Recognition of Individual Cattle by His and/or Her Voice

Yoshio Ikeda*, Yohei Ishii*
Foreign Member

Abstract

It was assumed that the voice of cattle is generated with the virtual white noise through the digital filter called the linear prediction filter, and filter parameters (prediction coefficients) were estimated by the maximum entropy method (MEM), using the sound signal of the animal. The feature planes were defined by the pairs of two parameters selected appropriately from these parameters. The cattle voices were divided into three levels, that is the high, medium and low levels according to their total power equivalent to the variances of the sound signal. It was found that the straight lines could be used for recognizing two cow and one calf for high level voices. For high and medium level voices, however, it was difficult or impossible to recognize individual cattle on the parameter planes.

Keywords: cattle voice, recognition, linear prediction filter, prediction coefficients, maximum entropy method

1. INTRODUCTION

It is said that experienced herdsman may identify problems through animal vocalization (Xin, H. et al, 1989). It is pointed out that the mechanization of livestock farming should be realized under considering animal welfare, for example without stress and invasion to animals (Fukukawa, K.,1994). One powerful technique for monitoring animal without invasion may be possible by using machine vision (Van der Stuyft, E. et al, 1991, Onyango,C.M. et al, 1995). Another method is to analyze the vocalization of animal since vocalization or call may contains information on animal (Kashiwamura,F. et al. 1985, Xin, H. et al, 1989, Ikeda, Y. et. al.,1991, Tamaki,K. et. al., 1993).

Variation of voice characteristics of animal has two aspects, one is between animals

*Graduate School of Agriculture, Kyoto University, Kyoto, Japan 606-8502

and other is within animal. The former could be used to recognize the individual animal, and the latter to monitoring of animal conditions such as hunger, oestrus and sickness. One of the final goal of voice analysis is to understand the intelligent behavior of the animals and synthesize the animal voice or vocalization for interchanging information with animal for breeding.

The objective of the present paper is, however, to analyze the voice of cattle to recognize the animal individually with their voice. As the first step of the research, it is necessary to clarify the characteristics of cattle voice itself. In this research, we measured the voice of six cattle, that is four adult cow and two male calves, analyzed the voice in time and frequency domains and tried to distinguish the cattle with their voice characteristics.

2. LINEAR PREDICTION MODEL AND POWER SPECTRUM

The discrete version of the general relationship between input signal $e(t)$ and output signal y(t) of linear system is in the digital version as

$$y(m)=-\sum_{i=1}^{p} a_i y(m-i)+\sum_{i=0}^{q} b_i e(m-i) \tag{1}$$

and this signal generating process is called auto-regressive moving average (ARMA) model (Morishita, I. et al, 1982). When b_i=0 for i=1,2,..,q, this model is called the auto-regressive model (or all-pole model or linear prediction filter) and is given as

$$y(m)=-\sum_{i=1}^{p} a_i y(m-i)+b_0 e(m) \tag{2}$$

Power spectrum of the output signal y(t) can be computed by the following equation, where the values of coefficients a_i could be estimated by the MEM (Hino,M. 1986.)

$$P(f)=\frac{\Delta t P_n}{\left|1+\sum_{k=1}^{n} a_k e^{j2\pi fk\Delta t}\right|^2} \tag{3}$$

where Δt is the sampling interval of the signal and P_n is the mean square error between the observed signal and the estimated signal.

Power spectra could be used to detect the formant frequency, and the parameters a_k(k=1,2,..,n) may give quantitative description of animal voice characteristics.

3. MEASURING METHOD AND MEASURED CATTLE

The voices of cattle were recorded from September 11 to 14 in 1995 at 7 to 8 o'clock in the morning before feeding, using the precision microphone(RION, NA-60) and digital audio tape recorder(Pioneer, DATA-05) set at a distance of 60 cm from the fence of pen (150 cm height) and at a height of 100 cm from the floor. The data sampling interval was 0.0002 s and data interval was 2 s during which overall voice was recorded. The weather were fine or cloudy, and so background noise was the low level sound of circulating fans and the metallic sound caused by bump of the fence and the cattle. Cattle might not be disturbed with the measuring devices. The cattle whose voices were measured are as shown in **Table 1**.

Table 1 Cattle(Japanese Black) used for Experiment

	Cattle No.					
	143	440	445	683	695	698
Age(Month)	133	78	78	13	9	9
Sex	Female	Female	Female	Female	Male	Male
Body Mass(kg)	717	540	661	380	253	227
Number of Voices	29	57	34	28	31	48

4. RESULTS AND DISCUSSION

Total Power of Sound Wave The voice signals could be divided roughly into three groups (high, medium, and low levels) according to the total power of sound, that is the variance of voice signal. In this report, we will discuss the characteristics of the voice of three cow, No. 143, No. 440 and No. 445, as well as one calf No. 697.

Order of Model The number p of terms in Eq. (2) should be so determined as the final prediction error (FPE) between the predicted and measured signals is minimum. In this experiment, the FPE values decreased quickly until p reached 10 to 15 and after that those values decreased gradually, then the value of p was decided to be fifteen.

Time-varying Power Spectra The amplitude of the sound signal of the animal change in time, then the short term power spectra calculated for the successive segments excised from the whole length of the signal changed with time as shown in **Fig. 1**. The data length of one short term was 0.4 s and the time shift to the next short term was 0.05 s. It is noticed that several components were involved in one voice, and the spectral structure of the voice changed with time and between cattle. For cow No. 440,the harmonic structure was recognizable. The frequencies of the main components were not changed during vocalization.

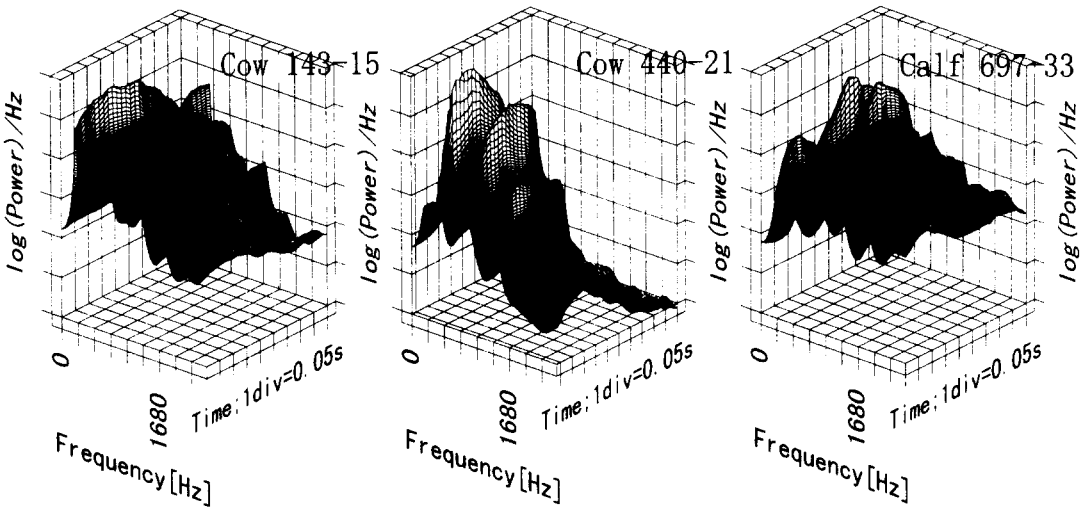

Fig. 1 Time-varying Power Spectra of Cattle Voice

Time-varying Variance For each short term segment, the integral of power spectral density and the variances of the voice signal were calculated. These curves are roughly similar to the amplitude envelopes of the sound signal, and it was confirm that the values of integrated power coincides with variances theoretically. In the subsequent analysis, the data segments of the maximum power for each voice were used.

Power Spectra for Maximum Power Segment In **Fig. 2**, power spectra for the data segment of the maximum power or variance are shown. It is noticed that the spectral structures are different between cattle. For cow No. 440 whose spectra is drawn with ■, the harmonic overtone can be seen clearly, the fundamental frequency is 500 Hz, and has the sparse harmonic(Hopp.S.L. et. al. Eds.,1998.) For cow 143(○) and calf 697(bold line), however, the harmonic structure can not be found. Each sound had three to four prominent components.

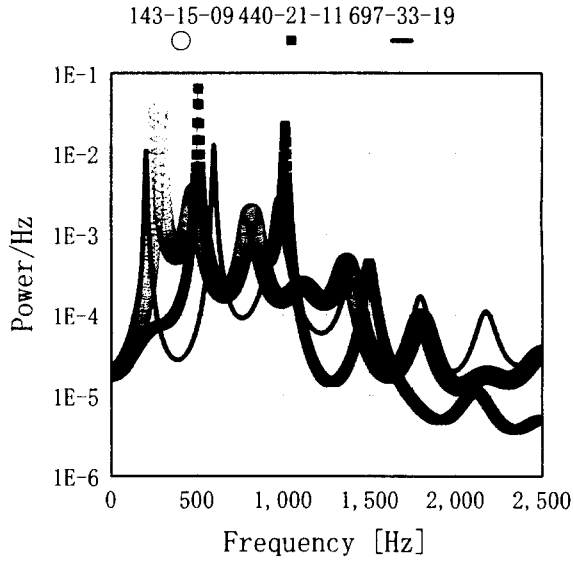

Fig. 2 PSD of Maximum Power Segment

Prediction Coefficients The differences in the spectral structure may be reflected on the prediction coefficients of the linear filter model of cattle voice. These coefficients are used to recognize the individual animal, defining the feature space with two appropriate variables selected from these

fifteen coefficients as the individual can be separated distinctly on the 2D feature plane with the simple line such as the straight line.

Separability of Individual Animal on Feature Plane In **Fig. 3**, it may be expecting that the 2D plane whose axes are coefficients a_1 and a_9 can be used for recognizing two

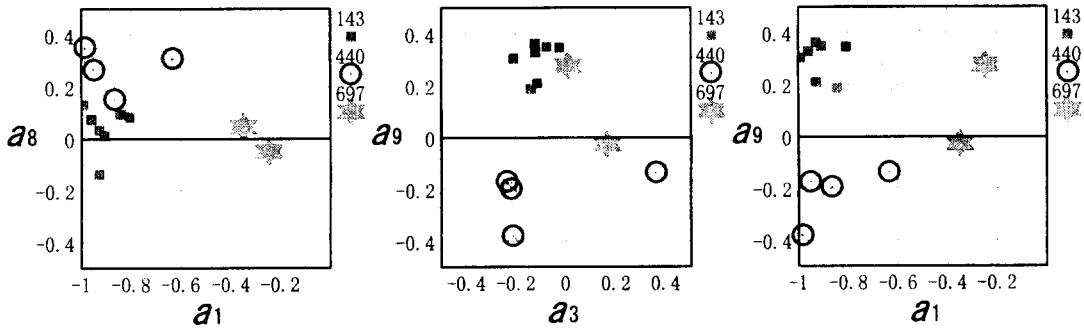

Fig. 3 Recognition of Two Cow and One Calf on Coefficients Space

cow and one calf. If three coefficients are used as the coordinates for the 3D feature space, it may be possible to recognize these animals more clearly. It was found that the recognition for the high and medium level voices of three cow and one calf was difficult or impossible. Including the medium level voices in features makes recognition unfeasible. The dimension of feature space should be more than three to recognize more cattle and highly advanced method for recognition technique such as dicriminant analysis or cluster analysis.

Estimation of Formant Frequency with All-pole Model The characteristic equation of the transfer function of all-pole model of order n has n complex conjugate roots or poles. The poles whose imaginary parts are positive give the formant frequencies (Deller,J.R. et. al.,1993.) This algebraic equation was solved numerically by the Durand-Kerner-Aberth method(Mori,M., 1987 .) The formant frequencies estimated by two methods, that is PSD and numerical solution, did not coincide considerably, and the frequencies computed from the all-pole models did not have the relation of integral multiplication. The reasons of these discrepancies between the measured and estimated values were not known, and should be explored including investigation of the order of the auto-regressive model, and furthermore modeling itself of voice generation system should be examined.

5. Conclusions

It was ascertained that we could recognize two cow and one calf by using the prediction coefficients of hypothetical signal generating system of the animal voices

and the simple linear discriminant function on the 2D feature plane. The next step of this research for recognizing animals may be as follows.

For the fundamental research for recognition of the individual animals, it may be necessary 1). To examine applicability of this technique to mute voices, 2). To investigate for more voices of more cattle, 3). To consider non-stationarity of voice signal concerning duration time and amplitude envelope pattern of the signals for selecting the qualified feature parameters, and 4). To develop the effective and stable algorithm for recognition.

For development and design of reliable and harmless monitoring system for animal welfare, it may be necessary to research voice characteristics under various physical conditions of animals including long term variation due to their growth.

Further, in order to realize precise recognition, it is necessary to utilize the image information such as face, markings and physique of body together.

References

Deller,J.R.,Proakis,J.G.,Hansen,J.H.L.; Discrete-Time Processing of Speech Signals, Macmillan, (1993)

→ (1994), Vol.56, No.2, 167-172

Fukukawa,K.;New Technology for Livestock Breeding and Management(in Japanese), J. of JSAM,→

Hino,M.;Spectral Analysis (in Japanese), Asakura Book Co., (1986)

Hopp,S.L.,Owren,M.J.,Evans,C.S.,Eds.;Animal Acoustic Communication, Springer, (1998)

Ikeda,Y.,Shimada,J.;Application of Acoustic Emission in Agricultural Mechanization -Frequency Characteristics of Livestock-(in Japanese), J. of JSAM Kansai Branch, (1991), No.70, 123-124

Kashiwamura,F.,Yamamoto,M.,;The Classification and Analytical Method of the Vocalization of Cattle (in Japanese), Livestock Management,(1985), Vol.21,73-83

Mori,M.;Programming for Numerical Computation(in Japanese), Iwanami Shoten,(1987)

Morishita,I.,Kobatake,H.,;Signal Processing (in Japanese), SICE Press, (1982)

Onyango,C.M.Marchant,J.A.,Ruff,B.P.,;Model Based Location of Pigs in Scenes,Comp. and Electronics in Agric., (1995), Vol.12, 261-273

→(1993), 1-11

Tamaki,K.,Susawa,K.,Otani,R.,Amano,K.,;Characteristics of Catle Voices and the Possibility of Their Discrimination (in Japanese), Research Bull. No.158 of the Hokkaido Nat. Agric. Exp. Station, →

Van der Stuyft,E.,Schofield,C.O.,Randall,J.M.,Wambacq,P.,Goedseels,V.;Development and Application of Computer Vision System for Use in Livestock Production, Comp. and Elec. in Agric., (1991), Vol.6, 243-265

→ASAE, (1989), Vol.32, No.6, 2181-2184

Xin,H.,DeShazer,J.A.,Leger,D.W.;Pig Vocalization Under Selected Husbandry Practices,Trans. of→