Amusement Research Project

Authors: Oniga Aris & Paulovici Andrei

The project presents a standalone desktop application capable of analyzing facial expressions and audio signals in real time to detect laughter and quantify amusement. The system transforms subjective human reactions into objective numerical scores, enabling structured comparison between different video categories and social contexts.

Research Questions

Q1: What category of videos do people find the funniest?

- Funny animals
- Human fails
- Trending Memes/Internet Humor
- Pranks/Hidden Camera clips
- Reaction Humor
- AI-powered clips
- Culture-Related Humor

Are there any differences between male and female participants?

Q2: How does the presence of others influence the amusement intensity?

Facial Analysis and Action Units

Facial expressions are analyzed using the Facial Action Coding System (FACS). The system focuses on Action Units most closely associated with smiling and laughter:

- **AU12 (Lip Corner Puller)** – indicates smile width
- **AU6 (Cheek Raiser)** – indicates genuine amusement
- **AU25 / AU26 (Mouth Opening / Jaw Drop)** – associated with laughter intensity

Facial Landmark Extraction

Facial landmarks are extracted using MediaPipe Face Mesh, which provides 468 real-time facial landmarks.

Sound Extraction

YAMNet audio is used to detect laughter-related sound activity, providing a complementary audio signal that strengthens amusement detection when combined with facial expression analysis.

Quantitative Formulas (Explicit)

Let

$$d(A, B) = \text{Euclidean distance between landmarks}$$

$$\text{Norm H} = d(\text{Left Eye Inner Corner, Right Eye Inner Corner})$$

$$\text{Norm V} = d(\text{Nose Tip, Chin Tip})$$

Action Units

AU12 — Lip Corner Puller (smile width)

$$AU12 = \frac{d(\text{Left Mouth Corner, Right Mouth Corner})}{\text{Norm\_H}}$$

AU6 — Cheek Raiser (eye–cheek compression)

Left cheek:

$$AU6\_L = 1 - \frac{d(\text{Left Cheek, Left Lower Eyelid})}{\text{Norm\_H}}$$

Right cheek:

$$AU6\_R = 1 - \frac{d(\text{Right Cheek, Right Lower Eyelid})}{\text{Norm\_H}}$$

Final AU6:

$$AU6 = \frac{(\text{AU6\_L} + \text{AU6\_R})}{2}$$

AU25 — Lips Part (lip separation)

$$AU25 = \frac{d(\text{Upper Lip Center, Lower Lip Center})}{\text{Norm\_V}}$$

AU26 — Jaw Drop

$$AU26 = \frac{d(\text{Lower Lip Center, Chin Tip})}{\text{Norm\_V}}$$

Audio activity is detected using YAMNet:

$$\text{AudioActivity} \in [0, 1] \text{ (where 1 indicates high probability of laughter)}$$

Scoring Model

$$\text{SmileScore} = 0.6 \cdot AU12 + 0.4 \cdot AU6$$

$$\text{LaughterScore} = 0.3 \cdot AU12 + 0.3 \cdot AU6 + 0.2 \cdot AU25 + 0.2 \cdot \text{AudioActivity}$$

$$\text{AmusementScore} = 0.6 \cdot \text{LaughterScore} + 0.4 \cdot \text{SmileScore}$$

All scores are normalized to the range [0, 1].

Video Data Collection

01

YouTube API Harvesting

02

Video Metadata Storage

03

Admin Screening

04

Approved Video Pool

05

Randomized Playlists for Experiments

- Curated search queries per category (animals, fails, pranks, memes, reaction humor, AI-powered clips, culture-related)
- Ensures diversity and reduces off-topic results
- Metadata stored: link, duration, category, status

Ensuring Accuracy & Validity

Content Quality Control

- Manual approve/deny review process
- Remove irrelevant and clickbait content
- Verify correct category assignment

Sampling Fairness

- Randomize selection from approved pool
- Avoid duplicate videos
- Maintain category variety

Data Integrity

- Store experiment-video score pairs
- Enforce consistent normalized scoring
- Maintain short video duration constraints

Recap / Methodology

The experiment is conducted in a quiet, controlled indoor environment to minimize background noise, as the audio module is sensitive and relies on clean sound input for accurate detection. Participants are instructed not to speak, chew, or exaggerate facial movements during video playback, ensuring that detected mouth activity corresponds to genuine amusement rather than intentional actions.

Participant Setup

30 participants took part in the study. Each participant watched a sequence of short humorous video clips selected randomly from predefined categories. To study social influence, some subgroups of 3-5 participants completed two sessions: one individually and one in a shared viewing environment where all 3-5 participants were present in the same room.

Results

Q1 — Funniest Video Categories

(normalized amusement, 0–1)

- Funny animals — 0.45
- Human fails — 0.44
- Trending memes / Internet humor — 0.41
- Reaction humor — 0.39
- Pranks / hidden camera — 0.39
- Culture-related humor — 0.37
- AI-powered clips — 0.36

Gender Comparison

Video Type	Female	Male
Animals	0.49	0.41
Reaction Humor	0.35	0.43
Culture-related Humor	0.38	0.37
Fails	0.41	0.48
Pranks	0.37	0.45
Memes	0.38	0.43
AI Clips	0.36	0.37

Differences are small to moderate (≈ 0.01 - 0.08)

Q2 — Social Context and Amusement

- Small group of friends (3–5) — 0.47
- One close friend — 0.45
- Alone — 0.40

Final Conclusion

This study demonstrates that amusement can be measured objectively using facial landmarks and lightweight audio cues. The results highlight clear trends in content preference and strong social effects on laughter, providing useful insights for content creation, marketing strategies, and emotion-aware systems.

Bibliography

- Tian, Y., Kanade, T., & Cohn, J. F. (2005). Facial Expression Analysis. In Handbook of Face Recognition (pp. 247–275). https://doi.org/10.1007/0-387-27257-7_12
- Girard, J. M., Cohn, J. F., & De La Torre, F. (2014). Estimating smile intensity: A better way. Pattern Recognition Letters, 66, 13–21. https://doi.org/10.1016/j.patrec.2014.10.004
- Valstar, Michel & Jiang, Bihan & Mehu, Marc & Pantic, Maja & Scherer, Klaus. (2011). The first facial expression recognition and analysis challenge. IEEE Trans. SMC B. 42. 921-926. 10.1109/FG.2011.5771374.