INTRODUCTION

- Properties of humor
  1. Producer + perceiver
  2. Positive emotional reactions (laughter)
  3. Highly individualistic & cultural specific

Lack of multimedia data annotated with humor

- Our approach
  - Collect videos potentially with humorous utterances
  - Use time-aligned user comments to automatically generate humor labels
  - Analyze the speech, text, and visual characteristics of humorous expressions
  - Train models for predicting humor

DATA COLLECTION

- Time-aligned comments on bilibili.com
  - Users can post comments about a specific scene while watching the video
  - ‘Papi酱’ - A Chinese online celebrity
    - Famous for discussing trending topics in a humorous way
  - 100 videos, 93593 comments

- Video segmentation
  - One-second unit level: 24K segments
  - Inter-pausal unit (IPU) level: 8K segments

- Laughing indicators
  - ‘233’ (internet meme) + ‘hh’ or ‘哈哈’ (onomatopoeia of laughter)

Audiences tend to respond to humor in videos with laughing
A high volume of laughing comments at a given time

HUMOR

CONSTRUCTING UNSUPERVISED LABELS

- Response time estimation
  - Users typically do not pause to comment
  - Response Time = reaction time + typing time

- An example of contextual smoothing and automatic labeling on IPU level

FEATURE EXTRACTION

- Acoustic-Prosodic features: pitch, intensity, speaking rate, etc.
- Transcript-based lexical features: word categories such as function words, affect words, social words, etc.
- Visual features: frame similarity, body pose, facial landmarks

FUTURE WORK

- Collecting videos with different types of humor
- Other live streaming websites with time-aligned user comments