Decorrelation and Distinctiveness provide with Human-Like Saliency

Antón Garcia-Diaz[1], Xosé R. Fdez-Vidal[1], Xosé M. Pardo[1], Raquel Dosil[1],

[1] Universidade de Santiago de Compostela, Grupo de Visión Artificial,
Departamento de Electrónica e Computación, Campus Sur s/n,
15782 Santiago de Compostela, Spain
{anton.garcia, xose.vidal, xose.pardo, raquel.dosil}@usc.es

Abstract. In this work we propose a new model for the extraction of bottom-up saliency. It presents low computational complexity compared to other models of the state of the art. It is based in biologically plausible mechanisms: the decorrelation and the distinctiveness of local responses. Decorrelation is achieved by applying principal component analysis over a set of multiscale low level features. Distinctiveness is measured using the Hotelling's T^2 statistic. It is conceived to be used in a machine vision system, in which attention would contribute to enhance performance together with other visual functions. We will show that our model is consistent with a wide variety of psychophysical phenomena that are referenced in the visual attention modeling literature, and that it outperforms other state of the art models in reproducing these phenomena.

Keywords: saliency, bottom-up, attention.

1 Introduction

The Human Visual System (HVS) has to face a huge computational complexity, as has been shown in visual search experiments [1]. It tackles this challenge through the selection of information using several mechanisms. In the basis of this selection process are the visual attention mechanisms, including the data-driven ones, leading to the so called bottom-up saliency. In the last decades, the interest in the understanding of this attentional component and the appraisal of its relative importance in relation to the top-down, knowledge-based mechanisms, has constantly raised. Correspondingly, an increasing number of approaches to its computational modeling is coming up. Besides, there exists an evident interest in the application of these models in the solution of technical problems requiring active vision approaches, ranging from robotics to image compression or object recognition.

Many of the models for bottom-up saliency proposed in the literature are based on abundant evidences from psychophysical experimentation. Nothdurft [2] has proposed that local feature contrast generally attracts gaze. In the same direction, Zetchsche [3] points out to the local contribution to the structure content as driving attention. In general, most models of bottom-up saliency assume that local distinctiveness is the basis for data-driven attention. Following this direction, the already classic model of Itti & Koch [4] proposed the iteration of center-surround competition processes to reach a powerful computational model of saliency. With a similar approach Le Meur et al. [5] posed a scheme using a multiscale and multioriented decomposition, along with contrast sensitivity, visual masking, center-surround competition and perceptual grouping. Recently, a different set of approaches to bottom-up saliency has been proposed based on similarity and local information measures. In these models local distinctiveness is obtained either from self-information [6][7], mutual information [8][9], or from dissimilarity [10], using different decomposition and competition schemes.

In a previous work [11] we studied the combination of scale information decorrelation with center-surround (c-s) differences. There, we compared the obtained performance in visual search experiments with the model of Itti & Koch [4]. In that approach only achromatic information was taken into account. Moreover, that model was unable to correctly reproduce some important psychophysical phenomena tackled here.

In this paper we propose a more efficient, simple and light approach to the problem of modeling bottom-up saliency. It is based solely in the decorrelation of scale information without the use of c-s differences. Hence, we turn back to the proposal of Olshausen & Field about the need of taking into account the decorrelation of neural responses when considering the behavior of a population of neurons subject to stimuli of a natural image [12]. That means considering neurons collectively, instead of individually. This is believed to be closely related to the important role of Non Classical Receptive Fields (NCRF) in the functioning of HVS. Therefore, we start from a classic multiscale decomposition on two main feature dimensions: local orientation energy and color. We obtain the decorrelated responses applying PCA to the multiscale features. Then, we measure the statistical distance of each feature to the center of the distribution as the Hotelling's T^2 distance. Finally, we apply normalization and Gaussian smoothing to gain robustness. The resulting maps are firstly summed, delivering local energy and color conspicuities, and then they are normalized and averaged, producing the final saliency map.

It is worth noting that we start -like probably most models- from a controlled decomposition, which retains important information. Thus, it is suitable for combination with top-down modulation approaches, like the incorporation of contextual influences or learning and recognition mechanisms. Unlike in the model of Bruce & Tsotsos [7], who use a decomposition based on independent components of patches from natural images, in our model it is very simple to actuate on (and from) scales, orientations or color components.

This approach reproduces a wide variety of psychophysical results, all of them closely related to the attentional function of the HVS. Hence, we will show how the model matches the nonlinearity against orientation contrast; the efficient (parallel) and inefficient (serial) search, the orientation asymmetry, the presence-absence asymmetry and Weber's law, the influence of background on color asymmetries, and the capability in the prediction of eye fixation data. Therefore the model achieves a degree of validation that outperforms other state of the art models.

The paper is developed as follows. Section 2 is devoted to describe the visual attention model. In Section 3 we present the experimental work carried out for validation of the model and the achieved results. Finally, Section 4 summarizes the paper and presents conclusions.

2 Model

Our model takes as input a color image codified using the Lab color model. In this way, each pixel is described by one luminance component (L), and two color opponent components: red/green (*a*) and blue/yellow (*b*). Unlike other implementations of saliency [8][13] this election is based on a widely used psychophysical standard. We decompose the luminance image by means of a Gabor-like bank of filters, in agreement with the standard model of V1. Since orientation selectivity is very weakly associated with color selectivity, the components *a* and *b* simply undergo a multiscale decomposition. Hence, we employ two feature dimensions -in the sense proposed by Wolfe [14]-: color and local energy. By decorrelating the multiscale responses, extracting from them a local measure of variability, and further performing a local averaging, we obtain a unified and efficient measure of saliency.

2.1 Local Energy and Color Maps

Local energy is extracted applying a bank of log Gabor filters [15] to the luminance components. The transfer function of the log Gabor filter takes the following expression

$$\mathrm{logGabor}(\rho,\alpha;\rho_i,\alpha_i)=e^{-\frac{(\log(\rho/\rho_i))^2}{2(\log(\sigma_{\rho i}/\rho_i))^2}}\,e^{-\frac{(\alpha-\alpha_i)^2}{2(\sigma_\alpha)^2}} \quad \textbf{(1)}$$

where (ρ, α) are polar frequency coordinates and (ρ_i, α_i) is the central frequency of the filter.

While the Gabor filter is non-zero for negative frequencies and presents a non-zero DC component, giving rise to artifacts, the log Gabor does not present this problem. Besides, it presents a symmetric profile in a logarithmic frequency scale. Hence, in a linear frequency scale it shows a long tail towards the high frequencies, providing a more localized impulse response. The impulse response is a complex valued function (with no analytical expression), whose components are a couple of functions in phase quadrature, *f* and *h*. Hence, the response of a log Gabor filter with scale *s* and orientation *o* to a luminance image L is:

$$\mathrm{Resp}_{so}(x,y)=(L*\mathrm{logGabor}_{so})(x,y)=L*f_{so}(x,y)+L*h_{so}(x,y)\,i \quad \textbf{(2)}$$

The modulus of the complex response of this filter is a measure of the local energy of the input associated to the frequency band with scale *s* and orientation *o* [16][17]

$$e_{so}(x,y) = \sqrt{(L * f_{so})^2 + (L * h_{so})^2} \quad \mathbf{(3)}$$

Regarding the color dimension, we obtain a multiscale representation for each of the opponent components *a* and *b*, from the responses to a bank of log Gaussian filters.

$$\text{logGauss}(\rho) = e^{-\frac{(\log(\rho))^2}{2(\log(2^n \sigma))^2}} \quad \mathbf{(4)}$$

Thus, for each scale and color opponent component we get a real valued response map:

$$\begin{aligned} \text{Resp}_{s\,a}(x,y) &= (a * \text{logGauss}_s)(x,y) \\ \text{Resp}_{s\,b}(x,y) &= (b * \text{logGauss}_s)(x,y) \end{aligned} \quad \mathbf{(5)}$$

The parameters used here were: 4 scales spaced by one octave, 4 orientations (for local energy), minimum wavelength of 4 pixels, angular standard deviation of σ_α = 37.5º, and a frequency bandwidth of 2 octaves.

2.2 Measurement of Distinctiveness

Variability and richness of structural content have been proven as driving attention in psychophysical experiments [3]. Here we have chosen a measure of distance between local and global structure to represent distinctiveness. But before estimating such distance, we need to preprocess the low level representation. Observations from neurobiology show decorrelation of neural responses, as well as an increased population sparseness in comparison to what can be expected from a standard Gabor-like representation [18]. To decorrelate the multiscale information of each sub-feature (orientations and color components) we perform a PCA on the corresponding set of scales. Once scales are decorrelated, we extract the statistical distance at each point as the Hotelling's T^2 statistic. Being x_{ij} a feature corresponding to pixel j and scale i, with i={1,...,*S*} and j ={1,...,*N*}, we compute the statistical distance T^2_j of each pixel in the decorrelated coordinates

$$\mathbf{x} = \begin{pmatrix} x_{11} & \cdots & x_{1N} \\ \vdots & \vdots & \vdots \\ x_{S1} & \cdots & x_{SN} \end{pmatrix} \rightarrow (\text{PCA}) \rightarrow T^2 = \left(T_1^2 \ , \ \cdots, \ T_N^2\right) \quad \mathbf{(6)}$$

T^2 is defined as follows, where $\mathbf{x_j}$ is a multiscale feature vector with *S* components and **W** is the covariance matrix.

$$T_j^2 = (\mathbf{x_j} - \bar{\mathbf{x}})' \mathbf{W}^{-1} (\mathbf{x_j} - \bar{\mathbf{x}}) \tag{7}$$

This should be viewed as the key point of our approach to the feature integration process. It provides an efficient mechanism for the deployment of pop-out effects, widely observed in psychophysics experiments, by means of a multivariate measure of the distance from a feature vector associated to a point in the image to the average feature vector of the global scene, that is, a measure of the local feature contrast.

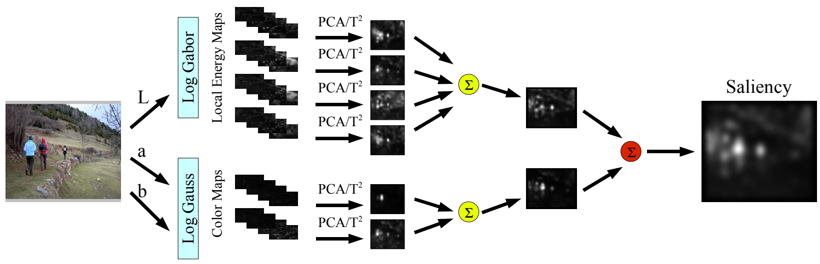

Fig. 1. Bottom-up Saliency Model.

Final Map. The final saliency map is obtained normalizing and smoothing the extracted maps, first within each feature dimension and next with the resulting local energy conspicuity and color conspicuity maps. In this way we obtain a unique measure of saliency for each point of the image.

Computational Complexity. The whole process involves two kinds of operations. Firstly, filtering for decomposition and smoothing has been realized in the frequency domain, as the product of the transfer functions of the input and the filters, using the Fast Fourier Transform (FFT) and its inverse (IFFT). This implies a computational complexity of O(k N log(N) + N), being k the number of operations, a constant independent of the image, and N the number of pixels of the image. The other operation is PCA with a complexity of $O(S^3 + S^2 N)$, being S the number of scales (dimensionality) and N the number of pixels (samples). There exist methods that allow to reduce this complexity in relation to the dimensions [19]. In our case, as the number of scales is relatively small and remains constant, we are interested in the dependency on the number of pixels, being O(N). Therefore, the overall complexity of the algorithm, against the resolution of the image, is established by the use of the FFT, being O(N log(N)).

3 Experimental Results and Discussion

In this work we focus in showing the consistency of the model with a number of outstanding psychophysical results related to bottom-up attention in the HVS. Most of these experiments are yet classic in literature related to visual attention. All of them have been employed to validate any of the state of the art models cited here. They are related to different behaviors observed in the study of human attention: nonlinearity against orientation contrast, efficient (parallel) and inefficient (serial) search, orientation asymmetry, presence-absence asymmetry and Weber's law, and influence of background on color asymmetries. Finally a ROC analysis shows how the model predicts eye fixations better than other models of the state of the art on an open access image dataset.

We start proving the nonlinear behavior of the model against orientation contrast. Hence, examining figure 2 we see how saliency increases quickly from 10º to 30º-35º, an then it remains constant at a saturation value. This is in agreement with the already classic psychophysical experiment conducted by Nothdurft [20]. Other models like the proposed by Harel et al. [10] or Bruce and Tsotsos [7] do not reproduce this result on our images. At least with the code publicly provided by the authors, using the default configuration. This is also the case of the model of Itti and Koch [4][8].

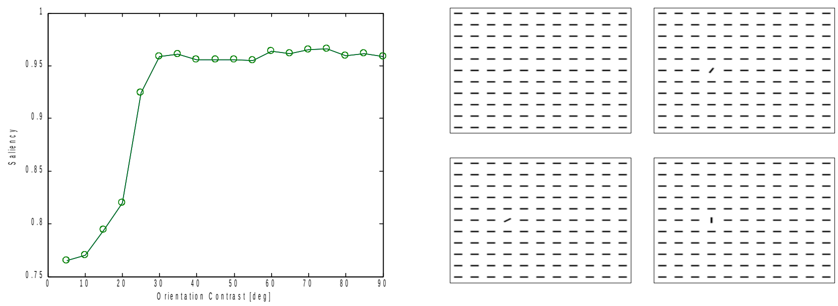

Fig. 2. Nonlinearity of saliency against orientation contrast. Four example images are shown.

Other main issue in the validation of a saliency model is related to the reproduction of the results of efficient search of certain feature singletons and the inefficient search of conjunction singletons. Of course, it is also very important what features give rise to this behavior. Thereby, we can see in figure 3 four typical examples to illustrate all of this. The first one is an image provided by Bruce with the code of his model of saliency [7], reproducing a typical example collected by Wolfe [14]. As we see, character "2" does not stand out among the fives, because it has not any different feature to produce pop-out nor parallel search. However, the tilted "5" does stand out, due to the unique orientation that it presents. In the same way, the red "5" stands out due to its unique color, and the smaller "5" stands out due to its unique size.

The three remaining images are also typical examples of color (second image) and orientation (third image) pop-out -this used by Itti & Koch [4]-, and serial search for a target differing from distractors in a unique conjunction of color and orientation (fourth image).

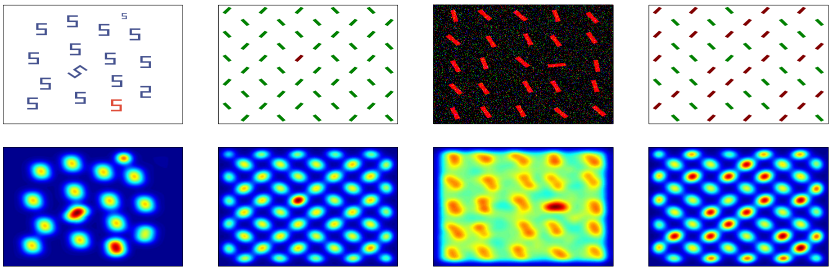

Fig. 3. Four examples related to efficient and inefficient search [14].

We show now how the saliency provided by the model allows to explain several psychophysical phenomena known as search asymmetries [21]. A couple of stimuli differing in a simple feature exhibit different detection times depending on which is the target and which is the distractor. Actually, this term encompasses phenomena of very different nature. It has been pointed, in most cases, that the name itself is not suitable. The reason is that the underlying assumption of symmetric design of the experiment is wrong.

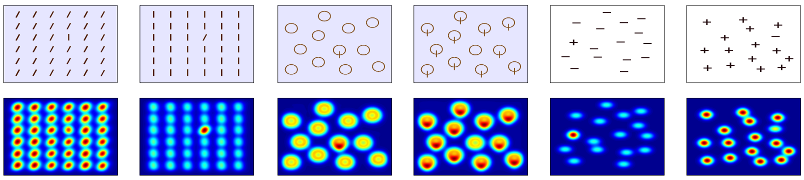

Fig. 4. In the upper row: images used to reproduce the orientation asymmetry (two first), and the presence-absence asymmetry (four remaining). In the lower row the corresponding saliencies are shown.

Orientation asymmetry seems to be however a real asymmetry, not related at all to an asymmetric design of the experiment. This asymmetry really indicates the existence of four privileged canonical orientations [21]. Thus the HVS is observed to present a different behavior depending on the orientation, therefore arising the asymmetry. Given that we have assumed the existence of four canonical orientations (like many other models), it should not be surprising that the model provides with the expected result. In this sense, we see in figure 5 that the relative saliency of a target tilted 80º within vertical distractors is clearly higher than that of a vertical target within 80º tilted distractors, so much so that in the first case occurs a pop-out, inexistent in the second case. Hence, the result provided by the model perfectly matches the observed in psychophysical experiments.

The model reproduces also the asymmetric behavior exhibited by the HVS when target and distractors only differ in the presence or absence of a simple element or feature. In the figure 4 we see two examples typically used to illustrate this fact. As we can see, when the target is the stimulus (circle, dash) with the additional vertical bar present, a pop-out is observed. However when the vertical bar is present in the distractors and absent in the target, there is no pop-out. This is again in agreement with the observed behavior of the HVS. The explanation in the frame of our model matches up with a fact pointed out by Zetchsche [3]: saliency is directly related to the structure content in the image. Therefore, the absence of structure doesn't contribute at all to the increase of saliency, but to its decrease. Or in other words, presence is not a feature. Then, this is not a true asymmetry since the underlying experiment design is not symmetric.

Another consequence of this presence/absence behavior is the so called Weber's law. This law states that an increase in the relative length in a given dimension gives place to a proportional increase of saliency [21]. In figure 5 we show four examples of the 20 images used to test this behavior in our model. The results provided present a very good match with the Weber's law. As we can see, saliency is linear respect to the relative extension of the target. This result is also reproduced by the models proposed by Harel et al. [10] or Gao et al [8] . The model of Bruce and Tsotsos does not reproduce so well this linear behavior. Except for Gao et al. these comparisons have been made with the code publicly provided by the authors, using the default configuration. On the other hand, the model of Itti and Koch fails in reproducing this law[4][8].

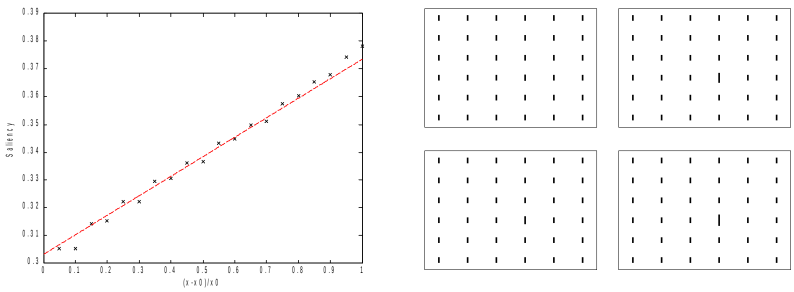

Fig. 5. Saliency against relative increase in length exhibits a linear behavior. Four of the 20 displays used are also shown.

Other important psychophysical result obtained by Rosenholtz et al. [22], shows the way in which background properties influence the color search asymmetries. Consider two stimuli of the same luminance differing only in color. These asymmetries consist in observing a different detection time in a visual search task when target and distractors are exchanged. In this context Rosenholtz et al observed that background properties (color and luminance) have direct effects on these asymmetry, to the point to be reversed, generated or suppressed. This challenges the denomination of asymmetry given to these phenomena, since it implies to forget the existence of the background, that breaks in fact the assumed asymmetry, explaining the results. In figure 6 we can see an example in which the model reproduces this asymmetry and its reversal under a change in background color, in the same way that was observed by Rosenholtz et al. Unlike Bruce and Tsotsos [7] -who employed images elaborated by their selves- we use here reproductions of the images employed by Rosenholtz et al. in their experiments [22]. Hence, in the figure 6 we can see how on a gray background the redder stimulus is more salient, explaining the lower detection time in visual search reported by Rosenholtz et al.. Meanwhile, when the color of the background is changed to red, the situation is reversed and the less red target becomes more salient than the redder target. This is in agreement with the detection time observed by in the experiments.

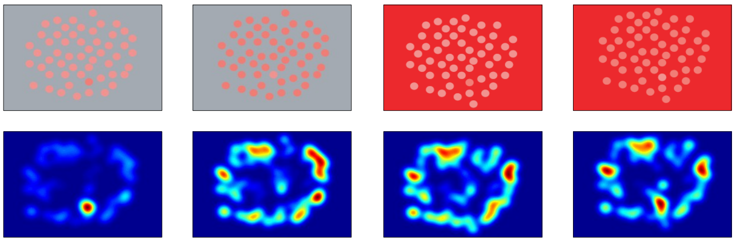

Fig. 6. Example of color asymmetry reversal by a change in color background.

Moreover, our model predicts a higher relative saliency of the redder target in the gray background than the less red target in the red background, again in agreement with the reported search detection times. Note that this is an asymmetry -both in experimental and model results-, not an antisymmetry. In contrast, in the example provided by Bruce and Tsotsos [7], their model seems to show an antisymmetric behavior. This criticism must be taken with caution since they employed their own images, similar but still different to those used in experiments with humans.

We have also observed, in agreement with human behavior, an influence of the contrast of stimuli relative to background in the magnitude of the asymmetry. But it remains out of the scope of this work to quantitatively measure of this effect. On the other hand, Rosenholtz et al. found that changes in color background can also generate or suppress a color asymmetry. In figure 7 we see how the result provided by our model matches well with this behavior of the HVS.

On a gray background, a bluer target is more salient than a less blue target, in agreement with the observed lower detection time. Meanwhile, on a red background the asymmetry almost disappears, with only a slightly higher saliency of the less blue target compared to the bluer. This result appears to be consistent again with the results reported by Rosenholtz et al. [22].

The model of Bruce and Tsotsos, which capable of capturing an antisymmetry (more than an asymmetry), on images of their own -not the ones used in psychophysical experiments-is not, however, able to reproduce correctly other important results like the suppression or generation of an asymmetry by a change in background color. On the other hand, their model provides with a measure of saliency much less graded than the model proposed here.

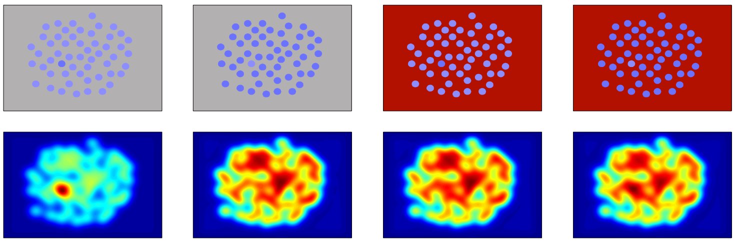

Fig. 7. Example of color asymmetry suppression by a change in color background.

Other state of the art models, like the model by Gao et al [8] or the model by Harel et al [10], have not shown either the capability of reproducing these results. The model proposed by Rosenholtz has been specifically designed to explain these results, and it is not clear how it would capture other important results like orientation or size pop-out. It remains also questionable its capability of reproducing other important results, specially with images constructed in different color spaces. In this sense, it is worth noting that we have used the Lab color space to decompose the image, which is different to the one employed by Rosenholtz et al. to synthesize their images. And this is very interesting, since even the stimuli symmetry may disappear when the color space to represent them is replaced. Therefore, we think that it could be interesting to do experiments similar to these but with stimuli that are symmetric in different color spaces, to see what computational models reproduce better the ensemble of results.

Finally we compare the performance of the model in predicting human eye fixations through ROC analysis. We use an open access image dataset, published by Bruce & Tsotsos. It is made up of 120 images, and of the corresponding fixation data for 20 different subjects. A detailed description of the eye-tracking experiment can be found in [6]. We can see in table 1 the obtained AUC value, that outperforms those obtained by the models of Bruce & Tsotsos and Gao et al. on the same image dataset.

Table 1. AUC values obtained from ROC analysis. (*published by the authors).

Model:	T^2-Based	Bruce and Tsotsos 2009* [7]	Gao et al. 2008* [9]
AUC:	0.791	0.781	0.769

The way in which ROC analysis is performed is often not explicit. In fact, they have been compared AUC values computed in different manners. We have opted here, like many authors [9][10][23], for computing a ROC curve for each image and next averaging the result. Bruce and Tsotsos [7] compute instead a unique curve for all of the images. Results are, in fact, very close using both methods, at least with this dataset. There are also several approaches in treating uncertainty. Hence, Harel et al. or Gao et al. do not provide with it. Bruce and Tsotsos use a procedure [24] that does not reflect the inter-scene variance, and only affects in practice the third decimal value. We think that this requires a deeper analysis, in the line pointed in [25]. It would be worth checking the influence of the type of scene in the result. In fact, the images in this dataset present mainly urban or indoor scenes, which lack of representativity of possibles contexts. We leave this task for a future work since it remains out of the scope of this paper, given the extension required.

4 Conclusions

In this work we have described a simple model of low computational complexity, that resorts to the decorrelation of the responses to a Gabor-like bank of filters. This mechanism is biologically plausible and could have an important role in the influence of NCRF when V1 cells are subjected to natural stimuli [11][17].

We have shown the agreement of our model with an important set of psychophysical phenomena. To our knowledge, none of the models of the state of the art cited here, have been validated with all of these results at once. On the other hand these results are highly relevant references in the literature related to visual attention [14][20][21][22]. Moreover, all of them have served to support the validity of any of the referred models.

Hence our model suitably reproduces the nonlinear behavior against orientation contrast; the efficient search phenomena on orientations, color and size, as well as the inefficient search of conjunctions of orientation and color; the orientation asymmetry; the presence/absence asymmetry and the Weber's law; and the influence of background in color search asymmetries (we expect that a quantitative comparison would reinforce this assessment). Finally by means of a ROC analysis we can claim that our model predicts human fixations better than other models of the state of the art on an open access image dataset.

On the other hand the computational complexity of our model is O(N log(N)). This value clearly improves the ones achieved by other models. For instance Harel et al. report a computational complexity of $O(N^4)$ for their model.

Finally, our model, like that of Bruce & Tsotsos [7], avoids any parameterization of the process, beyond the initial decomposition of the image. However we maintain an initial decomposition which is ordered and suitable for the incorporation, from the beginning, of top-down influences. This tunable design, in the line of many other approaches, makes the model more suitable for machine vision purposes.

Acknowledgments. This work has been financially supported by the Spanish Government through the research project AVISTA (TIN2006-08447), and by the Government of Galicia through the research project PGIDIT07PXIB206028PR.

References

1. Tsotsos, J. K.: Computational foundations for attentive Processes. In L. Itti, G. Rees, J.K. Tsotsos (eds) Neurobiology of Attention, pp.3--7 Elsevier Academia Press (2005)
2. Nothdurft, H.C. : Salience of Feature Contrast. In L. Itti, G. Rees, J.K. Tsotsos (eds) Neurobiology of Attention, pp. 233--239 Elsevier Academia Press (2005)

3. Zetzsche, C.: Natural Scene Statistics and Salient Visual Features. In L. Itti, G. Rees, J.K. Tsotsos (eds) Neurobiology of Attention, pp. 226--232 Elsevier Academia Press (2005)
4. Itti, L., Koch, C.: A saliency-based search mechanism for overt and covert shifts of visual attention. Vision Research, 40, 1489--1506 (2000)
5. Le Meur, O., Le Callet, P., Barba, D., Thoreau, D.: A coherent computational approach to model bottom-up visual attention. IEEE Transactions on Pattern Analysis and Machine Intelligence, 28, 802--17 (2006)
6. Bruce, N., Tsotsos, J.K.: Saliency Based on Information Maximization. Advances in Neural Information Processing Systems, 18, 155-162 (2006)
7. Bruce, N., Tsotsos, J.K.: Saliency, attention, and visual search: An information theoretic approach. Journal of Vision, 9(3), 1--24 (2009)
8. Gao, D., Mahadevan, V., Vasconcelos, N.: On the plausibility of the discriminant center-surround hypothesis for visual saliency, Journal of Vision, vol. 8, 13-- (2008)
9. Gao, D., Mahadevan, V., Vasconcelos, N.: The discriminant center-surround hypothesis for bottom-up saliency. In Proceedings of the Neural Information Processing Systems (NIPS) Conference, Vancouver, Canada, (2007)
10. Harel, J., Koch, C., Perona, P.: Graph-Based Visual Saliency. Advances in Neural Information Processing Systems, 19, 545-552 (2007)
11. Garcia-Diaz, A., Fdez-Vidal, X. R., Pardo, X. M., and Dosil, R.: Local energy variability as a generic measure of bottom-up salience, In Peng-Yeng Yin (ed) Pattern Recognition Techniques, Technology and Applications. In-Teh, Vienna, 1-24, (2008).
12. Olshausen, B. A., Field, D.J.: How Close Are We to Understanding V1?. Neural Computation, 17, 1665--1699 (2005)
13. Itti, L., Koch, C., Niebur, E.: A model of saliency-based visual attention for rapid scene analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 20 (11), 1254-59 (1998)
14. Wolfe, J. M., Horowitz, T. S.: What attributes guide the deployment of visual attention and how do they do it?. Nature Reviews. Neuroscience, 5(6), 495-501 (2004)
15. Field, D.J.: Relations Between the Statistics of Natural Images and the Response Properties of Cortical Cells. Journal of the Optical Society of America A, 4(12), 2379--2394 (1987)
16. Kovesi, P.: Invariant Measures of Image Features from Phase Information. Ph.D. Thesis, The University or Western Australia, (1996)
17. Morrone, M.C., Burr, D.C.: Feature Detection in Human Vision: A Phase-Dependent Energy Model. In: Proceedings of the Royal Society of London B, 235, pp. 221-245, (1988)
18. Vinje, W.E., Gallant, J.L.: Sparse coding and decorrelation in primary visual cortex during natural vision. Science, 287, 1273–1276, (2000).
19. Sharma, A., Paliwal, K.K.: Fast principal component analysis using fixed-point algorithm, Pattern Recognition Letters, 28, 1151-1155 (2007).
20. Nothdurft, H. C.: The conspicuousness of orientation and motion contrast. Spatial Vision, 7, 1993.
21. Treisman, A., Gormican, S.: Feature analysis in early vision: Evidence from search asymmetries. Psychological Review, 95, 15–48, (1988)
22. Rosenholtz, R., Nagy, A.L., Bell, N. R.: The effect of background color on asymmetries in color search. Journal of Vision 4, 224-240, (2004).
23. Tatler, B. W., Baddeley, R. J., and Gilchrist, I. D.:Visual correlates of fixation selection: Effects of scale and time. Vision Research, 45, 643–659, (2005).
24. Cortes, C., and Mohri, M.: Confidence intervals for the area under the ROC curve. Advances in Neural Information Processing Systems 17: Proceedings Of The 2004 Conference, MIT Press, p. 305 (2005)
25. Fawcett, T.: An introduction to ROC analysis, Pattern recognition letters, 27, 861-874 (2006)