#240

Browsing Visual Sentiment Datasets using Psycholinguistic Groundings

E-mail: kastnerm@murase.is.i.nagoya-u.ac.jp

Marc A. Kastner Ichiro Ide Yasutomo Kawanishi Takatsugu Hirayama Daisuke Deguchi Hiroshi Murase

Nagoya University, Japan

Motivation

Sentiment classification is a popular task in SNS

- Various applications can benefit from a better multi-modal sentiment understanding
- Sentiment datasets provide images and metadata for Adjective-Noun Pairs (ANPs)

Per image sentiment annotations are labor intensive

Can a textual psycholinguistic analysis help understanding per-image sentiments better?

Datasets

Images: Multilingual Visual Sentiment Ontology[1]

- 7.36M images annotated with:
  ANP and textual metadata (title, description, tags)
- Per-ANP sentiment ratings

Problem: Sentiment ratings are shared across images for the same ANP

Psycholinguistic measures: Glasgow Norms[2]

- 9 ratings each for 5,500 words:
  Arousal, Dominance, Valence, Imageability, Concreteness, Familiarity, Gender, Age of Acquisition, Semantic Size

Can we connect both for sentiment understanding?

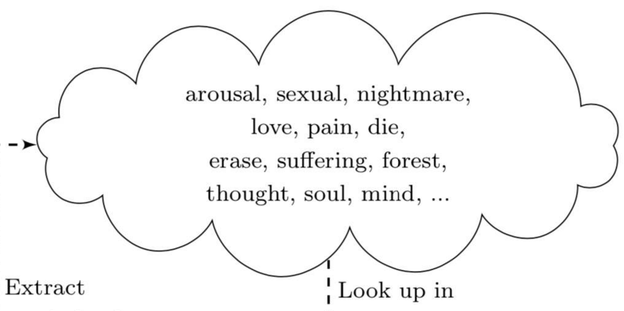
Psycholinguistic grounding

Compute per-image psycholinguistic ratings

1. Create word-cloud using textual metadata of an image
2. Look up psycholinguistic ratings from Glasgow Norms[2]
3. Calculate average score using top/bottom 5 words for each rating

+

Title: WALK OF SHAME
ANP: sexual abuse
Tags: abuse, uncle, damage, tears, morning, ...
Description:
... A mind too young to know what's wrong ...
... The pain and suffering he once felt ...
... Recalling nightmares from his sleep ...



Extract word cloud

Look up in psycholinguistics database

Word	Arousal	Valence	Concreteness	...
abuse	48	7	40	...
sexual	84	78	42	...
nightmare	59	10	36	...
...	...	...	...	...
average	59	51	57	...

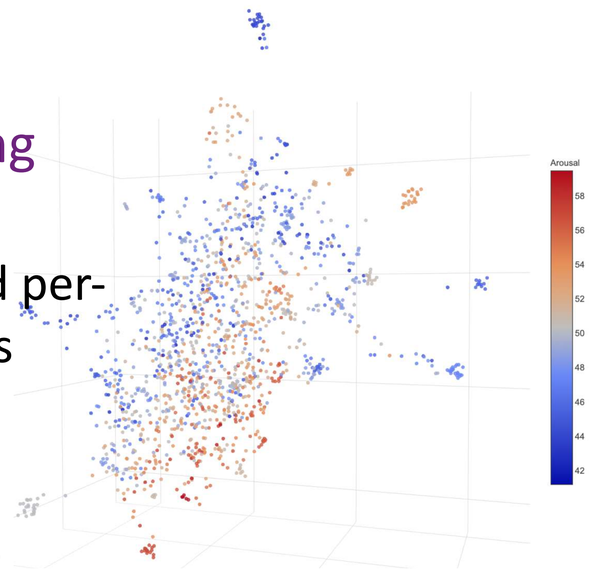
Visualization tool

Compute spatial embedding using UMAP[3]

- Using per-ANP sentiment and per-image psycholinguistic ratings

Visualize the embedding using Plot.ly Dash[4]

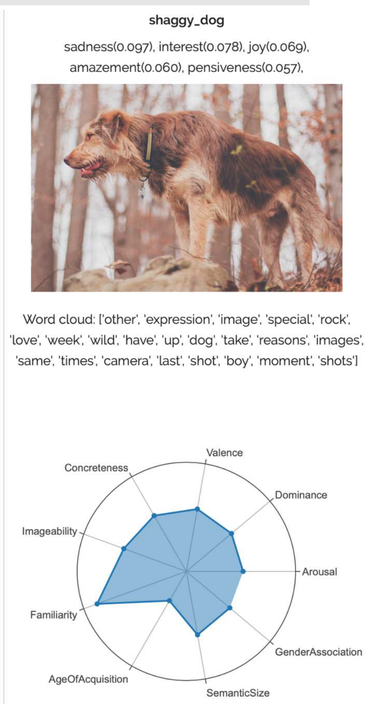
- Allows for interactive analysis
- Heat-map grading highlights psycholinguistic ratings



Filtering function allows comparing different nouns with same adjective, and vice-versa

Detailed view provides information per-image

- Psycholinguistic grounding as radar graph
- Meta-data including ANP, word-cloud and sentiment annotations



Designed to analyze the relationship between textual and visual sentiments across concepts

Discussion

Research direction 1: Perceptual clusters

- Use clustering on spatial embedding
- Find perceptually related images

Filtering per noun shows overlapping adjectives, suggesting similarly perceived images

Research direction 2: Visual characteristics

- Compare visual characteristics of related images

How do e.g. *lowly imageable* images look different from *highly imageable* ones?

Next steps for demo

More sophisticated approach for per-image ratings

- Use whole psycholinguistic vectors instead of average
- Include visual characteristics

Extend browser for multiple languages

- MVSO[1] is available for other languages, but Glasgow Norms[2] only contains English words
- Cross-language comparison can find cultural differences

[1] Jou et al. Visual Affect Around the World: A Large-scale Multilingual Visual Sentiment Ontology. ACMMM2015.
[2] Scott et al. The Glasgow Norms: Ratings of 5,500 words on nine scales. Behav Res Meth 2019.
[3] L. McInnes and J. Healy. UMAP: Uniform Manifold Approximation and Projection for dimension reduction. arXiv 2018.
[4] Plotly Technologies Inc. Collaborative data science. Montréal, QC, 2015. (https://plot.ly/dash)

MMM 2020