Your Reaction Suggest you Liked the Movie:

Automatic Content Rating Via Reaction Sensing

Navneet Kaur

IS- 709

03/09/2016

What is Rating?

* **** For Netflix Movie
* 87% red-tomato by Flixster
* 23 Likes for You-Tube video

These numbers are highly-lossy comparison

Problem Statement

* Eliciting a carefully considered rating from users is difficult
  * Due to lack of incentive
  * Providing a review can take a good amount of user's time
  * User not willing to make this time investment after watching a movie

Solution

* COTENT RATING SYSTEM-
  * Require minimal user participation
  * Provide rich, informative rating

Envisioned Movie Rating for Future

Pulse

* Pulse is a system that record user's reactions while watching a movie and automatically convert them into ratings.
* OPPORTUNITY
  * The opportunity arise form growing number of sensors especially in smartphone and tablet

Hypothesis

* When the user watches movie on these devices, a good fraction of their reactions leaves footprint on various sensing dimensions:
  * Movement of user head(detectable through front facing camera)
    * one could infer user lack of attention to that movie.

Hypothesis(cont.)

* Other kind of inference-
  * Laughter detection via microphone
  * Fast forwarding of the movie

Pulse learning

* Learns the mapping between the sensed reactions and these ratings
* Automatically compute their ratings, especially when they don’t provide one

Pulse Applications

* The timeline of the movie can be annotated with reaction labels (funny, intense, warm)
* Beneficial for advertising company
* Students can use reaction logs as case studies from real-world users-Educational Value

Challenges

* Translating the pulse to reality requires a number of challenges such as-
  * The viewer's head pose
  * Lip movement
  * Eye blinks

  Need to be detected and monitored over time to infer reactions
* Machine learning

Samsung Tablet

* Identified an opportunity
* Design Pulse
* Developed Pulse on Android based Samsung Galaxy Tablet

System Overview

* Reaction Sensing/Feature Extraction(RSFE)

All sensors are activated including the front facing camera, microphone, accelerometer, available location sensors

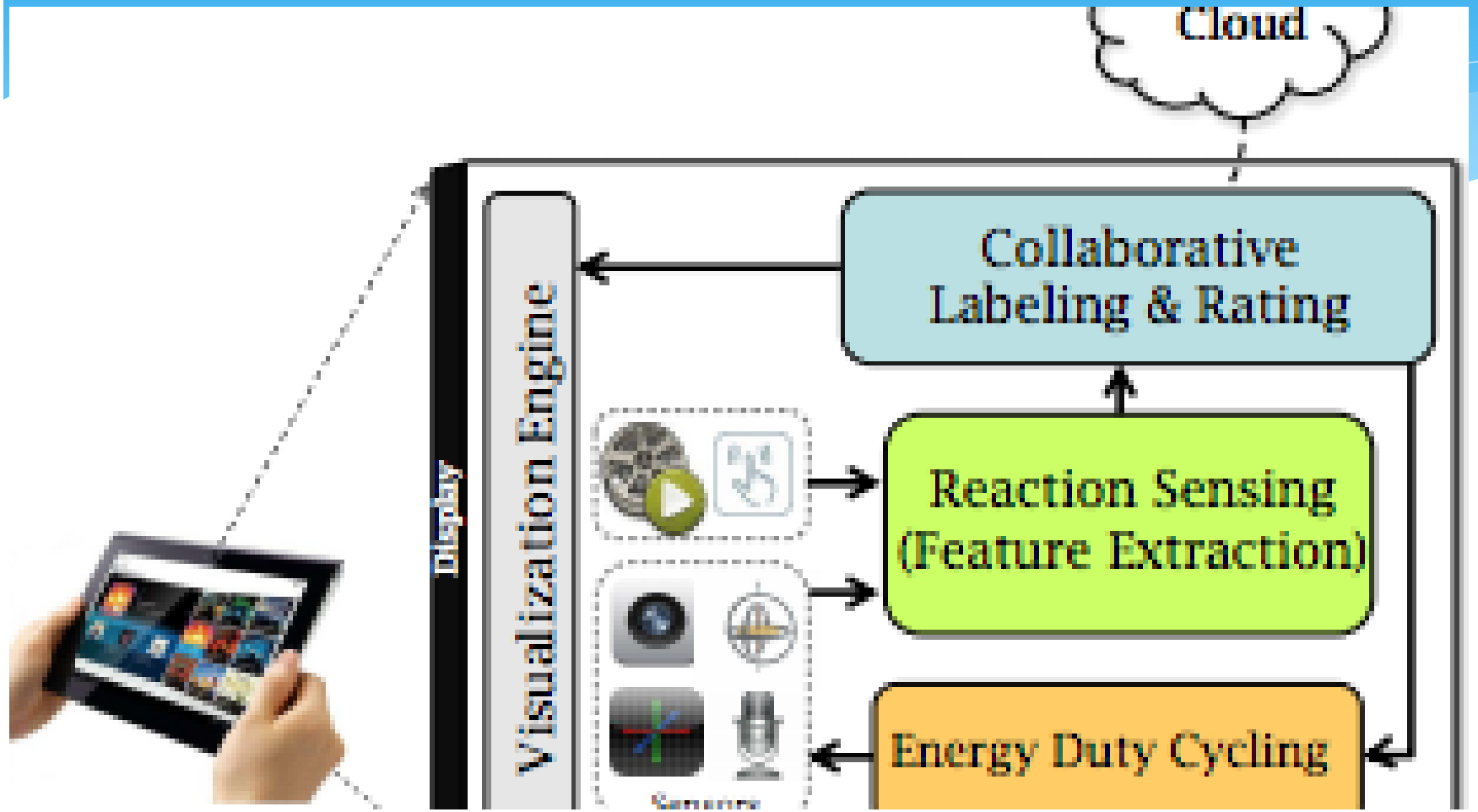
* Collaborating Labeling and Rating

Pulse employs Collaborative filtering methods

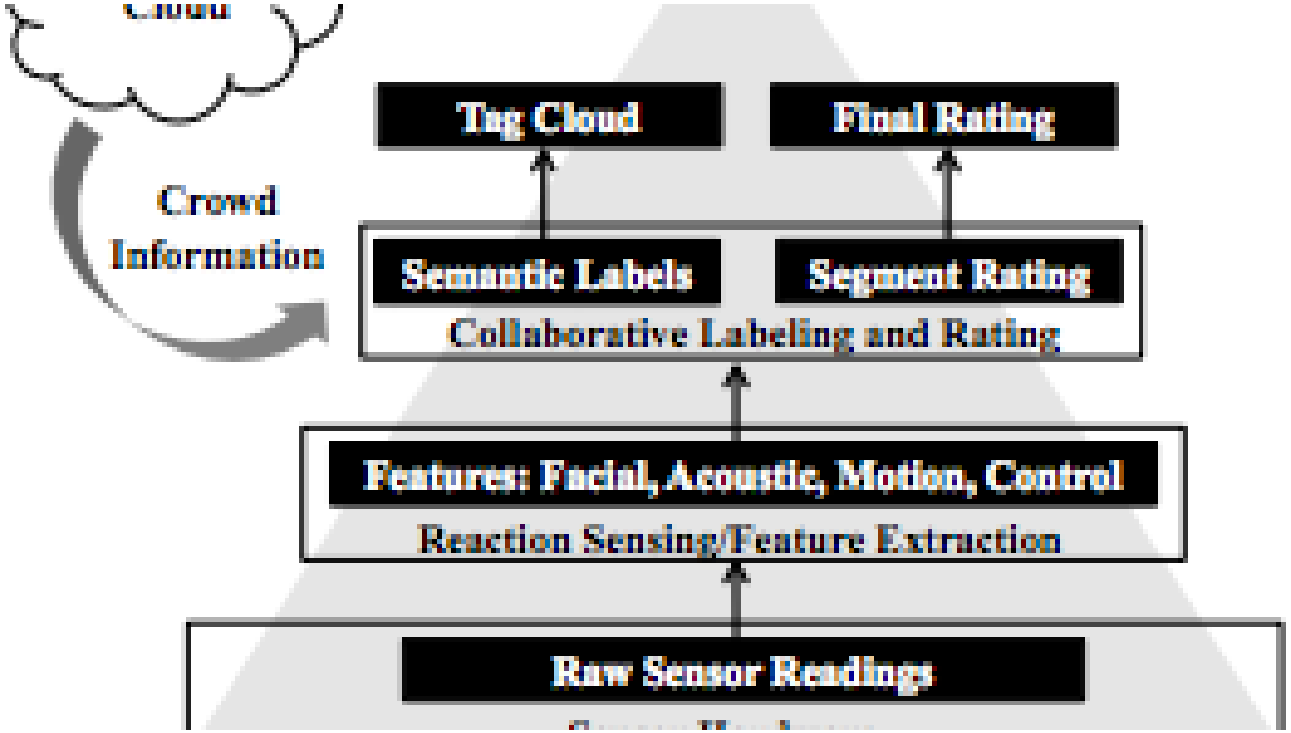
* Energy Duty-Cycling(EDC)

To minimize the energy consumption due to sensing, when running on smartphone

System Overview



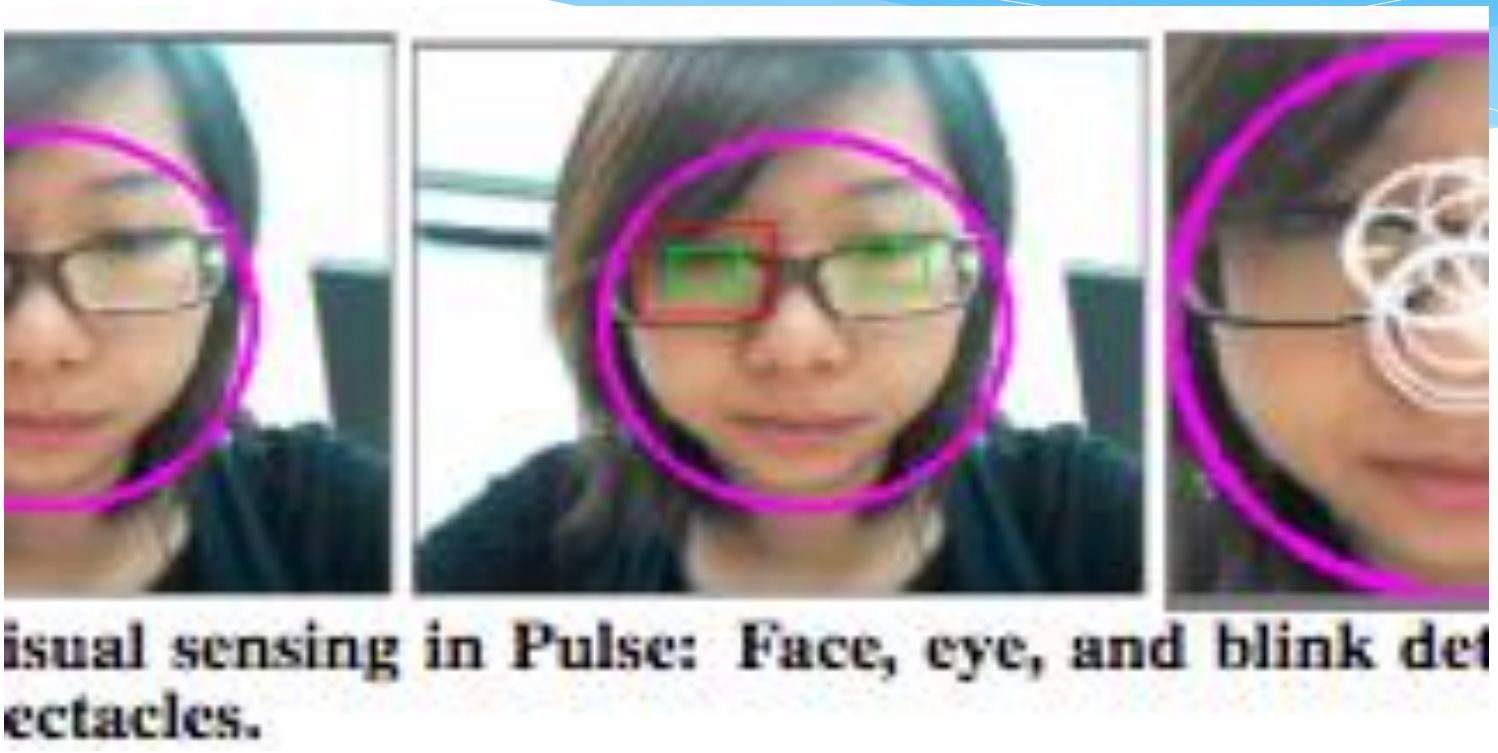
RSF and CLR Modules Distill Raw sensor reading to a rating, tag-cloud



Visual Sensing in Pulse

* Pulse continuously runs contour matching algorithm on each frame for face detection
* Then system runs contour matching for eye detection, lip detection
* If full face is not detected, Pulse track key-points(SURF key points)
* Pulse runs algorithm to perform blink-detection , and eye-tracking

Output of the algorithm



isual sensing in Pulse: Face, eye, and blink det
ectacles.

Pulse detects Face through the tablet camera , detects eyes using blink detection ,And finally track key points

Voice Detection(Comparing the power spectral densities)

* The original move soundtrack[-]
* The sound of the movie recorded through tablet microphone[-]
* The sound of the movie and human voice, recorded by tablet microphone[---]

Voice Detection

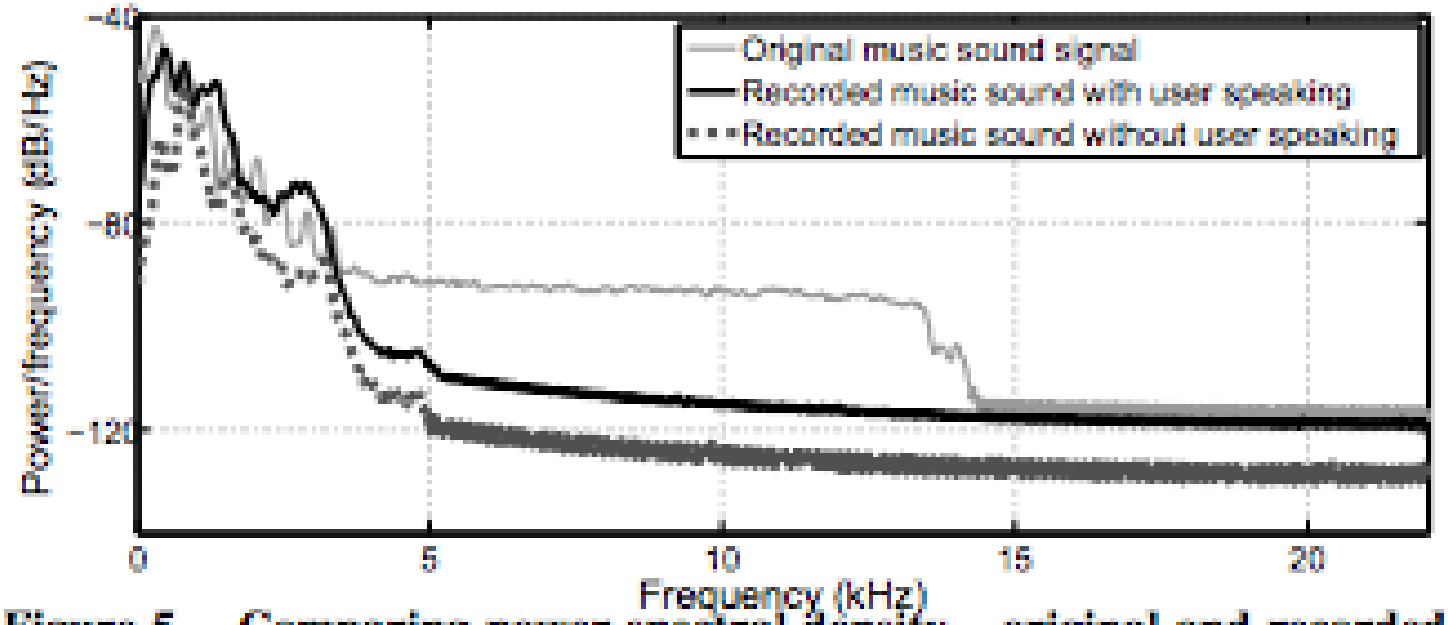


Figure 5. Comparing power spectral density – original and recorded soundtrack with human voice

Laughter Detection

* Classification of Human voice to two categories- speech, laughter
* Train Support vector machine on Mel frequency Cepstral Coefficients(MFCC) as the principle features.
    Pulse perform simple outlier detection to reduce false positive. So, if the frame is suspected as laughter, but 4 preceding and following frames are not, then theses outlier frames are eliminated.

Laughter Detection(cont.)
High accuracy and few false positive

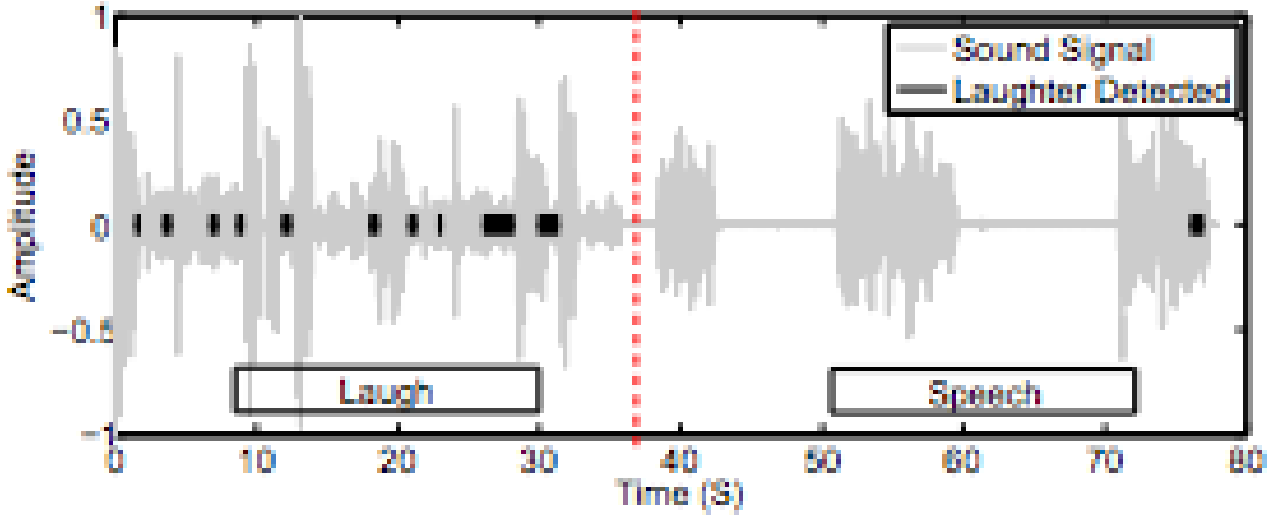


Figure 7. Discriminating laughter and speech from voice signals recorded by a tablet microphone.

Reaction Features

* Pulse observe a record different segments and assign lower/higher rating accordingly
  * Skipping boring segment
  * Roll back to interesting segment

Key Goal

* Key goals are to model the sensed data and use model to:

1. Estimate segment rating
2. Generate the final star rating from segment ratings
3. Estimate semantic labels
4. Generate the tag-cloud from semantic label

Segment Ratings

* For every short segment of a movie
* To compute the overall movie quality
* To select enjoyable segments

Key Challenge

* Ambiguity(Same user reaction for different scene of a movie)
* Pulse employs Collaborative Filtering and Gaussian Process Regression
* Uses weighted average function to convert segment rating to final rating

Semantic Label

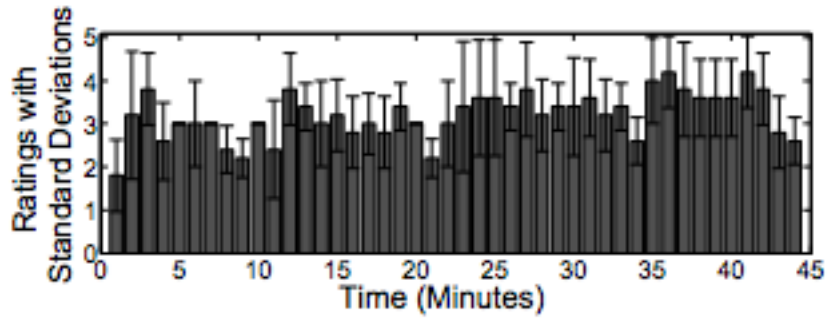
* CLR generates two types of label
  * Reaction label (laugh, smile, focused, distracted, nervous)
  * Perception label (funny, exciting, warm etc.)

Experimental Methodology

* Recruited 11 volunteers(4 females)
* Aged 24-28
* Provide volunteers with Android-based Samsung tablets preloaded with 6 movies
  * 3 comedies
  * 2 dramas
  * 1 horror
* Also provided software tool that allowed user to rate the movie soon after they watched it
  * 1 “did not like”
  * 5 “liked”

CHALLENGES

* Heterogeneity in users behavior
* Heterogeneity in environmental factors
* Heterogeneity in User tastes



High std. Dev. In ratings across users

Summary of Result

$$Precision = \frac{|\{\text{Human Selected} \cap \text{Pulse Selected}\}|}{|\{\text{Pulse Selected}\}|}$$

$$Recall = \frac{|\{\text{Human Selected} \cap \text{Pulse Selected}\}|}{|\{\text{Human Selected}\}|}$$

$$Fall - out = \frac{|\{\text{Non-Relevant} \cap \text{Pulse Selected}\}|}{|\{\text{Non-Relevant}\}|}$$

Summary Of Result

* Segment Rating: Pulse is able to capture enjoyable segment with
  * An average precision of 71%
  * An average recall of 63%
  * Minor fallout of 9%
* Final Rating
  * Pulse's overall star rating demonstrates an average error of 0.46 in 5 point scale.

Conclusion

* Guides an opportunity into an application that automatically rates the content on behalf of human users
* The core idea is to leverage device sensors such as: cameras, microphones, accelerometers to sense qualitative human reactions while watching video
* A movie automatically get tagged by
  * Star rating
  * Tag cloud of user reaction
  * Highlights of movie for different emotions

Insight Gain

* Building sensing based automatic system would be beneficial
* Pulse is still prototype with many limitations
* Pulse may raise privacy concerns
* Should explore additional optimizations in machine learning to improve performance

QUESTIONS?

?