Participatory Art Museum: Collecting and Modeling Crowd Opinions

Xiaoyu Zeng[1] **and Ruohan Zhang**[2]

TEXAS
The University of Texas at Austin

[1]Department of Art and Art History, The University of Texas at Austin, United States
[2]Department of Computer Science, The University of Texas at Austin, United States
*{edith.xiaoyu.zeng, zharu}@utexas.edu

Abstract

We collect public opinions on museum artworks using online crowdsourcing techniques. We ask two research questions. First, do crowd opinions on artworks differ from expert interpretations? Second, how can the museum manage large amount of crowd opinions, such that users can efficiently retrieve useful information? We address these questions through opinion modeling via semantic embedding and dimension reduction. We further demonstrate how to use dimension reduction algorithms to visualize the data.

Introduction

- Traditional art museums: A one-way communication in which non-expert visitors are perceived as a passive body of recipient and educatee.
- The Participatory Art Museum movement: To encourage public engagement, ranging from discussion-based gallery tours to crowdsourcing tagging and online transcription tasks [Simon, 2010, Ridge, 2014].
- Research goal: To facilitate the movement using recent crowdsourcing and natural language processing techniques.
- Research questions:
  1. What do non-experts think about museum artworks?
     – How are their opinions different from experts'?
  2. How can museum curators and/or artists efficiently collect opinions from the public?
     – Given a large amount of public opinions, how do they efficiently extract useful information?

Data Collection

- 21 artworks with distinct periodic styles
  – Famous artists' less well-known works
- 2,116 paragraphs from Amazon Mechanical Turk
  – Write a couple sentences to describe the mood of this image and your interpretation: how does this image make you feel? What do you think it is about? Why?
- 832 paragraphs from Wikipedia
- 603 paragraphs from museum catalogs or college-level textbooks

Opinion Clouds: Semantic Embedding and Visualization

Algorithm: Semantic Embedding and t-SNE

1. Break corpus down to sentence level.
2. Using tf-idf index, remove k least important tokens.
3. Calculate word2vec vectors for all tokens [Mikolov et al., 2013].
4. Average word vectors within each sentence to get a vector that represent the opinion of the sentence.
5. Project sentence vectors into 2D space using t-SNE [Van Der Maaten, 2014].

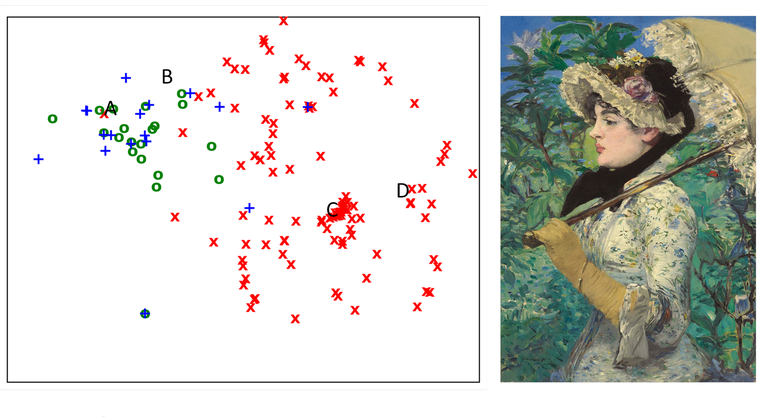

Figure 1: Opinion cloud for Édouard Manet's *Jeanne (Spring)*. Red: crowd responses. Green: Wikipedia. Blue: museum catalog from the J. Paul Getty Museum. The crowd opinions and expert opinions evidently form distinct clusters.

- For most artworks, the differences between the crowd and the experts are evident.
  – Differences on interests and perspectives:
    A (Wikipedia): *"Today, these are considered watershed paintings that mark the genesis of modern art."*
    C (Crowd): *"The woman in the painting is dressed immaculately and clearly on her way to be seen while taking a stroll on a lovely day."*
  – Differences on subjective feelings:
    B (Museum catalog): *"The painting's sensual handling and bright, vibrant palette evoke the pleasures of the season it celebrates."*
    D (Crowd): *"The mood I get from this image is one of resignation and reservation."*

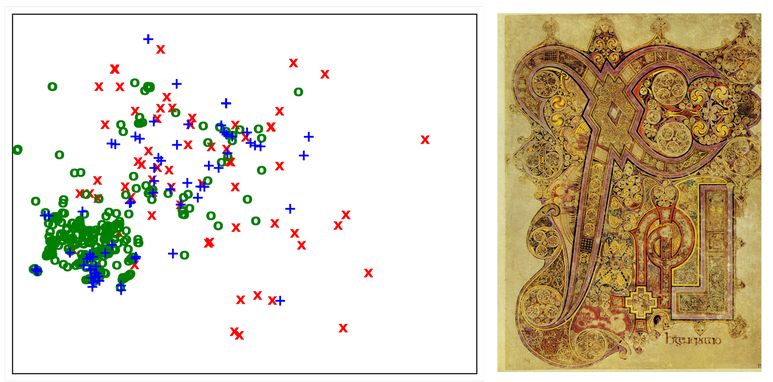

Figure 2: Opinion cloud for *Chi-Rho from the Book of Kells*, circa 800. Red: crowd responses. Green: Wikipedia. Blue: museum catalogs and books. The crowd opinions and expert opinions mix.

- This technique also presents an interesting way to view multiple artworks in semantic space of public interpretations:

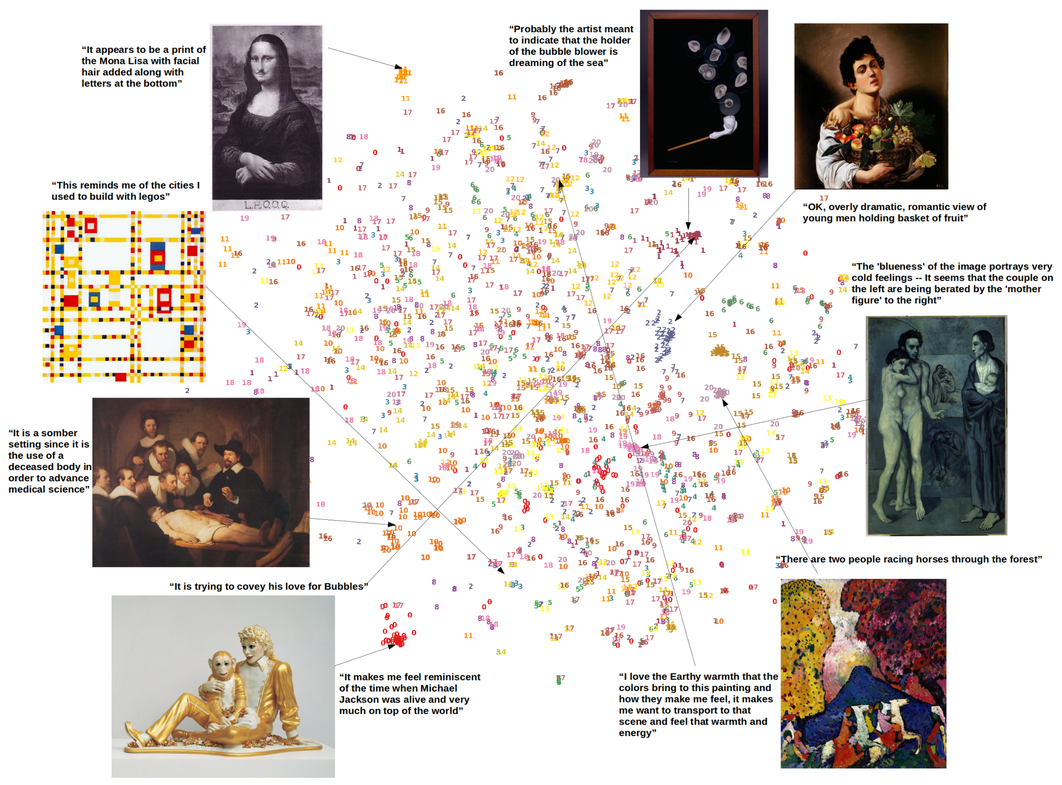

Figure 3: Opinion clouds for all artworks. Each point indicates the position of a crowd opinion in the projected semantic space. Different colors represent opinions on different artworks. Artworks: Marcel Duchamp, *L.H.O.O.Q.*; Piet Mondrian, *Broadway Boggie Woogie* ; Rembrandt van Rijn's *The Anatomy Lecture of Dr. Nicolae Tulp*; Jeff Koons, *Michael Jackson and Bubbles*; Joseph Cornell, *Soap Bubble Set*; Michelangelo Merisi di Caravaggio, *Boy with A Basket of Fruit*; Pablo Picasso, *La Vie*; Wassily Kandinsky, *Blue Mountain*.

Conclusions and Future Work

- We are in progress of collecting more data from crowdworkers, books, and online resources.
- What is a good quantitative evaluation metric to determine the quality of semantic embedding?
  – The difficulty is that the artwork interpretations are extremely subjective and open-ended.
- Fine-tune or retrain the word2vec model.
- Better sentence level semantic embedding methods.
- More specific analysis, e.g., sentiment analysis.

References

[Mikolov et al., 2013] Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., and Dean, J. (2013). Distributed Representations of Words and Phrases and Their Compositionality. In *NIPS*, pages 3111–3119.

[Ridge, 2014] Ridge, M. M. (2014). *Crowdsourcing Our Cultural Heritage*. Ashgate Publishing, Ltd.

[Simon, 2010] Simon, N. (2010). *The Participatory Museum*. Museum 2.0.

[Van Der Maaten, 2014] Van Der Maaten, L. (2014). Accelerating t-sne Using Tree-Based Algorithms. *JMLR*, 15(1):3221–3245.