3rd International Conference on
Spoken Language Processing (ICSLP 94)
Yokohama, Japan
September 18-22, 1994

Chaotic Characteristics of Voice Fluctuation and Its Model Explanation: Normal and Pathological Voices

Yuki Kakita and Hitoshi Okamoto

Dept. of Electronics, Kanazawa Institute of Technology, 7-1, Ohgigaoka, Nonoichi-machi, Kanazawa-Minami, 921, Japan

ABSTRACT

This paper reports on an analysis of vocal irregularity in pitch and amplitude in terms of (1) the "pseudo-phase-portrait," (2) the "fractal dimension" of the fluctuation, and (3) a model generating the fluctuation. A simple mathematical model generating the time series of fluctuation is proposed. Using this model, varying aspects of fluctuations as well as of the pseudo-phase-portraits are successfully described with a single control parameter. We also report that, by the simulation based on the model, a resultant intermittent type of fluctuation can manifest the chaotic behavior.

INTRODUCTION

In this paper, we analyzed the fluctuation of voice from the viewpoint of "Chaos." Chaos is defined as a condition of the system whose output manifests a random-like behavior in spite that the system is deterministic. The term "deterministic" means that the output is completely described by a rule which can be expressed, for example, by a mathematical function. Another significant characteristics of the chaos is that the chaos is very sensitive to the initial condition.

We use the "fractal dimension" for measuring the "randomness," or the "degree of complexity." Generally speaking, the value of fractal dimension in positive real number, but not in integer, suggests that the output could be chaotic [1].

To check the complexity of the fluctuation, we examined not only the time series but also the pseudo-phase-portraits. The phase portrait is a graphical representation of the attractor, which is represented as a set of trajectories and can be considered as a generator of a signal. Fig. 1 shows the relationship between the signal and its representation in the dynamical systems. Fig. 1(a) shows the original speech waveform, and (b) is called "the phase-portrait" reconstructed from the waveform. Fig. 1(c) shows the time series of the F_0 (fundamental frequency) (or amplitude) of the speech waveform, and (d) indicates "the pseudo-phase-portrait" reconstructed from the time series. The pseudo-phase-portrait is considered as a section of the phase-portrait as illustrated in Fig. 1(e).

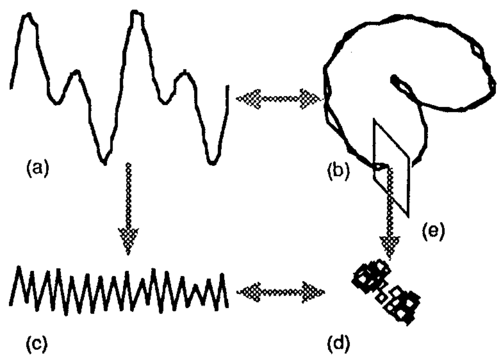

Fig. 1. The relation between the signal and its trajectories in phase-space. (a) Original speech waveform. (b) The phase-portrait. (c) The time series of the fluctuation. (d) The pseudo-phase-portrait. (e) The pseudo-phase-portrait is indicated as a section of the phase-portrait.

METHOD AND EXPERIMENT

Following is a brief explanation about how to obtain the value of fractal dimension. (See [2], [3] for the detail.)

(1) Reconstruct the pseudo-phase-portraits for d-fold (=1, 2, 3, ...) dimension by "embedding method." (2) Calculate the correlation dimension for each pseudo-phase-portrait in d-dimensional space. A saturated value for great d region of the correlation dimension is called the fractal dimension.

We computed the fractal dimension for the time series of F0 and amplitude which were extracted from the speech waveform of sustained utterances of a Japanese vowel /e/. For the normal case, 20 voice samples (4-2 repetitions x 8 subjects (5 male and 3 female)), and for the pathological (rough) cases, 39 voice samples (4-2 repetitions x 17 subjects (16 male and 1 female)) were analyzed. The pseudo-phase-portraits were classified and used for constructing a model which generats fluctuations.

RESULTS AND DISCUSSION

I. Pseudo-phase-portraits

Figure 2 shows typical examples of the time series of F_0 and corresponding pseudo-phase-portraits. Time series (a) is for a normal voice, whereas (c) and (e) for the "rough" voices.

Fig. 2(a) manifests a small amplitude fluctuation, and (b) indicates that the corresponding pseudo-phase-portrait is a "point." Fig. 2(c) indicates "period-doubling" (high-low pair) effect, and (d) indicates that the portrait is a pair of two-points. Fig. 2(e) indicates that the bursts of "period doubling" occur intermittently, and (f) shows corresponding portraits in "double-cycles." The change from (a) to (e) (through (c)) reminds us of a "bifurcation phenomenon" which means the "branching of states" or "repeat two-different values alternately" with respect to F_0. The same explanation can be done for the change in the shapes of the portraits from (b) to (f) (through (d)).

By observing the time series of fluctuation and the shape of the pseudo-phase-portrait, we can obtain information about the dynamical systems underlying the fluctuation.

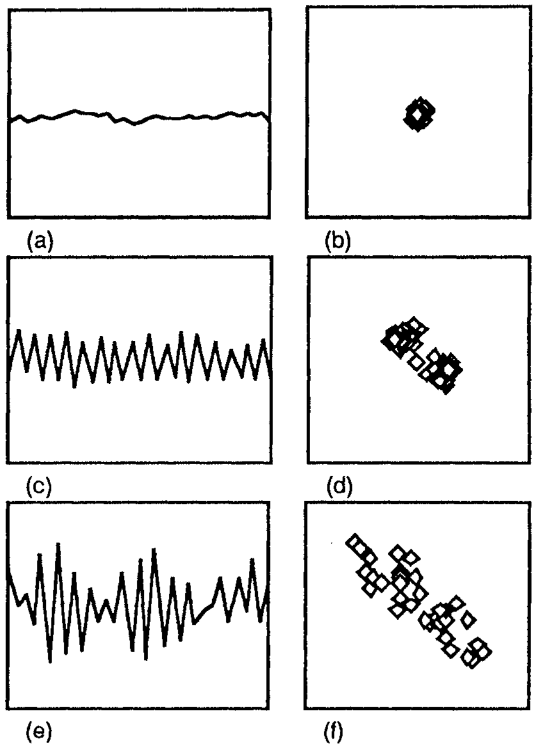

Fig. 2. Examples of the time series of F_0 fluctuation (a, c, e) and corresponding pseudo-phase-portraits (b, d, f).

II. Fractal dimensions

The median value of fractal dimension D is about 3.5 for the normal, and is about 4.4 for the "rough" voice. The values of the dimension correspond to the number of independent variables describing the dynamical system generating the particular fluctuation [1]. Therefore, assuming a single system that can generate the fluctuations both for the normal and the "rough," 5 (>4.4 >3.5) independent variables are necessary. The positive real (non-integer) number of fractal dimension can be an indication of a chaotic dynamical system.

III. Model for generating fluctuations

Now we construct a model for generating the time series of fluctuation in the deterministic sense. Why we make a deterministic model instead of a stochastic one is that we want to discriminate the patterns of fluctuation signal or its pseudo-phase-portrait in a specific manner, i.e. how doubles the period, and how far the two limit-cycles are separated, etc. This kind of approach leads us to obtain the specific information about the underlying physical system of the vocal fold vibration. In the previous section, we reported that the system for the fluctuation needs at least 5 variables. In this section, however, we propose a new model with 2 variables for simplicity. We will show that, even with 2 variables, significant characteristics can be explained. The model is represented as discrete equations because it generates time series of fluctuation which is inherently discrete.

To construct a new model of the fluctuation, we combined "the Hénon map model" and "the delayed logistic map model."

a. Hénon map

The Hénon map is represented as

$$X_{n+1}=Y_n+1-aX_n^2 \quad \text{..... (1a)}$$

$$Y_{n+1}=bX_n \quad \text{..... (1b)}$$

Basic geometric characteristic of the Hénon map is the "folding" (Fig. 3(a)) which causes the bifurcation observed as the "period doubling" [4]. The two sections in Fig. 3, (i) shows the pseudo-phase-portrait corresponding to the one shown in Fig. 2(b), and (ii) shows the pseudo-phase-portraits corresponding to that in Fig. 2(d).

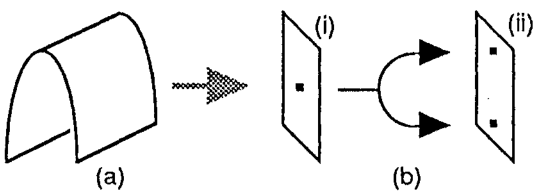

Fig. 3. Geometric characteristics of Hénon map. (a) the "folding" transformation. (b) the "period doubling" bifurcation. (i), (ii) the pseudo-phase-portrait before and after the bifurcation.

b. Delayed logistic map

The delayed logistic map is represented as

$$X_{n+1}=Y_n \quad \text{..... (2a)}$$

$$Y_{n+1}=aY_n(1-X_n) \quad \text{..... (2b)}$$

Basic geometric characteristic of the delayed logistic map is the "twisting" (Fig. 4(a)) which causes the bifurcation observed as "point to cycle" [5]. In Fig. 4(b), the section (i) shows the pseudo-phase-portrait (point) corresponding to that in Fig. 2(b), and the section (ii) shows the pseudo-phase-portrait (cycle).

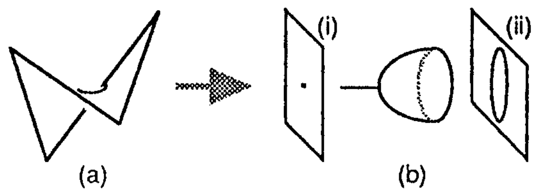

Fig. 4. Geometric characteristics of delayed logistic map. (a) "twisting" transformation. (b) "point to cycle" bifurcation. (i), (ii) the pseudo-phase-portrait before and after the bifurcation.

c. Our model

A new model which we propose has both the effets of "folding" and "twisting" transformations. (Fig. 5) Therefore, we assume that the model can generate both the bifurcation "double limit cycle" as well as of "double cluster."

Our model is represented as

$$X_{n+1}=X_n(Y_n^2-a) \quad \text{..... (3a)}$$

$$Y_{n+1}=X_n\,. \quad \text{..... (3b)}$$

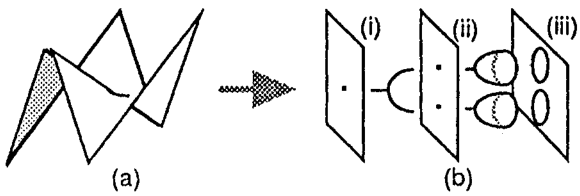

Fig. 5. Geometric characteristics of the proposed model. (a) in the representation of dynamical systems, both the "folding" and "twisting" are realized at the same time. (b) This model causes both the "double cluster" and "double limit cycle" bifurcations. The three sections shown in (i), (ii), (iii) correspond to pseudo-phase-portraits shown in Fig. 2(b), (d), (f) , respectively.

IV. Simulated pseudo-phase-portrait

Figure 6 shows the computer simulated pseudo-phase-portraits and the time series generated by our model. For the time series and the portrait for a simulated normal voice, Fig. 6((a), (b). parameter a =1.0) corresponds to Fig. 2(a), (b), respectively. Likewise, for a simulated "rough" voice, Fig. 6(c) and (d) (a =1.1) correspond to Fig. 2(c), (d). Likewise, Fig. 6(e) and (f) (a=1.6) correspond to Fig. 2(e) and (f), respectively.

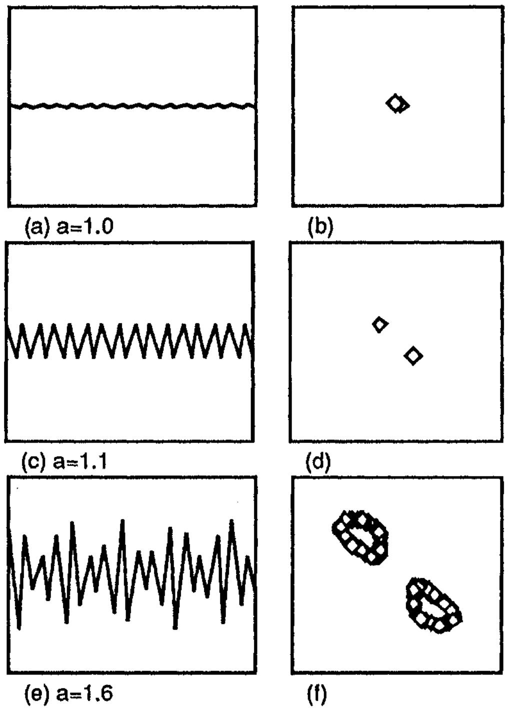

Fig. 6. The computer simulated time series (a, c, e) generated from our model and the pseudo-phase-portraits (b, d, f).

V. Prediction using our model

We present an example of the predicted time series of fluctuation using our model. Fig. 7(a) shows an output of our model suggesting that the output behaves like an "intermittent chaos." Fig. 7(b) shows its pseudo-phase-portrait suggesting that it is approximated as "connected double-cycles." The portrait is considered to be a further extended state of the "two-cycles" illustrated in Fig. 5. Fig. 7(c) shows an observed example whose behavior is close to that in (a). The pseudo-phase-portrait (d) looks similar to (b), except that the point intersection as indicated (i) in (b) is shown scattered in (d), probably due to a cyclic motion suggesting a slow drift of the baseline (or "trend"). On the other hand, outer cycles as indicated (ii) in (b) are not clearly seen in (d). This suggests that the speed of trajectory is too fast and/or the intermittent bursts occur less frequently, resulting the density of points being sparse. We also checked the sensitivity of the model to the initial condition. A slight change in the initial condition causes a variety of changes in the shape of fluctuation time series. This fact suggests that our model has a capability of generating chaotic behavior.

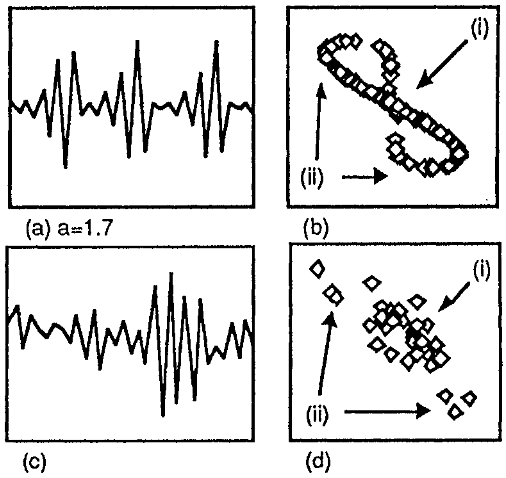

Fig. 7. A predicted behavior with our model. Computer simulated (a) time series and (b) the pseudo-phase-portrait using our model (a=1.7, in Eq. 3). Observed example of (c) the time series and (d) the derived pseudo-phase-portrait.

REFERENCES

[1] H. Takayasu, (1986). "Furakutaru (Fractal)" Asakura Shoten. (In Japanese)

[2] Y. Kakita, and H. Okamoto, (1992). "Fractal dimensional analysis of voice fluctuation: normal and pathological cases" Proceedings of the Fourth Australian International Conference on Speech Science and Technology (SST-92), Dec.1-3,1992. 130-133.

[3] P. Grassberger, and Procaccia, (1983). "Characterization of strange attractors" Phys. Rev. Lett. 50: 346-349.

[4] M. Hénon, (1976). "A Two-dimensional Mapping with a Strange Attractor" Commun. Math. Phys. 50, 69-77.

[5] D. G. Aronson, M. A. Chory, G. R. Hall, and R. P. McGehee, (1982). "Bifurcations from an Invariant Circle for Tow-Parameter Families of Maps of the Plane: A Computer-Assisted Study" Commun. Math. Phys. 83, 303-354.