MEMORABILITY OF NATURAL SCENES: THE ROLE OF ATTENTION

UMONS Université de Mons

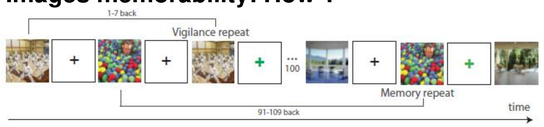
UNIVERSITÉ DE RENNES 1

Motivation: Investigate the influence of saliency-related features on images memorability

Images memorability: What ?

Probability of correctly detecting a repeat after a single view of an image in a long stream.

Images memorability: How ?



Memory game: 665 participants on Amazon's Mechanical Turk. (Isola et al. 2011).

Images memorability: Isola database

2222 images from SUN database (*Xiao et al. 2010*) with a memorability score from close to 0 (low memorability) to close to 1 (high memorability)

Image memorability prediction

Isola et al. proposed several features and a classifier to predict images memorability. The results are shown bellow:

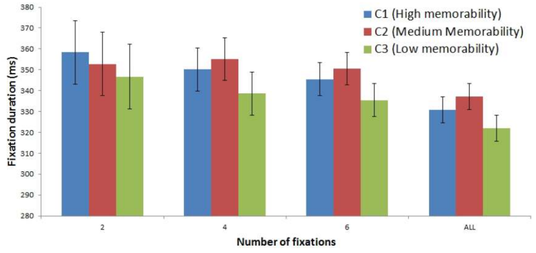
	Pixels	GIST	SIFT	SSIM	HOG 2x2	All Global Features
ρ	0.22	0.38	0.41	0.43	0.43	0.46

Eye-tracking experiment

Proposed dataset:

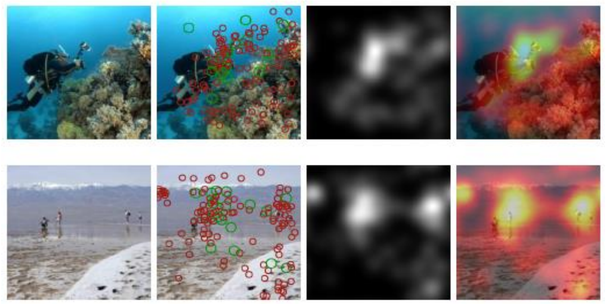
We extracted 3 classes of 45 images each from Isola et al. database. The first 45 (C1) are highly memorable images, the last 45 (C3) are the least memorable and the remaining 45 (C2) have an average memorability. The characteristics of C1, C2 and C3 are listed bellow:

Class	$Avg \pm STD$	t-test
$C1$	0.82 ± 0.05	$C1 vs C2, p << 0.001$
$C2$	0.68 ± 0.04	$C2 vs C3, p << 0.001$
$C3$	0.51 ± 0.08	–

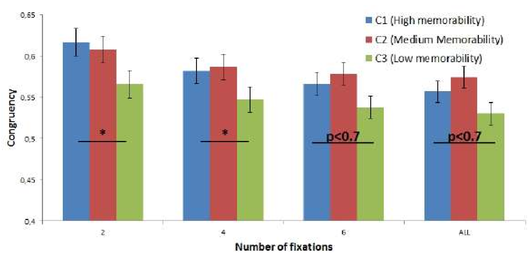


The fixation duration for the three image classes are shown for several viewing times: the 2 first fixations, the 4 first, the 6 first and all the fixations. The difference between C1 and C3 are every time statistically significant.

Example of eye-tracking results:



First row: high memorability image; Second row: low memorability image. First column: original pictures; Second column: fixation map (a green circle represents the first fixation of observers); Third column: Saliency map and Fourth column: heat map.

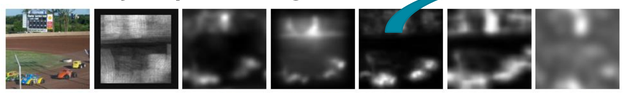


The congruency (agreement between viewers) is a second feature which is statistically different between C1 and C3.

Saliency-related memorability prediction

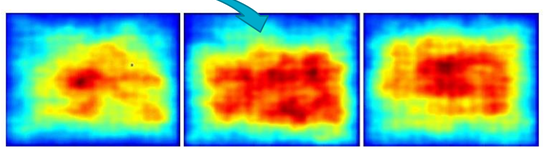
Two new saliency-related features for memorability prediction:

1/ Saliency maps coverage:

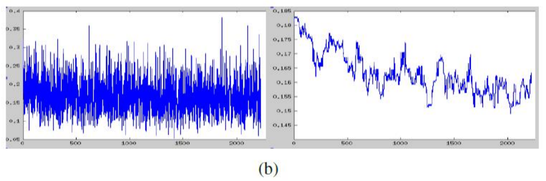


(a) Original (b) AIM (c) AWS (d) GBVS (e) RARE (f) Seo (g) LeMeur

Several saliency models were tested (above) and RARE 2012 was selected because the average coverage difference on several sets of images with different memorability was visible (Image (a) between left (high), middle (average) and right (high)). A coverage factor was computed (Graph (b) on the left on the 2222 images of Isola database. Right graph shows the feature after median filtering). The graph goes from low memorability to high memorability Images.



(a)



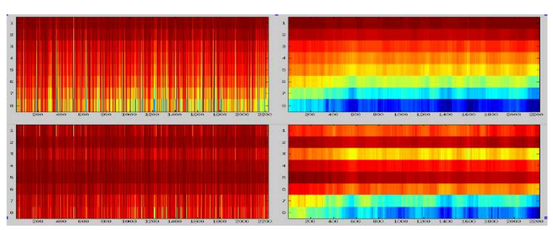
(b)

2/ Visibility features:

Low-pass filtering from I1 to I9 (pyramid-like): forgetting process.
Feature V1: the correlation between I1 and the others
Feature V2: the correlation between two successive filters

$$\forall i \in \{2,9\}, V1_i = |corr(I_1, I_i)|$$

$$\forall j \in \{1,8\}, V2_j = |corr(I_j, I_{j+1})|$$



Visibility feature vectors V1 and V2 computed for the whole 2222 images database. As for the coverage feature, the raw data both for V1 and V2 (left column) does not exhibit obvious differences. After median filtering (right column) differences between memorable (from the right) and less memorable images (from the left) are noticeable.

Eye-tracking data & saliency-related features available at:

- http;//tcts.fpms.ac.be/attention/?categorie26/images-memorability
- http://people.irisa.fr/Olivier.Le_Meur/publi/2013_ICIP

Conclusion: attention can play a role in memorability analysis !

Conclusion 1:

The fixation duration is longer for the most memorable images (especially for the very first fixations) which shows a higher cognitive activity for memorable images.

Conclusion 2:

The observers' congruency (agreement) is significantly higher for the most memorable images. This shows that when there are areas with high attraction on all viewers, this induces higher memorability.

Conclusion 3:

The use of coverage and visibility features (without any GIST) provides slight improvement compared to Isola 2011.

	Cov.	Vis.	Best (No GIST)	Best Isola
ρ	0.100	0.274	0.479	0.462

Conclusion 4:

The use of coverage and visibility features let us eliminate several other features, while keeping the same efficiency

	No Pixels	No SIFT	No HOG	No SSIM
ρ	0.476	0.474	0.470	0.468

Matei Mancas, Olivier Lemeur

matei.mancas@umons.ac.be | http://tcts.fpms.ac.be/attention -> Projects -> Attention and Images Memorability

olemeur@irisa.fr | http://people.irisa.fr/Olivier.Le_Meur/publi/2013_ICIP