Psychophysical Assessment of Perceived Interest in Natural Images: The ROI-D Database

Ulrich Engelke, Hans-Jürgen Zepernick

IEEE Visual Communication and Image Processing

2011 Tainan, Taiwan

Psychophysical Assessment of Perceived Interest in Natural Images: The ROI-D Database

Ulrich Engelke and Hans-Jürgen Zepernick
Blekinge Institute of Technology
School of Engineering
37179 Karlskrona, Sweden
Email : ulrichengelke@gmail.com, hans-jurgen.zepernick@bth.se

***Abstract*— We introduce a novel region-of-interest (ROI) database for natural image content, the ROI-D database. The database consists of ROI maps created from manual selections obtained in a psychophysical experiment with 20 participants. The presented stimuli were 42 photographic images taken from 3 publicly available image quality databases. In addition to the ROI selections, dominance ratings were recorded that provide further insight into the interest of the selected ROI in relation to the background. In this paper, the experiment is described, the resulting ROI database is analysed, and possible applications of the database are discussed. The ROI-D database is made freely available to the image processing research community.**

I. Introduction

At any instant in time, the human visual system (HVS) is confronted with an abundant amount of information to be processed. To reduce the complexity of scene analysis, attentional mechanisms [1] in the early visual system and the higher cognitive layers enable us to identify the most relevant visual information in a given context. Computational models that emulate these integral mechanisms of the HVS were found to be highly beneficial for optimisation of a wide range of image and video processing applications, including, source coding, segmentation, retargeting, and quality assessment.

A number of computational visual attention models [2] has been proposed in the literature that are typically designed based on gaze patterns recorded during eye tracking experiments. These models often focus on saliency-driven attention to visual content [3], neglecting higher order semantic information related to objects in the scene. The perceived level of interest to objects in a scene, however, has been shown [4] to have a comparably higher impact on visual attention than feature saliency. However, gaze patterns from eye tracking experiments do not necessarily reflect solely the interest of observers but include also bottom-up attention driven by feature saliency. It is thus imperative to conduct dedicated experiments to better understand the relative interest of a human observer to objects in a natural scene.

We therefore conducted a psychophysical experiment to investigate the perceived level of interest in photographic images. In this experiment, human observers were asked to manually select regions in a number of natural images, similar to the procedures followed in the establishment of the Berkeley Segmentation Dataset (BSD) [5]. However, the motivation here is very different. For the BSD, the participants were instructed to segment the different objects in the visual scene, not taking into account their relative interest. The resulting maps are instrumental for image segmentation tasks. In our experiment, we instructed the participants to select the regions that are of highest interest to them, the regions-of-interest (ROI). The participants were further instructed to rate the dominance of their ROI selection in relation to the background. The resulting ROI maps are instrumental as a ground truth for the design of ROI prediction models. Furthermore, the presented images were taken from three publicly available image quality databases [6]–[8]. Hence, together with the subjective quality ratings from these databases, the ROI maps are particularly beneficial to improve image quality prediction models.

In this paper, we discuss in detail the procedures of the experiment that we conducted and the post-processing of the ROI selections. We analyse the results through statistical analysis and highlight some interesting findings. To stimulate further research on perceived interest in natural images, we make the experiment outcomes available to the research community in the ROI-D database, with the 'D' indicating that dominance ratings are provided in addition to the ROI maps.

This paper is organised as follows. Section II introduces the ROI selection experiment and discusses the post-processing of the collected data. Section III discusses the outcomes of the experiment. Finally, conclusions are drawn in Section IV.

II. Region-of-interest experiment

A. Experiment details

The ROI experiment was conducted at the Blekinge Institute of Technology in Karlskrona, Sweden. The experiment took place in a laboratory environment under low light conditions. A total of 20 people from 13 different nationalities participated in the experiment. The age of the participants ranged from 23 to 37 with an average age of 29.7 years. Twelve of the participants were male and 8 were female.

The participants were shown the reference images from 3 well known image quality databases, with 14 images being contained in the MICT database [6], 10 images in the IRCCyN/IVC database [7], and 29 images in the LIVE database [8]. Given the overlap of 10 images between the MICT and the LIVE database, a total of 42 different images were chosen to be presented to the participants in our experiment. The names of the images and the related databases from which they

TABLE I

IMAGE NUMBERS (#) AND NAMES OF THE IMAGES IN THE IVC (I), LIVE (L), AND MICT (M) DATABASES (DB).

#	Name	DB	#	Name	DB
1	avion	I	22	lighthouse	L
2	barba	I	23	lighthouse2/kp21	L/M
3	boats	I	24	manfishing	L
4	clown	I	25	monarch	L
5	fruit	I	26	ocean/kp16	L/M
6	house	I	27	paintedhouse/kp24	L/M
7	isabe	I	28	parrots/kp23	L/M
8	lenat	I	29	plane/kp20	L/M
9	mandr	I	30	rapids	L
10	pimen	L	31	sailing1/kp06	L/M
11	bikes/kp05	L/M	32	sailing2	L
12	building2	L	33	sailing3	L
13	buildings/kp08	L/M	34	sailing4	L
14	caps/kp03	L/M	35	statue	L
15	carnivaldolls	L	36	stream/kp13	L/M
16	cemetry	L	37	studentsculpture	L
17	churchandcapitol	L	38	woman	L
18	coinsinfountain	L	39	womanhat	L
19	dancers	L	40	kp01	M
20	flowersonih35	L	41	kp07	M
21	house/kp22	L/M	42	kp12	M

were obtained are listed in Table I. Prior to the test images, 1 training image and 3 stabilisation images were shown to explain the ROI selection procedure and for the participants to get used to the selection process, respectively.

The images were presented to the participants on a 19" DELL LCD screen with a resolution of 1280×1024 and a black background behind the images. The viewing distance was approximately 50 cm, but the participants were allowed to look closer to identify details in the images. No time limits were imposed, but in general the participants took approximately 30-60 minutes to perform the ROI selections.

B. ROI selections

The participants were instructed to select an object or a region in the image that was of highest interest to them. If, and only if, there were multiple objects of highest interest, then the participants were allowed to select all of them. This was, for instance, often the case when two eyes were selected or multiple humans were selected. No limitations regarding the size of the ROI were imposed, except that the ROI needed to constitute a subset of the whole image. If the participants were unsatisfied with the accuracy or the choice of their selection, then it was deleted and a new selection was made.

The ROI were selected with a paint brush using Photoshop CS5. The brush had a circular shape and three different diameters for the participants to choose from; 20 pixels, 40 pixels, and 80 pixels. The small brush size was to be used for the selection of smaller objects and object contours whereas the larger brush size mainly facilitated fast filling of larger areas. The colour of the brush was chosen to be pink with the RGB values [255,0,240], as this colour is absent in all images and thus, easy segmentation of the ROI from the background

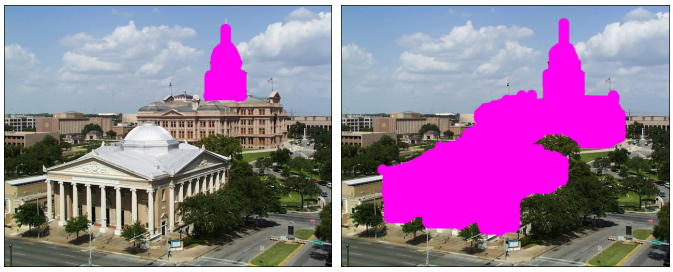

Fig. 1. Two example ROI selections for the image 'churchandcapitol'.

TABLE II

DOMINANCE RATING SCALE.

Dominance	Score
Very dominant	3
Dominant	2
Slightly dominant	1

is facilitated. Figure 2 illustrates two example ROI selections for the image 'churchandcapitol'.

C. Dominance ratings

In addition to the ROI selections, the participants were asked to provide a rating on a scale from 1 to 3 as to how dominant their interest in the ROI selection is in relation to the background. The rating scale is shown in Table II. The participants were instructed to give a rating of 3 if the selected ROI was very dominant compared to the remainder of the image. In case that there were other objects of interest that were not selected, then typically a rating of 2 was requested. If the participant struggled to decide between several objects of highest interest, but decided to select only a subset of them, then a rating of 1 was to be given as the other object(s) in the scene were of almost the same interest.

These dominance ratings, in the following denoted as r_D, serve to get a better understanding as to whether the selected ROI are dominating the image content or not. They can further be utilised as a weighting factor for the construction of ROI maps over all participants.

D. Post-processing into ROI maps

Binary ROI maps $ROI_{i,j}$ for the i^{th} image and j^{th} participant were created with 1's at each pixel location $[m, n]$ inside the ROI and 0's at each pixel location in the background. To highlight the perceived interest as an accumulation over all observers, these binary maps are combined over all participants as follows

$$ROI'_i = \sum_{j=1}^{20} ROI_{i,j} \tag{1}$$

and normalised into the final ROI map for each image

$$ROI_i = \frac{ROI'_i}{max(ROI'_i)} \tag{2}$$

with $ROI_i(m, n) \in [0 \ldots 1]$. To take into account the dominance ratings from the experiment, dominance weighted ROI

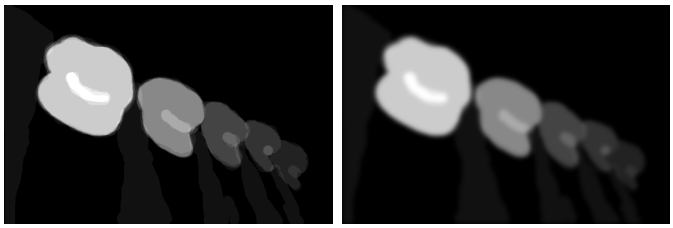

Fig. 2. ROI maps for image 'caps/kp03', ROI_{14} : (a) unfiltered, (b) filtered with a Gaussian kernel.

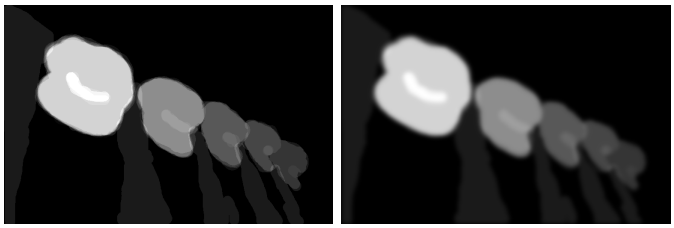

Fig. 3. Dominance weighted ROI maps for image 'caps/kp03', $ROI_{14,D}$: (a) unfiltered, (b) filtered with a Gaussian kernel.

maps are defined as

$$ROI'_{i,D} = \sum_{j=1}^{20} r_{D,j} \cdot ROI_{i,j}. \quad (3)$$

Similar to (2), these maps are transferred into normalised dominance weighted ROI maps, $ROI_{i,D}$. To smoothen the effect of imperfect ROI selections and for improved visualisation, these ROI maps can further be filtered using a Gaussian kernel. Examples of ROI maps, ROI_i, and dominance weighted ROI maps, $ROI_{i,D}$, for image 'caps/kp03' are shown in Fig. 2 and Fig. 3, respectively. The maps only differ in the relative dominance of the ROI but the general structures are very similar. This observation holds for most images in the test.

III. Data analysis

In the following, we briefly analyse the ROI selections, the dominance ratings, and their interrelationship. The analysis is conducted with respect to the numerical results presented in Figs. 4-6 and the example images presented in Fig. 7.

Figure 4 presents the maxima of each of the combined ROI maps before normalisation, ROI'_i. Hence, the maxima represent the number of overlapping ROI selections from all 20 observers. A higher number thus relates to a higher agreement between the participants regarding a particular ROI selection. It can be seen that the maxima of the ROI maps range from as low as 8 to as high as 19. In fact there were three images where 19 people agreed on an ROI selection. One of them is the image 'clown' (image 4) as presented in Fig. 7(a). Here, all but one selection were including the mirrored face of the clown. Similarly, all but one person selected the butterfly's wings in the image 'monarch' (image 25), as presented in Fig. 7(b). These images contain very distinct ROI in terms of a face and an animal. Both of them exhibit unusual and beautiful patterns, which appear to be of high interest to many

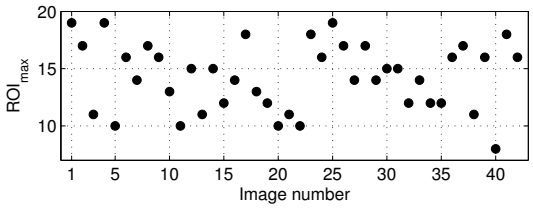

Fig. 4. Maxima of the combined ROI maps, ROI'_i.

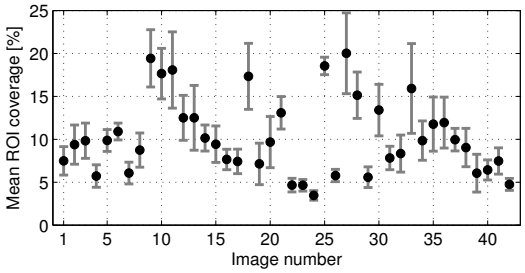

Fig. 5. ROI coverage averaged over all 20 participants.

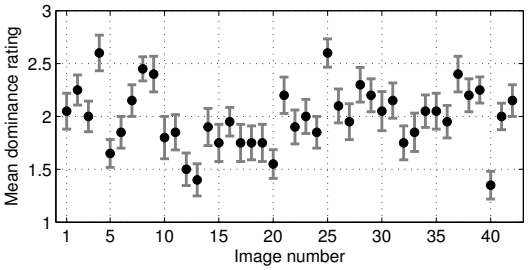

Fig. 6. Dominance ratings averaged over all 20 participants.

observers. On the other extreme is the image 'kp01' (image 40), as shown in Fig. 7(c), which only has 8 overlapping ROI selections. This is mainly due to several ROI being present in the image in terms of the door and the windows. Hence, different people selected different ROI, with an emphasis on the door and the three windows located towards the image center.

Figure 5 shows the ROI coverage for all images, averaged over all participants. The ROI coverage is the percentage of pixels within the ROI to the number of pixels in the image. It can be observed that the average ROI coverage varies strongly with the image content, with the smallest ROI coverage of approximately 4 % for the image 'manfishing' (image 24) and the largest ROI coverage of approximately 20 % for the image 'paintedhouse/kp24' (image 27). The largest variation amongst observers is exhibited for the image 'sailing3' (image 33), as indicated by the large standard error. Figure 7(d) shows that this image has a strong, small ROI on the people in the boat, but also exhibits very large ROI selections of the sails.

Figure 6 presents the dominance ratings for all images, averaged over all participants. The averaged dominance ratings range from approximately 1.4 to 2.6. It is interesting to note that the images 'clown' and 'monarch', which had the highest maxima in their ROI maps, have also received the

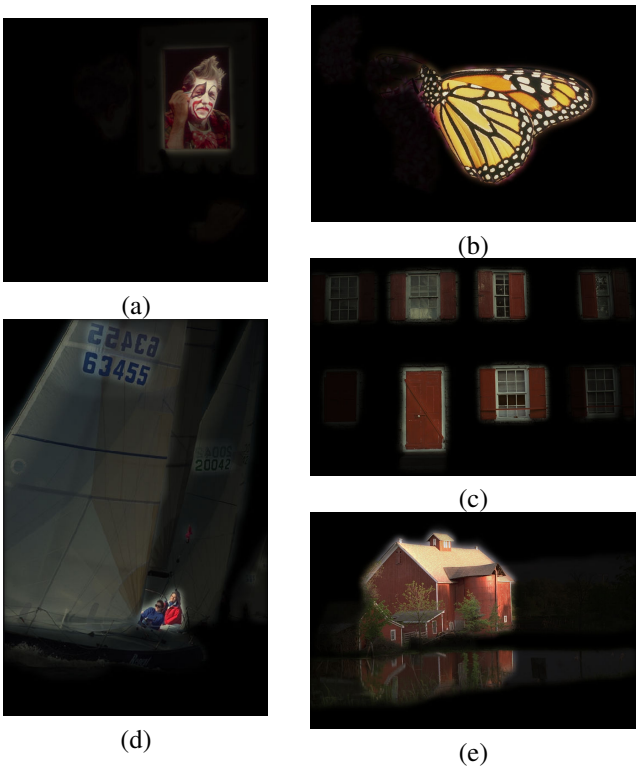

Fig. 7. ROI maps superimposed onto the images : (a) 'clown', (b) 'monarch', (c) 'kp01', (d) 'sailing3', (e) 'house/kp22'.

highest averaged dominance ratings. This is further evidence of the strong dominance of these ROI. Similarly, the image 'kp01' received both the lowest maxima in the ROI map and the lowest average dominance rating. These observations indicate an interrelationship between the ROI map maxima and the dominance ratings and indeed, these two factors exhibit a linear correlation coefficient of CC $= 0.595$. It is further observed that the ROI coverage is nearly uncorrelated (CC $= 0.017$) with the dominance ratings, indicating that the observers did not relate the size of the selected objects directly to their dominance over the background.

However, there are also exceptions to the rule, as the image 'house/kp22' (image 21) presented in Fig. 7(e) illustrates. This image exhibits a well above average dominance rating, but the maximum in the ROI map is well below average. The latter is caused by a segregation of the participants into two groups, of which one was mainly interested in the red house whereas the other group was mainly interested in its reflection in the lake. Both groups, however, consider their preference to be dominant above average. Such exceptions indicate that these interrelations need more in-depth investigation.

IV. Conclusions

We presented an ROI selection experiment for natural images and a statistical analysis of the resulting ROI maps and related dominance ratings. The ROI maps find deployment in a wide range of applications, of which some notable ones are the following. First and foremost, the ROI maps can be used for the design of computational ROI prediction algorithms [9]. These algorithms can in turn be deployed in real systems in order to bypass the need for manually selected ROI. Since the images presented in the experiment were taken from image quality databases, the ROI maps will prove to be useful to evaluate the potential benefits of incorporating ROI information into objective quality models [10]. Other applications include ROI-based image coding [11], retargeting [12], and retrieval [13]. Independent of the application, the ROI maps serve both to design and validate the performance of the according image processing systems.

Future work may include an extension of the database, by conducting further experiments in different laboratories. Furthermore, the interrelationship of the selected ROI with gaze patterns recorded from eye tracking experiments is of great interest and yet to be investigated.

We make the ROI selections and dominance ratings freely available in the ROI-D database. To obtain access, please send an email to the principal author of this paper Ulrich Engelke.

Acknowledgment

The authors would like to thank the participants for sparing their valuable time to help us out with the experiment.

References

[1] J. Wolfe, "Visual attention," in *Seeing*, K. K. D. Valois, Ed. Academic Press, 2000, pp. 335–386.

[2] S. Frintrop, E. Rome, and H. I. Christensen, "Computational visual attention systems and their cognitive foundations : A survey," *ACM Trans. on Applied Perception*, vol. 7, no. 1, Jan. 2010.

[3] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. PAMI*, vol. 20, no. 11, pp. 1254–1259, Nov. 1998.

[4] W. Einhäuser, M. Spain, and P. Perona, "Objects predict fixations better than early saliency," *Journal of Vision*, vol. 8, no. 14 :18, pp. 1–26, 2008. [Online]. Available : http://journalofvision.org/8/14/18/

[5] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, "Berkeley segmentation datasets and benchmarks 500," http ://www.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/resources.html, 2011.

[6] Z. M. P. Sazzad, Y. Kawayoke, , and Y. Horita, "Image quality evaluation database," http ://mict.eng.u-toyama.ac.jp/mict/index2.html, 2000.

[7] P. Le Callet and F. Autrusseau, "Subjective quality assessment IRCCyN/IVC database," http ://www.irccyn.ec-nantes.fr/ivcdb/, 2005.

[8] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, "LIVE image quality assessment database release 2," http ://live.ece.utexas.edu/research/quality, 2005.

[9] S. Pinneli and D. M. Chandler, "A Bayesian approach to predicting the perceived interest of objects," in *Proc. of IEEE Int. Conf. on Image Processing*, Oct. 2008, pp. 2584–2587.

[10] U. Engelke and H.-J. Zepernick, "A framework for optimal region-of-interest based quality assessment in wireless imaging," *Journal of Electronic Imaging, Special Section on Image Quality*, vol. 19, no. 1, Jan. 2010.

[11] K. H. Park and H. W. Park, "Region-of-interest coding based on set partitioning in hierarchical trees," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 12, no. 2, pp. 106–113, Feb. 2002.

[12] V. Setlur, T. Lechner, M. Nienhaus, and B. Gooch, "Retargeting images and video for preserving information saliency," *IEEE Computer Graphics and Applications*, vol. 27, no. 5, pp. 80–88, Sept. 2007.

[13] K. Vu, K. A. Hua, and W. Tavanapong, "Image retrieval based on regions of interest," *IEEE Trans. on Knowledge and Data Mining*, vol. 15, no. 4, pp. 1045–1049, July 2003.