

# Large Scale Image Annotation on Amazon Mechanical Turk (250,000 HITs = \$5,000)

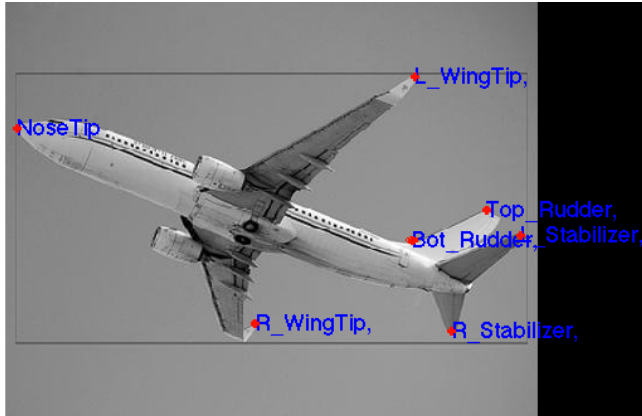
Subhransu Maji

[smaji@cs.berkeley.edu](mailto:smaji@cs.berkeley.edu)

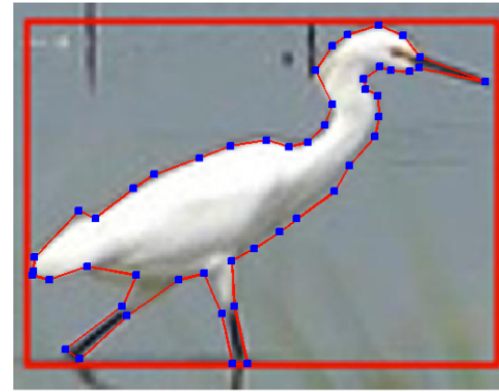
Computer Science Department  
University of California, Berkeley

# Things we did

## Marking Keypoints



## Object Boundaries



## Attributes of People

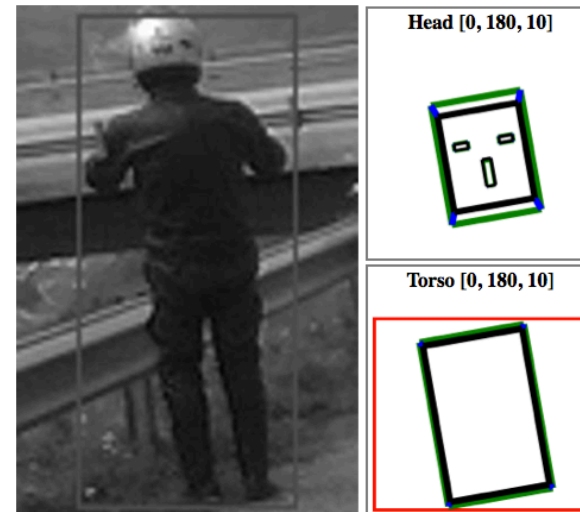


- ☐ shorts
- ☐ skirt
- ☐ jeans
- ☐ pants
- ☐ other

- ☐ shorts
- ☐ skirt
- ☐ jeans
- ☐ pants
- ☐ other

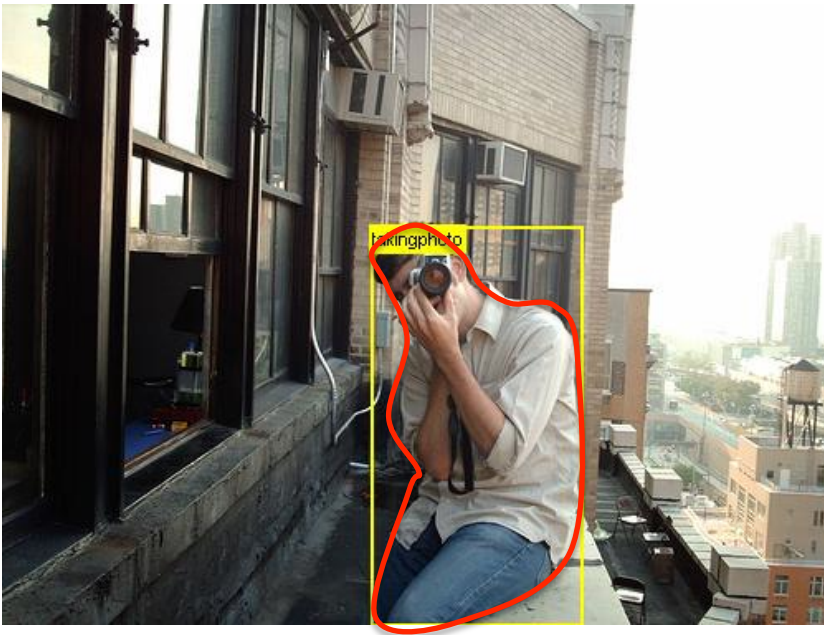
- ☐ shorts
- ☐ skirt
- ☐ jeans
- ☐ pants
- ☐ other

## 3d Pose



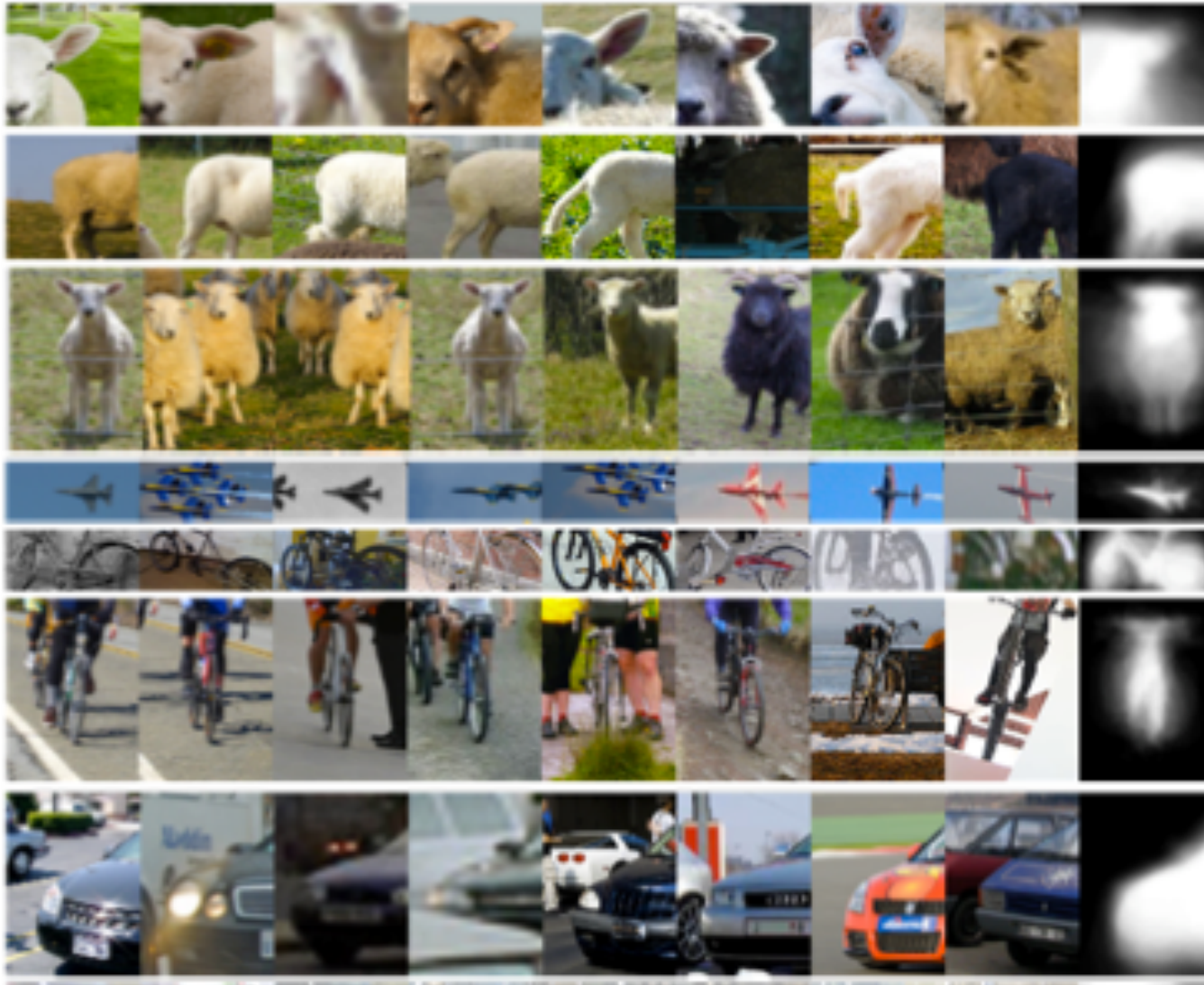
# Why did we do them?

- Describing People : detection, segmentation, pose, action, and other attributes



A man is **taking a photo**, wearing **jeans** and a **short sleeve shirt** and his head and body are **tilted slightly to the right**

# Poselets : Supervised Part Detectors



*Bourdev & Malik, 09*

*Bourdev, Maji, Brox & Malik, 10*



# Example Results (Submitted to CVPR'11)



Segmentation

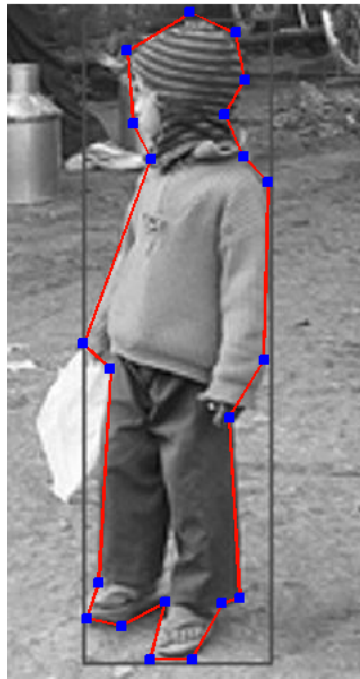


Figure 13. Results for **has-long-sleeves**. Top row: highest-ranked test examples. Bottom row: lowest-ranking.



3D pose estimation

# “X”vised



## Attributes

*Long Pants*

*Full Sleeves*

*Hat*

*Male*

*Wearing Sandals*

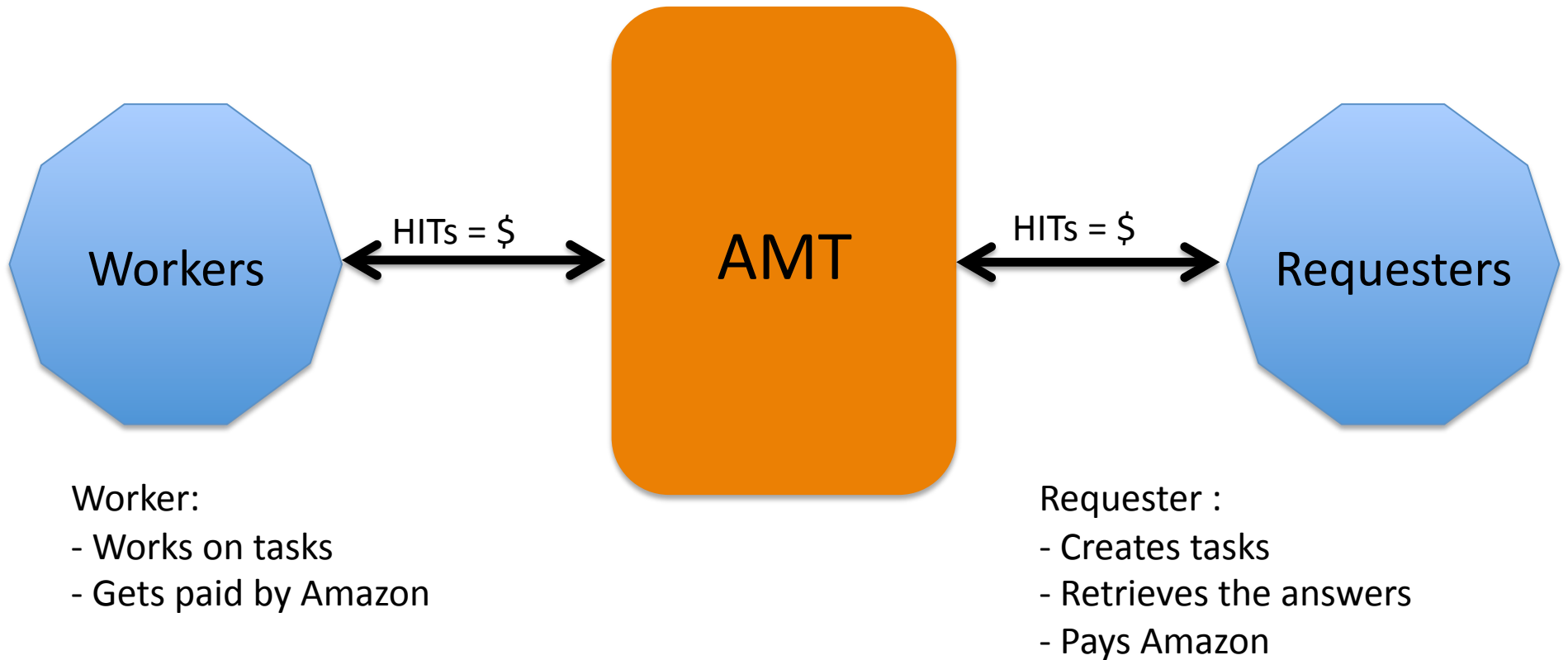
## Pose (yaw)

Head : -90 degrees

Torso : -80 degrees



# Amazon Mechanical Turk



# HIT : Human Intelligence Task

- “An Amazon Mechanical Turk HIT or Human Intelligence Task is the task you ask a Worker to complete. It may be a task that is inherently difficult for a computer to do.” – *AMT Best Practices Guide*

## All HITs

1-10 of 2165 Results

Sort by: 

Reward Amount (most first)

GO!

Show all details | 

Hide all details

Signup and Activate & Report Experience

Requester: PPD Inc.

HIT Expiration Date: Oct 16, 2010 (6 days 23 hours)

Reward: \$15.00

Time Allotted: 40 minutes

HITs Available: 1

FULL LENGTH Usability Test on www.tictocwatches.co.uk #215

Requester: TryMyUi Administrator

HIT Expiration Date: Oct 29, 2010 (2 weeks 5 days)

Reward: \$10.00

Time Allotted: 2 hours 43 minutes

HITs Available: 1

Take Q

Data Entry. Quick data entry verification.

Requester: Karen Burnow

HIT Expiration Date: Oct 10, 2010 (19 hours 44 minutes)

Reward: \$6.00

Time Allotted: 60 minutes

HITs Available: 1

Help test our Credit Report system and make \$5! Takes 2 minutes, costs you NOTHING.

Requester: Laurence Crawford

HIT Expiration Date: Nov 22, 2010 (6 weeks 2 days)

Reward: \$5.00

Time Allotted: 10 minutes

HITs Available: 1

Health System Recognition Survey

Requester: Deborah L Moore

HIT Expiration Date: Oct 21, 2010 (1 week 5 days)

Reward: \$4.50

Time Allotted: 30 minutes

HITs Available: 1

Take Q

Answer a survey about yourself

Requester: otilia

HIT Expiration Date: Oct 16, 2010 (6 days 23 hours)

Reward: \$4.00

Time Allotted: 1 hour 30 minutes

HITs Available: 1

UK English Voice Mail Messages

Requester: Karl Brown

HIT Expiration Date: Oct 14, 2010 (5 days 4 hours)

Reward: \$4.00

Time Allotted: 3 hours

HITs Available: 1

Not Qualified to w

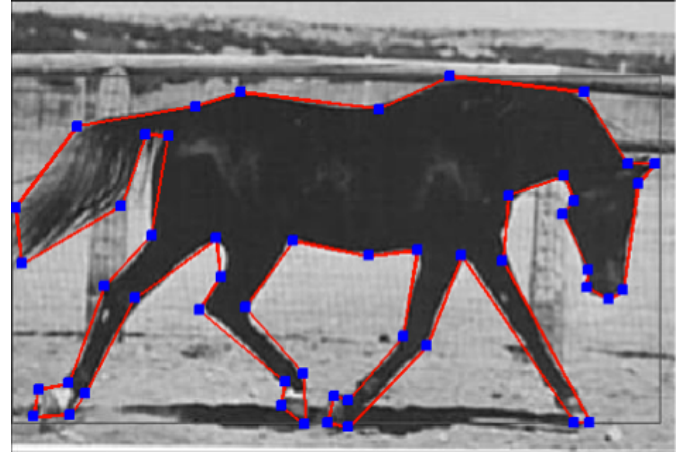
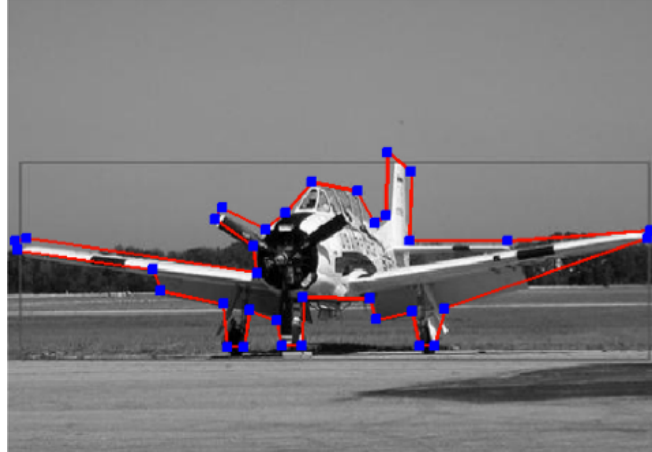


# Creating HITs

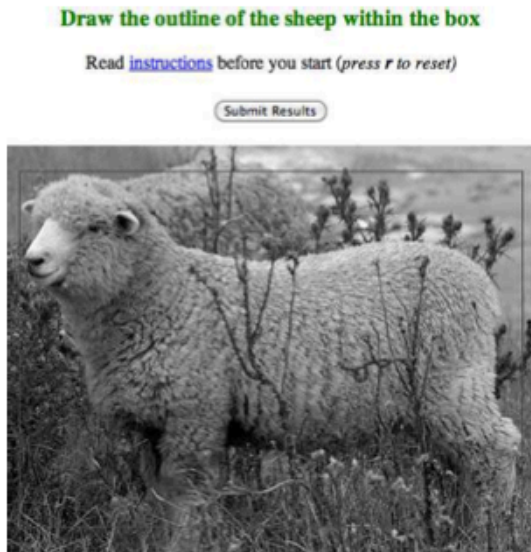
- Simple HITs
  - Surveys : same question
  - Simple text based : classify urls
  - Create HITs using Amazon's UI  
<http://docs.amazonwebservices.com/AWSMechTurk/latest/RequesterUI/>
- General HITs
  - GUI based : Image segmentation, etc
  - External HITS
- API
  - Java/C#/Perl/Ruby API
  - Command Line Interface

# Example : Image Segmentation

- Draw the boundaries of the object
- 20 categories, 23,000 images



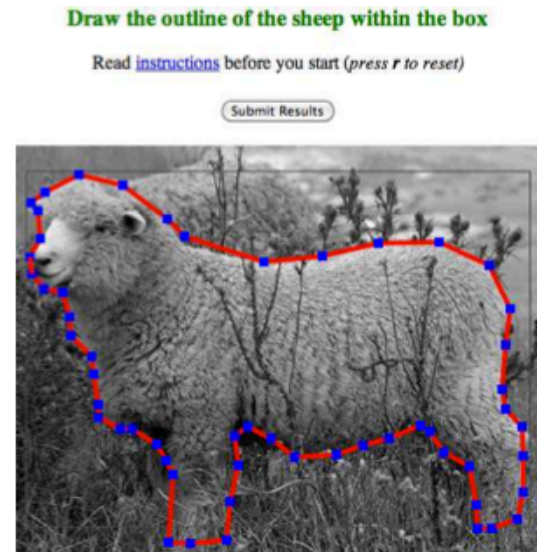
# HIT : User Interface



(a)



(b)



(c)

[http://www.cs.berkeley.edu/~smaji/mturk/segmentation/sheep.html?category-image=sheep,2010\\_000071\\_1.jpg](http://www.cs.berkeley.edu/~smaji/mturk/segmentation/sheep.html?category-image=sheep,2010_000071_1.jpg)

# HIT : User Interface

All HITS | HITS Available To You | HITS Assigned To You

Search for **HITS** containing  that pay at least \$ **0.00** for which you are qualified

**Timer:** 00:00:00 of 10 minutes

Want to work on this HIT?

Want to see other HITS?

**Total Earned:** \$1.75  
**Total HITS Submitted:** 116

Outline the boundary of the sheep

**Requester:** Subhransu Maji

**Reward:** \$0.02 per HIT

**HITS Available:** 0

**Duration:** 10 minutes

**Qualifications Required:** HIT approval rate (%) is greater than 75

**Draw the outline of the sheep within the box**

Read [instructions](#) before you start (press **r** to reset)

You must ACCEPT the HIT before you can submit the results.



# HIT : Instructions

## Example boundaries for sheep

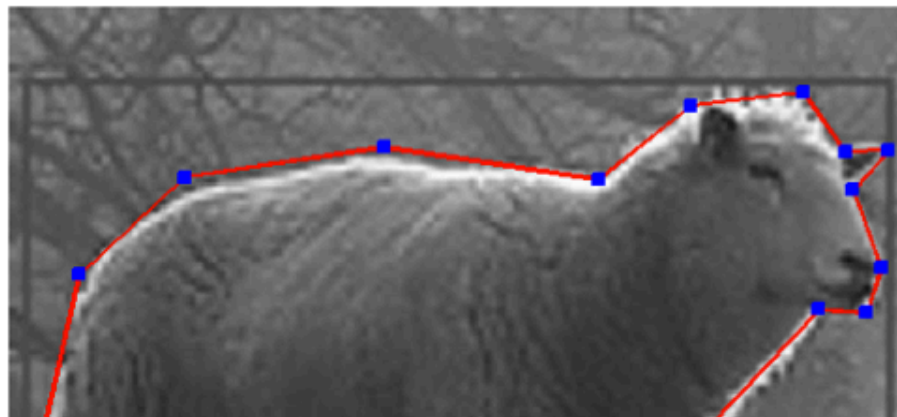
### Task Guidelines

Draw the outer boundary of the object within the gray box as accurately as possible.

### Interface instructions

- Remember to accept the hit before you proceed!
- Start by clicking somewhere and extend the polygon by clicking on the next point.
- You must close the polygon by clicking on the first point. The first point becomes larger when you come close.
- Once you close the polygon, you can move the control points to adjust the boundary.
- Press *r* to reset the annotation at any point.
- Press the 'submit button' at the top of the page to submit your results.

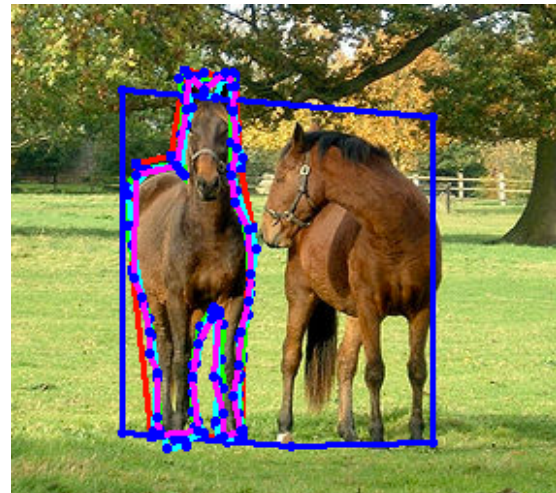
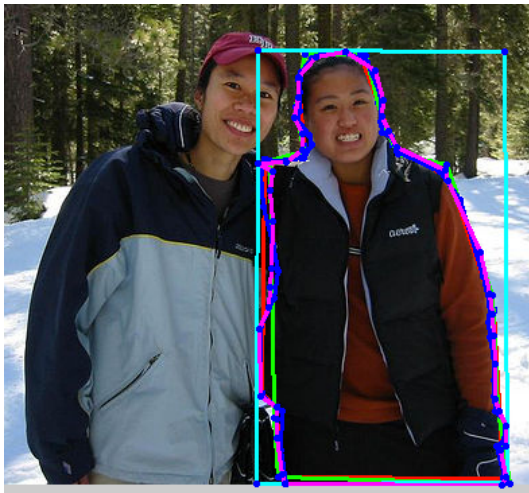
### Example Annotations



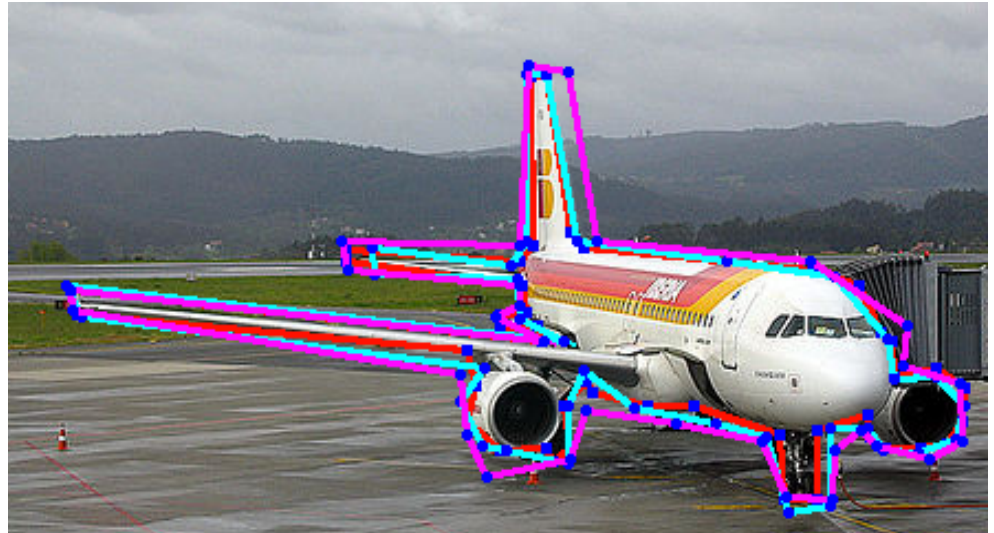
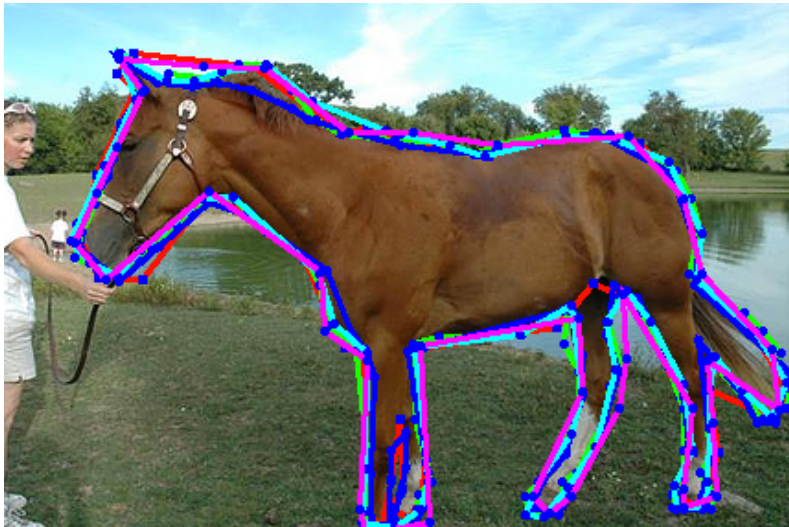


# HIT : Approving/Rejecting

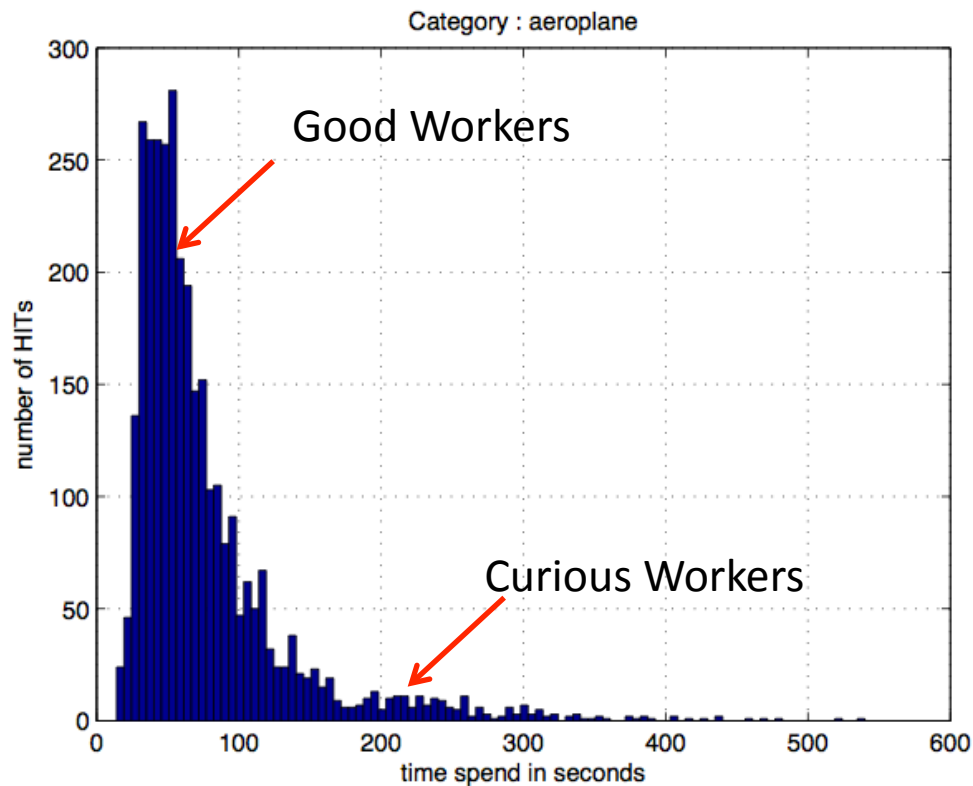
- Too many HITs to manually verify
- Automatic verification
  - Have each image done by 5 different people
  - Reject outliers



# Example Segmentations



# Some Statistics

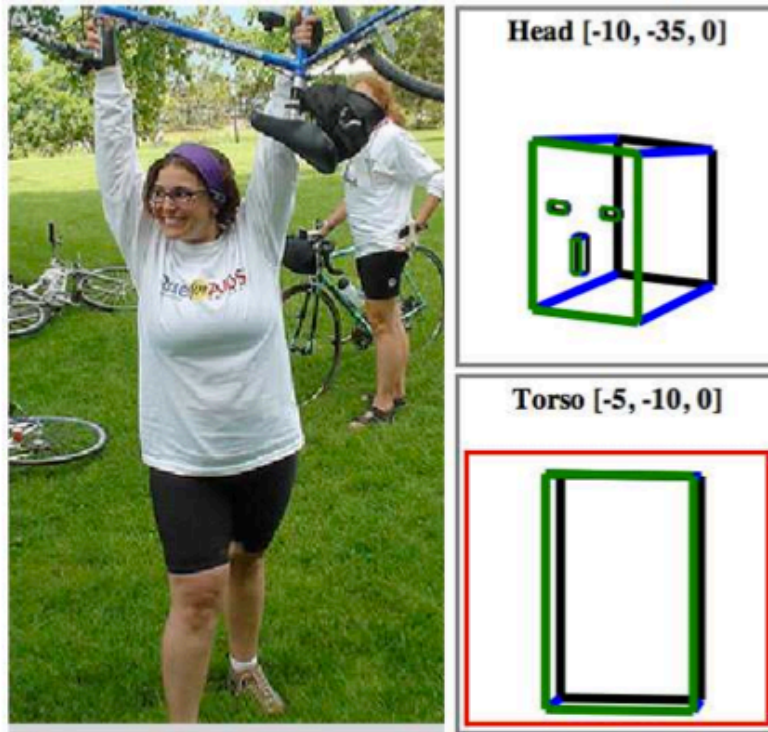


Average Submit Time : 77s (59s)  
Reward : 2 cents per annotation  
750 X 5 HITs ~ 3-4 hours

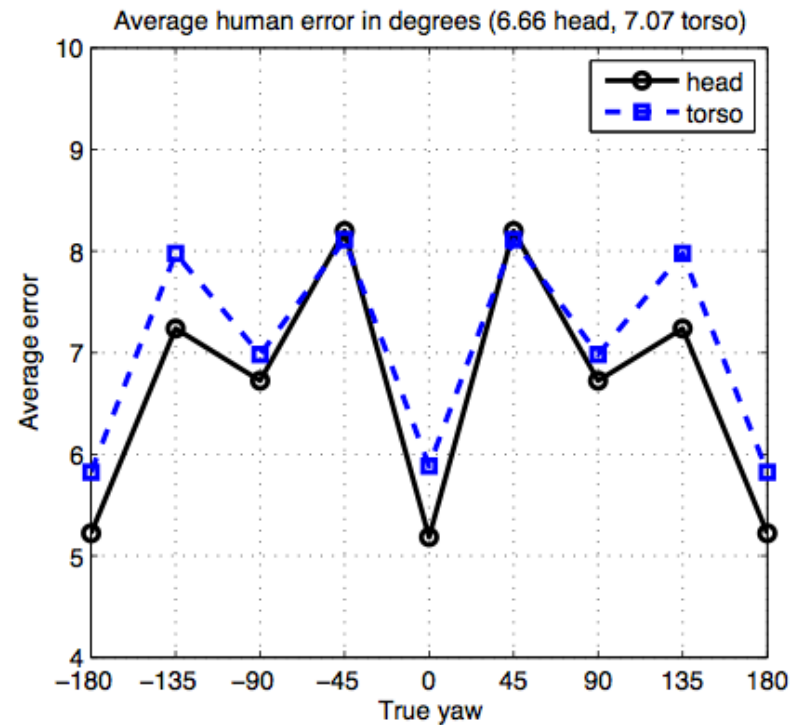
Roughly : \$1/hr



# “Bayes error” of the task



(a) AMT interface



(b) Human Error

# “Bayes error” of the task

**Which animal is it?**

*Identify which animal each of these patches are taken from*  
*Click on "Submit Results" button to check your score!*

1.  1. ☐ cat ☐ cow ☐ dog ☐ horse ☐ sheep

---

2.  2. ☐ cat ☐ cow ☐ dog ☐ horse ☐ sheep

---

3.  3. ☐ cat ☐ cow ☐ dog ☐ horse ☐ sheep

---

4.  4. ☐ cat ☐ cow ☐ dog ☐ horse ☐ sheep

---

5.  5. ☐ cat ☐ cow ☐ dog ☐ horse ☐ sheep

---

## Low level texture classification

Human performance is barely above chance!

Computer vision algorithms are probably better.



# “Bayes error” of the task

**Which animal is it?**

*Identify which animal each of these patches are taken from*  
*Click on "Submit Results" button to check your score!*

1.  1. ☐ cat ☐ cow ☐ dog ☐ horse ☐ sheep

---

2.  2. ☐ cat ☐ cow ☐ dog ☐ horse ☐ sheep

---

3.  3. ☐ cat ☐ cow ☐ dog ☐ horse ☐ sheep

---

4.  4. ☐ cat ☐ cow ☐ dog ☐ horse ☐ sheep

---

5.  5. ☐ cat ☐ cow ☐ dog ☐ horse ☐ sheep

---

## Low level texture classification

Human performance is barely above chance!

Computer vision algorithms are probably better.

**Can we do psychophysics?**

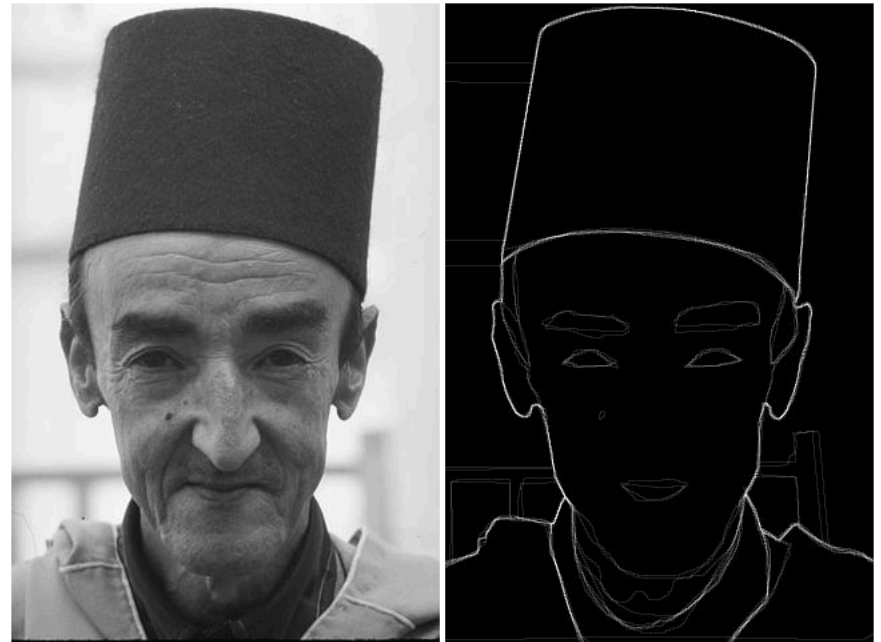
***Color spaces :***

***-metric learning with user constraints***

**Run social experiments?**

# Training vs. Testing

- Berkeley Segmentation Dataset (BSDS)
- 12,000 hand-labeled segmentations of 1,000 Corel dataset images from 30 human subjects.



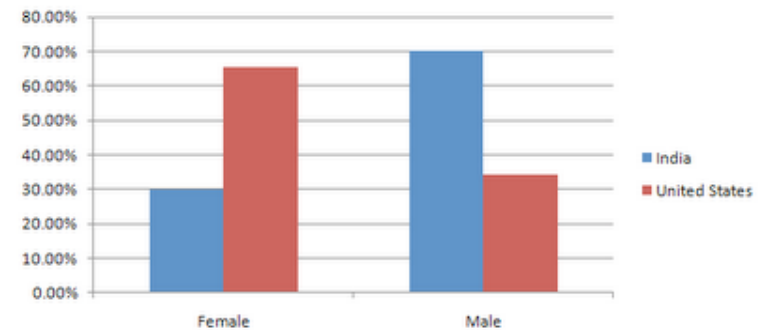
# Summary

- Three basic ingredients
  - **Interface**
    - Amazon UI builder/Internal HITS
    - External HITS
  - **Instructions**
    - Qualifications
    - Qualification Test
  - **Verification**
    - Outlier rejection
    - Secondary HITs to verify

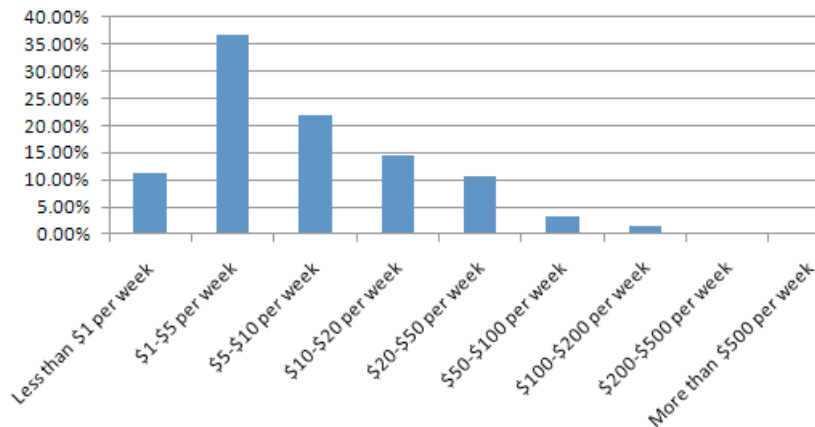
# Worker Demographics

- <http://behind-the-enemy-lines.blogspot.com/2010/03/new-demographics-of-mechanical-turk.html>
- By country:
  - United States: 46.80%
  - India: 34.00%
  - Miscellaneous: 19.20%

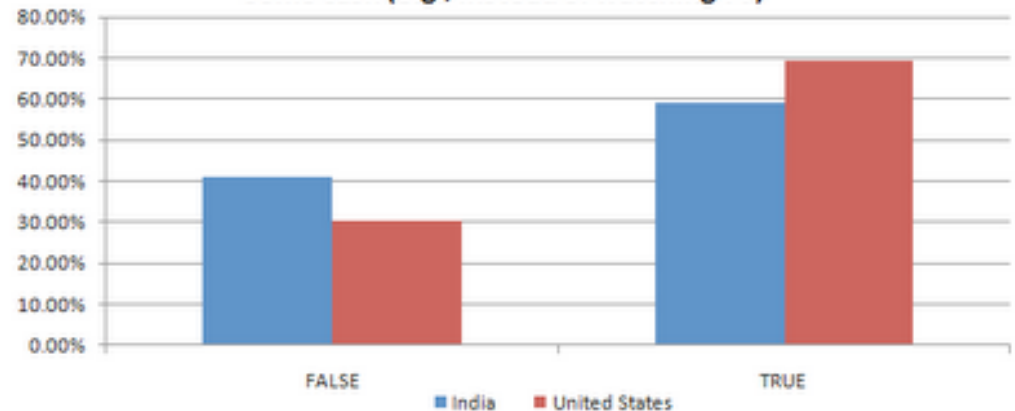
Gender Breakdown



Weekly Income from Mechanical Turk



Mechanical Turk is a fruitful way to spend free time and get some cash (e.g., instead of watching TV)



# Happy workers

- “thanks ... pls provide me the similar kind of task to perform please!”
- Thank you for all the paid work,  
I would love to continue working on this hits.
- thank you!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
- we like your work very much..it's very interesting  
and we enjoyed it..pl give us more hits,please....
- Thank you , Sir! I enjoy drawing the Saint Cows of  
India! I can draw all :p



# Questions?

- Resources :
  - [www.mturk.com](http://www.mturk.com)
  - <http://aws.amazon.com/mturk/>
    - sample code/user guide/tutorials/etc
  - [www.cs.berkeley.edu/~smaji/mturk](http://www.cs.berkeley.edu/~smaji/mturk)
  - Alex Sorokin homepage:  
<http://vision.cs.uiuc.edu/~sorokin2/>
  - Canvas 2d <http://dev.w3.org/html5/canvas-api/canvas-2d-api.html>