Visual Perception of Stochastic Resonance

Enrico Simonotto,[1,3] Massimo Riani,[1] Charles Seife,[2,*] Mark Roberts,[2] Jennifer Twitty,[3] and Frank Moss[3]
[1]*INFM-Unitá di Genova and Dipartimento di Fisica, Universitá di Genova, 16146 Genova, Italy*
[2]The Economist, *25 St. James's Street, London, SW1A 1HG, England*
[3]*Center for Neurodynamics, University of Missouri at St. Louis, St. Louis, Missouri 63121*

(Received 31 October 1996)

Stochastic resonance can be used as a measuring tool to quantify the ability of the human brain to interpret noise contaminated visual patterns. Here we report the results of a psychophysics experiment which show that the brain can *consistently and quantitatively* interpret detail in a stationary image obscured with time varying noise and that both the noise intensity and its temporal characteristics strongly determine the perceived image quality. [S0031-9007(97)02344-2]

PACS numbers: 87.10.+e, 05.40.+j

Internal noise has long been associated with the nervous system [1–7], thus prompting speculations that it may serve a useful role in neural computation [2–4,8], or signal averaging by summation across a population of neurons in parallel [9,10]. Though tantalizing, this idea remains undemonstrated in any biological experiment. What has been shown is that *external* noise added to a weak signal can enhance its detectability by the peripheral nervous systems of crayfish [11], crickets [12], rats [13], and humans [14,15] including possible medical applications [14,16], and within membranes [17] by the process of *stochastic resonance* (SR) [18–20]. Excepting a recent experiment which demonstrated SR in the human tactile system [21], the results of these works were obtained by computer analysis of neural recordings. But how does a complex organ such as a brain analyze similar weak and noisy signals?

SR has shown in several experiments that external noise added to a weak environmental signal can enhance the information content of evoked responses in the peripheral nervous system [11–16]. In these experiments, recordings of temporal sequences of neural action potentials were made and analyzed by computer for the signal-to-noise ratio [11,14–16], Shannon information rate, and the transinformation [12] or stimulus-response coherence or action potential timing precision [13]. Though noise enhanced information in the peripheral nervous system was demonstrated in all experiments, the question remains whether animals, including man, can make use of the enhancement. Specifically, could the computers previously used for signal and noise analysis in the physiological experiments be replaced by the human brain in a psychophysics experiment, and if so, would the results be comparable? We show here that the results are comparable, accurate, and repeatable and that the process is more efficient for a stationary image with time varying noise than for the same image with static noise.

Our experiment works with the human visual system [22–25] and derives from the simplest paradigm of SR: the nondynamical or threshold theory [26,27]. As shown in Fig. 1(a), the necessary components are a threshold, a subthreshold signal, and additive noise. The system is assumed capable of transmitting single bits of information, each of which marks a threshold crossing, as shown by the pulse train above. Figure 1(b) is a visual realization, where the subthreshold "signal" is an image digitized on a gray scale and depressed beneath

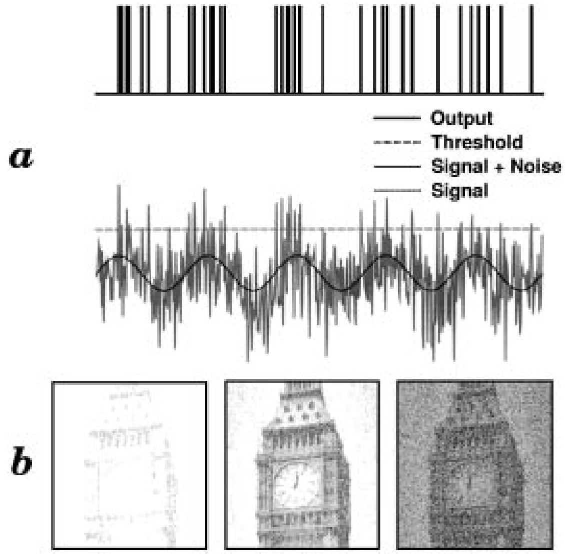

FIG. 1. (a) The threshold paradigm of SR. A subthreshold signal is shown by the sine wave plus Gaussian noise whose mean lies Δ below the threshold (horizontal line). Each positive going threshold crossing is marked by a standard pulse as shown above, the temporal sequence of which transmits the only information available about the signal through the system. (b) Visual images composed of a single signal—the picture of Big Ben—digitized on a 1 to 256 gray scale with a spatial resolution of 256 by 256 pixels. A random number ξ, from a Gaussian distribution with zero mean and standard deviation σ, is added to the original gray value I, in every pixel. Thus the noise in each pixel is incoherent with that in all other pixels though the standard deviation is the same for all. The resulting image is then threshold filtered according to the rule: if $I + \xi < \Delta$, the gray value in that pixel is replaced with 256 (white), otherwise with 1 (black), in this example. The pictures shown were made for $\Delta = 30$ and for $\sigma = 10$, 90, and 300 on the gray scale (left to right).

a threshold. Noise is added to the gray value in every pixel, and the result compared to a threshold value. Pixels containing a gray value above the threshold are painted white and the others black. Thus in every pixel one bit of information is retained, white marking a pixel wherein the threshold has been crossed and black otherwise. Figure 1(b) shows the result of adding noise of three intensities, increasing from left to right. As with SR, there is an optimal noise intensity [Fig. 1(b), center picture] which maximizes the information content.

Pictures of edifices are not suitable for a quantitative determination of image quality in psychophysics experiments. Instead, a pattern determined by a function of amplitude A, or contrast, with variable spatial frequency, as shown in Fig. 2(a) is used. A maximum contrast, noise free strip with gray values determined by the function in (a) and without threshold filtering is shown in the lower panel. Three example strips for a single noise intensity σ (near to the optimal value) and threshold Δ, but with contrast values decreasing from bottom to top are shown in Fig. 2(b). Subjects were presented with a sequence of images, each composed of a set of seven such strips. The threshold remained constant throughout the experiment, but the standard deviation of the noise for each presentation was chosen randomly from a set of ten values. Subjects were asked to count up from the bottom until they reached a strip for which they could no longer distinguish a specified fine feature, for example, one of the high frequency vertical bands toward the right of the strips. Thus they find their perceptive contrast threshold A_{th}, for that particular feature and noise intensity.

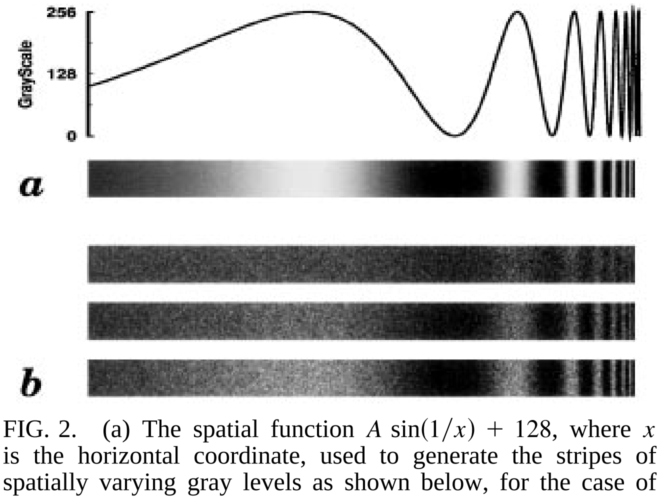

FIG. 2. (a) The spatial function $A \sin(1/x) + 128$, where x is the horizontal coordinate, used to generate the stripes of spatially varying gray levels as shown below, for the case of maximum contrast (amplitude) $A = 128$. (b) Three example stripes for decreasing contrast $A = 128$, 78, and 28 (bottom to top), threshold filtered with $\sigma = 250$ and $\Delta = 150$. The temporal development of the noise was generated by writing a new realization of the noise into each frame, threshold filtering, and presenting the images on a high speed computer monitor at a frame rate of 60 Hz (about 16.6 ms per frame), a time interval which is considerably faster than known averaging times in the visual system [29].

We emphasize that Figs. 1(b) and 2(b) show only the *static* effect of a single realization of the noise, the only possibility for presentation on a printed page. All results reported in this Letter were, however, obtained using images created with *temporal* variations of the noise, which, when switched on, result in a striking improvement in perceived picture quality. Subjects viewed computer generated animations which presented the seven strip image with time dependent noise [28].

A graph of perceptive threshold A_{th} versus noise intensity σ for a single session with an example subject is shown in Fig. 3(a). Subjects indicate high perceptive thresholds for both small and large noise intensities and

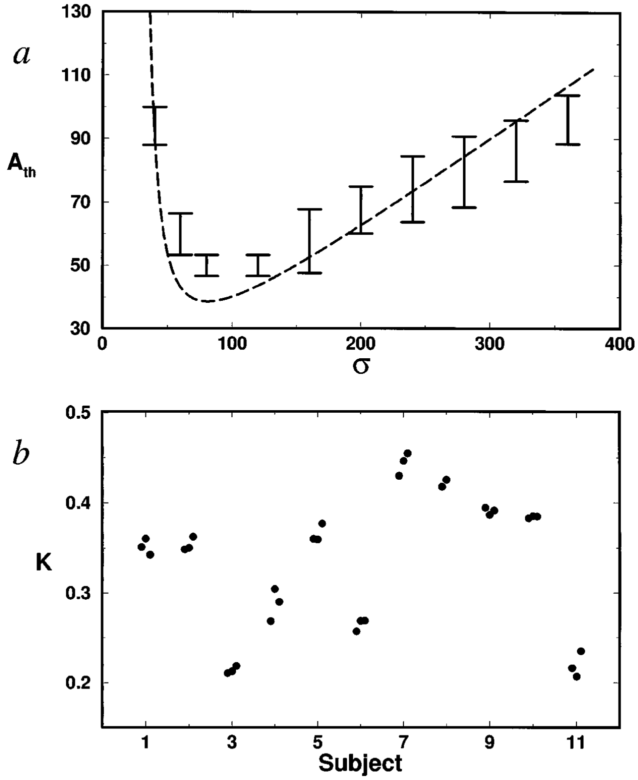

FIG. 3. (a) Perceptive contrast threshold A_{th} versus noise intensity σ for subject number 6. Each subject was presented with 10 different noise intensities chosen at random with each noise intensity visited (randomly) 10 times for a total of 100 presentations in a single session. The error bars are the standard deviations of the 10 determinations of A_{th} by the subject at each noise level. The solid line is Eq. (1) with $\Delta = 81$, and for $K = 0.289$ determined by nonlinear least squares fit. (b) The K values were obtained in three sessions each for each of 11 subjects: 5 males and 6 females between the ages of 18 and 26 with no known visual impairments other than eye glasses. The experiments were performed in a small, dimly illuminated room. The subjects viewed the monitor screen at a comfortable distance (about 40 cm). The total radiant power density from the monitor screen subtended at the approximate location of the subject's eyes was 5.2 $\mu\text{W/cm}^2$ for all sessions and varied less than 10% during a session over the entire range of presented noise levels. The room was quiet and free of distractions.

a minimum threshold at an optimal noise intensity. The solid curve is taken from threshold SR theory [26] by solving for the signal amplitude A, which results in a specific signal power density, proportional to K, as obtained from the power spectrum of a pulse train as in Fig. 1(a),

$$A_{\text{th}} = K\sigma \exp[\Delta^2/2\sigma^2]. \qquad (1)$$

This equation was fit to the data using K as the only adjustable parameter. A significant finding is that subjects respond to the *power* in the signal image rather than to an inherent, or subjective, signal-to-noise ratio [26]. The value of K is a quantitative measure of the sensitivity of the subject to the power contained in the signal part of the image and thus of the subject's ability to distinguish detail in the noise contaminated scene. We have found this to be a remarkably robust and repeatable measure as shown in Fig. 3(b) by the tight clusters of three K values for each subject, obtained in sessions often separated by a week or more.

The quality of the fit of Eq. (1) to the psychophysical data is surprisingly good. Equation (1) was derived from the power spectrum, and thus from the Fourier transform, of a train of identical pulses similar to neural action potentials determined by a combination of random and weakly coherent processes [26,27]. Thus this fit of the psychophysical data to a theory essentially based on the power spectrum of the pulse train suggests that the brain may make use of a similar computation when processing noisy images.

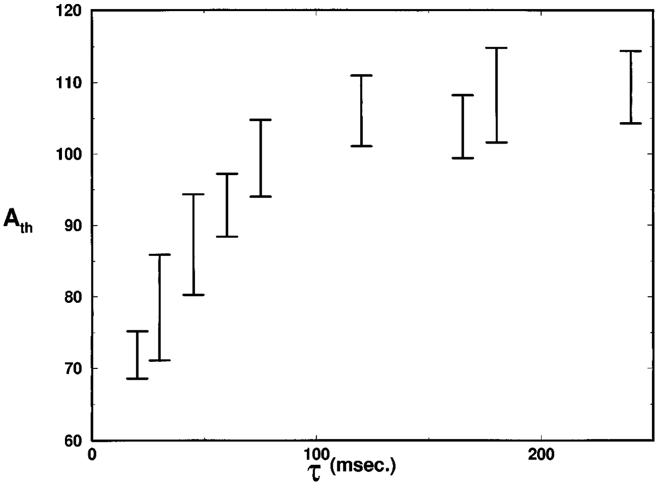

FIG. 4. Perceptive contrast threshold A_{th} versus noise correlation time τ for subject number 4. The noise in each pixel is now correlated in time with auto correlation function $\langle \xi(t)\xi(s)\rangle = (N^2/\tau)\exp[-|t-s|/\tau]$, where $N^2 = \tau\langle\xi^2\rangle \equiv \tau\sigma^2$ is the zero frequency amplitude of the noise power spectrum. The amplitude of the noise, and therefore the total noise power, is held constant at $\sigma = 320$ as the noise correlation time is changed. In this way, the noise correlation time determines only the mean *rate* at which random threshold crossings occur in each pixel and thus the mean sampling frequency of the underlying subthreshold image.

As a bridge between the behavior with temporally varying and static noise, we have performed perceptive threshold contrast experiments using the same protocol on the same set of subjects. In these experiments, however, the noise intensity was held constant while the noise correlation time was varied. The results for an example subject are shown in Fig. 4. The perceptive threshold is lowest for the shortest correlation time and approaches the static result (not shown) for long correlation times. The rise in the curve in the range of 40 to 60 ms is in good agreement with other measures of characteristic averaging times in the visual system [29].

These experiments have demonstrated the utility of SR as quantitative measure of the efficiency with which the visual system processes noisy information. The repeatability and stability of the measure for individual subjects suggests that it may become useful as a diagnostic tool in tracking or detecting visual impairments in humans or in selecting individuals with exceptional ability (small K) to perceive and interpret fine detail within noise contaminated images.

We are grateful to Martin Stemmler for stimulating discussions. This work is supported by the U.S. Office of Naval Research physics program. E. S. thanks the Consiglio Nazionale delle Ricerche, Italy, for financial support.

---

*Present address: 21 Circle Road, Scarsdale, NY 10583.

[1] G. L. Gerstein and B. Mandelbrot, Biophys. J. **4**, 41 (1964).
[2] G. P. Moore, D. H. Perkel, and J. P. Segundo, Annu. Rev. Physiol. **28**, 498 (1966).
[3] W. R. Adey, Int. J. Neurosci. **3**, 271 (1972).
[4] B. W. Knight, J. Gen. Physiol. **59**, 734 (1972).
[5] W. Denk and W. W. Webb, Hear. Res. **60**, 89 (1992).
[6] A. Longtin, A. Bulsara, and F. Moss, Phys. Rev. Lett. **67**, 656 (1991).
[7] (a) I. B. Stiebler and P. M. Narins, Hear. Res. **46**, 63 (1990); (b) P. M. Narins, J. H. Benedix, and F. Moss, J. Auditory Neurosci. (to be published).
[8] B. J. Gluckman, T. I. Netoff, E. J. Neel, W. L. Ditto, M. L. Spano, and S. J. Schiff, Phys. Rev. Lett. **77**, 4098 (1996).
[9] J. J. Collins, C. C. Chow, and T. T. Imhoff, Nature (London) **376**, 236 (1995).
[10] E. Pantazelou, F. Moss, and D. Chialvo, in *Noise in Physical Systems and 1/f Fluctuations*, edited by P. H. Handel and A. L. Chung (American Institute of Physics Press, New York, 1993), pp. 549–552. This work demonstrated SR in a system of independent elements connected in parallel and convergent on a summing center. Each element is subject to the same subthreshold signal. There may be a single external noise source [8,17] or each element may be subject to an internal noise incoherent with all other elements [9]. In the latter case, the signal power, after summation, is enhanced by a factor n, whereas the noise is enhanced only by $\sqrt{n}$, where n is

the number of elements in the array. Such an array may closely approximate a number of biological modalities, including voltage sensitive ion channels in cell membranes [17], and, in humans, sensory neurons in the median nerve [14] and muscle spindles [15]. See also F. Moss and X. Pei, Nature (London) **376**, 211 (1995).

[11] (a) J.K. Douglass, L. Wilkens, E. Pantazelou, and F. Moss, Nature (London) **365**, 337 (1993); (b) X. Pei, J. Wilkens, and F. Moss, J. Neurophysiol. **76**, 3002 (1996).

[12] (a) J.E. Levin and J.P. Miller, Nature (London) **380**, 165 (1996); (b) C. Heneghan *et al.*, Phys. Rev. E **54**, R2228 (1996).

[13] (a) J.J. Collins, T.T. Imhoff, and P. Grigg, J. Neurophysiol. **76**, 642 (1996); (b) X. Pei, L. Wilkens, and F. Moss, Phys. Rev. Lett. **77**, 4679 (1996).

[14] F. Chiou-Tan *et al.*, Intern. J. Bifurc. Chaos **7**, 1389 (1996).

[15] P. Cordo *et al.*, Nature (London) **383**, 769 (1996).

[16] R.P. Morse and E.F. Evans, Nature Medicine **2**, 928 (1996). This work shows that SR is quite close to a practical medical application: the enhancement by noise of encoded vowel sounds in cochlear implants for the profoundly deaf. See also F. Moss, F. Chiou-Tan, and R. Klinke, Nature Medicine **2**, 860 (1996).

[17] S.M. Bezrukov and I. Vodyanoy, Nature (London) **378**, 362 (1995).

[18] K. Wiesenfeld and F. Moss, Nature (London) **373**, 33 (1995).

[19] F. Moss and K. Wiesenfeld, Sci. Am. **273**, No. 2, 66 (1995).

[20] A. Bulsara and L. Gammaitoni, Phys. Today **49**, No. 3, 39 (1996).

[21] J. Collins, T.T. Imhoff, and P. Grigg, Nature (London) **383**, 770 (1996).

[22] M. Stemmler, M. Usher, and E. Niebur, Science **269**, 1877 (1995).

[23] D. Chialvo and V. Apkarian, J. Stat. Phys. **70**, 375 (1993).

[24] M. Riani and E. Simonotto, Phys. Rev. Lett. **72**, 3120 (1994).

[25] M. Riani and E. Simonotto, Nuovo Cimento **17D**, 903 (1995).

[26] G. Gingl, L. Kiss, and F. Moss, Europhys. Lett. **29**, 191 (1995). This simple theory begins with a classical formula for the positive threshold crossing rate (Fig. 1) from S.O. Rice in *Selected Papers on Noise and Stochastic Processes*, edited by N. Wax (Dover, New York, 1954), pp. 133–294. The calculation neglects correlations between crossing events and thus assumes a signal amplitude A, much smaller than the noise. An additional approximation (adiabatic) assumes that the frequency of the signal is much smaller than the mean threshold crossing rate. The power spectrum of the spike train (of identical pulses) is then a delta function at the signal frequency riding on a white noise background. The signal-to-noise power ratio is the ratio of the strength of the delta function S^2 (its power) to N^2, the noise power in a 1 Hz bandwidth, $(S/N)^2 = (2f_0\Delta^2A^2/\sigma^4\sqrt{3})\exp[-\Delta^2/\sigma^2]$, where f_0 is the noise bandwidth (rectangular power spectrum). See also F. Moss *et al.*, Intern. J. Bifurc. Chaos **4**, 1383 (1994). A qualitatively similar expression has been obtained by P. Jung, Phys. Rev. E **50**, 2513 (1994) where the correlations are also considered. However, subjects detect A rather than S/N. Inverting the above equation and using $\sigma^2 = N^2f_0$ results in Eq. (1) where the constants are lumped into $K = (\sqrt{3}/2)^{1/2}(S/\Delta)$, a constant proportional to the signal *power*, S.

[27] G. Gingl, L. Kiss, and F. Moss, Nuovo Cimento **17D**, 795 (1995).

[28] An animation demonstrating the effects of the noise intensity for both static and temporal realizations can be viewed on the World Wide Web, at URL site: http://www.ge.infm.it/~simon/sr.html.

[29] See, for example, R. Bowen, J. Pola, and L. Matin, Vision Res. **14**, 295 (1974). Flicker-fusion experiments find characteristic averaging times in the range 40 to 80 ms in the human visual system.