Outline

1. What is the automatic cameraman
2. What is Machine Learning?
3. TAC Technologies
4. What about ML at UCSD?
1. What is the automatic cameraman
2. What is Machine Learning?
3. TAC Technologies
4. What about ML at UCSD?
About The Automatic Cameraman (TAC)

Hardware

- **What?** A touchless interactive display
- **How?** We are able to create an interactive experience by analyzing the audio and video streams.
Step #1: Roughly point camera at a sound source if present.
Step #2: Use a face detector to center the face in the field of view
Step #3: Display an onscreen button for the user to “press” to start a recording (and then a stop button afterwards).
Webpage for Recordings

- Webpage shows the movies that have been recorded.
- Greeting cards to your family and friends!
- Lots of new data!

**Video Recordings from Cameraman**

If you'd like a video of you to be removed please contact -- eettinger AT cs DOT ucsd DOT edu.

Next 20

<table>
<thead>
<tr>
<th>Movie</th>
<th>Date of Recording</th>
<th>Size</th>
<th>Log</th>
<th>Movie</th>
<th>Date of Recording</th>
<th>Size</th>
<th>Log</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.jpg" alt="Video 1" /></td>
<td>02/06/09 12:17:09</td>
<td>2.1 MB</td>
<td>1233951429.log</td>
<td><img src="image2.jpg" alt="Video 2" /></td>
<td>02/05/09 13:51:39</td>
<td>0.72 MB</td>
<td>1233870699.log</td>
</tr>
<tr>
<td><img src="image3.jpg" alt="Video 3" /></td>
<td>02/05/09 09:56:49</td>
<td>17.67 MB</td>
<td>1233856609.log</td>
<td><img src="image4.jpg" alt="Video 4" /></td>
<td>02/05/09 09:52:49</td>
<td>3.16 MB</td>
<td>1233856369.log</td>
</tr>
<tr>
<td><img src="image5.jpg" alt="Video 5" /></td>
<td>02/05/09 09:40:59</td>
<td>6.18 MB</td>
<td>1233855659.log</td>
<td><img src="image6.jpg" alt="Video 6" /></td>
<td>02/04/09 19:03:58</td>
<td>8.12 MB</td>
<td>1233803038.log</td>
</tr>
<tr>
<td><img src="image7.jpg" alt="Video 7" /></td>
<td>02/04/09 14:07:10</td>
<td>10.22 MB</td>
<td>1233785230.log</td>
<td><img src="image8.jpg" alt="Video 8" /></td>
<td>02/04/09 14:06:13</td>
<td>1.27 MB</td>
<td>1233785173.log</td>
</tr>
<tr>
<td><img src="image9.jpg" alt="Video 9" /></td>
<td>02/04/09 14:01:17</td>
<td>4.83 MB</td>
<td>1233784877.log</td>
<td><img src="image10.jpg" alt="Video 10" /></td>
<td>02/04/09 11:32:15</td>
<td>1.37 MB</td>
<td>1233775935.log</td>
</tr>
</tbody>
</table>
I will introduce you to the automatic cameraman (TAC) and a bit about machine learning.

Halfway through we will take groups up to the 4th floor to see the system in action.

During the second half of the class we will talk about the technology behind TAC

Your homework this week is based on interacting with TAC and thinking about how it could be used.
Audio localization - able to determine the location of sounds.
Face detection & tracking - can find a face in an image (video stream).
Skin detection - can find where skin regions are in an image.
Face recognition - whose face is it? (coming soon...)
Audio keyword spotting - recognize a limited audio vocabulary: good, bad, yes, no, left, right, up, down, etc. (coming soon...)
Ok...that’s cool. But what can this be used for?

Game Playing

Video Conferencing

Information Kiosks
World Problem: Assisted Living

- Grandma not ready for assisted living, but needs a safety net.
- Put TAC (or something like it) in her home to monitor and ensure her safety!
  - A central place to contact help/family hands free
  - Additional microphones placed throughout the home can track people and alert when something appears to be wrong.
  - A person can call for help anywhere in the home and the system will call 911 or OnStar like central help service.
- Grandma can still live independently.
On average, assisted living costs are $3500 per mo.

Approximately 40 million people currently get social security now, with 1 million more each year.

Assume we can get 1% of the market, or 10000 customers per year.

If we charge $10000 installation and $250 per mo.

$100 \text{ M} + 30 \text{ M} \times N$ after $N$ years (assuming no deaths)

Easily 250M a year business, more if larger market share.
What technologies do we need to make TAC work?

- How do you tell a computer what a face looks like?
- How do you tell it what skin color looks like?
- How does the computer know where sounds are coming from?

**Answer**: You train it on lots and lots of examples of faces, skin, and sounds. From this data the computer learns these “patterns”.

This is exactly what artificial intelligence (AI) and statistical machine learning are (sub-disciplines of computer science).
What technologies do we need to make TAC work?

- How do you tell a computer what a face looks like?
- How do you tell it what skin color looks like?
- How does the computer know where sounds are coming from?

**Answer:** You train it on lots and lots of examples of faces, skin, and sounds. From this data the computer learns these “patterns”.

- This is exactly what artificial intelligence (AI) and statistical machine learning are (sub-disciplines of computer science).
What technologies do we need to make TAC work?

How do you tell a computer what a face looks like?
How do you tell it what skin color looks like?
How does the computer know where sounds are coming from?

**Answer:** You train it on lots and lots of examples of faces, skin, and sounds. From this data the computer learns these “patterns”.

This is exactly what artificial intelligence (AI) and statistical machine learning are (sub-disciplines of computer science).
What technologies do we need to make TAC work?

How do you tell a computer what a face looks like?
How do you tell it what skin color looks like?
How does the computer know where sounds are coming from?

**Answer:** You train it on lots and lots of examples of faces, skin, and sounds. From this data the computer learns these “patterns”.

This is exactly what artificial intelligence (AI) and statistical machine learning are (sub-disciplines of computer science).
What technologies do we need to make TAC work?

- How do you tell a computer what a face looks like?
- How do you tell it what skin color looks like?
- How does the computer know where sounds are coming from?

**Answer:** You train it on lots and lots of examples of faces, skin, and sounds. From this data the computer learns these “patterns”.

- This is exactly what artificial intelligence (AI) and statistical machine learning are (sub-disciplines of computer science).
Outline

1. What is the automatic cameraman
2. What is Machine Learning?
3. TAC Technologies
4. What about ML at UCSD?
A Thought Experiment

Imagine explaining to a robot who is fresh off the assembly line what a human face looks like. How would you describe it?
Imagine explaining to a robot who is fresh off the assembly line what a human face looks like. How would you describe it?

- You: Well, a face typically has two eyes, a nose and a mouth. The eyes are above the nose, which is above the...
- Robot: Wait a minute...what do eyes look like?
- You: Oh of course. Eyes are round circular shapes that have a dark circle in the middle surrounding by a whitish color.
- Robot: Are they exactly circular?
- You: Well no, the shape varies from person to person.
- Robot: What is the tolerable variation in size and shape? How small can they be and how large? How far can they deviate from a circular shape?
- You: Oh boy...I give up. You better learn this on your own.
Imagine explaining to a robot who is fresh off the assembly line what a human face looks like. How would you describe it?

- **You:** Well, a face typically has two eyes, a nose and a mouth. The eyes are above the nose, which is above the...
- **Robot:** Wait a minute...what do eyes look like?
  - **You:** Oh of course. Eyes are round circular shapes that have a dark circle in the middle surrounding by a whitish color.
  - **Robot:** Are they exactly circular?
  - **You:** Well no, the shape varies from person to person.
  - **Robot:** What is the tolerable variation in size and shape? How small can they be and how large? How far can they deviate from a circular shape?
  - **You:** Oh boy...I give up. You better learn this on your own.
Imagine explaining to a robot who is fresh off the assembly line what a human face looks like. How would you describe it?

- **You:** Well, a face typically has two eyes, a nose and a mouth. The eyes are above the nose, which is above the...
- **Robot:** Wait a minute...what do eyes look like?
- **You:** Oh of course. Eyes are round circular shapes that have a dark circle in the middle surrounding by a whitish color.
- **Robot:** Are they exactly circular?
- **You:** Well no, the shape varies from person to person.
- **Robot:** What is the tolerable variation in size and shape? How small can they be and how large? How far can they deviate from a circular shape?
- **You:** Oh boy...I give up. You better learn this on your own.
Imagine explaining to a robot who is fresh off the assembly line what a human face looks like. How would you describe it?

**You:** Well, a face typically has two eyes, a nose and a mouth. The eyes are above the nose, which is above the...

**Robot:** Wait a minute...what do eyes look like?

**You:** Oh of course. Eyes are round circular shapes that have a dark circle in the middle surrounding by a whitish color.

**Robot:** Are they exactly circular?

**You:** Well no, the shape varies from person to person.

**Robot:** What is the tolerable variation in size and shape? How small can they be and how large? How far can they deviate from a circular shape?

**You:** Oh boy...I give up. You better learn this on your own.
Imagine explaining to a robot who is fresh off the assembly line what a human face looks like. How would you describe it?

- **You:** Well, a face typically has two eyes, a nose and a mouth. The eyes are above the nose, which is above the...
- **Robot:** Wait a minute...what do eyes look like?
- **You:** Oh of course. Eyes are round circular shapes that have a dark circle in the middle surrounding by a whitish color.
- **Robot:** Are they exactly circular?
- **You:** Well no, the shape varies from person to person.
- **Robot:** What is the tolerable variation in size and shape? How small can they be and how large? How far can they deviate from a circular shape?
- **You:** Oh boy...I give up. You better learn this on your own.
Imagine explaining to a robot who is fresh off the assembly line what a human face looks like. How would you describe it?

- You: Well, a face typically has two eyes, a nose and a mouth. The eyes are above the nose, which is above the...
- Robot: Wait a minute...what do eyes look like?
- You: Oh of course. Eyes are round circular shapes that have a dark circle in the middle surrounding by a whitish color.
- Robot: Are they exactly circular?
- You: Well no, the shape varies from person to person.
- Robot: What is the tolerable variation in size and shape? How small can they be and how large? How far can they deviate from a circular shape?
- You: Oh boy...I give up. You better learn this on your own.
Imagine explaining to a robot who is fresh off the assembly line what a human face looks like. How would you describe it?

- **You:** Well, a face typically has two eyes, a nose and a mouth. The eyes are above the nose, which is above the...
- **Robot:** Wait a minute...what do eyes look like?
- **You:** Oh of course. Eyes are round circular shapes that have a dark circle in the middle surrounding by a whitish color.
- **Robot:** Are they exactly circular?
- **You:** Well no, the shape varies from person to person.
- **Robot:** What is the tolerable variation in size and shape? How small can they be and how large? How far can they deviate from a circular shape?
- **You:** Oh boy...I give up. You better learn this on your own.
Some problems we’d like to solve are very hard to solve well by hand.

Machine learning attempts to devise ways in which computers can learn how to solve these kinds of problems on their own.

How do they learn? By giving them many many examples of inputs and what your desired output is.

For example, to teach the robot what a face looks like you’d give it many many examples of faces and non-faces and it would learn what a face looks like on its own.

Then, hopefully it could identify new faces well in the real world.
What is Machine Learning?

- Purpose: To find useful patterns in data.
- What do you mean by “data”? 

Audio

Image/Video Biology Data

Text

Image/Video

Biology Data
Machine Learning (ML) is typically broken down into 3 types of problems:

- **Supervised Learning** - Predict some “label” associated with each datum.
  
  - e.g. E-mail spam. Given an e-mail, decide whether it is spam or not.

- **Unsupervised Learning** - Cluster the data into “similar” groups.
  
  - e.g. Animal image clustering. Pictures of dogs are closer to each other than a dog and a bee. No labels are needed to identify this.

- **Semi-supervised learning** - Some data has a label, some data comes unlabeled.
  
  - e.g. Given some e-mails that are labeled as spam/not spam, and also many many other e-mails without labels.

Many other ways to dice up ML - active learning, reinforcement learning, online learning, etc. but the above are the big 3.
Let’s look closer out our e-mail spam example...

- Somebody hands us a set of e-mails and tells us which ones are spam, and which ones are not.

**SPAM**

- ED pills $1.10-$1.50/each, PharmacyChecker Certified. Tens of Thousands of Happy Customers! gapp 99
- Vicodin, Hydrocodone, Phentermin, Codene, NoPreSCRIPTION + UP TO 80% OFF. Your #1 source for buying Vicodin online at a fraction of U.S. prices. Order Now! gex mtv. Stem ER
- Paying Too Much for HardMeds? Let’s Start SAVING now!

**NOT SPAM**

- Evan Ettinger (UCSD)
- Jessica Bagley (University Programs  | Deepa X)
- To "Georgia Vanghose" <vanghose@ucsd.edu>
Spam: Feature Extraction

- We extract numerical features from the e-mails e.g. How many dollar signs in the e-mail? Does it come from a reputable sender? How many times are ED terms mentioned? etc.
- After extraction, the e-mail below becomes just some numbers with its human given label: (3, 0, 4, SPAM)
We feed 100s/1000s/millions of labeled, feature extracted, examples into a learning algorithm which tries to find patterns among the numerical features.

The learning algorithm outputs a decision rule that we can use on new e-mails to predict whether they are spam or not.

The final decision rule tries to be as good a predictor of spam as possible i.e. minimize false-negatives and false-positives.
So what...are there really any important problems that ML can help with?

- Computer vision - face detection, person tracking, supermarket items, hand-written digits, cancerous cells/legions, etc.
- Speech recognition
- Machine translation
- Disease prediction from DNA, fMRI, or other medical tests.
- Search engines
- Spam and document classification
- Game playing
- Recommendation systems - amazon, online dating, netflix, music, etc.
- Ad placement
- Credit fraud detection
- Stock market analysis
- Skynet
- ...
Are there any company’s that care?
1. What is the automatic cameraman
2. What is Machine Learning?
3. TAC Technologies
4. What about ML at UCSD?
Arrival Delay

- A fixed location sound source will have an arrival delay for each pair of mics
An audio signal is just a series of numbers that sample at regular intervals how much sound pressure is being applied to the mic.

We can estimate the delay by trying to find the best alignment of two audio signals.
When you can measure multiple delays...

- Source locations correspond to the intersection of multiple curves (actually hyperbolas) when several delays are known.
- System of nonlinear eqns to solve... much like triangulation! Actually called multilateration.
What about face detection?

An image to the computer is just a series of numbers (like audio!). Image sizes are described in pixel units. Each pixel can be a single color e.g. 640x480 pixels. To fully describe the color of a pixel, how much red, green and blue (RGB) are needed to create the color is required. So an image is just a series of triples (R,G,B) where each is a number between 0 and 255.
How do you extract features from an image?

- Take a small square “patch” of the image. We want to know whether a face appears in the patch.
- We must do “Feature Extraction” from this patch, e.g. What colors appear in this patch? How much blue, red, etc.? What types of edges are there?
- An edge is just a large variation in color among neighboring pixels!
How do you extract features from an image?

- Take a small square “patch” of the image. We want to know whether a face appears in the patch.
- We must do “Feature Extraction” from this patch, e.g. What colors appear in this patch? How much blue, red, etc.? What types of edges are there?
- An edge is just a large variation in color among neighboring pixels!
Feature Extraction (cont.)

Feature Extraction

(100,0,10, SPAM)
(0,1,0, NOT SPAM)
(10,1,100, SPAM)

etc.

SPAM NOT
SPAM

If # of ! > 10 AND viagra > 2
OUTPUT SPAM
Otherwise
OUTPUT NOT SPAM

Detection Mask for Face Detection

HOG features
= 16 * 8 = 128

Skin color features
= 16 * 2 = 32

Total number of features
= 128 + 32 = 160
We use a learning algorithm called “AdaBoost” which gives a score to each patch...NEG = is a face, POS = not a face.
Learning Algorithm - Stage 1

We use a learning algorithm called “AdaBoost” which gives a score to each patch...NEG = is a face, POS = not a face.
Learning Algorithm - Stage 1

We use a learning algorithm called “AdaBoost” which gives a score to each patch...NEG = is a face, POS = not a face.
Once we learn a face detector, we deploy it by sliding windows of various sizes across the entire image.

Approx. 300000 different locations/sizes to examine in 640x480 image!

But, once you've found a face in a particular frame do you really need to scan the entire image in the next time?

No! The face probably didn't go too far. Tracking allows us to identify the most likely places a face could be next.

Reduces the number of locations down to about 10K or 30x speed up.
Outline

1. What is the automatic cameraman
2. What is Machine Learning?
3. TAC Technologies
4. What about ML at UCSD?
Cool ML people at UCSD

CSE ML People

- Yoav Freund - Learning theory, biology, TAC.
- Sanjoy Dasgupta - Learning theory.
- Charles Elkan - Data mining, biology, text mining.
- Lawrence Saul - Data visualization, audio, ML applications.
- Serge Belongie - Computer vision.
- David Kriegman - Computer vision.
- Pavel Pevzner - Bioinformatics.
- Vineet Bafna - Bioinformatics.
ML courses you can take!

ML CSE Courses

- CSE 151 - intro to AI
- CSE 152 - intro to computer vision
- CSE 182 - biological databases
- CSE 190 - cognitive modeling
- And grad courses like...
  - CSE 250A/B - artificial intelligence series
  - CSE 254 - machine learning
- Also statistics courses are highly related!
TAC is an interactive display that YOU can play with on the 4th floor.
Based on ideas from Machine Learning.
ML is an area of applied statistics, with many applications...
and many interesting job opportunities!
Thanks! Questions?