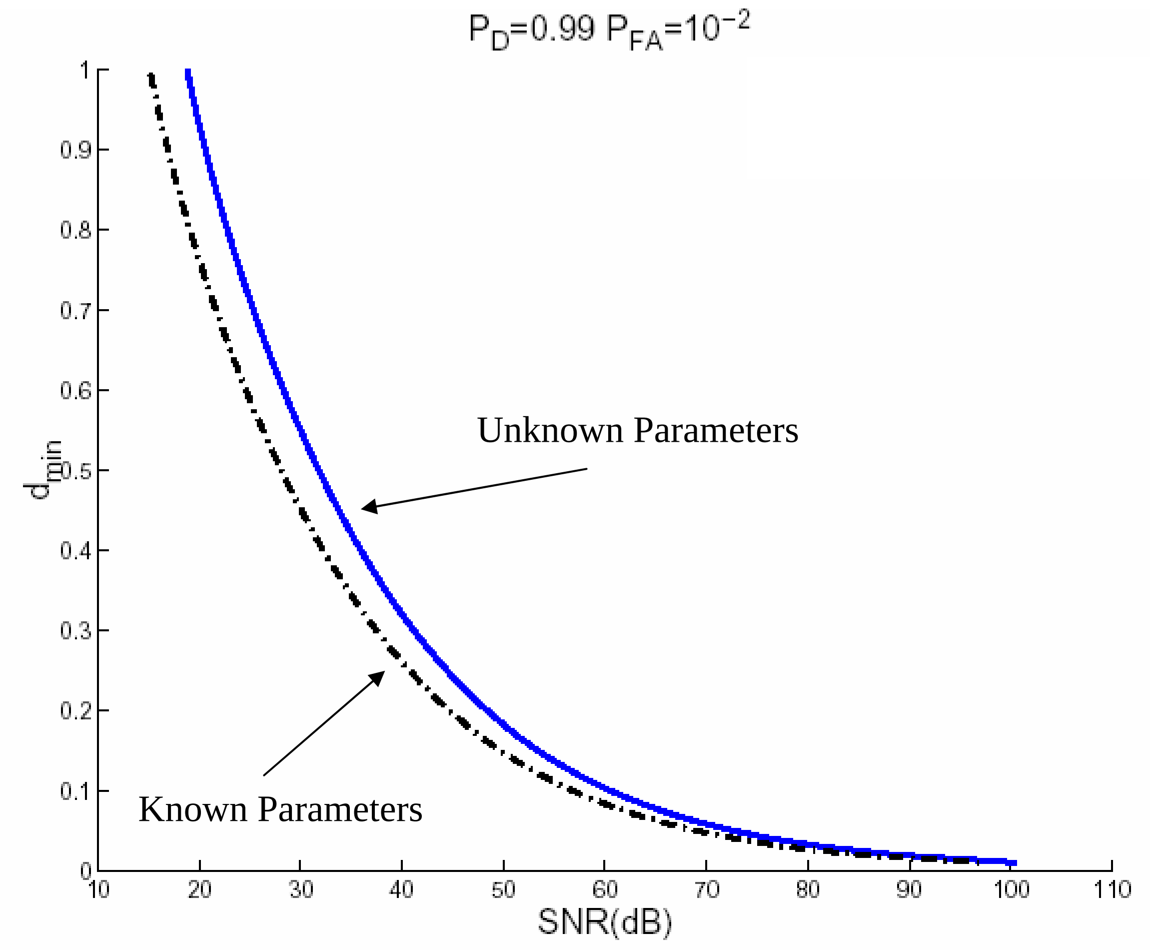
$$\boxed{SNR} \approx \frac{C}{N\boxed{d^4}}$$

Number of samples at focal plane array

M. Shahram, and P. Milanfar, "**Imaging Below the Diffraction Limit: A Statistical Analysis**", *IEEE Transactions on Image Processing, vol. 13, no. 5, pp. 677-689, May 2004*

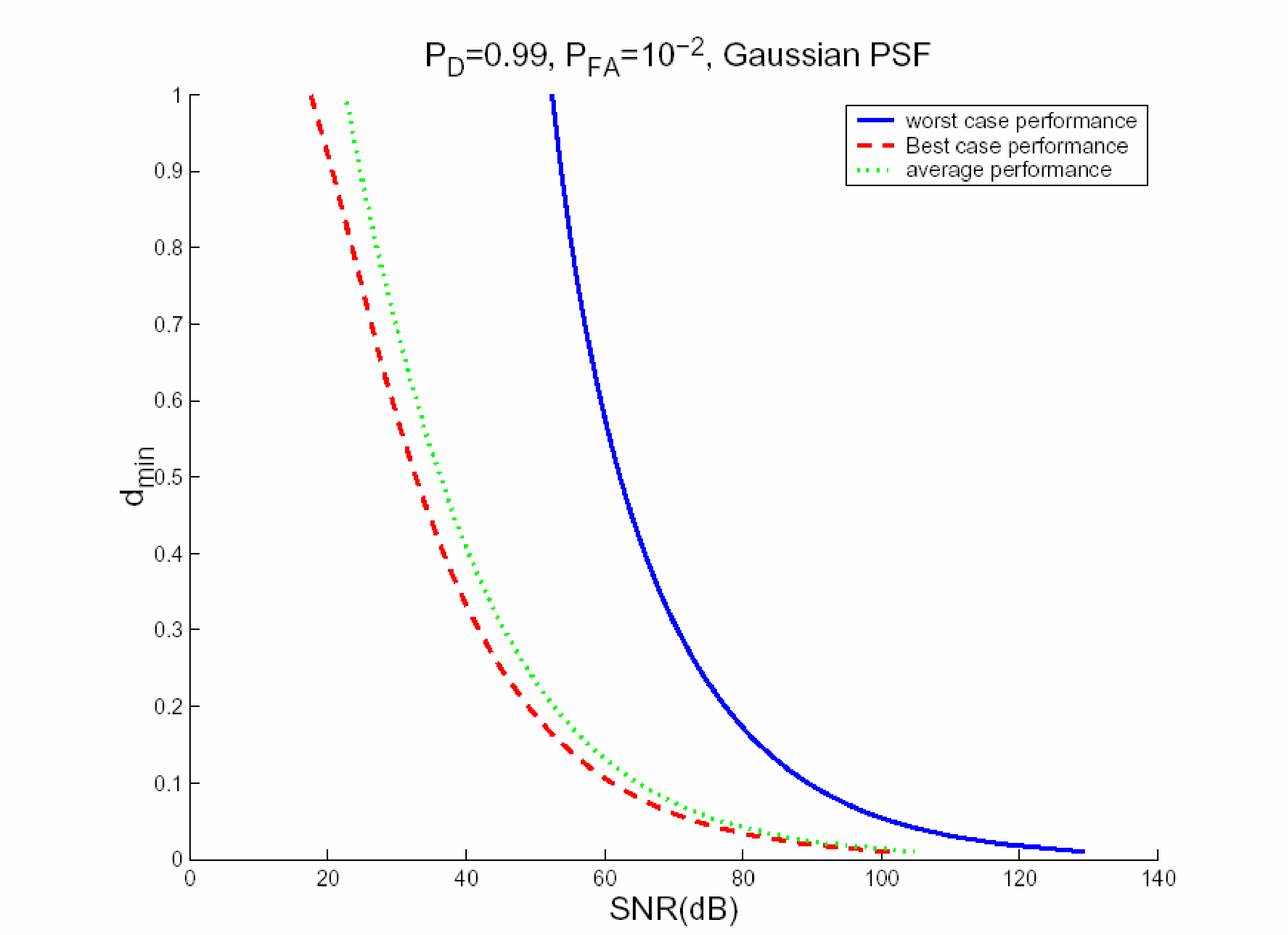
M. Shahram, and P. Milanfar,"**Statistical and Information-Theoretic Analysis of Resolution in Imaging**", to *appear in IEEE Transactions on Information Theory*

Example for super-critical sampling



What happens when there is aliasing?

Example for sub-critical sampling (50% below Nyquist, two frames)



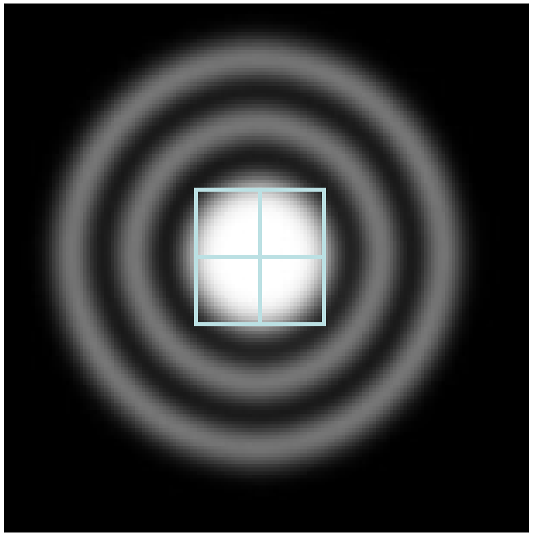
Message: Things can get a lot worse, but not impossible!

Milanfar et al. EE Dept, UCSC

Processing :Multi-frame Resolution Enhancement (Super-resolution)

Why Spatial Resolution Enhancement?

- To obtain an alias-free, “diffraction limited” image we need 4 pixels covering the Airy disk:



- That is: radius of the Airy disk must match the pixel dimensions.

Motivation: SuperResolution Goes to Hollywood

CBS Program “Numb3rs” Episode from March 11, 2005

Overcoming Sensor Limitations by Processing

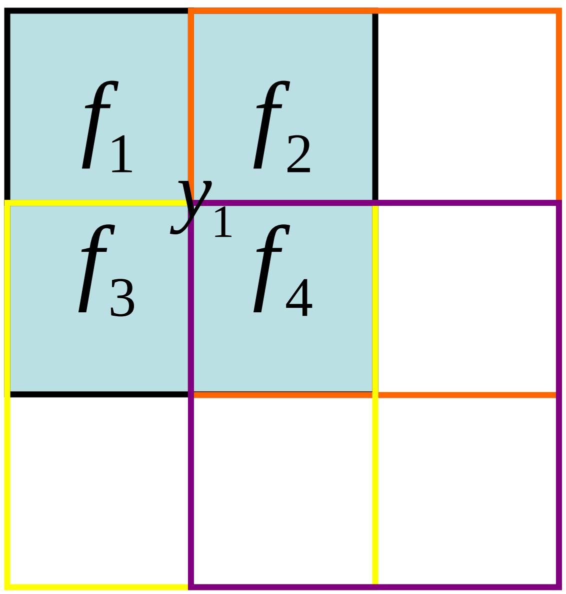
The Idea: “Diversity” + Aliasing

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

Data Courtesy USAF

Resolution Enhancement Model

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



"PSF"

$$y_1 = a_1 f_1 + a_2 f_2 + a_3 f_3 + a_4 f_4 + e_1$$

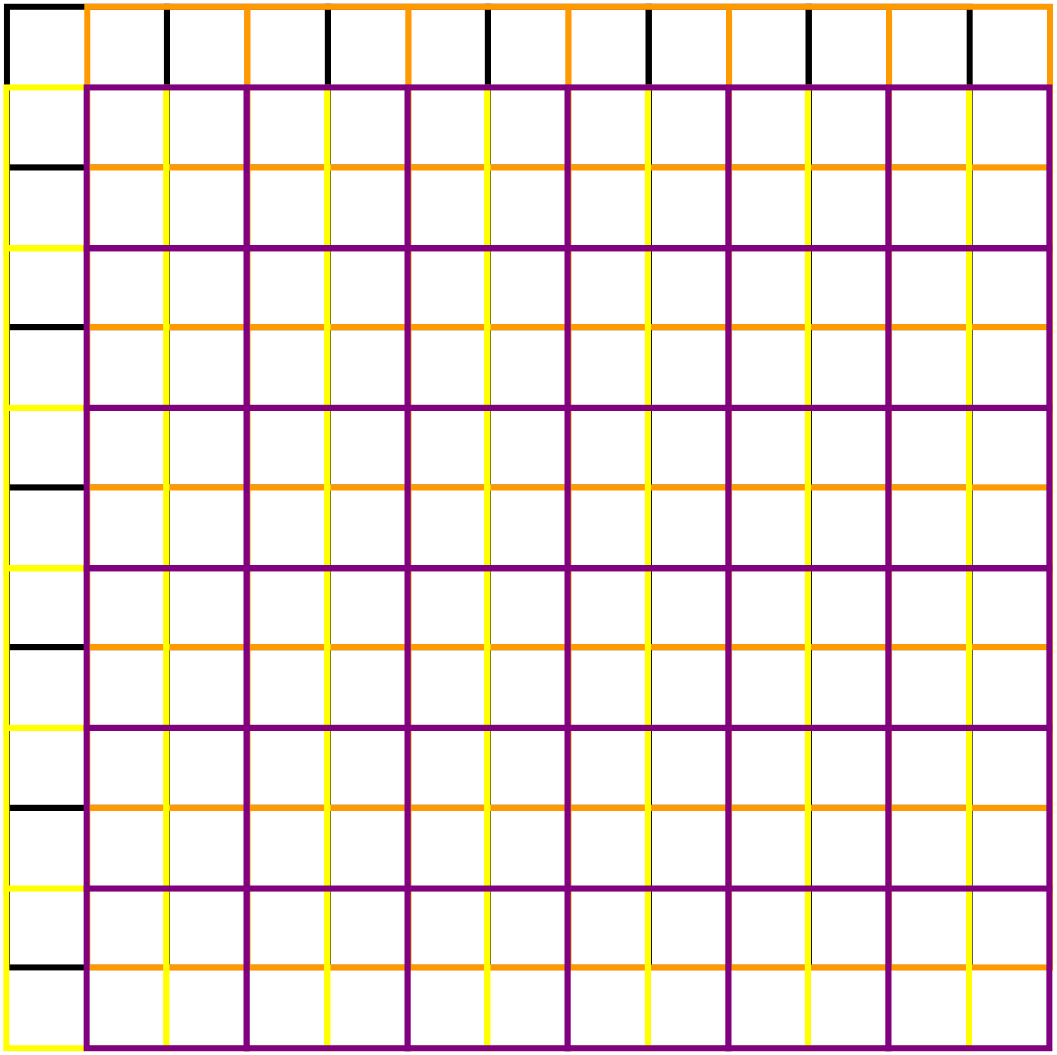
$$\boxed{y_2 = 0 \cdot f_1 + a_1 f_2 + 0 \cdot f_3 + a_3 f_4 + e_2}$$

$$\boxed{y_3 = 0 \cdot f_1 + 0 \cdot f_2 + a_1 f_3 + a_2 f_4 + e_3}$$

$$\boxed{y_4 = 0 \cdot f_1 + 0 \cdot f_2 + 0 \cdot f_3 + a_1 f_4 + e_4}$$

Low vs High Res Pixels

x2 enhancement:
Need 4 frames.



The Mathematical Model

k-th frame $\longrightarrow$ $\underline{\mathbf{y}}_k = A_k \underline{\mathbf{f}} + \underline{e}_k \quad for \quad 1 \le k \le p$

$$A_k = DHF(v_k)$$

Downsampling

Blurring

Warping

$$\underline{\mathbf{y}} = A(\underline{\mathbf{v}})\underline{\mathbf{f}} + e$$

- Statistical estimation problem
- The system is typically underdetermined and ill-conditioned.
  - Need N^2 frames for factor of N enhancement.
- Model is uncertain, and sensitive to unknown parameters.
- Computational complexity is a major concern

The Optimization Problem

Data Info: Builds robustness to model uncertainty

$$\left\{\underline{\hat{\mathbf{f}}}, \underline{\hat{\mathbf{v}}}\right\} = \underset{\underline{\mathbf{f},\mathbf{v}}}{\operatorname{argmin}}\Big[\ \overbrace{\left\| A(\mathbf{v})\underline{\mathbf{f}} - \underline{\mathbf{y}} \right\|_1} + \underbrace{\lambda \sum_{l=-P,}^{P} \sum_{m=-P}^{P} \alpha^{|m|+|l|} \left\| \underline{\mathbf{f}} - S_x^l S_y^m \underline{\mathbf{f}} \right\|_1}\ \Big]$$

$$0 < \alpha < 1$$

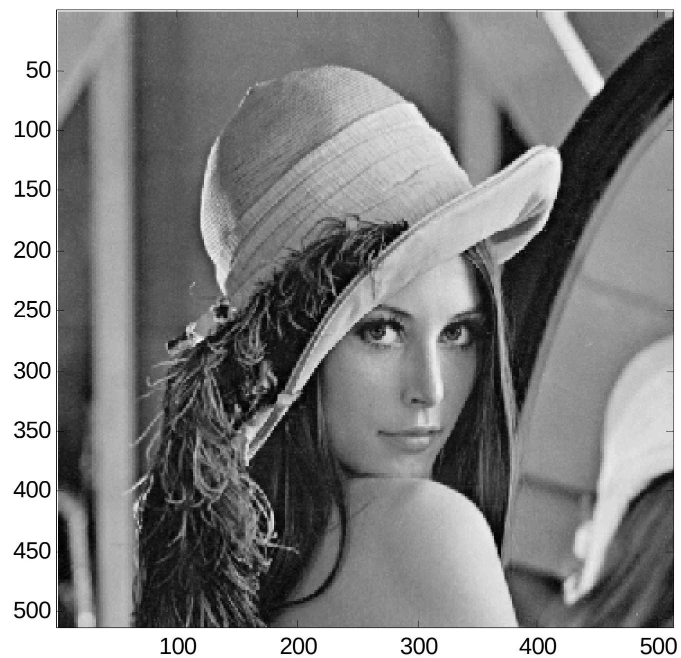
L$_1$ Prior: Incorporates multiscale model of edges

S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "**Fast and Robust Multi-frame Super-resolution**", *IEEE Transactions on Image Processing, vol. 13, no. 10, pp. 1327-1344 , October 2004*

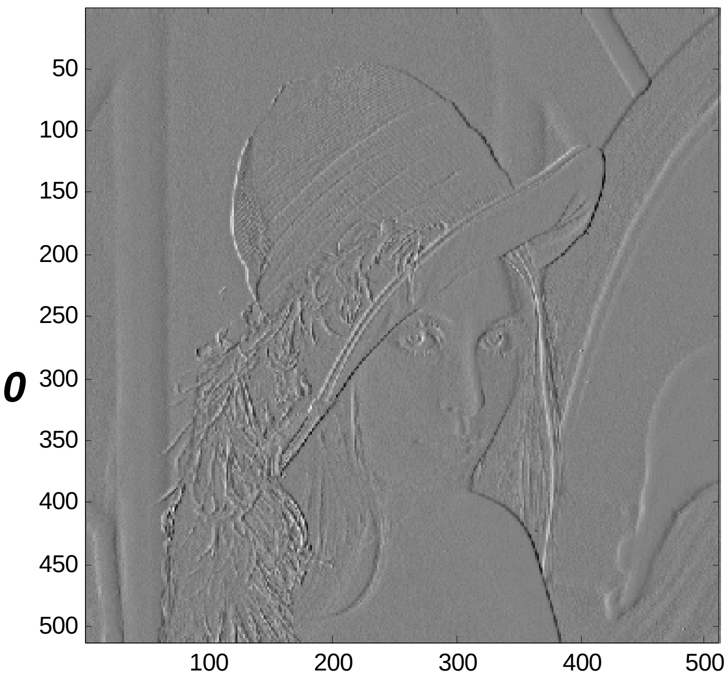
Why this L_1 prior?

- Let's look at pixel differences across scales $$\mathbf{I}_{l,m} = \mathbf{f} - S_x^l S_y^m \mathbf{f}$$

$\boldsymbol{f}$

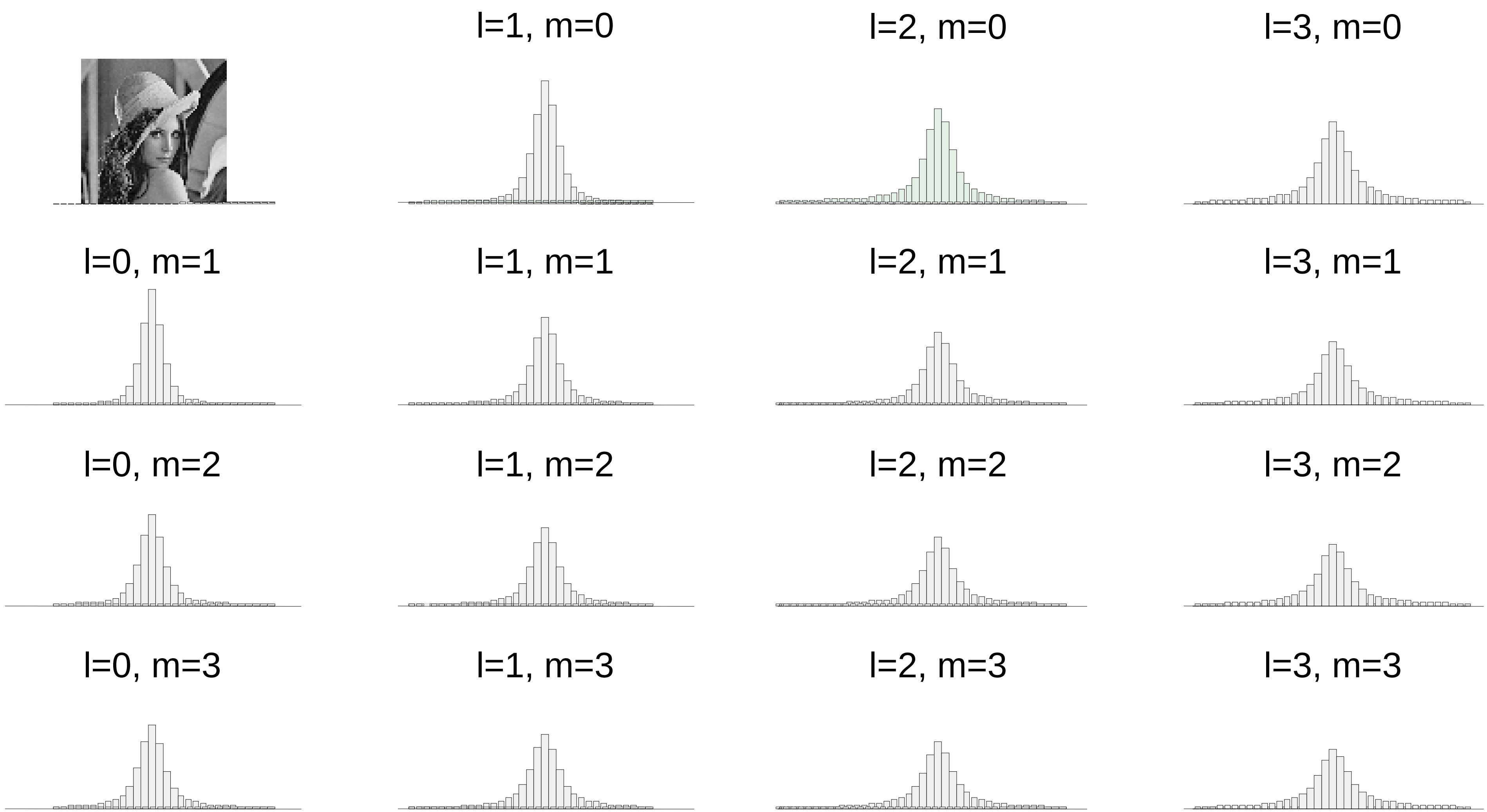


$\boldsymbol{I}_{1,0}$

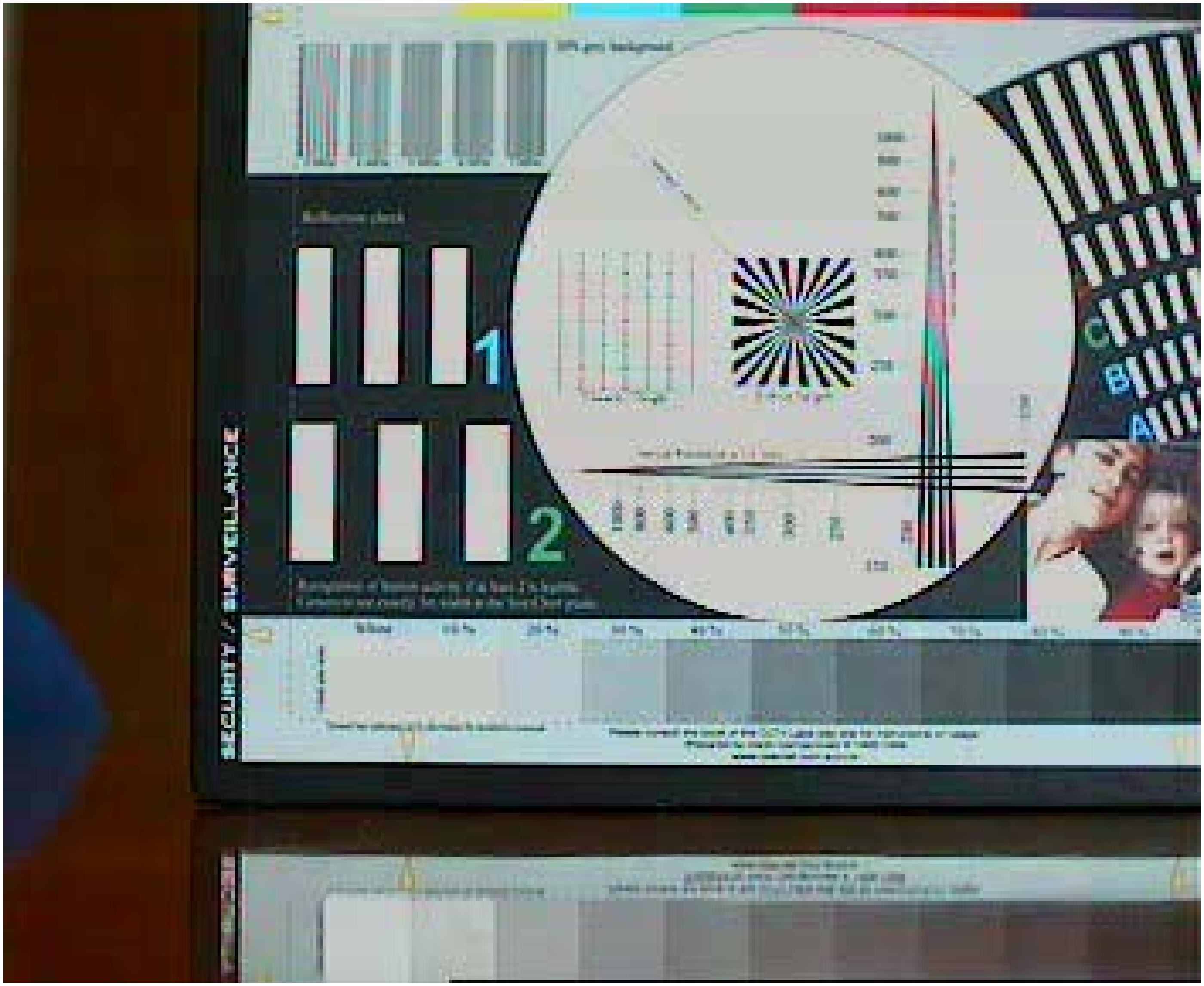


P. Milanfar, and D. Odom, "**Modeling Multiscale Differential Pixel Statistics with Applications**", SPIE Electronic Imaging Symposium: Conference on Computational Imaging, January 2006, San Jose, CA

Histograms of $\mathbf{I}_{l,m} = \underline{\mathbf{f}} - S_x^l S_y^m \underline{\mathbf{f}}$

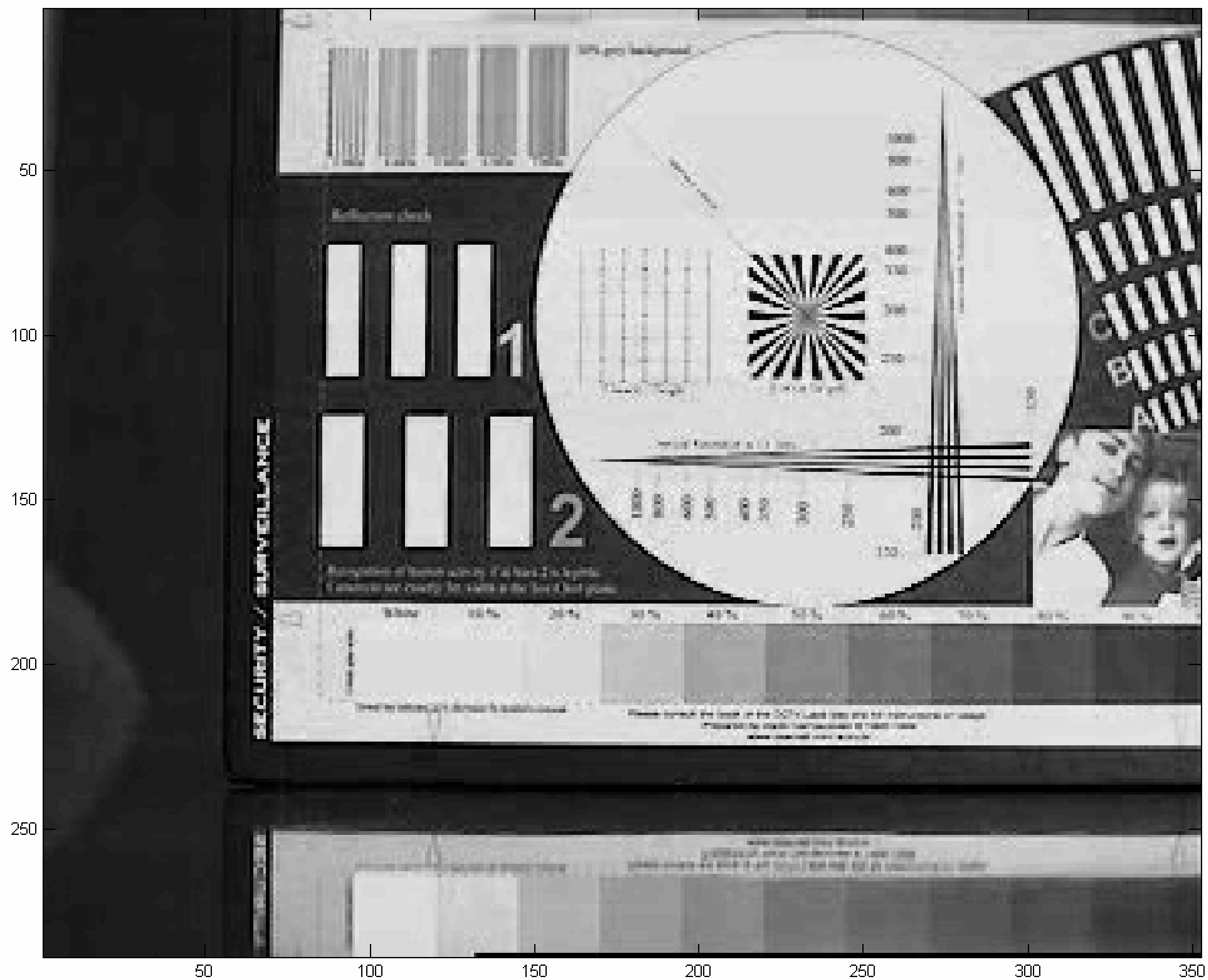


Milanfar et al. EE Dept, UCSC



Data Courtesy Vigilant Technology

Before

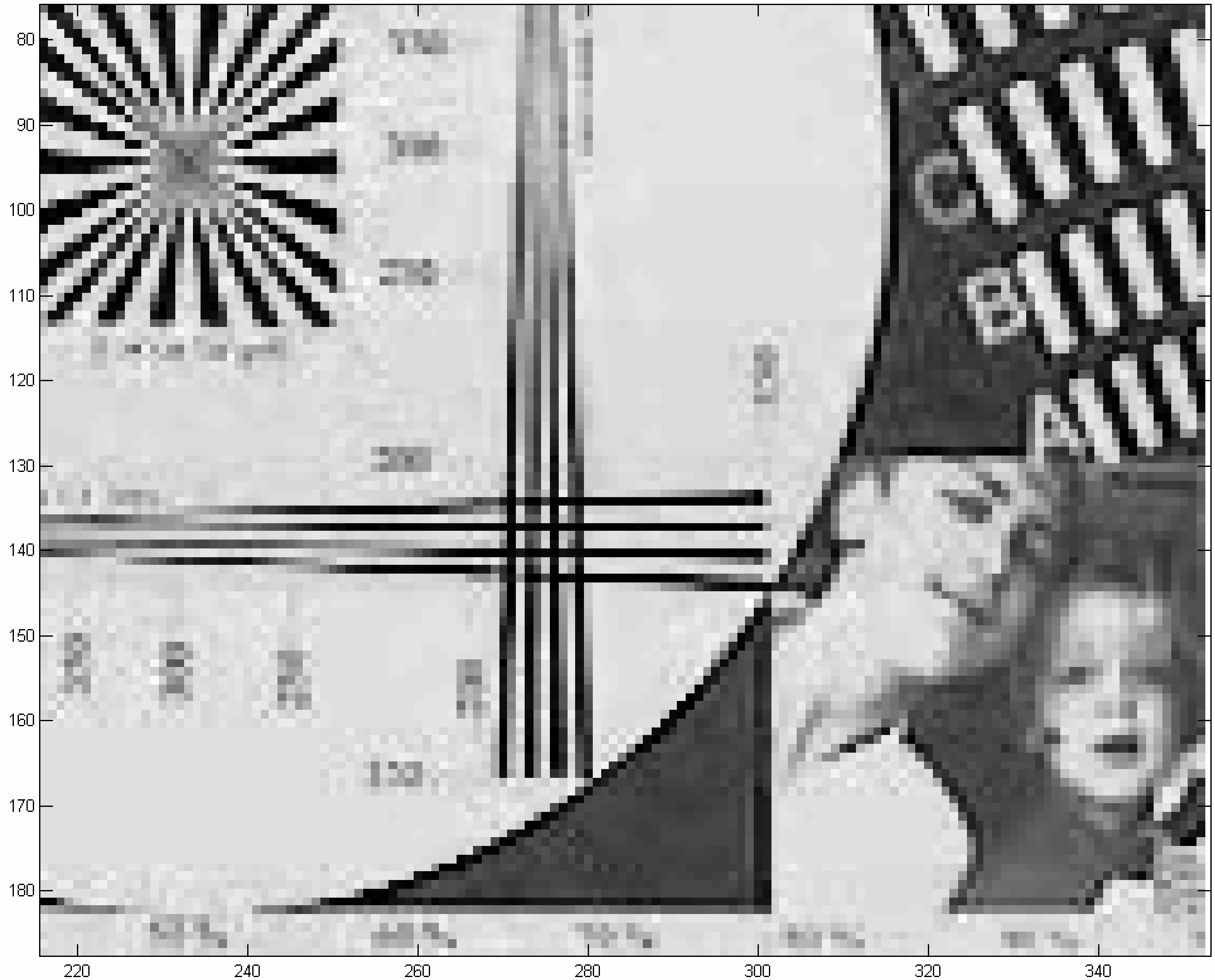


50
100
150
200
250
50
100
150
200
250
300
350
SECURITY / SURVEILLANCE
1
2

After: 4x

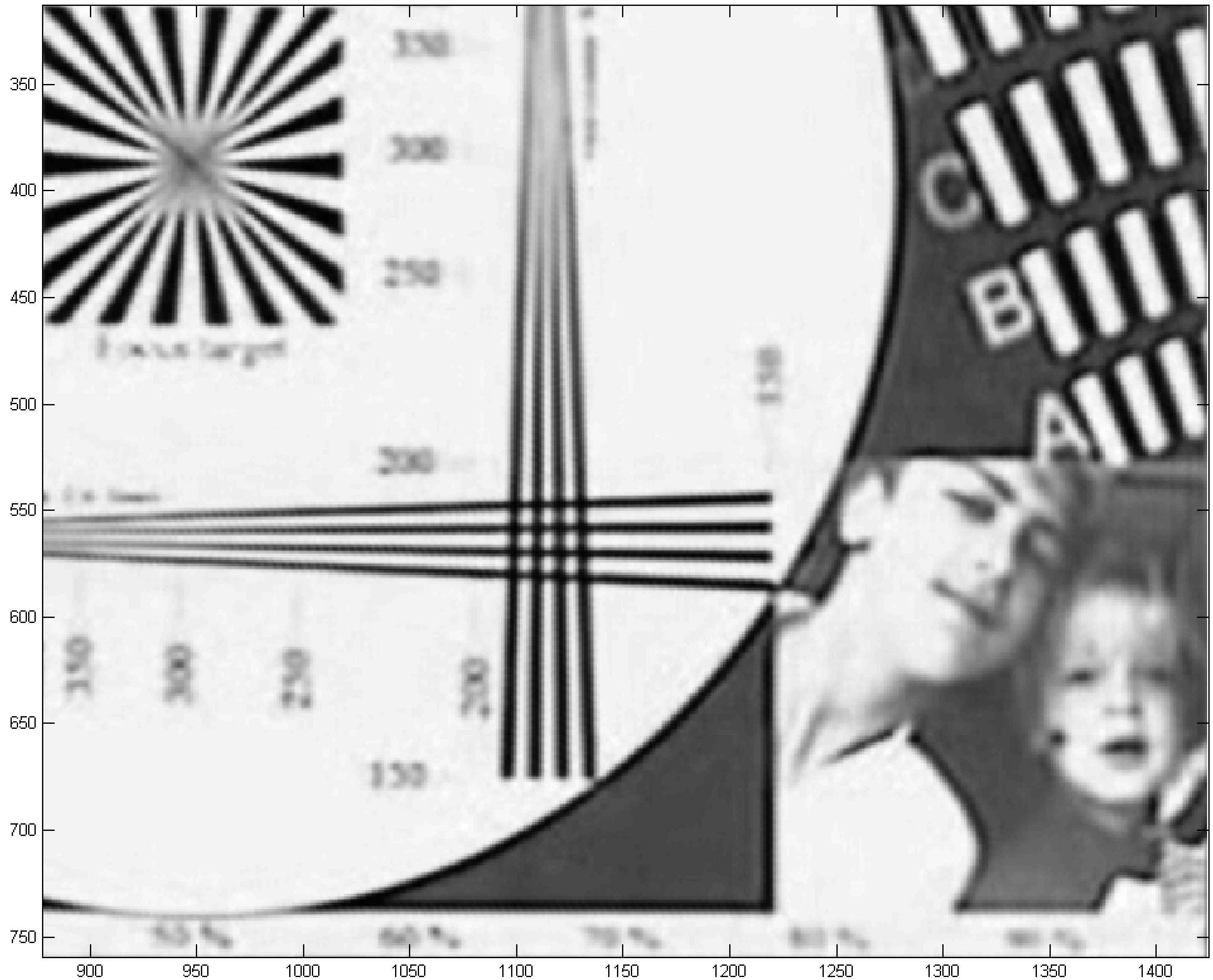

100
200
300
400
500
600
700
800
900
1000
1100
200
400
600
800
1000
1200
1400
SECURITY / SURVEILLANCE
1
2
A
B
C

Detail Before



Data Courtesy Vigilant Technology

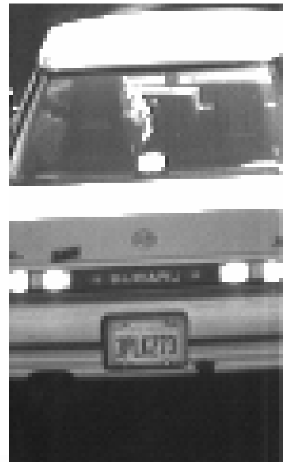
Detail After

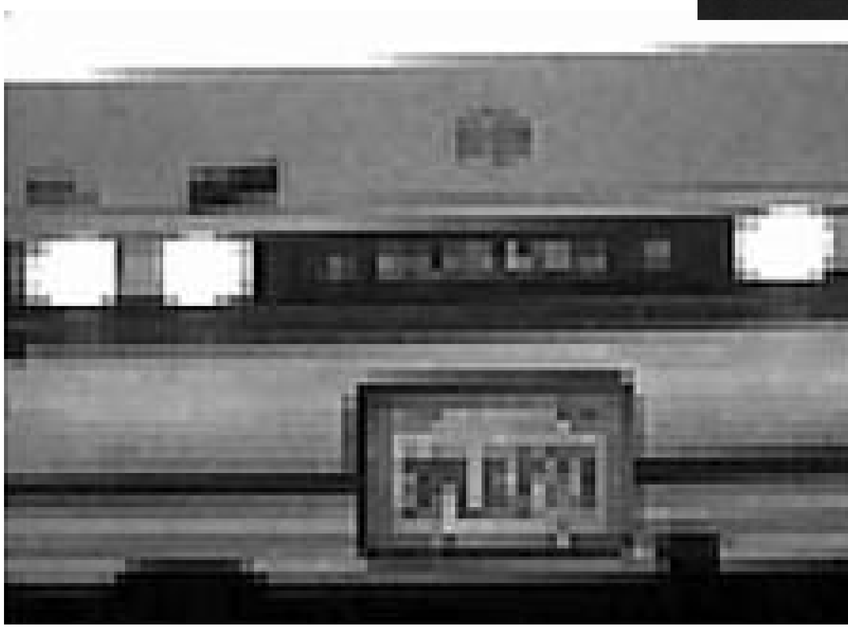


Data Courtesy Vigilant Technology

Security Camera (before/after)

60 input frames





Milanfar et al. EE Dept, UCSC

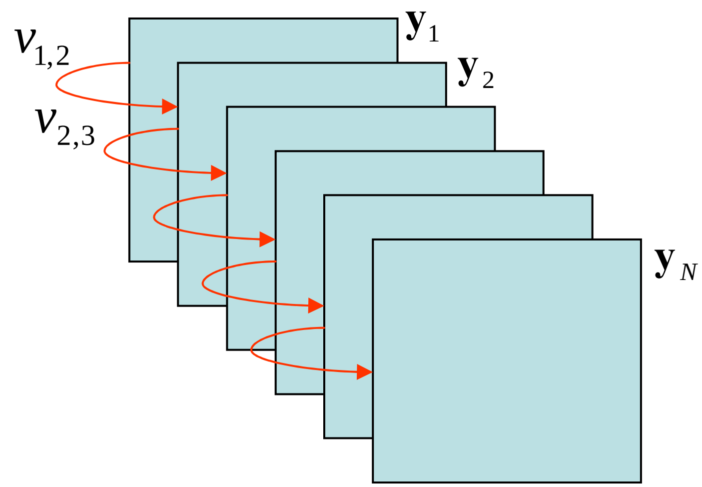
Processing Limits: Statistical Bounds on Super-Resolution Performance

Review Basic Formulation

- Consider a sequence of noisy, translating images $\{\mathbf{y}_k\}$ over time.

$$\mathbf{y}_k = \text{Translate}\,(\mathbf{y}_{k-1}, v_{k-1,k}) + \text{error}$$

Frame-to-frame motion vectors



- Image formation model:

$$\mathbf{y}_k = \text{Sample}[\,f(x, y, t_k) * h(x, y)] + \text{noise}$$

Aliasing

Point-spread function

Fusion of Multiple Video Frames

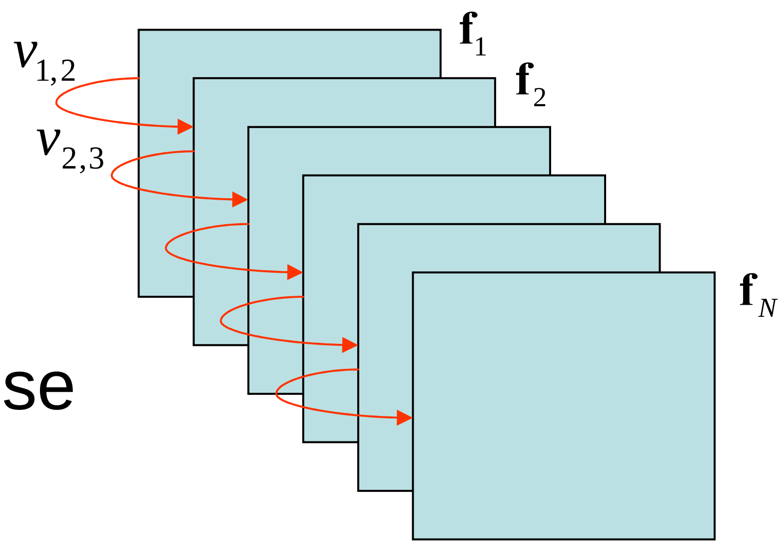
- Reconstruction Problem: Given the frames, estimate the high resolution image $f(x, y, t)$. (Superresolution)
  - Implicit problem: Estimate the motion vectors from aliased images

Nuisance Parameters

$$\mathbf{f}_k = \text{Translate}\,(\mathbf{f}_j, \boxed{v_{j,k}}) + \text{error}$$

$$\mathbf{f}_k = \text{Sample}[\boxed{f(x, y, t_k)} * h(x, y)] + \text{noise}$$

Desired unknowns



Milanfar et al. EE Dept, UCSC

Registration of Multiple Aliased Images

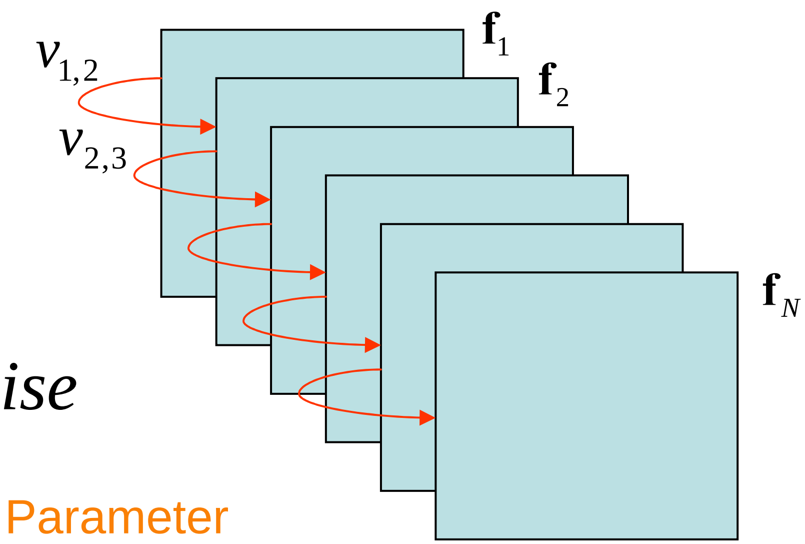
- Motion Problem: Given the frames, estimate vectors $\{v_{j,k}\}$
  - Implicit problem: Estimate underlying high resolution image from aliased data

Desired unknowns

$$\mathbf{f}_k = \text{Translate}\,(\mathbf{f}_j, \boxed{v_{j,k}}) + \text{error}$$

$$\mathbf{f}_k = \text{Sample}[\boxed{f(x,y,t_k)} * h(x,y)] + noise$$

Nuisance Parameter



Registering Aliased Images : A Very Poorly Understood Problem

How well can the problem be solved?

Estimation approach: Look at the Fisher Information (hence the Cramer-Rao bound).

Registration Information

Information "Correlation"

$$\mathbf{J}\left(\{v_{j,k}\}, f\right) = \begin{bmatrix} \mathbf{J}_{vv} & \mathbf{J}_{fv} \\ \mathbf{J}_{fv}^{T} & \mathbf{J}_{ff} \end{bmatrix}$$

Reconstruction Information

$\mathbf{J}_{vv}$ - Depends on the <u>set of motions</u> (sampling offsets) and the amount of <u>texture energy</u> in the signal

$\mathbf{J}_{ff}$ - Depends only on the <u>set of motions</u>

CRB for Aliased Image Registration

Using Schur decomposition, the CRB for aliased image registration is:

$$\text{Cov}\left(\{v_{j,k}\}\right) \geq \left(\mathbf{J}_{vv} - \mathbf{J}_{fv}\mathbf{J}_{ff}^{-1}\mathbf{J}_{fv}^{T}\right)^{-1}$$

Registration Information

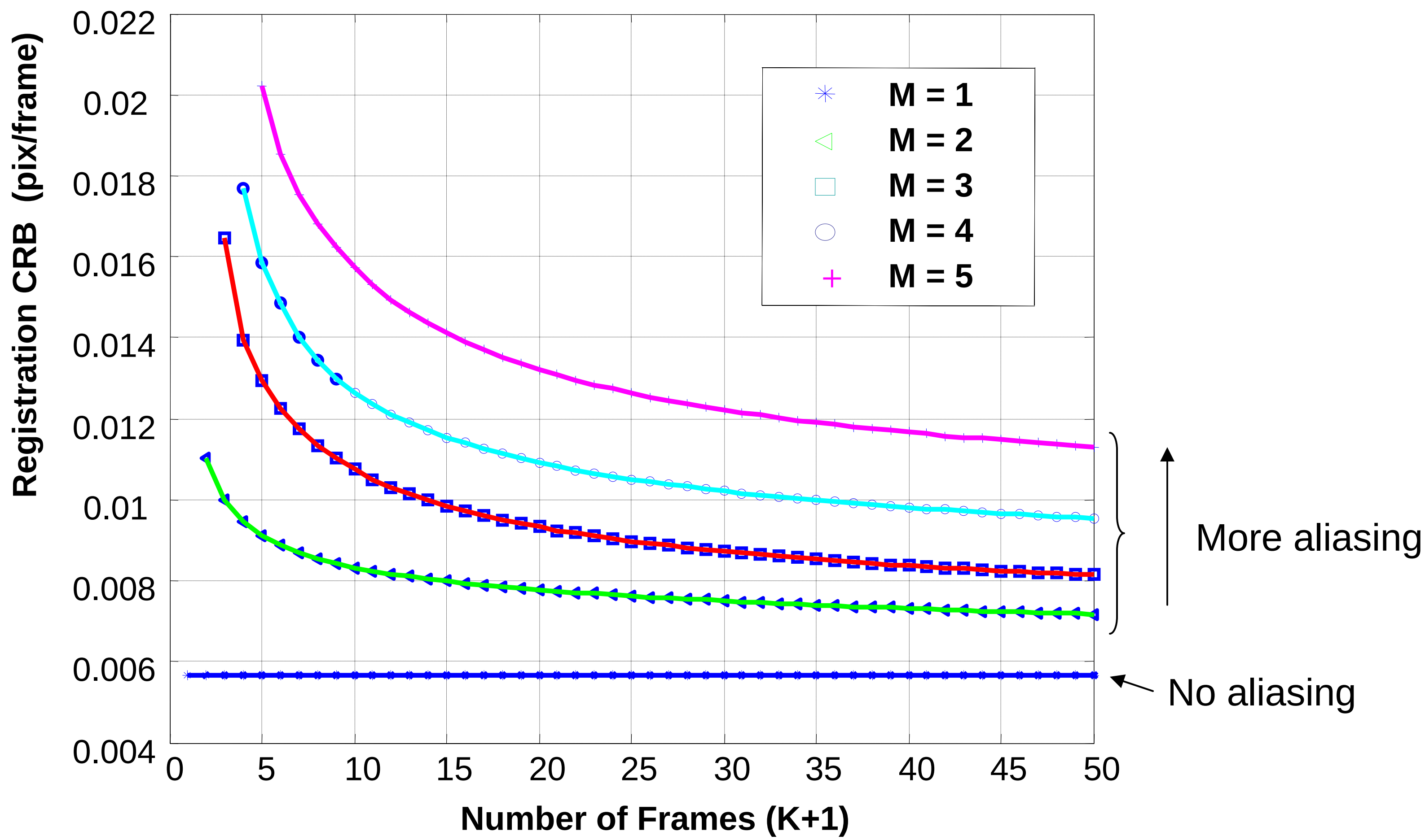
Information Loss due to uncertainty about the high resolution image.

With just a pair of aliased images, the FIM is generically singular, hence unbiased <u>pairwise</u> registration of aliased images is essentially impossible. (Not so in absence of aliasing!)

D. Robinson, and P. Milanfar, **"Statistical Performance Analysis of Super-resolution",** To appear in *IEEE Transactions on Image Processing*

Registering <u>Sets</u> of Images

CR Bound (per frame) for multi-frame image registration.

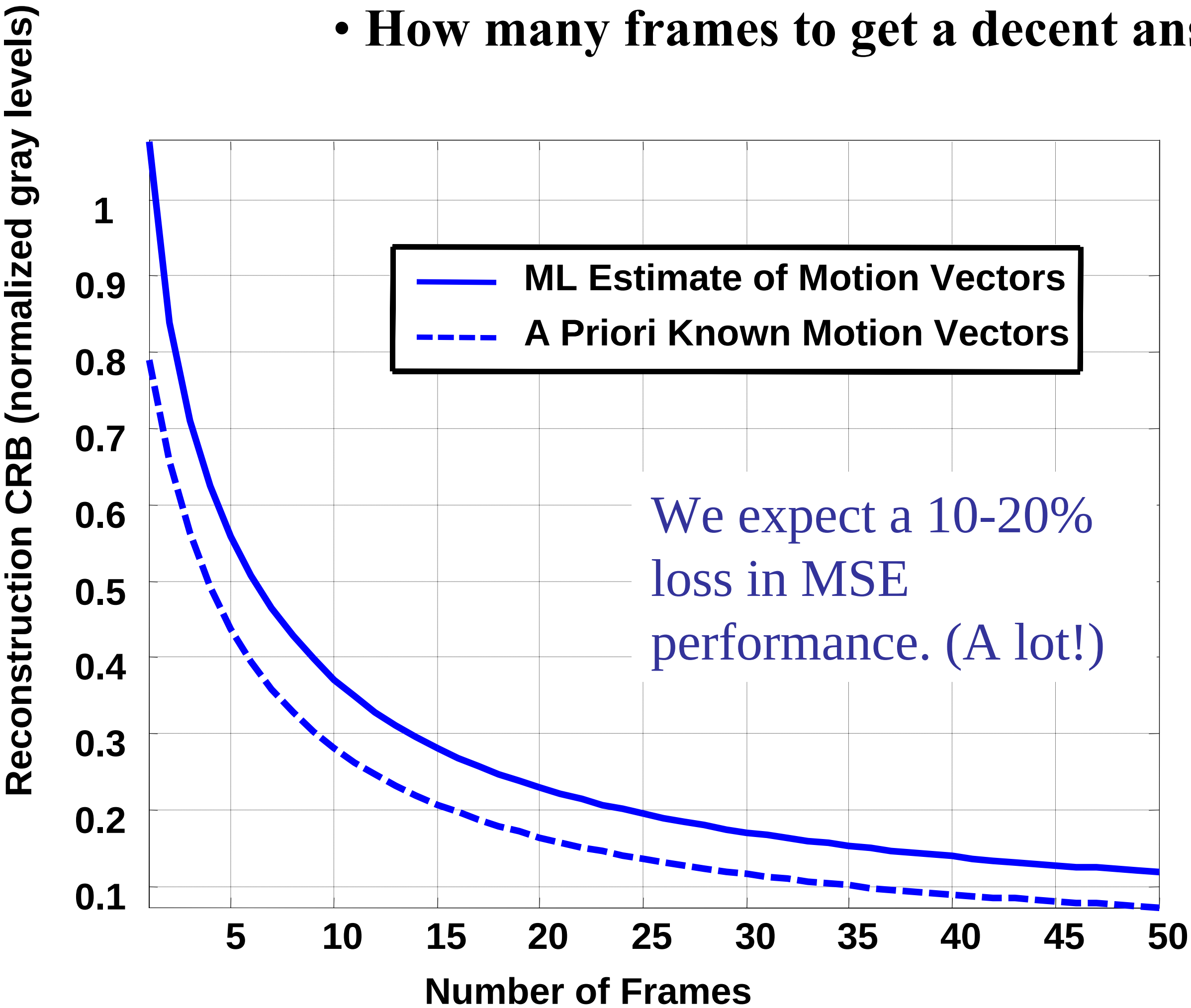


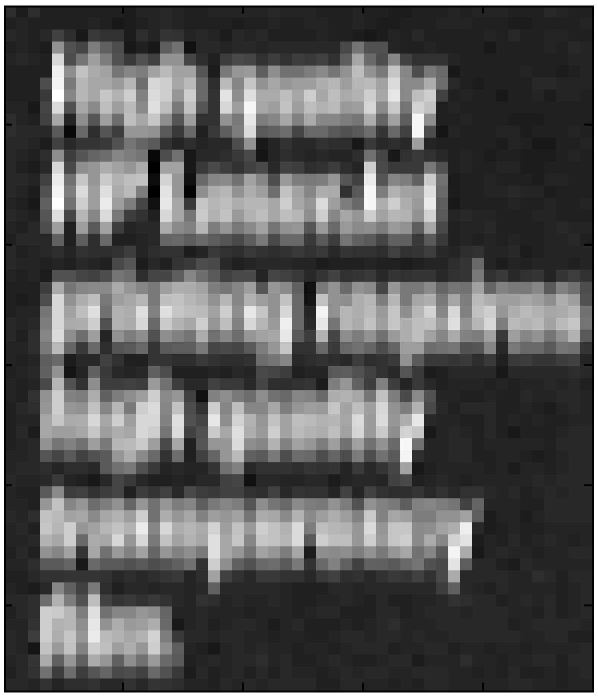
Milanfar et al. EE Dept, UCSC

Insights gained:

• **How much information is lost by needing to estimate the motion vectors?**

• **How many frames to get a decent answer?**





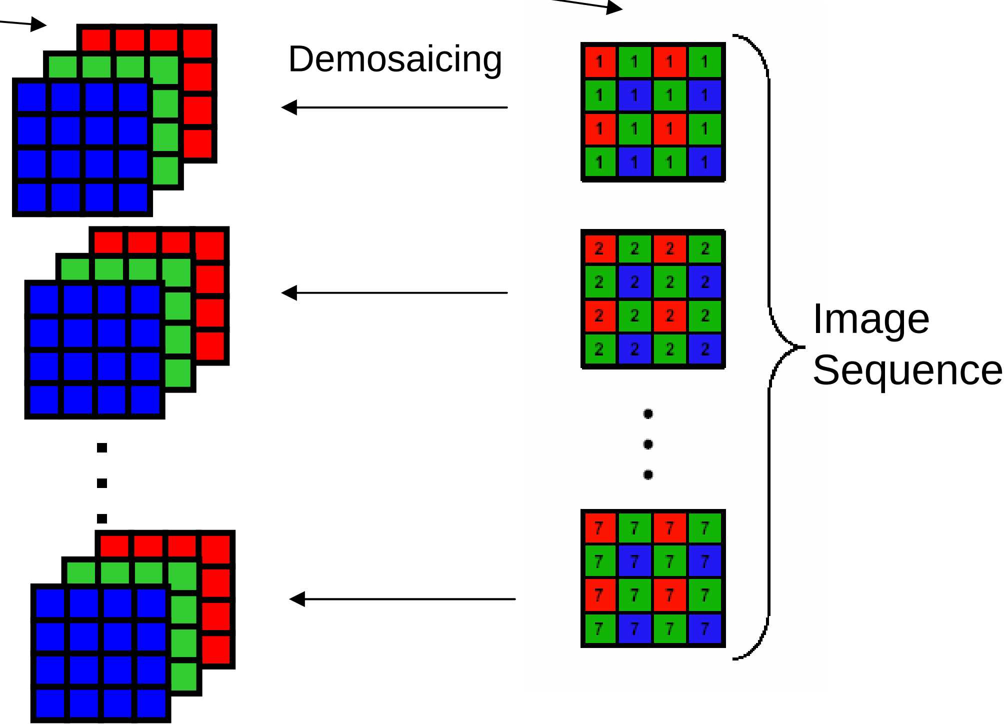
Milanfar et al. EE Dept, UCSC

Further Extensions: Color, Dynamics, Algorithmic Improvements

Color Super-Resolution

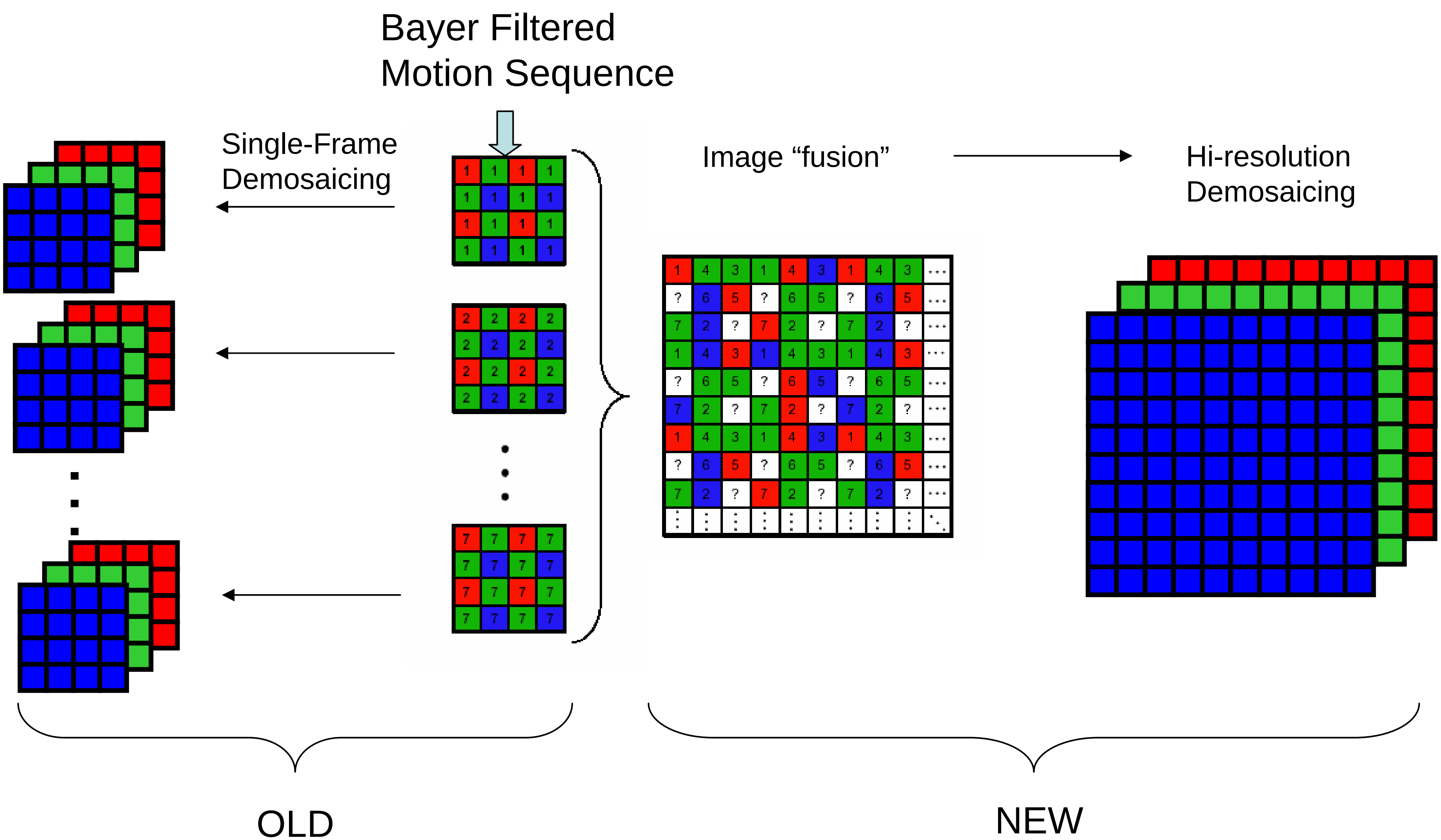
- Two types of input to consider:
  – Raw CFA data
  – Full RGB fields

- Unified Treatment



Milanfar et al. EE Dept, UCSC

Simultaneous Demosaicing and Super-Resolution



Milanfar et al. EE Dept, UCSC

Characteristics of Color Algorithm

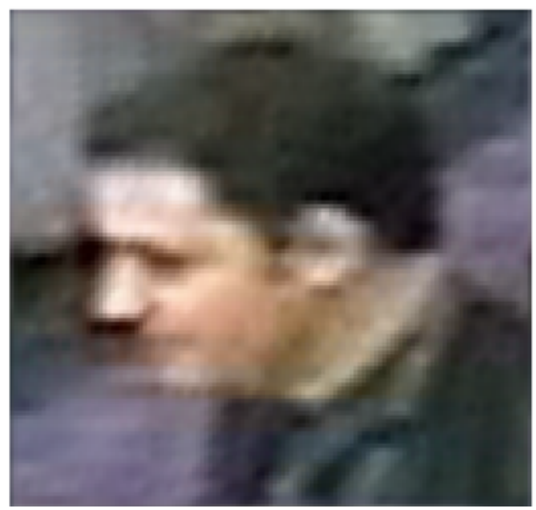
I. Robust to the data noise and motion estimation errors (L_1 Norm).
II. Sharp edges in luminance component (L_1).
III. Minimize artifact in the chrominance component (L_2 Norm).
IV. Similar edge location-orientation in all color bands.

S. Farsiu, M. Elad, and P. Milanfar, "**Multi-Frame Demosaicing and Super-Resolution of Color Images**", To appear in *IEEE Trans. on Image Processing*

S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "**Advances and Challenges in Super-Resolution**", *International Journal of Imaging Systems and Technology, Special Issue on High Resolution Image Reconstruction, vol. 14, no. 2, pp. 47-57, 2004.*

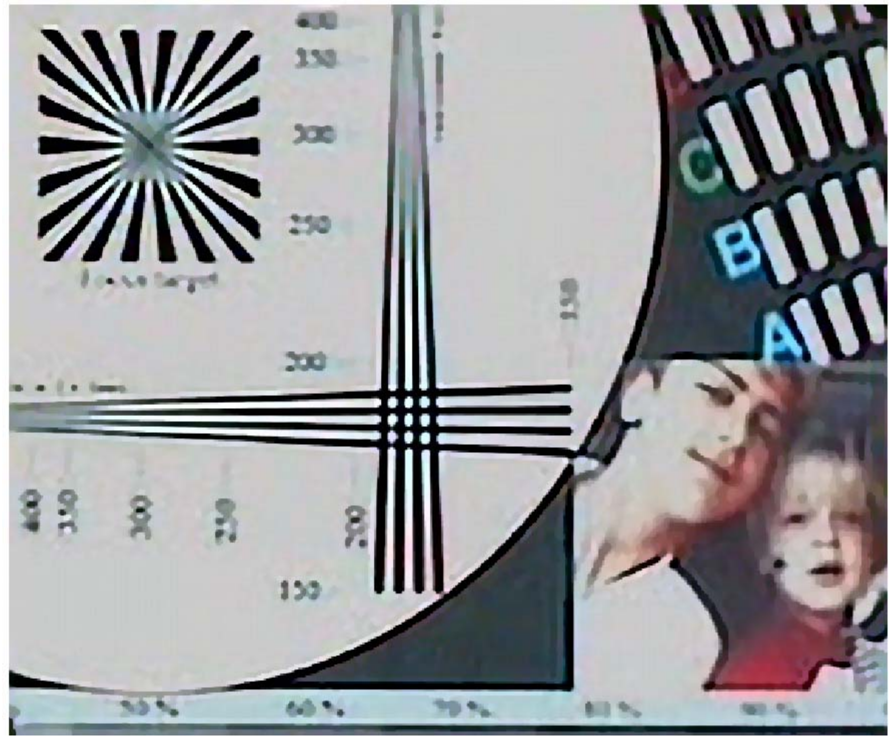
RGB Color Security Camera

24 input frames



Milanfar et al. EE Dept, UCSC

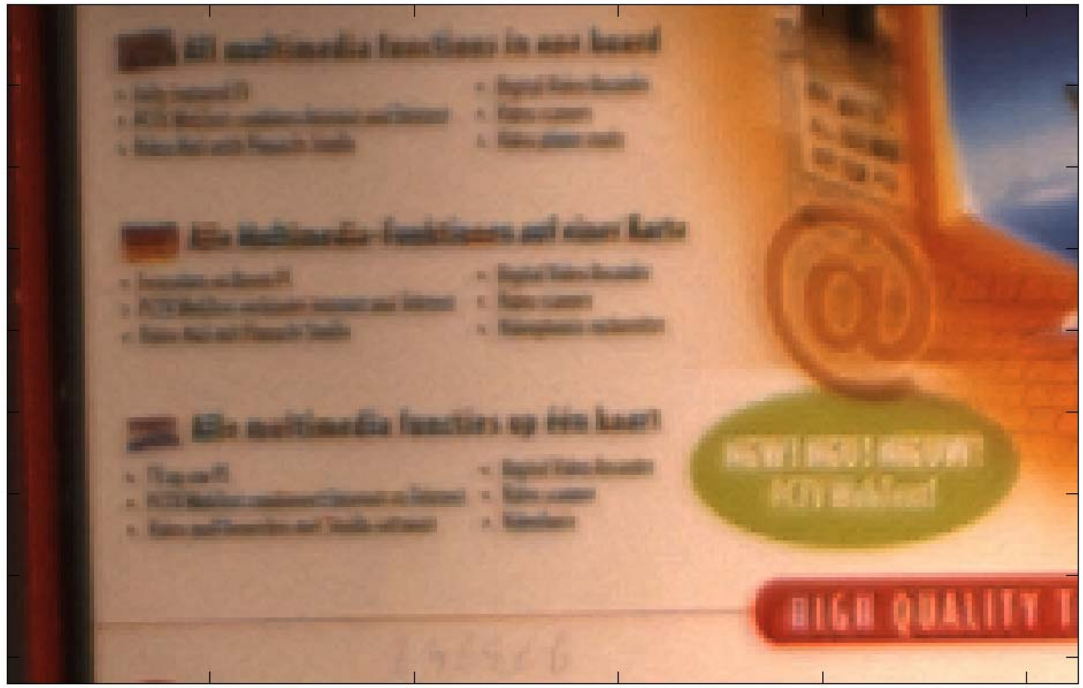
RGB Color Super-Resolution



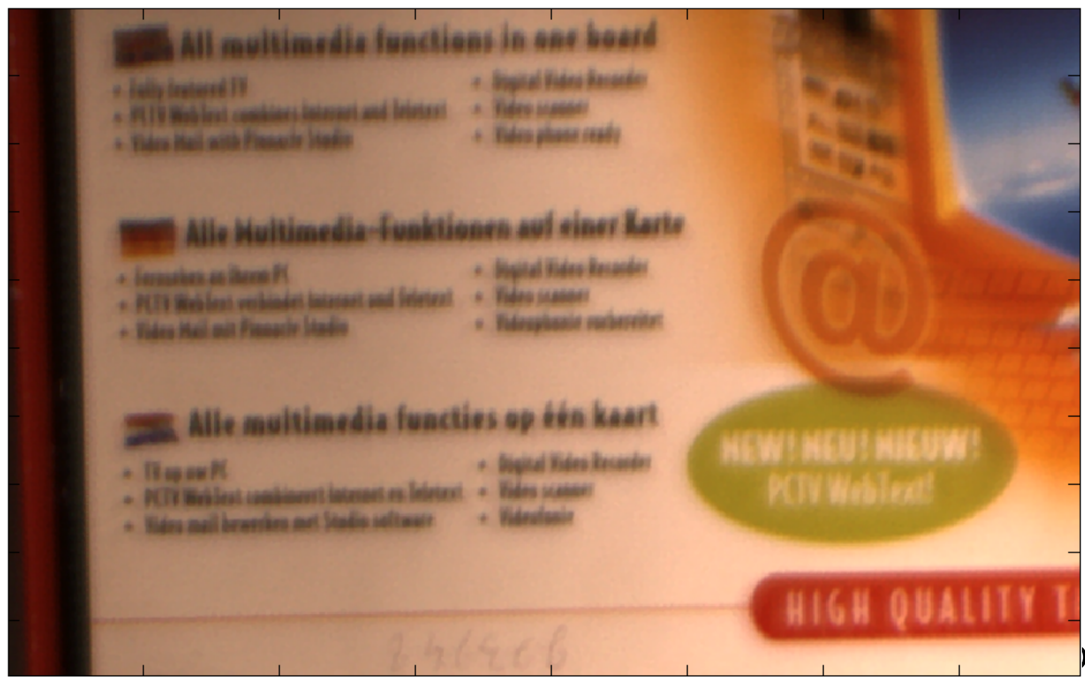
40 input frames, resolution enhancement factor of x4

Milanfar et al. EE Dept, UCSC

Demosaiced from 1-CCD
CFA Camera



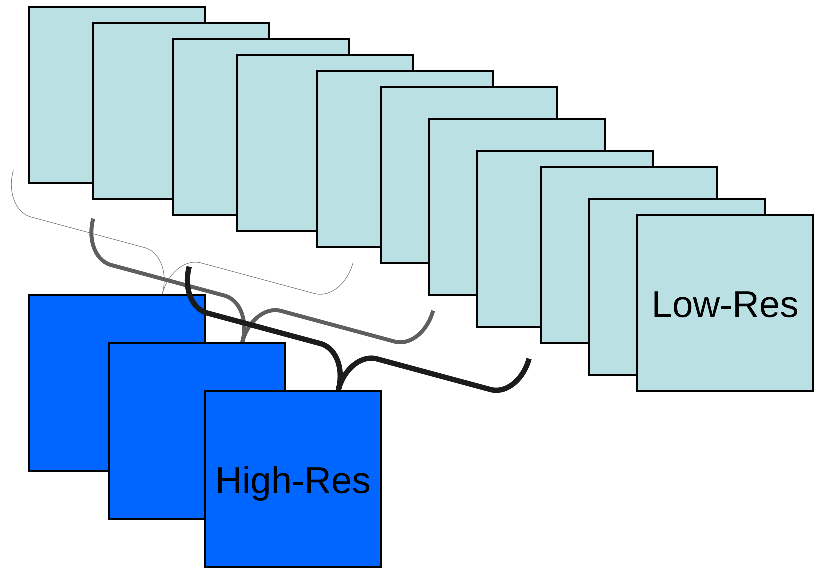
24-Frame Demosaicing
and Reconstruction x3



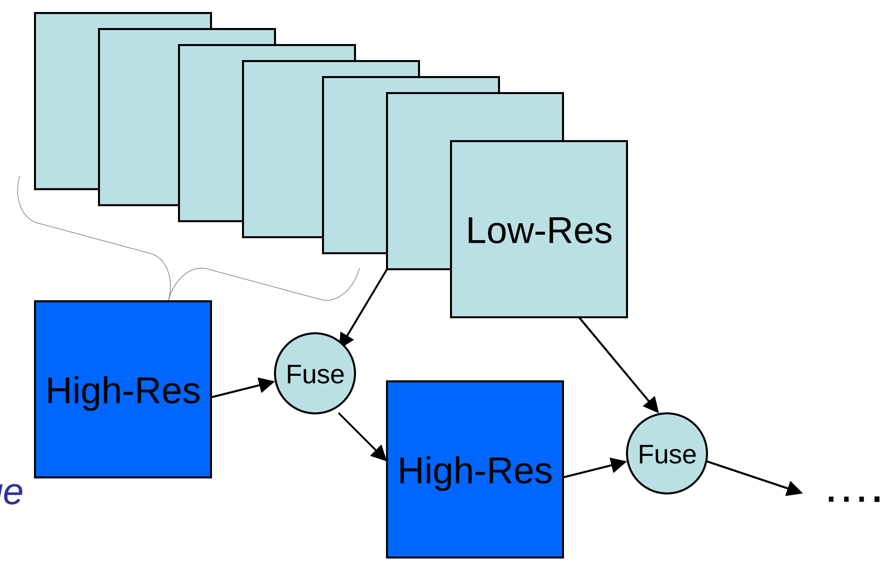
Data courtesy of Technion

Dynamic Super-Resolution

- Naïve Approach:



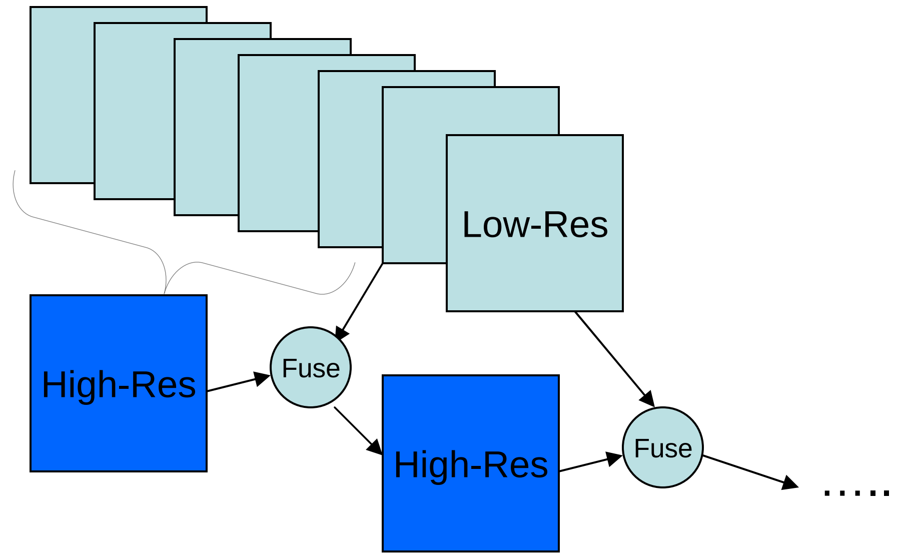
- Right Approach:



M. Elad and A. Feuer, “**Super-Resolution Restoration of Continuous Image Sequence – Adaptive Filtering Approach**”, *IEEE Trans. on Image Proc.*, Vol. 8. Number 3, pp. 387-395, March 1999

Dynamic Super-Resolution

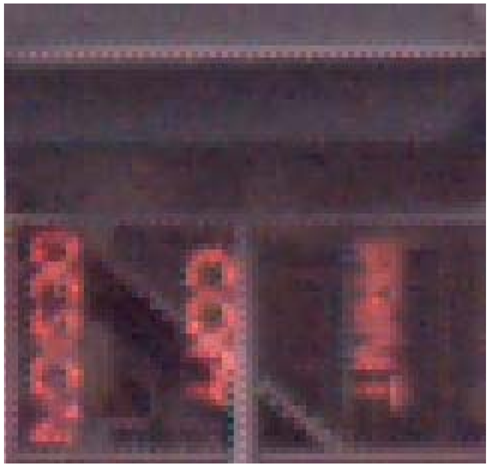
- Adapted for color
- Improved robustness
- Different implementation



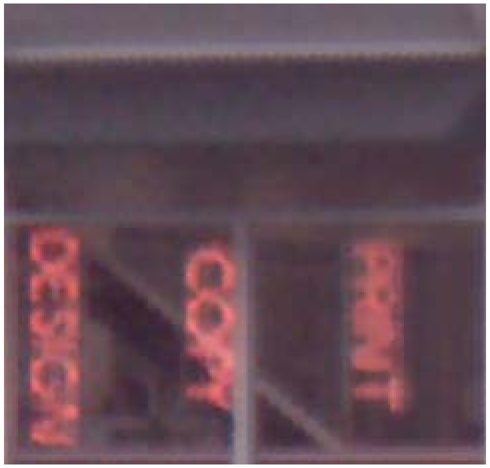
S. Farsiu, M. Elad, and P. Milanfar, "**Video-to-Video Dynamic Superresolution for Grayscale and Color Sequences**," To appear in *EURASIP Journal of Applied Signal Processing, Special Issue on Superresolution Imaging*

Video-to-Video Example I

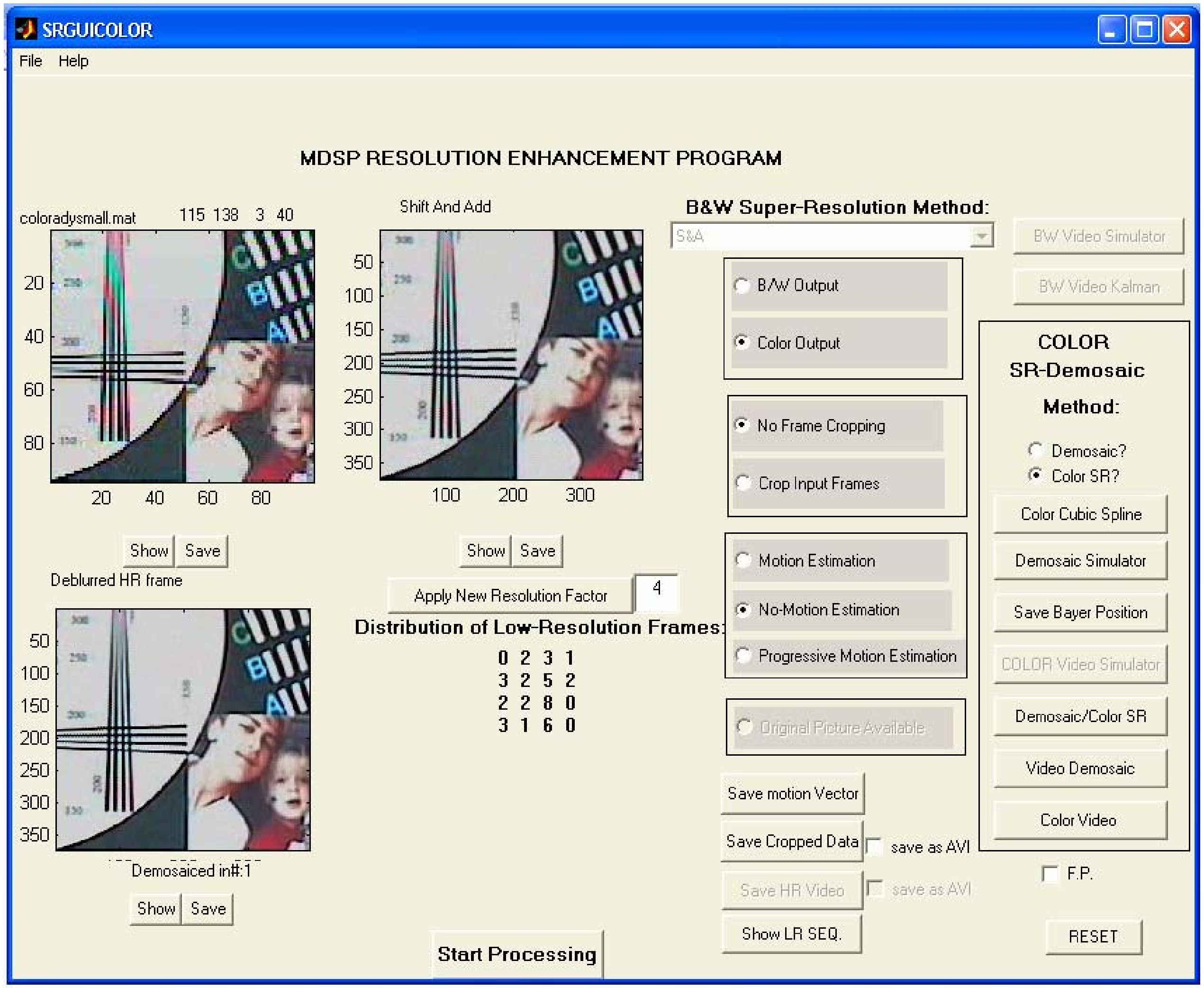
LR Video



HR Outcome



Software



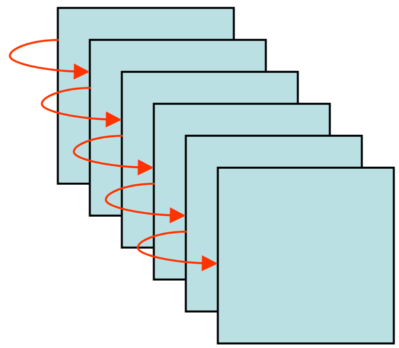
MotionDSP (Milanfar, Farsiu, Elad)

Further Refinements

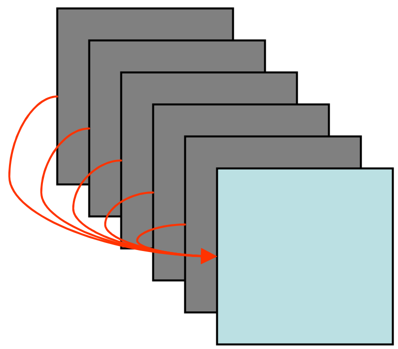
- Computationally, it still makes the most sense to solve the motion/fusion problems in series.
  - Need extremely accurate motion estimation.
  - Need excellent filtering, interpolation.

Better Motion Estimation

- Almost all motion estimation algorithms today deal with the case of only two (consecutive) frames at a time.



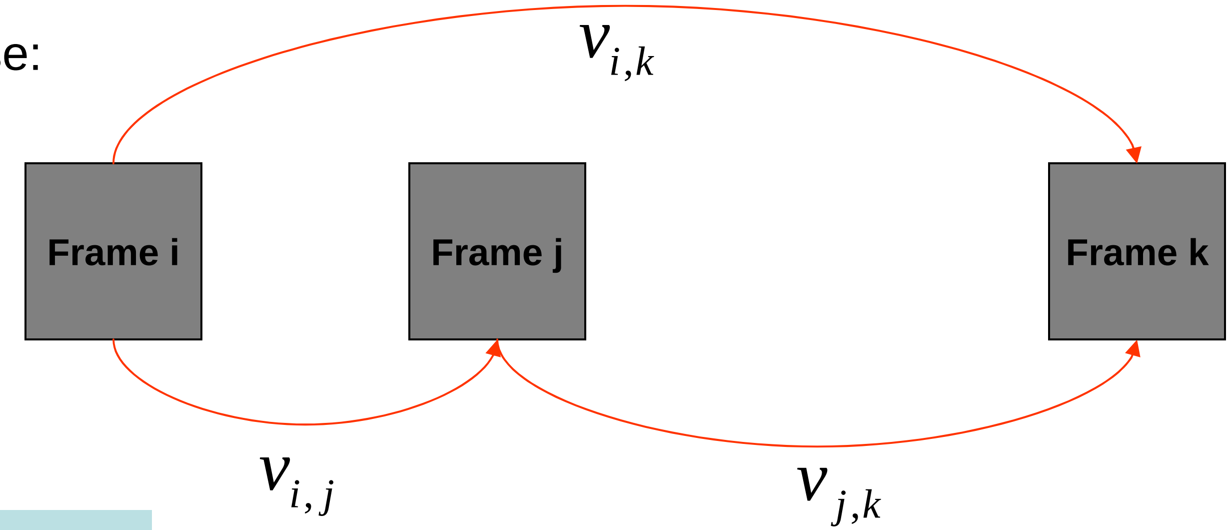
Pairwise estimation
("Progressive")



Fixed reference estimation
("Anchored")

Constraints on Motion Vectors Across Time

Translation case:



$$v_{i,k} = v_{i,j} + v_{j,k}$$

$$v_{k,j} = -v_{j,k}$$

$$v_{i,i} = 0$$

- Linear set of constraints imply that the motion vectors live in a subspace.
- For more general motion, you have nonlinear constraints, → group structure.
- Applicable to any core algorithm.

S. Farsiu, M. Elad, P. Milanfar, "**Constrained, Globally Optimal, Multi-frame Motion Estimation**," Proc. of the 2005 IEEE Workshop on Statistical Signal Processing, Bordeaux, France, July 2005

Better Filtering and Interpolation: The Kernel Regression Idea

Data:

$$y_i = z(\mathbf{x}_i) + \varepsilon_i, \qquad \mathbf{x}_i = [x_1, x_2]_i^T$$

Measurement

Regression function

Local Polynomial Kernel Regression:

Kernel function

$$\arg\min_{\{\beta_n\}} \sum_{i=1}^{P} \Big[y_i - \underbrace{\beta_0 - \boldsymbol{\beta}_1^T (\mathbf{x}_i - \mathbf{x}) - \boldsymbol{\beta}_2^T \operatorname{vech}\left\{ (\mathbf{x}_i - \mathbf{x})(\mathbf{x}_i - \mathbf{x})^T \right\} - \cdots}_{\text{Taylor expansion of } z(\mathbf{x}_i) \text{ near x}} \Big]^2 K_{\mathbf{H}}(\mathbf{x}_i - \mathbf{x})$$

Smoothing matrix

$z(\mathbf{x})$

Milanfar et al. EE Dept, UCSC

Even Better: Adaptive Kernel Regression

Consider the Denoising Problem First: $\mathbf{x} = \mathbf{x}_j$

Spatial kernel   Radiometric kernel

$$\arg\min_{\{\beta_n\}} \sum_{i=1}^{P} \left[y_i - \beta_0 - \boldsymbol{\beta}_1^T \left(\mathbf{x}_i - \mathbf{x}_j \right) - \cdots \right]^2 \underbrace{K_{\mathbf{H}}(\mathbf{x}_i - \mathbf{x}_j)}_{\text{Spatial kernel}} \underbrace{\boxed{K_g(y_i - y_j)}}_{\text{Radiometric kernel}}$$

Special Case (order=0):

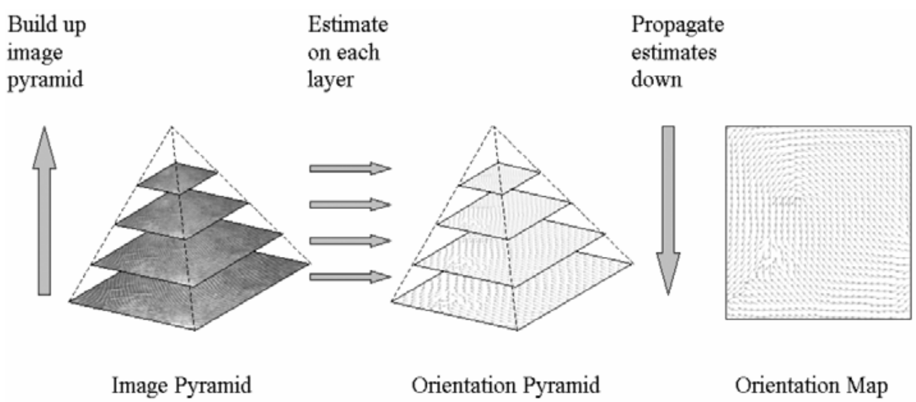
$$\hat{z}\left(\mathbf{x}_j; 0, \mathbf{H}, g\right) = \frac{\sum_{i=1}^{P} K_{\mathbf{H}}\left(\mathbf{x}_i - \mathbf{x}_j\right) K_g\left(y_i - y_j\right) y_i}{\sum_{i=1}^{P} K_{\mathbf{H}}\left(\mathbf{x}_i - \mathbf{x}_j\right) K_g\left(y_i - y_j\right)}$$

$\longrightarrow$ Bilateral filter

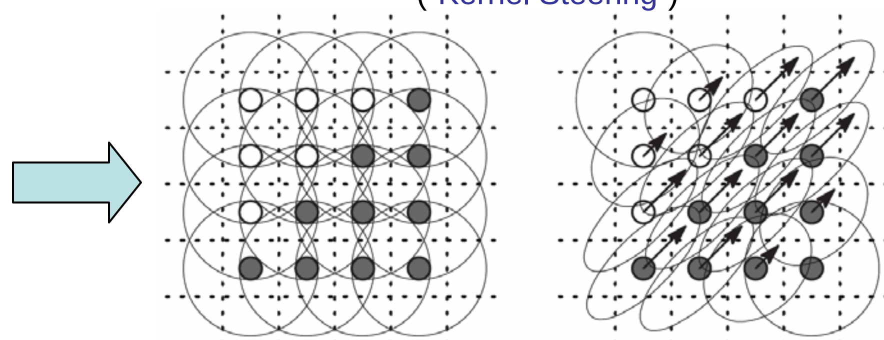
(Gaussian Kernels)
Tomasi ('98)
Elad ('01)

Choice of Radiometric Kernel

- K_g implicitly exploits the local gradient information. (e.g. BF, $K_g = \exp\left(-(y_i - y_j)^2 / \sigma_g^2\right)$)
- Improved solution is possible by explicit incorporation of *orientation info*.



Spread the kernels along local orientations ("Kernel Steering")



Xiaoguang, Feng, P. Milanfar, "**Multiscale Principal Components Analysis for Image Local Orientation Estimation**", Proceedings of the 36th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA. November 2002

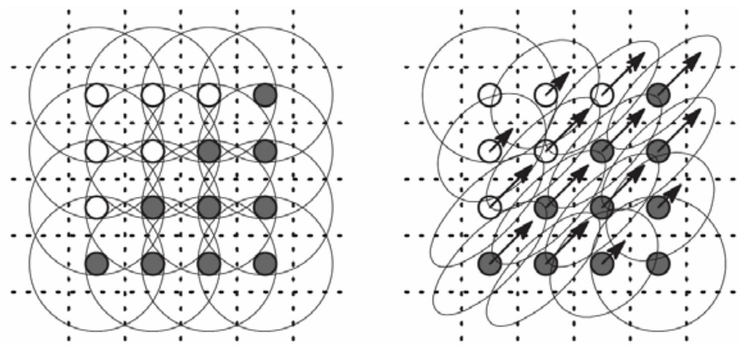
Adaptive Kernel Regression

- Steerable Kernel Regression

Local Steering matrix

$$\arg\min_{\{\beta_n\}} \sum_{i=1}^{P} \left[y_i - \beta_0 - \boldsymbol{\beta}_1^T (\mathbf{x}_i - \mathbf{x}) - \cdots \right]^2 \frac{1}{h\, s_i} K\left\{ \frac{(\mathbf{x}_i - \mathbf{x})^T \mathbf{S}_i (\mathbf{x}_i - \mathbf{x})}{h^2} \right\}$$

Global Scaling parameter



H. Takeda, P. Milanfar, "**Image Denoising by Adaptive Kernel Regression**" Proceedings of the Asilomar Conference on Signals and Systems, Oct. 2005, Pacific Grove, CA

Image Denoising

Original image

Noisy, sigma = 25

Milanfar et al. EE Dept, UCSC

Denoising Results

Bilateral filter,
RMSE = 8.66

Adaptive kernel,
Steered
order = 0
RMSE = 7.04

Standard kernel,
order = 2,
RMSE = 10.21

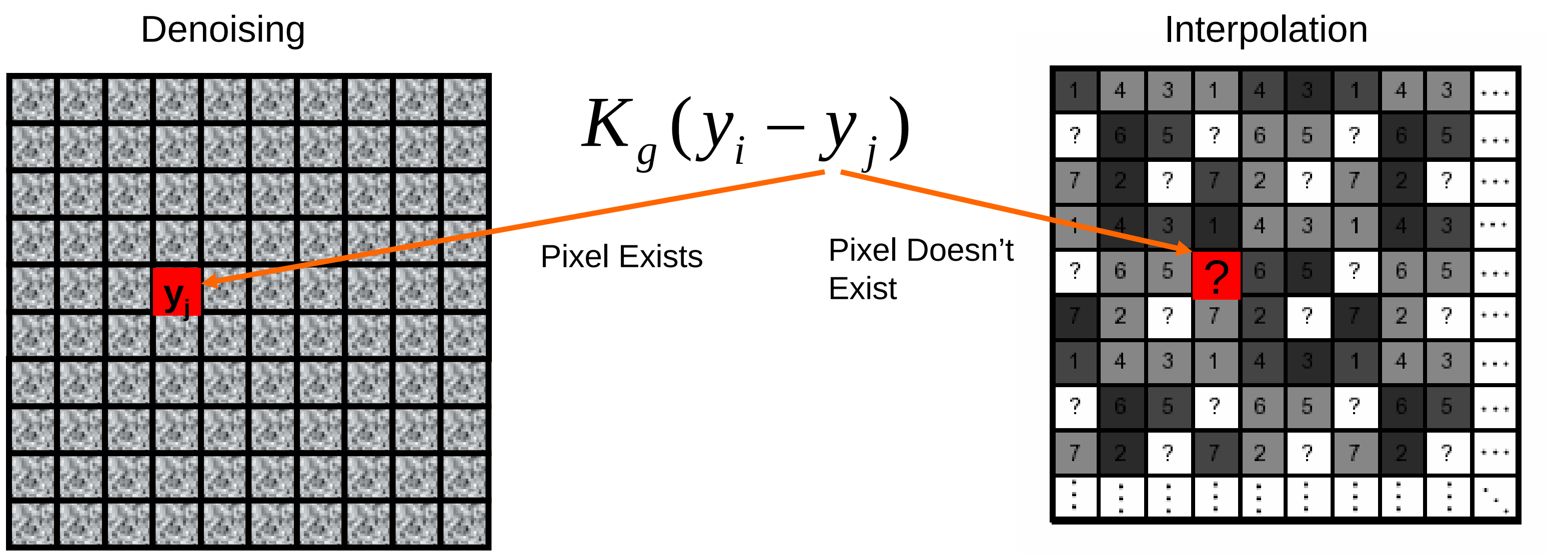
Adaptive kernel,
Steered,
order = 2
RMSE = 7.03

Gaussian kernel

Milanfar et al. EE Dept, UCSC

What About Fully Adaptive Kernel-based Interpolation?

The Problem: At the location x where we wish to interpolate, there is no pixel value (yet)



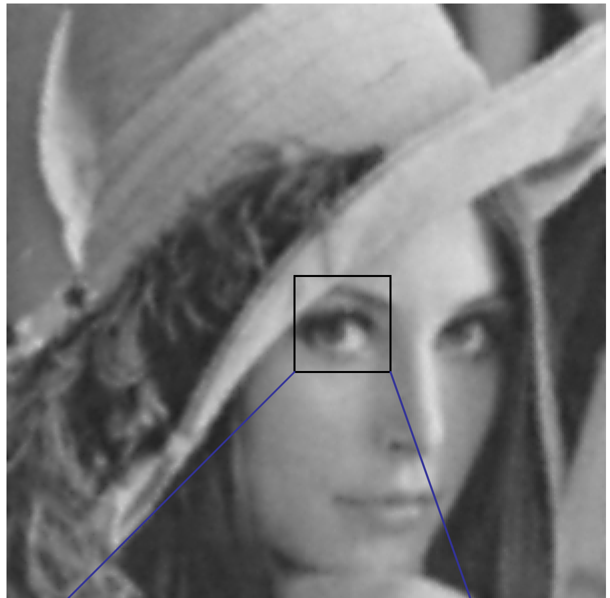
The Solution: Produce a "pilot", low-complexity, estimate of the pixel ? then apply the more sophisticated adaptive kernel techniques described earlier.

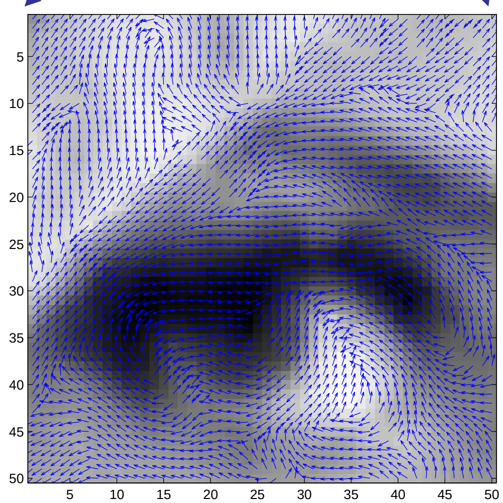
The process can in fact be iterated for further improvement.

Illustration:

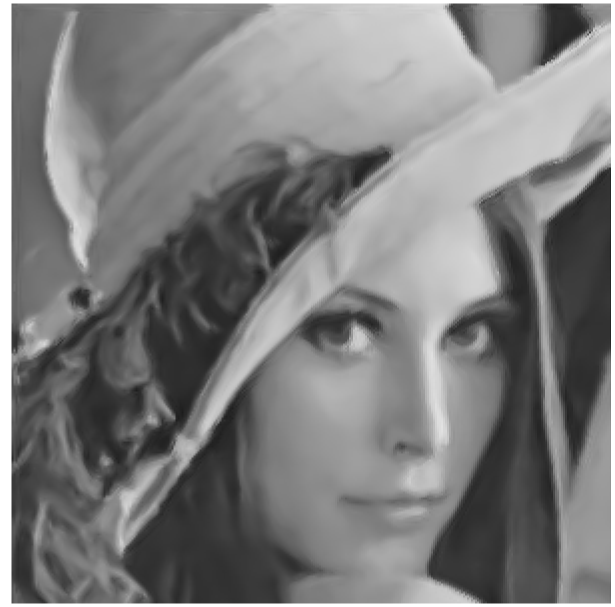
30% of Pixels Retained

Local Constant Kernel Estimate





Adaptive Kernel Estimate



Milanfar et al. EE Dept, UCSC

Upsampling Example: Interpolation from Regular Samples

Downsampling

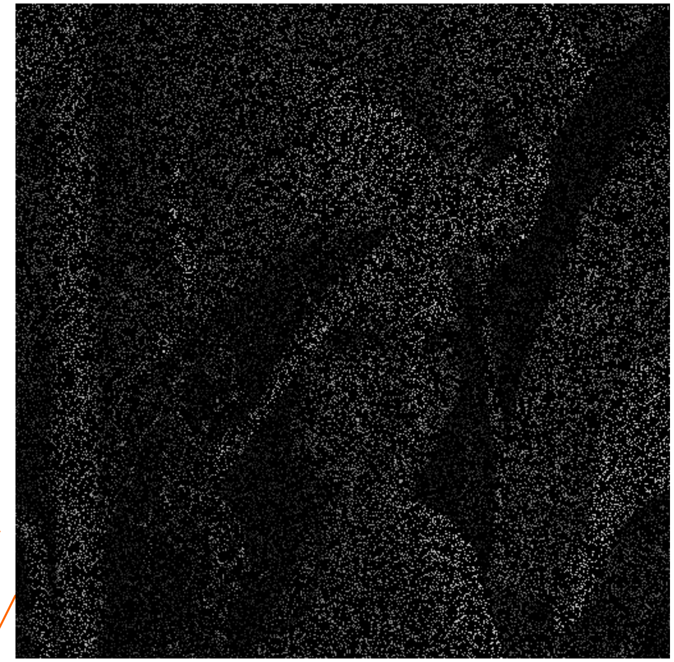
$\downarrow$ 3x

Standard kernel,
order = 2
RMSE = 8.32

Steerable kernel,
order = 2
RMSE = 7.59

Interpolation from Irregular Samples

Random
Downsampling
85%



Standard kernel,
order = 2
RMSE = 9.35

Steerable kernel,
order = 2
RMSE = 8.38

SR Example

Low resolution video, 8 frames

Local constant estimator, order=0

Local quadratic estimator, order=2

Estimated scene

Super-resolution Example

Motion Estimation

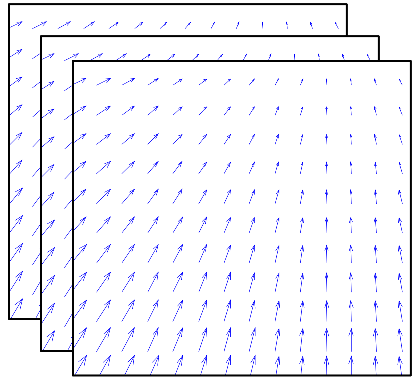


Image Reconstruction

Resolution enhancement from video frames captured by a commercial webcam
(3COM Model No.3719)

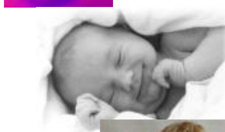
Some Final Remarks

- SR is an idea whose time has come.
- Time to seriously consider applications.
- **A prediction**: in 5-7 years, SR will be used routinely in consumer products.

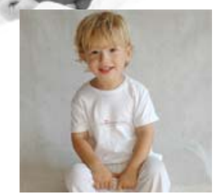
The Life of Super-Resolution so-far

- A pregrant idea: Super-Res is conceived
  - Yen (‘56) and Papoulis (’77) Sampling Theorems

- Birth: A first super-resolution algorithm
  - Tsai and Huang (’84)



- Toddler: Back-projection methods
  - Peleg, Keren, Schweitzer (’87), Peleg and Irani (‘90)



- Early Education: Some formal signal processing
  - Bose (’90), Tekalp et al (’92)
- Pre-teen: The facts of life
  - Elad (’95), Katsaggelos (’95), Schultz (’95), Foroosh (’95)

- Teenager: Getting good with numbers, and learning to learn
  - Nguyen,Milanfar,Golub (’98), Baker (’99)

- College: Color, compression, stability, learn to adapt better
  - See Special Issue of EURASIP JASP

- **TODAY: SR has recently graduated from college.**
  - **<u>Time to get a job</u> and become useful.**
  - **Or go to graduate school….**