

Prediction of the Inter-Observer Visual Congruency (IOVC) and Application to Image Ranking

O. Le Meur, T. Baccino and A. Roumy

Univ. of Rennes 1

<http://www.irisa.fr/temics/staff/lemeur/>

ACM Multimedia 2011

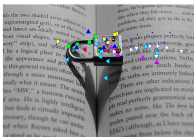
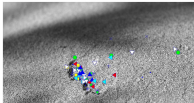
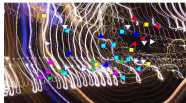
December 8, 2011

Problematic and context

Definition (Inter-observer visual congruency)

Inter-observer visual congruency (IOVC) reflects the visual dispersion between observers or the consistency of overt attention (eye movement) while observers watch the same visual scene.

Do observers look at the scene similarly?

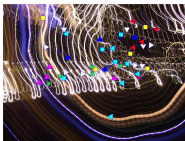


Problematic and context

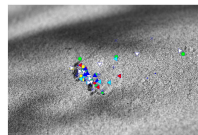
Two issues:

- 1 Measuring the IOVC by using eye data

LOW



HIGH



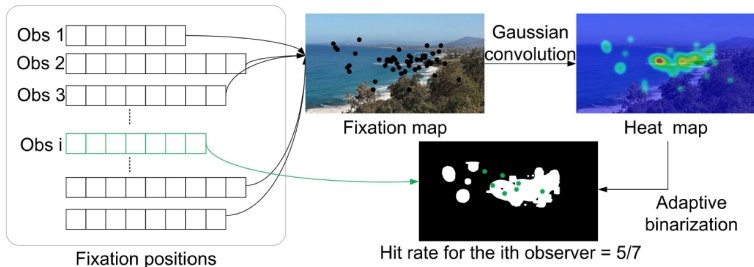
- 2 Predicting this value with a computational model

- 1 Problematic and context
- 2 Measuring the IOVC by using eye data
- 3 IOVC computational model
- 4 Performance
- 5 Application to image ranking
- 6 Conclusion

Measuring the IOVC

To measure the visual dispersion, we use the method proposed by [Torralba et al., 2006]:

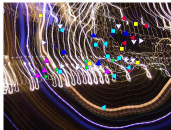
one-against-all approach



- $IOVC = 1$, strong congruency between observers.
- $IOVC = 0$, lowest congruency between observers.

Measuring the IOVC

Examples based on MIT eye tracking database



(a) 0.29



(b) 0.23

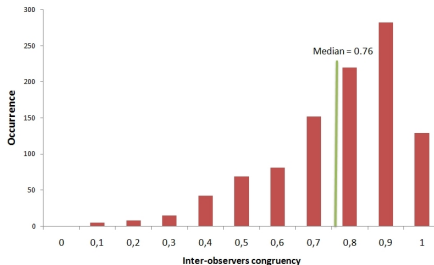


(c) 0.72



(d) 0.85

For the whole database:



Measuring the IOVC

Factors impacting the visual dispersion:

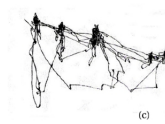
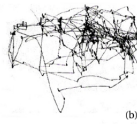
Bottom-up factors (stimulus-dependent)

- Scene complexity
- Scene consistency
- Human face
- Color
- ...

Top-down factors (observer-dependent)

- Task at hand
- Cultural difference
- Prior knowledge
- Prior experience
- ...

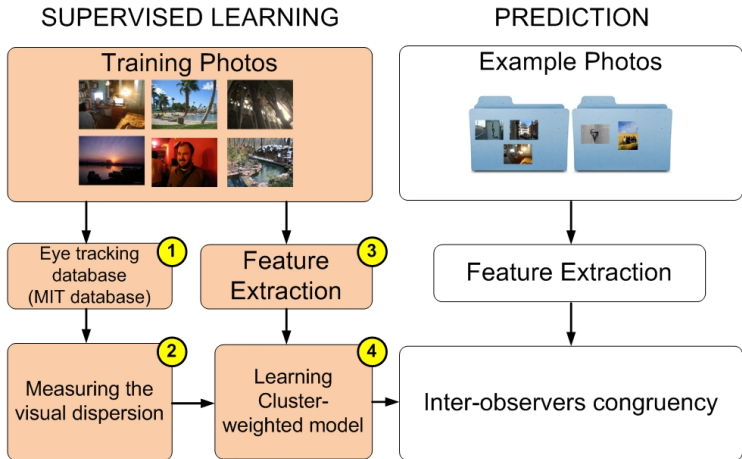
Yarbus, 1967



- 1 Problematic and context
- 2 Measuring the IOVC by using eye data
- 3 IOVC computational model**
- 4 Performance
- 5 Application to image ranking
- 6 Conclusion

IOVC computational model

Global architecture of the proposed approach:



Feature extraction

6 visual features are extracted from the visual scene:

Mixture between low and high-level visual features

- Color harmony [Cohen et al., 2006]
- Scene complexity:
 - Entropy E [Rosenholtz et al., 2007]
 - Number of regions (Color mean shift)
 - Amount of contours (Sobel detector)
- Faces [Lienhart & Maydt, 2002] (*Open CV*)
- **Depth of Field**



$E = 14.67 \text{ dit/pel}$, $\text{NbReg} = 103$



$E = 14.72 \text{ dit/pel}$, $\text{NbReg} = 72$

Feature extraction

Depth of Field computation inspired by [Luo et al., 2008]:

The horizontal/vertical derivatives histogram is modified after a blurring operation

I the input picture and f_k the blurring kernel of size $k \times k$ ($k = \{3, 5, 7\}$).

- 1 The vertical and horizontal derivatives are given by:

$$p_{xk} \propto \text{hist}(I * f_k * d_x)$$

$$p_{yk} \propto \text{hist}(I * f_k * d_y)$$

where $d_x = [1 \ -1]$ and $d_y = [1 \ -1]^T$.

- 2 For a pixel (i, j) and for a kernel k ,

$$D_k(i, j) = \sum_{(n, m) \in W_{ij}} KL(p_{xk} | p_{x1})(n, m) + KL(p_{yk} | p_{y1})(n, m)$$

- 3 The DoF value is finally computed as follows:

$$DoF = \sum_{(i, j) \in I} \sum_k D_k(i, j)$$

Feature extraction

Depth of Field computation inspired by [Luo et al., 2008]:



(a) Original



(b) DoF=0.12



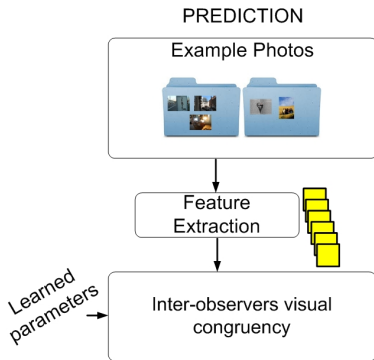
(c) Original



(d) DoF=0.75

Feature extraction

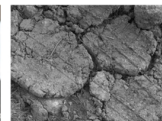
After the learning, we can now predict:



Predicted IOVC:



IOVC=0.96



IOVC=0.52



IOVC=0.94



IOVC=0.52



IOVC=0.82



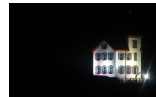
IOVC=0.64

- 1 Problematic and context
- 2 Measuring the IOVC by using eye data
- 3 IOVC computational model
- 4 Performance**
- 5 Application to image ranking
- 6 Conclusion

Performance

Two tracking databases are used to evaluate the performance:

- MIT's database (1003 pictures, 15 observers)
- Le Meur's database (27 pictures, up to 40 observers)



(a) 0.94, 0.96

We compute the Pearson correlation coefficient between predicted IOVC and 'true' IOVC:

- MIT's database:
 $r(2004) = 0.34, p < 0.001$
- Le Meur's database:
 $r(54) = 0.24, p < 0.17$



(b) 0.89, 0.94



(c) 0.91, 0.91

Performance

Some cases for which the model fails...



(a) (0.83,0.59)



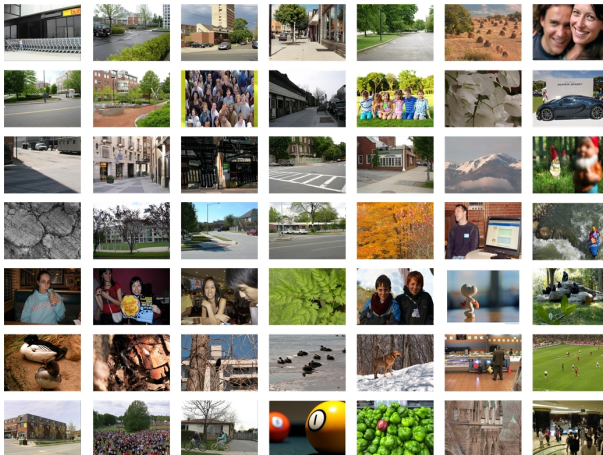
(b) Heat map

Figure: (.,.), true IOVC and predicted IOVC

- 1 Problematic and context
- 2 Measuring the IOVC by using eye data
- 3 IOVC computational model
- 4 Performance
- 5 Application to image ranking**
- 6 Conclusion

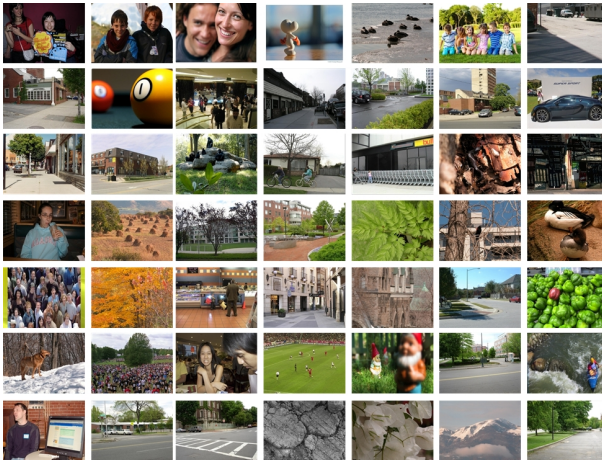
Application to image ranking

Goal: to sort out a collection of pictures in function of IOVC scores.



Application to image ranking

Goal: to sort out a collection of pictures in function of IOVC scores.



Application to image ranking

Goal: to sort out a collection of pictures in function of IOVC scores.

Top five pictures:



IOVC=0.98



IOVC=0.96



IOVC=0.94



IOVC=0.85

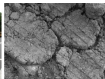


IOVC=0.82

Last five pictures:



IOVC=0.46



IOVC=0.52



IOVC=0.52



IOVC=0.64



IOVC=0.64

- can be combined with other factors to evaluate the interestingness or attractiveness of visual scenes.

- 1 Problematic and context
- 2 Measuring the IOVC by using eye data
- 3 IOVC computational model
- 4 Performance
- 5 Application to image ranking
- 6 Conclusion

Conclusion

Contributions:

- Main contribution: A new approach to evaluate the visual dispersion between observers
 - based on low-level visual features
 - it outperforms the visual clutter approach [Rosenholtz et al., 2007]
- Side contributions:
 - A new method to estimate the DoF
 - IOVC-based image ranking

Conclusion

This study is a start !

A number of tracks for the future:

- Time-dependent IOVC
- Spatio-temporal IOVC to deal with video sequences

Rooms to improve applications such as:

- Video compression
- Computational model of visual attention
- Retargeting
- Estimating the attractiveness of picture...

References

- [Cohen et al., 2006] D. Cohen-Or, O. Sorkine, R. Gal, T. Leyvand, and Y. Xu. Color harmonization. In ACM Transactions on Graphics (Proceedings of ACM SIGGRAPH), volume 56, pages 624-630, 2006.
- [Le Meur et al., 2006] O. Le Meur, P. Le Callet, D. Barba and D. Thoreau. A coherent computational approach to model the bottom-up visual attention. IEEE Trans. on Pattern Analysis and Machine Intelligence, 28(5), pp. 802-817, 2006.
- [Lienhart & Maydt, 2002] R. Lienhart and J. Maydt. An extended set of haar-like features for rapid object detection. In ICIP, volume 1, pages 900-903, 2002.
- [Luo et al., 2008] Y. Luo and X. Tang. Photo and video quality evaluation: focussing on the subject. In ECCV, pages 386-399, 2008.
- [Rosenholtz et al., 2007] R. Rosenholtz, Y. Li, and L. Nakano. Measuring visual clutter. Journal of Vision, 7(2), March 2007.
- [Torralba et al., 2006] A. Torralba, A. Oliva, M. Castelhano, & J. Henderson. Contextual guidance of eye movements and attention in real-world scenes: the role of global features in object search. Psychological review, 113(4):766-786, 2006.