Memorability of Image Regions

Aditya Khosla Jianxiong Xiao Antonio Torralba Aude Oliva

Massachusetts Institute of Technology

CSAIL

RememberMe.csail.mit.edu

Overview

Background:

Image memorability is an ***intrinsic*** and ***measurable*** property! [Isola et al, CVPR 2011]

Goal:

Find memorability of image regions automatically without manual annotation

Method:

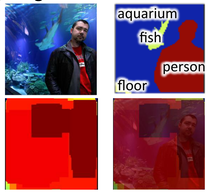
An ***interpretable*** model of memory as a noisy process composed of image regions

Advantages over manual annotation:

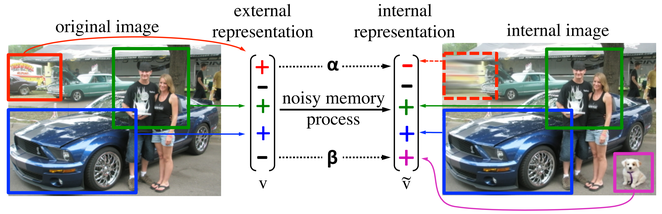
- Manual annotation of segments is expensive
- Granularity of annotations is not well defined

Result:

- Same prediction performance as using ground truth image segments!
- Automatically generated memorability maps that correspond well with manually annotated images



Algorithm



Expected distance of image j:

$$E(D_j|v_j) = \sum_{i=1}^{N} \alpha_i^{v_j(i)} * \beta_i^{1-v_j(i)} = v_j^T \vec{\alpha} + (\neg v_j)^T \vec{\beta} \propto_{rank} -s_j$$

Expected distance of all images:

$$E(D|v) = \begin{pmatrix} — v_1^T — & — \neg v_1^T — \\ \vdots & \vdots \\ — v_M^T — & — \neg v_M^T — \end{pmatrix} \cdot \begin{pmatrix} \vec{\alpha} \\ \vec{\beta} \end{pmatrix} \propto_{rank} -\vec{s}$$

Ordinal Rank Regression with α, β ∈ [0, 1]!

original image: image shown to observer

internal image: image retained in observers' brain

red: forgotten image region

blue, **green**: correctly retained image region in memory

pink: hallucinated image region

external representation (v): observed image in terms of image region types, $v \in \{0, 1\}^n$

internal representation ($\tilde{v}$): image retained in memory in terms of image region types, $\tilde{v} \in \{0, 1\}^n$

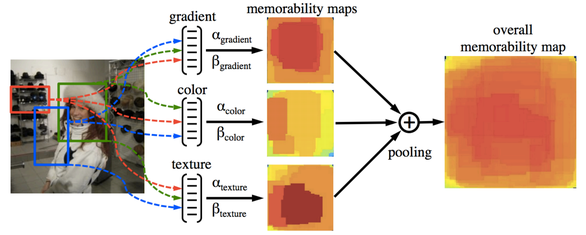
α_i: probability of forgetting image region of type i

β_i: probability of hallucinating image region of type I

s: image memorability score

D: distance between internal and external representation i.e. $\|v - \tilde{v}\|_1$

Image Region Attributes



Gradient
SIFT and HOG

Texture
Local binary pattern

Shape
Image self-similarity

Color
Color Naming

Saliency
Eye-tracking based

Semantic
Object Bank

Experiments

Dataset: Image memorability dataset with 2222 images, ~80 scores/image, 25 train/test splits
(images sampled from SUN database [Xiao et al, CVPR 2010])

Learning image region types: k-means clustering with 256 dictionary size

Memorability Prediction

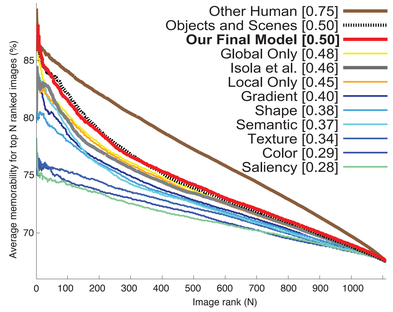
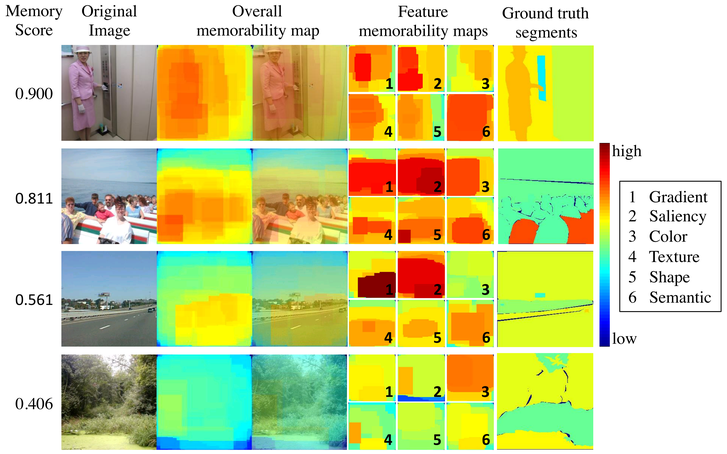
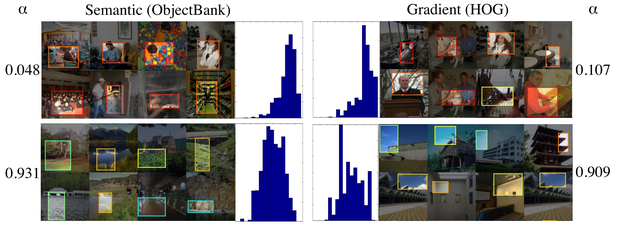
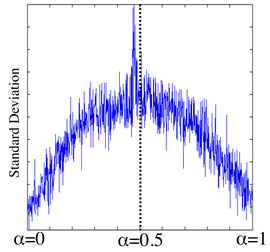


Image Memorability Maps

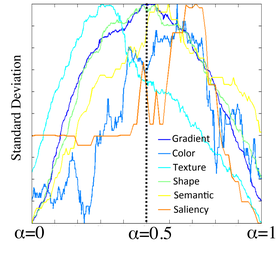


Analysis

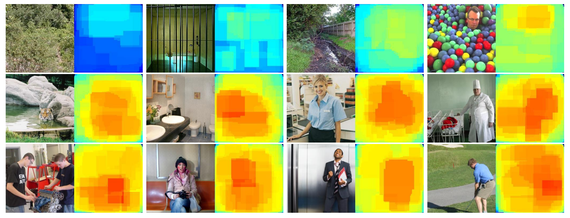




Standard deviation of memorability scores of image regions and their corresponding α



Additional Image Memorability Maps



Conclusion

- We demonstrate an effective yet interpretable framework to automatically discover the memorability of image regions
- Future development of such automatic algorithms of image memorability could have many exciting and far-reaching applications in computer science, graphics, media, designs, gaming and entertainment industries in general