What makes an image memorable?

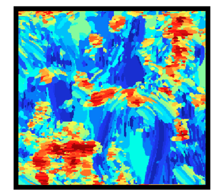
Phillip Isola, Jianxiong Xiao, Antonio Torralba, Aude Oliva, MIT

Questions

Intrinsic memorability?

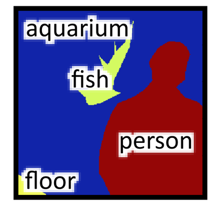
Are some images consistently more memorable than others, even across different observers and contexts?

Can we predict it?

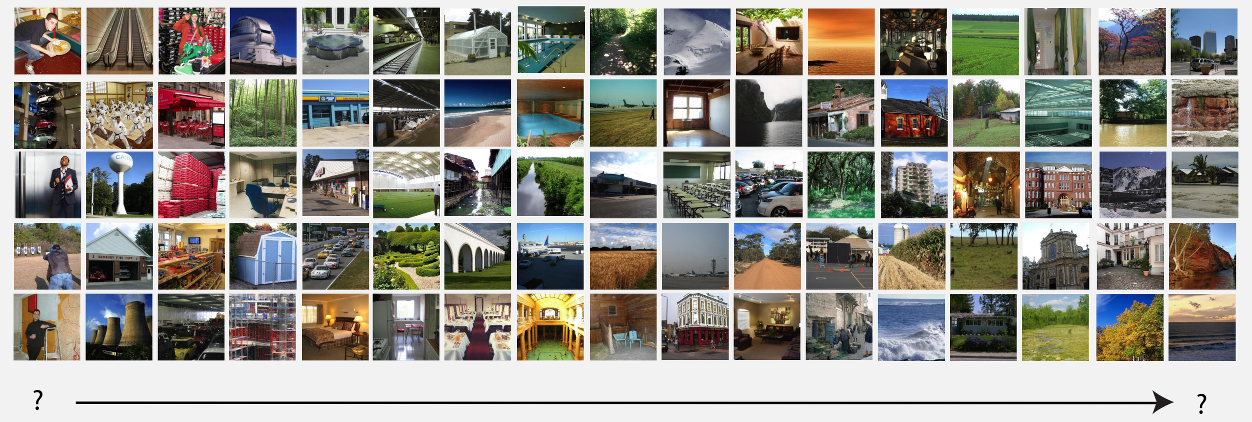


Can we automatically predict an image's memorability from its features?

What image content matters?



What image content is driving memorability?

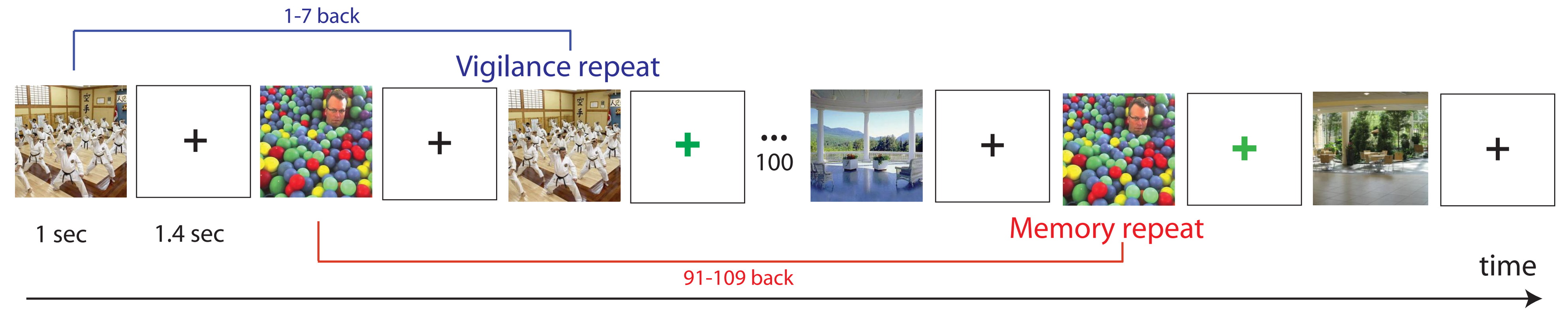


Database: 10442 photographs sampled from SUN database (*Xiao et al. 2010*).
2222 target images scored.
Images annotated with LabelMe (*Russell et al. 2008*).

Memorability = probability of correctly detecting a repeat after a single view of an image in a long stream.

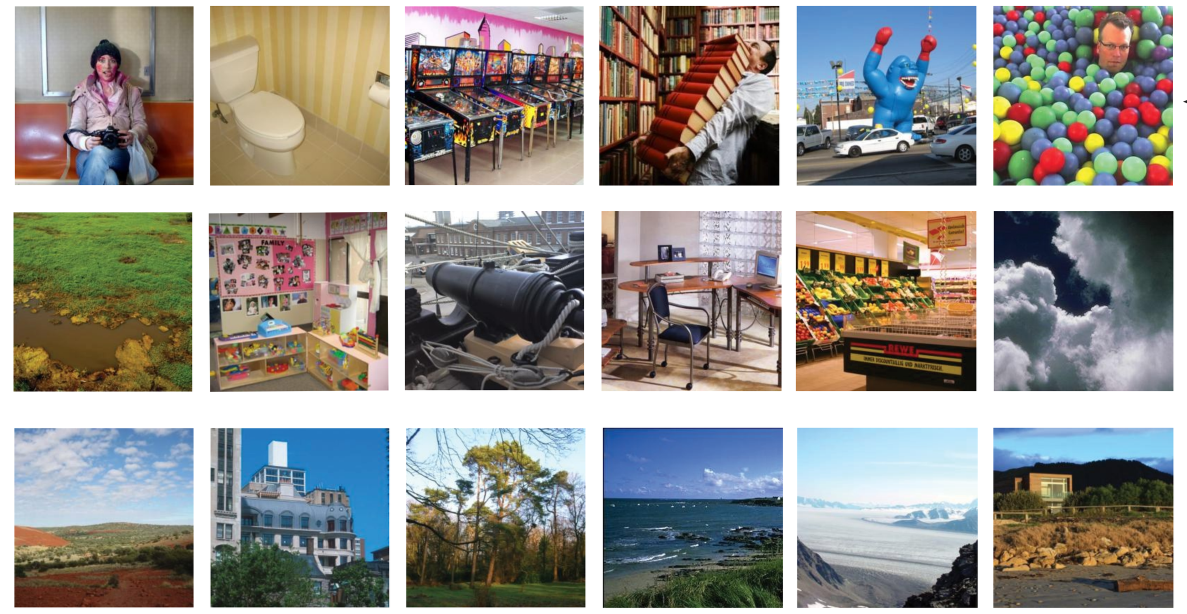
Memory Game

665 participants on Amazon's Mechanical Turk.



Memorability (hit rate)
Mean: 67.5%
SD: 13.6%

False alarm rate
Mean: 10.7%
SD: 7.6%

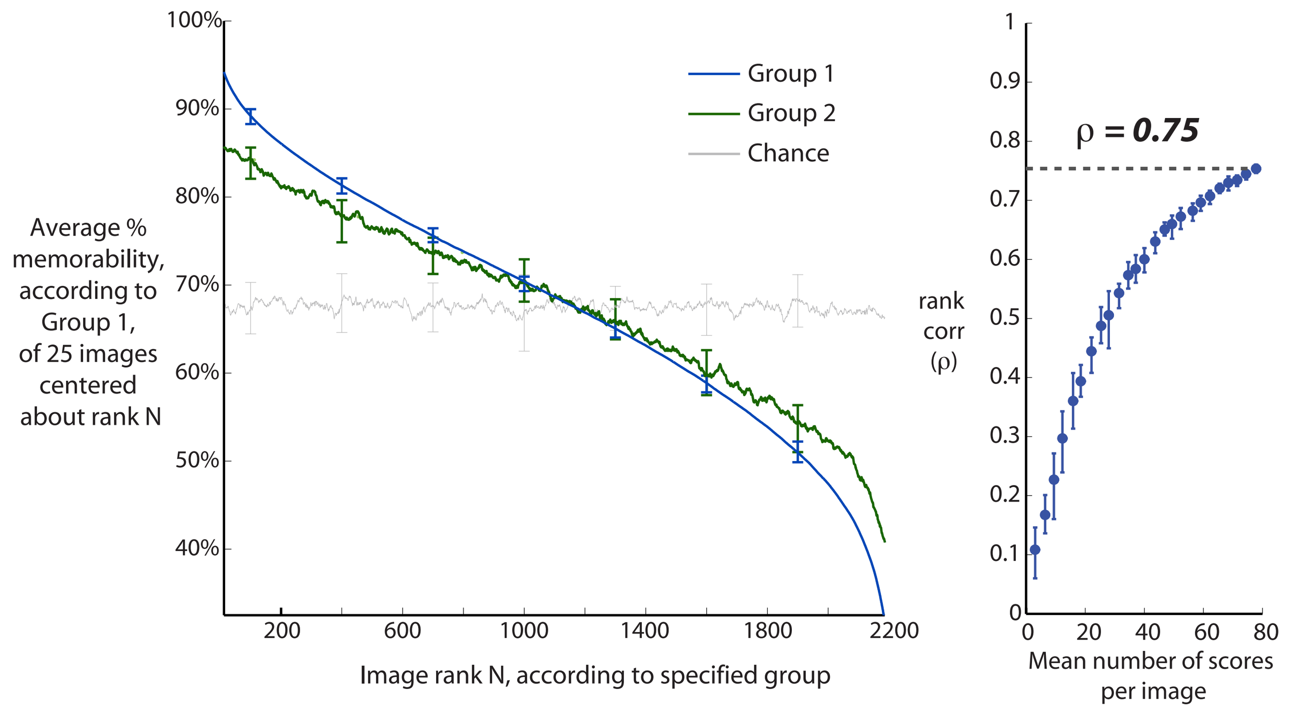


Memorable
Hit rate: 67/70
False alarm rate: 4/80

Average
Hit rate: 59/81
False alarm rate: 7/92

Forgettable
Hit rate: 21/68
False alarm rate: 3/82

Intrinsic memorability?



Inter-subject consistency

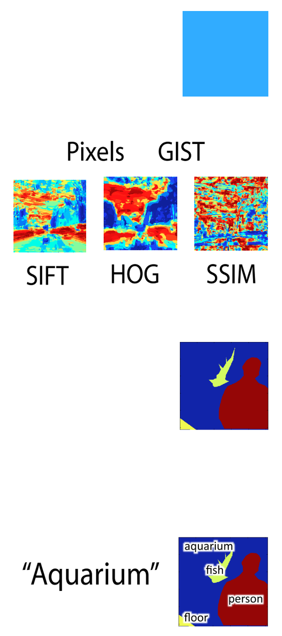
The memorability we measured is a stable property of an image that is shared across many observers and contexts.

Sizeable effect

Some images are consistently much more memorable than others.

What classes of information predict memorability?

Prediction algorithm: SVM Regression with non-linear kernels on the following feature sets.

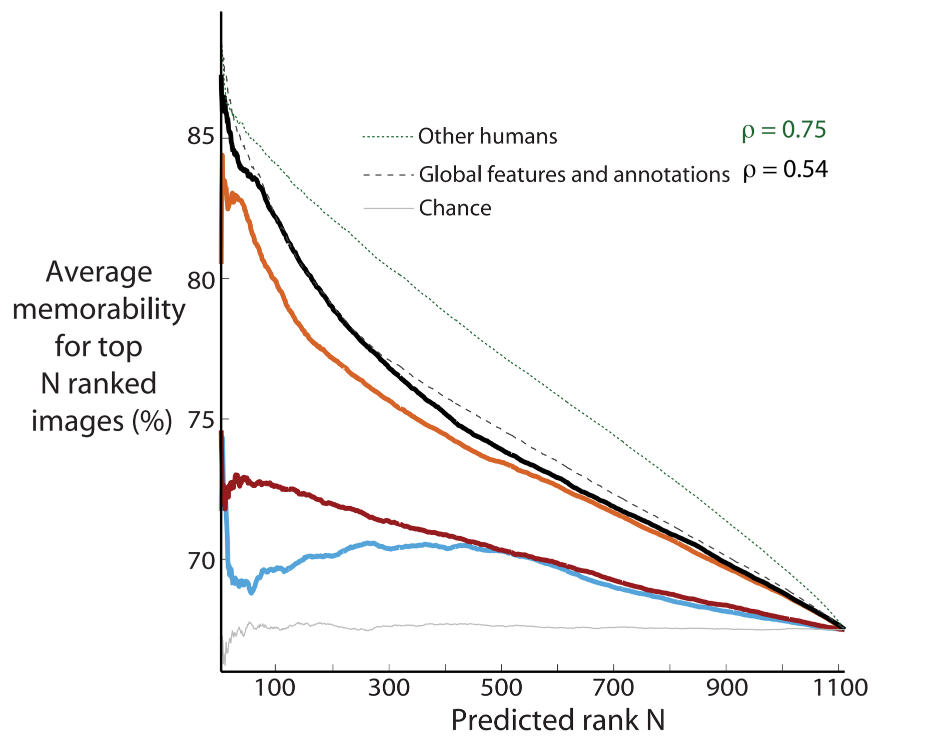


1) Simple image stats? $\rho = 0.16$
e.g. mean hue, brightness, number of objects

2) Global image features? $\rho = 0.46$
pixel histograms, GIST, SIFT, HOG, SSIM

3) Object segmentation statistics? $\rho = 0.20$
histograms of object segments counts, sizes, and rough positions

4) Object and scene semantics? $\rho = 0.50$
number, size, and rough position of each object class; scene category



Predictable from image

Predicted memorable
Predicted average
Predicted forgettable

Rich features necessary

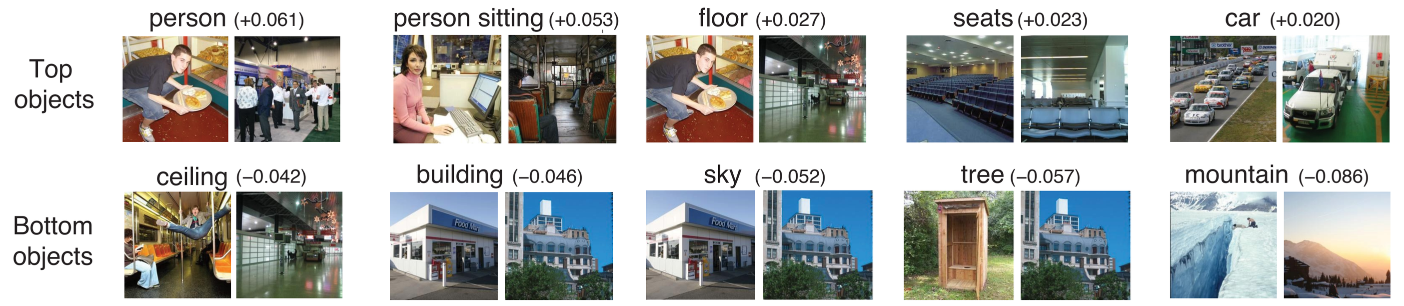
Simple features not enough.
Multivariate global features do better.
Semantic annotations do best.

What content makes an image memorable?

Object score = (prediction when object included in image's feature vector) - (prediction when object removed)

Objects shaded according to object score (computed per image)

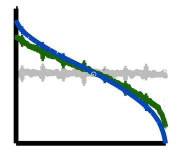
Common objects ranked according to object score (averaged across images)



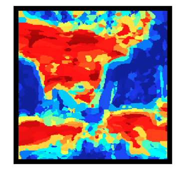
Common scenes ranked according to their average memorability

Conclusions

Stable intrinsic memorability



Predictable from rich image features



People, close-ups, and human-scale objects predict memorability

This work is in press as Isola, Xiao, Torralba, and Oliva, CVPR 2011, and online at web.mit.edu/phillipi/Public/WhatMakesAnImageMemorable