From SD to HD
Improving Video Sequences Through Super-Resolution

Michael Elad

Computer-Science Department
The Technion - Israel

final Workshop - November 10, 2011

In this Talk …

we shall describe a fascinating technology called

Super-Resolution

which suggests ways to take (very) poor quality images and fuse them to high-quality image

Few Interesting Facts:

- "יש מאין"? Not here!
- This field is ~25 years old
- Israeli scientists play a major role in it
- Despite its appeal, there were no industrial applications
- … until recently !

Agenda

1. **Introduction**

   Basic Concepts: Image, Pixel, Resolution

2. **Classic Super-Resolution**

   How does it Work?

   Problems and (Many) Limitations

3. **New Era of Super-Resolution**

   Changing the Basic Concept

   New Results

4. **Summary**

Agenda

1. **Introduction**

   Basic Concepts: Image, Pixel, Resolution

2. **Classic Super-Resolution**

   How does it Work?

   Problems and (Many) Limitations

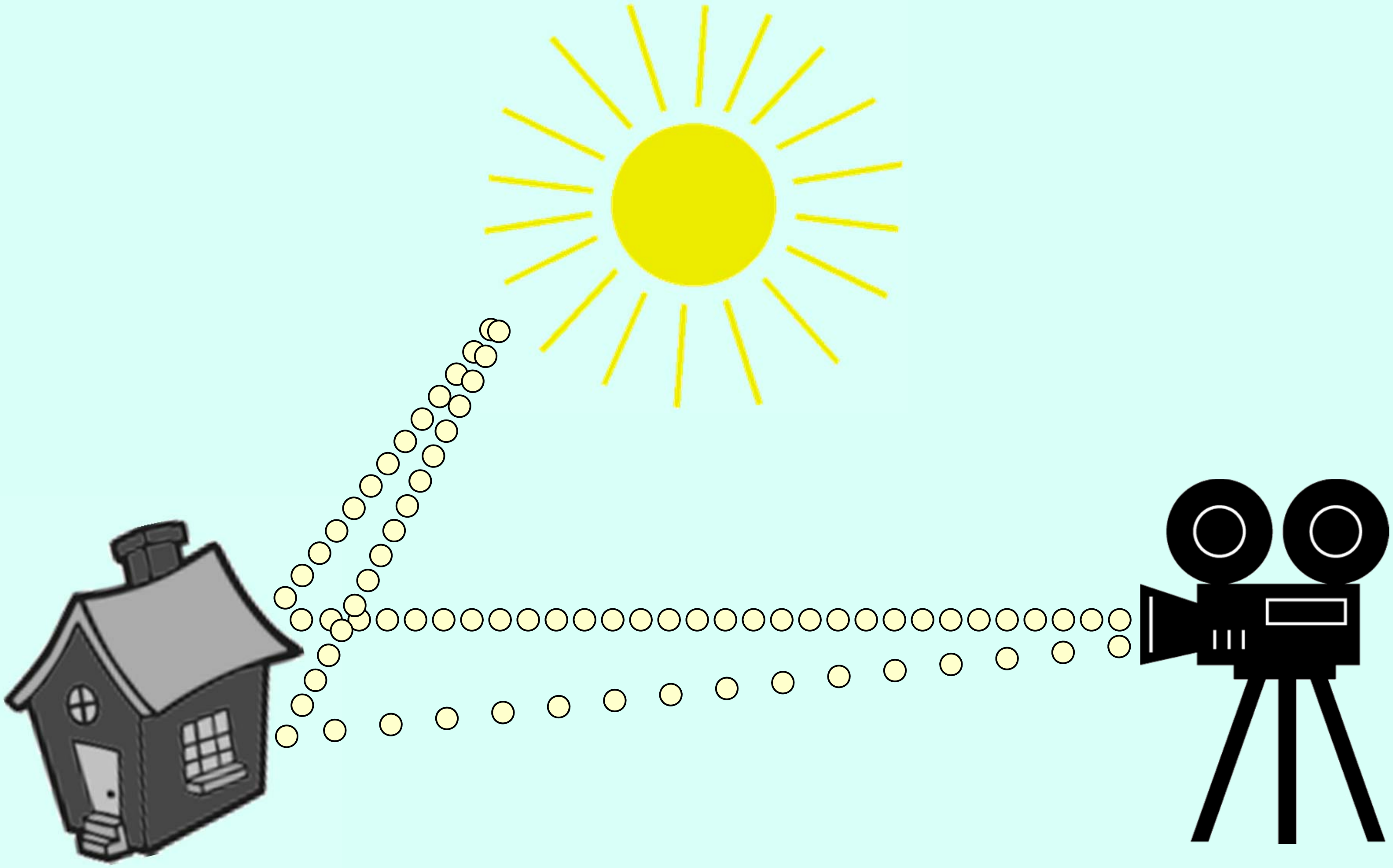
3. **New Era of Super-Resolution**

   Changing the Basic Concept

   New Results

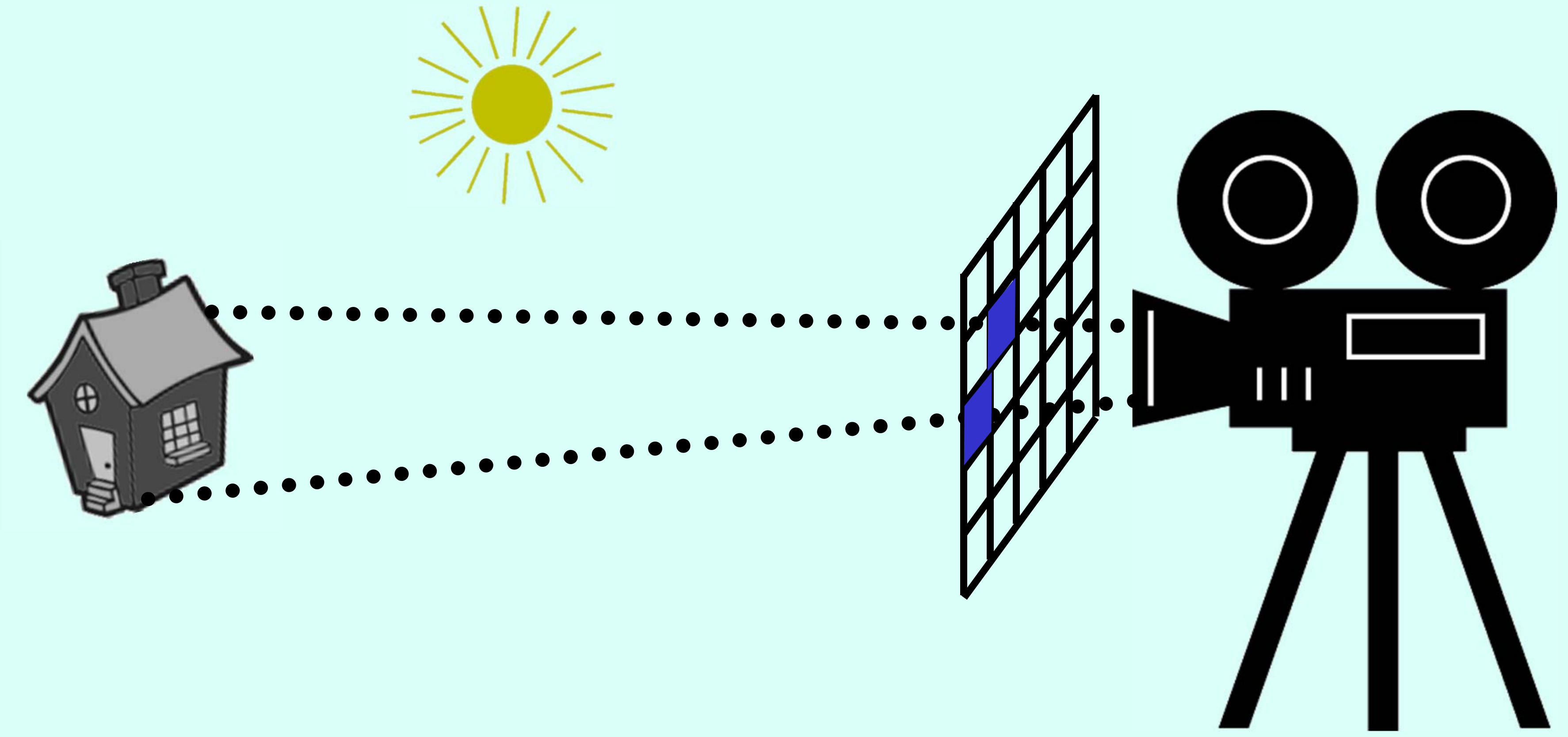
4. **Summary**

What is an Image?



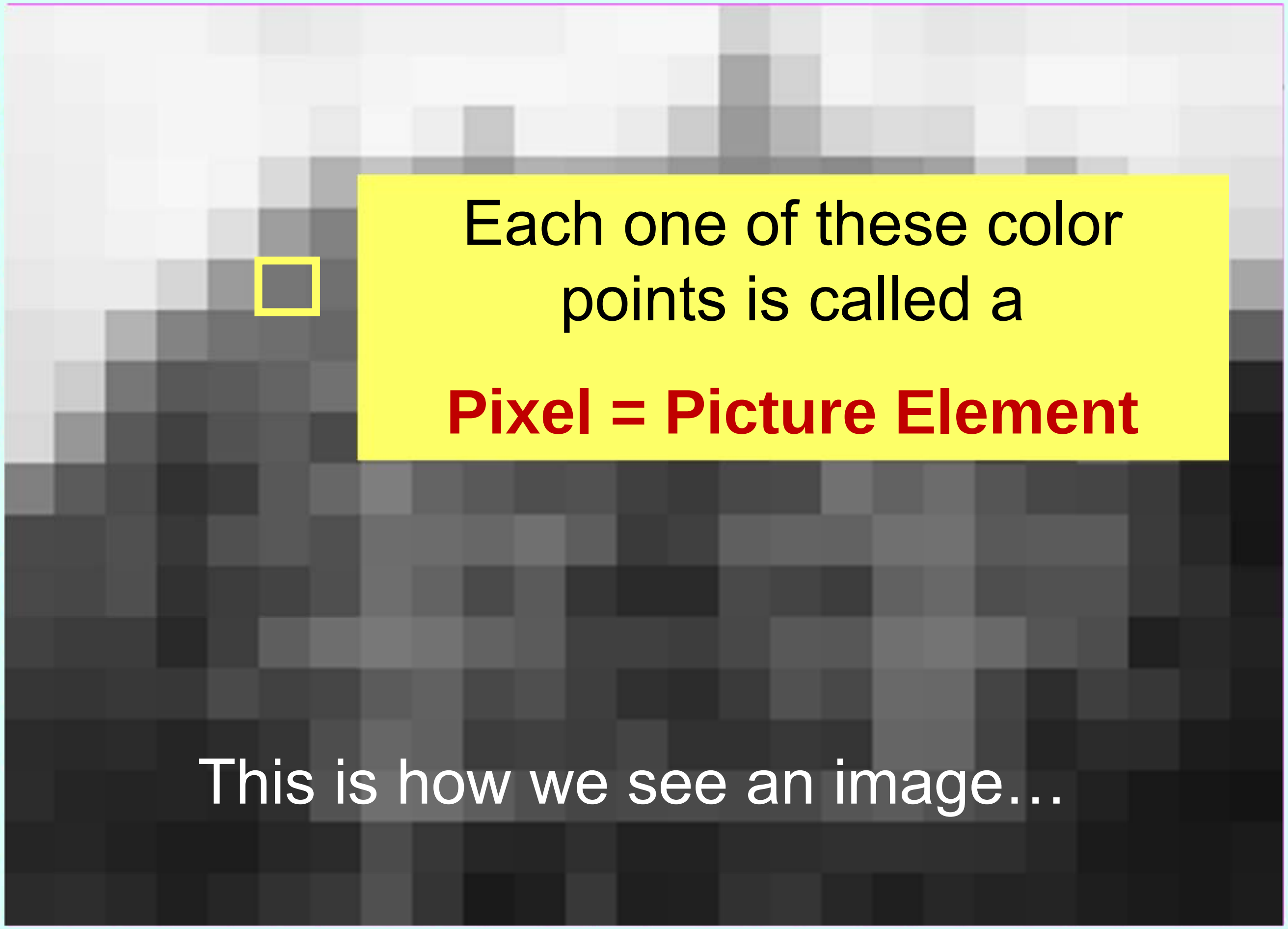
What is an Image?

The camera counts the number of photons…



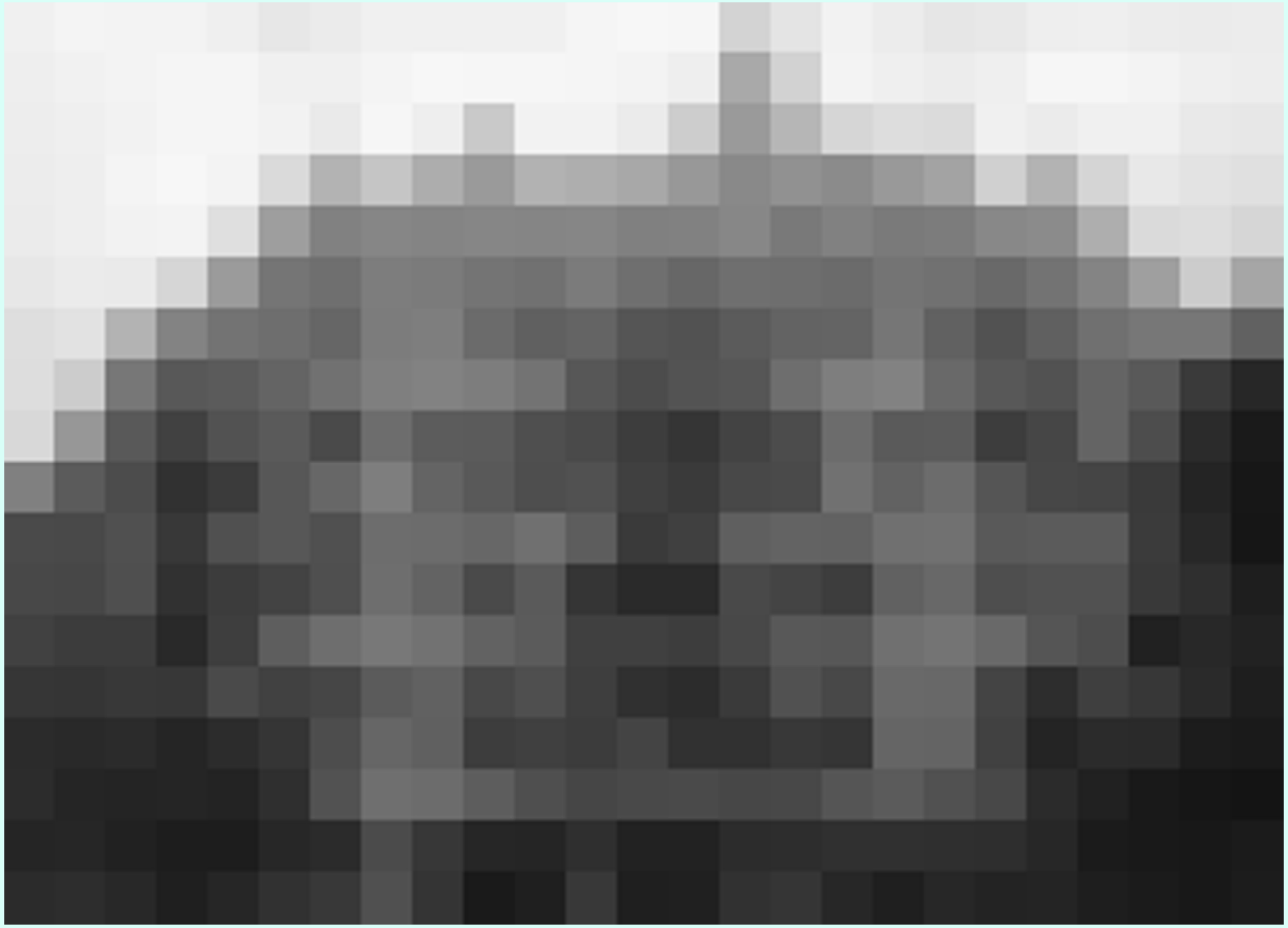
What is an Image?

To the computer, an image is nothing but a table of numbers



Resolution

This image is 18 by 25 pixels



Resolution

This image is 36 by 50 pixels

Resolution

This image is 72 by 100 pixels

Resolution

This image is 144 by 200 pixels

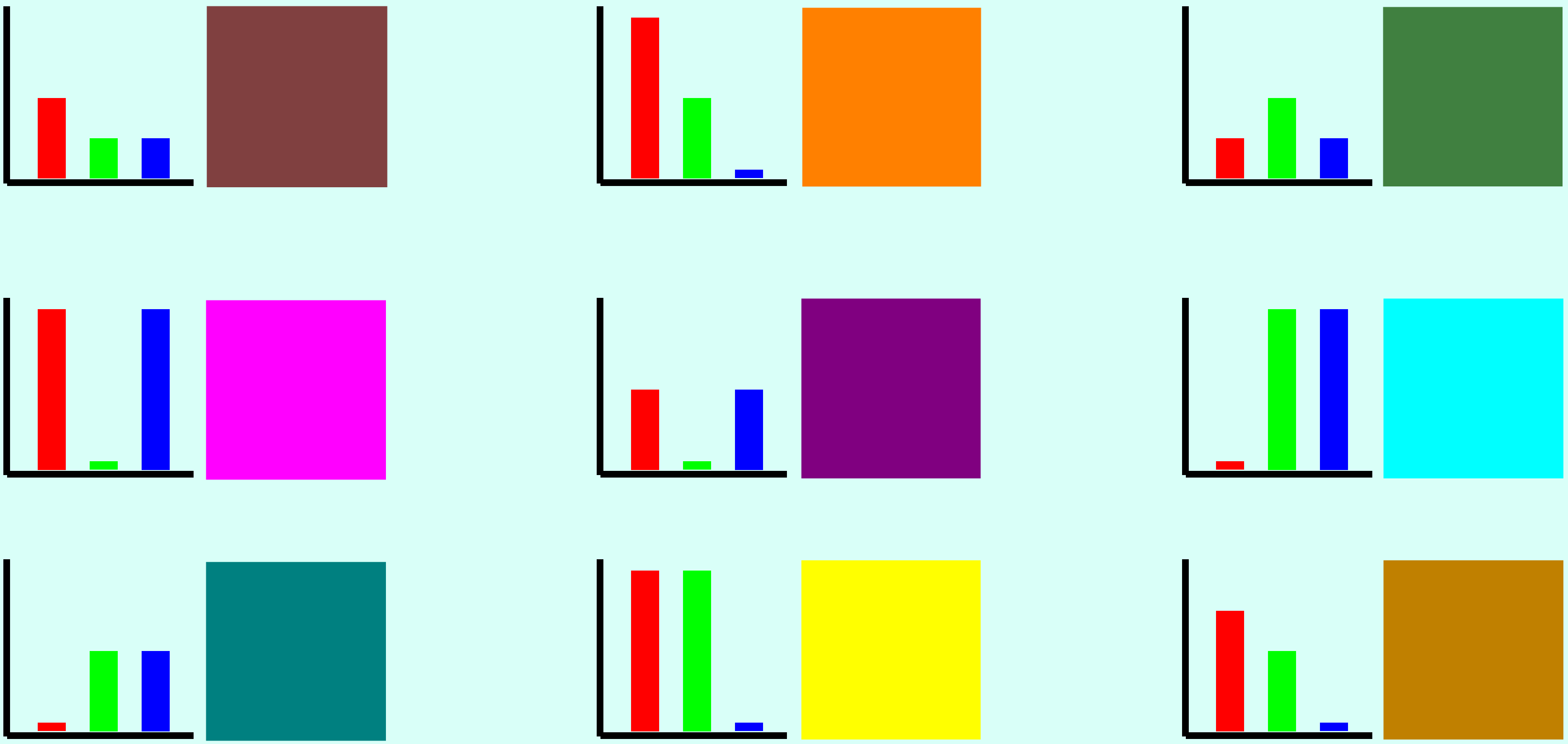
Resolution

This image is 288 by 400 pixels

Resolution: the size of the smallest detail that can be seen in the image

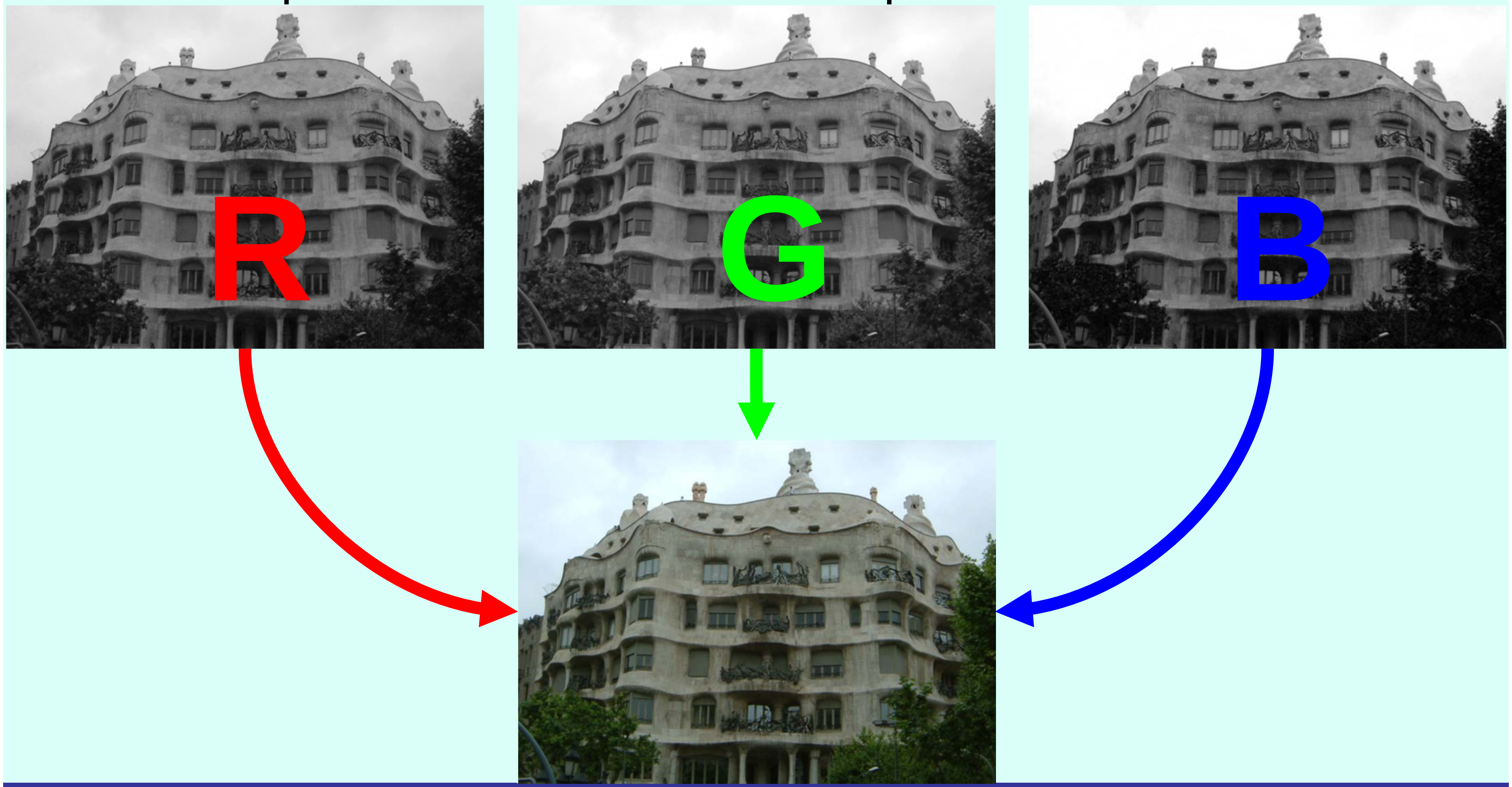
Color

Colors can be created by a mixture of Red , Green and Blue



Color Images

Each pixel counts the number of photons in each color



Back to Resolution

Screens are composed of pixels as well

What happens when the image has a

different number of pixels from the screen?

Resolution Disparities

Case 1 (easy!): Image has more pixels than screen

Option 1: Crop

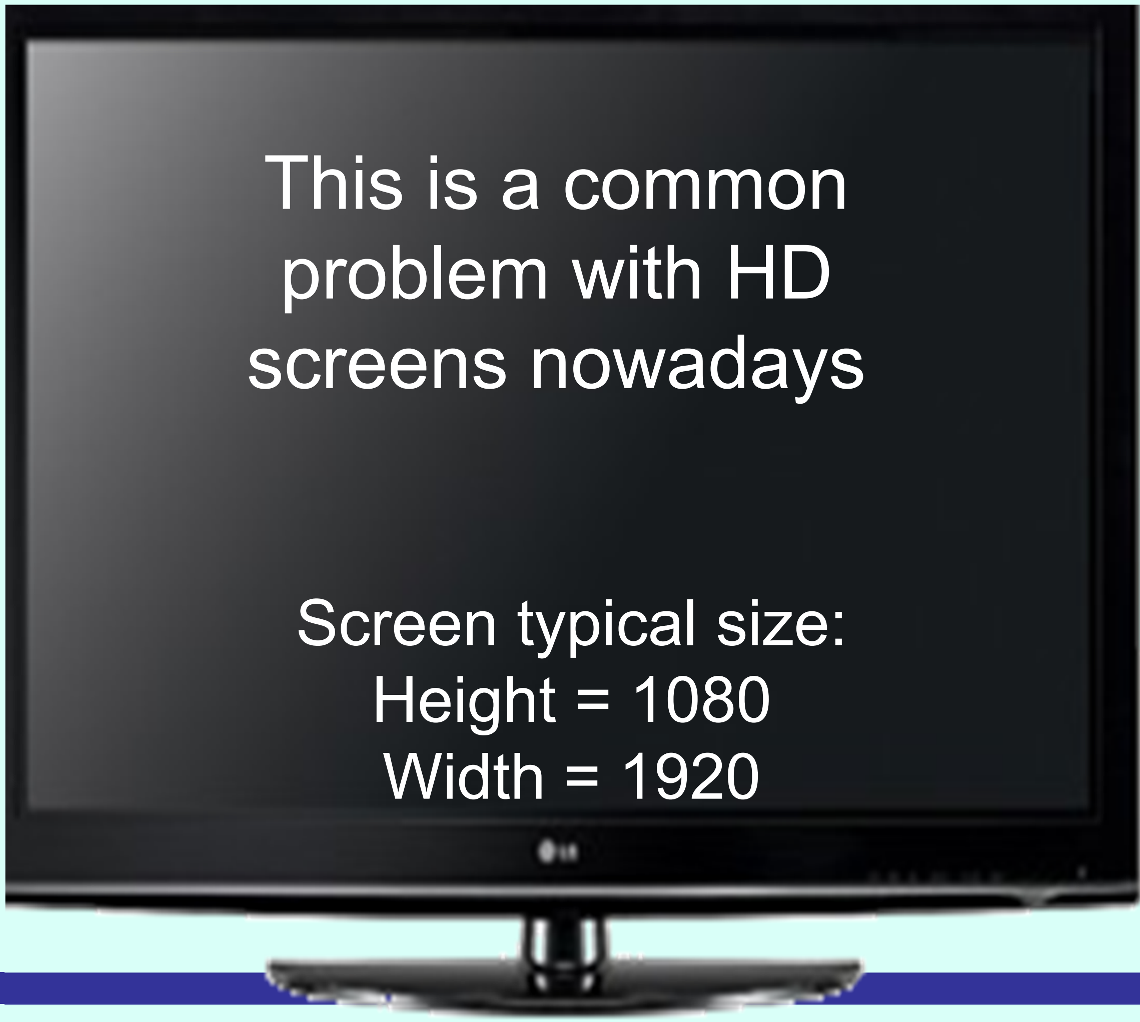
Resolution Disparities

Case 1 (easy!): Image has more pixels than screen

Option 2: Shrink

Resolution Disparities

Case 2 (problem!): screen has more pixels than image



Video typical size:
Height = 480
Width = 640

Why Isn't Everything HD in the First Place?

- Various reasons:
  - Old material: Sequences shot before HD cameras (HD came to the market ~3 years ago …)
  - Cost: HD video technology is too expensive (but this is fast-changing):
    - HD camera and equiment is more expensive
    - HD is more expensive to broadcast: More pixels $\rightarrow$ "wider pipes"
    - HD requires more storage space
- Bottom Line: much of the available video material is of inadequate resolution for display on HD screens.
- Comment: The above problem is relevant to other needs (security, military, medical, entertainment, …)

Resolution Disparities

Case 2 (problem!): screen has more pixels than image

Option 1: Put image in the middle of the screen

Resolution Disparities

Case 2 (problem!): screen has more pixels than image

Option 2 : Interpolation
- Guess pixel values between known pixels

Interpolation Methods

Magnify 4:1

- Interpolation possibilities:
  - Nearest neighbor
  - Bilinear interpolation
  - Bicubic interpolation

  Spatially invariant meth

  - There are also edge-adaptive techniques that give more sharpness
- None of the above lead to a true increase in the optical (true) resolution

Resolution Disparities

Case 2 (problem!) : screen has more pixels than image

Agenda

1. Introduction

   Basic Concepts: Image, Pixel, Resolution

2. **Classic Super-Resolution**

   How does it Work?

   Problems and (Many) Limitations

3. New Era of Super-Resolution

   Changing the Basic Concept

   New Results

4. Summary

Joint work with

Arie Feuer

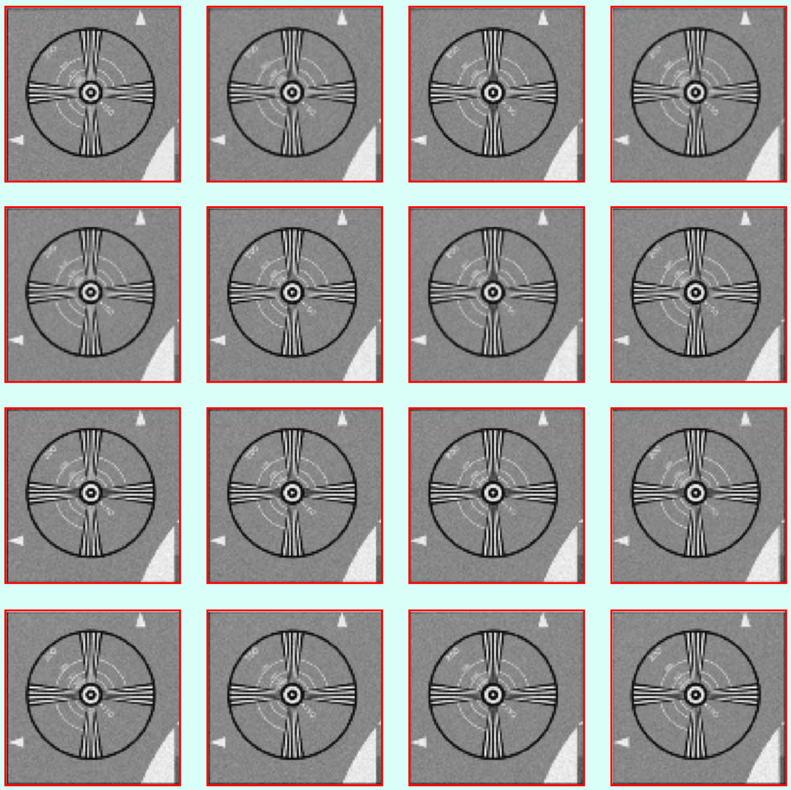
Jacob Hel-Or

Peyman Milanfar

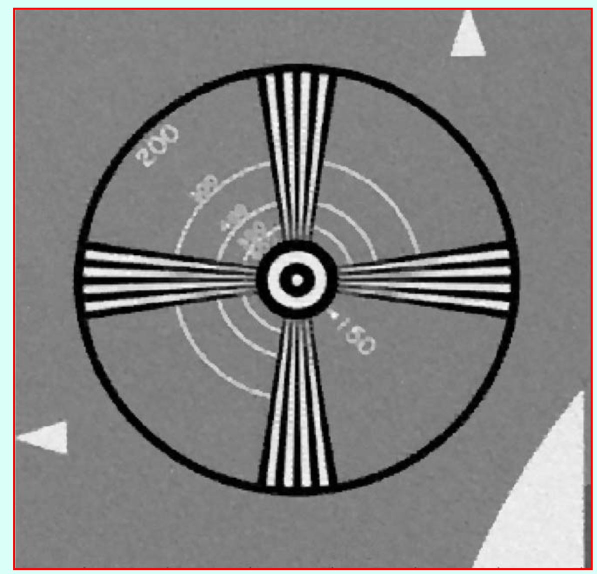
Sina Farsiu

Our Objective

Given: A set of degraded (warped, blurred, decimated, noised) images:

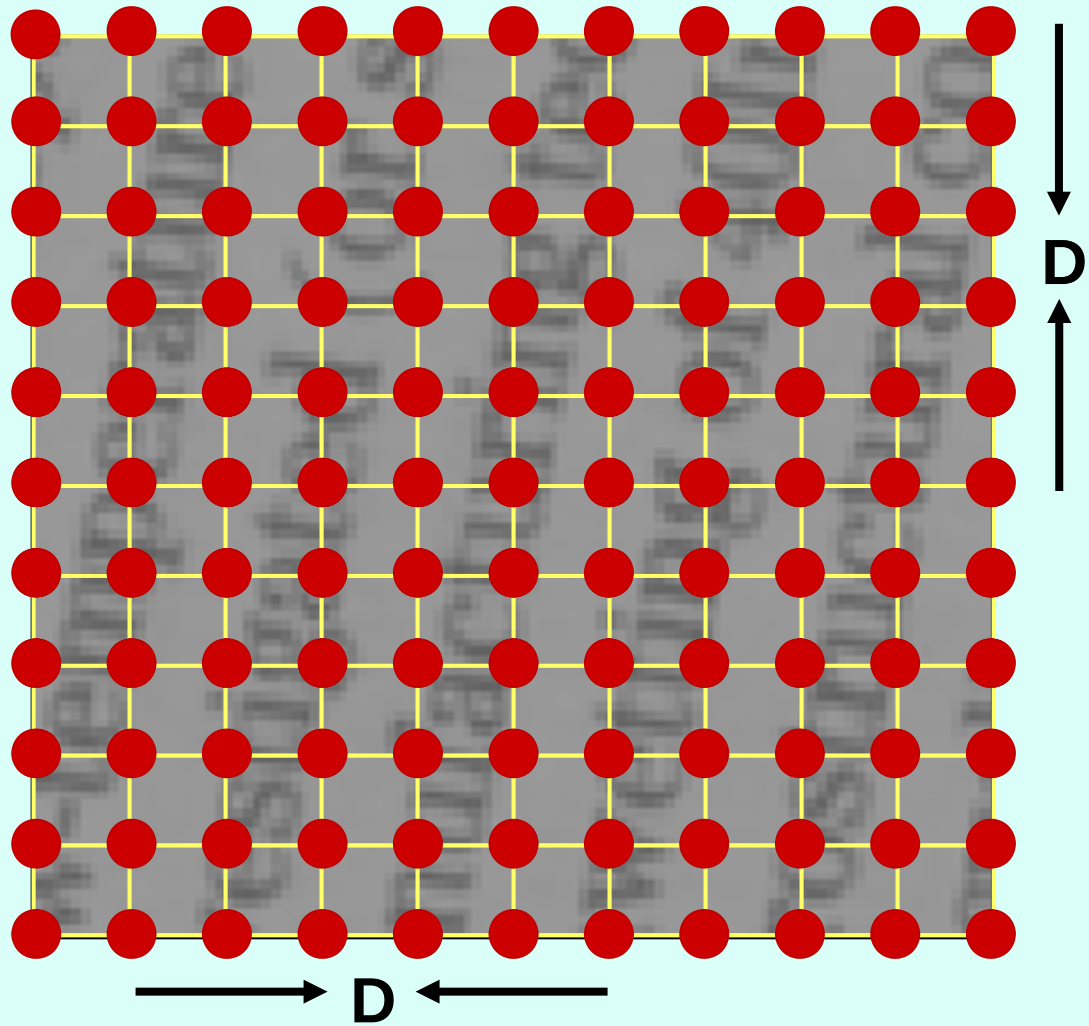


Required: Fusion of the given images into a higher resolution image/s



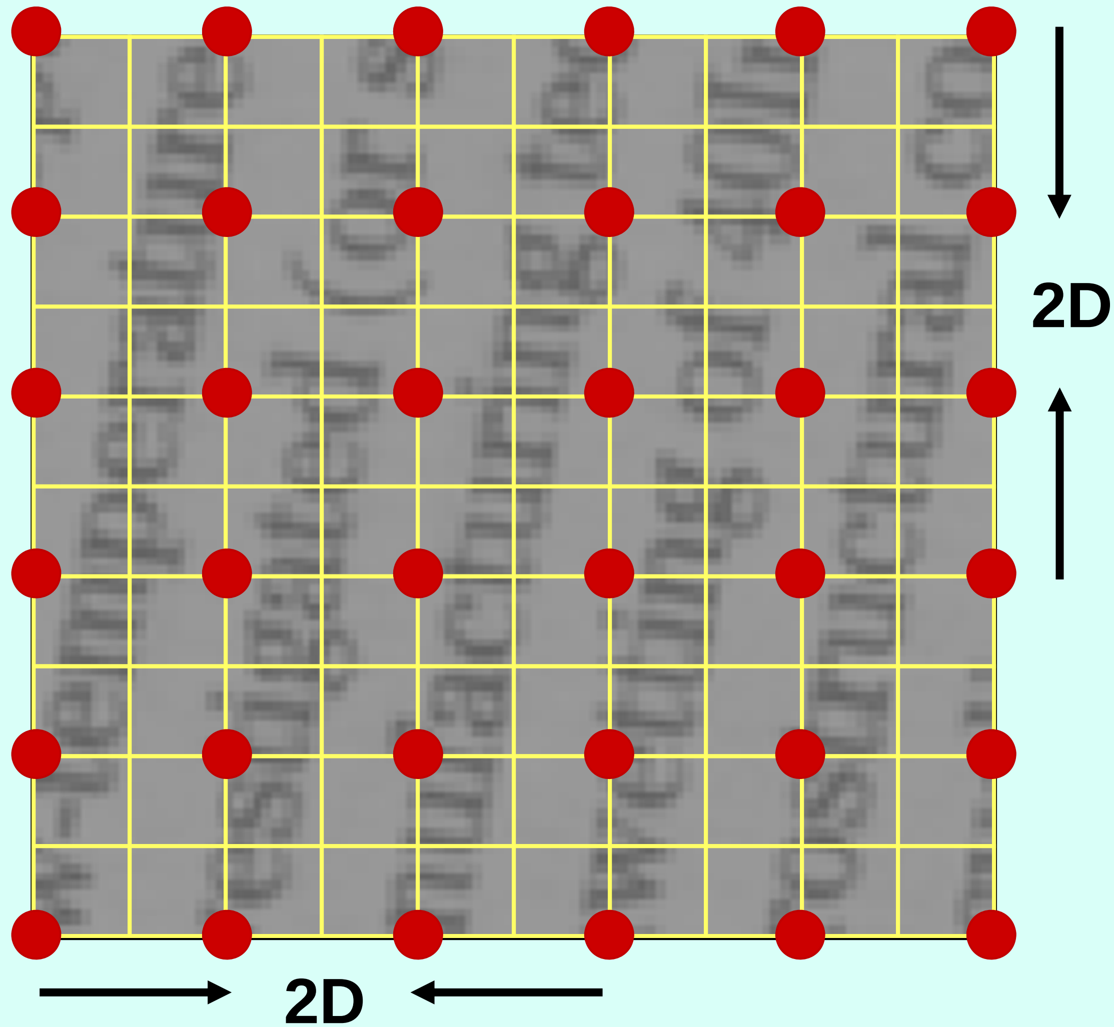
Intuition

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



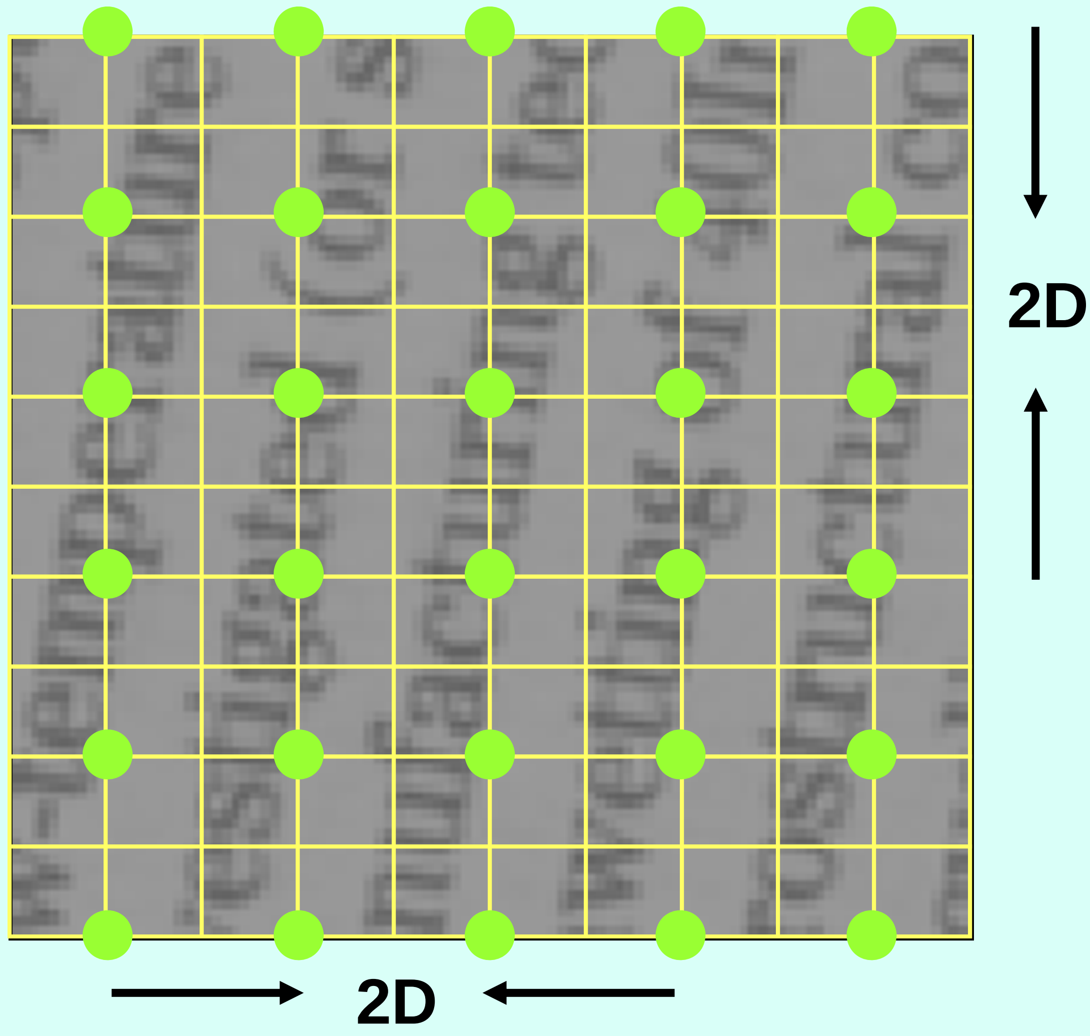
Intuition

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



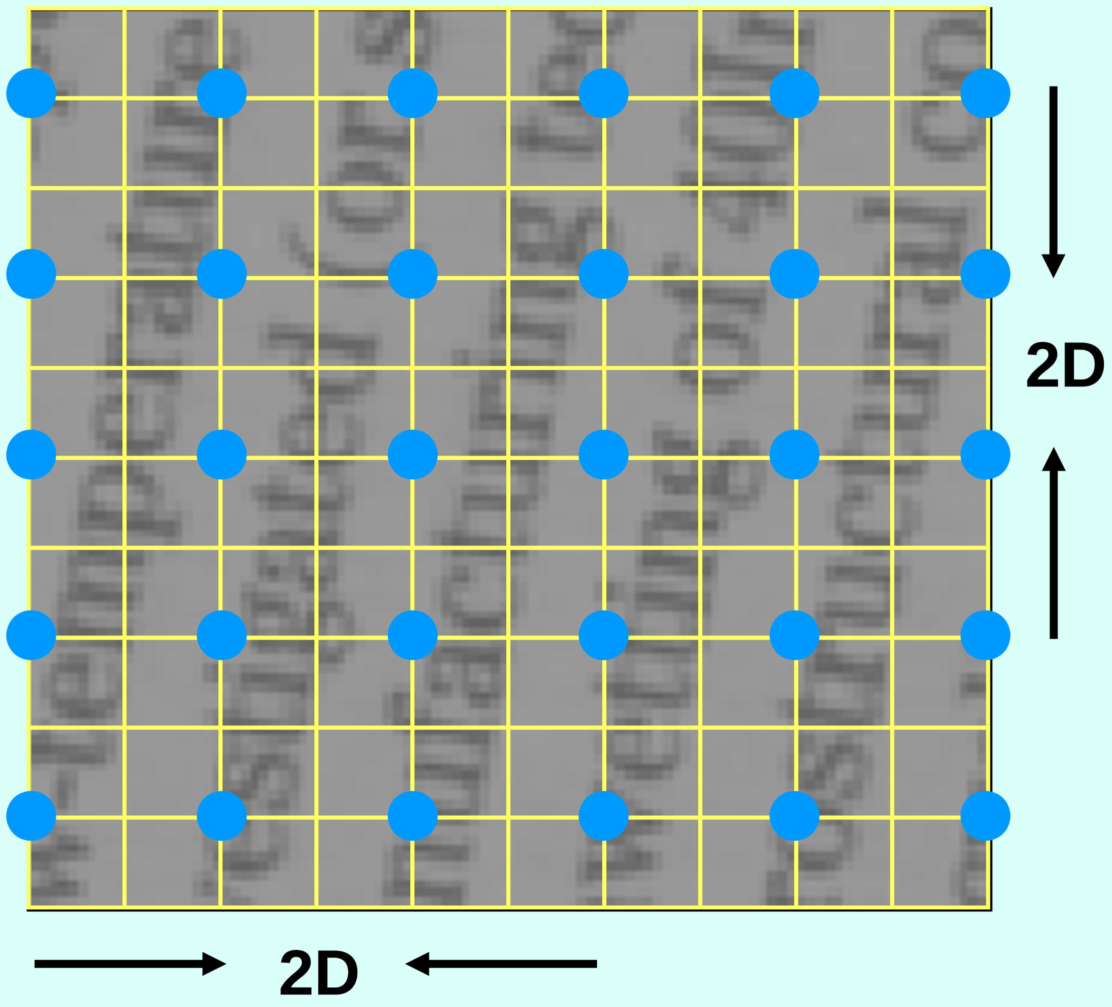
Intuition

However, we are allowed to take a second picture and so, shifting the camera 'slightly to the right' we obtain



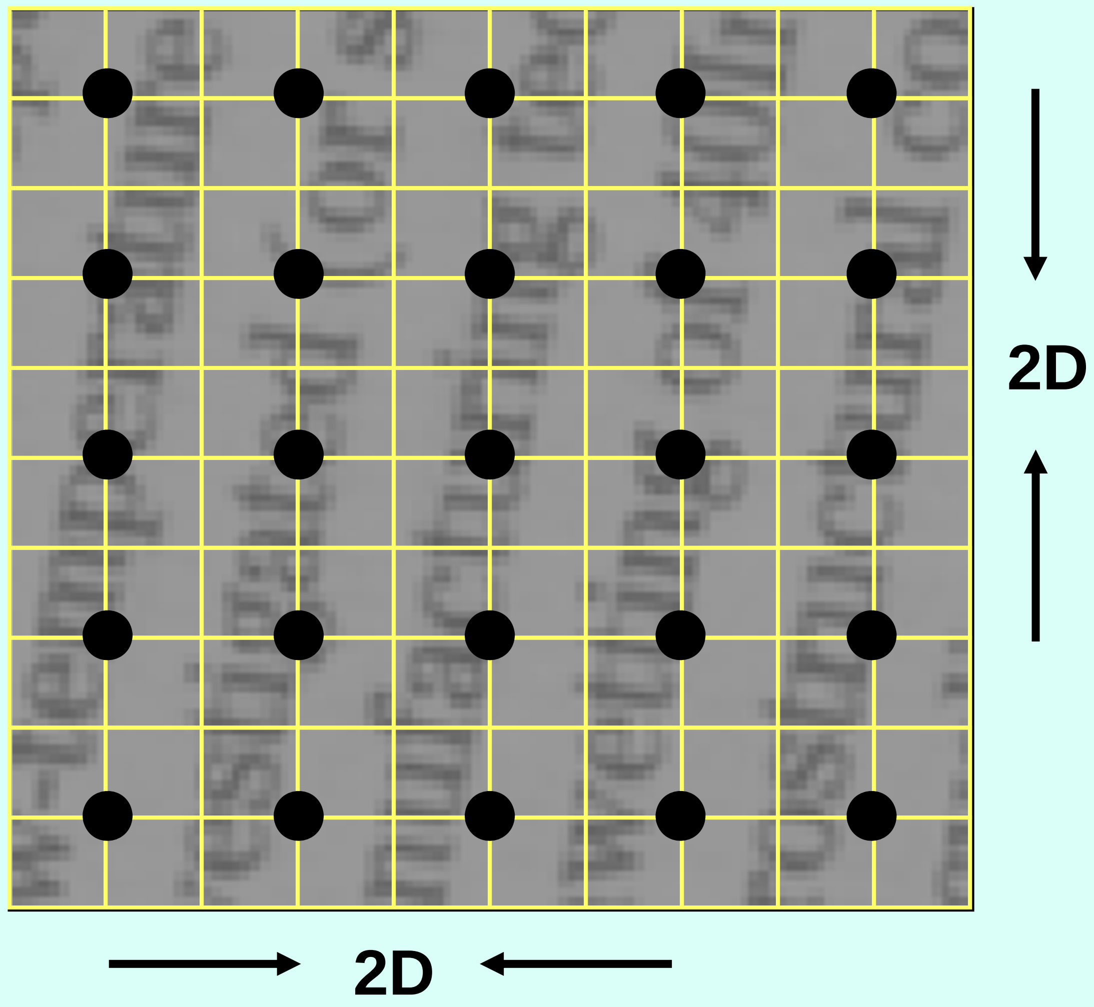
Intuition

Similarly, by shifting down we get a third image



Intuition

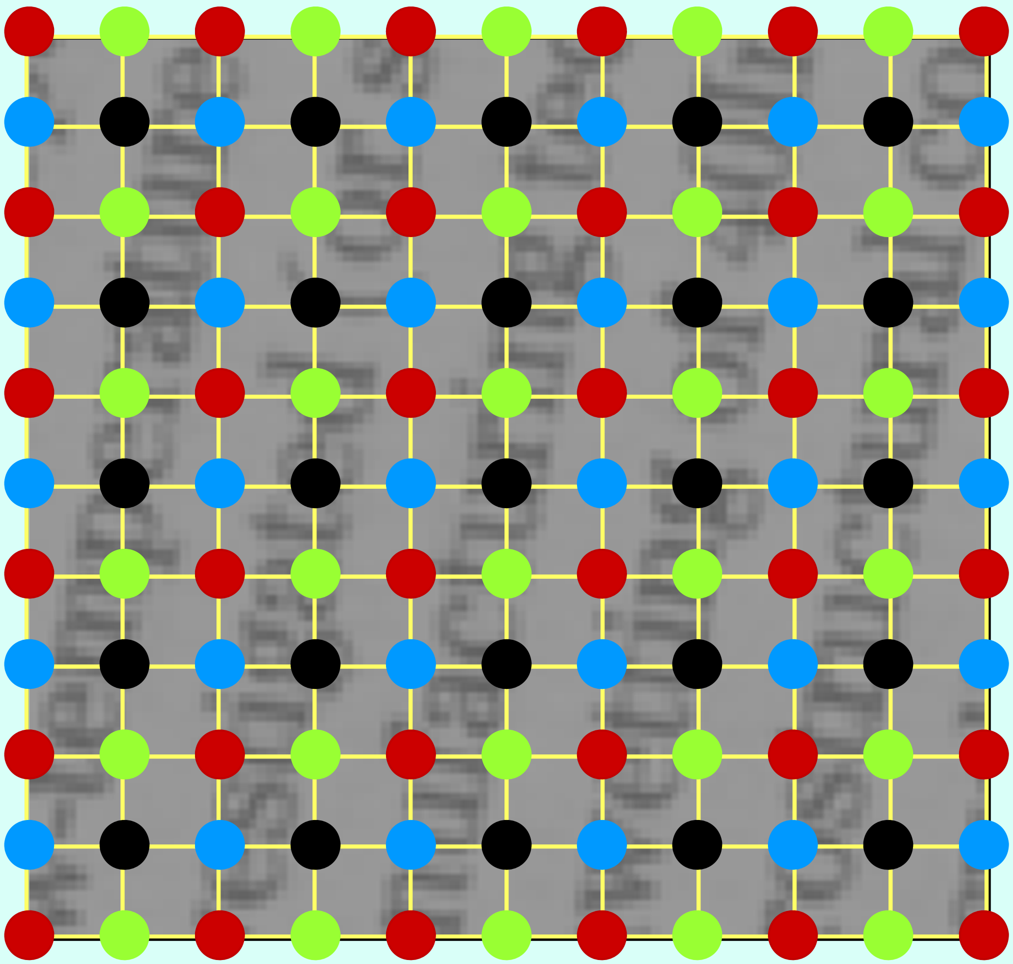
And finally, by shifting down and to the right we get the fourth image



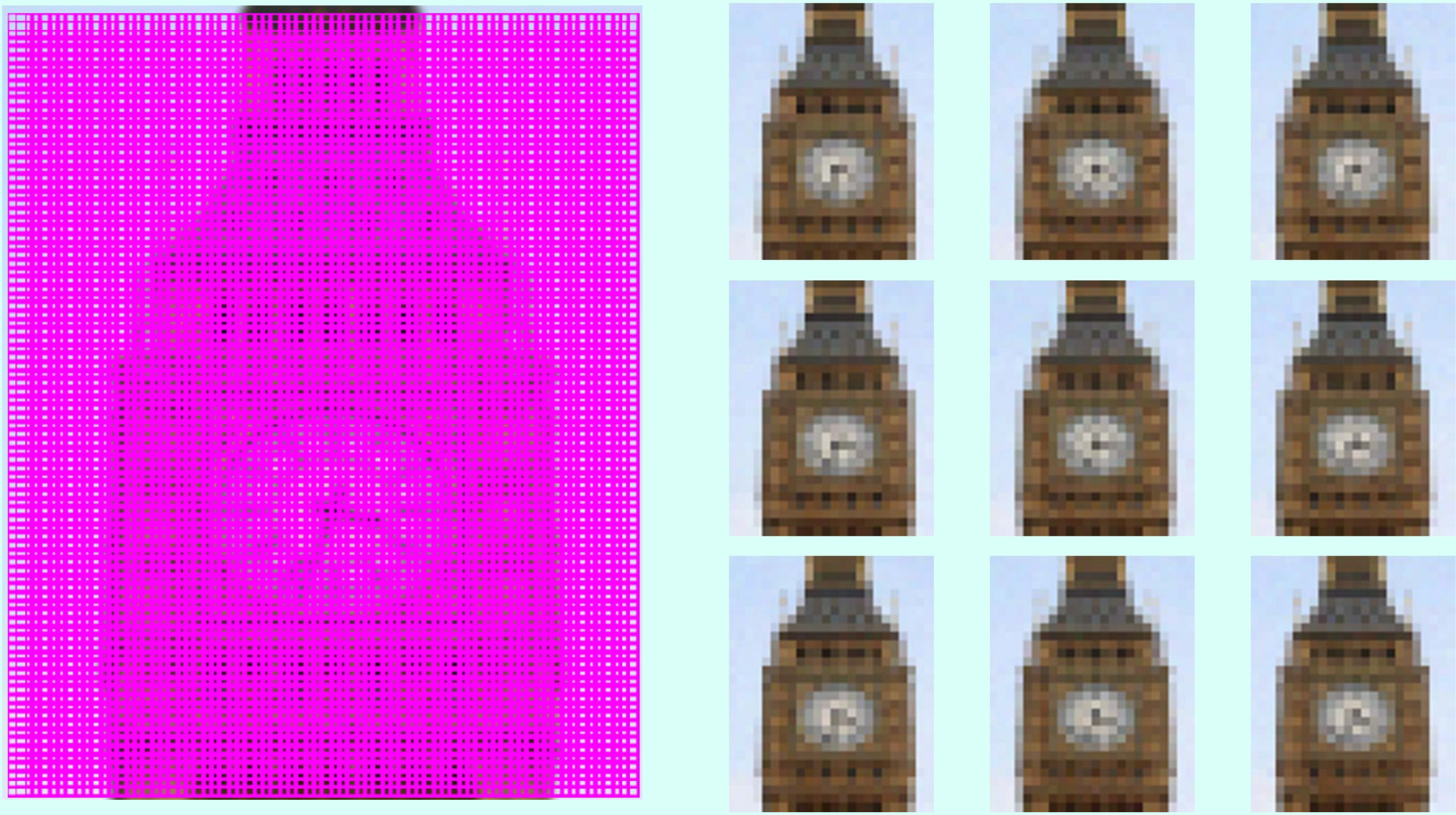
Intuition

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

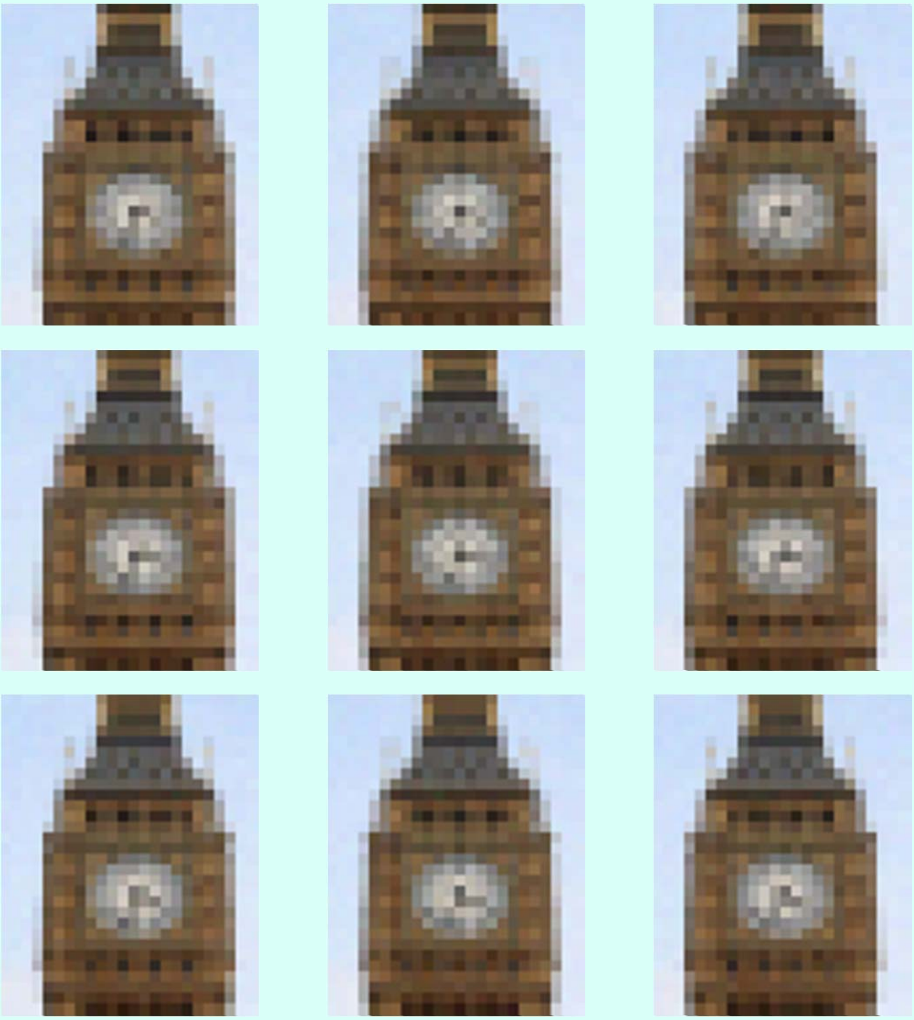
This is Super-Resolution in its simplest form



Super-Resolution: The Simplest Example



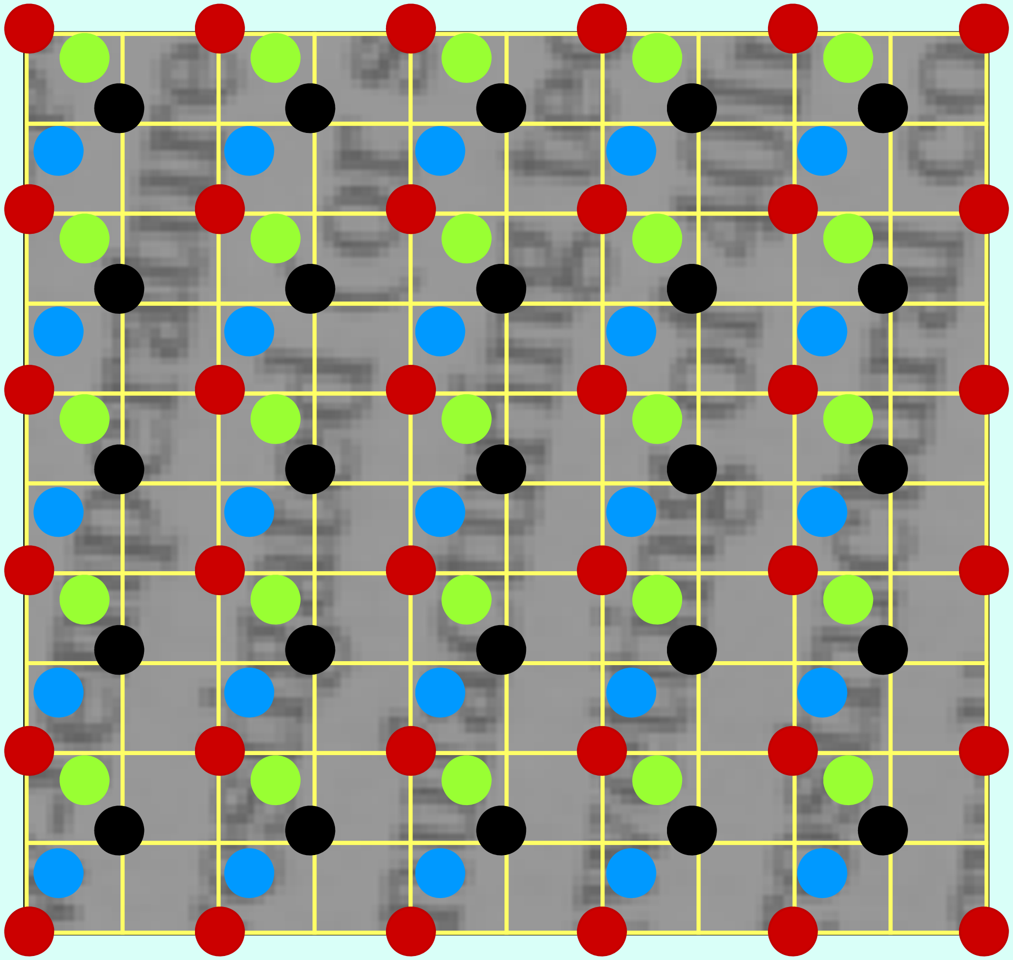
Super-Resolution: The Simplest Example



Intuition – Complicating the Story

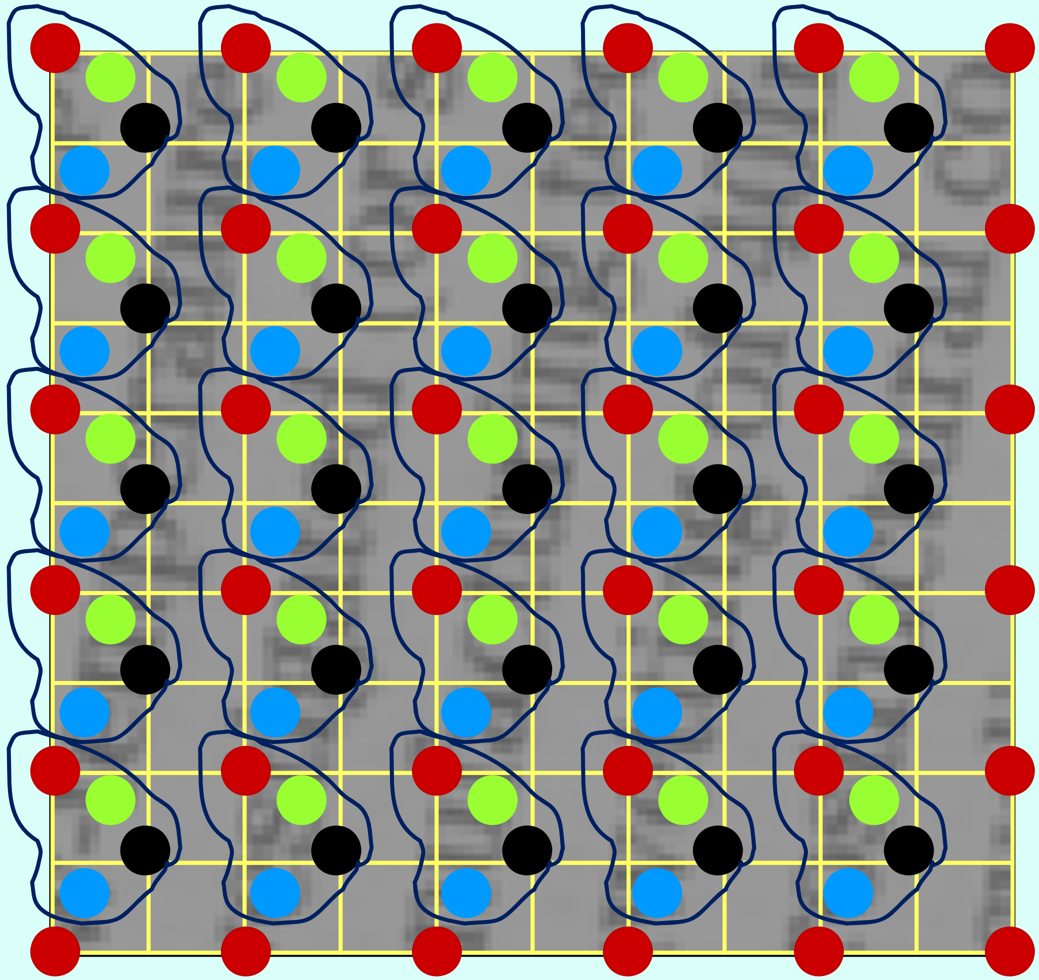
In the previous example we counted on exact movement of the camera by D in each direction

What if the camera displacement is uncontrolled?



Intuition – Complicating the Story

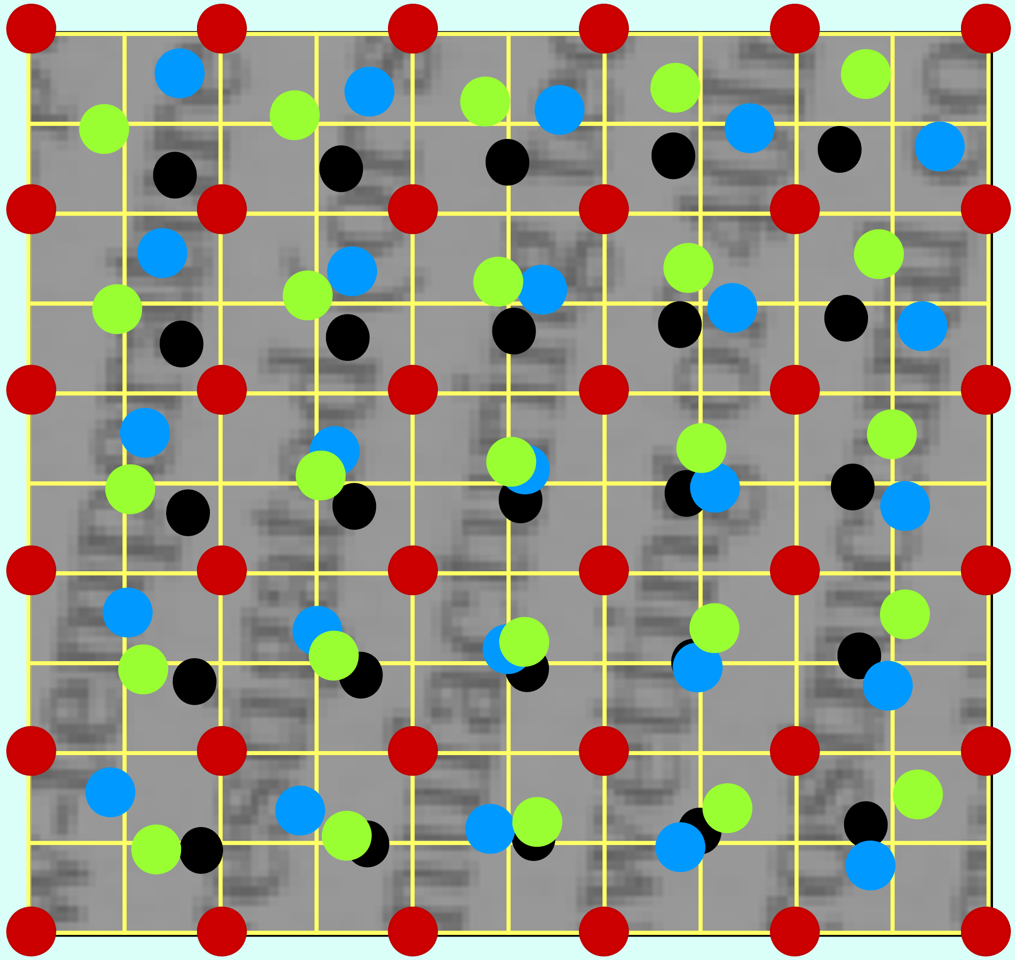
It turns out that there is a sampling theorem due to Yen (1956) and Papulis (1977) for this case, guaranteeing perfect reconstruction for periodic uniform sampling if the sampling density is high enough



Intuition – Complicating the Story

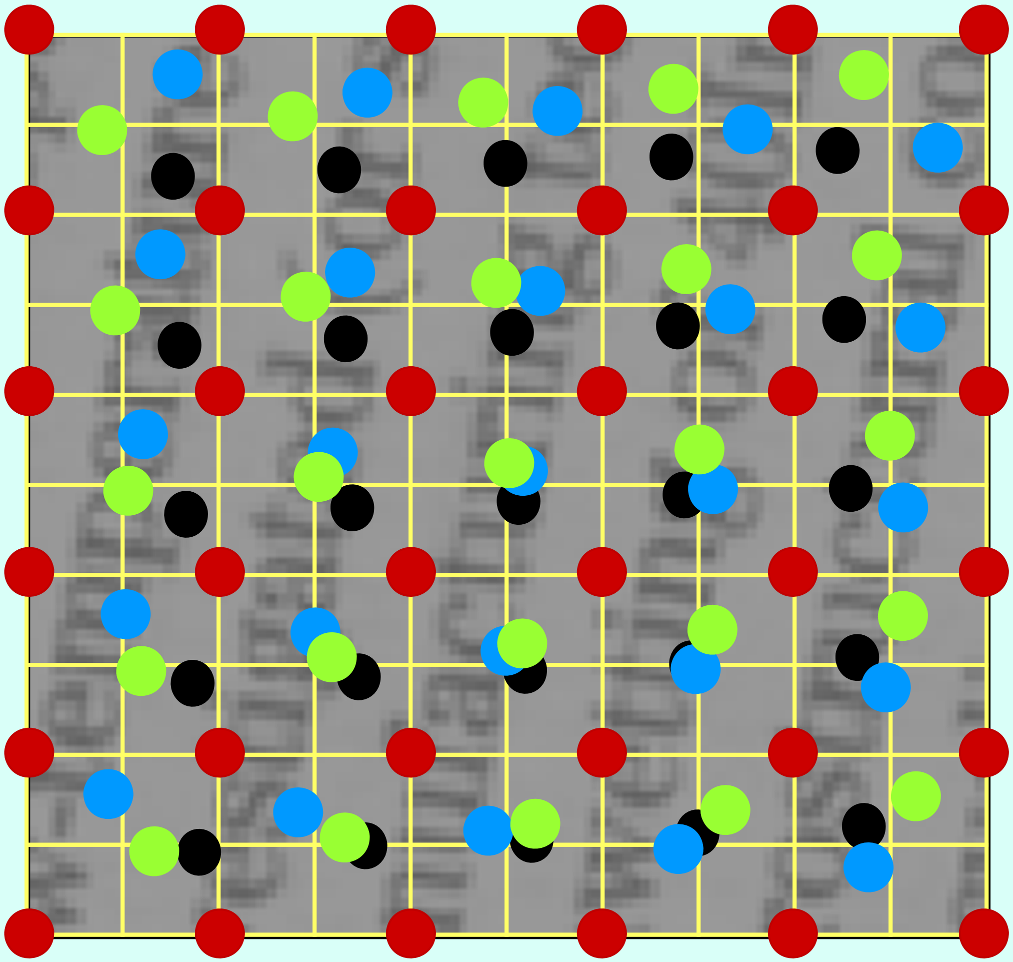
In the previous examples we restricted the camera to move horizontally/vertically parallel to the photograph object.

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



Intuition – Complicating the Story

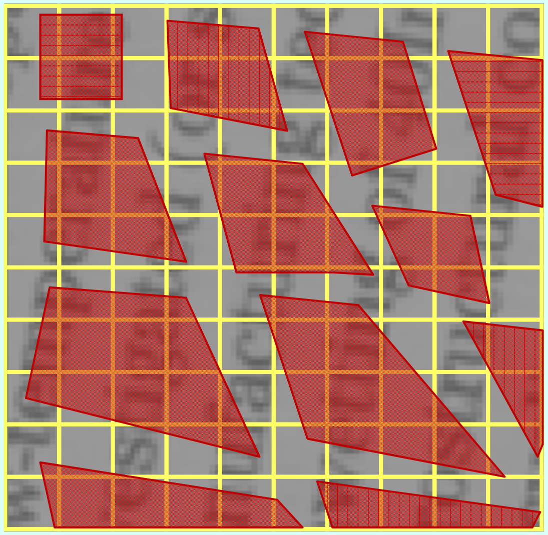
There is no sampling theorem covering this case !



The Problem is Actually More Difficult

Problems:

1. Sampling is not a point operation – there is a blur
2. Motion may include perspective warp, local motion (moving objects), etc.
3. Samples may be noisy – any reconstruction process must take that into account

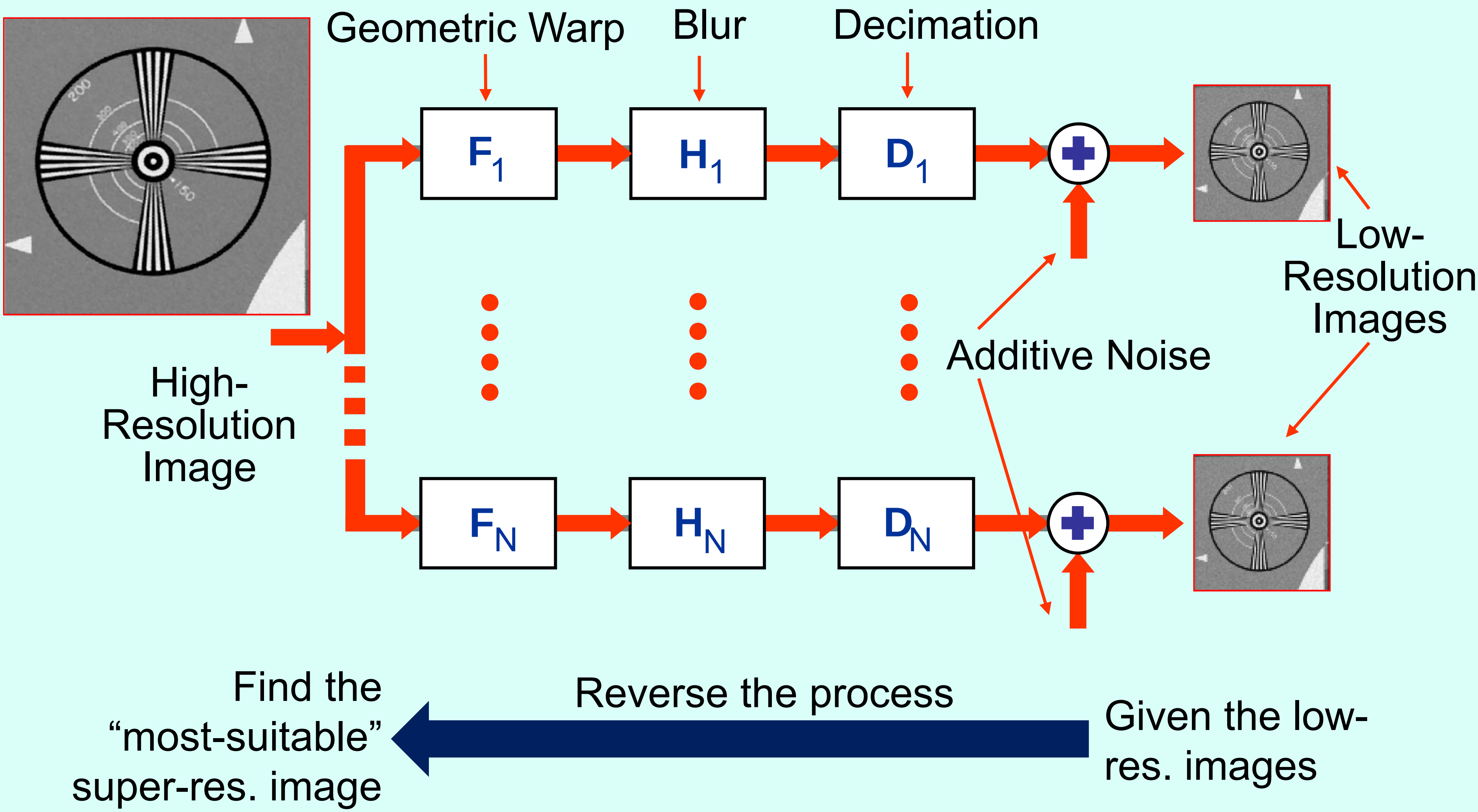


Solution:

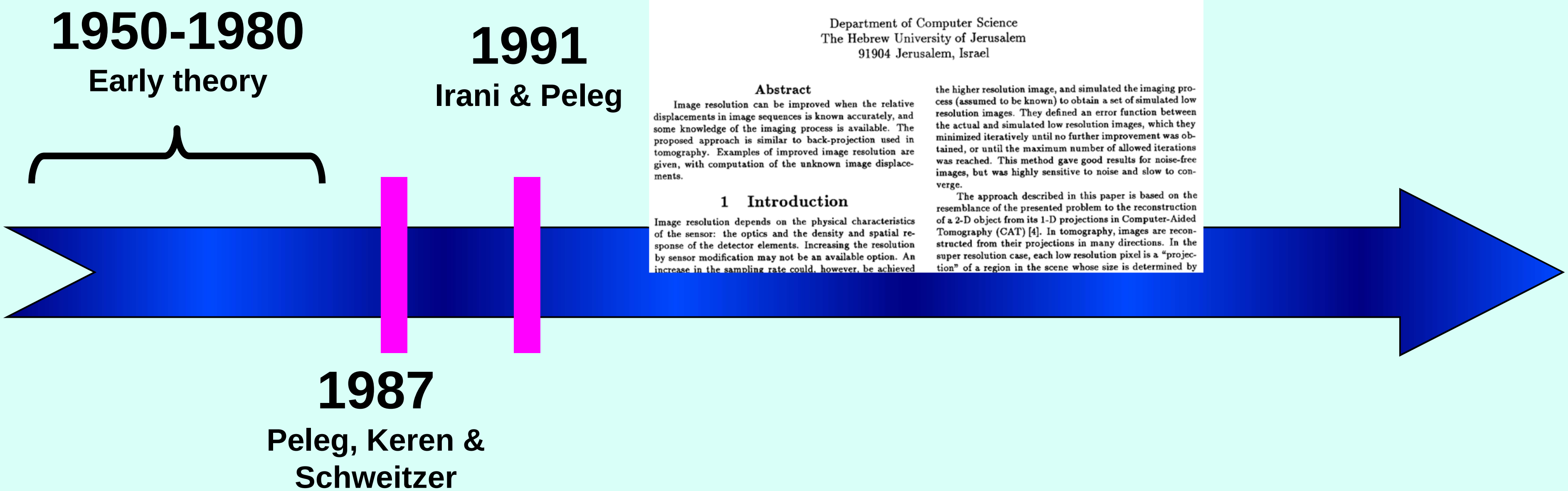
1. Change the treatment: Forget about the "sampling" interpretation, and replace it with an inverse problem point of view
2. A consequence: We need to estimate the motion between the different images in a very accurate (sub-pixel accuracy) way

Inverse Problem?

Elad & Feuer (1997)



Evolution of Super-Resolution



Super Resolution From Image Sequences

Michal Irani and Shmuel Peleg*

Department of Computer Science
The Hebrew University of Jerusalem
91904 Jerusalem, Israel

Abstract

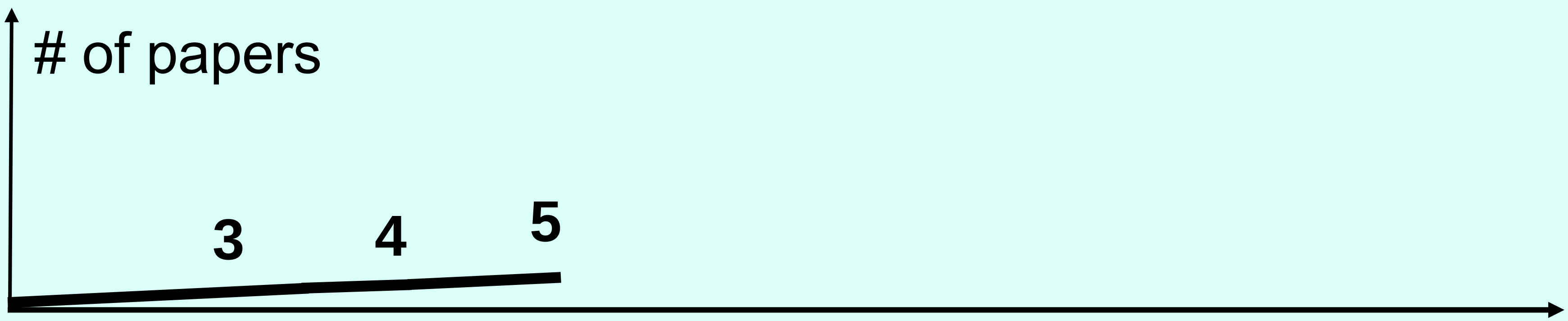
Image resolution can be improved when the relative displacements in image sequences is known accurately, and some knowledge of the imaging process is available. The proposed approach is similar to back-projection used in tomography. Examples of improved image resolution are given, with computation of the unknown image displacements.

1 Introduction

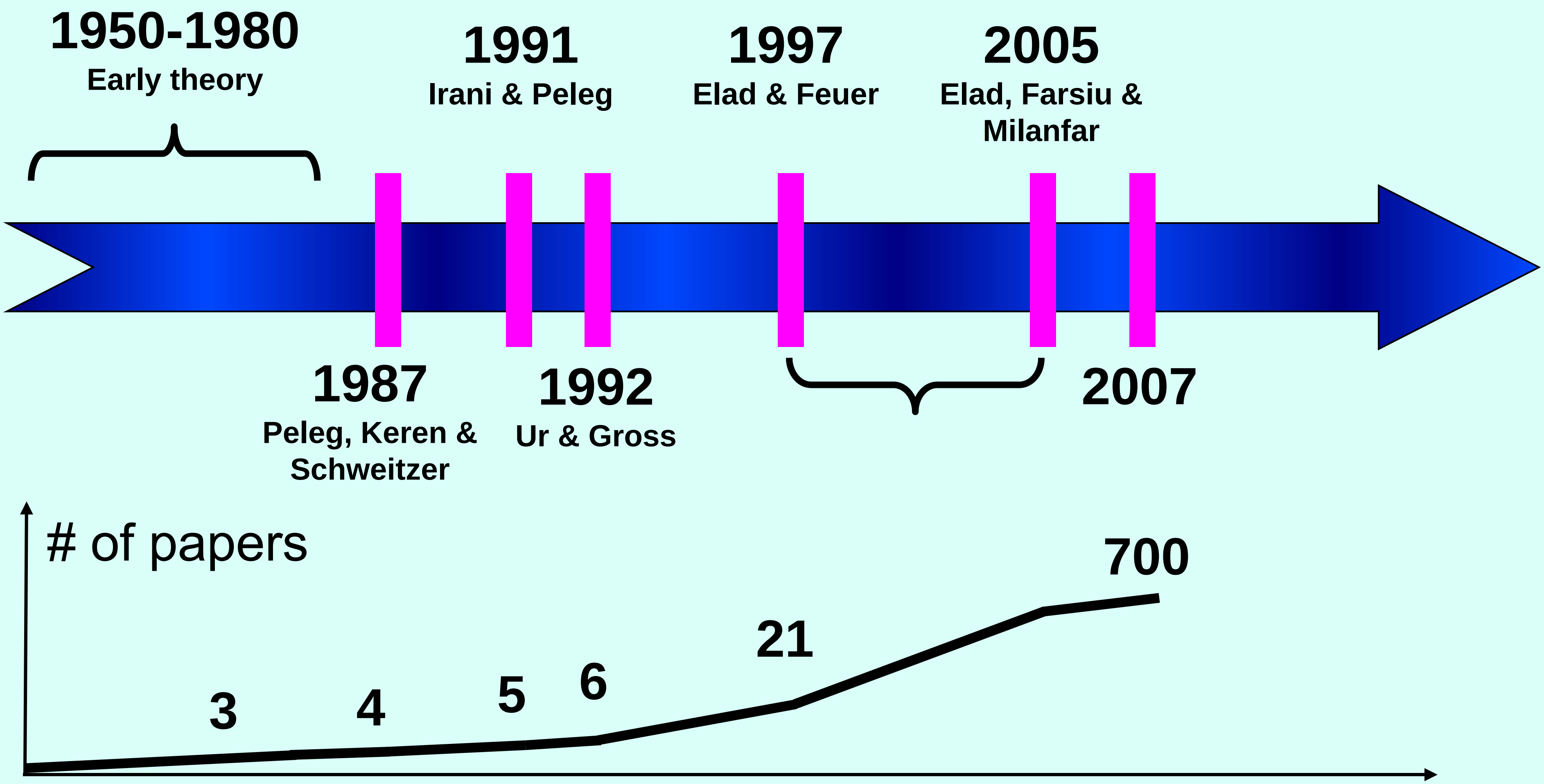
Image resolution depends on the physical characteristics of the sensor: the optics and the density and spatial response of the detector elements. Increasing the resolution by sensor modification may not be an available option. An increase in the sampling rate could, however, be achieved

the higher resolution image, and simulated the imaging process (assumed to be known) to obtain a set of simulated low resolution images. They defined an error function between the actual and simulated low resolution images, which they minimized iteratively until no further improvement was obtained, or until the maximum number of allowed iterations was reached. This method gave good results for noise-free images, but was highly sensitive to noise and slow to converge.

The approach described in this paper is based on the resemblance of the presented problem to the reconstruction of a 2-D object from its 1-D projections in Computer-Aided Tomography (CAT) [4]. In tomography, images are reconstructed from their projections in many directions. In the super resolution case, each low resolution pixel is a "projection" of a region in the scene whose size is determined by

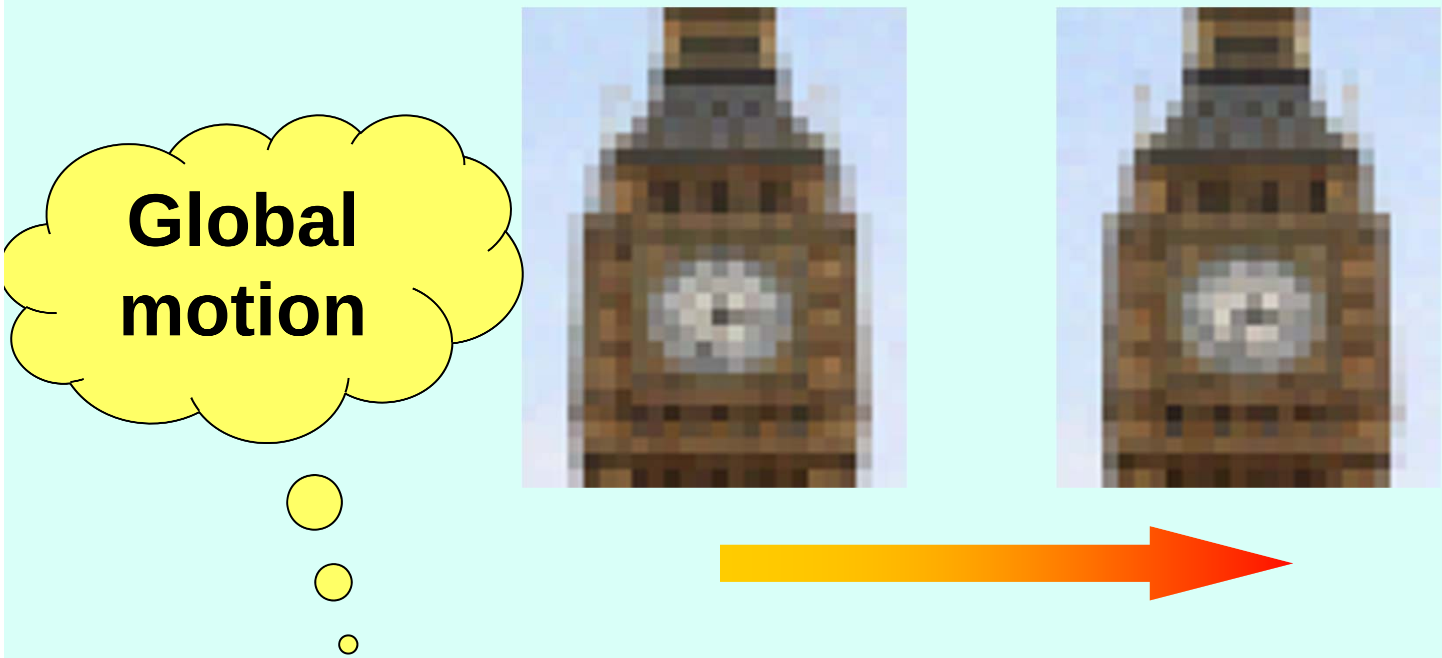


Evolution of Super-Resolution



We Rely on Motion Estimation

If we don't know how the camera moved, we
must figure this out from the low-resolution images themselves



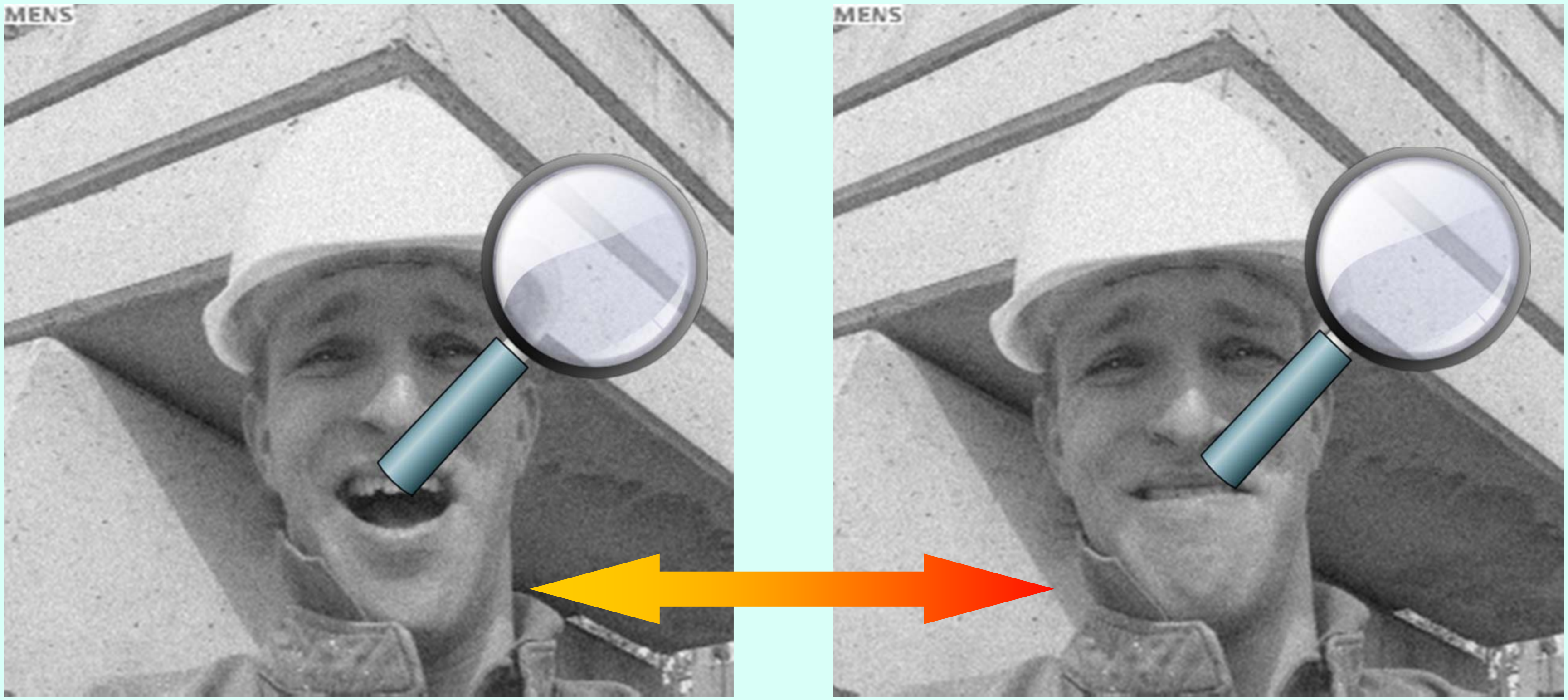
All pixels moved exactly 1/3 of a pixel to the right

Motion Estimation – Another Example

All pixels moved 2 pixels to the right and 1 pixel up

Motion Estimation: Global versus Local

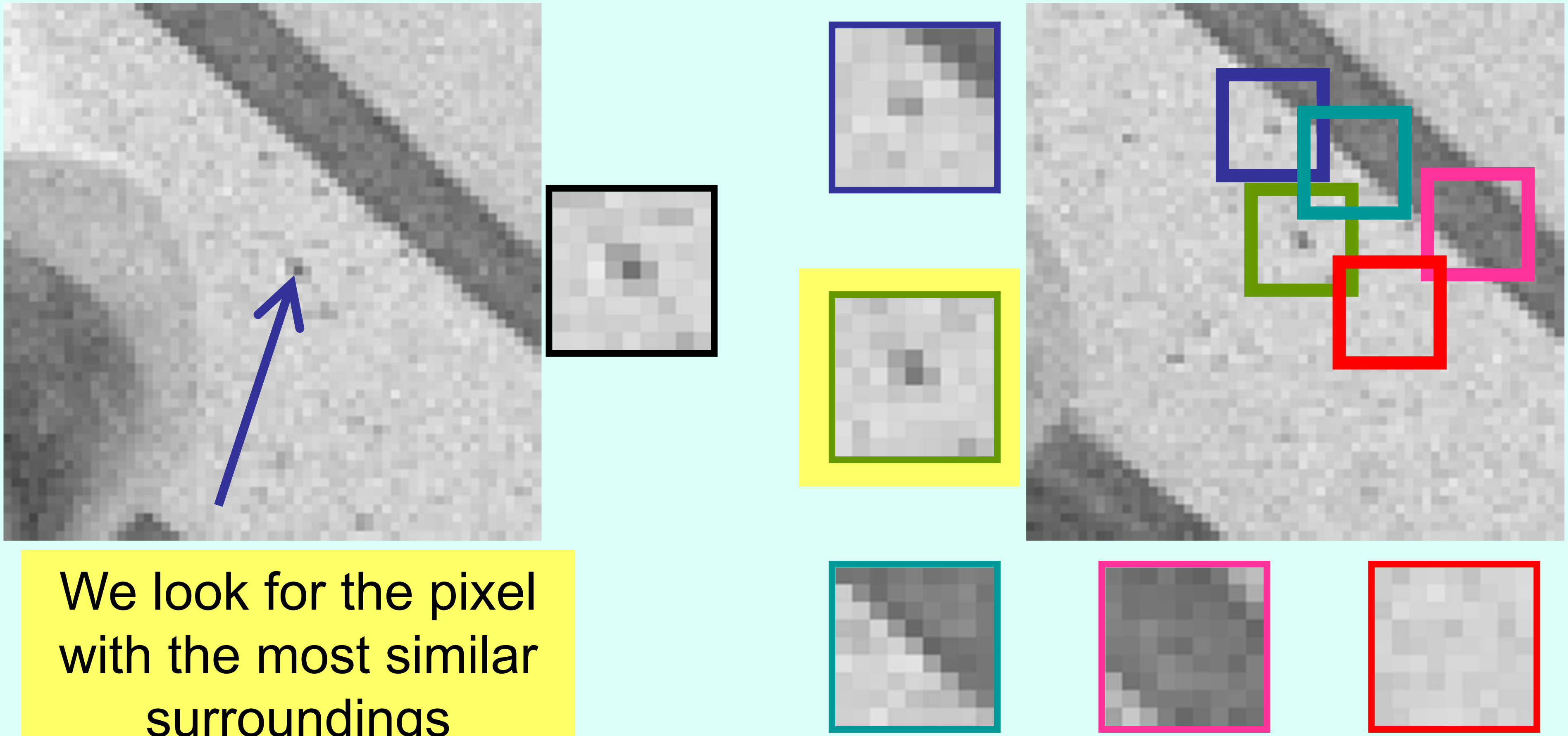
What about here?



For SR, we need the accurate motion of **every** pixel

Motion Estimation – Basic Technique

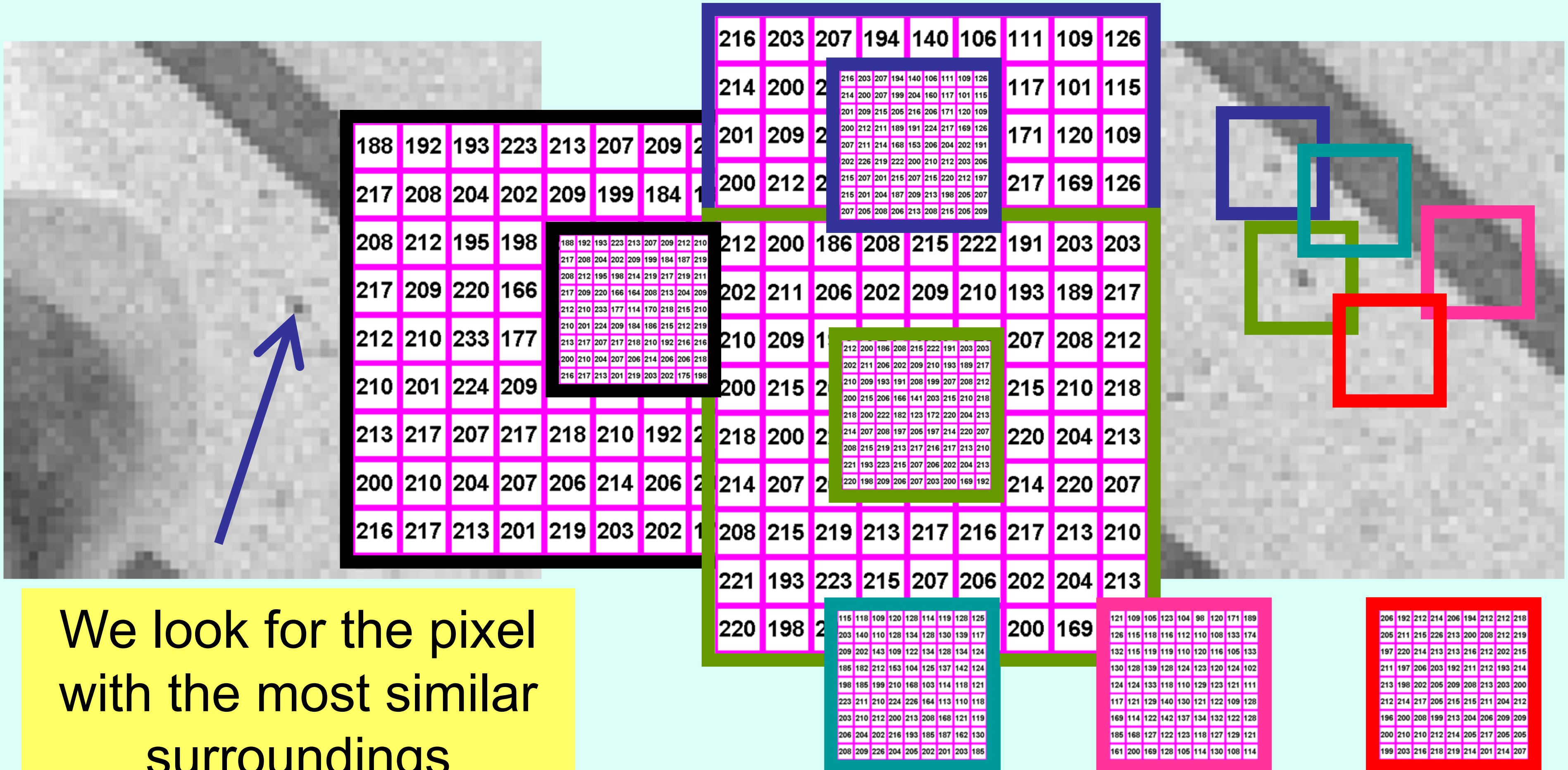
Computing the motion for a **specific pixel**



We look for the pixel with the most similar surroundings

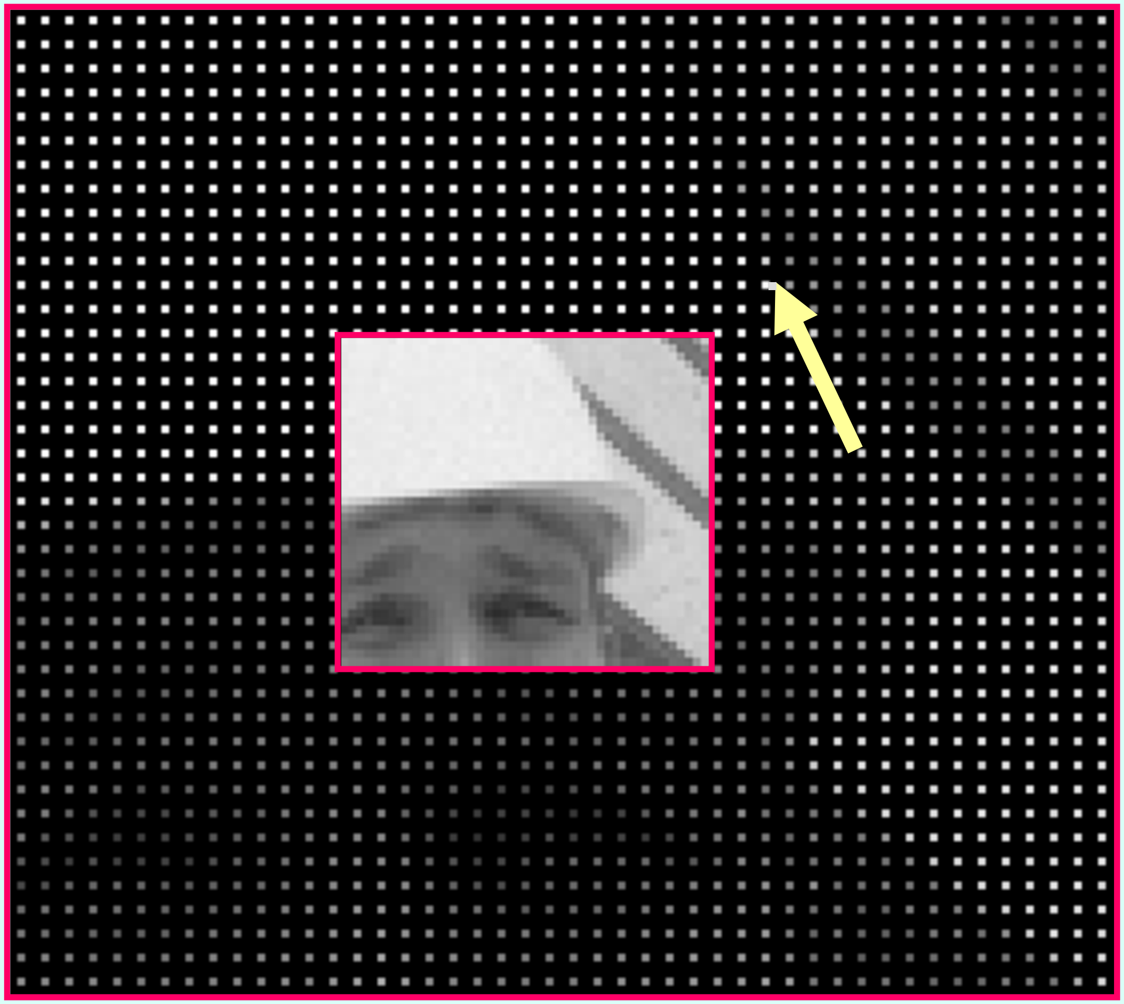
Motion Estimation – Basic Technique

Computing the motion for a **specific pixel**

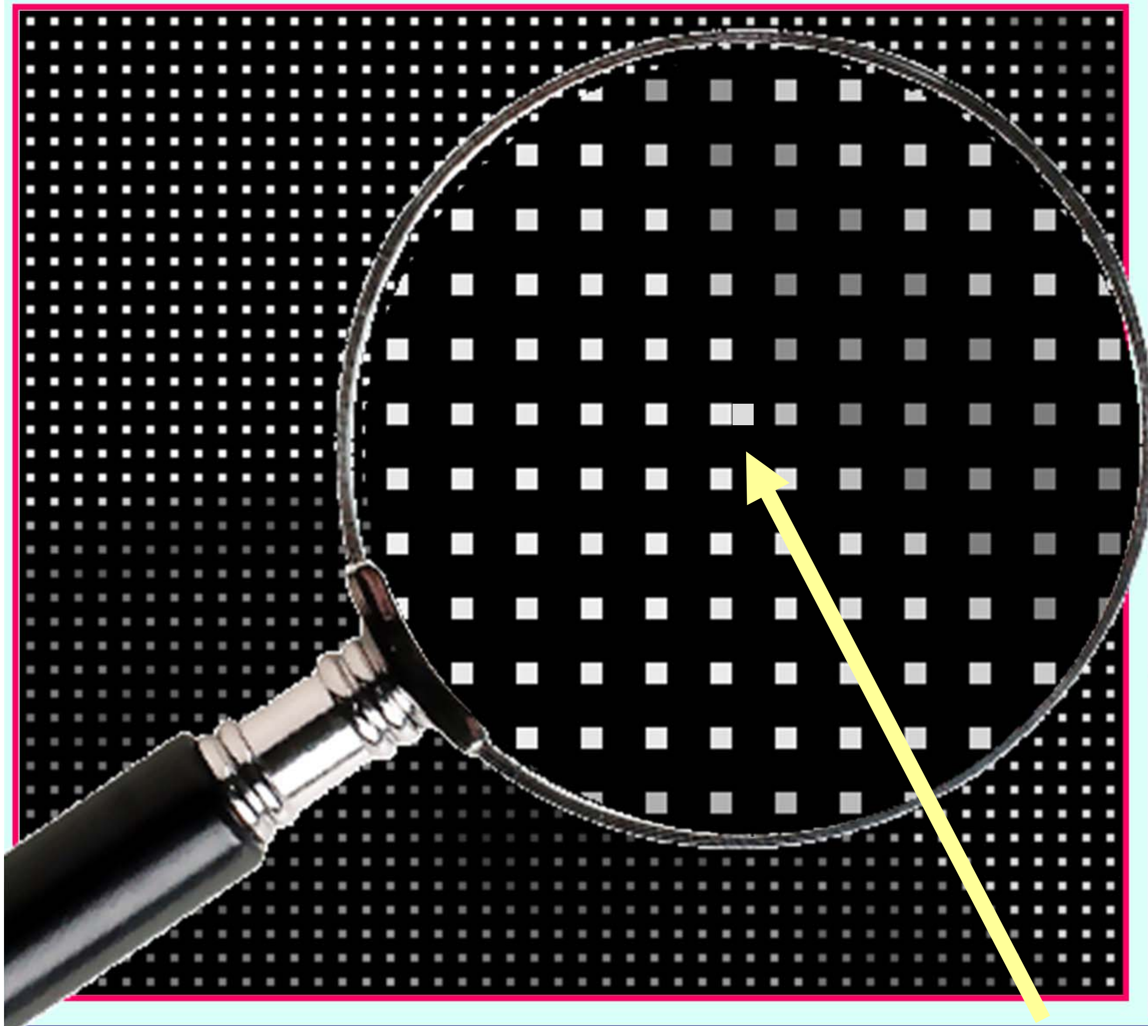


We look for the pixel with the most similar surroundings

SR with Motion Estimation



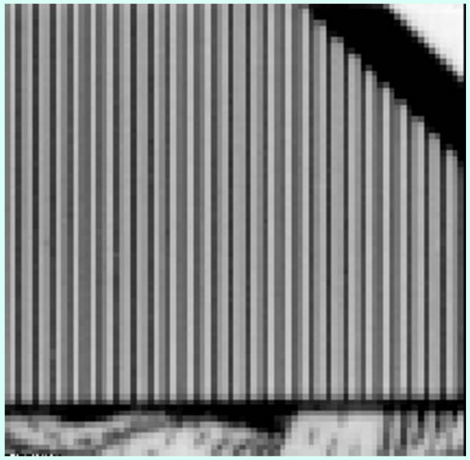
SR with Motion Estimation

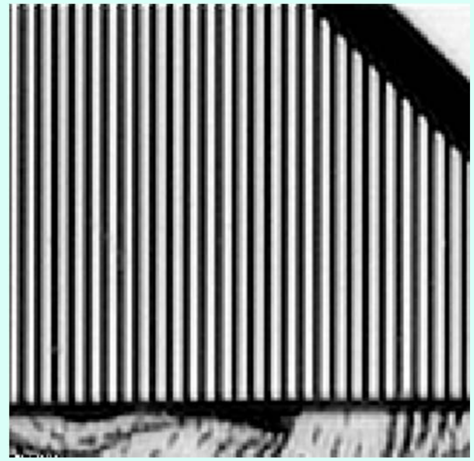


SR With Motion Estimation – Results (Scanner)

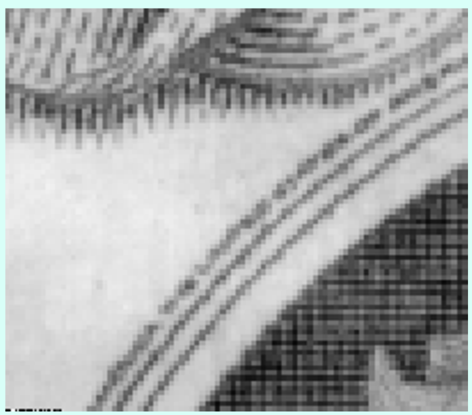
16 images, ratio 1:2 in each axis

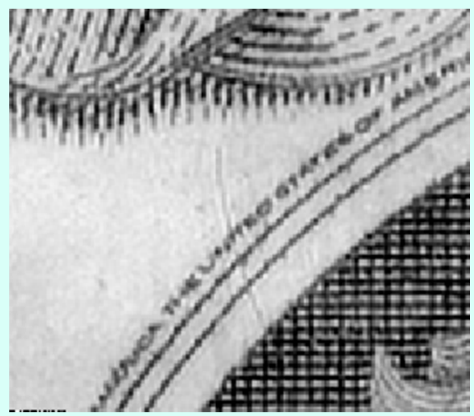
Taken from one of the given images





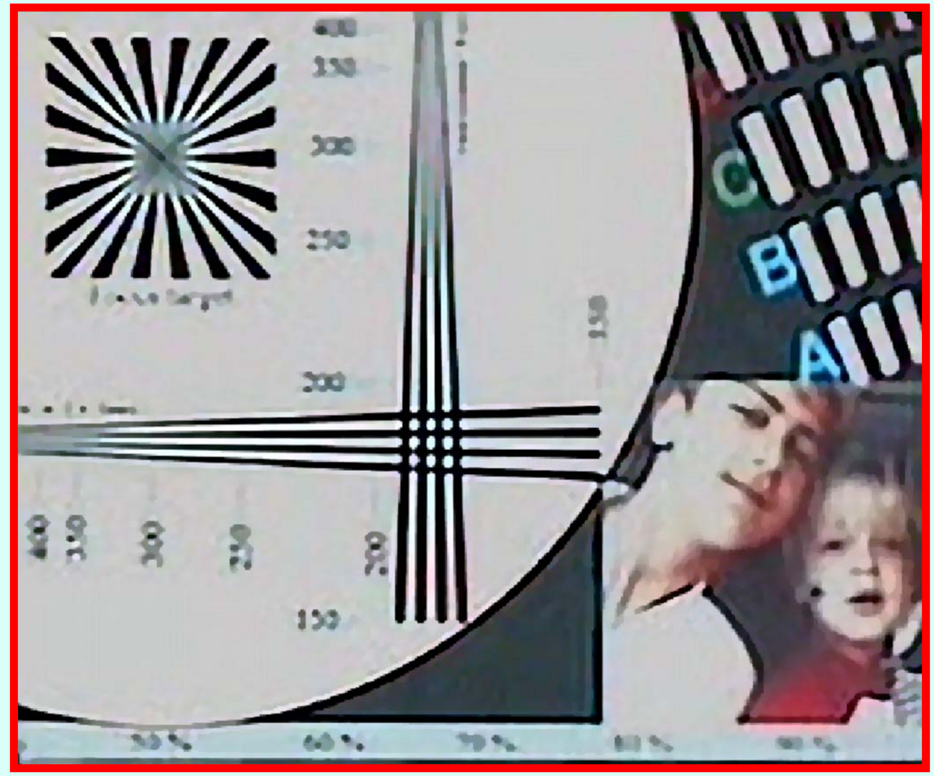
Taken from the reconstructed result





Elad & Hel-Or (1999)

SR With Motion Estimation - Results



Farsiu, Elad & Milanfar (2004)

SR With Motion Estimation - Results

Farsiu, Elad & Milanfar (2004)

SR With Motion Estimation - Results

40 images
ratio 1:4

30 images
ratio 1:4

Farsiu, Elad & Milanfar (2004)

SR With Motion Estimation - Results

Farsiu, Elad & Milanfar (2005)

SR With Motion Estimation

- ❑ Has been practiced for about 20 years, since ~1987 and till 2007 along the above lines
- ❑ This process requires knowing the motion very accurately
  - Very difficult to obtain when motion is not global
  - Output result is HORRIBLE when estimation makes mistakes
- ❑ Therefore, SR has been limited to movies with simple motion …
- ❑ This may explain why we have not seen this technology coming into play in the industry

SR Doesn't Work Here …

Original

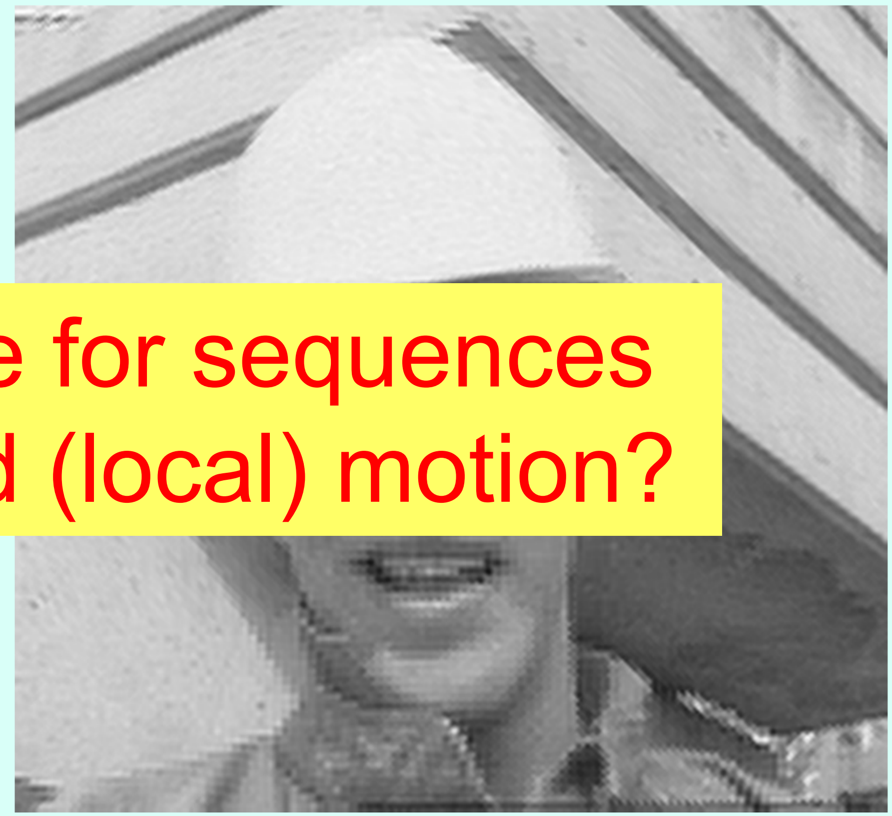
Interpolation

Super-Resolution
with motion estimation

And a Snapshot …

True high-quality

Super-Resolution with motion estimation



Is there no hope for sequences with complicated (local) motion?

Agenda

1. Introduction

   Basic Concepts: Image, Pixel, Resolution

2. Classic Super-Resolution

   How does it Work?

   Problems and (Many) Limitations

3. **New Era of Super-Resolution**

   Changing the Basic Concept

   New Results

4. Summary

Joint work with

Matan Protter

Principles of the New SR Approach

- The solution is based on:
  - Probabilistic motion estimation
  - Using self–motion that exploits spatial redundancy
  - Local processing
- Our mathematical formulation (for those who care):

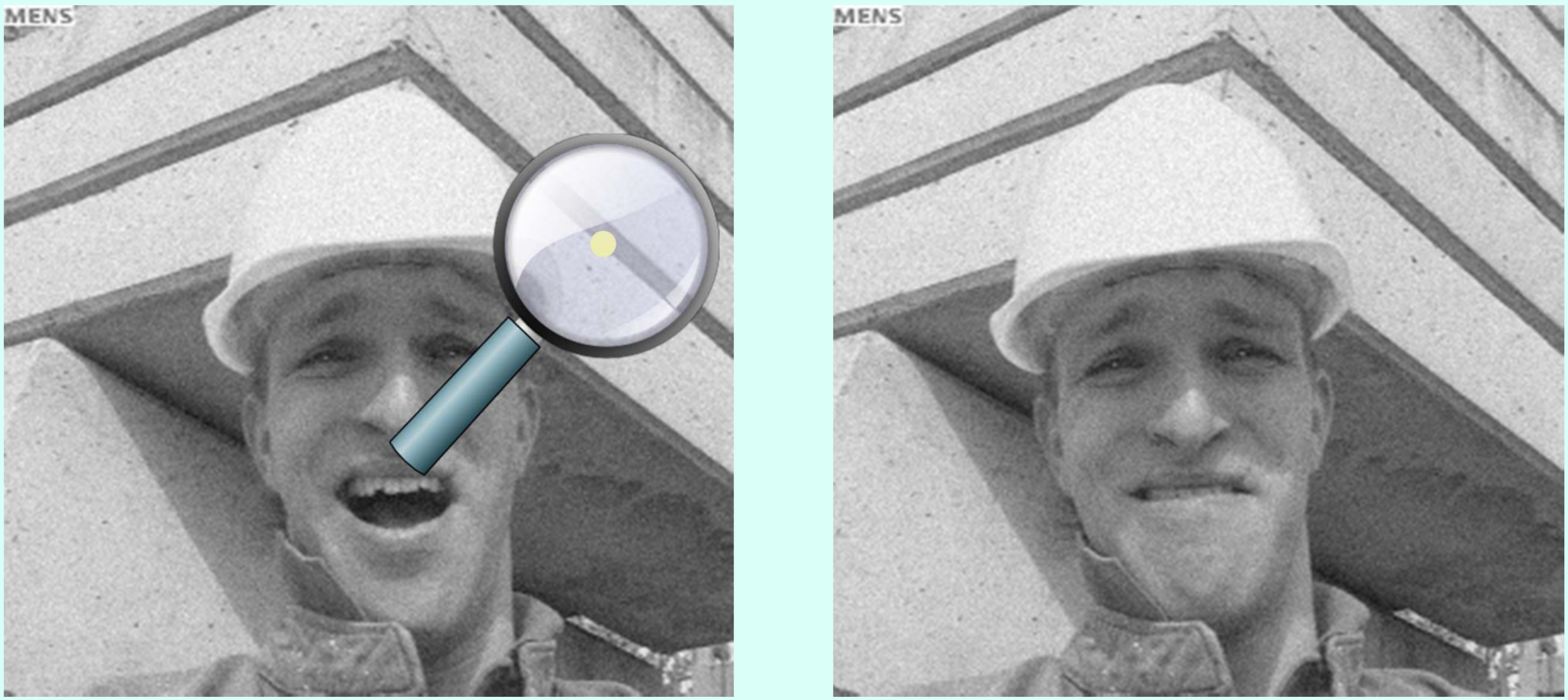
$$\eta_{SR}(\mathbf{X}) = \sum_{(k,l)\in\Omega} \sum_{t} \sum_{(i,j)\in N(k,l)} \mathbf{w}[k,l,i,j,t] \cdot \left\| \mathbf{D}_{\mathbf{p}} \mathbf{R}_{kl}^{\mathbf{H}} \mathbf{H} \mathbf{X} - \mathbf{R}_{ij}^{\mathbf{L}} \mathbf{y}_t \right\|_2^2$$

- This formulation leads to a family of algorithms. We will discuss the simplest of them …

The Problem

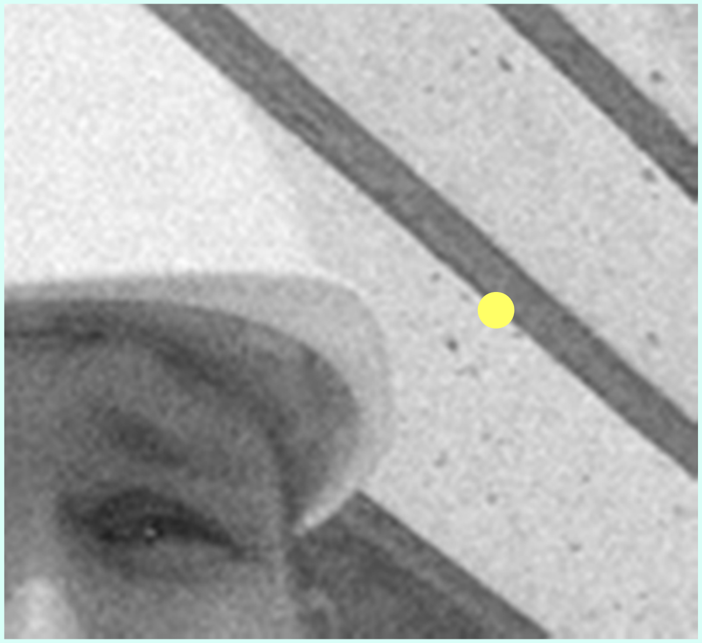
Probabilistic Motion Estimation

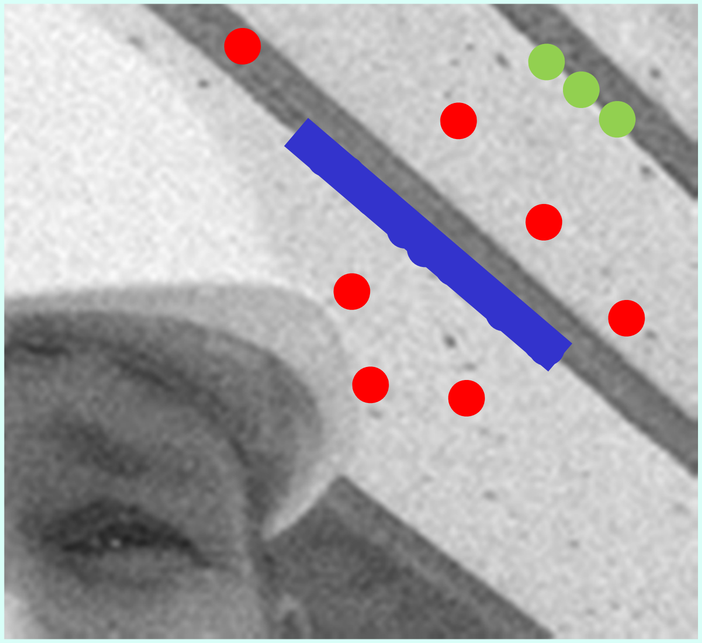
Where does this pixel go?



Probabilistic Motion Estimation

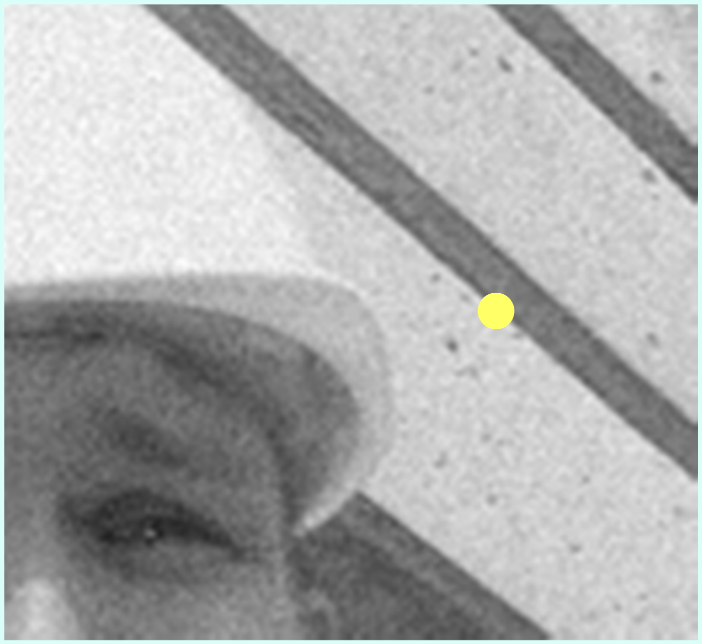
Where does this pixel go?

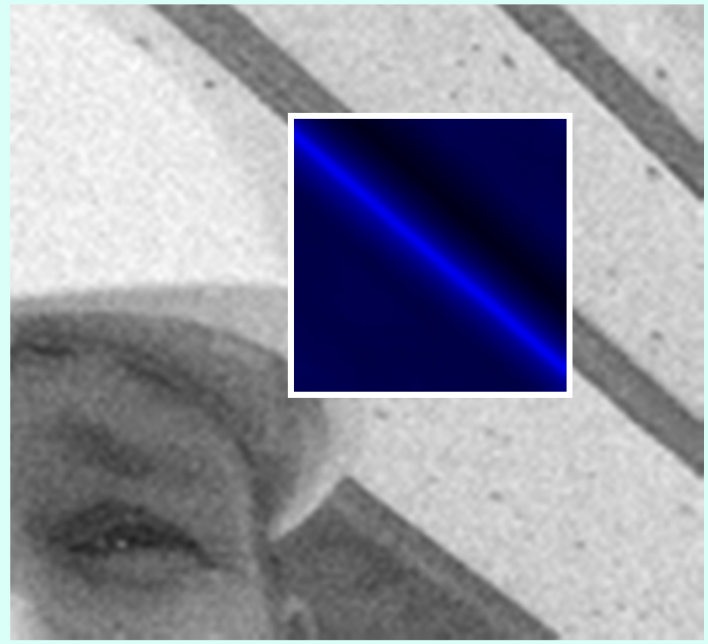




Probabilistic Motion Estimation

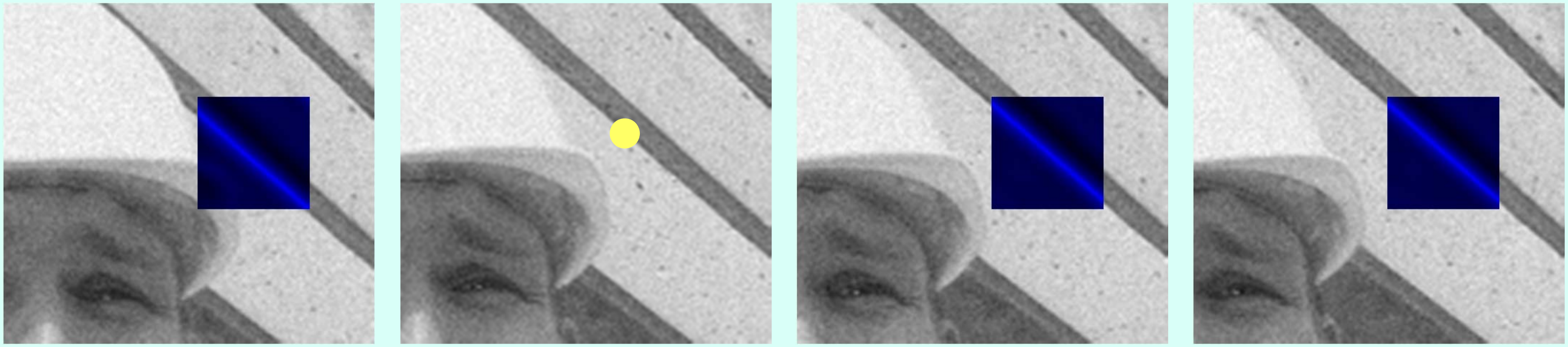
Where does this pixel go?





Probabilistic Motion Estimation

Where does this pixel go?

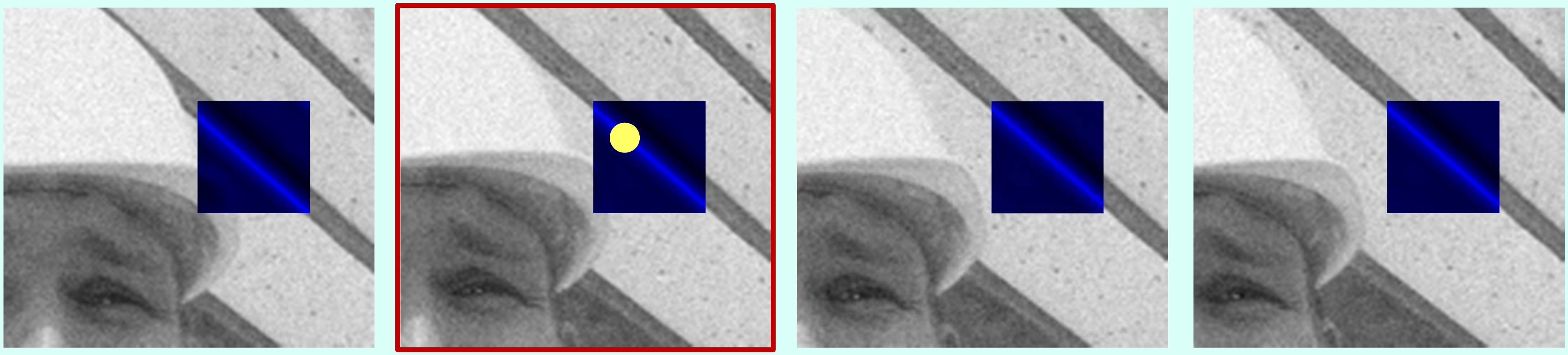


In probabilistic motion estimation, all motions are possible

Some are just more likely than others

Exploiting Self-Motion

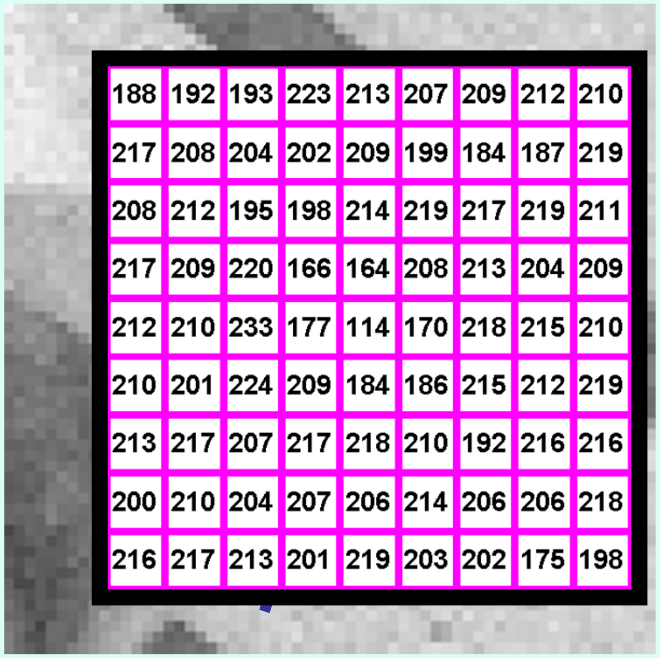
What about similar patches within the same image?



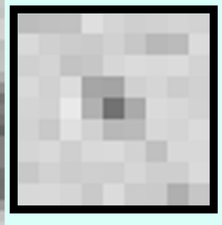
- When seeking the relevant matches, we consider matches within the same image just as well
- Thus, even one image can boost itself to SR if it contains self-similarities

Computing The Probabilities

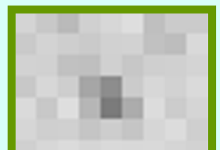
Computing probabilities for a **specific pixel**



188	192	193	223	213	207	209	212	210
217	208	204	202	209	199	184	187	219
208	212	195	198	214	219	217	219	211
217	209	220	166	164	208	213	204	209
212	210	233	177	114	170	218	215	210
210	201	224	209	184	186	215	212	219
213	217	207	217	218	210	192	216	216
200	210	204	207	206	214	206	206	218
216	217	213	201	219	203	202	175	198

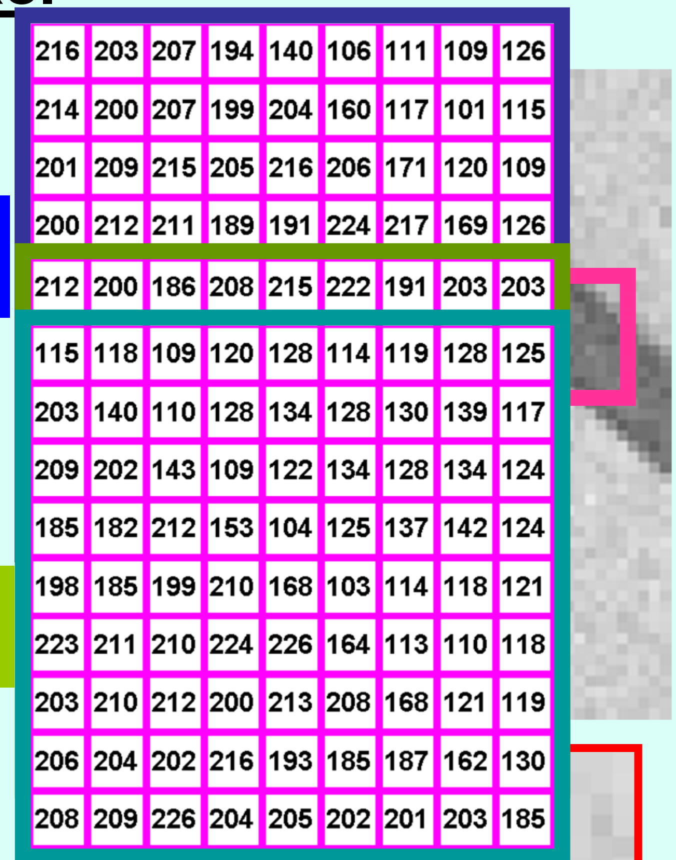


Probability: low



Probability: high

Probability: zero



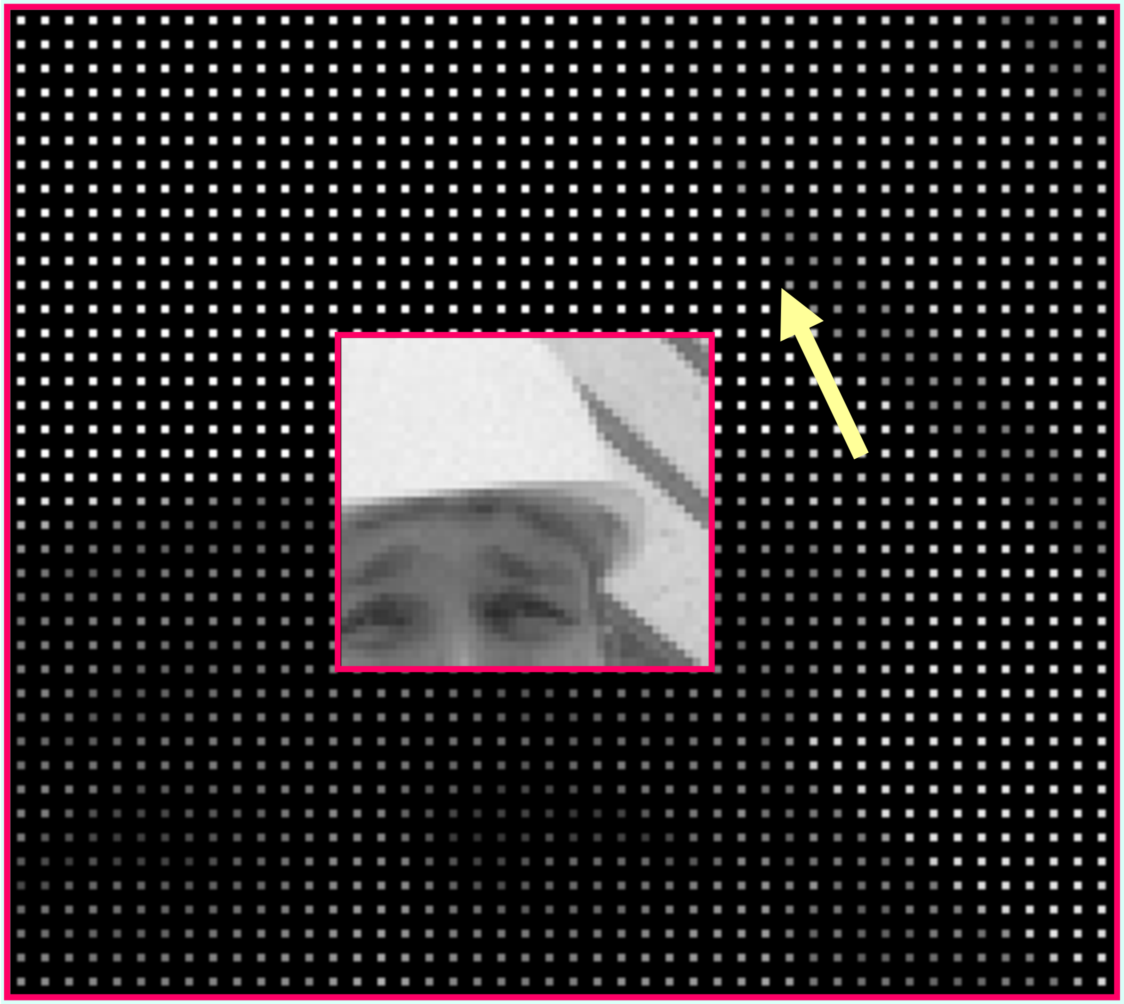
216	203	207	194	140	106	111	109	126
214	200	207	199	204	160	117	101	115
201	209	215	205	216	206	171	120	109
200	212	211	189	191	224	217	169	126
212	200	186	208	215	222	191	203	203
115	118	109	120	128	114	119	128	125
203	140	110	128	134	128	130	139	117
209	202	143	109	122	134	128	134	124
185	182	212	153	104	125	137	142	124
198	185	199	210	168	103	114	118	121
223	211	210	224	226	164	113	110	118
203	210	212	200	213	208	168	121	119
206	204	202	216	193	185	187	162	130
208	209	226	204	205	202	201	203	185

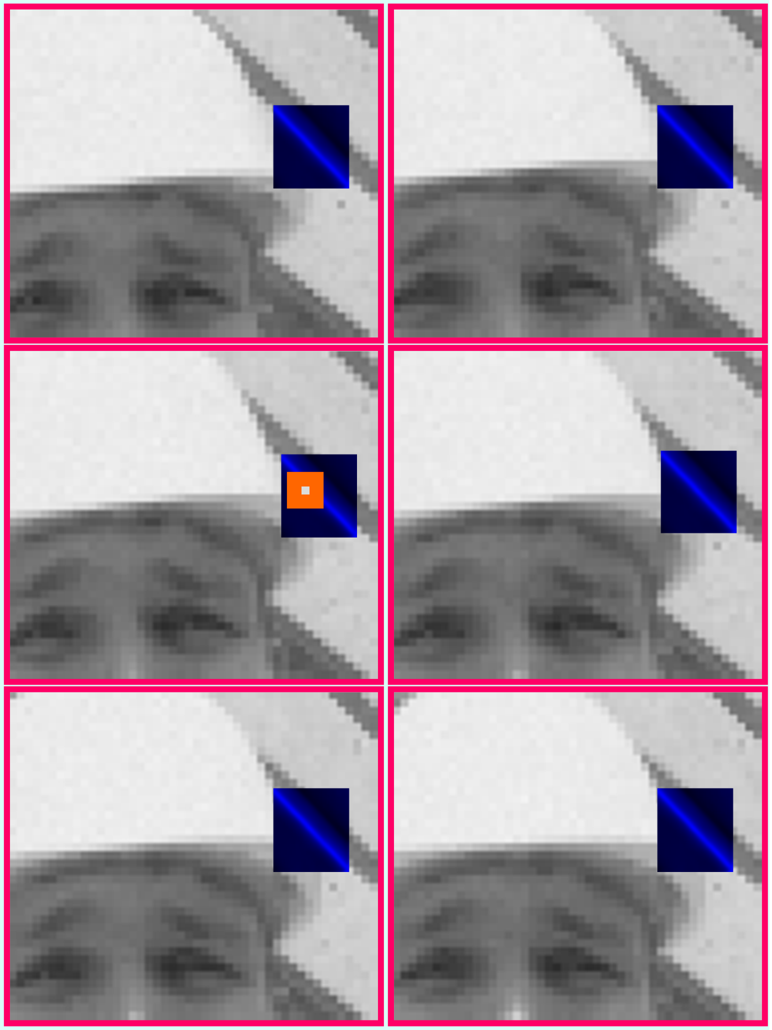
Probability: zero

Probability: low

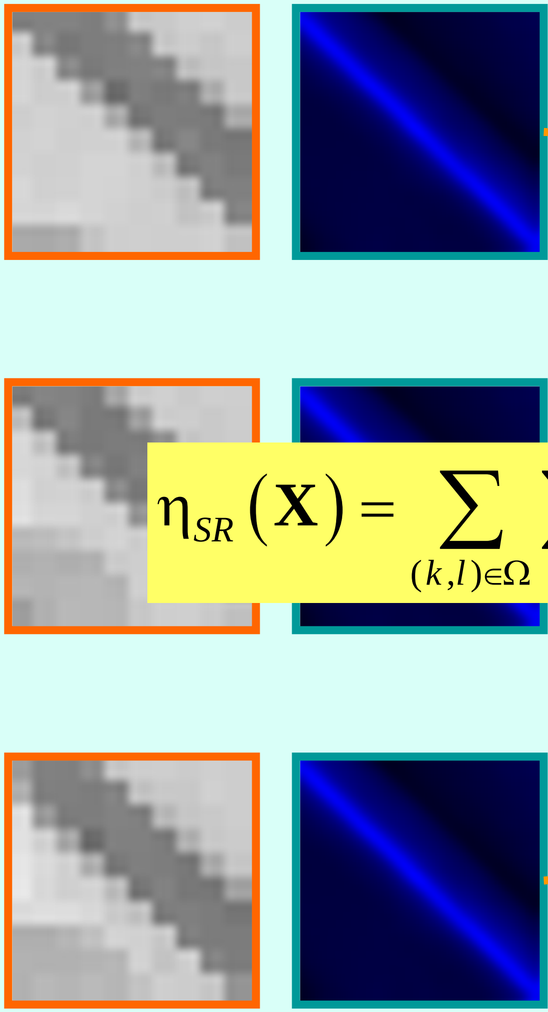
We use similarity of surroundings to measure the probability

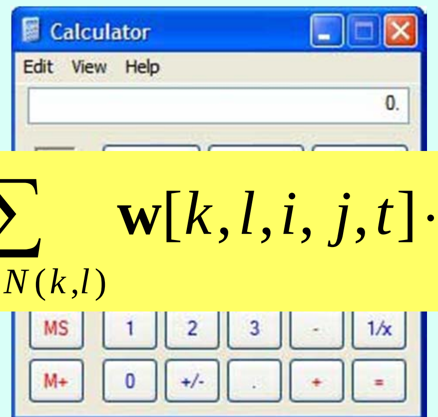
SR with Probabilistic Motion Estimation



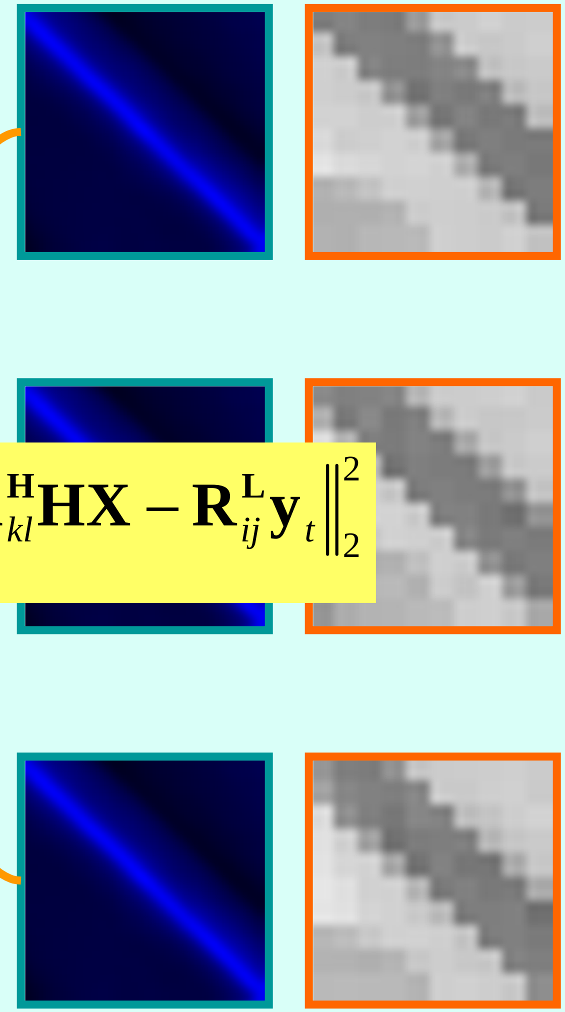


SR with Probabilistic Motion Estimation





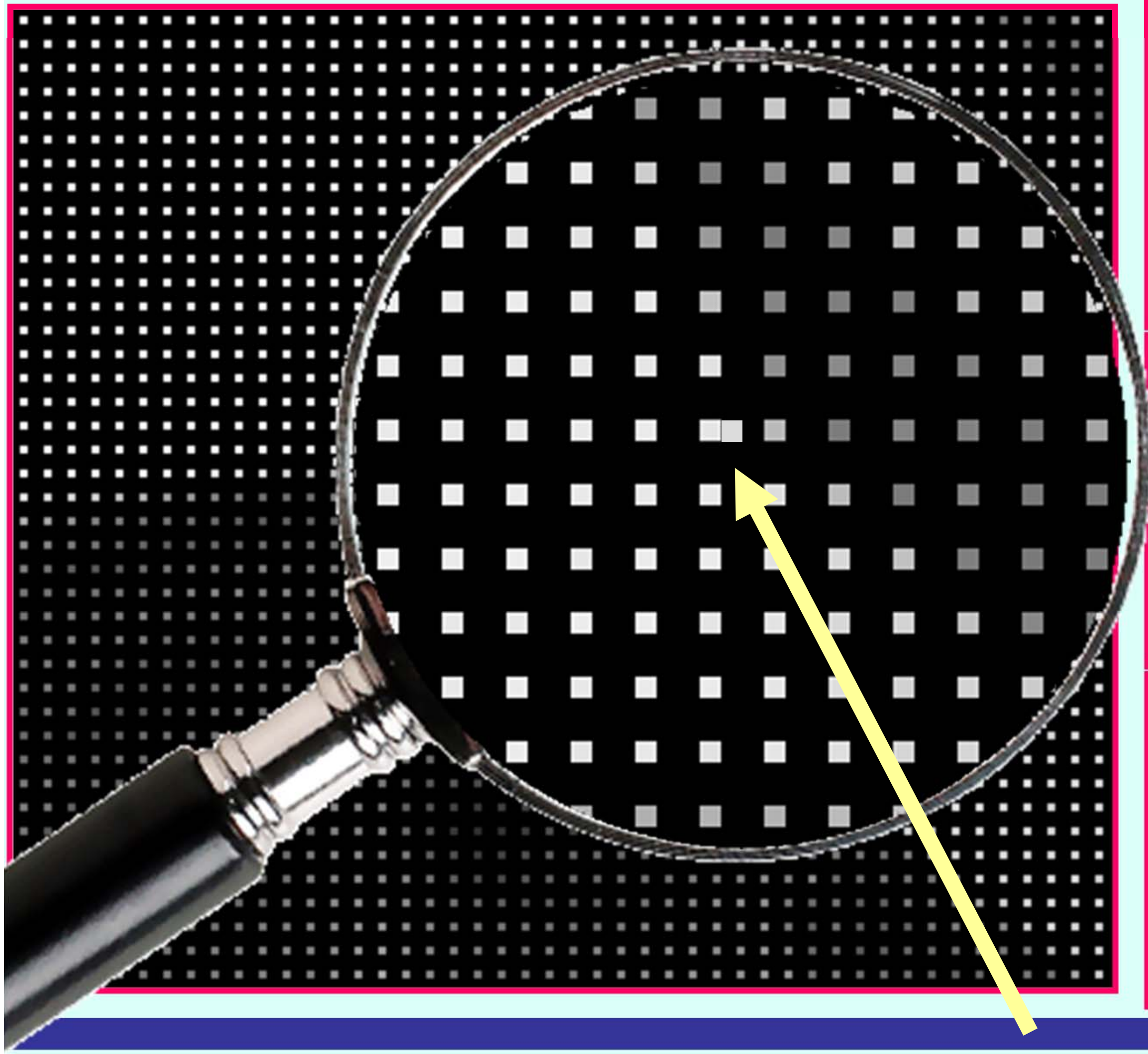
$$\eta_{SR}(\mathbf{X}) = \sum_{(k,l)\in\Omega} \sum_{t} \sum_{(i,j)\in N(k,l)} \mathbf{w}[k,l,i,j,t] \cdot \left\| \mathbf{D}_{\mathbf{p}} \mathbf{R}_{kl}^{\mathbf{H}} \mathbf{H}\mathbf{X} - \mathbf{R}_{ij}^{\mathbf{L}} \mathbf{y}_t \right\|_2^2$$



Result Pixel Value:
132

SR with Probabilistic Motion Estimation

SR with Probabilistic Motion Estimation



Results: Miss America

Input Sequence
(30 Frames)

Created from original high-res. sequence using 3x3 uniform blur, 3:1 decimation, and noise with std = 2

Original Sequence
(Ground Truth)

Lanczos
Interpolation

Algorithm Result

Results: Foreman

Input Sequence (30 Frames)

Original Sequence (Ground Truth)

Lanczos Interpolation

Algorithm Result

Results: Foreman

Classic Super-Resolution | **New Super-Resolution**

Results: Salesman

Input
Sequence
(30 Frames)

Original
Sequence
(Ground Truth)

Lanczos
Interpolation

**Algorithm
Result**

Results: Suzie

Input Sequence (30 Frames)

Original Sequence (Ground Truth)

Lanczos Interpolation

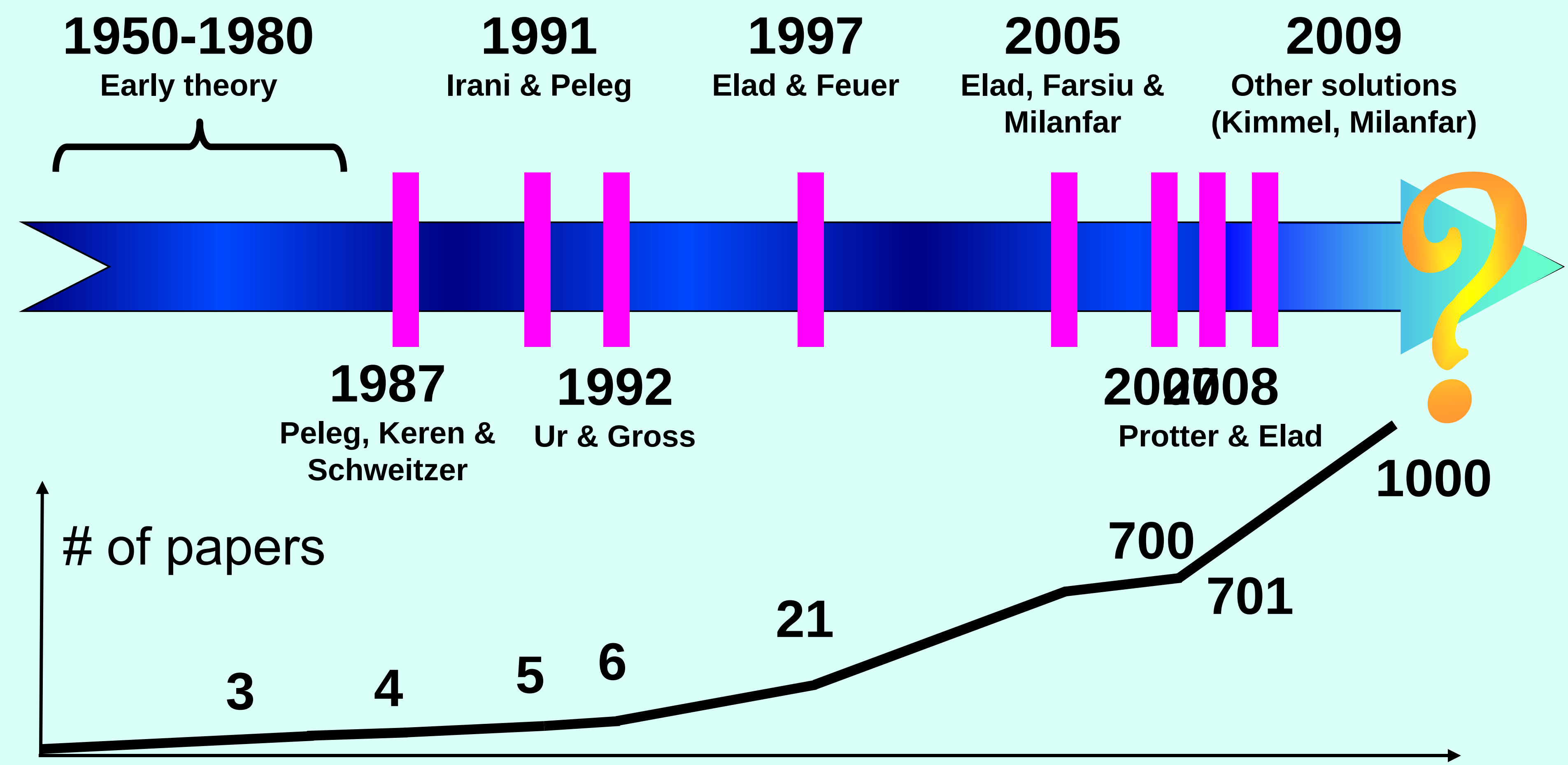
Algorithm Result

Results : SD to HD

Agenda

1. Introduction
   - Basic Concepts: Image, Pixel, Resolution
2. Classic Super-Resolution
   - How does it Work?
   - Problems and (Many) Limitations
3. New Era of Super-Resolution
   - Changing the Basic Concept
   - New Results
4. **Summary**

Evolution of Super-Resolution



Limitations in The Proposed Approach

Two main limitation:

- The improvement is not always large
  - Especially if the movie has been “played around” with (e.g. deep compression)
  - It does not look worse, though
- Computation time is huge
  - We are still far from improving movies as they are played (real-time processing) by factor 20:1 (with GPU implementation)

Summary

Problem:
displaying low-quality
sequences on high-quality

Solution:
Super-Resolution can
improve the quality of the

Implications:

- No limitation on type of sequence
- Typically much improved quality
- Can be applied to other tasks
- Many attempts to improve/extend/compete with this approach

B

Intro
motion estimation into
Super-Resolution

:

urate
motion estimation; only
possible in few sequences

Thank You For Inviting Me

Questions?