

What makes a picture memorable?

Aude Oliva, Phillip Isola, Aditya Khosla, and Wilma A. Bainbridge

Computer vision techniques predict whether an image will be remembered or forgotten.

‘Mental sports’ have become a new trend in self-improvement, with video games designed to improve mental fitness. At the World Memory Championships, athletes compete to recall massive amounts of information; contestants must memorize and recall sequences of abstract images and the names of people whose faces are shown in photographs. While these tasks might seem challenging, our research suggests that images that possess certain properties are memorable. Our findings can explain why we have all had some images stuck in our minds, but ignored or quickly forgotten others.

Although image memorability seems subjective and hard to quantify, our recent work¹⁻⁵ shows that it is not an inexplicable phenomenon. We found that visual memorability is largely intrinsic to the image and reproducible across a diverse population. This means that despite varied experiences, individuals tend to remember and forget the same images. Using experimental data detailing the types of images people remember or quickly forget, we developed an algorithm that automatically predicts whether an image will be memorable.

To determine the intrinsic features that make an image memorable, we first asked 665 individuals to participate in a computer memory game. During each level of the game, for up to 30 levels, participants viewed a stream of images and then pressed the space bar whenever they saw one of those images repeated in a subsequent sequence. In total, the image database contained 2222 repeated images and 8220 unrepeated images that included faces, interior-design photos, nature scenes, streetscapes, and others. We found that photographs with people or central objects were memorable, whereas landscapes—that one might expect to be memorable—were among the most forgettable (see Figure 1).¹

Next, we assigned a ‘memorability score’ to each image, which was defined as the percent of correct detections by participants in the study. On average, 78 participants scored each

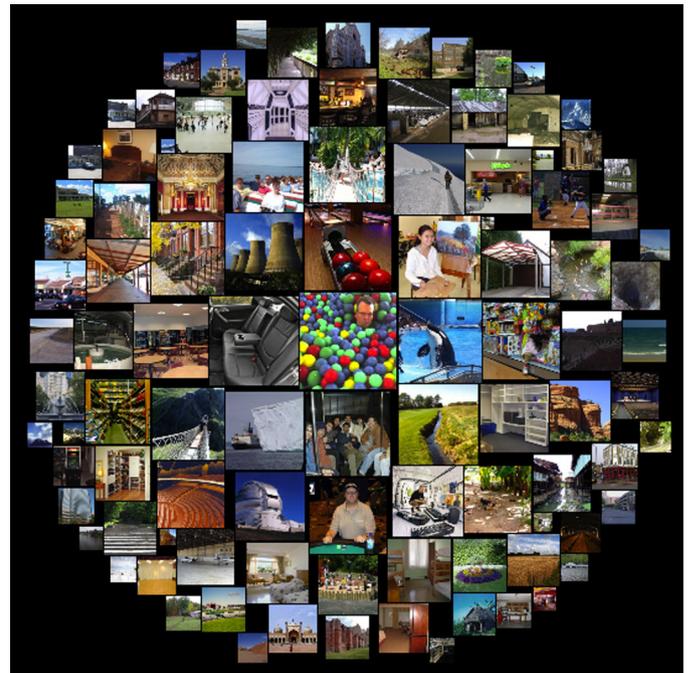


Figure 1. A sample of the photos used to develop an algorithm that predicts image memorability.¹ The photos are scaled according to their experimentally measured memorability, with the most memorable photos appearing larger and in the center.

image. We then investigated the features of the more memorable images, including color, object statistics (e.g., number of objects or amount of space occupied by objects in the image), object semantics (e.g., what type of object appeared in an image, such as an animal or car), and scene semantics (e.g., the place the image represented, such as a kitchen or landscape). Using computer vision techniques,¹ we developed an image-ranking algorithm to automatically predict the memorability of images. To do this, we trained a support vector regressor to map from features to memorability scores using only features algorithmically extracted from the images. The algorithm learned from the memorability scores calculated from the memory game.

Continued on next page

We used half of the images in the database to train the algorithm and tested its performance with the remaining half. The algorithm correctly identified images with people as most memorable, indoor scenes and large objects as slightly less memorable, and outdoor landscapes as the least memorable.¹ We also found that atypicality and aesthetic beauty attributes did not explain much of the variation we observed in scene memorability.² For instance, landscapes—despite being beautiful—were often forgotten, whereas generic photos of social scenes—for example, a dinner party, subway car, office—were almost always remembered.

We have furthered this work by developing a framework for predicting image memorability that accounts for how the memorability of image regions and different types of features fade over time.³ Research on human visual memory has shown that observers typically remember visual details attached to objects that have a specific semantic label or distinctive interpretation.⁶ For example, observers remember different types of cars by tagging each car with a different brand name, but often confuse different types of apple that differ by color.⁷ This suggests that different features, objects, and regions in an image might have different effects on memorability. To examine this idea, we developed a method for creating automated memorability maps that display which local information in an image is memorable and which is forgettable.³

Predicting image memorability lends itself to a wide variety of applications. We live in an age of data deluge, and memorability prediction could provide a method for summarizing and condensing the onslaught of visual data we encounter. For example, a photo album could be summarized using a few memorable photographs that convey the overall story. In education, textbook diagrams could be created to stick in students' minds, teachers could select memorable examples to illustrate concepts, and memorable cartoons could be used as mnemonic aids to make learning easier. Memorability could also find applications in user-interface design. For example, memorable icons could clarify a messy desktop, and mnemonic labels could be attached to pill containers or entryways in retirement homes. In addition, understanding memorability might lead to intelligent systems that preferentially store information based on its memorability, making sure to prioritize important information that humans will likely forget.

Memorability research could be especially applicable within the domain of face memorability.⁴ As social creatures, faces are a key part of our lives, and we often struggle to make a strong first impression and to be memorable. Indeed, in future work, we will be looking into algorithms that enable us to modify a portrait in subtle ways to enhance or reduce its memorability,

while maintaining other facial traits like identity, attractiveness, and facial expression. Perhaps within the next few years smartphone applications will be developed that can select the most memorable photograph for a profile picture or that can help you apply makeup to boost your memorability. Additionally, therapeutic technologies could be realized to train people to focus on key memorability-determining facial features to help those with social processing and memory-related disorders, such as autism, prosopagnosia, or Alzheimer's disease.

A common factor across disciplines, memorability represents a fairly general quantification of the utility of visual information. Memorability varies from image to image, yet remains largely constant across multiple people viewing the same picture.^{1,2,4,5} We have found that we can predict whether people will remember an image based on its components.^{1,3} In forthcoming work, we are extending the paradigm beyond photographs; we are investigating the intrinsic memorability of data visualizations, artwork, and even English words. With this base understanding of memorability in place, our work might encourage machine vision and artificial intelligence researchers to consider not only what the world is about, but what humans consider meaningful: what they remember.

The authors thank the National Science Foundation (Grant No. 1016862), Google, and Xerox for partly supporting this research.

Author Information

Aude Oliva, Phillip Isola, Aditya Khosla, and Wilma A. Bainbridge

Massachusetts Institute of Technology (MIT)
Cambridge, MA

Aude Oliva is a principal research scientist in the Computer Science and Artificial Intelligence Laboratory at MIT. Her research lies at the interface between human perception, cognition, neuroscience, and computer vision. She is the recipient of a National Science Foundation CAREER Award.

Phillip Isola is a graduate student in the Department of Brain and Cognitive Sciences at MIT. He works on human and computer vision, and is the recipient of a National Science Foundation Graduate Research Fellowship.

Continued on next page

Aditya Khosla is a graduate student in the Electrical Engineering and Computer Science Department at MIT and is interested in computer vision and machine learning. He is the recipient of the Facebook Fellowship 2013–2014.

Wilma A. Bainbridge is a graduate student in the Department of Brain and Cognitive Sciences at MIT working on the human neuroscience of perception and memory. She is the recipient of a National Defense Science and Engineering Graduate Fellowship.

References

1. P. Isola, J. Xiao, A. Torralba, and A. Oliva, *What makes an image memorable?*, **IEEE Conf. Comput. Vision Pattern Recognit.**, pp. 145–152, 2011. doi:10.1109/CVPR.2011.5995721
2. P. Isola, D. Parikh, A. Torralba, and A. Oliva, *Understanding the intrinsic memorability of images*, **Adv. Neural Inf. Process. Syst.** **24**, pp. 2429–2437, 2011.
3. A. Khosla, J. Xiao, A. Torralba, and A. Oliva, *Memorability of image regions*, **Adv. Neural Inf. Process. Syst.** **25**, pp. 305–313, 2012.
4. W. A. Bainbridge, P. Isola, I. Blank, and A. Oliva, *Establishing a database for studying human face photograph memory*, **Proc. 34th Annu. Meeting Cognit. Sci. Soc.**, pp. 1302–1307, 2012.
5. A. Khosla, J. Xiao, P. Isola, A. Torralba, and A. Oliva, *Image Memorability and Visual Inception*, **SIGGRAPH Asia 2012 Tech. Briefs** **35**, pp. 1–4, 2012. doi:10.1145/2407746.2407781
6. T. F. Brady, T. Konkle, G. A. Alvarez, and A. Oliva, *Visual long-term memory has a massive storage capacity for object details*, **Proc. Nat'l Acad. Sci.** **105** (38), pp. 14325–14329, 2008. doi:10.1073/pnas.0803390105
7. T. Konkle, T. F. Brady, G. A. Alvarez, and A. Oliva, *Conceptual distinctiveness supports detailed visual long-term memory for real-world objects*, **J. Exp. Psychol. Gen.** **139** (3), pp. 558–578, 2010. doi:10.1037/a0019165