Ubiquitous and Mobile Computing
CS 528:
TagSense: A Smartphone-based Approach to Automatic Image Tagging

Bo Peng

Computer Science Dept.
Worcester Polytechnic Institute (WPI)
Introduction

- What is image tagging? (Facebook)

- Face Recognition
Introduction (cont’d)

- Any problems?
  - Pictures and videos are exploded
  - Online content warehouses
  - Difficult to search and browse

- Any solutions?
  - Multi-dimensional and out-of-band sensing
  - Main idea?
Main Idea

- Sketch flow of TagSense:

- When - Where - Who - what
Scope of TagSense

- Not a complete solution
- AT LEAST one of the sensing dimensions
- Electronic footprint required! (Image of objects, animals, people without phones, oops...)

Comparison with Face Recognition

- Complementary!!

<table>
<thead>
<tr>
<th></th>
<th>Face Recognition</th>
<th>TagSense</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lighting surrounded</td>
<td>Good lighting</td>
<td>Bad lighting</td>
</tr>
<tr>
<td>Physical features</td>
<td>Yes (curious about twins)</td>
<td>Not really</td>
</tr>
</tbody>
</table>
System Overview

- Camera phone triggers sensing in participants
- Gathers the sensed information
- Determine who is in the picture
Who are in the picture

- Accelerometer based motion signature
  - Move into a specific posture in preparation
  - Stay still during the picture-click
  - Move again to normal behavior
Who are in the picture (cont’d)

- Complementary compass directions
  - Poses do not reflect on accelerometer
  - Solve the problem
    - Assumption: roughly face the direction of the camera
    - personal compass offset (PCO)

\[
UserFacing = (CameraAngle + 180) \mod 360
\]

\[
PCO = ((UserFacing + 360) - CompassAngle) \mod 360
\]
Who are in the picture (cont’d)

- Complementary compass directions
  - Does it work? (50 pictures, all facing the camera)

- Does not work 😞
Who are in the picture (cont’d)

- Complementary compass directions
  - Recalibrating the PCO

- Alice is posing, computing the PCO
- Alice is changing the direction of the phone
- Alice is posing, compute a new PCO
Who are in the picture (cont’d)

- Motion correlation across visual and accelerometer/compass
  - When clicking, several snapshots following
  - Motion vector
  - Optical flow (Matlab, detect direction and velocity)
Who are in the picture (cont’d)

- Defects
  - Can not pinpoint people in a picture
  - Can not identify kids (No phones!)
  - Compass based method assumes people are facing the camera
What are they doing

- Accelerometer
  - Standing, Sitting, Walking, Jumping, Biking, Playing
- Acoustic
  - Talking, Music, Silence
Where is the picture taken

- Indoor? Outdoor?
  - Variation of light intensity measured 400 different times
Performance

- Tagging people
Performance (cont’d)

- Tagging people
Performance (cont’d)

- Tagging activities and context
  - Assessment by human

\[
\text{precision} = \frac{|\text{Tags by Humans} \cap \text{Tags by TagSense}|}{|\text{Tags by TagSense}|}
\]

\[
\text{recall} = \frac{|\text{Tags by Humans} \cap \text{Tags by TagSense}|}{|\text{Tags by Humans}|}
\]

![Bar chart showing average precision and recall for different reviewer IDs.](image)
Performance (cont’d)

- Tagging based image search (200 pictures)
  - Volunteer look at 20 pictures and come up with query string
Future of TagSense

- Smartphones are becoming context-aware with personal sensing
- Smartphones may have directional antennas
- The granularity of localization will approach a foot
- Smartphones are replacing point and shoot cameras
Related Work

- ContextCam
  - Wear a device... (Not practical)

- SensingCam
References


Reference (cont’d)


Reference (cont’d)


