The Role of Complexity in Visual Perception: Some Results and Perspectives

Vittoria Bruni
Dept. of SBAI
University of Rome La Sapienza
Rome, Italy
e-mail: vittoria.bruni@uniroma1.it

Domenico Vitulano
Institute for Calculus Applications (IAC)
CNR
Rome, Italy
e-mail: d.vitulano@iac.cnr.it

***Abstract*— In the last years, there has been a change of perspective in approaching some information processing problems: from functional to human perception perspective. In this work, some information theoretic concepts concerning information complexity will be revised and used as formal model for human visual perception based approaches that have been used for solving some image and video processing problems, as for example restoration, detection, tracking and visual quality assessment.**

Keywords-Kolmogorov complexity; asymptotic equipartition property;just noticeable detection threshold; human perception.

I. INTRODUCTION

The vision process has much in common with compression and more in general with the concept of information complexity. The idea behind compression is to reduce the statistical redundancy of the data, for example for storage purposes. In particular, the compressor is an algorithm whose aim is to produce a string of bits whose length is less than the string of bits that are necessary for representing the original data. On the other hand, the Kolmogorov complexity [1] of a signal is the length of the shortest computer program that outputs it. Several neurological studies proved that in the observation of a scene (early vision) very few points are used by human observers to understand scene content [2]. These points are foveated, i.e., human attention decreases as one moves away from them, and they are the ones showing independence between local luminance mean and contrast. In other words, few fixation points are necessary to code scene information and to understand (learn) its content. Vision process can be then modeled as an encoder/decoder system, where human eye is the decoder (final receiver of image information) or the transmission channel. As a result, information theory concepts can be used for coding visual information in order to

- design algorithms for objective evaluation of quality in a way that is consistent with subjective human evaluation;
- develop automatic algorithms adapted for optimal perceptual quality.

This strategy provides a new perspective on image content representation where image pixels are no longer seen as simple probabilistic data but should be accounted for by Human Visual System (HVS) limits and rules.

Next section provides some examples while the last section draws the conclusions.

II. SOME NOTES ON VISUAL COMPLEXITY

Human perception can guide digital restoration according to a new paradigm: reducing the visual contrast of image anomalies till they are masked by surrounding information. As a result, according to the task, image degradation is not removed but its contribution is hidden in the image according to the visual contrast masking effect; in other words, it represents negligible information and does not contribute to signal complexity/quality [3]. Figure 1 shows an example: noise is equally distributed in the image but it is not perceived in the textured region, as correctly measured by the Structural SIMilarity index (SSIM); therefore, denoising can be applied only in regions where noise is visible in order to avoid the introduction of visible artifacts due to oversmoothing or misalignments --- the latter are caused by a not correct motion estimation in case of video denoising. The same concepts have been successfully applied for tuning and balancing quantization errors in image compression.

In this context, looking at the scene inspection as a random walk, image anomalies capture human eye attention at first sight as they are perceived as foreign objects in the scene, independently of scene complexity. As a result, the automatic detection of image anomalies, as well as a moving target, is allowed by looking at them as those resulting visually different from the remaining image content. In other words, the anomaly represents a "surprise" and then the code of visualized information increases whenever the "surprise" occurs. More in general, human perception offers new ways for the representation of image content; for example, in a hierarchical way: from the most visible to the less one (saliency maps); or using perception-based sampling rules (fixation points). It seems based on a measure of object complexity in the sense of Kolmogorov complexity, i.e., the probability of finding that object in the nature. It is obvious that this kind of interpretation generalizes the common concept of surprise which is measured by the Shannon entropy. Therefore, the challenge is to define novel mathematical tools able to directly account for these concepts. The minimum description length represents a useful tool in this sense as it provides a formal equivalence between coding and learning [4]. It is based on the Occam razor concept [5] for which the simplest solution is more likely to be the best, not necessarily the real one. However, some refinements are required as it represents, in some

 ISBN: 978-1-61208-724-5

sense, a computable interpretation of Kolmogorov complexity which, in contrast, is not computable even though more close to the real scenario. Classical information theory concepts can be then used not only for coding but also for learning: simple and few information is enough for describing a more complex one. In this setting, feature extraction is a way of coding and if the characteristics of the decoder (human eye) are embedded in the model, feature extraction becomes a way of coding based on human vision. An example is shown in Figure 2 where fixation points are used for assessing image quality, or where human perception is used for segmenting dermoscopy images.

III. Conclusions and future work

The simulation of human visual system represents the key issue for developing novel and effective solutions for some classical visual information-based processing problems. In this paper, the double role of complexity, in the sense of information compression, has been discussed. On the one hand, complexity provides a theoretical tool useful for formal modeling. Conversely, the mechanisms regulating human perception can offer the way of defining novel paradigms for information coding. The goal of future research is then to provide formal models and mathematical methods able to represent and process such visual information. In particular, since compact representation is the key issue in several applications, the definition of a visual perception-inspired multiscale transform represents an interesting and promising challenge.

References

[1] R. Cilibrasi and P. M. B. Vitányi, "Clustering by compression," IEEE Transactions onInformation Theory, vol. 51 no. 4, pp. 1523-1545, 2005.

[2] R. A. Frazor and W. S. Geisler, "Local luminance and contrast in natural images," Vision Research, vol. 46, pp. 1585-1598, 2006.

[3] H. R. Sheikh, A. C. Bovik, and G. De Veciana, "An information fidelity criterion for image quality assessment using natural scene statistics," IEEE Trans. Image Process., vol. 14, no. 12, pp. 2117–2128, 2005

[4] V. Bruni, E. Rossi, and D. Vitulano, "On the equivalence between Jensen-Shannon divergence and Michelson contrast," IEEE Trans. on Information Theory, vol. 58, no. 7, pp. 4278-4288, 2012

[5] P. D. Grünwald, "A tutorial introduction to the minimum description length principle," Advances in Minimum Description Length: Theory and Applications, I. J. Myung, P. D. Grünwald, and M. A. Pitt, Eds. Cambridge, MA: MIT Press, 2004.

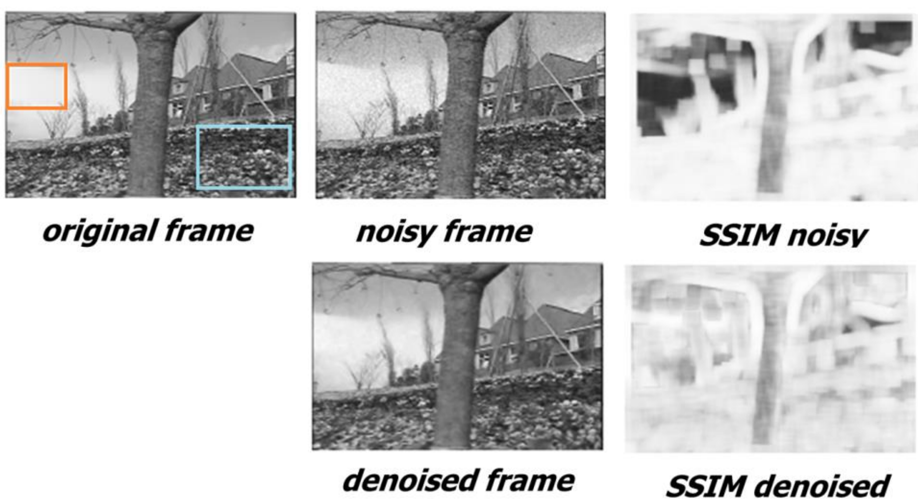

Figure 1. Top) Noise is differently perceived in the image: more visible in the red box than in the cyan box. Accordingly, SSIM in the red box is darker than in the cyan box. Bottom) Denoising applied to the noisy image increases the visual quality in the red box but decreases the one in the cyan box. Accordingly, SSIM of the denoised image in the cyan box is darker than the one of the noisy one in the same box.

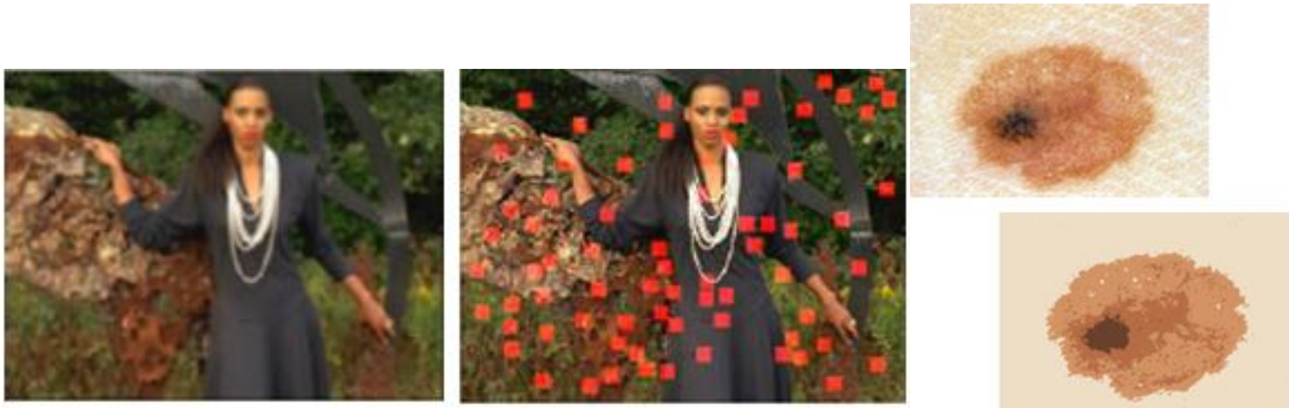

Figure 2. Left) Blurred image. Middle) Red boxes are the points necessary for assessing image quality in a way which is consistent with HVS. Right) Dermoscopic image and its human perception based-quantized version using the least number of quantization bins.

 ISBN: 978-1-61208-724-5