

CS 4518 Mobile and Ubiquitous Computing

Lecture 20: Movie Rating

Emmanuel Agu





**Your Reaction Shows You Liked the
Movie**

The Problem: Rating Movies & Videos

Your reactions suggest you liked the movie: Automatic content rating via reaction sensing, X Bao, S Fan, A Varshavsky, K Li, R Roy Choudhury, in Proc UbiComp 2013



● Current Rating System:

1. Today's ratings are mostly 1-5 rating, inadequate
2. Eliciting more in-depth, careful rating from users is difficult, requires incentives



Figure 1: Rating of Avatar from rotten tomatoes



Key Observations

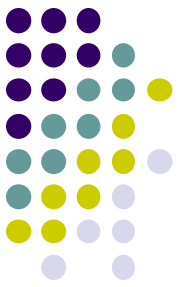
- Smartphone sensors can be used to infer user rating while users watch YouTube videos
 - Laughter detected (microphone) => Funny
 - Stillness while watching (accelerometer) => Intense drama
 - Head turn (front facing camera) + talk (microphone) => Lack of interest
 - Fast forwarding movie => Lack of interest



- **Paper Goal** : Research and Develop movie rating system called **Pulse**
 - Learns mapping between the sensed reactions and ratings
 - Automatically computes users' ratings.

Pulse Vision

- Movie's playback timeline can be annotated with reaction labels (e.g., funny, intense, warm)
- Senses user reactions and translates them to an overall system rating.
- In future, tag-cloud of these sensed user reactions can augment movie ratings

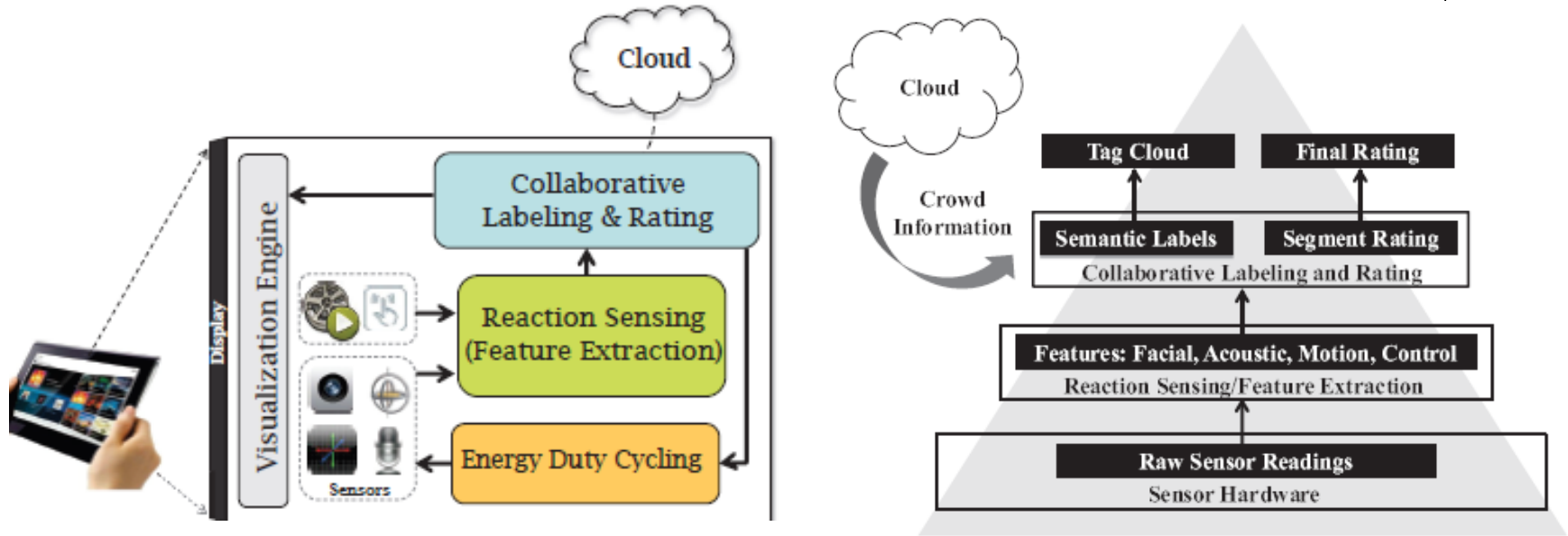


Pulse Vision

SYSTEM OVERVIEW



- **Main modules** : Reaction Sensing and Feature Extraction (RSFE), Collaborative Labeling and Rating (CLR), and Energy Duty-Cycling (EDC).

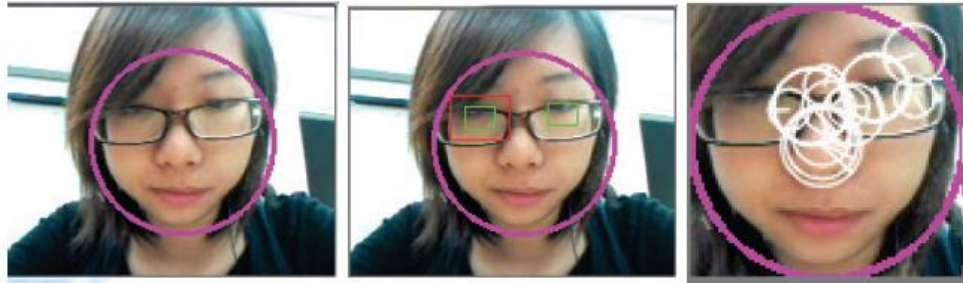


- **RSFE**: processes the raw sensor readings and extracts features to feed to CLR.
- **CLR**: The CLR module processes each (1 minute) movie segment of the movie to create “semantic labels” + “segment ratings”.
 - Segment ratings are merged to yield the final “star rating”
 - Semantic labels are combined to create a tag-cloud.
- **EDC**: minimizes energy consumption due to sensing.

System design: RSFE



- **Visual:** Pulse detects the face through camera, detects eyes using blink detection, generates visual features and tracks key points (face, eyes, lip)



- **Acoustic:**
 - **Voice Detection:** Activates microphone, records ambient sounds, separates user's voice
 - **Laughter Detection:** Pulse assumes that acoustic reactions during a movie are either speech or laughter
 - Once human voice is detected, classified as speech or laughter
 - Support vector machine (SVM) classifier using Mel-Frequency Cepstral Coefficients (MFCC) as features.
- **Control operations:** Users skip boring movie segments, rewind interesting segments
- Visual, acoustic features and control operations forwarded to CLR module



Pulse Evaluation Methodology

- 11 volunteers, 6 new movies, watch movies using Pulse video player
- After watching: rate segments, perception label, final “star” rating

Challenges

Predicting human judgment, minute by minute, is quite difficult.

- **Heterogeneity in users behavior**
Some users naturally fidgety, others still
- **Heterogeneity in environment factors**
Eg: Same user may watch same movie differently at *office VS. at home*
- **Heterogeneity in user tastes**
Different users may rate same movie differently

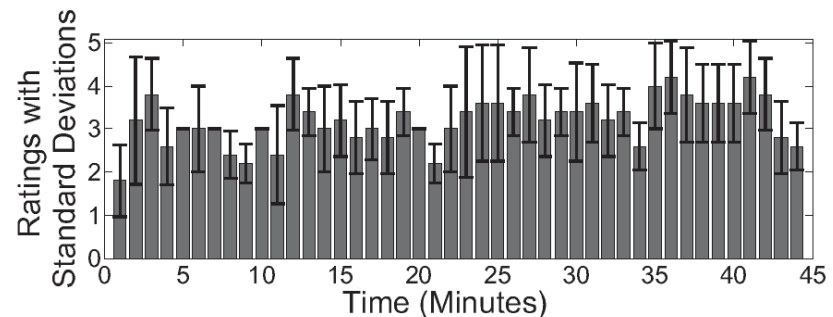


Figure 11. High Std. Dev. in ratings across users.



Final Results

- **Performance of Final “Star” Rating**

Average error of 0.46 on a 5 point scale.

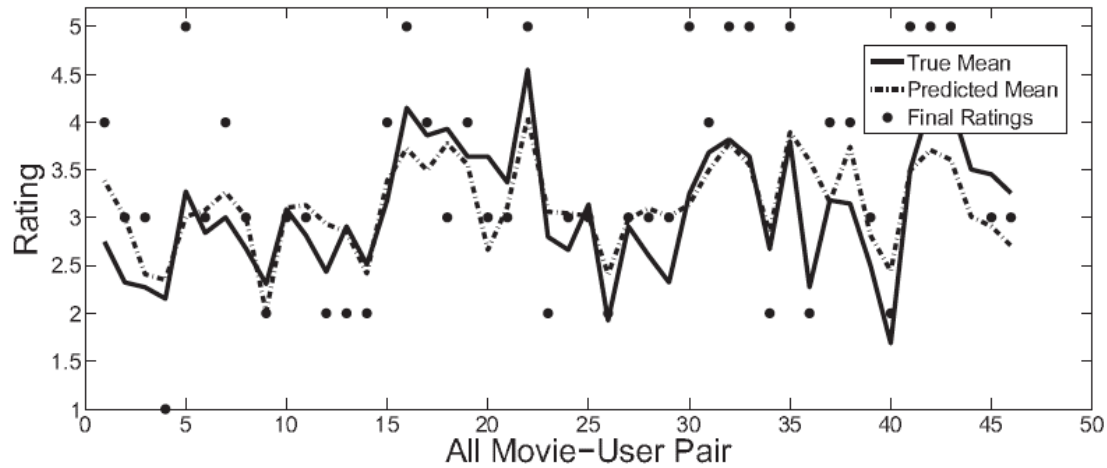


Figure 18. (a) Mean segment ratings and corresponding users' final ratings.



What Else Sensed?

Other Sensable Behaviors



- Mood (happy, sad, etc)
 - Predictors: e.g. late night browsing (sad)
- Boredom of Smartphone User
- Addicted Smartphone Usage