The 2009 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 11-15, 2009 St. Louis, USA

Development of Emotional Tremor-based Vision System

Shogo Yonekura Yasuo Kuniyoshi Yoichiro Kawaguchi
The University of Tokyo,
7-3-1, Hongo, Bunkyo-ku, Tokyo, JAPAN, 113-0033
E-mail: yonekura@iii.u-tokyo.ac.jp

Abstract— **We aim to investigate biological function of emotional bodily movements while building a robot in the sense of survival. In this particular paper, we build a novel active vision system which detects weak signals by generating tremor actively. Proposed active vision system consists of a noise generator and a neural system. As a noise generator, we use motors with a decentered weight. And as a neural system in order to detect visual signals, we use ensemble of FitzHugh-Nagumo neurons. We first show that our robotic platform can generate several kinds of tremors; tremor around 40 Hz, 50Hz, and tremor which does not involve specific peak in its power spectrum. And, in order to evaluate the effect of tremor, we prepare clear-films as visual stimuli which move aperiodically in a dark environment. We show that as a result of sensory noise induced by tremor, and as a result of stochastic resonance, neural system successfully detects movements of clear films.**

I. Introduction

Recent social robots are designed to express their internal states and desires with their face [11] or by bodily gesture [12]. It may be a consensus that emotional expressions of a robot is of great service to man-robot interaction. Meanwhile, compared to social functionality of emotion, little is known about functionality and mechanisms of emotion in the sense of survivability. Therefore, it seems to be impossible so far to build emotional robots in the sense of survival.

Our research goal is to synthesize emotional robot in the sense of survival. For this purpose, we start paying attention to emotional bodily movements such as trembling, piloerection and eye-tremor during strong arousal, by placing our hypothesis on that these bodily movements provide survival functions. And we try to build a robot which can generate emotional bodily movements, while on the other hand trying to address functionalities of such emotional bodily movements.

In this particular paper, we focus on emotional tremor/ trembling during fear, anger and joy. And we aim to show that emotional tremor can be functional in the framework of sensory-motor coordination (i.e., bodily-movements in order to realize perception or recognition [1]). First, we build a simple robotic platform which can generate tremor, and we analyze the effects in terms of robot's perception. We show that bodily tremor is a component to realize a novel kind of sensory-motor coordinations; detection of weak sensory stimulus by the exploitation of tremor-induced sensory noise.

Hereafter, we briefly describe possible scenario concerning how tremor can provide functions.

II. Tremor induced stochastic resonance

This section provides prospective scenario concerning functions of tremor; tremor as active noise generator in order to induce stochastic resonance (SR) of a neural system.

A. Tremor and sensory noise

Tremor is addressed as small and rapid muscle and body vibration. In strong muscle contraction, tremor's frequency reaches more than 40 Hz, though in muscle relaxation the frequency is around 5 Hz [14][15]. Particularly, eye-tremor during fixational eye movements, is around 200 Hz [13].

It is clear that several sensory streams are effected by tremor. Not to mention the visual stream, but even tactile and auditory stream can be disturbed and modified by bodily tremor. Point would be that if tremor frequency is rapid enough, then, tremor-induced sensory signal may be assumed as sensory noise.

B. Stochastic resonance

Traditionally, SR is addressed as a phenomenon in which noise enables a nonlinear system to detect subthreshold signals [19][5][3]. Examples of SR are fish exploiting noisy turbulent water in order to detect faint signals of predators [19]. Cricket likewise exploits noise to detect low-frequency air-signals from predators [22]. SR is also found to be used in the control of muscle-lengths [24]. And some works have showed the existence of SR in visual perception [7][8][9].

As for noise property necessary necessary for SR, it is known that not only white noise but also temporally or spatially colored noise can induce SR [21][23][20][6]. Furthermore, SR is observable even when noise magnitude is too strong [4], and when signals are supra-threshold [10][25].

C. Hypothesis: tremor-induced stochastic resonance

To authors' knowledge, meanwhile, power spectra of tremor during emotion is not investigated unfortunately. However, taking into account that our skeletal muscle is strongly contracted in response to stress, it may be reasonable to assume that during fear and strong arousal state tremor is quick (e.g., more than 40 Hz) rather than slow.

Under the assumption that frequency of emotional tremor is high enough, it may be reasonable to split sensory streams into "original sensory data" and "additive noise", even though validation of this assumption is left as our future work. And, as far as we can separate original sensory data and additive noise, SR would be observed at the level of neural behaviors.

III. Mono streamed Tremor vision system

This paper describes two implementations of tremor vision to investigate effects of tremor onto visual perception. Similarity of our implementations and brain structure is discussed in section VIII. In this section, we describe the former mono-streamed tremor vision system.

A. Model and implementation

Temporal differentiation of pixel $p_i(t)$ of a camera represents motion-information of camera-subjects, $s(t)$, as far as the camera is quiescent. Contrarily, $p_i(t)$ includes motion-information of the camera itself $m_c(t)$, if a camera moves. Here, if the camera motion is rapid enough, then, $m_c(t)$ can be assumed as "noise" $n(t)$ of motion information $s(t)$. That is, $\Delta p_i(t) = s(t) + n(t)$. As additive noise enhances perceptive ability of a nonlinear system by stochastic resonance, nonlinear system will be able to detect and represent $s(t)$ even if it is too weak and subthreshold.

The architecture depicted by Fig. 1 consists of mainly three parts; visual filter, neural system, and tremor generator.

B. Camera specifications

As a camera device, we use PGR IEEE1394 camera, *Firefly MV*, which can capture images of pixel 160×120 at more than 200FPS. We configured "Brightness", "Exposure", "Gamma", "Shutter speed" and "Gain" to fixed values, so that we can suppress noise derived from camera's automatic parameter-reconfiguration.

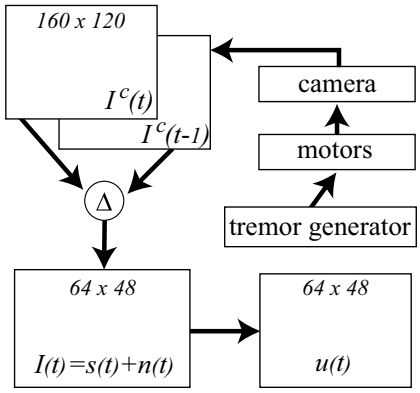

Fig. 1. Architecture of mono-streamed tremor vision system: $I^c(t)$ means gray-scaled image sampled by a camera. Motion-information I^t is calculated from camera image $I^c(t)$, by temporal subtraction. Nonlinear neural system receives down-sampled $I(t)$ as input. If camera tremor is rapid enough, $I(t)$ can be separated into the form $I(t) = s(t) + n(t)$, and then nonlinear neural system $u(t)$ may detect $s(t)$ by exploiting SR, even $s(t)$ is weak and subthreshold. In this article, $s(t)$ is aperiodic consistently (i.e., visual stimulus is 'randomly' generated).

C. Design of visual image filters

Taking the necessity of real-time process into account, we selected simple temporal subtraction of camera images, as image filter to get motion-information. Given that $I^c(t)$ is gray-scaled camera image at time t,

$$I_i(t) = I_i^c(t) - I_i^c(t-1). \quad (1)$$

The size of $I^c(t)$ and $I(t)$ is 160×120.

D. Ensemble of nonlinear neurons

As a neural system, we consider an ensemble of nonlinear neurons of FitzHugh-Nagumo, based on the literature [4][5]. The dynamics of ith neuron is described as;

$$\dot{u_i} = c\left(-\frac{u_i^3}{3} + u_i - v_i + I_i\right), \quad (2)$$

$$\dot{v_i} = u_i - bv_i + a, \quad (3)$$

where $a = 0.7$, $b = 0.8$, $c = 10.0$. u_i is fast variable representing spike of a neuron, whereas v_i is slow recovery variable. Stepsize is synchronized to that of camera frames (i.e., approximately 0.0046 sec).

E. Robotic platform

As motors to generate tremor, we use motors with a decentered weight. Furthermore, in order to generate tremor around both of axis x and y, we attached two vibrator motors as Fig.2. Amplitude of tremor is adjusted by switching off the current to the vibrator motors at the probability $(1-\gamma)$, and by controlling PWM Duty ratio D. Control cycles are 16MHz for γ and 0.16MHz for D, respectively.

Fig. 2. Schematic platform. v denotes a vibration motor. c denotes a camera. Four vibration motors are attached. Dashed line in the left figure denotes wire in order to fix the position of camera and to adjust spring and damping effect of the system.

Meanwhile, due to a lot of mechanical friction, gear-mechanism is not suitable to transfer tremor/trembling. Taking this point into account, we employ wire-driven active vision system. One of advantages of wire-driven system, for the purpose of tremor transfer, will be that we can control spring / damping effect by controlling tension of the wires. Schematic model and photographs of the robotic platform is shown in Fig.2.

F. Noise properties

Noise variance σ and power spectrum depends on the probability γ and PWM duty ratio D to some extent (in a more precise sense, σ is a function of environmental illumination and texture, γ and D). Fig.3 shows distribution of noise variance with respect to γ and D, and two typical power spectra of noise in the experimental environment described in section IV. Peak spectrum around 40 - 50 Hz is observed for $\gamma \approx 0.8$ and $D \approx 50$, $\gamma \approx 40$ and $D \approx 100$, and broad spectrum for otherwise. Later, we discuss difference of SR effect in terms of noise power spectrum, based on several analyses.

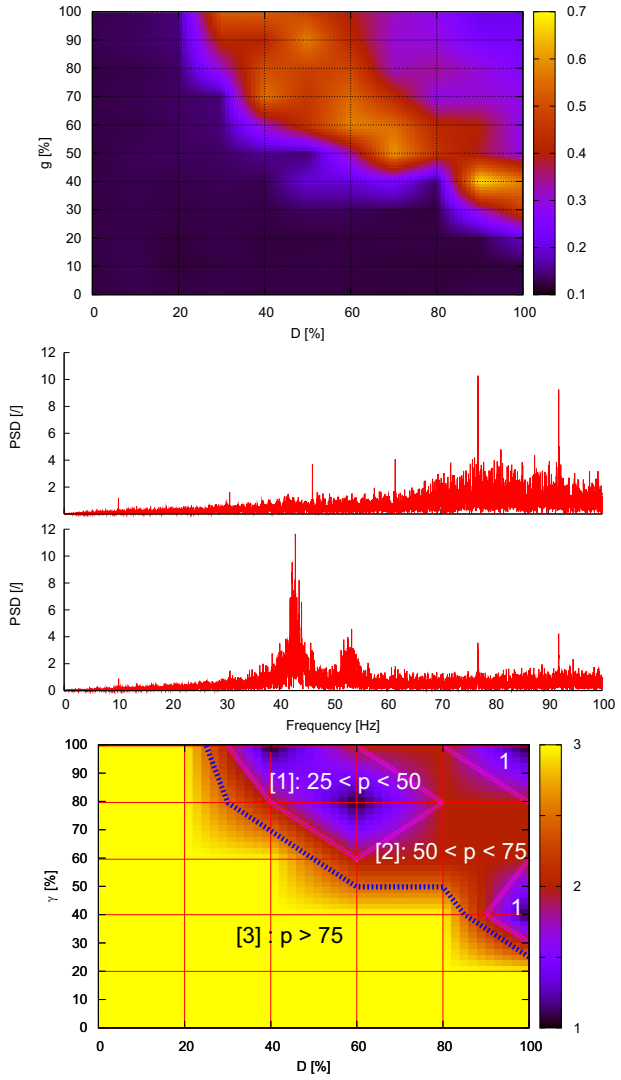

Fig. 3. Top: Distribution of noise variance with respect to motor-current switching probability γ and motor-current PWM duty D. Horizontal axis is D [%] and vertical axis is γ [%]. Middle: typical power spectrum (i.e., broad spectrum and peak around 40 Hz) of noise. Bottom: phase diagram of noise power spectrum(p is peak frequency [Hz]).

IV. Experimental Setup

A. Basic design

The primary purpose of the experiments is to clarify that tremor can help detection of subthreshold weak signals. For this purpose, we examine system performance with several system parameters such as tremor motor current probability γ, PWM duty D and neural parameter g.

In order to evaluate system performance (system output with respect to visual stimulus), we selected cross-correlation measure.

B. Schematic model of the experimental environment

The model of the experimental environment is shown in Fig.4. To give the camera aperiodic visual stimuli, which should be hard too "see", we prepared a "windmill" with clear-thin films. And behind the windmill, we put a picture

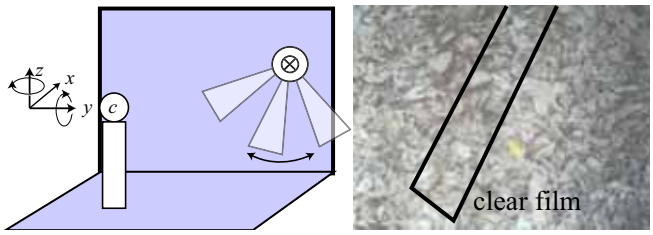

Fig. 4. Schematic model of the environment configuration. A "windmill" consisting of clear films is attached to a desk-wall, and a camera is fixed on a desk. Windmill rotation gives visual stimuli to the camera. Vibration-motors rotate around axes x and z.

of fallen leaves. One reason of this background texture is that it has a lot of frequency components. The images captured by the camera (640 $\times$ 480 pixel) is shown in the left side of Fig.4. Visual stimuli are provided by moving the windmill randomly (i.e., aperiodically). It should be noted that in order to avoid strong flicker noise around 100 Hz of flour lamps, we alternatively prepare darkroom. The experimental environment is illuminated with DC-LEDs.

C. Method to estimate system performance

We measure effects of stochastic resonance by the following normalized cross-correlation power norm C;

$$C = \frac{\overline{[s_e(t) - \overline{s_e}][r(t) - \overline{r}]}}{[\overline{\{s_e(t) - \overline{s_e}\}^2}]^{1/2}[\overline{\{r(t) - \overline{r}\}^2}]^{1/2}}, \tag{4}$$

with the overbar denoting an averaging over time. Here, $r(t)$ is the mean to represent firing rate constructed as follows;

$$r(t) = \sum_i y_i(t), \tag{5}$$

$$y_i(t) = \begin{cases} 1 & (u_i(t) > 0) \\ 0 & (u_i(t) \leq 0) \end{cases}. \tag{6}$$

And $s(t)$ is time-series of visual stimuli being input to neural system. Since it is impossible to have exact dynamics of $s(t)$, we estimate and re-construct $s(t)$ based on the input to motor devices (let's say $s_e(t)$). It should be noted, therefore, that in cases of motor backlash and unexpected movements of the films, $s_e(t)$ does not represent $s(t)$ exactly.

V. Results: Mono Streamed Tremor Vision

A. Successful detection by Tremor

We add tremor by the configuration $\gamma = 0.5$ and $D = 50$. Fig. 5 shows that neurons successfully "fire" and detect subthreshold weak signals, due to the presence of noise generated by bodily-tremor. The coherence measure C was 0.153. Meanwhile, firings without visual stimuli can be observed (around 37 sec for example). This implicates existence of too strong spatio-temporal correlation of tremor because of mechanical resonance or some other reasons.

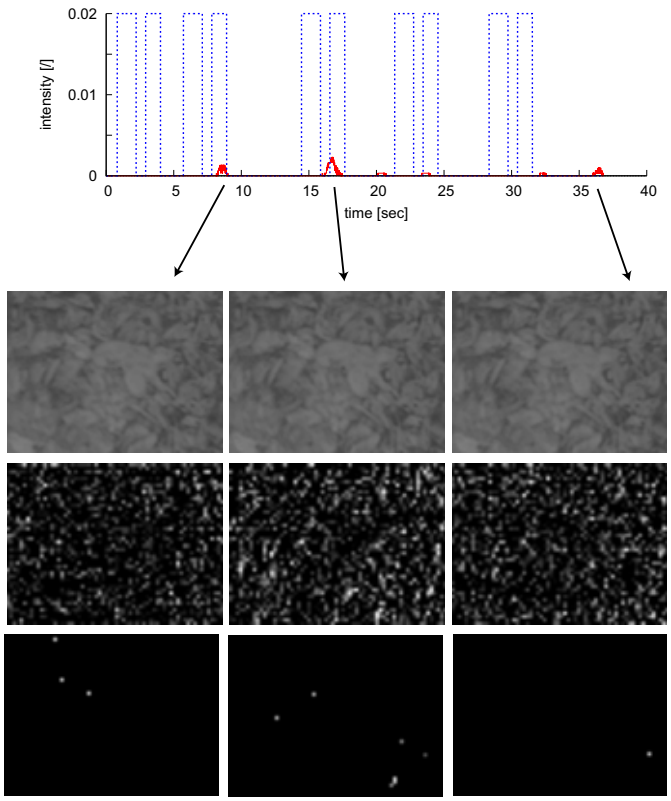

Fig. 5. From top to bottom, time series of $s_e(t)$ and $r(t)$, camera image, $I(t)$, and response of the neural system $y(t)$. The presence of tremor-noise leads to the detection of stimuli information. The right side of $y(t)$ shows miss-firing due to the slow and long-term correlation of noise (see Fig.3). The left and middle sides of $y(t)$ show successful detection. of the shape of the clear film.

B. Correlation Analysis

We took statistics of noise σ and C, while separating cases of $C \geq 0$ and $C < 0$ (this is because $C < 0$ is derived by long-term correlated noise, which is undesirable to evaluate SR). In cases except $g = 750$, we can observe bell-shaped curve of C. This implicates the existence of traditional SR; under the presence of noise adequate intensity, the sensitivity of nonlinear neurons is maximized. Note that the curve of C with $g = 750$ starts approximately from $\sigma = 0.25$. This is because of camera's internal noise (probably thermal noise). Note also that C becomes higher by the exploitation of bodily tremor with low input gain g such as $g = 300$ or 450 than $g = 600$ or $g = 750$.

VI. Dual streamed tremor vision system

A. Issue with respect to the noise controlling

Previous experiments implicate that in order to exploit the effect of SR it is necessary to adjust noise intensity within an adequate range. Meanwhile, in our platform, noise variance σ is a function of environmental illumination and texture (let' say ξ), input stimulus s as well as gain g, motor-current probability γ and duty D. Therefore, in order to control σ, it is required to solve; $(\sigma) \mapsto (\xi, s, g, \gamma, D)$. This problem,

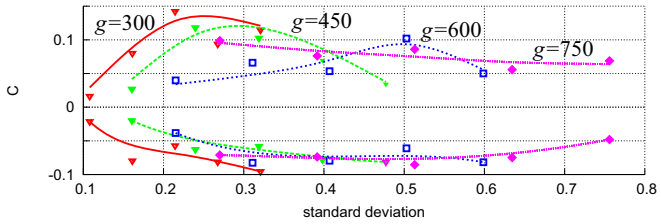

Fig. 6. Correlation versus noise deviation σ, under $g = 300, g = 450, g = 600$ and $g = 750$. For most of the parameter sets, we can observe bell-shaped curve of C, implicating the existence of traditional SR.

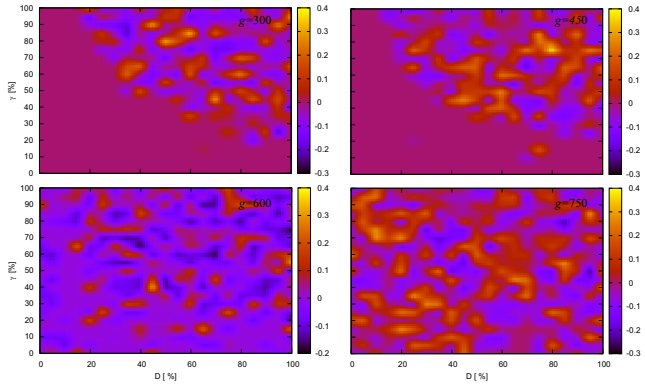

Fig. 7. Correlation C versus γ and D. Under weak g such as 300 and 450, C is enhanced in slow noise phases ([1] and [2] in Fig.3), whereas C is enhanced in broad noise phase ([3]) under higher g.

however, is inverse problem and may be impossible to solve in most of systems.

Alternatively, we choose not to control noise intensity but to minimize decrease of SR effect with respect to the noise strength. Fortunately, the literature of SR provides us a theory with respect to this issue; when multiple neurons receives a common input signal, and independent noise, then, averaged neural activity represents well the input signal regardless of noise intensity. In order to exploit this theory, we build a model in which neurons receives dual streamed input, that is, coarse and fine sensory input.

B. Dual stream model of tremor vision

The concept of the model is to split sensory signal into fine scale input I^0 and down-sampled coarse scale input I^1. Total input to neurons $I(t)$ is described as;

$$I(t) = I^0 \oplus I^1, \tag{7}$$

where $\oplus$ is operator which consists of point-wise addition, and interpolation and resize of I^0 and I^1 to the size of the neural network. In our implementation, I^0 and I^1 is 120×100 and 16×12 respectively. Approximately 60 neurons thus receive common input derived from the down-sampled sensory signal and independent noise signals.

VII. Results : Dual Streamed Tremor Vision

A. Success of detection by tremor

We add tremor by the configuration $\gamma = 0.86$ and $D = 80$. Fig. 9 shows that neurons successfully "fire" and detect

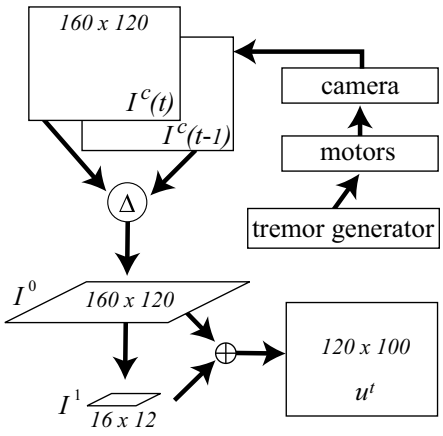

Fig. 8. Architecture of dual streamed model. ΔI^c is decomposed to fine scaled I^0 (160×120) and coarse scaled I^1 (16×12). $I = I^0 \oplus I^1$is input to the neural system.

weak signals, due to the presence of noise generated by bodily-tremor. The coherence measure C was 0.531. A big difference with the result shown in Fig.5 is that noise of fine scale is convoluted with noise of rough scale.

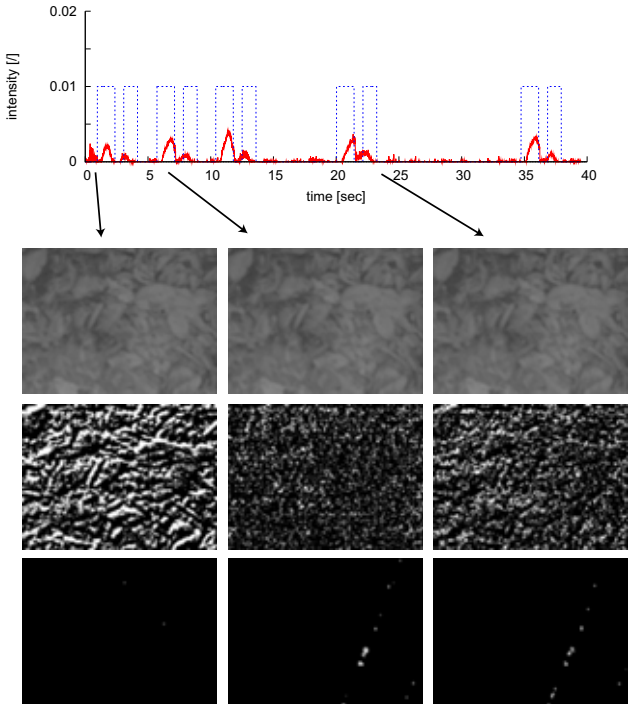

Fig. 9. Top: time series of $s(t)$ and $r(t)$. Bottom: $I^c(t)$(upper), $I(t)$(middle) and $u(t)$ (lower).

B. Correlation analysis

Correlation analyses implicate benefit of dual-streamed model. It is clearly shown that decrease of C by large σ is suppressed due to the averaging effect of the model. However, it should be noted that this model, too, is not able

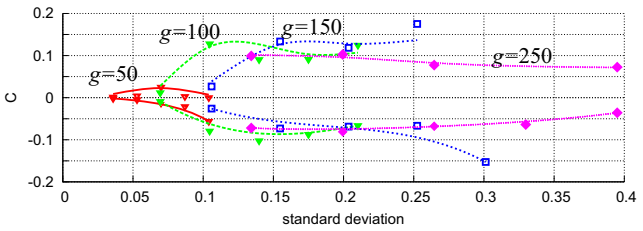

Fig. 10. Correlation C versus noise deviation σ, under $g = 50, 100, 150$ and 250. It should be noted that bell-shaped curve of C disappears and thus we can observe suppression of C decrease for large σ.

to suppress miss-firing due to the strong correlated tremor (see Fig.9 $C < 0$).

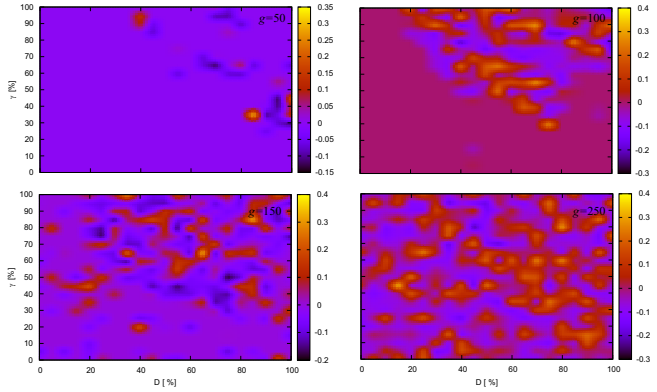

Fig. 11. Correlation C versus γ and D. Under weak g such as 50 and 100, C is enhanced in slow noise phases ([1] and [2] in Fig.3), whereas C is enhanced in broad noise phase ([3]) under higher g.

VIII. Discussion and Conclusion

A. Suppression of long-termed strong noise correlation

The primary reason of negative value of C (i.e.,$C < 0$) may be mechanical resonance, which leads to long-termed strong noise correlation being harmful to SR. Suppression of such harmful resonance may drastically improve the system performance. One direction for such improvements may be re-designing of robotic platform, in which more strong spring and weak dampers are implemented. And the other direction may be implementation of feed-back controller to suppress strong noise-correlation.

B. Roles of internal and embodied noise

In our work, two noise sources are involved; camera's internal (and thermal) noise and tremor-oriented embodied noise (Phases [1] and [2] in Fig.3 are mainly derived from bodily tremor and [3] in Fig.3 may be internal thermal noise). Experimental results shown in Fig.7 and Fig.11 implicate that for low gain g, embodied noise alone can help signal detection, and for high gain, internal noise, too, can help signal detection. Interesting result is that we can get better C by the exploitation of tremor with low g than by simple increase of g (see Fig.10). Taking into account that SR is more effectively induced by correlated noise [21][23], it is

not plausible hypothesis that animal exploits bodily-tremor in order to detect weak-stimulus.

C. Necessity of probabilistic information processing

Though SR enables a neural system to detect weak and subthreshold signals, some new framework may be required in order to generate rational agent's behavior based on the result of information processing being associated with SR. This is because, SR-based neural output is noisy and may involve a lot of miss-firings. Additionally, to make things worse, a subjective agent would not be able to know whether the result of such information-processing is correct or not, as far as stimulus is not clear. So far, we are expecting that probabilistic framework such as Bayesian interface [2] is helpful to address these issues concerning indeterminacy.

D. Links to biological emotion

We hope our model links to biological emotion in the following two senses;

1. Similarity of phenomenon
2. Similarity of structure

The former is phenomenological similarity in that our platform 'trembles' as animals trembles during strong arousal states such as fear, anger and joy. This similarity, however, should be precisely investigated further because power spectrum of tremor frequency of biological emotion is not investigated, and thus, tremor by biological emotions would not induce SR. The latter is structural similarity of our dual-streamed model and brain structure. That is, I^0 can be assumed as LGN (lateral geniculate nucleus), I^1 can be assumed as limbic system (amygdala for example), and subjective neural system u can be assumed as visual cortex.

E. Emotion in perspective of sensori-motor coordination

Bodily trembling due to fear, anger and joy has long been assumed as useless in itself though trembling provides *'somatic marker'* when being perceived by cerebral cortex [16][18][17]. Contrarily, what we show in this paper is that bodily trembling directly modulates and deforms sensory streams and thus can be a direct bias onto perception and cognition; trembling can be a bodily movements in order to realize sensori-motor coordination [1]. We hope a lot of other emotional bodily movements which seems to be useless such as horripilation, piloerection, wriggling, screaming, laughter and jumping, can be functional in the scope of sensori-motor coordination.

F. Conclusion

In this paper, we reported our work toward the synthesis of emotional robot within the perspectives of survivability. In particular, we reported a novel tremor-based active vision system which exploits stochastic resonance by actively generating tremor-based embodied noise.

Our work is addressed as the first challenge toward the understanding of emotion within the perspectives of sensori-motor coordination. So far, we succeeded to provide a sound finding that tremor leads to stochastic resonance of nonlinear neural system. In addition we succeeded to build a novel robotic platform which is able to detect weak signals by trembling. We hope investigation of functions of emotional bodily movements based on robotic paradigm would be fruitful.

ACKNOWLEDGMENT

This work is supported by JST CREST.

REFERENCES

[1] R.Pfeifer and C.Sceier, *Understanding Intelligence*, MIT Press, Cambridge, MA, 1999
[2] W.J.Ma, J.M.Beck, P.E.Latham and A.Pouget, Bayesian interface with probabilistic population codes, *Nature Neuroscience*, 9(11):1432-1438, 2006
[3] M.D.McDonnell and D.Abbott, What is stochastic resonance?, *PLos Computat Biol*, 5(5):e1000348, 2009
[4] J.J.Collins,C.C.Chow and T.T.Imhoff, Stochastic resonance without tuning, *Nature*, 376(20):236-238, 1995
[5] J.J.Collins, C.C.Chow and T.T.Imhoff, Aperiodic stochastic resonance in excitable systems, *Phys. Rev. E*, 52(4):1063, 1995
[6] H.Yasuda, T.Miyaoka, J.Horiguchi,A.Yasuda,P.Hänggi and Y.Yamamoto, Novel class of neural stochastic resonance and error-free information transfer, *Physical Review Letters*, 118103, 2008
[7] K.Ghosh, S.Sarkar and K.Bhaumik, A possible mechanism of stochastic resonance in the light of an extra-classical receptive field model of retinal ganglion cells, *Biological Cybernetics*, 100:351-259,2009
[8] K.Kitajo,D.Nozaki,L.M.Ward and Y.Yamamoto, Behavioral stochastic resonance within the human brain, *Physical Review Letters*, 218103, 2003
[9] E.Simonotto, M.Riani, C.Seife, M.Roberts, J.Twitty and F.Moss, Visual perception of stochastic resonance, *Physical Review Letters*, 78(6):1186-1189, 1997
[10] M.D.McDonnell, Information capacity of stochastic pooling networks is achieved by discrete inputs, *Physical Review E*, 79, 041107, 2009
[11] C.Brezeal and B.Scassellati, Infant-like social interactions between a robot and a human caretaker, Adaptive Behavior, 8(1): 49-74, 2000
[12] N.Endo, S.Momoki, M.Zecca,M.Saito,Y.Mizoguchi,K.Itoh, and A.Takanishi, Development of whole-body emotion expression humanoid robot, ICRA, 2140-2145, 2008
[13] S.M.Conde, S.L.Macknik, and D.H.Hubel, The role of fixational eye movements in visual perception, *Nature Neuroscience*, 5:229-240, 2004
[14] J.H.McAuley, J.C.Rothwell, and C.D.Marsden, Frequency peaks of tremor, muscle vibration and electromyographic activity at 10 Hz, 20 Hz and 40 Hz during human finger muscle contraction may reflect rhythmicities of central neural firing, *Exp Brain Res*, 114:525-541, 1997
[15] D.Günther, J.Raethjen, M.Lindemann, and P.Krack, The pathophysiology of tremor, *Muscle & Nerve*, 24:716-735, 2001
[16] C.Darwin, *The Expression of the Emotions in Man and Animals, 3rd ed.*, Oxford, 1996
[17] W.James, What is an emotion ?, *Mind*, 9(34): 188-205, 1884
[18] A.R.Damasio, *Descartes' Error: Emotion, Reason, and the Human Brain*, Grosset/Putnam, 1994
[19] K.Wiesenfeld, and F.Moss, Stochastic resonance and the benefits of noise: from ice ages to crayfish and SQUIDs, *Nature*, 373(5):33-36, 1995
[20] K.Hamaguchi, M.Okada, S.Kubota, K.Aihara, Stochastic resonance of localized activity driven by common noise, *Biological Cybernetics* 92:438-444, 2005
[21] D.Nozaki, J.J.Collins, and Y.Yamamoto, Mechanism of stochastic resonance enhancement in neuronal models driven by 1/f noise, *Physical Review E*, 60(4):4637-4644, 1999
[22] S.Mitaim, and B.Kosko, Adaptive stochastic resonance, *Proceedings of the IEEE*, 86(11):2152-2183, 1998
[23] S.Wang, F.Liu, and W.Wang, Impact of spatially correlated noise on neural firing, *Physical Review E*, 69, 011909, 2004
[24] J.B.Fallon, R.W.Carr and D.L.Morgan, Stochastic resonance in muscle receptors, *J.Neurophysiol.* 91:2429-2436, 2004
[25] H.Sasaki, S.Sakane, T.Ishida, M.Todorokihara, T.Kitamura and R.Aoki, Suprathreshold stochastic resonance in visual signal detection, *Behavioural Brain Research*, 193:152-155, 2008