

# Modeling Aesthetics and Emotions in Visual Content: From Vincent van Gogh to Robotics and Vision

James Z. Wang

College of Information Sciences and Technology  
The Pennsylvania State University  
University Park, Pennsylvania  
jwang@ist.psu.edu

## ABSTRACT

As inborn characteristics, humans possess the ability to judge visual aesthetics, feel the emotions from the environment, and comprehend others' emotional expressions. Many exciting applications become possible if robots or computers can be empowered with similar capabilities. Modeling aesthetics, evoked emotions, and emotional expressions automatically in unconstrained situations, however, is daunting due to the lack of a full understanding of the relationship between low-level visual content and high-level aesthetics or emotional expressions. With the growing availability of data, it is possible to tackle these problems using machine learning and statistical modeling approaches. In the talk, I provide an overview of our research in the last two decades on data-driven analyses of visual artworks and digital visual content for modeling aesthetics and emotions.

First, I discuss our analyses of styles in visual artworks [2]. Art historians have long observed the highly characteristic brushstroke styles of Vincent van Gogh and have relied on discerning these styles for authenticating and dating his works. In our work, we compared van Gogh with his contemporaries by statistically analyzing a massive set of automatically extracted brushstrokes. A novel extraction method is developed by exploiting an integration of edge detection and clustering-based segmentation. Evidence substantiates that van Gogh's brushstrokes are strongly rhythmic.

Next, I describe an effort to model the aesthetic and emotional characteristics in visual contents such as photographs [1, 3, 4, 6]. By taking a data-driven approach, using the Internet as the data source, we show that computers can be trained to recognize various characteristics that are highly relevant to aesthetics and emotions. Future computer systems equipped with such capabilities are expected to help millions of users with unimagined ways.

Finally, I highlight our research on automated recognition of bodily expression of emotion [5, 7]. We propose a scalable and reliable crowdsourcing approach for collecting in-the-wild perceived emotion data for computers to learn to recognize the body languages of humans. Comprehensive statistical analysis revealed many interesting insights from the dataset. A system to model the emotional expressions based on bodily movements, named ARBEE

(Automated Recognition of Bodily Expression of Emotion), has also been developed and evaluated.

## CCS CONCEPTS

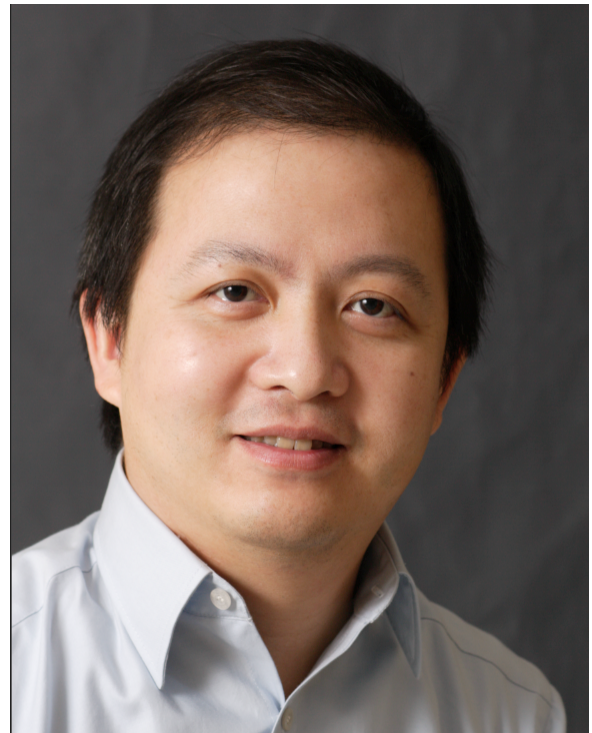
• **Applied computing** → *Fine arts*; **Psychology**; • **Computing methodologies** → *Scene understanding*.

## KEYWORDS

Visual aesthetics; evoked emotions; bodily expressed emotion understanding; oil paintings; visual artworks

## ACM Reference Format:

James Z. Wang. 2020. Modeling Aesthetics and Emotions in Visual Content: From Vincent van Gogh to Robotics and Vision. In *Joint Workshop on Aesthetic and Technical Quality Assessment of Multimedia and Media Analytics for Societal Trends (ATQAM/MAST'20)*, October 12–16, 2020, Seattle, WA, USA. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3423268.3423586>



(Photo by Justina J. Wang)

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

ATQAM/MAST'20, October 12–16, 2020, Seattle, WA, USA

© 2020 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-8154-3/20/10.

<https://doi.org/10.1145/3423268.3423586>

## BIOGRAPHY

James Z. Wang is a Professor of Information Sciences and Technology at The Pennsylvania State University. He received the bachelor's degree in mathematics and computer science *summa cum laude* from University of Minnesota, and the MS degree in mathematics, the MS degree in computer science, and the PhD degree in medical information sciences, all from Stanford University.

His research primarily focuses on the modeling of objects, concepts, aesthetics, and emotions in big visual data, with applications in biomedicine, visual art, meteorology, psychology, and social media. He was a recipient of the National Science Foundation Career Award (2004) and Amazon Research Awards (2018 and 2019).

He was a Visiting Professor with the Robotics Institute, Carnegie Mellon University, from 2007 to 2008. He was a Lead Special Section Guest Editor for *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2008). In 2011 and 2012, he served as a Program Manager at the Office of the Director of the National Science Foundation where he managed international collaboration programs covering all sciences and engineering disciplines.

## ACKNOWLEDGMENTS

The author would like to thank his collaborators and students for valuable discussions and carrying out parts of the research. This material is based upon work supported by the National Science Foundation under Grant Nos. 1921783, 1110970, and 0347148, and the

Amazon Research Awards program (2018 and 2019). The research is also supported by the National Endowment for the Humanities and Nvidia Corporation. This research used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation Grant No. ACI-1548562.

## REFERENCES

- [1] Ritendra Datta, Dhiraj Joshi, Jia Li, and James Z Wang. 2006. Studying aesthetics in photographic images using a computational approach. In *Proceedings of the European Conference on Computer Vision*. Springer, 288–301.
- [2] Jia Li, Lei Yao, Ella Hendriks, and James Z Wang. 2011. Rhythmic brushstrokes distinguish van Gogh from his contemporaries: findings via automated brushstroke extraction. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34, 6 (2011), 1159–1176.
- [3] Xin Lu, Reginald B Adams, Jia Li, Michelle G Newman, and James Z Wang. 2017. An investigation into three visual characteristics of complex scenes that evoke human emotion. In *Proceedings of the 7th International Conference on Affective Computing and Intelligent Interaction*. IEEE, 440–447.
- [4] Xin Lu, Poonam Suryanarayan, Reginald B Adams Jr, Jia Li, Michelle G Newman, and James Z Wang. 2012. On shape and the computability of emotions. In *Proceedings of the 20th ACM international conference on Multimedia*. 229–238.
- [5] Yu Luo, Jianbo Ye, Reginald B Adams, Jia Li, Michelle G Newman, and James Z Wang. 2020. ARBEE: Towards automated recognition of bodily expression of emotion in the wild. *International Journal of Computer Vision* 128, 1 (2020), 1–25.
- [6] Lei Yao, Poonam Suryanarayan, Mu Qiao, James Z Wang, and Jia Li. 2012. OSCAR: On-site composition and aesthetics feedback through exemplars for photographers. *International Journal of Computer Vision* 96, 3 (2012), 353–383.
- [7] Jianbo Ye, Jia Li, Michelle G Newman, Reginald B Adams, and James Z Wang. 2017. Probabilistic multigraph modeling for improving the quality of crowdsourced affective data. *IEEE Transactions on Affective Computing* 10, 1 (2017), 115–128.