Attentive Visual Efficiency

Computer Vision Group[1*]

[1]Department of Computer Science and Artificial Intelligence, University of Granada, CITIC-UGR, 18071 Granada, Spain

[*]To whom correspondence should be addressed; E-mail: jags@decsai.ugr.es

The human visual system allocates different amounts of processing resources to different portions of the visual field which provides a trade-off between resources and time. On the one hand, attention can be shifted to a new location through a saccadic eye movement. On the other hand, the photoreceptor density that decreases between the fovea and the periphery induces no-uniform processing capability over the entire field. In fact, the conclusion is still more surprising: features will only be perceived if they success in attracting attention. This text focuses on a computational approach to the human attentional mechanism and its possible applications mainly for predicting visual efficiency in the field of computer vision. It is intended to be an open gate on further research.

Axiomatic Approach to Computational Attention

In Reference (*1*), we describe, in terms of a decision problem, any situation in which a computational system will be forced to allocate attention at any time to one spatial location to improve the reconstruction fidelity on a neighborhood of the chosen point. The result is a rational model of computational attention in which a multi-bitrate attention map will provide us with the attention score for each spatial location at high and low quality versions of the image reconstruction. At any time a rational system should choose, even though without any outside knowledge, among alternative spatial locations in such a way as to avoid certain forms of behavioral inconsistency. We compare the performance between a rational approach of computational attention and various models for predicting visual target distinctness, using scenes that represent military vehicles in complex rural backgrounds.

Information Visibility Using Transmission Methods

In Reference (*2*), we use the new paradigm to the evaluation of the visual efficiency of image information when it is reconstructed at high and low fidelity using different transmission methods. To this aim, we apply a rational model of computational attention in which a multi-bitrate attention map will provide us with the attention score for each spatial location at high and low quality versions of the image reconstruction. The rational approach of attention does not purport to describe the ways in which the Human Visual System (HVS) actually do behave in making choices among possible locations of interest for allocating attention. Instead we are interested in the aspects of rationality that seem to be present in the decision making of the HVS. From both rate-attention and rate-distortion curves, we conclude which transmission method is the overall winner according to the new paradigm.

Comparative Visual Efficiency of Image Fusion Methods

The objective of Reference (*3*) is to represent relevant information from multiple individual images in a single image. Some fusion methods may represent important visual information more distinctively than others, thereby conveying it more efficiently to the human observer. In this paper, we propose to rank order images fused by different methods according to the computational attention value of their visually important details. First we compute for each of the fused images a multi-bitrate attention map, following a rational model of computational attention. From this attention map, we then calculate the average attention score within areas of interest (e.g., living creatures, man-made objects, and terrain features), for each bitrate. A high computed mean attention value within the areas of interest at any reconstruction fidelity corresponds to a high computational saliency of the areas of interest. The computational results agree with human observer performance, making the approach valuable for practical applications.

Comparative Visibility Analysis of Advertisement Images

Is there any advertisement in a particular dataset more visually efficient than the rest? In Reference (*4*) we propose that advertisement images may be rank ordered based on their important information visibility using computational attention.

For each one of the advertisement images we firstly compute a multi-bitrate attention map following a rational model of computational attention. Next, based on the attention map, we calculate the average attention score, for each bitrate, within the areas of interest either provided by the publicist or by the use of automated detection. A high value of the mean attention within the areas of interest at any reconstruction fidelity corresponds to a high saliency of these areas.

Thus, for each advertisement, we calculate a rate-attention curve as given by the normalized mean attention score within the areas of interest across bitrates. Each image is decoded at different bitrates of picture quality using a coding method. Unsupervised learning can then be used to perform the clustering of the advertisements into subsets so that images in the same cluster are similar in the rate-attention sense. In the experiments one advertisement has appeared to be more visually efficient than the rest of images in a dataset of example.

Attention-based Advertisement Images

In advertising, some images will be more visually efficient than others, since are more able to provide its main message to the potential customer. We have also that there exist high saliency regions which may be detected by any human user because involuntarily attract attention. Our aim in Reference (*5*) is to increase the visual efficiency of advertisements by blending the original advertisement with a high saliency image which predicts what regions of the advertisement involuntarily attract attention. A computational model will be used to perform the automated detection of such a high saliency version for any advertisement.

Different levels of opacity between the original advertisement and its high saliency version can be used in order to improve the blending. In a rate-attention sense, we can conclude which blending with an optimal level of opacity is more able than the rest to provide its main message to the potential customer.

Attention-based Peak Signal-to-Noise Ratio

Finally, in Reference (*6*) we present a measure of image quality based on computational attention, in which pixel errors are computed on high saliency locations at which humans might perceive features in the images since they involuntarily attract attention.

In the first part of Reference (*6*), the scheme of relevant feature detection will be discussed and evaluated in terms of their performance for a visual processing task. The model is based on the use of a multi-bitrate attention map which provides us with a computational attention score for each spatial location at high and low quality versions of the image reconstruction. In the second part of this paper, examples are provided of object detection applications that illustrate the ability of the induced error measure to predict the detectability of objects as well as their perceptual capabilities.

Web Platform

All software (with documentation) developed in the papers may be accessed on the Internet site `http://decsai.ugr.es/cvg`. All material is made available to other researchers for academic use only. The programs were not optimized to a commercial application level.

We have developed a web platform with two applications illustrating the concepts inside the papers. `http://cvg.ugr.es/attention` for evaluating the impact of advertisement images including some very useful tools for publicists in advertising agencies, and `http://cvg.ugr.es/fusion` for evaluating different methods of image fusion.

To use them, you must be registered (separately) in the web site providing your email and other professional information. This is for free and the submitted information is only used for statistical academic purposes.

1 Comparative Visibility Analysis of Advertisement Images

This web application (http://cvg.ugr.es/attention) allows you to classify an advertisement image according to his visual efficiency. In order to do that, you must determine a region of interest (ROI) over your input image.

You must provide an image to work with and one or more regions of interest over the image. There are two ways to do that and the results will be the same. Please note that the calculations are very slow (they may take several minutes depending on the images). Here we describe the two possible approaches to provide images and to obtain the corresponding results.

1.1 Using an external editor

You can create a binary mask (ROI) with an external editor, and submit the original image together with the ROI mask. Figure 1 shows this interface. Once you have selected both images, you must press the “Send” button.

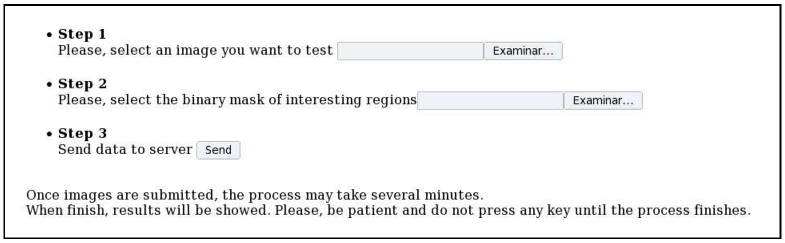


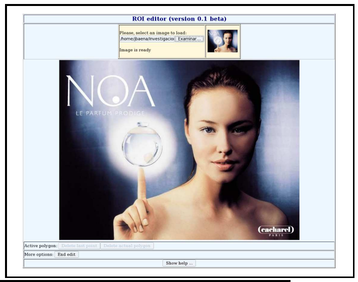
Figure 1: Web interface: submitting image and ROI.

1.2 Using the interactive integrated editor

Instead of using an external editor to get the ROI, you can use the interactive editor integrated inside the application. In the Fig. 2 you can see two screenshots. On the left, one image has been loaded (top button of the window). The application displays a reduced version if the image is too large to work with. Once loaded, you can use the left button of the mouse to create polygons over the image. Each click adds a new point or, if you drag the mouse while pressing it, you will get a smoother contour. To close the polygon you must double click the left mouse button. It is possible to create as many polygons as you need. They are drawed in a blue tone over the image. The red one is the polygon you are creating just now. All of them together define the ROI mask.

When you have the desired ROI mask, press the “End edit” button (at the bottom of the window). Then a preview of the image and the ROI are displayed to ensure they are correct. Now, if you want, you have the opportunity to save the ROI mask image to your hard disk by pressing the right button over the image (contextual menu of the web navigator) and selecting “Save image as ...”. Note that, probably, images are displayed smaller that the original one: no matter with this, real size images will be created and processed.

At this point you must press the “Send images” button and calculations will start.

[Advertisement image is loaded.] 

[ROI are being created.]

Figure 2: Web interface: interactive editing of ROI.

1.3 Obtaining results

Once images are submitted by either of the two described methods, calculations will proceed. This task may take several minutes. When the calculations are finished, results are showed as you can see on figure 4. The graphics show the rate-attention curves from a reference database of advertisement images. The line with '+' marks, corresponds to the rate-attention for the submitted image and ROI. Images are clustered and displayed using a different color for each cluster. Finally, you can see a textual classification for the impact of your image: Very low (1 star), low (2 stars), medium (3 stars), high (4 stars) and very high (5 stars). This quantifies the visual impact of your advertisement image according to the selected ROI.

2 Image fusion

This web platform is designed to analyze the quality of different methods of image fusion (http://cvg.ugr.es/fusion). Through the interface, you will submit, at least, four images (see figure 5):

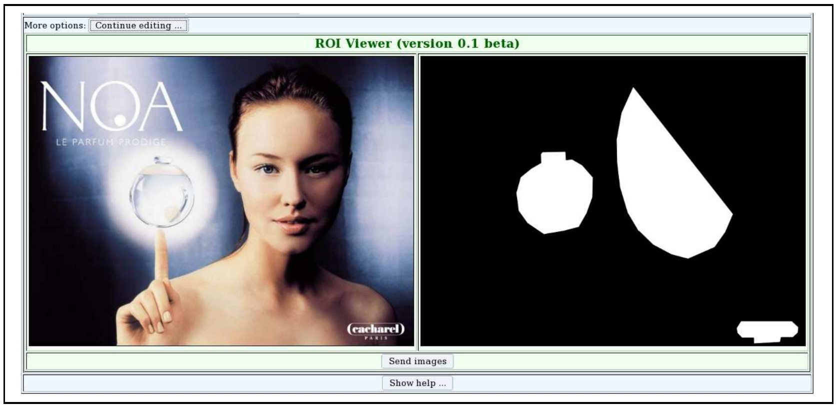


Figure 3: Web interface: Previewing image and ROI.

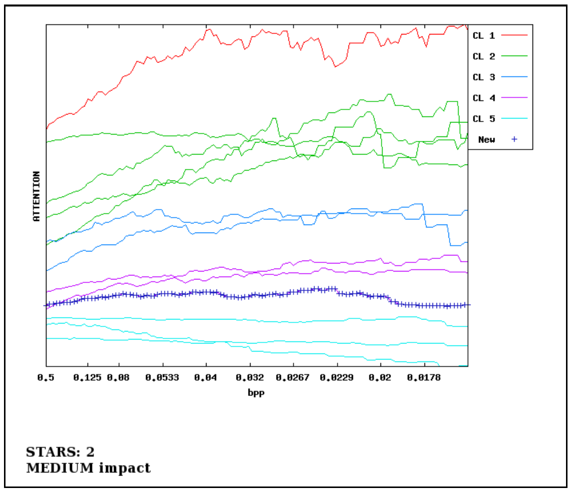


Figure 4: Web interface: Results.

- Visual image (image in the visual spectrum).
- IR image (infrared image).
- Two fused images. By pressing the “Add fused image” button you can submit more than two images.

Fused images are obtained by different methods and then submitted here to be compared.

Once images are submitted, they are processed (again, this will take several minutes) by applying the algorithm described in (*4*). Two automated ROI’s are extracted, one from the

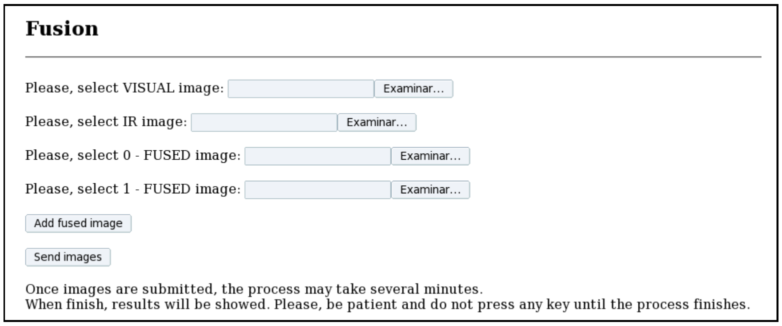


Figure 5: Web interface: Image input.

visual image and the other from the infrared image. Then, for each ROI, we calculate the rate-attention curve for each fused image. In the results page we can see:

- Original images (visual and infrared) together with their automated ROI's. See figure 6.
- Fused images ordered from best to worse using ROI from visual image. See figure 7-(a).
- Fused images ordered from best to worse using ROI from infrared image. See figure 7-(b).

Note that each fused image is numbered according to the input order used.

References and Notes

1. J.A. Garcia, R. Rodriguez-Sanchez, and J. Fdez-Valdivia. Axiomatic Approach to Computational Attention. *Pattern Recognition*, vol. 43 (4), pp. 1618-1630. (2010)

2. J.A. Garcia, R. Rodriguez-Sanchez, J. Fdez-Valdivia, and J. Martinez-Baena. Information Visibility Using Transmission Methods. *Pattern Recognition Letters*, Vol. 31, pp. 609-618. (2010)

3. J.A. Garcia, R. Rodriguez-Sanchez, J. Fdez-Valdivia, and Lex Toet. Comparative Visual Efficiency of Image Fusion Methods. Submitted to *Information Fusion*, (2010)

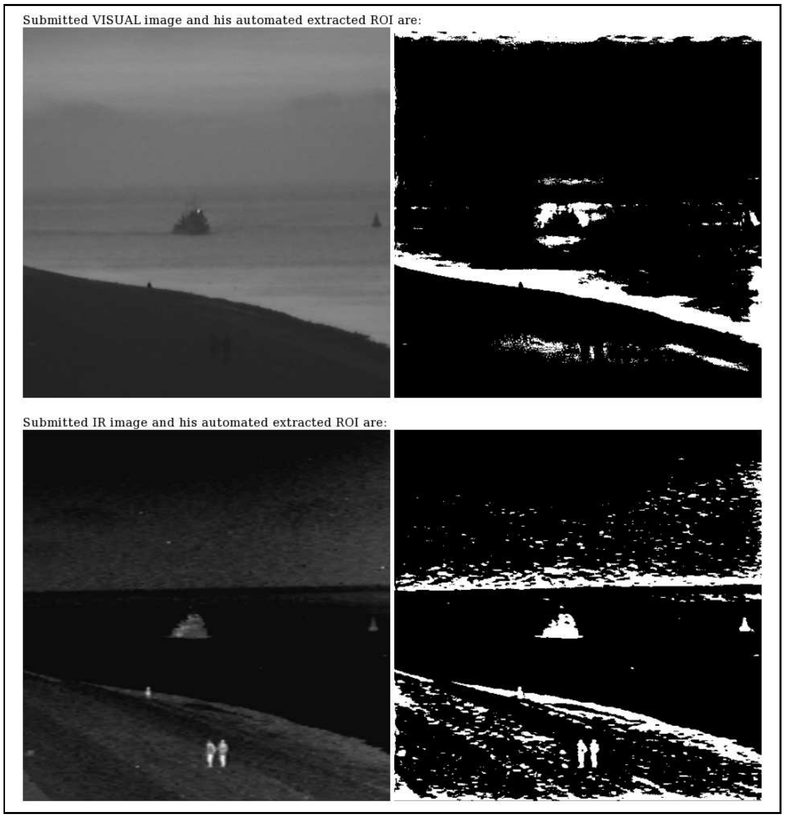


Figure 6: Web interface: Results (visual and infrared with ROI's).

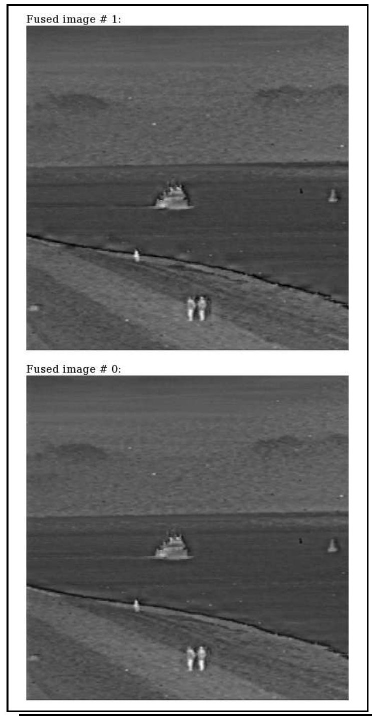
4. J.A. Garcia, R. Rodriguez-Sanchez, J. Fdez-Valdivia, and J. Martinez-Baena. Comparative Visibility Analysis of Advertisement Images. Submitted to *Signal Processing: Image Communication*, (2010)

5. J.A. Garcia, R. Rodriguez-Sanchez, and J. Fdez-Valdivia. Attention-based Advertisement Images. Submitted to *Pattern Recognition Letters*, (2010)

6. J.A. Garcia, R. Rodriguez-Sanchez, and J. Fdez-Valdivia. Attention-based Peak Signal-to-Noise Ratio. Submitted to *Information Sciences*, (2010)

7. http://cvg.ugr.es/attention. Web Platform for applications of the Computational Attention Model. (2010)

[Fused images ordered with visual ROI.]



[Fused images ordered with infrared ROI.]

Figure 7: Web interface: Results (ordered fused images).

Acknowledgments. This paper was sponsored by the Spanish Board for Science and Technology (MICINN) under grant TIN2010-15157.