Compressive Imaging

Aswin Sankaranarayanan
(Computational Photography – Fall 2017)
Traditional Models for Sensing

- Linear (for the most part)
- Take as many measurements as unknowns

\[ x \xrightarrow{\text{sample}} y \]
Traditional Models for Sensing

- Linear (for the most part)
- Take as many measurements as unknowns

Typically, \( M \geq N \)
Under-determined problems

Fewer measurements than unknowns!

An infinite number of solutions to such problems
Under-determined problems

Fewer measurements than unknowns!

An infinite number of solutions to such problems

Is there any “hope” of solving these problems?
Complete the sentences

I cnt blv I m bl t rd ths sntnc.

Wntr s cmng. n wt, wntr hs cm.

Hy, I m slvng n ndr-dtrmnd lnr systm.

how: ?
Complete the matrix

How: ?
Complete the image

Model ?
Real data has structure

Image gradients are sparse!

Image credit: David W Kennedy (Wikipedia)
Real data has structure

Real world images: Only a few non-zero coefficients in a transformation
A toolset to solve under-determined systems by exploiting additional structure/models on the signal we are trying to sense.
Suppose measurement matrix $A$ satisfied certain conditions:

- $M \geq c_1 K \log(N/K)$
- All $K$-sparse signals $x$ can be recovered
  - In the absence of noise, the recovery is exact!

[Candes and Tao, 2004]
Compressive Sensing: Big Picture

• **If signal has structure, exploit it to solve under-determined problem**

• **Structure**: Refers to a lower-dimensional parametrization of the signal class
  – Sparsity in a basis (like Fourier or wavelets)
  – Sparsity of gradients
  – Low-rank, low-dim smooth manifold
  – Any set with a projection operator

• Number of measurement is often proportional to the dim of the low-dim parameters

• Range of recovery techniques
• (Take **18-898G** next semester for a deep dive)
High-speed videography using CS

Key papers
Veeraraghavan et al., *Coded strobing, PAMI 2011*
Reddy et al., *P2C2, CVPR 2011*
Hitomi et al., Coded exposure, *ICCV 2011*
Low-speed capture works well for static scenes
High-speed scenes

33 ms
High-speed scenes

Blurring in dynamic areas

High spatial resolution in static areas
High speed scenes

Image credit: Boston.com
Spatio-Temporal Resolution Tradeoff

Single image

Spatial Resolution = 1 Megapixel
Temporal Resolution = 1 fps
Bandwidth = 1 Megapixel/s

Slide credit: Mohit Gupta (Hitomi et al. 2011)
Spatio-Temporal Resolution Tradeoff

Interpolated movie

Spatial Resolution = 1/4 Megapixel
Temporal Resolution = 4 fps
Bandwidth = 1 Megapixel/s

Slide credit: Mohit Gupta (Hitomi et al. 2011)
Spatio-Temporal Resolution Tradeoff

Interpolated movie

Spatial Resolution = 1/36 Megapixel
Temporal Resolution = 36 fps
Bandwidth = 1 Megapixel/s

Slide credit: Mohit Gupta (Hitomi et al. 2011)
Spatio-Temporal Resolution Tradeoff

Challenges
1. Bandwidth of data
2. Light throughput

Slide credit: Mohit Gupta (Hitomi et al. 2011)
From this photo ...
... to this one

Credit: Edgerton

Idea 1: Multiplexing in Time

33 ms
Idea 1: Multiplexing in Time

**Benefits**
1. Bandwidth of data remains the same
2. Light throughput is not significantly reduced
Idea 1: Multiplexing in Time

Challenge:
More unknowns than measurements
How do we recover?
Idea 2: Signal Models

• Real-world signals are highly redundant
Sparsity

$N$ pixels

$K < N$
large wavelet coefficients
(blue = 0)

$N$ wideband signal samples

$K < N$
large Gabor (TF) coefficients

frequency
time
Idea 2: Signal Models

• Real-world signals are highly redundant

• Models
  – Sparse gradients
  – Sparse in transform: Wavelets, Fourier
  – Low rank: PCA, Union-of-subspaces

• Key idea: **Constrain** the solution space!
  – Number of *degrees of freedom* significantly lesser than ambient dimensionality
Periodic signals

Bottling line

Toothbrush

Credit: Veeraraghavan et al, 2011
High-speed Camera

Nyquist Sampling of $x(t)$ – When each period of $x$ has high frequency variations, Nyquist sampling rate is high.

$$P = 10\text{ms}$$

$$T_s = \frac{1}{2f_{\text{Max}}}$$

Periodic signal has regularly spaced, sparse Fourier coefficients. Is it necessary to use a high-speed video camera? Why waste bandwidth?
Solving for the video

Camera observations at a pixel

\[
\begin{align*}
\text{Frame 1} & \quad \text{Frame Integration Period } T_s \\
\vdots & \quad \vdots \\
\text{Frame } M & \\
\end{align*}
\]

\[
y = Ax
\]

N unknowns

Coded Strobing

Frame Integration Period

\( T_s \)
Solving for the video

\[ \mathbf{x} = \mathbf{Fb} \]

Fourier Basis

Non-zero elements

Basis Coeff
Solving for the video

\[
\begin{pmatrix}
\text{y} \\
\end{pmatrix} = \begin{pmatrix}
\text{A} \\
\end{pmatrix}
\]
Solving for the video

\[
\begin{bmatrix}
\end{bmatrix}
= \begin{bmatrix}
\end{bmatrix}
\begin{bmatrix}
\end{bmatrix}
\end{align*}

\[
\begin{bmatrix}
\end{bmatrix}
= \begin{bmatrix}
\end{bmatrix}
\begin{bmatrix}
\end{bmatrix}
\end{align*}

\[
y = A F S
\]
Implementation

PGR Dragonfly2 (25 fps)

FLC Shutter
Can flutter at 250us
Toothbrush (simulation)

20fps normal camera

20fps coded strobing camera

Reconstructed frames

1000fps hi-speed camera
Mill Tool

Mill tool rotating at 50Hz

Normal Video: 25fps

Mill tool rotating at 50Hz

Coded Strobing Video: 25fps

Mill tool rotating at 50Hz

Reconstructed Video at 2000fps
Optical super-resolution

Key papers
Duarte et al., Single pixel camera, *SPM* 2008
Wang et al., *LiSens*, *ICCP* 2015
Chen et al., *FPA-CS*, *CVPR* 2015
Video sensing in infrared

- Sensing in infra-red has applications in night-vision, astronomy, microscopy, etc.
- Materials that sense in certain infrared bands are extremely costly
  - A 64 x 64 sensor costs upwards of USD 2000
  - 1 Megapixel sensor costs > USD 100k

Table 1. Approximate per-pixel price of detector elements in various spectral bands.

<table>
<thead>
<tr>
<th>Spectral band</th>
<th>Detector technology</th>
<th>Approx. per-pixel price ($/pix)</th>
</tr>
</thead>
<tbody>
<tr>
<td>mmW/THz</td>
<td>Multiple</td>
<td>$10^2-10^4$</td>
</tr>
<tr>
<td>LWIR</td>
<td>HgCdTe</td>
<td>$&lt;10^1$</td>
</tr>
<tr>
<td></td>
<td>Bolometer</td>
<td>$10^{-2}$</td>
</tr>
<tr>
<td></td>
<td>InSb/PbSe</td>
<td>$10^{-1}$</td>
</tr>
<tr>
<td>MWIR</td>
<td>InGaAs/PbSe/Si</td>
<td>$10^{-1}$</td>
</tr>
<tr>
<td>SWIR</td>
<td></td>
<td>$&lt;10^{-6}$</td>
</tr>
<tr>
<td>NIR/VIS/NUV</td>
<td>Si (thinned)</td>
<td>$&lt;10^{-3}$</td>
</tr>
<tr>
<td>MUV</td>
<td>Si-PIN/CdTe/Si (thinned)</td>
<td>$10^2-10^3$</td>
</tr>
<tr>
<td>EUV</td>
<td>Si-PIN/CdTe</td>
<td>$10^{-2}$</td>
</tr>
<tr>
<td>Soft-xray</td>
<td>Multiple</td>
<td>$10^2-10^4$</td>
</tr>
<tr>
<td>Hard-xray/gamma</td>
<td>Multiple</td>
<td></td>
</tr>
</tbody>
</table>
Can we super-resolve a low-resolution sensor?

- Spatial light modulation
  - Introduce a high-resolution mask between scene and sensor

![Diagram of optical system](image.png)

- Photo-detector
- Digital micro-mirror device
Single pixel camera

- Each pattern of micro-mirrors yield **ONE** compressive measurement

- A **single** photo-detector tuned to the wavelength of interest

- Resolution of the camera is that of the DMD, and not the sensor

*with Kelly lab, Rice University*
CS-MUVI on SPC

Single pixel camera setup
CS-MUVI: IR spectrum

**InGaAs Photo-detector** (Short-wave IR)

SPC sampling rate: 10,000 sample/s

Number of compressive measurements: $M = 16,384$

Recovered video: $N = 128 \times 128 \times 61$. Compression = 61x
Results

- Real data acquired using a single pixel camera
- Sampling rate: 10,000 Hz
- Number of compressive measurements: 65536
- Total duration of data acquisition: 6 seconds
- Reconstructed video resolution: 128x128x256

Final estimate (6 different videos)
Motivation

SPC has very **low measurement rate**

**DMD** --- $R_{DMD}$ **patterns/sec** (typically, in 10s kHz)

**ADC** --- $R_{ADC}$ **samples/sec** (typically, in 10s MHz)

Measurement rate of the SPC = $\min(R_{ADC}, R_{DMD})$
Parallel Compressive Imaging

- Use multiple pixels or a low-resolution sensor array

- How do we decide the specifications of the low-resolution sensor?
  - Number of pixels, geometry, etc...

Diagram:

- Objective lens
- Relay lens
- Low-resolution array
- Digital micromirror device (DMD)
- ADC
Measurement rate

Measurement rate vs. number of pixels, $F$.

- $R_{\text{ADC}}$ (ADC measurement rate)
- $R_{\text{DMD}}$ (DMD measurement rate)
- $F = 1$
- $F = 10^6$

**Conventional sensor**

$F = 1$

$F = 10^6$
Measurement rate

![Graph showing measurement rate](image)

- $R_{ADC}$
- $R_{DMD}$
- $F \cdot R_{DMD}$

- Conventional sensor

**# of pixels, $F$**

- $F = 1$
- $F = 10^6$
Optimal # of pixels

$F_{optimal} = \frac{R_{ADC}}{R_{DMD}}$

$F_{optimal}$ is typically in 1000s of pixel for today’s DMDs and ADCs.

Implications.
Measurement rate of a conventional sensor but with a fraction of the number of pixels! (less than 0.1% pixels)
Two Prototypes

- **Focal plane array-based CS (FPA-CS)**
  - **SWIR**
  - Map DMD onto a low-resolution 2D sensor
  - Each pixel on the sensor observes a 2D patch of micromirrors on the DMD

- **Line-sensor based Compressive Imager (LiSens)**
  - Map DMD onto a line-array sensor
  - Each pixel on the sensor observes a line of micromirrors on the DMD
FPA-CS Results

Scene (seen in a visible camera)

Image seen by 64x64 SWIR sensor

Super-resolved image by FPA-CS architecture

Chen et al. CVPR 2015
**Line-Sensor-based compressive camera (LiSens)**

1. Use a linear array of pixels (a line-sensor)
2. Add a cylindrical lens
Hardware prototype

Measurement rate: 1 MHz
Comparison against SPC

Capture duration: **880ms** 440ms 220ms 110ms

LiSens
0.8M meas.

SPC
16k meas.
Compressive Light Fields

Key Papers

• Marwah et al., *Compressive coded apertures*, SIGGRAPH 2013
• Tambe et al., *Compressive LF videos*, ICCV 2013
• Ito et al., *Compressive epsilon photography*, SIGGRAPH 2014
Epsilon Photography

Capture stack of photographs by varying camera parameters **incrementally**
Ex 1 - Epsilon photography applied to exposure

Exposure Bracketting for HDR

Slide credit: Verma and Mon-Ju. “High Dynamic Range Imaging”
Ex 2 – Epsilon photography applied to focus

Focus stack

Slide credit: dpreview.com
Ex 3 – Epsilon photography applied to aperture and focus

Confocal stereo
Per-pixel depth estimation

Hasinoff and Kutulakos, ECCV 2006
Confocal Stereo

Confocal stereo
Per-pixel depth estimation

Hasinoff and Kutulakos, ECCV 2006
Pros and Cons

• Pros
  – Per-pixel operations (for the most part)

• Cons
  – Too many images
  – Need texture (problem for everybody passive)
  – Align?
Epsilon Photography

Capture stack of photographs by varying camera parameters incrementally

Extremely slow!
### Compressive Epsilon Photography

<table>
<thead>
<tr>
<th>ft</th>
<th>1.5</th>
<th>1.75</th>
<th>2</th>
<th>2.5</th>
<th>3.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>m</td>
<td>0.44</td>
<td>0.5</td>
<td>0.6</td>
<td>0.8</td>
<td>1</td>
</tr>
</tbody>
</table>

**Varying Focus**
<table>
<thead>
<tr>
<th>ft</th>
<th>1.5</th>
<th>1.75</th>
<th>2</th>
<th>2.5</th>
<th>3.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>m</td>
<td>0.44</td>
<td>0.5</td>
<td>0.6</td>
<td>0.8</td>
<td>1</td>
</tr>
</tbody>
</table>

Compressive Epsilon Photography

Varying Focus
Redundancies in Focus-Aperture Stacks
Redundancies in Focus-Aperture Stacks

Approx. accuracy (in %)

Subspace dimension

cluster mean

Approx. accuracy (in %)

Subspace dimension

cluster mean
Per-pixel models

• Key idea: Model intensity variations observed at an *individual pixel*

• Advantages
  – No smoothing. *Spatial resolution can be preserved*
  – Parallel recovery at each pixel

• Disadvantages
  – Lack of spatial constraints
Gaussian Mixture Models

Observation: Structure of EP intensity profiles tied to depth at a pixel
Given a few images captured with pre-selected parameters

+ per-pixel GMM of intensity variations

recover the entire epsilon photography intensity profile at each pixel.

Linear inverse problem
Lots of solvers
We use a maximum likelihood estimator
Advantages of the GMM model

Analytical bounds on performance.

Can greedily pre-select camera parameters that maximize average reconstruction performance.
Advantages of the GMM model

Small aperture leads to large DOF and provides textural cues

Large aperture leads to small DOF and provides depth cues
Chess
Fluffy
Animals
CS Summary

• Three questions
  – Is sensing **costly**? (how?)
  – Is there a sparsifying/parsimonious representation?
  – Acquire some sort of randomized measurements?
A simple case study: MRI

MRI obtains samples in Fourier space.

Taking lesser samples

==

higher speed of operation, less time etc.

Lustig et al., 2008
MRI without a signal model

From 10 times lesser number of measurements

Lustig et al., 2008
MRI + CS

*with* signal model

From 10 times lesser number of measurements

The recovery is *exact*, provided some conditions are satisfied

Lustig et al., 2008
Summary

• CS provides the ability to sense from far-fewer measurements than the signal’s dimensionality

• Implications
  – Fewer pixels on the sensor
  – Shorter acquisition time
  – Slower rate of acquisition