2008 International Symposium on Nonlinear Theory and its Applications
NOLTA'08, Budapest, Hungary, September 7-10, 2008

Inferring the intention of a Crying Baby

- Characterization of Painful Cry by Pitch Variation -

Tomomasa Nagashima‡ , and Hidenori Tanaka†

‡Department of Computer Science and Systems Engineering, Muroran Institute of Technology
27-1, Mizumoto-cho, Muroran-shi, Hokkaido 050-8585, Japan
† Satellite Venture Business Laboratory, Muroran Institute of Technology
27-1, Mizumoto-cho, Muroran-shi, Hokkaido 050-8585, Japan
Email: nagasima@epsilon2.csse.muroran-it.ac.jp, tanaka@mail.svbl.muroran-it.ac.jp

Abstract– Cry is the only way for baby to express his /her intention to the others. In contrast to adult, baby does not have ability to operate any language for communication. So, it is an interesting general question how mother knows the intention of her baby, when baby cries. Did baby cry due to hunger, unpleasantness in diapers or pain?

In the acoustic study of voices, it has been developed along the two directions i.e., 1) recognizing the speech (contents) or speaker, and 2) how voices affect listeners impression. For the latter issue, we have previously investigated the relation between pitch variation of baby's cry and listening impression by focusing the instability of fundamental period of baby's cry.

In this paper, we investigate the other problem of baby's cry, i.e., emotion recognition of baby, which can be considered to be a part of speech recognition/speech understanding. Especially, we focus our attention to baby's cry coming from pain. We analyzed the irregular variation of wave forms based on new index ePPQ and discussed the possibility to discriminate the painful cry from the cries with other causes. By comparing two cries coming from pain and other cause, it is suggested, based on ePPQ, that there is a possibility to discriminate baby's cry resulting from pain from other cries. It will be important to discriminate the painful cry from the cries with other causes, like hunger, unpleasantness or fear, because the baby's cry resulting from pain might be a signal indicating some unknown disease.

1. Introduction

Recently, working couple becomes popular in younger generation of Japan. Accordingly, child raising becomes one of the social problems, because for many young parents, it is not easy task to make parenting, especially for those who have never had such experiences. They must rear their child by themselves without the aid of relatives. It is pointed out that failure of child raising may induce child abuse in the worst case. The cause of child abuse comes down to failure of communication between parents and children. So, it becomes an important social issue to construct proper strategy to prevent the problem. In this context, it is desired to develop a system-technology supporting the communication between parent and child. In the communication between human and human, the voice is one of the most useful tool. It is the case even for babies whose faculties of speech are not developed sufficiently. But, for baby, it is more than important, because cry is the only tool for baby to communicate.

While there is few research for the voices of baby's cry, in a previous paper, we have confirmed that baby's cry is nonlinear phenomena. Therefore, it will be interesting subjects in nonlinear systems [1]. In the communication through cries from babies to parents, emotion of the parents listening to the cries is very important. Mukai *et al*. suggested that the cries of baby who have a structural disease in the vicinity of throat give listener the unpleasant impression and this type of cry will be one of the causes of child abuse [3, 4].

Previously, we have made an investigation on the relation between listeners' impressions and acoustic features of babies' cry by proposing the extended Pitch Perturbation Quotient (ePPQ) to quantify "the instability of the fundamental periods" , and suggested that the ePPQ was a candidate to evaluate the listeners' unpleasantness for cry [5, 6].

In this paper, on the contrary, we investigate the recognition problem of emotion of baby by analyzing the voices of crying baby. Among all, we focus our attention to the emotion centered on pain. It will be important to discriminate the cry coming from pain from the cries with other causes, because it may contain the cases of some unknown diseases.

2. Some features on present status in analyzing Baby's voices

In a previous paper, we have shown, based on the method of surrogate data, that baby's voices including cry should be generated as nonlinear oscillations [1].

While it is still difficult to understand the whole phenomena of cries of baby by any existing theory, it might be said that one of the characteristic feature of voices of crying babies will be the existence of a kind of irregularity in the wave forms (oscillations) shown in Figure1.

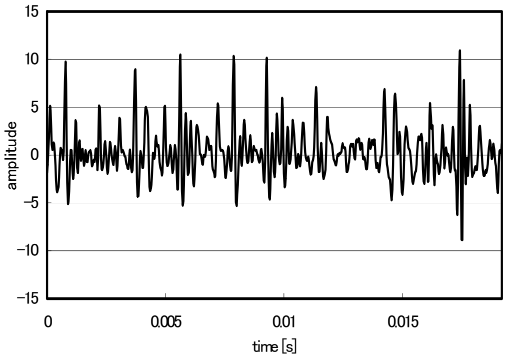


Fig.1 An example of random variation in baby's cry.

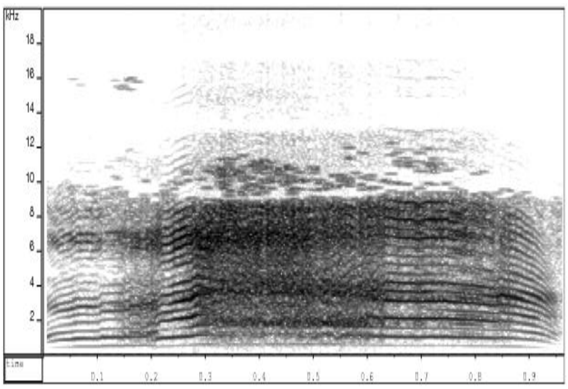


Fig.2 An example of Sound spectrum

The most common tool today to analyze the complex nature of voices is provided by the sound spectrum which provides the structure of a time-dependent frequency spectrum of the corresponding object (voice). An example of sound spectrum is shown in Fig.2.

As shown in Fig.2, in the sound spectrum, irregularity of oscillation may be manifested by the existence of vague line structure or diminishing line structure.

Recently, Arakawa et al. [7, 8] has reported that in contrast to the case of the cry coming from hungry, the power spectrum of the cries coming from discomfort and pain look similar with each other. So, it is difficult to discriminate those cries by power spectrum.

Although the sound spectrum has been powerful tool and successfully used in analyzing the nature of the voices, it is required to develop the objective description/criterion for understanding the results of sound spectrum in applying the computer processing of voices, because it has been usual to understand/interpret the results in a subjective manner.

In the analysis of the voice of babies, the preprocessing of voice data is important. In general, a baby cry succeeds for the period more than one breadth. In analyzing the baby's cry, it is appropriate to separate sound data of a baby's cry into the data consisting of one breadth, more specific, one expiration.

In the following, we use this separated data as one data of baby's cry, and describe an essential feature of our method to evaluate the instability of fundamental period of baby's cry, which will quantify the irregularity of oscillations in baby's cry.

3. Extension of Pitch Perturbation Quotient: ePPQ

The ePPQ is an index to quantify the instability of fundamental periods. This was made by extension of PPQ that had been proposed by Kikuchi and Kasuya [9].

Here, the algorithm for calculation of ePPQ is expounded.

ePPQ is calculated by substituting time series of fundamental periods $p(n)$ into the following equation (1) and (2). The time series of fundamental periods was obtained by calculating the fundamental periods at a regular time interval in the sound data.

$$ePPQ = \frac{1}{N-2k}\sum_{i=k}^{N-k-1} A(i), \qquad (1)$$

$$A(i) = \left| 1 - \frac{p(i)}{\frac{1}{2k+1}\sum_{n=-k+i}^{k+i} p(n)} \right|. \qquad (2)$$

Equation (1) denotes that ePPQ is the average of $A(i)$. We label $A(i)$ as "perturbation value". Equation (2) denotes that the perturbation value $A(i)$ is an absolute value of the difference between 1 and current fundamental period divided by the average of fundamental periods for $2k+1$ points which center on the current time point.

To calculate ePPQ, first of all, the time series of fundamental period must been given. In this study, we applied the method of Average Magnitude Difference Function (AMDF) [10] to extract the fundamental period.

In this method, to begin with, the following $r_t(l)$ was calculated as an AMDF for the part of waveform data immediately after the time point t at which fundamental period should be obtained and a part of waveform data after l points delay.

$$r_t(l) = \sum_{m=t}^{t+M-1} \left| x(m) - x(m+l) \right|, \qquad (3)$$

where $x(m)$ denotes the signal data of the sound and M was the length of the compared data. Then, the threshold r_{th} was calculated from the minimum value $r_{\min}$ and the average value r_{av} of $r_t(l)$ by the following equation,

$$r_{th} = r_{\min} + (r_{av} - r_{\min}) / 4. \qquad (4)$$

Finally, the fundamental period was calculated by multiplying sampling period by minimal l which made $r_t(l)$ less than or equal to the threshold r_{th}. The time series of the fundamental periods was obtained by calculating the fundamental periods at the regular time interval in the same way. The ePPQ was defined by substituting the time series of the fundamental period $p(n)$ into the equation (2).

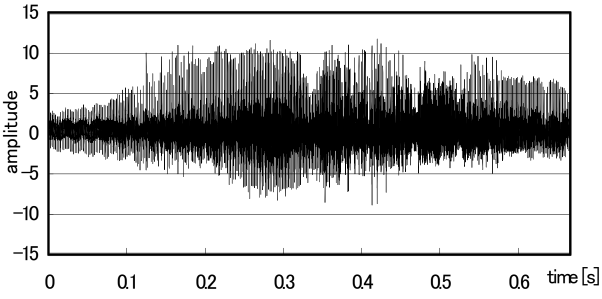
3. Preliminary experiment: Application of ePPQ to real cries with pain and other causes

In this chapter, we show an example of calculating ePPQ from babies' cry practically.

Here, we used the data of crying baby that had been disclosed by Dr. Mukai in his homepage (http://www02.so-net.ne.jp/~s-mukai/). These cry data were uttered by pinching lightly at the foot of a 1-month-old boy. To record the cries, MINIDISC RECORDER (MDS-102, SONY) was used. Microphone was set 15 cm above mouth of the babies. The data were downloaded from homepage and resampled to 40 [kHz]. Then the voiced parts were cut out from them by the visual observation. Here, we used a voiced part which was cut out from the cries with pain data 1) and a voiced part which was cut out from the cries without pain (data 2) to calculate ePPQ.

The parameters in calculating ePPQ were set as below. In setting the time series of fundamental period, the fundamental periods were calculated at the time interval Δt. Δt was fixed to 20 points (i.e., 0.5 [ms]). In equation (1) and (2) of previous chapter, k was fixed to 1 and N was set to the length of time series of fundamental periods for whole target sound data. In equation (3), the compared data length M was set to 128 and delay l ranged from 40 to 280 (i.e., $l = 40, 41, \ldots, 280$).

Fig.3 and Fig.4 show the waveform and the time series of fundamental periods of data 1 and data 2 respectively. In each figure, the upper graph shows the waveform, where the horizontal axis shows time [s] and the vertical axis shows amplitude. And the lower graph shows the time series of fundamental period, where the horizontal axis shows time [s] and the vertical axis shows fundamental period [ms].



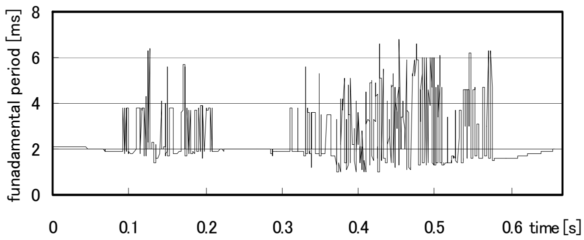
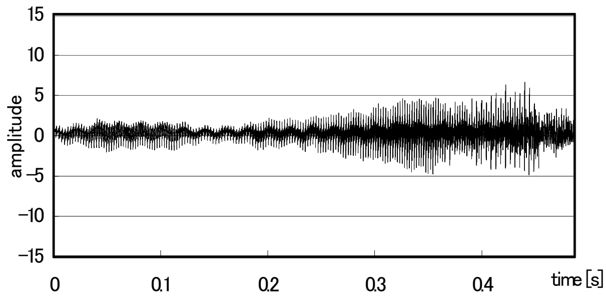


Fig.3: Waveform (upper) and fundamental periods (lower) of data 1.



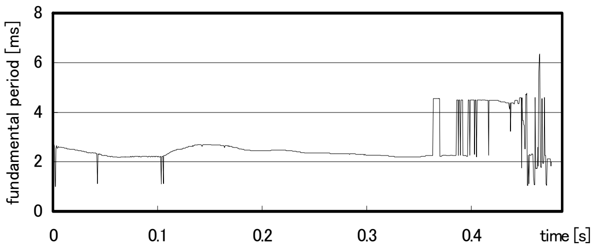


Fig.4: Waveform (upper) and fundamental periods (lower) of data 2

As a result of calculating the ePPQ, ePPQ for data 1 was 8.09×10^{-2} and ePPQ for data 2 was 2.49×10^{-2}. Consequently, the cry with pain has higher ePPQ than the cry without pain.

4. Conclusions and Discussions

In this paper, we have made a basic analysis for inferring the emotion of crying babies. For this purpose, we have used the previously introduced index ePPQ which quantifies the instability of the fundamental period of voices. As a preliminary experiment, we have compared two baby's cries with different causes, a cry with pain and the other cry without pain.

The result revealed that the ePPQ's for two cries take significantly different magnitude, implying a possibility to discriminate the emotions of crying baby. Specifically, it implies a possibility to discriminate the cries with pain from other causes. As our result in this paper has been drawn based on a preliminary experiment, in order to confirm our result suggested in this paper, it will be necessary to conduct a further precise experiment with a sufficient number of crying data of babies.

Acknowledgments

We would like to thank Dr. Susumu Mukai for providing us with the materials and valuable advices. We also thank to the financial supports from JSPS.

References

[1] J. Imai, H. Tanaka, and T Nagashima, "Analysis of Nonlinearities in Nonverbal Voices - Toward Extraction of Kansei Information via Nonverbal Voices -," *Kansei Engineering International*, Vol.5, No.3, pp.7-14, 2005.

[2] K. Michelsson, S. Paajanen, A. Rinne, H. Tervo, I. Kunnamo and T. Kiminkinen, "Mothers' Perceptions of and Feelings towards Their Babies' Crying," *Early Child Development and Care*, vol.65, pp. 109-116, 1990.

[3] S. Mukai, S. Nagasugi, Y. Higashino, "Sound Analysis through Subjective Evaluation of Colic Crying," *The Japan Journal of Logopedics and Phoniatrics*, vol.38, pp.1-10, 1997 (in Japanese).

[4] S. Mukai and S. Nagasugi, "Sound Analysis through Subjective Evaluation of Colic Crying - Second Report: Colic -," *The Japan Journal of Logopedics and Phoniatrics*, vol.40, pp.349-356, 1999 (in Japanese).

[5] H. Tanaka, J. Imai, I. Saitoh, N. Kasuga and T. Nagashima, "Index for Unpleasantness of Baby's Crying: Proposition of Extended Pitch Perturbation Quotient and Evaluation of Unpleasantness," *Journal of Japan Society of Kansei Engineering*, vol.5, no.2, pp. 71-76, 2005 (in Japanese).

[6] H. Tanaka, T. Nagashima and S. Mukai, "Listening impressions and pitch variation of baby's cry - Evaluation by extended Pitch Perturbation Quotient -," *Proceedings of the 8th International Conference on Humans and Computers* (*HC-2005*), pp. 108-113, 2005.

[7] K. Arakawa, A. Horikawa, T. Yoshida, and S. Itoh, "Characteristics of the Crying Voice of Babies in Painful State", *IEICE Technical Report*, SIP2007-94, SIS2007-30, SP2007-56, pp65-67, 2007 (in Japanese).

[8] Y. Mima and K. Arakawa, "Cause Estimation of Young Babies' Cry from the Frequency Analyses of the Voice-Classification of Hunger, Sleepiness, and Discomfort-", *IEICE Technical Report*, SIP2006-79, SIS2006-42, SP2006-65, pp.43-45, 2006 (in Japanese).

[9] Y. Kikuchi, and H. Kasuya, "Development of Voice Evaluation System for Clinical Application," *Trans. IEE of Japan*, vol.108-C, no.10, pp. 873-878, 1988 (in Japanese).

[10] M. J. Ross, H. L. Shaffer, A. Cohen, R. Freudberg and H. J. Manley, "Average Magnitude Difference Function Pitch Extractor," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol.ASSP-22, no.5, pp.353-36, 1974.