The Role of Noise in Sub-threshold Detection(しきい値下信号の検出におけるノイズの役割)

著者	Marisciel Tayam Litong
号	207
発行年	2001
URL	http://hdl.handle.net/10097/12902

	マリシエル・タヤム・リトン
氏　名（本籍）	Marisciel Tayam Litong
学位の種類	博士（情報科学）
学位記番号	情　博　第　207　号
学位授与年月日	平成14年1月10日
学位授与の要件	学位規則第4条第1項該当
研究科、専攻	東北大学大学院情報科学研究科（博士課程）システム情報科学専攻
学位論文題目	The Role of Noise in Sub−threshold Detection （しきい値下信号の検出におけるノイズの役割）
論文審査委員	（主査） 東北大学教授　鈴木陽一　東北大学教授　中島康治 東北大学教授　川又政征　東北工業大学教授　沢田康次 東北大学助教授　早川義徳 （理学研究科）

論　文　内　容　要　旨

1 Introduction

This work revolves around the central theme that noise can be beneficial specifically in signal detection. There are two types of detectors considered in this investigation; the first system is a bithreshold device and the second system is the human hearing system. The most significant outcome of this work is in demonstrating how to enhance the usefulness of noise in detection and in showing that this noise-aided detection occurs spontaneously in the human hearing system. The noise in a system is hardly ever beneficial. Recently, there has been a surge of reports on the beneficial aspects of noise. This phenomenon is collectively known as stochastic resonance(SR) [1]. The first physical observation of SR is reported for a Schmitt trigger [2]. In biological systems, reports of SR-like behaviour has been shown for the mechanosensory systems of crickets [3, 4]. Several works have alluded to a positive role of background noise in the human hearing system[5]. In most studies, SR has been formulated using Gaussian white noise. Real noise sources however are not always Gaussian and are correlated. Noise correlation and noise distribution effects on SR should also be studied for a more complete picture of this noise-aided phenomenon. The numerical simulations in this work is motivated with this purpose. The other main motivation is to see whether SR is exhibited in the human hearing system. The previous works that relate SR to the human auditory system are involved at hair cell level, wherein the sound pressure levels involved are nominal. The interest of this work, however, is at ordinary levels and at ordinary situations at which human speak and listen.

2 The Methods for Enhancing Stochastic Resonance in a Bithreshold System

This section describes the detection parameters, sub-threshold signal and different types of noise considered, corresponding to different cases of distribution and coherence. The system under study is a symmetric bithreshold (see Fig. 1) trigger device described by;

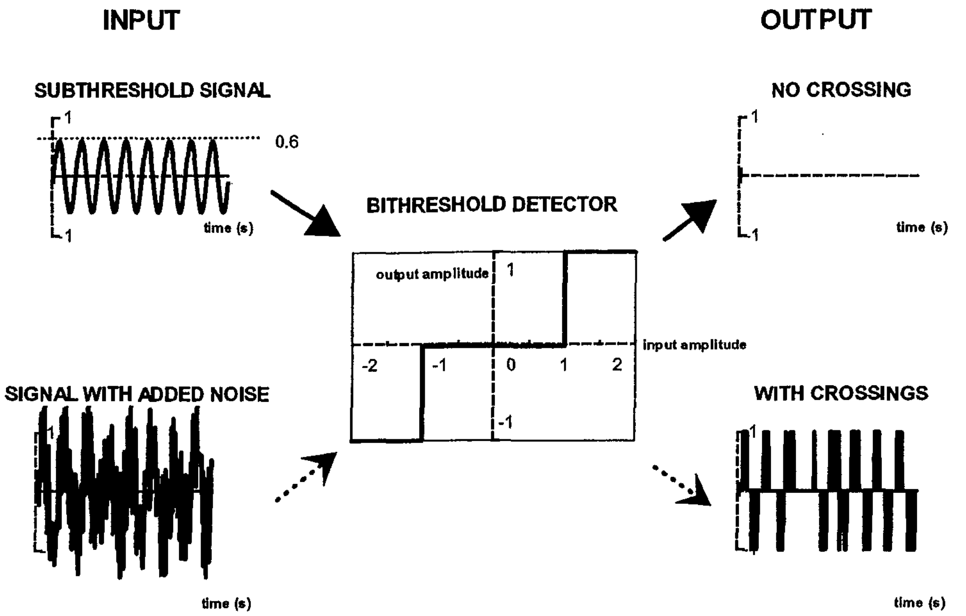

Figure 1: The subthreshold sinusoid does not have crossings with the threshold. Essentially, it is undetectable. Upon adding noise, crossings are induced that makes the sinusoid detectable.

$$y = \begin{cases} +1 & : \quad x \geq B \\ 0 & : \quad -B < x < B \\ -1 & : \quad x \leq -B \end{cases} \tag{1}$$

where B=0.5 arbitrary units (au).

The input $x[n]$ to the detector is composed of the weak periodic modulation plus the noise term as follows;

$$x[n] = A_s \cos(2\pi f_s n(\Delta T_s)) + \xi^n \tag{2}$$

where $f_s = 100Hz$ is the signal frequency and $A_s = 0.3au$ is the signal amplitude. The term ξ^n represents the normalized noise added to the subthreshold sinusoid that is needed to induce threshold-crossings. The normalization is based on equal mean threshold crossing rate for the noise alone.

1. *Gaussian White Noise* (ξ_{gwn}) and *Uniform White Noise* (ξ_{uwn}) The tuning parameter is D, which is defined as the standard deviation for the Gaussian noise and the maximum absolute value for the uniform noise.

2. *Time-correlated noise* (ξ_{tcn}) and *shuffled version* (ξ_{stcn}) The tuning parameter T_c is the correlated time of the series. To remove the correlation in ξ_{tcn} , the entire sequence is shuffled before it is added to the weak periodic signal.

3. *Periodic noise* (ξ_{pn}) and *shuffled version* (ξ_{spn}) Another periodic signal is used with a suprathreshold amplitude $A_{pn} \geq |B|$ is the amplitude and $f_{pn} \neq f_s$ is the frequency, which serves as the tuning parameter.

2.1 Measure of performance

Performance is based on how distinguishable the true frequency stands against the other components in the power spectrum, which is computed from the output of the detector. This measure Q is defined as;

$$Q = \frac{P_{f_s}}{P_{var}} \tag{3}$$

P_{f_s} is the power of the signal frequency f_s and P_{var} is a measure of the variation of the other frequencies about the mean.

2.2 Results and Discussion

The effect of distribution on the SR performance is shown clearly by comparing the performance of the Gaussian noise ξ_{gwn} and uniform noise ξ_{uwn}. A better Q performance was obtained using ξ_{uwn}. This suggests that changing the noise distribution can optimize the SR performance. A look at the histogram 2 for the noise cases considered illustrates this point. The Uniform noise has values equally distributed along the amplitude axis. Hence, it is well-represented in the optimal range (shaded area), which allows more threshold crossings. The Uniform noise can then be considered the optimal distribution. On the contrary, the Gaussian noise has values concentrated at zero and away from the optimal range. Hence it can induce less crossings. The Gaussian noise is considered suboptimal. An interesting case of noise distribution is the periodic noise case. Most of its values are located at the optimal range. The Periodic noise can be considered to have a hyperoptimal distribution. Since the periodic noise has a hyperoptimal distribution, it is expected to give the best Q performance. Figure 3(c), however, shows that on the contrary,ξ_{pn} has the worst performance. The best Q value is obtained with its corresponding shuffled version ξ_{spn}. For this case, correlation in the noise series is not helping. However, for the exponentially correlated noise ξ_{tcn}, removing the correlation by shuffling the series or making the correlation time approach zero decreased the Q to values that are comparable with that of the white noise sources. For this case, correlation in the noise series is helping. These results show that the role of noise correlation is two-fold. It is helpful only when the noise distribution is suboptimal, and is harmful when the noise distribution is already optimal (or even hyperoptimal).

3 Listening Performance and the Effect of Background Noise in a Cocktail Party Setting

In this section, the ability of a human subject to listen to or track a particular voice of interest is examined. This ability is related to the so called Cocktail Party Effect (CPE), which describes the situation in which humans are able to converse with one another even in the presence of other interfering speeches.

3.1 Design and Procedure

Speech samples for eight different speakers, consisting of four males and four females, are used. Of the four male and four females, one male and one female are designated as the target voice speeches while the other 6 speakers are the non-target speakers. Each speech sample follows a basic sentence structure as

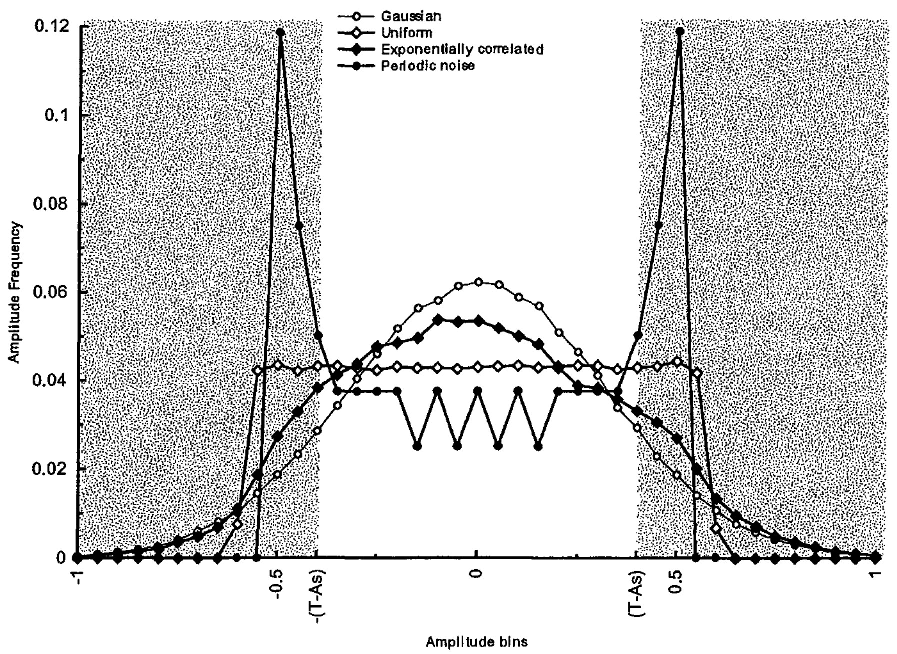

Figure 2: Computed histogram for ξ_{gwn},ξ_{uwn},ξ_{tcn} and ξ_{pn}. The optimal range of noise values are shown by the shaded area.

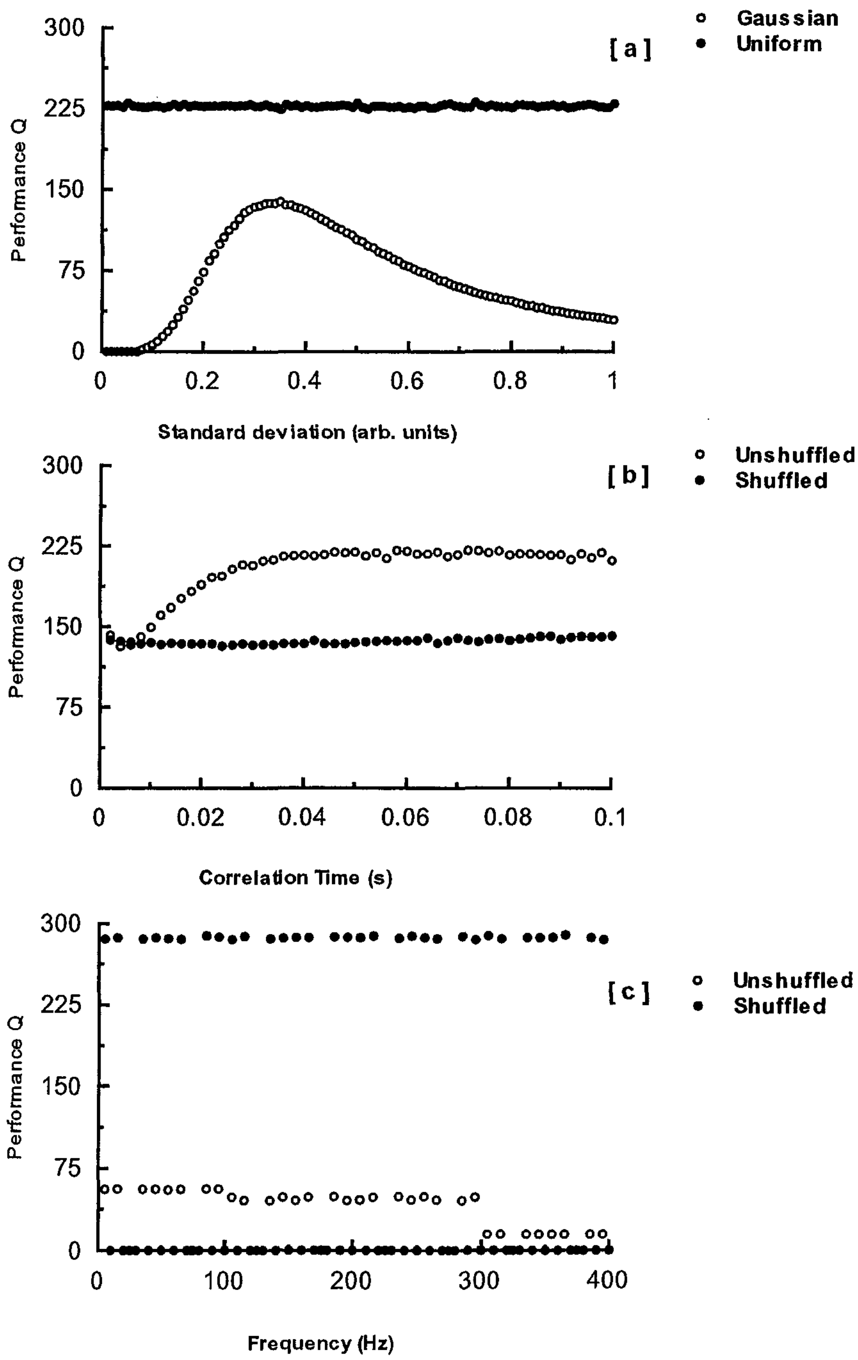

Figure 3: The Q performance when adding (a)ξ_{gwn} and ξ_{uwn}; (b)ξ_{tcn} and ξ_{stcn}; and (c) ξ_{pn} and ξ_{spn}.

Table 1: Table of values for the characteristic number M and pre-factor A.

	M		A	
	SYNC	ASYNC	SYNC	ASYNC
Male	18.7	31.0	0.991	0.967
Female	15.3	27.3	0.989	1.048

NAME go to COLOR, NUMBER now. The procedure consists of three stages: the training of the subject, the pretesting of the subject and the actual experiments. During the training, the subject was presented several speech samples for the target voice, which either male or female at a time. The subject familiarized himself with the target voice. The signals are mixed synchronously (same starting times) or asynchronously (delayed starting times). Before mixing, the speech sources were normalized by the total power of each speech sample.

3.2 Results and Discussion of Experiment

The measure of performance used in the listening experiments (voice tracking) are based on the probability of the correct response $p(C)$ of the human subjects. The subject must decide whether the target voice is present or absent in the presented signal. The performance is shown in Figure 4. Qualitatively, the performance plots shows the following trend: (a) a decrease in performance as the number of mixed components is increased; (b) better performance for the asynchronous mode of mixing; and (c) comparable performance for both the male or female target voice cases. This plot is the averaged performance obtained for 3 subjects with 100 trials per data point.

3.2.1 Trend Analysis

To compare the performance results for the different experimental cases, the data are analyzed for trends. Based on the Pearson product-moment coefficient and the error values, the exponential decay function is the best curve to describe the trends in the given data. The values for the prefactor A and the reciprocal to the decay factor M corresponding to the $p(C)$ plots are summarized in 1. The values indicate that the accuracy, implied in A is not significantly different whether the target voice is male or female or whether the mixing mode is synchronous or asynchronous. The M parameter shows, however, a significant difference in the performance between the two cases of mixing, with the asynchronous mode being characterized by slower decay.

3.2.2 Performance Normalization

In Figure 4, the performance plot for an ideal machine detector is also indicated. This performance plot is based on the assumption that a machine can perfectly detect the target signal given a signal-to-noise S/N ratio equal to 1, the case which corresponds to the number of mixed components $m = 2$. At $m = 2$, one of the components of the mixed signal is the target and the other component is considered the noise. Since the power of each component has been normalized before mixing, and the signal-to-noise ratio is computed in amplitude (or square root of the power) domain, then the signal-to-noise ratio varies inversely as the the square root of m. Under this assumption, the $p(C)$ performance for the human subjects can be normalized with respect to the machine performance to give a comparative measure. This normalization with extrapolated values is illustrated

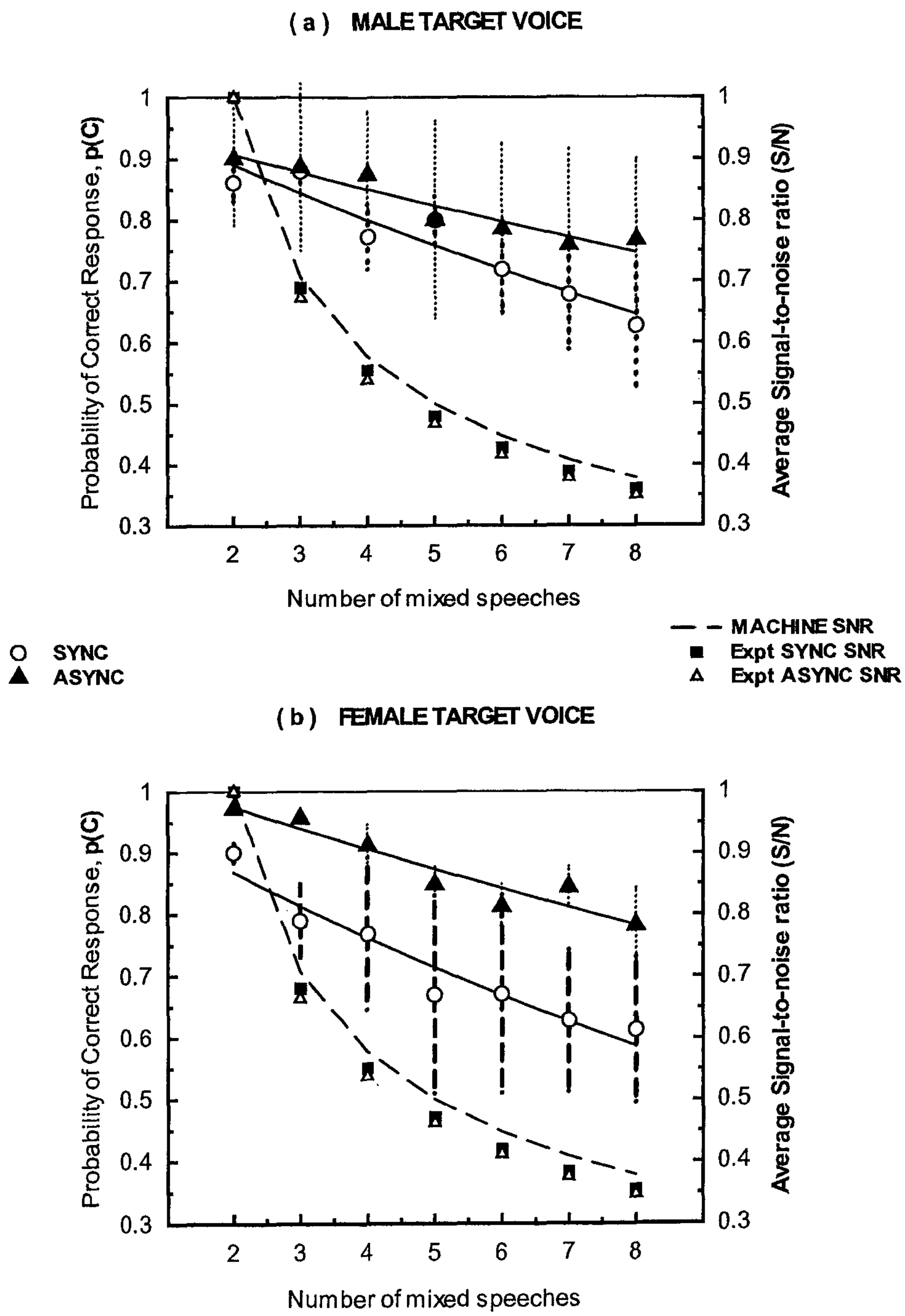

Figure 4: The probability of correct responses, $p(C)$.

in Figure 5. It is interesting to note that this curve is reminiscent of the stochastic resonance signature; i.e. an optimal peak for an intermediate noise level.

4 CONCLUSION

The most significant outcome of this work is in demonstrating that the usefulness of noise as exhibited in a (physical model) bithreshold detector is also effectively exhibited in a biological detector that is the human ears. This is based on the stochastic resonance-like curve obtained when the human listening performance, which is exponentially decaying, is compared to an ideal machine performance that depends only on the signal-to-noise ratio, which degrades as the $\sqrt{m}$. The m is the number of mixed components or the number of simultaneous speakers in the cocktail party setting. For the bithreshold detector, I showed that the classical stochastic resonance, i.e. the stochastic resonance when Gaussian white noise is used, can be enhanced in two ways; one way is to vary the noise correlation and the other way is to optimize the noise distribution.

References

[1] Luca Gammaitoni, Peter Hänggi, Peter Jung, and Fabio Marchesoni. Stochastic resonance. *Rev. Mod.Phys.*, 70(1):223–287, 1998.

[2] S. Fauve and F. Heslot. *Physics Letters*, 97A(5), 1983.

[3] J. E. Levin and J.P. Miller. Broadband neural encoding in the cricket cercal sensory system enhanced by stochastic resonance. *Nature (London)*, 380:165–168, March 1996.

[4] T. Shimozawa, T. Kumagai, and Y. Baba. Mobilities of the cercal wind-receptor hairs of the cricket. *Journal of COmputational Physiology A*, 183:7–21, 1998.

[5] Sebastien Camalet, Thomas Duke, Frank Jülicher, and Jacques Prost. Auditory sensitivity provided by self-tuned critical oscillations of hair cells. *PNAS*, 97(7):3183–3188, 2000.

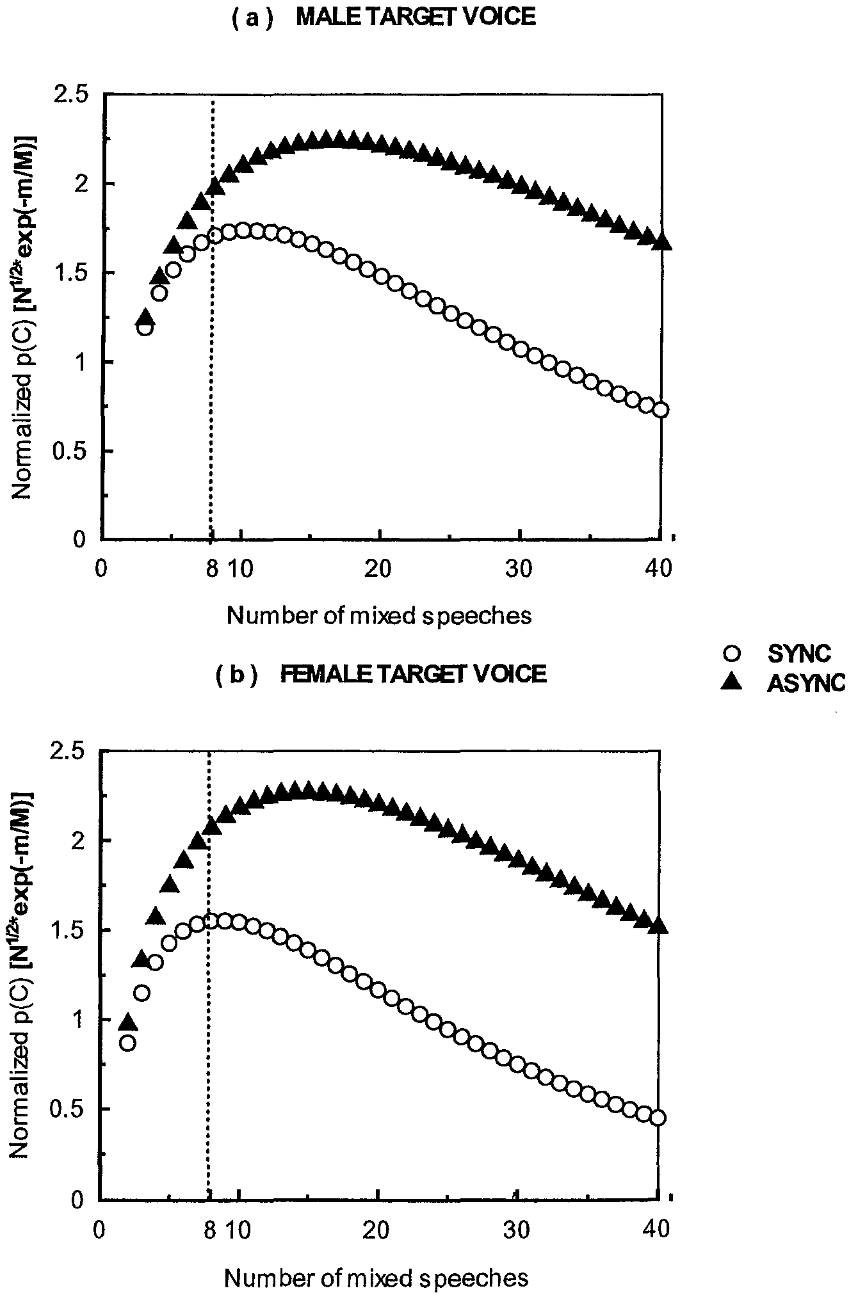

Figure 5: The normalized $p(C)$ performance plots based on the machine performance assumption.

論 文 審 査 の 結 果 の 要 旨

近年，検出器の閾値以下の信号を検出する場合に果たすノイズの役割が確率共鳴現象と呼ばれ，その基本的性質と生体への応用に興味がもたれている。これまでの確率共鳴に関する研究はガウス的ノイズに関するものに限られていたが，著者は，ノイズの特徴が確率共鳴に与える影響を詳細に研究すると共に，多数話者の音声が共存する場合の人の音声聴取能力（カクテルパーティ効果）の測定を行い，これまで応用対象がコオロギなどの下等動物のセンサ機能に限られていた確率共鳴の概念の拡大可能性を検討した。

本論文は，これらの成果をまとめたもので全文 4 章よりなる。

第 1 章は序論であり，研究の背景と目的を明らかにしている。

第 2 章では，ランダムなガウス分布を持つノイズのみならず，ノイズの特徴を一般化して確率共鳴を更に効率化する方策について研究した結果について述べている。ノイズとしてガウス分布であるが時間相関を持つ場合，一様分布をもつ場合，カオス時系列で異常分布を持ちかつ時間相関を持つ場合を比較することにより，ガウス分布のように確率共鳴に有効な成分が少ない場合は，ノイズの相関は確率共鳴の効率を増強するが，有効成分が多い分布の場合は逆に相関は効率を減少させることを明らかにした。これは，確率共鳴を効率化する条件を初めて明らかにしたもので，評価できる。

第 3 章では，複数話者が同時に発話している状態において，特定の話者の音声信号を正しく認識する割合を測定した結果について論じている。話者数の増加とともに，認識率は一般に指数関数的に減少することが示された。更に機械的な検出器の SN 比が話者数の平方根に逆比例して減少することを考慮すると，人の聴覚の相対的認識率は，10～20 名程度の話者数で最大になり，閾値検出器で示される確率共鳴と類似の振る舞いを示すことを明らかにした。これは雑音中における人間の音声聴取能と確率共鳴の関連を示唆する興味ある結果である。

第 4 章は結論である。

以上要するに本論文は，閾値検出器のしきい値以下の信号に対する確率共鳴現象に対するノイズの分布と時間相関の影響を明らかにすると共に，多数話者の音声信号が共存する場での特定話者に対する認識率が確率共鳴と類似の効果を示すことを明らかにしたもので，信号処理工学，システム情報科学の発展に寄与するところが少なくない。

よって，本論文は博士（情報科学）の学位論文として合格と認める。