

Influence de la structure spatiale des images lors de la quantification de la résonance stochastique dans la perception visuelle

Agnès DELAHAIES, David ROUSSEAU, Jean-Baptiste FASQUEL, François CHAPEAU-BLONDEAU

Laboratoire d'Ingénierie des Systèmes Automatisés (LISA), Université d'Angers
62 avenue Notre Dame du Lac, 49000 Angers, France

agnes.delahaies@univ-angers.fr, david.rousseau@univ-angers.fr,
jean-baptiste.fasquel@univ-angers.fr, chapeau@univ-angers.fr

Résumé – Nous étudions l'influence des structures spatiales des images sur les effets non linéaires de bruit utile, ou résonance stochastique, pour la transmission et le traitement de l'information dans le domaine de l'imagerie. Nous illustrons notre propos à partir d'un système imageur non linéaire minimal, similaire à celui historiquement utilisé lors de la première démonstration d'un effet de bruit utile dans la perception visuelle. Pour quantifier l'influence bénéfique du bruit dans la perception visuelle, des mesures de similarité usuelles ne prenant pas en compte les structures spatiales des images ont jusqu'ici été utilisées. Nous exposons en quoi ces mesures de similarité n'expliquent pas complètement la perception visuelle de résonance stochastique, et nous montrons dans ce contexte l'intérêt de mesures de similarité inspirées du système visuel comme les descripteurs locaux de type SIFT. Cette contribution montre une quantification nouvelle, par des SIFT, de la résonance stochastique dans les images et ouvre des perspectives pour l'étude des mécanismes psychovisuels en jeu dans les effets non linéaires de bruit utile en imagerie.

Abstract – We study the influence of spatial structures of images on a nonlinear useful noise effect, also called stochastic resonance, on the transmission or process of information in the domain of imaging. We provide illustrations with a minimal nonlinear imaging system, similar to the one historically introduced in the first demonstration of a useful noise effect in visual perception. Until now, the constructive influence of noise in visual perception has quantitatively been assessed with measures of similarity that do not take into account the spatial structures in images. We explain why such measures of similarity do not fully explain the stochastic resonance visually perceived and we demonstrate in this context the interest of measures of similarity inspired from the visual system like local feature descriptor of SIFT type. This constitutes a new quantification, with SIFT, of stochastic resonance in images and opens perspectives for the study of the psychovisual mechanism at work in the nonlinear useful noise effects in imaging.

1 Introduction

Il est désormais un fait établi (voir [1] pour une récente synthèse) que la présence d'un niveau de bruit non nul dans certains systèmes non linéaires peut profiter au traitement de l'information par ces systèmes. Introduits en physique dans les années 1980, puis enregistrés dans de nombreux autres domaines dont ceux des neurosciences et des sciences cognitives, ces effets de bruits utiles, aussi désignés sous le terme rassembleur de résonance stochastique, ont été largement étudiés en sciences de l'information (voir également [1] et [2, 3] pour le GRETSI). L'injection délibérée de bruit, à des niveaux contrôlés et prévus via les analyses théoriques de la résonance stochastique, est utile dans des situations contraintes où un système non linéaire en charge de tâches de traitement de l'information est d'emblée, i.e. à bruit natif, positionné de façon défavorable pour la transmission de l'information. Si la compréhension des différents mécanismes par lesquels la résonance stochastique opère a beaucoup progressé ces dernières années, notamment via la modélisation théorique et l'expérimentation sur des systèmes physiques, il n'en demeure pas moins que certaines questions restent ou-

vertes [1] et notamment la question centrale de savoir si le système nerveux, qui traite de grandes quantités d'information en présence de bruit et de fortes non linéarités intrinsèques, utilise in vivo la résonance stochastique. Cette question peut être abordée à différentes échelles partant du neurone isolé, en passant par des associations de neurones jusqu'au niveau d'intégration maximale par des expériences psychosensorielles. Dans ce travail, nous nous plaçons au niveau des expériences psychosensorielles. Nous revenons sur les pas d'une des premières démonstrations psychovisuelles de résonance stochastique en imagerie [4] et nous montrons en quoi la compréhension des mécanismes mis en jeu dans ces expériences nécessite d'être approfondie.

2 Perception visuelle de la résonance stochastique

On considère la situation de référence étudiée dans [4] et [5] où une image initiale x ajoutée à un bruit n est transmise par

un quantifieur $g(\cdot)$ à deux niveaux selon $\mathbf{y} = g(\mathbf{x} + \mathbf{n})$. La Fig. 1(a–d) montre une traduction visuelle d’un effet de transmission non linéaire d’image aidée par le bruit. En complément de la perception visuelle impliquée dans la Fig. 1, on considère une évaluation de la transmission non linéaire par des mesures quantitatives objectives. Les mesures de similarité entrée–sortie entre les images \mathbf{x} et \mathbf{y} sont des mesures statistiques pixel à pixel. Nous utilisons les trois mesures de similarité utilisées dans [5] avec l’écart quadratique moyen

$$E(\mathbf{x}, \mathbf{y}) = \sqrt{\langle (x - y)^2 \rangle}, \quad (1)$$

le coefficient d’intercovariance normalisé

$$C(\mathbf{x}, \mathbf{y}) = \frac{\langle xy \rangle - \langle x \rangle \langle y \rangle}{\sqrt{\langle x^2 \rangle - \langle x \rangle^2} \sqrt{\langle y^2 \rangle - \langle y \rangle^2}}, \quad (2)$$

et l’indice SSIM

$$S(\mathbf{x}, \mathbf{y}) = \frac{4(\langle xy \rangle - \langle x \rangle \langle y \rangle) \langle x \rangle \langle y \rangle}{(\langle x^2 \rangle - \langle x \rangle^2 + \langle y^2 \rangle - \langle y \rangle^2)(\langle x \rangle^2 + \langle y \rangle^2)}, \quad (3)$$

où $\langle \cdot \rangle$ est une moyenne d’ensemble sur les pixels des images. Nous ajoutons aux trois mesures E, C, S l’information mutuelle de Shannon

$$I(\mathbf{x}, \mathbf{y}) = H(\mathbf{y}) - H(\mathbf{y}|\mathbf{x}), \quad (4)$$

où $H(\cdot)$ est l’entropie standard de Shannon

$$H(\mathbf{y}) = \int_{\mathbf{y}} -dy p_{\mathbf{y}}(y) \log_2[p_{\mathbf{y}}(y)], \quad (5)$$

et

$$H(\mathbf{y}|\mathbf{x}) = \int_{\mathbf{x}} dx p_{\mathbf{x}}(x) \int_{\mathbf{y}} -dy p_{\mathbf{y}|\mathbf{x}}(y) \log_2[p_{\mathbf{y}|\mathbf{x}}(y)] \quad (6)$$

avec la densité conditionnelle définie par $p_{\mathbf{y}|\mathbf{x}}(y) dy = \Pr\{\mathbf{y} \in [y, y + dy] | \mathbf{x} = x\}$, et la densité marginale qui est $p_{\mathbf{y}}(y) = \int_{\mathbf{x}} dx p_{\mathbf{x}}(x) p_{\mathbf{y}|\mathbf{x}}(y)$. On observe l’évolution des mesures de similarités E, C, S et I sur la Fig. 3 lorsque le niveau σ du bruit augmente. Sur la Fig. 3, l’évolution non monotone des mesures traduit un effet de résonance stochastique ou de transmission non linéaire d’image aidée par le bruit, avec des niveaux de bruits optimaux non nuls proches pour les différentes mesures et situés autour du niveau de bruit de la Fig. 1(c) en accord avec la perception visuelle à $\sigma_{opt} \approx 0.5$. On considère à présent la séquence d’images de la Fig. 2(e–h). Sur la Fig. 2(f–h), notamment la Fig. 2(g), la perception visuelle n’enregistre pas de transmission d’image aidée par le bruit, a contrario de la situation de la Fig. 1(b–d). Toutefois, sur la Fig. 2(e–h) comme sur la Fig. 1(a–d), on a affaire à des images initiales \mathbf{x} en (a) et en (e) qui présentent la même proportion $p_0 = 0.16$ de pixels noirs. Comme conséquence, les mesures quantitatives pixel à pixel de la Fig. 3 caractérisent également la situation de la Fig. 2(e–h), autant que celle de la Fig. 1(a–d). Sur la base de ces mesures de la Fig. 3, on enregistre un effet de résonance stochastique équivalent sur la Fig. 2(e–h) comme sur la Fig. 1(a–d). Alors que la perception visuelle sur la Fig. 2(e–h) n’enregistre

pas de résonance stochastique. La perception visuelle sur la Fig. 1(a–d) opère sur des réalisations uniques du bruit \mathbf{n} sur chaque image successive, au contraire des mesures quantitatives de la Fig. 3 qui sont des moyennes d’ensemble sur le bruit \mathbf{n} . Dans l’étude [4], il est choisi de considérer des images aux structures spatiales minimales (un sinus wobulé) soumis à un panel d’observateurs et avec une variation temporelle du bruit dans l’image. Les effets de résonance stochastique sont ainsi mesurés dans [4] en moyenne. Pourtant, comme le montre la Fig. 1(a–d), un effet de transmission non linéaire aidée par le bruit est perceptible sur une seule réalisation. Si la perception visuelle identifie un effet de bruit utile sur la Fig. 1(a–d) et non sur Fig. 2(e–f), c’est probablement en exploitant des structures spatiales présentes dans l’image de la Fig. 1(a–d) et absentes dans l’image de la Fig. 2(e–h). La prise en compte des structures spatiales dans les images constitue ainsi une voie d’investigation nouvelle dans l’analyse quantitative de la résonance stochastique que nous nous proposons d’aborder maintenant.

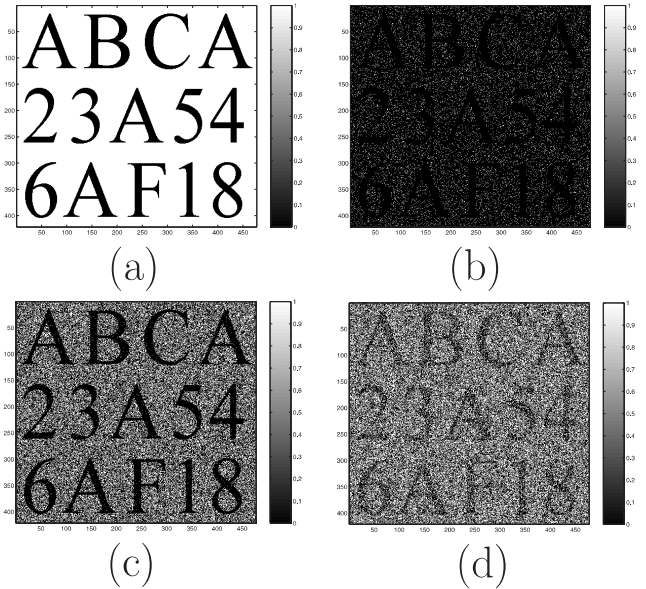


FIG. 1: (a) image binaire initiale \mathbf{x} . (b)–(d) images binaires \mathbf{y} en sortie d’un quantifieur à deux états de seuil $\theta = 1.1$, avec le bruit \mathbf{n} pris gaussien centré d’amplitude rms $\sigma = 0.07$ (b), $\sigma = 0.49$ (c) et $\sigma = 1.5$ (d).

3 Influence des structures spatiales des images

Dans cette contribution, nous travaillons, comme sur la Fig. 1, avec des bruits non corrélés spatialement et nous étudions l’influence des structures spatiales de l’image \mathbf{x} porteuse d’information sur la perception visuelle de la résonance stochastique. Les mesures de similarité utilisées sur la Fig. 3 réalisent des statistiques d’ensemble qui donnent à chaque pixel le même poids et ne permettent ainsi pas de rendre compte d’éventuelles

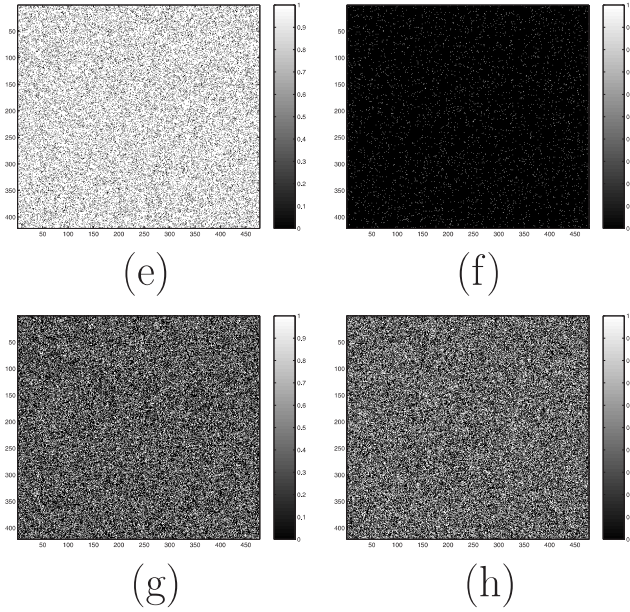


FIG. 2: Expérience de résonance stochastique visuelle similaire à celle de la Fig. 1 avec comme image binaire initiale (e) ayant même distribution de pixels blancs et noirs que Fig. 1(a). Les panneaux (f)–(h) sont respectivement obtenus avec les niveaux de bruits identiques à ceux des panneaux Fig. 1(b)–(d)

zones saillantes dans les structures des images. Or, ces zones saillantes en terme d'orientation ou de contraste de niveaux de gris sont connues pour fixer l'attention [6] c'est-à-dire conférer un poids plus important à certaines zones d'un point de vue informationnel. Nous proposons ici de considérer des mesures de similarité inspirées du fonctionnement du système visuel. Pour illustration, nous utilisons les descripteurs locaux tels que les SIFT (Scale Invariant Feature Transform) que nous mettons en œuvre dans leur forme originelle introduite dans [7], détaillée pédagogiquement dans [8], et pour lesquels des implantations logicielles didactiques sont librement disponibles sur [9]. Les SIFT sont basés sur l'extraction de descripteurs locaux invariants en échelle et par rapport à l'orientation des contours des images. Nous mesurons le nombre de descripteurs locaux communs entre une image de sortie y et l'image initiale x en ne conservant que les descripteurs associés à des points homologues dans les deux images. Sur la Fig. 4, l'évolution non monotone de cette mesure traduit un effet de résonance stochastique uniquement pour l'image initiale Fig. 1(a) structurée spatialement, alors que le nombre de descripteurs locaux identifiés dans l'image initiale Fig. 2(e) non structurée reste à un niveau relatif artefactuel, bas et quasi constant. Pour l'image initiale Fig. 1(a) structurée spatialement, la mesure de similarité de la Fig. 4 basée sur les SIFT est maximisée pour un niveau optimal σ_{opt} non nul de bruit. La résonance stochastique est enregistrée via les SIFT, de plus $\sigma_{opt} \approx 0.5$ est inchangé par rapport aux mesures de similarité de la Fig. 4. La comparaison des Figs. 3 et 4 montre la pertinence de notre approche pour expliquer quantitativement et en terme de mécanisme psy-

chovisuel de saillance locale la différence observable entre les Fig. 1(a–d) et Fig. 2(e–f).

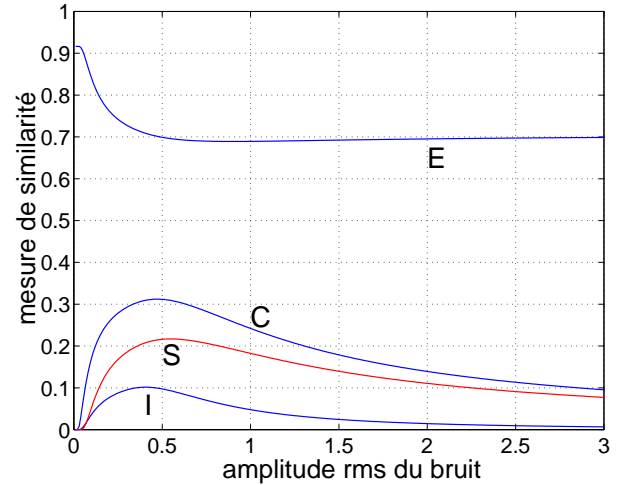


FIG. 3: Mesures de similarité entre l'image binaire initiale x Fig. 1(a) ou Fig. 2(e) et l'image binaire y en fonction du niveau σ de bruit n pris gaussien centré. L'écart quadratique moyenne rms $E(x, y)$, le coefficient d'intercovariance normalisé $C(x, y)$, l'indice SSIM de [5] $S(x, y)$, et l'information mutuelle de Shannon $I(x, y)$ sont identiques pour les Fig. 1(a) et Fig. 2(e).

4 Conclusion et perspectives

Nous avons montré la nécessité de poursuivre les analyses des effets non linéaires de bruit utile dans la perception visuelle afin de comprendre l'influence des structures des images. Pour ce faire, nous avons proposé de reprendre un schéma classique de résonance stochastique avec un capteur à seuil et un bruit additif gaussien comme banc d'essai pour des mesures psychovisuelles. Cette approche ouvre de multiples voies d'investigation que nous développons actuellement dans les directions suivantes.

Il existe de nombreux descripteurs locaux [10] avec des propriétés d'invariance diverses. Les SIFT utilisés ici possèdent par exemple des propriétés d'invariance en rotation que nous n'avons pas mobilisées dans ce rapport et il serait intéressant de tester ces propriétés dans le contexte de la résonance stochastique. Des mesures psychovisuelles de cartes de saillance [6] modélisant l'attention peuvent également être évaluées pour quantifier la résonance stochastique. Au niveau de l'œil lui-même, les effets de résonance stochastique ont été étudiés expérimentalement et en simulation numérique [3] à des échelles de temps de l'ordre de la microseconde où l'on enregistre des saccades aléatoires des positions du regard sur la rétine. La résonance stochastique a également été mise en évidence avec des stimuli visuels sur des durées de l'ordre du dixième de seconde [13]. Nous expérimentons actuellement des échelles temporelles intermédiaires en étudiant le comportement du re-

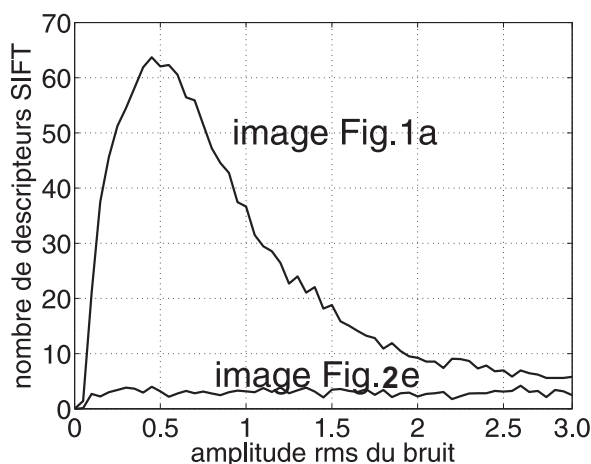


FIG. 4: Nombre de descripteurs SIFT homologues selon [7] entre l'image binaire initiale x Fig. 1(a) ou Fig. 2(e) et l'image binaire y de sortie en fonction du niveau σ de bruit n pris gaussien centré. Contrairement aux mesures de similarités de la Fig. 3 le nombre de descripteurs SIFT homologues diffèrent pour les Fig. 1(a) et Fig. 2(e).

gard au cours du temps face à la résonance stochastique dans les images au moyen d'un dispositif d'eyetracking, où des points de fixations de l'œil sont enregistrés à l'échelle de la milliseconde [11]. La perception visuelle est également impactée par l'échelle d'observation des images [12]. Nous étudions ainsi l'influence de la distance d'observation des images de résonance stochastique sur le niveau de bruit optimal.

Le modèle de résonance stochastique utilisé ici met en jeu un couplage additif du bruit comme celui que l'on rencontre dans les techniques de dithering [1]. D'autres couplages image-bruit existant au niveau même des capteurs ont montré qu'ils pouvaient donner lieu à des effets de résonance stochastique. C'est le cas en particulier du bruit multiplicatif de speckle [2] typique de l'imagerie cohérente qu'il conviendrait de tester avec les différentes mesures psychovisuelles évoquées ici. Une spécificité des patterns de speckle est qu'ils présentent eux-mêmes des structures spatiales sous forme de grains. Il serait ainsi important d'étudier l'influence de la taille de ces grains sur la perception des structures des images dans la résonance stochastique.

À un niveau plus large de considération, des effets psychosensoriels de la résonance stochastique ont été enregistrés sur d'autres sens que la vision avec, en particulier, l'audition et le toucher [1, 13]. Dans ces domaines de la perception, les mesures théoriques utilisées pour quantifier la résonance stochastique ont été, jusqu'à présent, des mesures ne prenant pas en compte les structures temporelles des signaux acoustiques ou tactiles appliqués et, il serait donc intéressant de considérer une transposition de l'approche que nous proposons ici pour les images à ces autres domaines où notre système nerveux, pris dans son ensemble, se montre capable de percevoir des effets bénéfiques du bruit pour des tâches variées de traitement de l'information.

Références

- [1] M. D. McDonnell, N. G. Stocks, C.E. M. Pearce, D. Abbott, "Stochastic Resonance: From Suprathreshold Stochastic Resonance to Stochastic Signal Quantization", *Cambridge University Press*, Cambridge, 2008.
- [2] S. Blanchard, D. Rousseau, D. Gindre, F. Chapeau-Blondeau, "Transmission d'image assistée par le bruit en imagerie cohérente et incohérente", *Proc. 21^e Colloque GRETSI*, pp. 813-816, 11-14 Sept. 2007, Troyes, France.
- [3] S. Zozor, C. Duchêne, P.-O. Amblard, "Échantillonnage aléatoire de la rétine et micro-saccades : amélioration du traitement de l'information par le bruit?", *Proc. 21^e Colloque GRETSI*, pp. 241-244, 11-14 Sept. 2007, Troyes, France.
- [4] E. Simonotto, M. Riani, C. Seife, M. Roberts, J. Twitty, F. Moss, "Visual perception of stochastic resonance", *Physical Review Letters*, vol. 78, pp. 1186-1189, 1997.
- [5] D. Rousseau, A. Delahaies, F. Chapeau-Blondeau, "Structural similarity measure to assess improvement by noise in nonlinear image transmission", *IEEE Signal Processing Letters*, vol. 17, pp. 36-39, 2010.
- [6] L. Itti, C. Koch, E. Niebur, "A model of saliency-based visual attention for rapid scene analysis", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, pp. 1254-1259, 1998.
- [7] D. G. Lowe, "Object recognition from local scale-invariant features", *Proceedings of the International Conference on Computer Vision*, vol. 2, pp. 1150-1157, 20-25 Sept. 1999, Corfu, Greece.
- [8] D. G. Lowe, "Distinctive image features from scale-invariant keypoints", *International Journal of Computer Vision*, vol. 2, pp. 91-110, 2004.
- [9] <http://www.vlfeat.org/~vedaldi/code/sift.html>.
- [10] J. Li, N. M. Allison "A comprehensive review of current local features for computer Vision", *Neurocomputing*, vol. 71, pp. 1771-1787, 2008.
- [11] A. T. Duchowski, "Eye Tracking Methodology: Theory and Practice," *Springer* Berlin, 2007.
- [12] A. Delahaies, D. Rousseau, F. Chapeau-Blondeau "Joint acquisition-processing approach to optimize observation scales in noisy imaging", *Optics Letters*, vol. 36, pp. 972-974, 2011.
- [13] F. Moss, L. M. Ward, W. G. Sannita, "Stochastic resonance and sensory information processing: A tutorial and review of application", *Clinical Neurophysiology*, vol. 115, pp. 267-281, 2004.