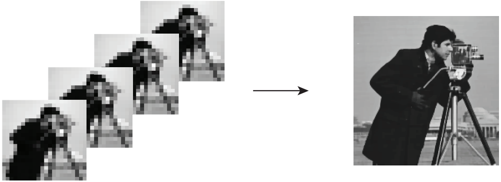
Super-resolution with Structured Motion

Gabby Litterio, David Lizarazo-Ferro,
Pedro Felzenszwalb, Rashid Zia

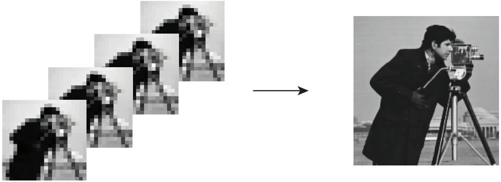
Brown University

Overview



- Super-resolution: reconstruct high-resolution image from one or more low-resolution images.

Overview



- Super-resolution: reconstruct high-resolution image from one or more low-resolution images.
- Focus on imaging constraints, low-level priors and sensor/camera motion.
- Connection between super-resolution and box filters.
- We recover high-resolution images using motion blur.

Motion + Low-resolution $\rightarrow$ High-resolution

Motion + Low-resolution → High-resolution

Multiple sub-pixel translations

1 of 64 images

Interlaced data

Deconvolution

Motion + Low-resolution → High-resolution

Multiple sub-pixel translations

1 of 64 images

Interlaced data

Deconvolution

Motion blur

Motion + Low-resolution → High-resolution

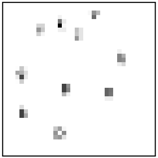
Multiple sub-pixel translations


1 of 64 images


Interlaced data


Deconvolution

Motion blur



Motion + Low-resolution → High-resolution

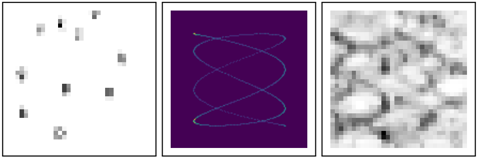
Multiple sub-pixel translations

1 of 64 images

Interlaced data

Deconvolution

Motion blur



Motion + Low-resolution → High-resolution

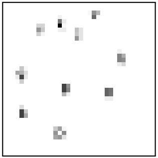
Multiple sub-pixel translations

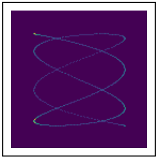

1 of 64 images

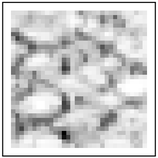

Interlaced data


Deconvolution

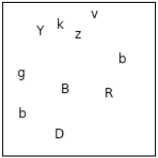
Motion blur


Static Image


Trajectory



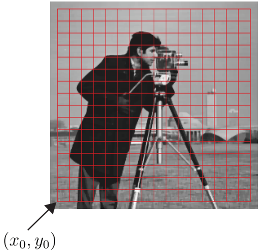
Image



Reconstruction

Imaging Model

Camera with a translating sensor (or equivalent).



Sensor integrates brightness $g(x, y)$ over square pixels.

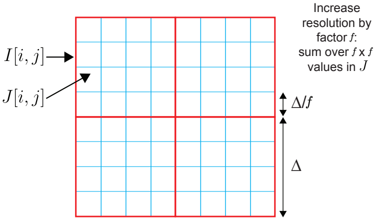
Static exposure

$$I[i,j] = \int_0^{\Delta} \int_0^{\Delta} g(x_0 + i\Delta + x, y_0 + j\Delta + y) dxdy.$$

Moving exposure

$$I[i,j] = \int_0^{T} \int_0^{\Delta} \int_0^{\Delta} g(x_0(t) + i\Delta + x, y_0(t) + j\Delta + y) dxdydt.$$

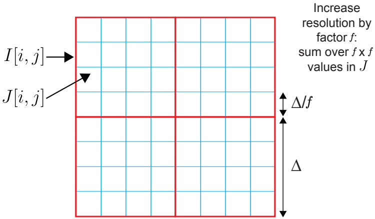
Box Convolution



$$I = (J \otimes B) \downarrow_f$$

f: resolution increase factor
J: high-resolution image
I: low-resolution image
B: $f \times f$ box filter

Box Convolution



$$I = (J \otimes B)\downarrow_f$$

f: resolution increase factor
J: high-resolution image
I: low-resolution image
B: $f \times f$ box filter

Capturing a super-resolution image

1) $f \times f$ subpixel shifts $\rightarrow$ collection of images $\{I_{k,l}\}$.
2) Interlacing $\{I_{k,l}\} \rightarrow H$.

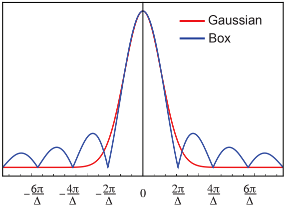
$$H = J \otimes B$$

3) Recovering J involves a deconvolution.

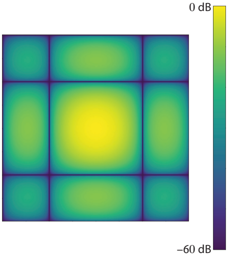
Box Filter

$$H = J \otimes B$$

- Luckily, a box filter is a bad low-pass filter.
- H determines most of the Fourier coefficients of J.
- In this setting sparse signals can be be recovered with convex optimization [Donoho 89] [Candes, Romberg, Tao 06] [F 25].



FT of 1D Gaussian vs box



DFT of a 4x4 box

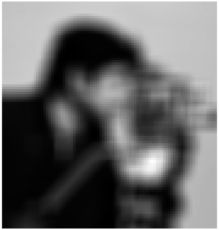
Revisiting the Limits of Super-Resolution [Baker and Kanade]

[Baker and Kanade]
The Limits of Super-Resolution and How to Break them, 2002.

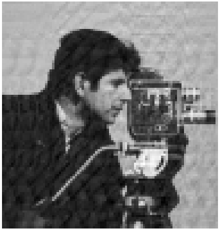
Revisiting the Limits of Super-Resolution [Baker and Kanade]

[Baker and Kanade]
The Limits of Super-Resolution and How to Break them, 2002.

Two (particular) solutions J with $H = (J \otimes B)$



H



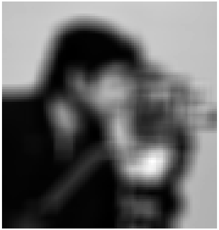
Quadratic prior [BK]

TV prior

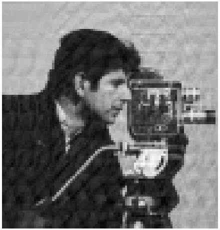
Revisiting the Limits of Super-Resolution [Baker and Kanade]

[Baker and Kanade]
The Limits of Super-Resolution and How to Break them, 2002.

Two (particular) solutions J with $H = (J \otimes B)$



H



Quadratic prior [BK]

TV prior

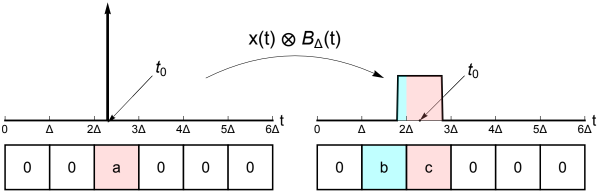
Sparse images can be recovered via convex optimization:

$$J^* = \operatorname*{argmin}_J ||H - (J \otimes B)||_2^2 + \lambda TV(J)$$

$$TV(J) = \sum_{i,j} |J[i,j] - J[i+1,j]| + |J[i,j] - J[i,j+1]|$$

Motion Blur Can Help

Localizing a point source in one dimension

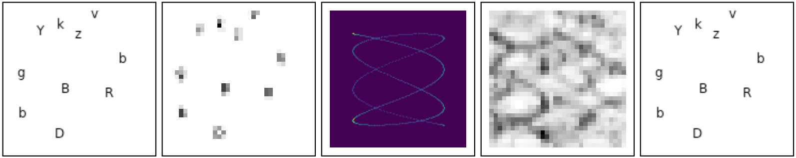


$$a = b + c,$$

$$t_0 = (k + 0.5)\Delta + c/(b + c)).$$

Motion Blur for Super-Resolution

We can use *large* motions to superimpose multiple low-res images.



Target | Static Image | Trajectory | Image | Reconstruction

$$I = (J \otimes Q \otimes B) \downarrow_f$$

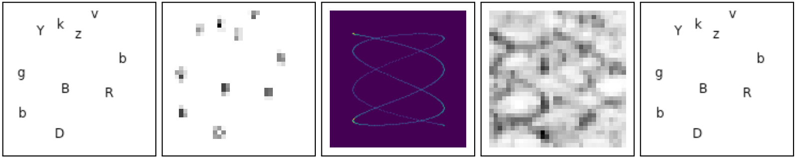
Q: occupancy mask (blur kernel)

Reconstruction:

$$J^* = \underset{J}{\operatorname{argmin}} \, ||I - (J \otimes Q \otimes B) \downarrow_f ||_2^2 + \lambda ||J||_1$$

Motion Blur for Super-Resolution

We can use *large* motions to superimpose multiple low-res images.



Target | Static Image | Trajectory | Image | Reconstruction

$$I = (J \otimes Q \otimes B) \downarrow_f$$

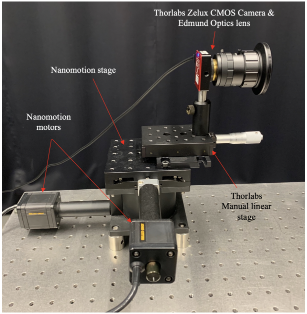
Q: occupancy mask (blur kernel)

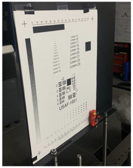
Reconstruction:

$$J^* = \underset{J}{\operatorname{argmin}} \, ||I - (J \otimes Q \otimes B) \downarrow_f ||_2^2 + \lambda ||J||_1$$

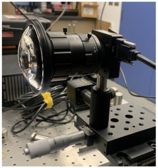
Blind deconvolution (solving for Q/J simultaneously) can increase resolution

Experimental Setup

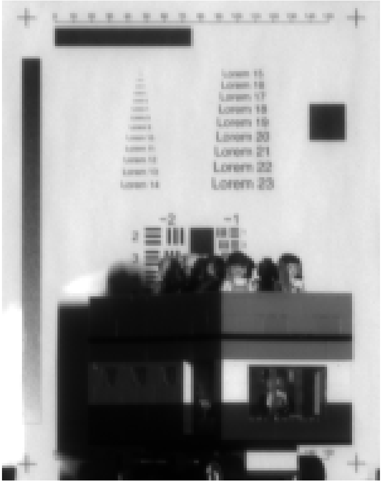




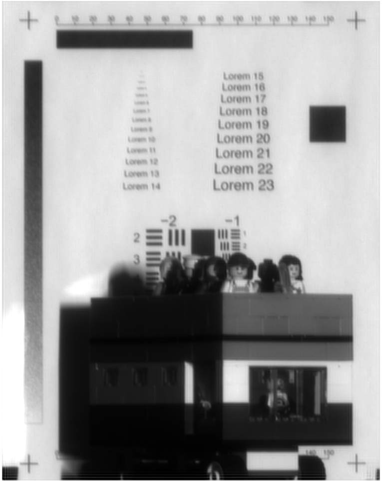
~ 3 meters



Multiple Shifts

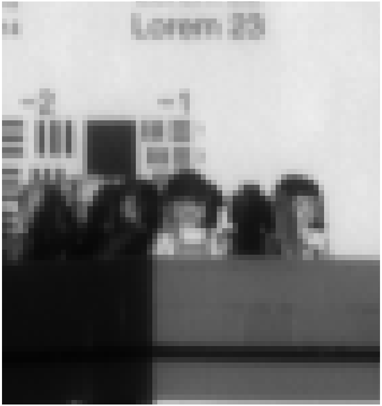


Static image (1 of 64)

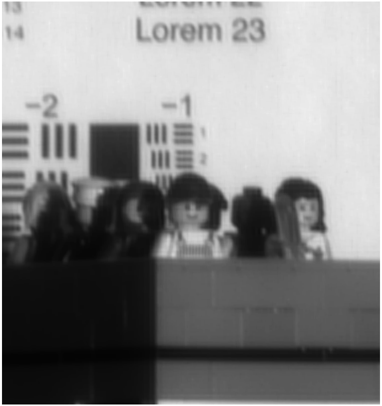


TV reconstruction

Multiple Shifts



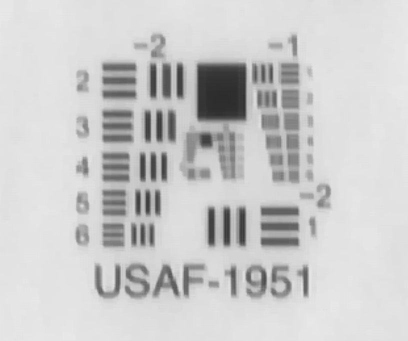
Static image (1 of 64)



TV reconstruction

Scanning Motion

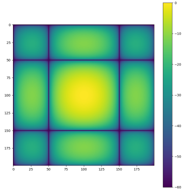
Static image



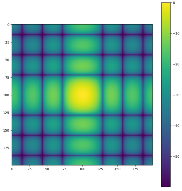
TV reconstruction

Summary and Future Work

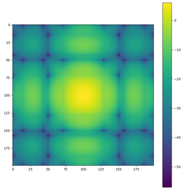
- Sparse image priors resolve the ambiguity of super-resolution under translation.
- We can recover high-resolution images *using* motion blur.
- Other implementations: pixel-shift camera, nano-positioning stage, satellite/drone, etc.
- Multiple pixel sizes (magnifications) can remove all ambiguities.



$|\,\mathrm{DFT}(B_4)|$



$|\,\mathrm{DFT}(B_7)|$



$|\,\mathrm{DFT}(B_4)| + |\,\mathrm{DFT}(B_7)|$