CSE 527: Introduction to Computer Vision

# System for Automatic Development of a Visual Attribute Vocabulary

Presented By: Ankush Gulati Prasad Prabhu Goal of the Project

To develop an automated system for developing a **vocabulary of visual attributes** of objects and scenes using **human responses**.

# Why Human Responses?

• To get more intuitive information than just the low level features.



#### Similar histogram distribution but different image contents

Source: http://www.cs.columbia.edu/CAVE/projects/histograms/

# Challenges in using Human Responses

• IRRELEVANT, WRONG or PARTIALLY ACCURATE tags.



TAGS: X100, Street Photography, Brooklyn, NYC, New York City Courtesy: Flickr

#### Work Breakup

- **1. Setup** a mechanism to get the images tagged by humans.
- **2.** Build a database of images.
- 3. Clean noisy data.
- 4. Train the dataset.
- **5. Build** computer vision routines for training as well as testing images.

### Work Breakup

- 1. Setup a mechanism to get the images tagged by humans.
- 2. Build a database of images.
- 3. Clean noisy data.
- 4. Train the dataset.
- 5. Build computer vision routines for training as well as testing images.

#### Human Response Accumulation System

• Gather ample human responses for the image collection to create a sufficiently large training data set.



### Human Response Accumulation System

- An automated system that allows us to feed images and returns tagged images.
  - Create simple Human Intelligence Task(H.I.T), a template that accepts tags for images at Amazon Mechanical Turk.
  - Automated script to build as many HIT as required and run it over Turk.

## Amazon's Mechanical Turk



### Amazon's Mechanical Turk (contd.)



Load HITs (Human Intelligence Tasks) on MTurk

Users complete tasks & Earn Money

You get Quality Human Responses

### Work Breakup

- **1. Setup** a mechanism to get the images tagged by humans.
- **2.** Build a database of images.
- 3. Clean noisy data.
- 4. Train the dataset.
- **5. Build** computer vision routines for training as well as testing images.

# Specialized Image Crawler

 Build a comprehensive database of images with varied objects & scenes.



# Snapshot of a Sample HIT

#### Instructions

Previous

Not

- 1. You will be provided an image
- 2. Interpret the contents of the image as you may find it to be.
- Then, provide appropriate tags to respective images. "You have to provide 3 tags in the space provided as a text box. This is necessary for your job to be accepted.
- 4. So in all 3 tags per image. At the end of tagging current image, click "next" (at the bottom-left of the page) to get the next image. You may move backwards to change your answer before you press the final SUBMIT. However after clicking submit your results will be final and it will be sent to Turk for approval.



#### Tag words should include proper values

- Should not contain any special characters 18 # \$1%\*&\*()\*=+[)\*:..7()|\*\*:<>? For e.g. Begal tags are birdSwings, postal-lefter, etc.
- Spaces within words or trailing are not allowed. Use single words.
  For e.g. "helio workd" (space in middle) or "helioworld "(trailing spaces) are (legal tags.

Image # 1./2

# **Implementation Details**

- API Used:
  - Amazon Mechanical Turk API
- Language Used:
  - Python
  - HTML
  - JavaScript
- Database
  - MYSQL 5

Entire source code, HIT templates, installation details & tutorials available online:

http://recognition.cs.stonybrook.edu/~ykou/website/



#### Future Work

- 1. Setup a mechanism to get the images tagged by humans.
- 2. Build a database of images.
- 3. Clean noisy data.
- 4. Train the dataset.
- 5. Build computer vision routines for training as well as testing images.



Thank you, note, pin, paper