

# Ubiquitous and Mobile Computing

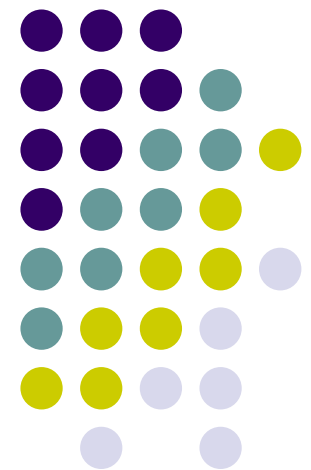
CS 528:

## *TagSense: A Smartphone-based Approach to Automatic Image Tagging*

Bo Peng

*Computer Science Dept.*

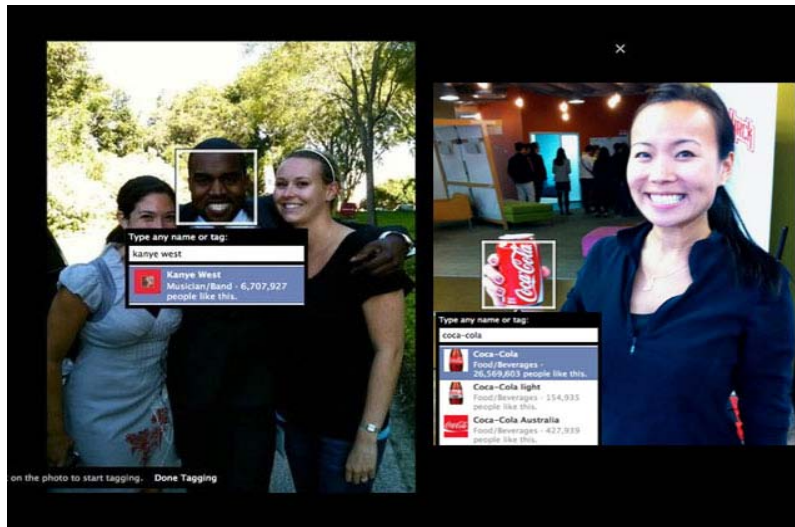
*Worcester Polytechnic Institute (WPI)*





# Introduction

- What is image tagging? (Facebook)



- Face Recognition



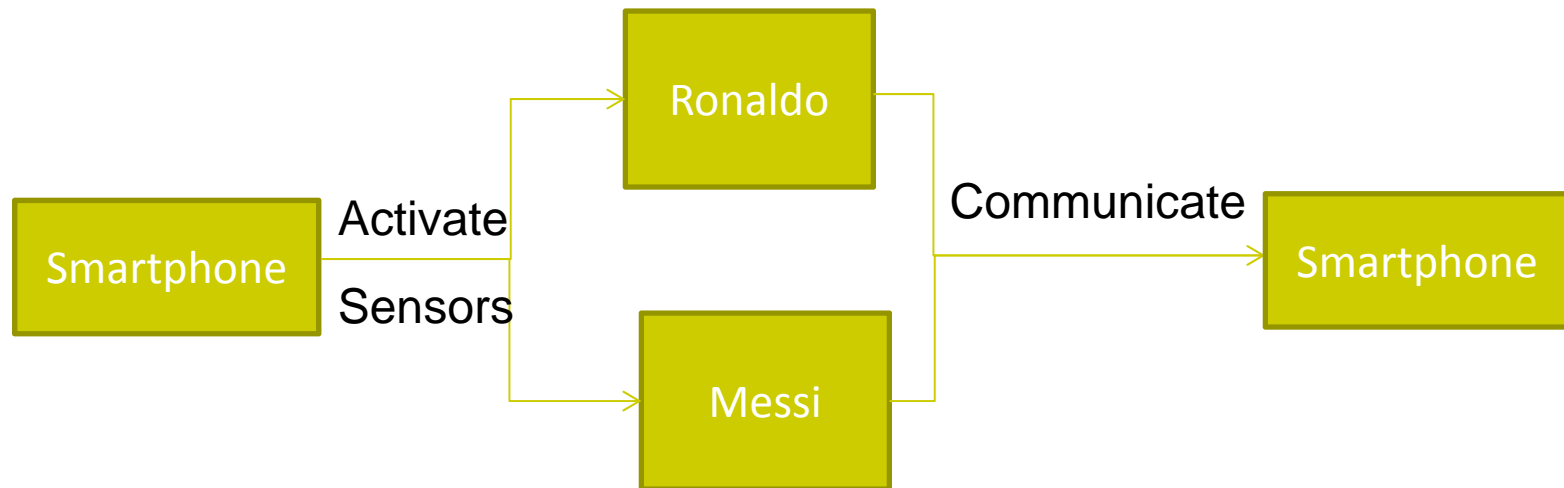
## Introduction (cont'd)

- 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?



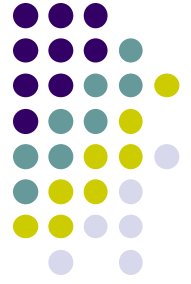
# Main Idea

- Sketch flow of TagSense:



- When - Where - Who - what

# Scope of TagSense



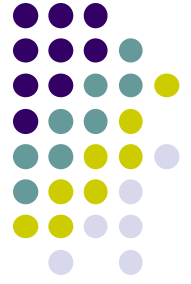
- Not a complete solution
- AT LEAST one of the sensing dimensions
- Electronic footprint required! (Image of objects, animals, people without phones, oops...)

# Comparison with Face Recognition



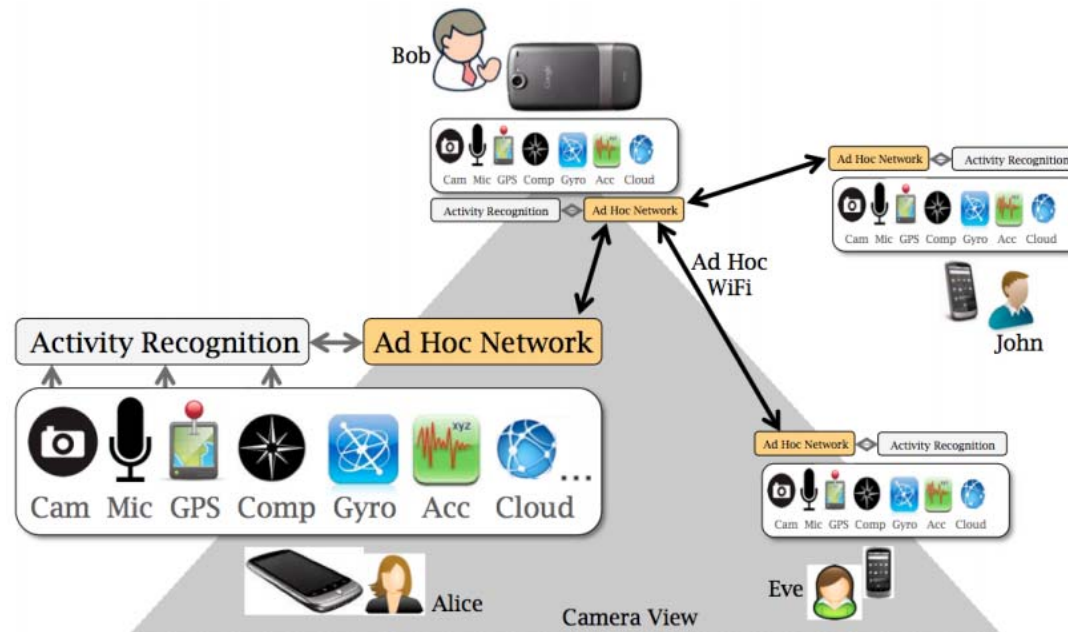
- Complementary!!!

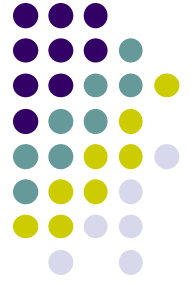
	Face Recognition	TagSense
Lighting surrounded	Good lighting	Bad lighting
Physical features	Yes (curious about twins)	Not really



# System Overview

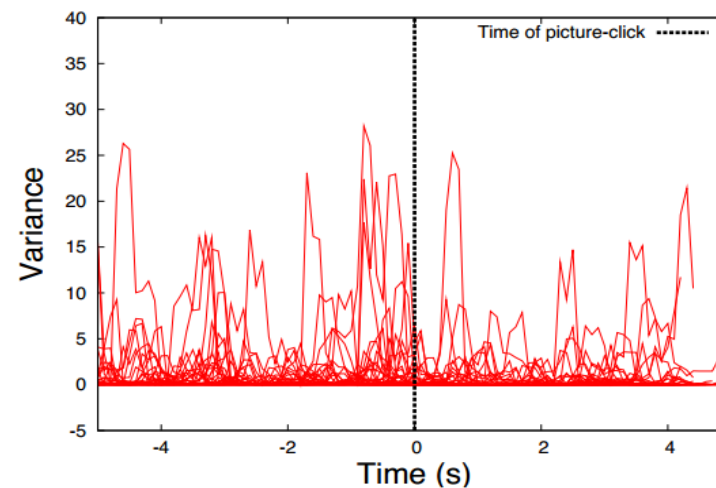
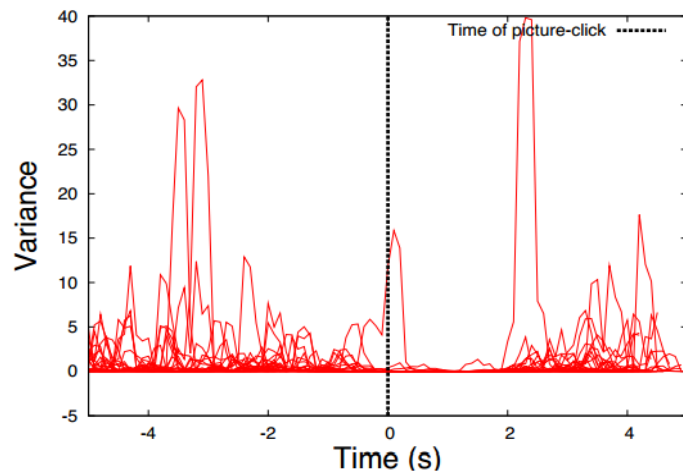
- Camera phone triggers sensing in participants
- Gathers the sensed information
- Determine who is in the picture





# 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

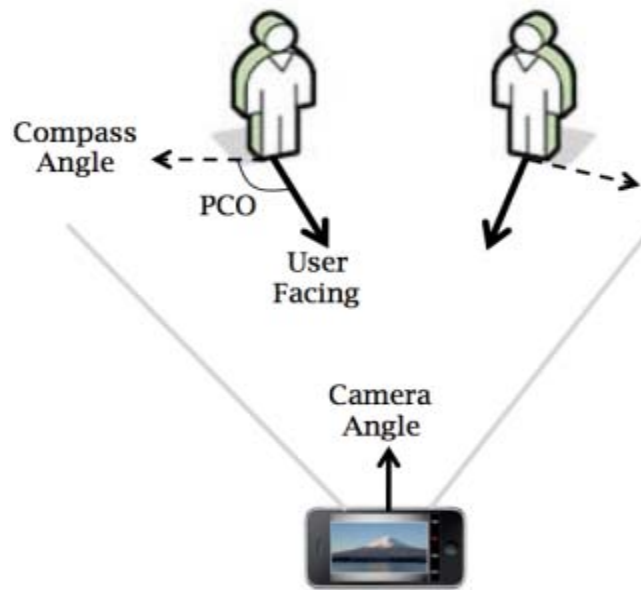






# Who are in the picture (cont'd)

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



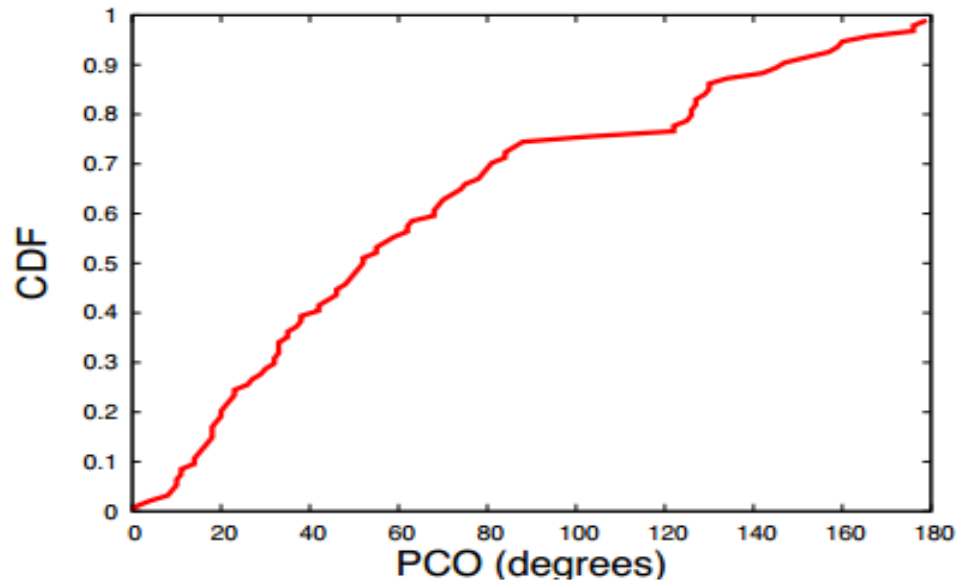
$$UserFacing = (CameraAngle + 180) \text{ mod } 360$$

$$PCO = ((UserFacing + 360) - CompassAngle) \text{ mod } 360$$



# Who are in the picture (cont'd)

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

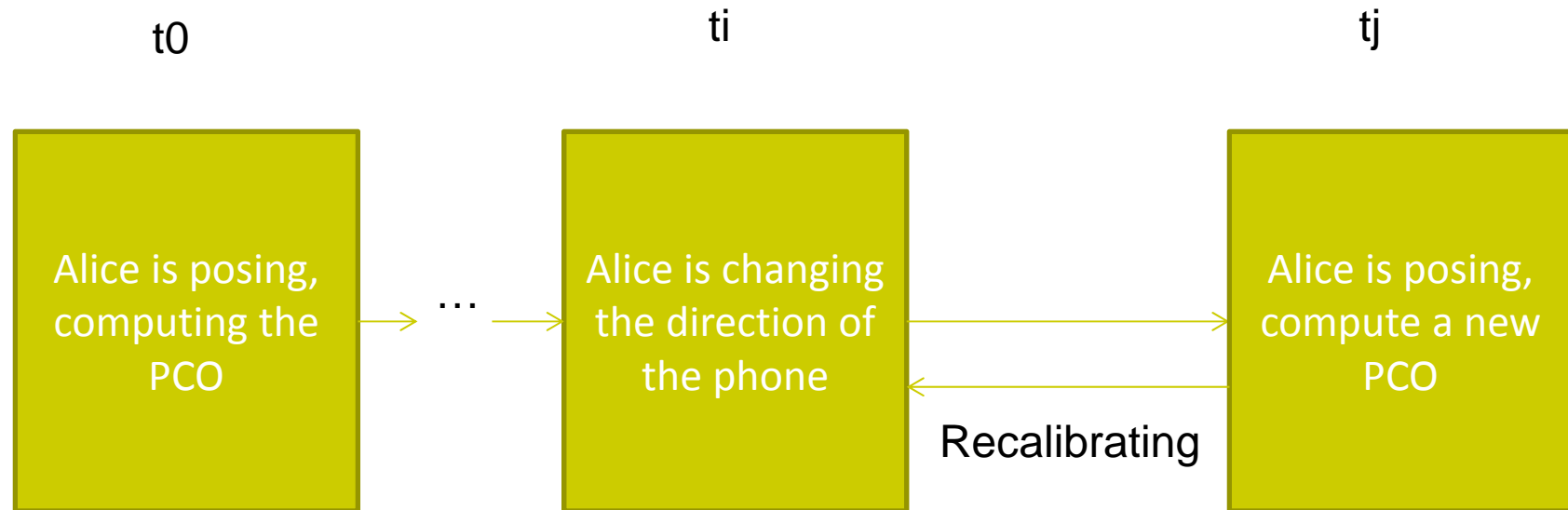


- Does not work 😞

# Who are in the picture (cont'd)



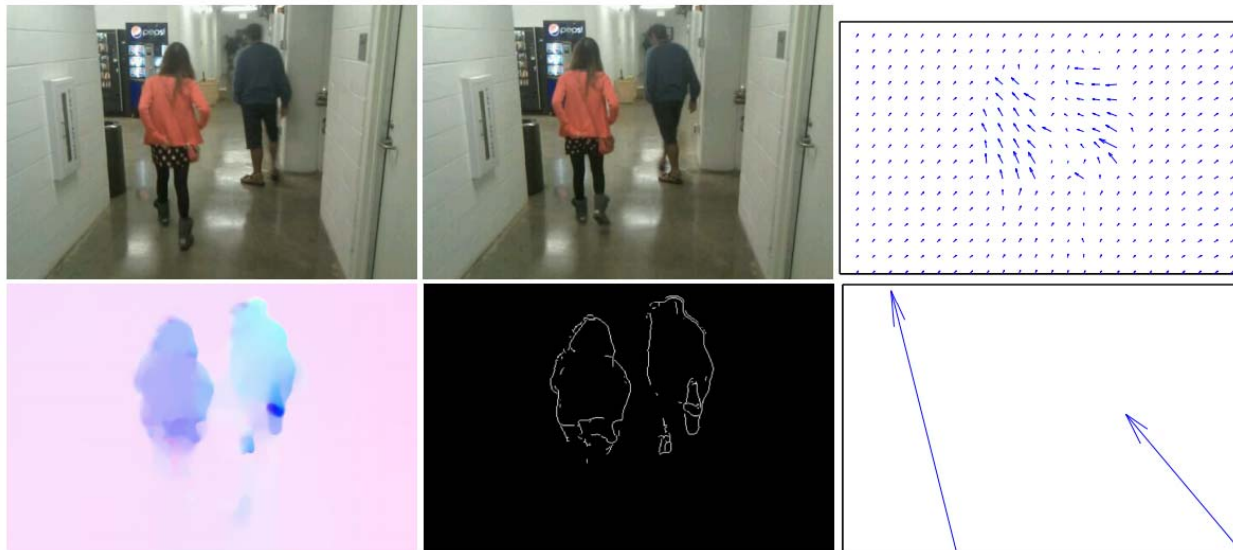
- Complementary compass directions
  - Recalibrating the PCO



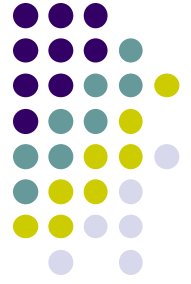


# Who are in the picture (cont'd)

- 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

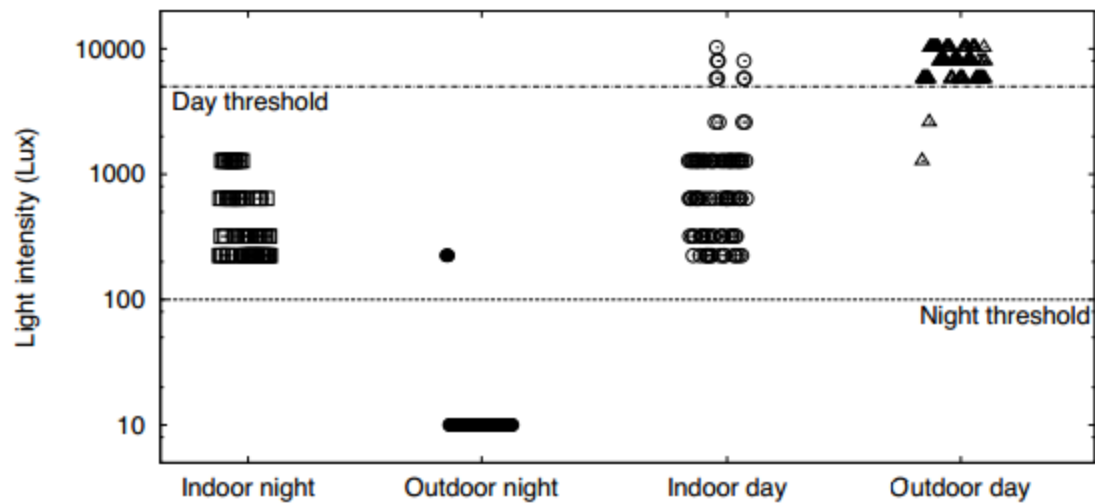


- Accelerometer
  - Standing, Sitting, Walking, Jumping, Biking, Playing
- Acoustic
  - Talking, Music, Silence



# Where is the picture taken

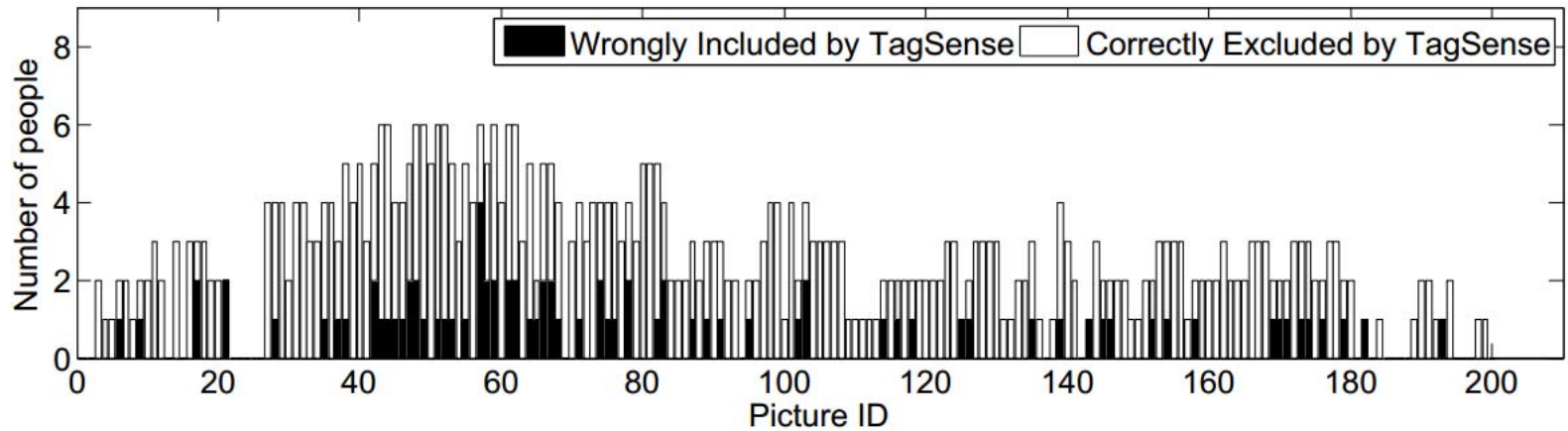
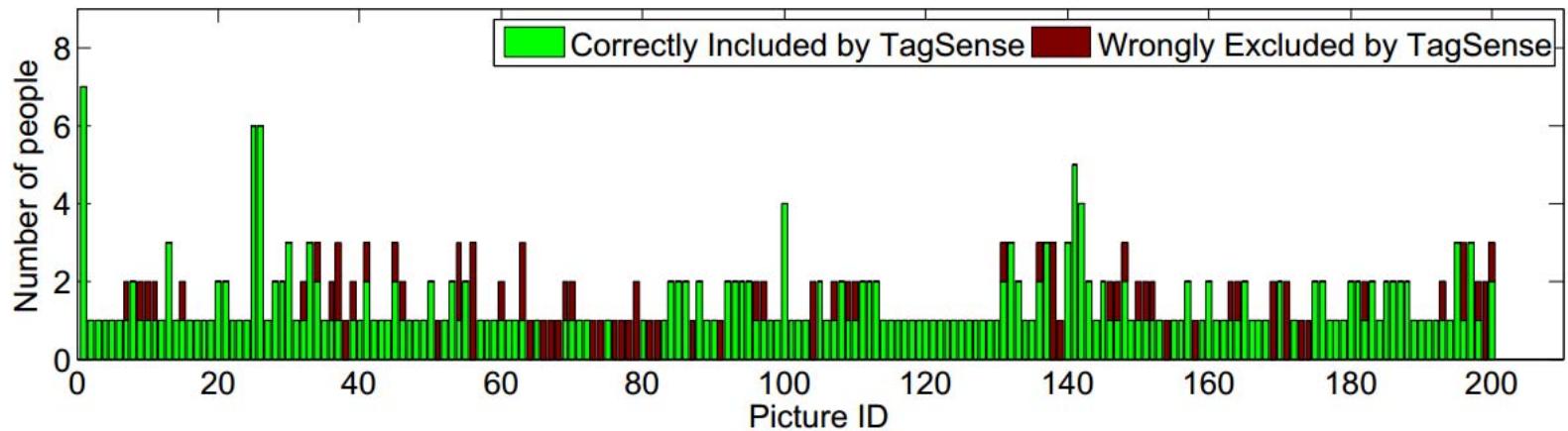
- Indoor? Outdoor?
  - Variation of light intensity measured 400 different times



# Performance



- Tagging people

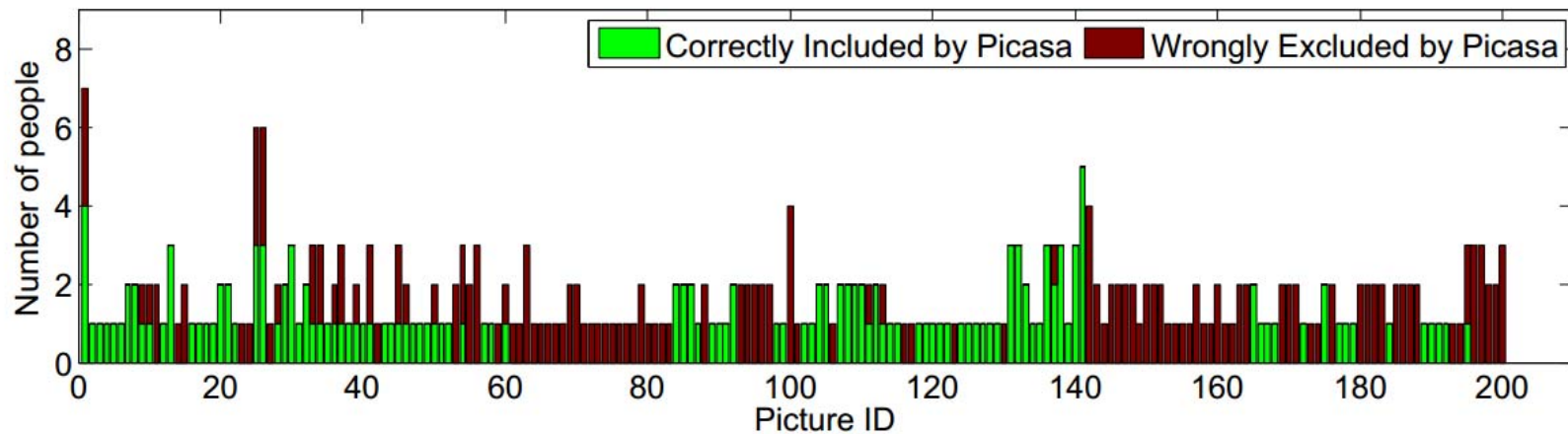
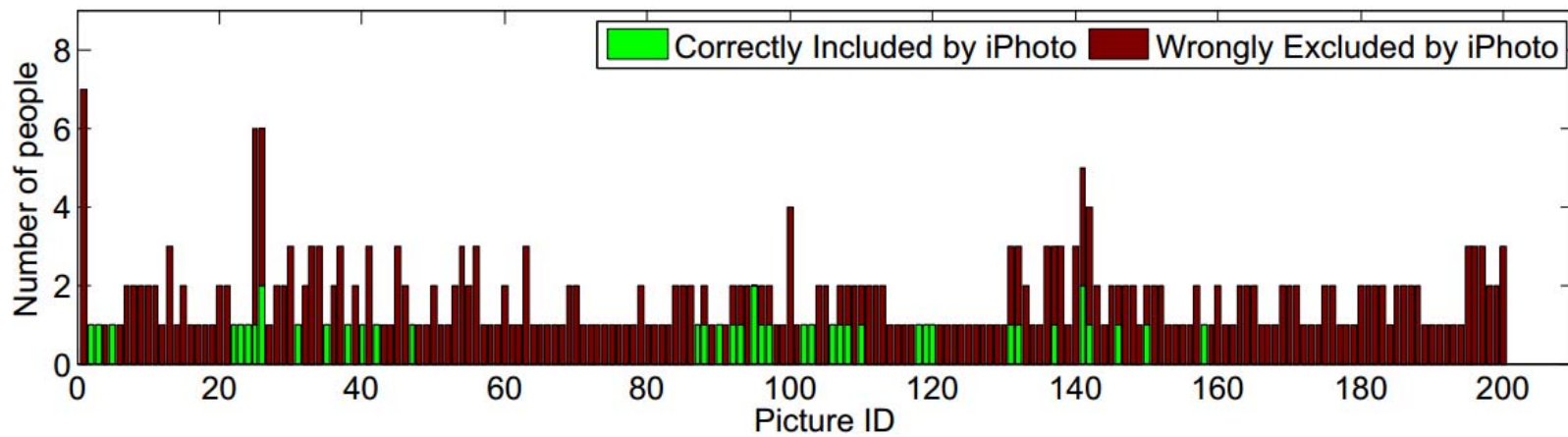






# Performance (cont'd)

- Tagging people



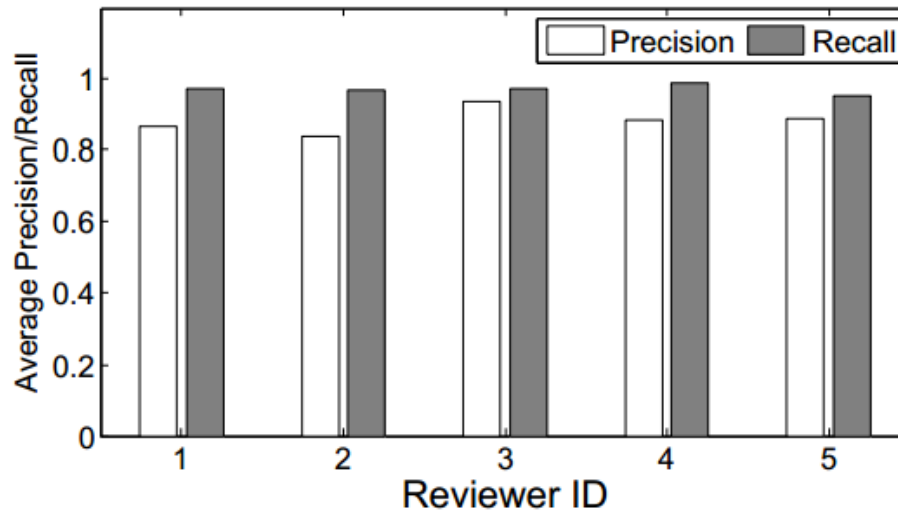


# Performance (cont'd)

- Tagging activities and context
  - Assessment by human

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

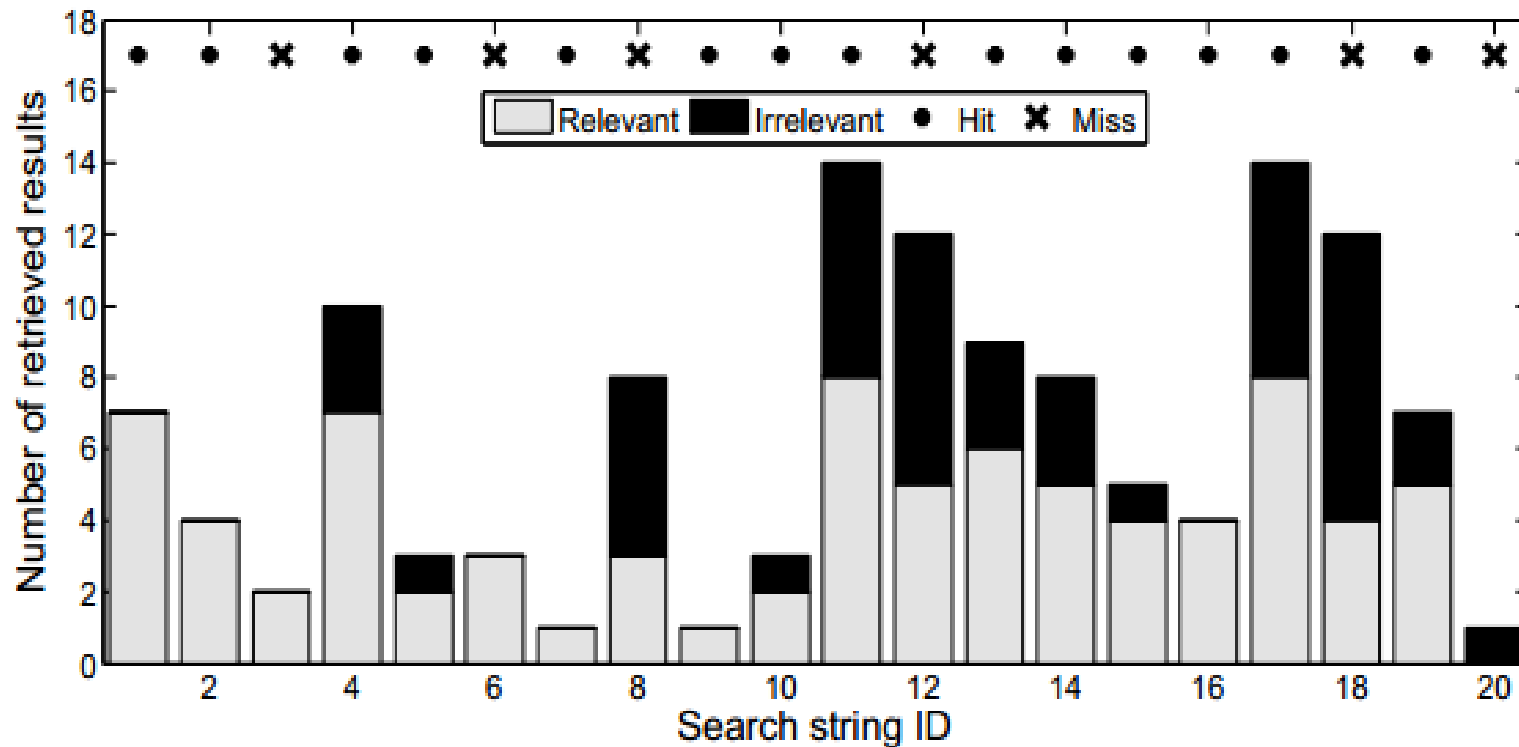
$$\text{recall} = \frac{|\text{Tags by Humans} \cap \text{Tags by TagSense}|}{|\text{Tags by Humans}|}$$





# 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



- [1] Tingxin Yan, Deepak Ganesan, and R. Manmatha, “Distributed image search in camera sensor networks,” ACM SenSys, pp. 155–168, Nov 2008.
- [2] Amazon, “Amazon Mechanical Turk,” <https://www.mturk.com/mturk/welcome>.
- [3] Google Image Labeler, “<http://images.google.com/imagelabeler/>,” .
- [4] L. Von Ahn and L. Dabbish, “Labeling images with a computer game,” in ACM SIGCHI, 2004.
- [5] Tingxin Yan, Vikas Kumar, and Deepak Ganesan, “Crowdsearch: exploiting crowds for accurate real-time image search on mobile phones,” in ACM MobiSys, 2010.
- [6] T. Nakakura, Y. Sumi, and T. Nishida, “Neary: conversation field detection based on similarity of auditory situation,” ACM HotMobile, 2009.
- [7] H. Lu, W. Pan, N. D. Lane, T. Choudhury, and A. T. Campbell, “SoundSense: scalable sound sensing for people-centric applications on mobile phones,” in ACM MobiSys, 2009.
- [8] A. Engstrom, M. Esbjornsson, and O. Juhlin, “Mobile collaborative live video mixing,” Mobile Multimedia Workshop (with MobileHCI), Sep 2008.
- [9] Google Goggles, “<http://www.google.com/mobile/goggles/>,” .
- [10] L. Bao and S.S. Intille, “Activity recognition from user-annotated acceleration data,” Pervasive Computing, 2004.



## Reference (cont'd)

- [11] D.H. Hu, S.J. Pan, V.W. Zheng, N.N. Liu, and Q. Yang, "Real world activity recognition with multiple goals," in ACM UbiComp, 2008.
- [12] M. Azizyan, I. Constandache, and R. Roy Choudhury, "SurroundSense: mobile phone localization via ambience fingerprinting," in ACM MobiCom, 2009.
- [13] C. Liu, "Beyond Pixels: Exploring New Representations and Applications for Motion Analysis," in Doctoral Thesis MIT, 2009.
- [14] E. Miluzzo, N. D. Lane, K. Fodor, R. Peterson, H. Lu, M. Musolesi, S. B. Eisenman, X. Zheng, and A. T. Campbell, "Sensing Meets Mobile Social Networks: The Design, Implementation and Evaluation of CenceMe Application," in ACM Sensys, 2008.
- [15] M. Braun and R. Spring, "Enkin," <http://enkinblog.blogspot.com/> .
- [16] E. Aronson, N. Blaney, C. Stephan, J. Sikes, and M. Snapp, "The jigsaw classroom," Improving Academic Achievement: Impact of Psychological Factors on Education, 2002.
- [17] A.A. Sani, L. Zhong, and A. Sabharwal, "Directional Antenna Diversity for Mobile Devices: Characterizations and Solutions," in ACM MobiCom, 2010.
- [18] K. Chintalapudi, A. Padmanabha Iyer, and V.N. Padmanabhan, "Indoor localization without the pain," in ACM Mobicom, 2010.



## Reference (cont'd)

- [19] C. Peng, G. Shen, Z. Han, Y. Zhang, Y. Li, and K. Tan, "A beepbeep ranging system on mobile phones," in ACM SenSys, 2007.
- [20] Nokia Siemens Networks, "Unite: Trends and insights 2009," 2009.
- [21] Sam Grobart, "In Smartphone Era, Point-and-Shoots Stay Home," New York Times, Dec 2010.
- [22] R. Datta, D. Joshi, J. Li, and J.Z. Wang, "Image retrieval: Ideas, influences, and trends of the new age," ACM CSUR, 2008.
- [23] Gustavo Carneiro, Antoni B. Chan, Pedro J. Moreno, and Nuno Vasconcelos, "Supervised learning of semantic classes for image annotation and retrieval," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, pp. 2007, 2007.
- [24] Alipr, "Automatic Photo Tagging and Visual Image Search ," [http:// alipr. com/](http://alipr.com/) .
- [25] Mor Naaman, Ron B. Yeh, Hector Garcia-Molina, and Andreas Paepcke, "Leveraging context to resolve identity in photo albums," in Proc. of the 5th ACM/IEEE-CS joint conference on Digital libraries, 2005, JCDL '05.
- [26] Risto Sarvas, Erick Herrarte, Anita Wilhelm, and Marc Davis, "Metadata creation system for mobile images," in ACM MobiSys, 2004.
- [27] Shwetak N. Patel and Gregory D. Abowd, "The contextcam: Automated point of capture video annotation," in Proc. of the 6th International Conference on Ubiquitous Computing, 2004.
- [28] R. Want, "When cell phones become computers," IEEE Pervasive Computing, IEEE, 2009.





## Reference (cont'd)

- [29] R.K. Balan, D. Gergle, M. Satyanarayanan, and J. Herbsleb, "Simplifying cyber foraging for mobile devices," in ACM MobiSys, 2007.
- [30] D.H. Nguyen, G. Marcu, G.R. Hayes, K.N. Truong, J. Scott, M. Langheinrich, and C. Roduner, "Encountering SenseCam: personal recording technologies in everyday life," in ACM Ubiquitous computing, 2009.
- [31] P. Mohan, V. N. Padmanabhan, and R. Ramjee, "Nericell: Rich monitoring of road and traffic conditions using mobile smartphones," in ACM SenSys, 2008.
- [32] J. Lester, B. Hannaford, and G. Borriello, "ÒAre You with Me?Ó-Using Accelerometers to Determine If Two Devices Are Carried by the Same Person," Pervasive Computing, 2004.
- [33] T. van Kasteren, A. Noulas, G. Englebienne, and B. Krose, "Accurate activity recognition in a home setting," in ACM Ubicomp, 2008.
- [34] M. Leo, T. D'Orazio, I. Gnoni, P. Spagnolo, and A. Distanto, "Complex human activity recognition for monitoring wide outdoor environments," in IEEE ICPR, 2004.
- [35] B. Logan, "Mel frequency cepstral coefficients for music modeling," in ISMIR, 2000.
- [36] S. Baker, D. Scharstein, JP Lewis, S. Roth, M.J. Black, and R. Szeliski, "A database and evaluation methodology for optical flow," in IEEE ICCV, 2007.
- [37] Joshua J. Romero, "Smartphones: The Pocketable PC," IEEE Spectrum, Jan 2011.