Participatory Art Museum: Collecting and Modeling Crowd Opinions

Xiaoyu Zeng[1] and **Ruohan Zhang**[2]
[1]Department of Art and Art History, The University of Texas at Austin, Austin, TX, 78712, USA
[2]Department of Computer Science, The University of Texas at Austin, Austin, TX, 78712, USA
[1]edith.xiaoyu.zeng@utexas.edu (512-998-4072), [2]zharu@utexas.edu
https://goo.gl/JyZYbq

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 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.

Introduction

As a secular temple of high culture, the traditional art museum has been dominated by expert-generated contents that help visitors to "decipher" fine art: labels and docent-led tours, while providing necessary guidance, often establish a one-way communication in which non-expert visitors are perceived as a passive body of recipient and educatee. With the rise of the general public as an anonymous "crowd" in the past decade, many professionals began to re-evaluate the role of non-expert audiences in museum experiences (Antrobus 2010). More art museums are now encouraging means for public engagement, ranging from discussion-based gallery tours to crowdsourcing tagging and transcription tasks online (Simon 2010; Ridge 2014).

The main purpose of our work is to facilitate the Participatory Museum movement (Simon 2010) through introducing recent technologies from crowdsourcing and natural language processing. We propose that collecting interpretations of artworks from the crowd can benefit both experts and non-experts. For the connoisseurs, crowd judgments help to correct homogeneity-generated biases by introducing diverse perspectives. For the non-expert viewers, the crowd interpretation presents an interesting alternative that invites an open and dynamic dialogue.

The lack of crowd participation and difficulties in organizing collected opinions remain issues in many art museums. In this work, we propose to use current online crowdsourcing tools to collect crowd opinions on artworks. We then model and visualize collected data as **opinion clouds** in semantic space. There are two motivations for doing this. First, the traditional practice to have museum staff read and summarize visitor feedback does not scale to the amount of data collected via online crowdsourcing. Second, museum curators may look for different things in the data: some are interested in popular opinions, while others look for outliers – potentially innovative ideas that contribute to art historical analyses. Modeling and visualizing crowd opinion as a distribution in semantic space can make information easily accessible for both needs.

Data

We crowdsourced non-expert interpretations of 21 artworks by artists with distinct periodic styles. To reduce prejudice, all artwork used for this experiment were famous artists' relatively less well-known works with factual information hidden from participants. We ran the task on Amazon Mechanical Turk for its large and representative subject pool (Paolacci, Chandler, and Ipeirotis 2010). Each participant received an average of $0.096 for answering these questions:

> 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?

We collected a total of 2116 sentences from crowd. For all 21 works, we also collected corresponding articles from Wikipedia pages. Then, we asked art history experts to provide canonical articles from books and museum catalogs. This results in a total of 832 sentences from Wikipedia, and 603 sentences from scholarly writings.

Semantic Embedding and Visualization

Our goal is to model the crowd opinions as distributions in semantic space, so that one can tell the distance between expert and crowd opinions, and can then easily identify popular interpretations by clusters and unique interpretations by outliers. We follow Algorithm 1 to produce the opinion clouds. The word2vec model (Mikolov et al. 2013) we use is a pre-trained one on Google News corpus. The t-SNE algorithm (Van Der Maaten 2014) is initialized using PCA.

Figure 1 Left shows opinion clouds for crowd responses to 5 artworks (Marcel Duchamp, *L.H.O.O.Q.*; 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*), where different colors represent different

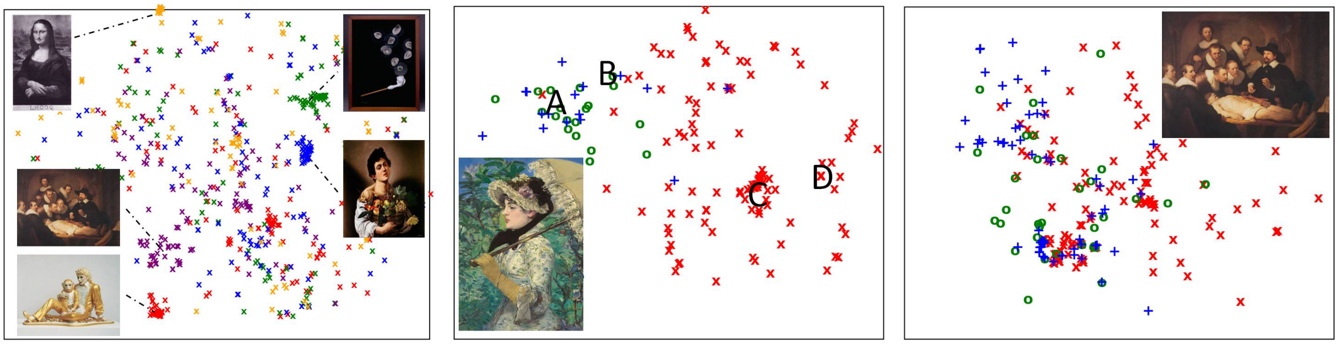

Figure 1: **Left:** Opinion clouds for 5 artworks. Each point indicates the position of a crowd opinion in the projected semantic space. Different colors represent opinions on different artworks. **Middle:** Opinion cloud for Édouard Manet's *Jeanne (Spring)*. Red: crowd responses. Green: Wikipedia. Blue: analyses from books and museum labels. Crowd and experts form distinct opinion clouds. **Right:** Opinion cloud for *The Anatomy Lecture of Dr. Nicolae Tulp*. Crowd and expert opinions mix.

Algorithm 1 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.
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.

artworks. Intuitively, clusters formed for different artworks suggest that the embedding and projection are reasonable.

Figure 1 Middle shows result for *Jeanne (Spring)*, where crowd and experts opinions form distinct clusters. The discrepancy reveals different interests and perspectives: while experts focus more on historical significance of an artwork, the crowd focus more on its content. For example:

A (Wikipedia): *"Today, these are considered watershed paintings that mark the genesis of modern art."*[1]
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."*

The difference could also arise from subjective feelings:

B (Museum catalog): "*The painting's sensual handling and bright, vibrant palette evoke the pleasures of the season it celebrates."*[2]
D (Crowd): *"The mood I get from this image is one of resignation and reservation."*

Figure 1 Right shows an example where crowd and experts agree more. Using opinion clouds, museum staff or visitors can visualize these differences as spatial distance, and quickly search for their opinions of interest.

By examining results for all artworks, we find that crowd opinions typically cover a larger semantic space than experts'. For many artworks, significant amount of crowd opinions are distant from the cluster of expert opinions. Results for all the artworks can be found on our project page.

[1] "Édouard Manet." Wikipedia. https://goo.gl/Lxl6p4.
[2] "Jeanne." The J. Paul Getty Museum. https://goo.gl/MIVKvo.

Conclusions and Future Work

In this work, we propose a method for museums to collect and model large amount of public opinions about artworks. Our results show that in semantic space, crowd opinions typically differ from expert opinions. Through our interface, museum professionals and visitors can more efficiently find opinions of interest.

For future work, it is desirable to have a quantitative evaluation metric to determine the quality of semantic embedding. The main challenge is that the artwork interpretations are extremely subjective and open-ended. We plan to recruit art historians to manually label each crowd opinion as novel or not, given Wikipedia and scholarly writings as reference. We are in the progress of collecting more data, so that we can fine-tune or retrain the word2vec model for our domain. In addition, averaging word vectors to get sentence semantics can be replaced with more sophisticated method, from weighted average methods to doc2vec models (Le and Mikolov 2014). We can compare different embedding methods once the evaluation metric is developed.

References

Antrobus, C. 2010. From Audiences to Users: Changing Art Galleries and Museums Together. *London: NESTA, Clore Leadership Programme.*

Le, Q. V., and Mikolov, T. 2014. Distributed Eepresentations of Sentences and Documents. In *ICML*, volume 14, 1188–1196.

Mikolov, T.; Sutskever, I.; Chen, K.; Corrado, G. S.; and Dean, J. 2013. Distributed Representations of Words and Phrases and Their Compositionality. In *NIPS*, 3111–3119.

Paolacci, G.; Chandler, J.; and Ipeirotis, P. G. 2010. Running Experiments on Amazon Mechanical Turk. *Judgment and Decision making* 5(5):411–419.

Ridge, M. M. 2014. *Crowdsourcing Our Cultural Heritage*. Ashgate Publishing, Ltd.

Simon, N. 2010. *The Participatory Museum*. Museum 2.0.

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