Application of a Fuzzy Set Model to the Identification of Sounds

佐藤, 教昭
産業医科大学生体情報研究センター

岩宮, 眞一郎
九州大学大学院芸術工学研究院音響部門

https://doi.org/10.15017/2794911

出版情報：芸術工学研究. 4, pp.1-6, 2005-12-20. 九州大学大学院芸術工学研究院
バージョン：
権利関係：

研究論文

Application of a Fuzzy Set Model to the Identification of Sounds

佐藤教昭、岩宮眞一郎
SATOH Noriaki, IWAMIYA Shin-ichiro

スペクトル構造と立ち上がり時間を手がかりとした音の識別過程にファジィ集合の概念を適用し，聴覚系の音響情報処理過程のモデル化を行った．ファジィ積集合モデルを導入することにより，２つの物理量を手がかりとした場合の識別実験結果が，片方ずつを手がかりとした実験結果から予測できることを示した．この結果は，音の識別という音響心理学的分野へのファジィ集合の適用が可能であることを示唆している．

1. INTRODUCTION

Fuzzy set theory (Zadeh, 1965) has enabled the production of human subjective output in the fields of engineering and science that have traditionally required strict judgment. In the fields of control engineering and information science, applications of fuzzy set theory have been able to solve problems that had not been settled in traditional set theory. Through applications to electrical products such as washing machines, vacuum cleaners, and refrigerators, the usability of fuzzy set theory has been confirmed. Recently, fuzzy set theory has also been applied to matters in the humanities fields and the social sciences, including psychology (Matt et al., 2003; Takayanagi et al., 2000; Iwamiya et al., 1989).

We can recognize the types and states of those sounds even if they are in various environments. It is thought that the hearing process in such instances is conducted by verifying the current sound with sound images that have been memorized in the cerebrum ever since birth. This hearing process does not verify the sounds strictly but ambiguously. In other words, since the verification process is fuzzy, identification becomes possible (i.e., if the verification is strict, on the other hand, then voices of friends, etc., are only identifiable in their respective specific conditions).

In the present study, a fuzzy set model is applied to the identification of sounds, and the relationship between the physical parameters of prescribed sounds and the "likelihood of the sounds" is examined, taking note of the fact that the acoustical signals are recognized ambiguously when they are heard.

All stimuli in the experiments utilized synthesized sounds generated with a computer in order to exclude the effects of subjects' experiences or memories for specific sounds, such as those of musical instruments or speech, and in order to examine only the effects of acoustical clues on the identification of sounds.

2. EXPERIMENT I: IDENTIFICATION OF SOUNDS WHEN CLUE IS ONE OF THE SPECTRUM OR ATTACK PATTERN

2.1. Method and procedure

Five men served as subjects in all experiments. Their ages ranged from 24 to 30 years, and they had received technical listening training. They were tested individually in a double-walled sound attenuating chamber.

All stimuli were generated digitally, played by 16-bit digital-to-analog converters at a sampling rate of 20kHz, and low-pass filtered at 4kHz. The fundamental frequency of the stimuli was 440Hz. The stimuli consisted of five harmonic components. The duration of the stimuli was 1,500ms for all experiments, and all stimuli had the same offset, which was shaped by a linear function giving a fall time of 100ms. The stimuli had a level of 74.7dB SPL.

The onset of all stimuli was shaped by a linear function giving a rise time of 250ms in the case of identifying tones when the clue was the spectrum pattern. The standard tones were the spectrum patterns of -4, 0, and +4dB/oct. Fig. 1 shows examples for the spectrum patterns of the stimuli. First, subjects repeatedly listened to the standard tones so that they could perfectly identify them. Second, subjects heard one of 15 spectrum patterns, or -7, -6, -5, -4, -3, -2, -1, 0, +1, +2, +3, +4, +5, +6, +7dB/oct, changed every 1dB step. Subjects identified which of the three standard tones had been presented, and the degree of reliability of each response varied from 0% to 100% for each 10% change in the steps.

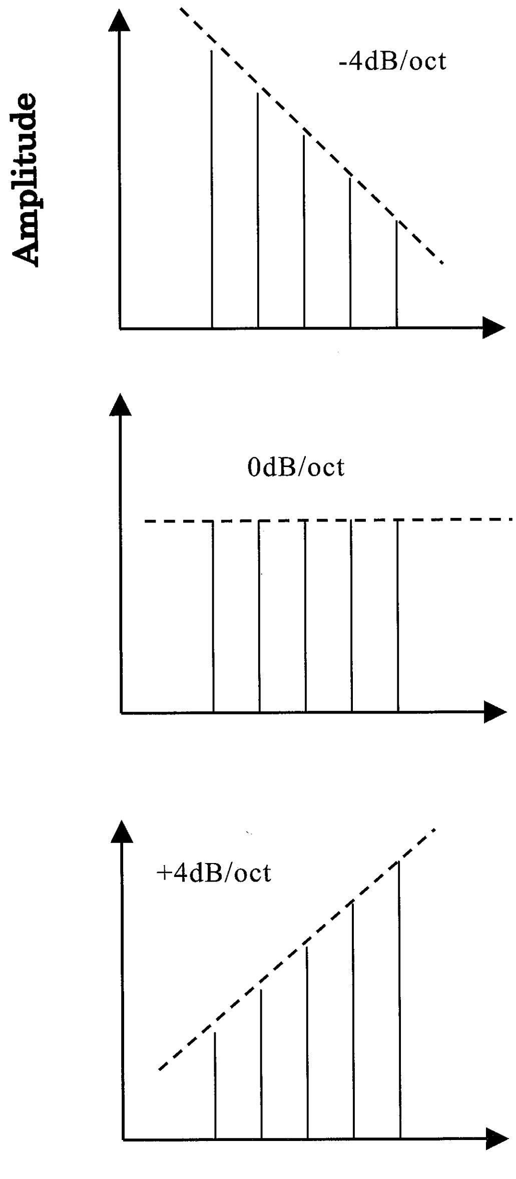

Fig. 1 Example for the spectrum patterns of the stimuli. The amplitudes of five harmonic components vary with the slope of the spectrum envelope.

In the case of the identification of tones when the clue was the attack pattern, the stimuli of the spectrum pattern were fixed at 0dB/oct. The onsets of the standard tones were 125, 250, and 500ms. The onsets of the test stimuli were 74, 88, 105, 125, 149, 177, 210, 250, 297, 354, 420, 500, 595, 707, and 841ms that had logarithmically equal intervals. Subjects identified the test tones using the same procedure as described above when the clue was the spectrum pattern.

2.2. Results and application of fuzzy set model

Fig. 2 shows the weighted response rates for all subjects in the case of the identification of tones as a function of the (a) spectrum pattern and (b) attack pattern. The weighted response rates are defined as the number of response for each standard tone multiplied by the degree of reliability. The weighted response rates for each standard tone are shown as different symbols. Not only true three standard tones but also other test stimuli were judged as standard tones. The weighted response rates reached a maximum when the test tones were the standard tones, and they decreased as the test tones moved away from the standard tones. Such a tendency suggests that subjects' memories of standard tones form a fuzzy set.

It is assumed that S_i is a fuzzy set of the "likelihood of being a standard tone" (i.e. how it seems to be a standard tone) as a function of spectrum pattern and that A_j (i,j=1,2,3) is that of a function of attack pattern. When each test stimulus is judged as a standard tone, the response rate is interpreted in terms of its degree of "likelihood of being a standard tone" and correspondence with the membership function. Membership functions of spectrum pattern (x) and attack pattern (y) are represented as $m_{S_i}(x)$ and $m_{A_j}(y)$, respectively, where $m_{S_i}(x)$ and $m_{A_j}(y)$ are independent. A fuzzy set C_{ij} of the identification of tones that forms in the case of clues of both the spectrum and attack patterns can be represented as a product set of S_i and A_j ($S_i \cap A_j$). In fuzzy set theory, the membership function of a product set is generally defined as

$m_{C_{ij}}(x,y) = \min[m_{S_i}(x), m_{A_j}(y)]$ (Zadeh, 1965).

Using this equation, results for the identification of tones in the case of clues to both the spectrum and attack patterns could be predicted.

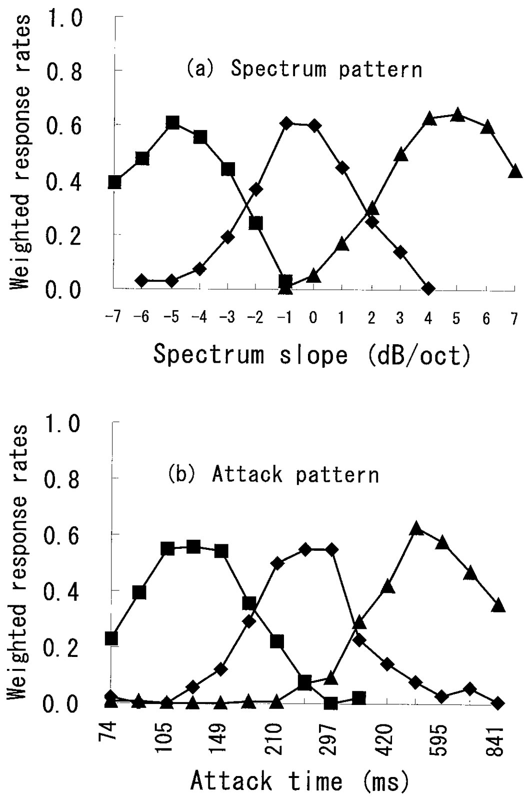

Fig. 2 Averaged weighted response rates as a function of (a) spectrum and (b) attack pattern.

3. EXPERIMENT II: IDENTIFICATION OF SOUNDS IN THE CASE OF CLUEs TO BOTH THE SPECTRUM AND ATTACK PATTERNS AND PREDICTION USING A FUZZY PRODUCT SET MODEL

3.1. Method and procedure

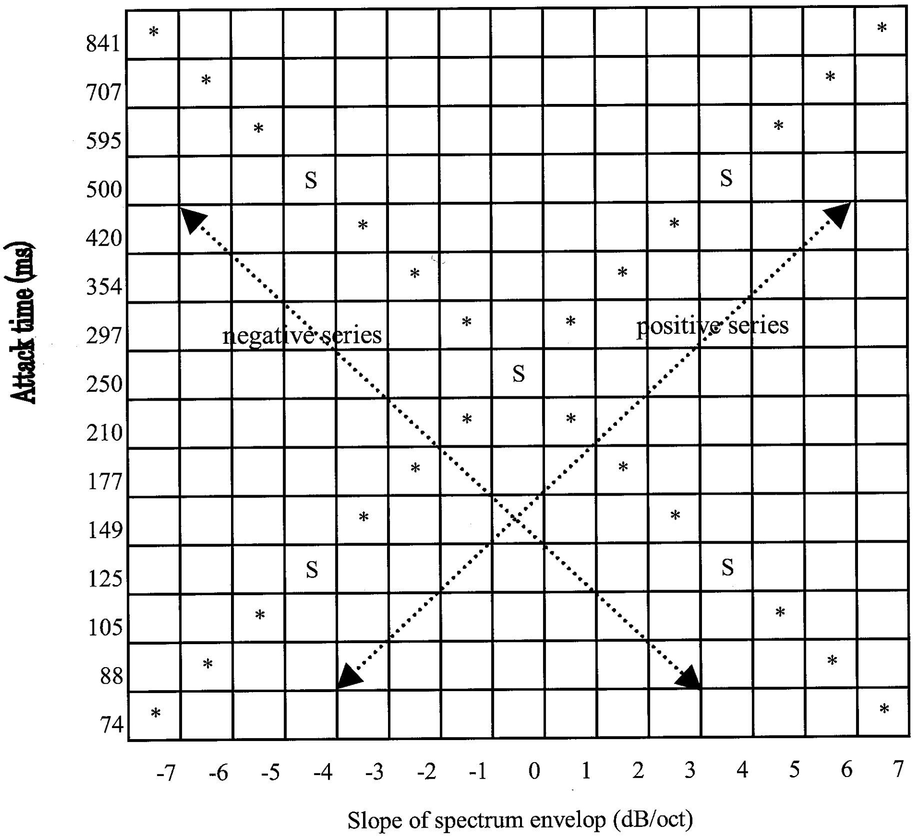

Fig. 3 Acoustical Properties of stimuli using Experiment II. Asterisks(*) show the test tones and the symbol "S" shows the standard tones.

The method and procedure used in Experiment I was used with the following changes in increasing the number of standard tones. First, the test stimuli consisted of a series of combinations in which the slope of the spectrum envelope increased as the attack time increased (positive series) and another series of combinations in which the slope of the spectrum envelope increased as the attack time decreased (negative series). The positive series consisted of -7dB/oct, 74ms to +7dB/oct, 841ms. The negative series consisted of -7dB/oct, 841ms to +7dB/oct, 74ms. Thus, the test stimuli consisted of 29 different tones. Second, the five standard tones were (-4dB/oct, 125ms), (-4dB/oct, 500ms), (0dB/oct, 250ms), (+4dB/oct, 125ms) and (+4dB/oct, 500ms). Fig. 3 shows the acoustical properties of the stimuli used in this experiment.

3.2. Results

Fig. 4 shows the weighted response rates for all subjects in the identification of tones as belonging to the (a) positive series or (b) negative series. The weighted response rates for each standard tone are shown by the solid lines with different symbols. The solid lines without symbols show the predicted values using the results of Experiment I based on fuzzy product set theory. The identification properties of tones for the combination of the slope of spectrum envelope and attack time form a peak near the standard tones and are the same as the results obtained when using only one clue. Comparing these measurements with the predicted values, the weighted response rates were found to generally correspond with the fuzzy membership values.

4. DISCUSSION

In the present study, the membership

function of the product set applied the definition by Zadeh (1965), but Goguen (1969) has also shown the possibility for different rules. According to Goguen's article, there are two possible rules regarding the truth values for the conjunction of the compound proposition (consistent with the product set). These two rules are expressed by

$t(A \cap B)=\min[t(A),t(B)]$

$t(A \cap B)=t(A) \cdot t(B)$,

where the truth value for a proposition A is t(A).

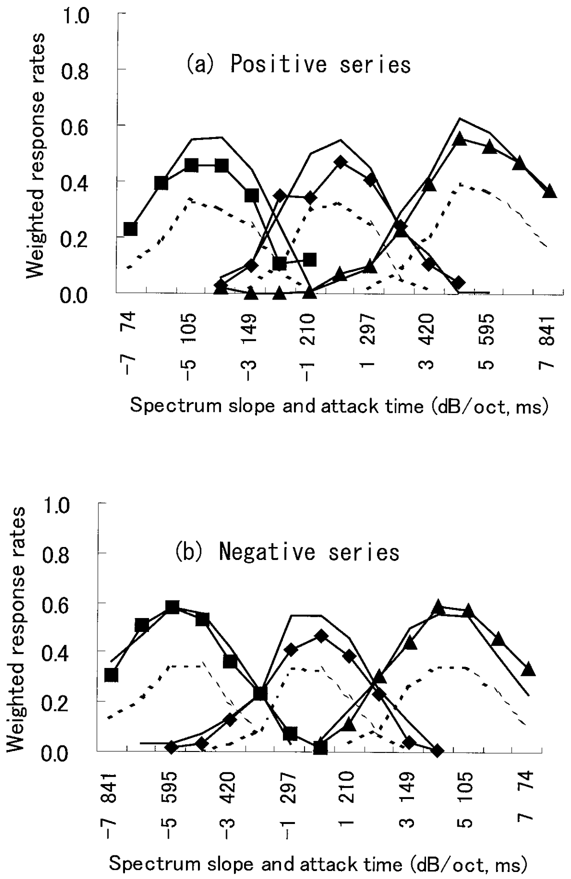

Fig. 4 Averaged weighted response rates in the case of combined spectrum and attack patterns in the (a) positive and (b) negative series. The solid lines without symbols represent the predictions by the minimum rule, and the dashed lines represent the predictions by the multiplying rule.

The former rule is called the minimum rule (Zadeh's definition) and the latter the multiplying rule. Regarding these rules, Yamashita et al. (1989) have shown through experiments in rules for conjunction that subjects judged the truth value of a combination of propositions, such as "a sparrow is a bird" and "a bat is a bird," consists of the combination of the perceived truth values of each single proposition. Moreover, they noted that (1) human judgment on truth value is not only affected by the minimum value in an element proposition, but also by the multiplying value, and also that (2) the truth value of a conjunction does not become less against a prediction of the multiplying rule and that the effect of a multiplying value has some limitations.

Based on the present data, the predicted response rates according to the multiplying rule are incorporated into Fig. 4 and shown as dashed lines. The Matching degree between the measurements and the predicted values was tested by calculating the root-mean-squared deviations between the measurements and the predicted values using the following equation.

$$A_g = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(M_i - P_i)^2}$$

In the above equation, M_i is the measurements, P_i is the predicted values using the minimum rule or the multiplying rule, and n is the number of values for each standard tone. The higher A_g means less matching degree of two curves. If the two curves correspond perfectly, however, then the value of A_g is zero.

Table 1 shows the matching degree between the measurements and the predicted values for the minimum rule and the multiplying rule from the positive and negative series. The measurements show better agreement with the predicted values using the minimum rule than those using the multiplying rule for all standard tones. The application effect of the multiplying rule is not contradicted, but it is not necessary to consider the multiplying rule in the present study.

The disagreement observed for the multiplying rule might be caused by the fact that the present study dealt with the identification process of sounds prescribed by relatively simple physical parameters, whereas Yamashita et al., for example, required highly advanced judgments of their subjects.

It can thus be suggested that the application of a fuzzy set model to the sound identification process using two acoustic clues is reasonable.

Table 1 Matching degree between the measurements and the predicted values for the minimum rule and the multiplying rule from the positive and negative series.

Positive	Minimum	Multiplying
(-4, 125)	0.059	0.223
(0, 250)	0.052	0.213
(+4, 500)	0.029	0.278

Negative	Minimum	Multiplying
(+4, 125)	0.027	0.131
(0, 250)	0.052	0.086
(-4, 500)	0.042	0.140

5. CONCLUSIONS

It can be suggested from the results of the present experiments that the application of a fuzzy set model to the identification of sounds is reasonable. In this model, clues to the spectrum and attack patterns of harmonic complex tones consist of five components. The present study shows the possibility of an application of fuzzy set theory to the psychoacoustics of real sounds such as musical instrument sounds or speech sounds.

REFERENCES

Goguen JA, The logic of inexact concepts, Synthese, 1969, 19, 325-373.

Iwamiya S, Kawanishi T and Satoh N, Identification of spectrum and attack pattern of sound -Application of the fuzzy set model for identification of sound-, The 1989 Spring Meeting of the Acoustical Society of Japan, 1989, 363-364 (in Japanese).

Matt GE, Turingan MR, Dinh QT, Felsch JA, Hovell MF and Gehrman C, Improving self-reports of drug-use: numeric estimates as fuzzy sets, Addiction, 2003, 98, 1239-1247.

Takayanagi S and Cliff N, An examination of graduate students' statistical judgments: statistical and fuzzy set approaches, Psychological reports, 2000, 81, 243-259.

Yamashita T and Yamashita K, Fuzzy logic for compound propositions, Japanese Journal of Psychology, 1989, 60, 312-315 (in Japanese).

Zadeh LA, Fuzzy Sets, Information and Control, 1965, 8, 338-353.