

# 音声信号におけるゆらぎ値等の特徴量と 感性的印象との相関性に関する研究

香川大学大学院工学研究科 博士後期課程信頼性情報システム工学専攻 <b>博 士 論 文</b>		
修了年度		平成26年度（2014年度）
指導教員		服部 哲郎
審査 担当 教員	主査	服部 哲郎
	副査	堀川 洋
		丹治 裕一

香川大学大学院 工学研究科  
博士後期課程 信頼性情報システム工学専攻

川上 裕介

平成26年6月26日

# Research on Correlation between Feature Quantity of Sound Signal and Kansei Impression

## Abstract

This research investigates the correlation between feature quantity of sound signal and human Kansei impression or affective impression. In this research, we regard the feature quantity as three kinds of values, that is, *fluctuation value*, *intercept*, and *sum of squared errors* that are obtained by making a regression analysis of sound signal in the Fourier domain. Also, we analyse the correlation not only for all frequency domain, but also for each of three divided frequency domains, that is, *low frequency band*, *middle frequency band*, *high frequency band* in the Fourier domain. Moreover, we examine as to how the human impression changes depending on the cases where the aforementioned feature quantity is artificially varied by sound processing. And we examine as to how the Kansei impression changes of the sound signal at the time of performing the HMGD processing. As its application, we present HMGD processing example to colour image.

## Keyword

Sound processing, fluctuation value, intercept, sum of squared errors, regression analysis, Kansei impression, HMGD

## あらまし

本研究では、音声信号のフーリエ領域における回帰分析によって算出される、ゆらぎ値、切片、残差二乗和の3つの特徴量等と音声信号の感性的印象との相関性を調査する。また本研究では、全周波数帯域、低域、中域、高域における各々の特徴量と音声の感性的印象との関係についても詳細に分析する。更には、特徴量を人為的に変更した場合に応じた感性的印象の変化についても調査分析を行う。また、HMGD処理を行った際の音声信号の感性的印象の変化も調査し、応用としてカラー画像へのHMGD処理例を示す。

## Keyword

音声処理, ゆらぎ値, 切片, 残差二乗和, 回帰分析, 感性的印象, HMGD

# 目次

1	序論	1
2	信号解析における数学的準備	4
2.1	フーリエ級数	4
2.1.1	複素フーリエ級数	5
2.2	フーリエ変換	6
2.2.1	離散フーリエ変換	6
2.3	回帰分析	8
2.4	ゆらぎとゆらぎ値随伴量の定義	10
2.5	相関係数	13
2.6	Ward 法	13
2.7	Wave ファイルの構造	15
3	音声信号における特徴量等の相関性調査	18
3.1	予備調査	18
3.1.1	ゆらぎ値随伴量算出・変更プログラム	18
3.1.2	ゆらぎ値随伴量間の相関関係	20
3.1.3	ゆらぎ値随伴量と音声の変化	23
3.2	本調査概要	28
3.3	ゆらぎ値随伴量と楽曲の感性的印象（調査 1）	29
3.4	ゆらぎ値随伴量の人為的変更と楽曲の感性的印象（調査 2）	31
4	調査結果と考察	36
4.1	ゆらぎ値随伴量と楽曲の感性的印象（調査 1）	36
I.	AF（全域）	39
II.	LF（低域）	40
III.	MF（中域）	41
IV.	HF（高域）	41
4.2	ゆらぎ値随伴量の人為的変更と楽曲の感性的印象（調査 2）	54

5	HMGD 処理の試験的適用	57
5.1	ヒストグラムマッチングの原理	57
5.2	音声信号への適用	60
5.3	カラー画像への適用	63
5.4	曲率によるヒストグラムピークの自動検出	78
6	結論	81
	謝辞	85
	参考文献	86
	業績	89

# 第1章

## 序論

近年、様々な領域の信号に対して、人間が心理的安らぎを覚える「癒し効果」を有するものとして、信号のパワースペクトルが周波数  $f$  に対して  $1/f$  の関係となるような  $1/f$  ゆらぎの存在が指摘され、研究されて来ている [1-6]。  $1/f$  ゆらぎは脳波における  $\alpha$  波と密接に関連しており、大脳の働きを抑制させリラックスした状態が得られることが知られている。

$1/f$  ゆらぎを有する物として、自然界では小川のせせらぎ、木々の葉がざわめく音、小鳥のさえずり等が挙げられる。これらの音声や映像は CD や DVD に収録した商品として市販されており、癒し効果がある事が謳われている。一方、音楽に着目してみれば、モーツァルトの音楽に同様の効果が表れる事が知られている。彼の音楽は主題（基本となる旋律）を繰り返し、アレンジするのみならず、直感的かつ不規則に別の主題を盛り込む等、規則性と不規則性を有している。このような規則性、及び不規則性は先述した自然界における音、或いは蝋燭の炎の揺らめきにおいても観測されており、  $1/f$  ゆらぎは規則性と不規則性が入り混じった状態であるとも考えられる。

しかしながら、従来の音声信号の  $1/f$  ゆらぎに関する研究では [7, 8]、ゆらぎ値のみに着目した解析・評価であり、同じ  $1/f$  ゆらぎを有しながらもパワースペクトル (PS) 分布が異なる場合の解析・評価までには至っていない。即ち、小川のせせらぎとピン

クノイズを聞き比べたとき，双方の音声信号おけるゆらぎ値は同じであるが，PS 分布の違いによって感性的印象が異なる場合がある．そのため，ゆらぎ値以外の要素も感性的印象に強く影響しているのではなかろうか，という疑問が生じる [9, 10]．

そこで本論文では，先ず，音声信号からのゆらぎ値を求める際に同時算出される随伴量を感性的特徴量として捉え，それらと感性的印象との対応関係を調査する．ゆらぎ値を含め同時算出される量としては，ゆらぎ値，切片，残差二乗和の 3 種類がある為，本論文ではこれらの特徴量をゆらぎ値随伴量と称する．ゆらぎ値随伴量の算出方法として，音声信号のフーリエ変換した後，周波数  $f$  の対数を横軸に取り，PS の対数を縦軸に取ったグラフ上において，回帰分析（直線回帰）の手法に基づいた方式を用いる [11, 12]．

音声信号と感性的印象との相関性分析については，アンケート評価による統計調査に基づいた分析を行う．本論文では，音声信号に対する感性的印象を詳細に分析するため，音声信号に対する被験者の感性的印象の傾向を Ward 法によってクラスター化し，グループ毎に音声信号のゆらぎ値算出随伴量と感性的印象との調査分析を行う．

一方，画像信号に対しても，感性的印象効果を高める自動補正処理の研究が広く行われている．その自動補正処理に関する研究は途上であるが，これまでの研究においてヒストグラムマッチング [22-28] を利用した方式が提案されている．

そこで，ガウス分布にヒストグラムをマッチングさせる，Histogram Matching based on Gaussian Distribution（以下，HMGD）処理 [13, 14, 20] が画像信号と音声信号の双方に有効ではないかと考え，試験的にその処理を適用して感性的印象効果を調査した．特に音声信号に対しては，その特徴量の変化も試験的に調査した．

本論文の構成を述べる．第 2 章では，フーリエ級数，フーリエ変換，回帰分析等の信号解析における数学的知識，及び Wave ファイルの構造について述べる．

第 3 章において，始めに事前調査について述べる．具体的には，本論文の調査で用いる 10 曲の楽曲のゆらぎ値算出随伴量を算出し，ゆらぎ値算出量の成分間における

---

相関性を調査する。また、ゆらぎ値随伴量を人為的に変更した際における PS の変化について述べる。続く本調査として、音声信号における特徴等との相関性についての調査概要及び方法を述べる。尚、本調査においては音声信号に対する感性的印象を詳細に分析する為、被験者の感性的印象をグループ化し、グループ毎にゆらぎ値算出随伴量と感性的印象の調査分析を行っている。

第 4 章では、ゆらぎ値随伴量と楽曲の感性的印象の関係、ゆらぎ値随伴量を変更した際の感性的印象の変化についての結果及び考察を述べる。

第 5 章では、画像信号に対する感性的印象効果を高める自動補正処理として提案されている、ヒストグラムマッチングの原理を説明する。我々はその中でもガウス分布にマッチングさせる HMGD (Histogram Matching based on Gaussian Distribution) 処理が画像信号及び音声信号の両面にて有効ではないかと考えており、試験的に適用した際の感性的印象効果を調査する。特に音声信号に対しては、その特徴量の変化についても試験的に調査している。

第 6 章の結論において、本研究において得られた成果をまとめ、今後の課題及び研究の応用展開について述べる。

## 第2章

# 信号解析における数学的準備

本論文では、フーリエ変換、回帰分析等の数学的手法及びクラスター分析等の統計的手法を用いて解析を行っている。本章では、これらの数学的概念、定理、理論を述べると共に、音声信号における特徴量であるゆらぎ値随伴量の定義を述べる。更に WAVE ファイルの構造についても述べる。

### 2.1. フーリエ級数

周期  $T$  の信号を  $f(t)$  と定義する。このとき、 $f(t)$  は (2.1) 式に示すフーリエ級数として示すことができる。

$$\begin{aligned} f(t) &= \frac{1}{2} a_0 + a_1 \cos \omega t + a_2 \cos 2\omega t + \dots + b_1 \sin \omega t + b_2 \sin 2\omega t + \dots \\ &= \frac{1}{2} a_0 + \sum_{n=1}^{\infty} (a_n \cos n\omega t + b_n \sin n\omega t) \end{aligned} \quad (2.1)$$

ここで、角周波数の正弦波成分を  $n$  次高調波という。また  $a_n, b_n$  をフーリエ係数と呼び、三角関数の直交性より (2.2) 式で示される。特に  $a_0/2$  は信号  $f(t)$  の直流成分である。



$$a_n = \frac{2}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(t) \cos n\omega t \, dt, \quad n = 0, 1, 2, \dots$$

$$b_n = \frac{2}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(t) \sin n\omega t \, dt, \quad n = 1, 2, \dots$$
(2.2)

### 2.1.1. 複素フーリエ級数

2.1 に示すフーリエ級数を複素数の領域へ拡張する．オイラーの公式を用いれば，余弦関数・正弦関数は，以下の(2.3)，(2.4)式で示される．

$$\cos n\omega t = \frac{1}{2} (e^{jn\omega t} + e^{-jn\omega t})$$
(2.3)

$$\sin n\omega t = \frac{1}{2j} (e^{jn\omega t} - e^{-jn\omega t})$$
(2.4)

(2.3)，(2.4)式を(2.1)式に代入すれば，

$$\begin{aligned} f(t) &= \frac{1}{2} a_0 + \sum_{n=1}^{\infty} (a_n \cos n\omega t + b_n \sin n\omega t) \\ &= \frac{a_0}{2} + \sum_{n=1}^{\infty} \left\{ \frac{1}{2} (e^{jn\omega t} + e^{-jn\omega t}) a_n + \frac{1}{2j} (e^{jn\omega t} - e^{-jn\omega t}) b_n \right\} \\ &= \frac{a_0}{2} + \sum_{n=1}^{\infty} \left\{ \left( \frac{a_n - jb_n}{2} \right) e^{jn\omega t} + \left( \frac{a_n + jb_n}{2} \right) e^{-jn\omega t} \right\} \end{aligned}$$

である．従って複素フーリエ級数は以下の(2.5)式で示される．

$$f(t) = \sum_{n=-\infty}^{\infty} c_n e^{jn\omega t}$$

$$\therefore c_0 = \frac{1}{2} a_0, \quad c_n = \frac{a_n - jb_n}{2}, \quad c_{-n} = \frac{a_n + jb_n}{2}$$
(2.5)

## 2.2. フーリエ変換

フーリエ変換とは、周期的・非周期的な時間領域の関数  $f(t)$  を、周波数領域の関数  $F(\omega)$  に変換することである。

始めに、周期が無限大（非周期的）である時間領域の関数を  $f(t)$  と定義し、周波数領域の関数  $F(\omega)$  を (2.6) 式にて定義する。

$$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t} dt \quad (2.6)$$

従って、 $f(t)$  は以下の (2.7) 式にて求められる。

$$f(t) = \int_{-\infty}^{\infty} F(\omega)e^{j\omega t} d\omega \quad (2.7)$$

(2.6)、(2.7) 式はフーリエ変換対と呼ばれている。また、(2.6) 式は時間領域の関数  $f(t)$  のフーリエ変換であり、(2.7) 式はその逆変換である。

$F(\omega)$  にオイラーの公式を適用すれば、(2.6) 式は以下の (2.8) 式としても示す事が出来る。

$$F(\omega) = |F(\omega)|e^{j\theta(\omega)} = |F(\omega)|\cos\theta(\omega) + j|F(\omega)|\sin\theta(\omega) \quad (2.8)$$

(2.8) 式において、 $|F(\omega)|$  を振幅スペクトル、 $\theta(\omega)$  を周波数スペクトルと呼び、振幅スペクトルの二乗  $|F(\omega)|^2$  をパワースペクトルと呼ぶ。

### 2.2.1. 離散フーリエ変換

2.2 節で述べたフーリエ変換は、関数  $f(t)$  が時間的に連続である場合において定義されている。しかし、実際のコンピュータシステム等における信号処理では、入力されたアナログ信号をサンプリングし、量子化を行ったうえで離散信号（デジタル信号）として取り扱っている。従って、離散値におけるフーリエ変換対を定義する必要があるが生じてくる。

$f(t)$  をサンプリング周期  $T$  でサンプリングしたときの信号を  $f_s(t)$  とすれば、これは

単位インパルス列  $\delta(t)$  との積によって得られるから,

$$f_s(t) = \sum_{n=-\infty}^{\infty} f(nT) \delta(t - nT) \quad (2.9)$$

$f_s(t)$  のフーリエ変換を  $F_s(\omega)$  と定義すれば, (2.9) 式より,

$$F_s(\omega) = \int_{-\infty}^{\infty} \left\{ \sum_{n=-\infty}^{\infty} f(nT) \delta(t - nT) \right\} e^{-j\omega t} dt = \sum_{n=-\infty}^{\infty} f(nT) \int_{-\infty}^{\infty} \delta(t - nT) e^{-j\omega t} dt$$

である. 単位インパルス関数の定義を利用すれば,  $F_s(\omega)$  は以下の (2.10) 式のように変形される.

$$F_s(\omega) = \sum_{n=-\infty}^{\infty} f(nT) e^{-jn\omega T} \quad (2.10)$$

更に  $f(nT) = f(n)$  と置くと, 離散フーリエ変換は以下の (2.11) 式で定義される.

$$F_s(\omega) = \sum_{n=-\infty}^{\infty} f(n) e^{-jn\omega T} \quad (2.11)$$

(2.6) 式より,  $F_s(\omega)$  の逆離散フーリエ変換は, 以下の (2.12) 式で求めることができる.

$$f(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} F_s(\omega) e^{jn\omega T} d(\omega T) \quad (2.12)$$

以上, (2.11), (2.12) 式を離散フーリエ変換対という.

### 2.3. 回帰分析

回帰分析とは、近似直線との残差 $e_i$ の二乗和 $s$ が最小となる回帰直線を導出することである。 $s$ の定義を(2.13)式に示し、図2.1に回帰直線の例を示す。(2.13)式において、 $y$ は目的変数、 $Y$ は回帰推定値である。

$$s = \sum_i e_i^2 = \sum_i (y_i - Y_i)^2 \quad (2.13)$$

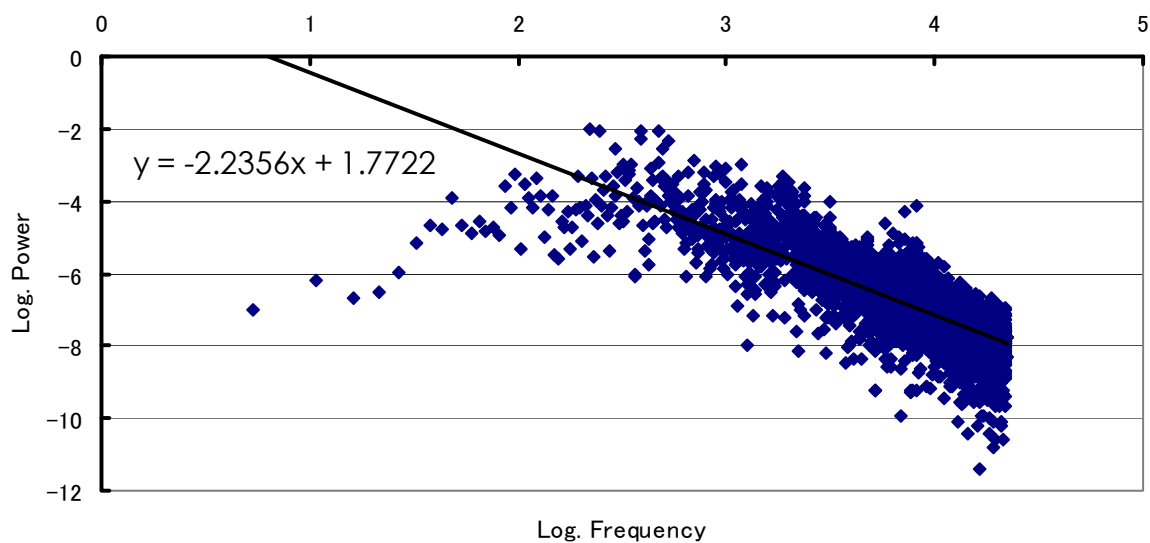


図 2.1 回帰直線の例

一般的には、(2.14)式に示す $Y$ の $n$ 次元線形回帰モデルについて、 $s$ が最小となる回帰係数 $a$ を求めることにより行う。即ち、(2.15)式が満たされればよいことになる。

$$Y = a_1 x_1 + \cdots + a_m x_m + a_0 \quad (2.14)$$

$$\frac{\partial s}{\partial a_i} = 0 \quad (2.15)$$

(2.15)を計算すれば、

$$\begin{aligned}\frac{\partial s}{\partial a_0} &= -2 \sum_i (y_i - Y_i) = 0 \quad \Rightarrow \quad \sum_i y_i = \sum_i Y_i \\ \frac{\partial s}{\partial a_1} &= -2x_1 \sum_i (y_i - Y_i) = 0 \\ \frac{\partial s}{\partial a_j} &= -2x_j \sum_i (y_i - Y_i) = 0 \quad \Rightarrow \quad \sum_i x_j y_i = \sum_i x_j Y_i \quad \because j=1, \dots, m\end{aligned}$$

ここで、 $x_0 = 1$  と定義すると、

$$\sum_i y_i = \sum_i Y_i \quad \Rightarrow \quad \sum_i x_0 y_i = \sum_i x_0 Y_i \quad \sum_i x_j y_i = \sum_i x_j Y_i \quad (j=1, \dots, m)$$

即ち、

$$\begin{aligned}\sum_i x_j y_i &= \sum_i x_j Y_i \quad (j=0, 1, \dots, m) \\ \sum_{i=1}^n x_j y_i &= \sum_{i=1}^n x_j (a_1 x_1 + a_2 x_2 + \dots + a_m x_m + a_0 x_0)\end{aligned} \tag{2.16}$$

ここで、 $n$  はデータ数であり、 $m$  は次元数を表す。

続いて、 $\vec{A}, \vec{X}, \vec{Y}$  を以下の(2.17)式で定義する。

$$\vec{A} = \begin{Bmatrix} a_1 \\ \vdots \\ a_m \\ a_0 \end{Bmatrix} \quad \vec{X} = \begin{Bmatrix} x_{11} & x_{12} & \dots & x_{m1} & x_{01} \\ \vdots & \vdots & & \vdots & \vdots \\ x_{1n} & x_{2n} & \dots & x_{mn} & x_{0n} \end{Bmatrix} \quad \vec{Y} = \begin{Bmatrix} y_1 \\ \vdots \\ y_n \end{Bmatrix} \tag{2.17}$$

(2.16)式に(2.17)式を適用すれば、

$$\vec{X}^t \vec{Y} = \vec{X}^t \vec{X} \vec{A} \quad \therefore \vec{A} = \left( \vec{X}^t \vec{X} \right)^{-1} \vec{X}^t \vec{Y} \tag{2.18}$$

以上から、回帰係数 $a_i$ を求めることができる。

本論文で扱う音声信号は一次元であるから、 $m=1$ である。従って(2.17)式は、(2.19)式のように表す事が出来る。

$$\vec{A} = \begin{Bmatrix} a_1 \\ a_0 \end{Bmatrix} \quad \vec{X} = \begin{Bmatrix} x_{11} & x_{01} \\ \vdots & \vdots \\ x_{1n} & x_{0n} \end{Bmatrix} \quad \vec{Y} = \begin{Bmatrix} y_1 \\ \vdots \\ y_n \end{Bmatrix} \quad (2.19)$$

## 2.4. ゆらぎとゆらぎ値算出随伴量の定義

ゆらぎの内、よく知られた  $1/f$  ゆらぎとは、パワースペクトル（以下、PS）分布が周波数と逆比例関係にあり、周波数  $f$  が高くなるにつれて PS が  $1/f$  の割合で小さくなることを意味しており、人間が心地よいと感じる効果を持つことが指摘されている。その理由としては、人間の生体リズムが  $1/f$  ゆらぎを持つが故、ある事象に  $1/f$  ゆらぎが含まれていれば、人間はリズムの合致を感じ、その事象に快感を覚えるためとされている [11, 15]。また、自然界にある現象の大半が  $1/f$  ゆらぎ傾向を示す為であるとも指摘されている。

図 2.2 に PS の概念図を示す。また、図中の PS 曲線を (2.20) 式にて定義する。尚、 $Y(f)$ 、 $\alpha(f)$  はそれぞれ PS の強度及び誤差を示す。

$$Y(f) = \frac{k}{f^a} + \varepsilon(f) \quad (2.20)$$

両辺の対数を取れば、

$$\begin{aligned} \log Y(f) &= \log \left\{ \frac{k}{f^a} + \varepsilon(f) \right\} = \log \left\{ \frac{k}{f^a} (1 + \tilde{\varepsilon}(f)) \right\} \\ &= \log \left( \frac{k}{f^a} \right) + \log \{ 1 + \tilde{\varepsilon}(f) \} \end{aligned} \quad (2.21)$$

従って、(2.21) 式は以下の (2.22) 式で示される。

$$\log Y(f) = -a \log f + \log k + \log \{ 1 + \tilde{\varepsilon}(f) \} \quad (2.22)$$

更に、 $y(f) = \log Y(f)$ 、 $b = \log k$ 、 $\hat{\varepsilon} = \log \{ 1 + \tilde{\varepsilon}(f) \}$  と定義すれば、(2.22) 式は (2.23) 式として表す事が出来る。

$$\begin{aligned} y(f) &= -a \log f + \log k + \log\{1 + \tilde{\varepsilon}(f)\} \\ &= -a \log f + b + \hat{\varepsilon}(f) \end{aligned} \quad (2.23)$$

よって、図 2.2 の PS 曲線は図 2.3 として表される。図 2.3 は、横軸に周波数  $f$  の対数を取り、縦軸に PS の対数  $y(f)$  を取った音声信号の PS の例である [15]。

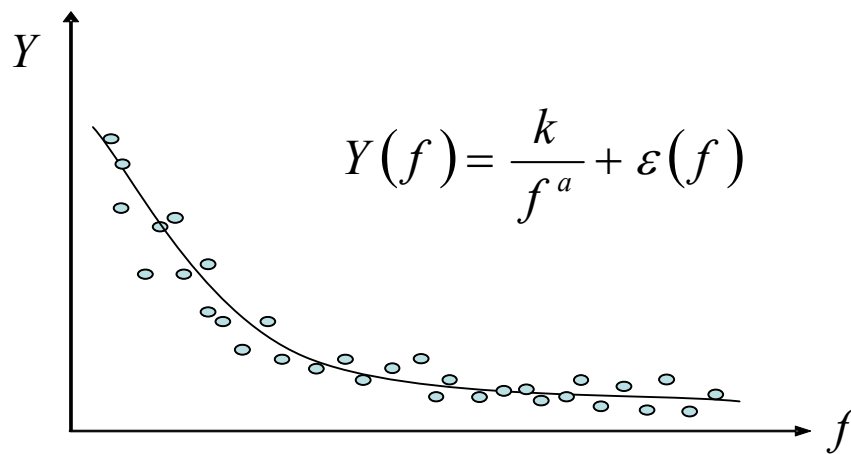


図 2.2 PS 曲線の例

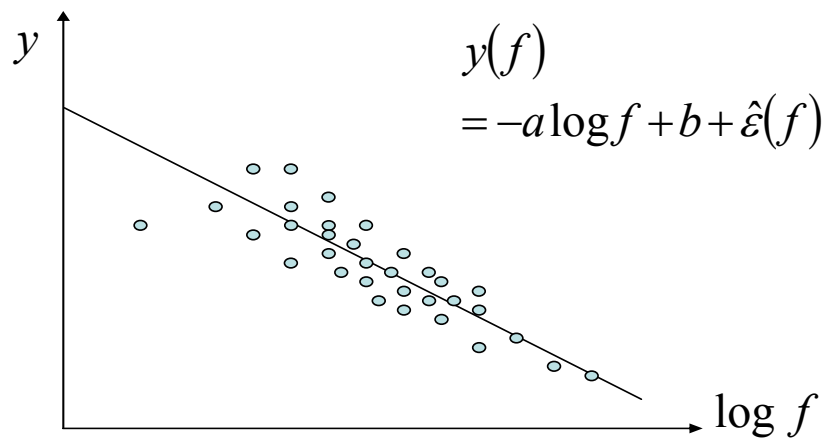


図 2.3 PS の回帰直線の例 [15]

本論文では、傾斜の絶対値  $a$  を『ゆらぎ値』と定義し、 $y$  切片  $b$  を『切片』と定義する。更に、PS の実測値  $y_i$  と回帰直線上の理論値  $Y_i$  の残差  $\varepsilon_i$  の平方和  $s$  を『残差二乗和』と定義する。残差二乗和の定義式は(2.13)式に示す通りである。

また、残差のスケーリングを行う事によって切片と残差二乗和を変更する事が出来る。スケーリング設定値を  $A$  及び  $B$  とすれば、スケーリング後の残差  $e_i'$  は(2.23)式で求められる。

$$e_i' = Ae_i + B \quad (2.23)$$

上式より、 $A$  の変更によって回帰直線に対する残差の振幅を調整し、 $B$  の変更によって切片を調整出来る事が分かる。図 2.4 にゆらぎ値算出随伴量を変更した際の、PS の変化例を示す。

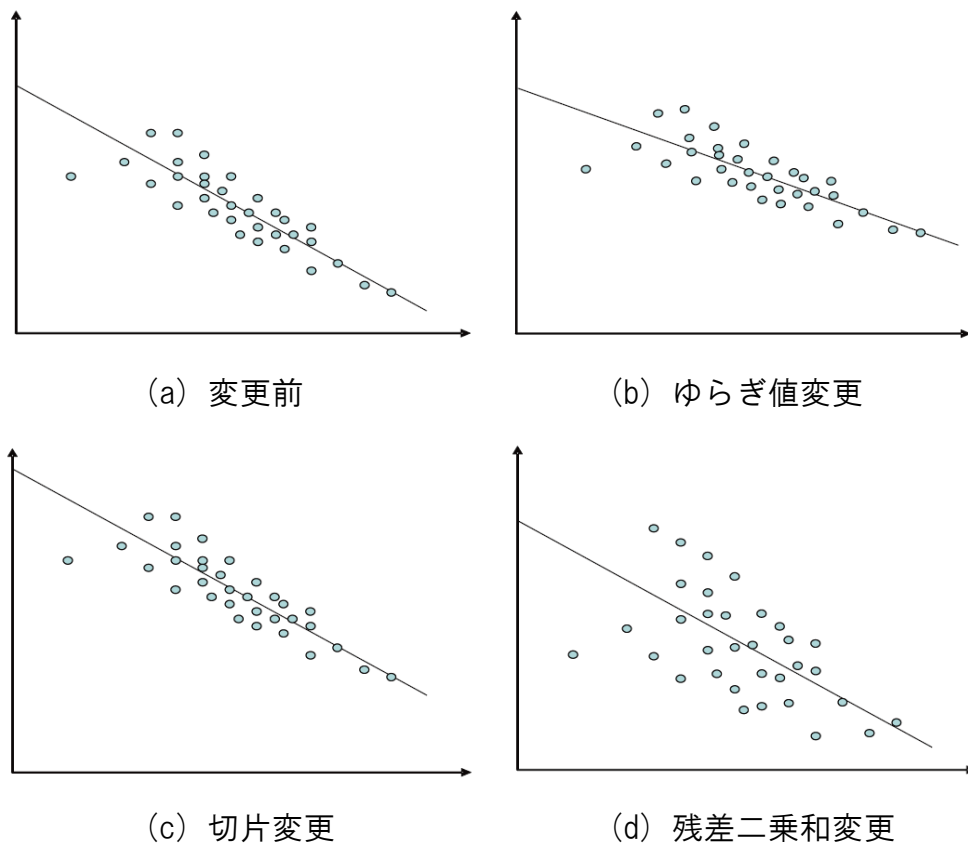


図 2.4 PS の変化例[15]



## 2.5. 相関係数

相関係数とは、2組の確率変数の中の相関（類似性の度合い）を示す統計学的指標である。相関係数の取り得る範囲は $-1$ 以上 $+1$ 以下の実数である。

2組のデータ系列 $(x_i, y_i)$  ( $i=1, 2, \dots, n$ )を定義すると、相関係数 $C$ は以下の(2.24)式で定義される。

$$C = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (2.24)$$

但し $\bar{x}$ ,  $\bar{y}$ はデータ $x_i$ ,  $y_i$ の相加平均である。また(2.24)式の分子はデータ $x_i$ ,  $y_i$ の共分散であり、分母はデータ $x_i$ の標準偏差と $y_i$ の標準偏差の積である。

一般的に $C \geq |0.8|$ であるとき、2つのデータ系列の間には強い相関があると言われている。尚、 $C > 0$ であるときは正の相関があり、 $C < 0$ であるときは負の相関がある。 $C = 0$ であるときは無相関である。本論文では、ゆらぎ値、切片、残差二乗和の各パラメータと、その感性的印象の関係の度合いを示す指標として相関係数を用いる。

## 2.6. Ward 法

本論文ではWard法を用いたクラスター分析を行っている。

Ward法は、クラスター $A$ と $B$ の距離を、それらを融合した時のクラスター内変動の増加分を(2.25)式にて定義し、距離の小さなクラスターから融合している方法である。但し、 $d(x, y)$ はユークリッド距離、 $\mu_{AB}$ はクラスター $A$ と $B$ を融合したクラスターの平均ベクトル、 $\mu_A$ と $\mu_B$ はクラスター $A$ と $B$ それぞれの平均ベクトルである。 $S$ は平均からの変動である。

$$D(A, B) = \sum_{x \in A, B} d(x, \mu_{AB})^2 - \left( \sum_{x \in A} d(x, \mu_A)^2 + \sum_{x \in B} d(x, \mu_B)^2 \right) = S_{AB} - (S_A + S_B) \quad (2.25)$$

Ward 法による融合過程の概念図を図 2.5 に示す. 3つのクラスター  $C_i, C_j, C_k$  がそれぞれ平均ベクトル  $\mu_i, \mu_j, \mu_k$  を中心にクラスターを形成しているものとする. 尚, クラスター  $C_i$  と  $C_k$  の距離は, 二つのクラスターに属するデータ全体の平均  $\mu_{ik}$  を用い, (2.25) 式にて求めたクラスター内変動の増加分である. また,  $C_j$  と  $C_k$  の距離は, 平均  $\mu_{jk}$  を用いて計算したクラスター内変動の増加分である. 図 2.3 では,  $C_j$  と  $C_k$  の距離の方が小さいので, これらが融合される.

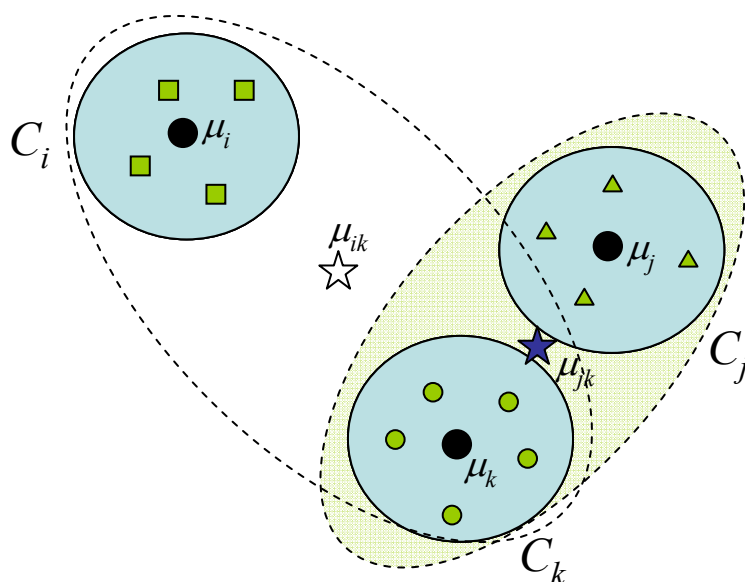


図 2.5 Ward 法によるクラスター融合過程の概念図[16]

クラスター  $A$  と  $B$  が融合されて  $C$  となった時, ほかのクラスター  $X$  との距離は (2.26) 式によって示される.

$$D(C, X) = \frac{N_A + N_X}{N_A + N_B + N_X} D(A, X) + \frac{N_B + N_X}{N_A + N_B + N_X} D(B, X) - \frac{N_X}{N_A + N_B + N_X} D(A, B) \quad (2.26)$$

本論文では、三次元随伴パラメータ空間中のデータ間の距離はユークリッド距離において定義されるものとする。従って三次元のベクトル群においてクラスターが分類されることになる。しかしながら、視覚的に見易くする為ゆらぎ値算出随伴量の内の2種を軸とする二次元平面に投射している。

## 2.7. Wave ファイルの構造

Wave ファイルはマルチメディア用ファイル形式で用いられる、リソース変換ファイル形式 (RIFF ; Resource Interchange File Format) の構造によってデータを格納している。RIFF フォーマットのファイルではチャンク (Chunk ; 塊) と呼ばれる、纏まった単位でデータを表現する。図 2.6 に Wave ファイルの構造を示す。

RIFFヘッダ	R' I' F' F'	4byte
	ファイルサイズ(byte)	4byte
WAVEヘッダ	W' A' V' E'	4byte
fmtヘッダ	f' m' t'	4byte
	fmtチャンクサイズ(byte)	4byte
	フォーマットID	2byte
	チャンネル数	2byte
	サンプリング周波数(Hz)	4byte
	データ速度(byte/sec)	4byte
	ブロックサイズ(byte/sample×ch)	2byte
dataチャンク	ビット数	2byte
	d' a' t' a'	4byte
	波形データサイズ(byte)	4byte
	波形データ	nbyte

図 2.6 Wave ファイルの構造

Wave ファイルではファイルフォーマットの定義やサンプリング周波数等を宣言するため、専用のチャンクが付加されている。専用のチャンクとしては、fmt チャンク、fact チャンク、data チャンク、LIST チャンクの4種類が定義されている。

fmt チャンクは Wave ファイルの先頭に位置し、ファイルサイズ、フォーマット ID 等を示すチャンクである。チャンクの先頭には文字列 'f' 'm' 't' ' ' (スペース) が記されている。『フォーマット ID』は Wave ファイルに記録されているデータフォーマットの種類を数字にて表記したものである。代表的なフォーマット ID を表 2 に示す。

表 2 フォーマット ID 一覧 (抜粋) [17]

ID	フォーマット
0x0000	不明
0x0001	PCM
0x0002	MS ADPCM
0x0005	IBM CSVD
0x0006	A-Law
0x0007	$\mu$ -Law
0x0011	IMA/DVI ADPCM
0x0014	ADPCM (G.723)
0x0020	YAMAHA ADPCM
0x0022	TrueSpeech
0x0030	AC2 (Dolby)
0x0031	GSM 6.10
0x0033	ANTEX ADPCM
0x0034	VQLPC (Control Resources)
0x0035	DIGIREAL
0x0036	DIGIADPCM
0x0037	CR10 (Control Resources)
0x0040	ADPCM (G.721)

『チャンネル数』は、モノラルであれば 1、ステレオであれば 2 である。『データ速度』は (サンプリング周波数)  $\times$  (サンプルあたりのバイト数)  $\times$  (チャンネル数) で表される。本論文では 44.1KHz, 16bit(2bit) のモノラル音声を扱うため、88200[byte/sec]となる。尚、PCM フォーマット以外の場合は、図 2.4 の他に拡張部分が付加されることがある。

fact チャンクは全サンプル数を表すチャンクであり、波形データの前に位置する

---

が必須ではない為、本論文ではこれを用いていない。

data チャンクは波形のサイズ及び波形データを示すチャンクである。『波形データサイズ』に波形データのサイズを記し、以降は波形データが記録されている。

LIST チャンクは著作権情報等を格納したチャンクであるが、fact チャンクと同様に必須でない為、本論文では取り扱わない。

## 第3章

# 音声信号における特徴量等の相関性調査

### 3.1. 予備調査

#### 3.1.1. ゆらぎ値随伴量算出・変更プログラム

本論文の調査実験で用いた、ゆらぎ値随伴量算出及び変更を行うプログラム、『SoundProc.exe』の処理過程を図3.1に示す。

SoundProc.exeは、コマンドライン引数において、処理対象のWaveファイル名、ゆらぎ値随伴量の設定値、パワースペクトル(PS)の保存の有無及び随伴パラメータを求める帯域の分割数を入力し、実行する。引数の設定内容を表3.1に示す。

表3.1 SoundProc.exeのコマンドライン引数

書式: SoundProc.exe [arg1] [arg2] [arg3] [arg4] [arg5] [arg6]

引数名	設定内容
[arg1]	処理対象となるWaveファイル名(拡張子は入力不要)
[arg2]	ゆらぎ値の設定値(noneを入力すると変更しない)
[arg3]	切片の設定値(noneを入力すると変更しない)
[arg4]	残差二乗和の設定倍率(noneを入力すると変更しない)
[arg5]	PSの保存の有無(y/n)
[arg6]	随伴パラメータを求める帯域数

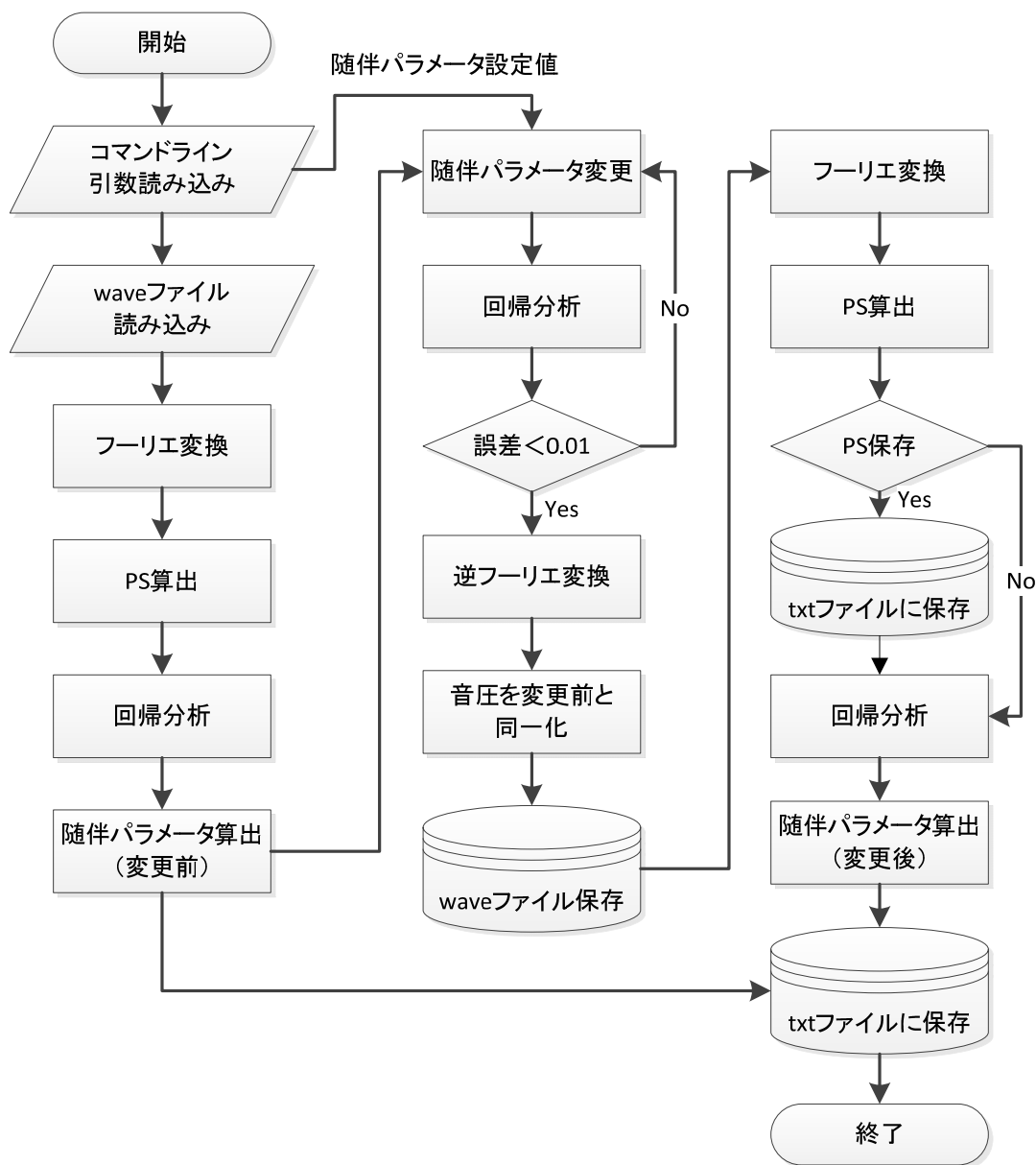


図 3.1 SoundProc.exe の処理過程

例えば、以下のようにコマンドプロンプトに入力すれば、『test.wav』のゆらぎ値を1に変更してPSを保存する。更に、LF(低域:0~300Hz)・MF(中域:300~1000Hz)・HF(高域:1000~22050Hz)の3つの帯域におけるゆらぎ値随伴量が算出される。

SoundProc.exe test 1 none none y 3

また、ゆらぎ値、切片、残差二乗和の何れかを変更した際、変更前の Wave ファイルの音圧と変更後の Wave ファイルの音圧が変化することが考えられる。そこで、本プログラムでは変更前と変更後の音圧を計測して倍率を求め、この倍率を変更結果に乗算し、両者の音圧を同一にしている。なお、本プログラムは C 言語 (Borland C++ Compiler 5.5.1) で作成し、許容誤差を  $10^{-6}$  以下としている。

### 3.1.2. ゆらぎ値随伴量間の相関性

表 3.2 に示す楽曲に対してゆらぎ値、切片、残差二乗和を求め、これらのゆらぎ値随伴量間の相関性を調査した。本論文で用いる楽曲はサンプリング周波数 44.1kHz、ビット数 16bit の Wave ファイルである [11]。結果を図 4.1～図 4.3 に示す。

表 3.2 Wave ファイルの一覧 [11]

番号	ファイル名	ジャンル
1	Another_Sky.wav	イーजीリスニング
2	ロンドンデリーエア.wav	クラシック
3	君を信じて.wav	イーजीリスニング
4	漂流.wav	イーजीリスニング
5	Down_by_the_Riverside.wav	ジャズ
6	Space_Odessey3_Revelation.wav	イーजीリスニング
7	TOMORROW.wav	ポップ
8	古いフランスの歌.wav	クラシック
9	自由.wav	ポップ
10	赤い川の谷間(brass).wav	ジャズ

ゆらぎ値と切片の関係を示した散布図 (図 3.2) より、ゆらぎ値が小さい楽曲では切片が低くなり、ゆらぎ値が大きい楽曲は切片が高くなっていることが分かる。このとき、相関係数は 0.991 と強い相関を示したことから、ゆらぎ値と切片は相関関係が存在すると考えられる。

切片と残差二乗和の関係を示した散布図 (図 3.3) より、切片が小さい楽曲ほど残



差二乗和が小さく、切片が大きい楽曲ほど残差二乗和が大きくなる傾向が見られる。このとき、相関係数は0.852と強い相関を示したことから、切片と残差二乗和は相関関係が存在すると考えられる。

ゆらぎ値と残差二乗和の関係を示した散布図（図3.4）より、残差二乗和が同等の値を示している部分に注目すれば、残差二乗和が同等にも係わらず、ゆらぎ値が大きく異なっている楽曲が見受けられる。相関係数は0.786であり、ゆらぎ値と残差二乗和にも相関が見られる。

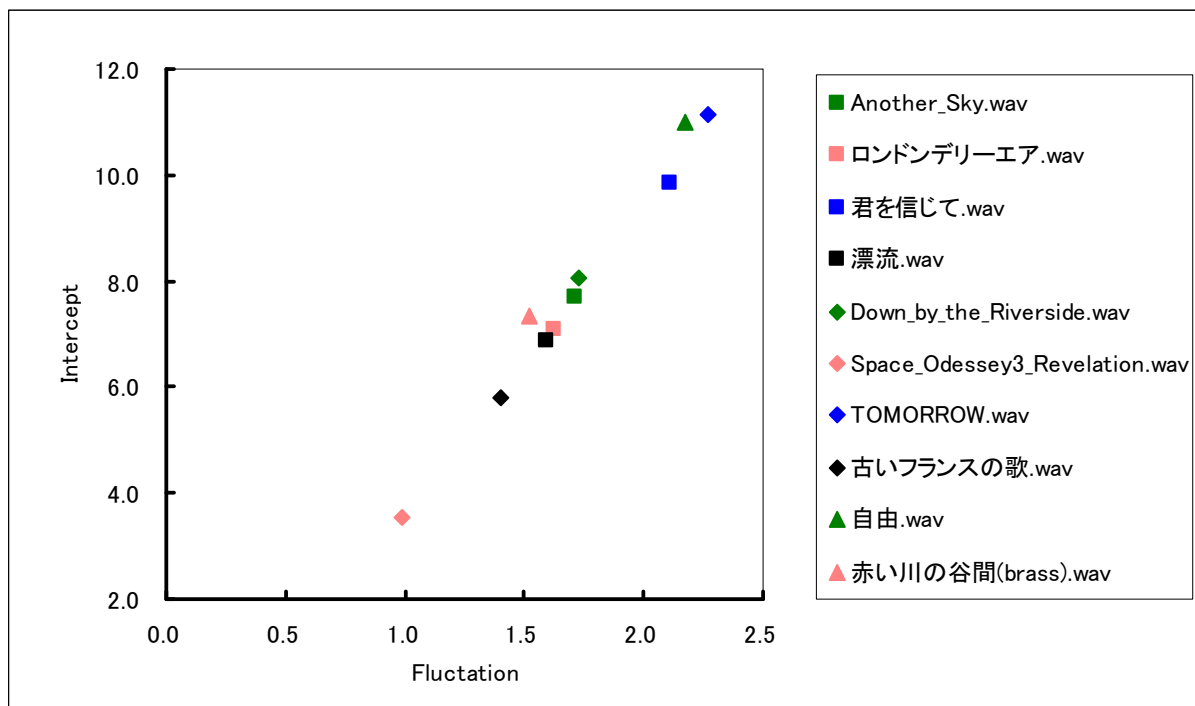


図3.2 ゆらぎ値随伴量の相関性（ゆらぎ値と切片）

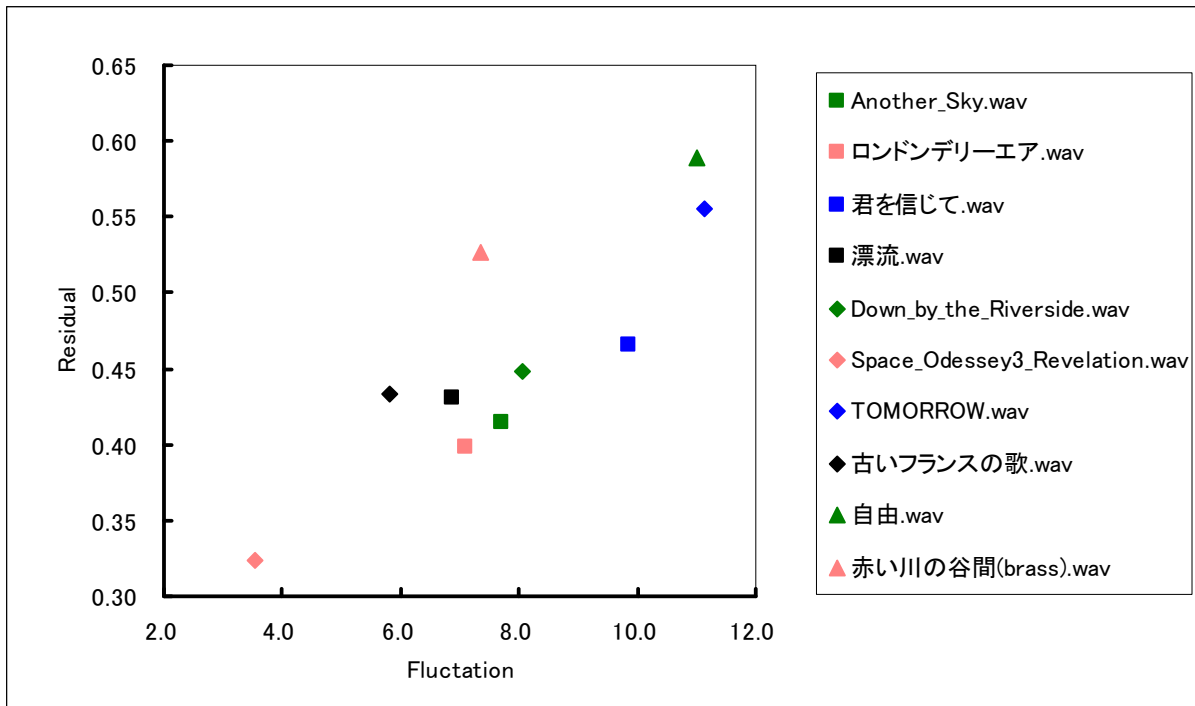


図 3.3 ゆらぎ値随伴量の相関性 (切片と残差二乗和)

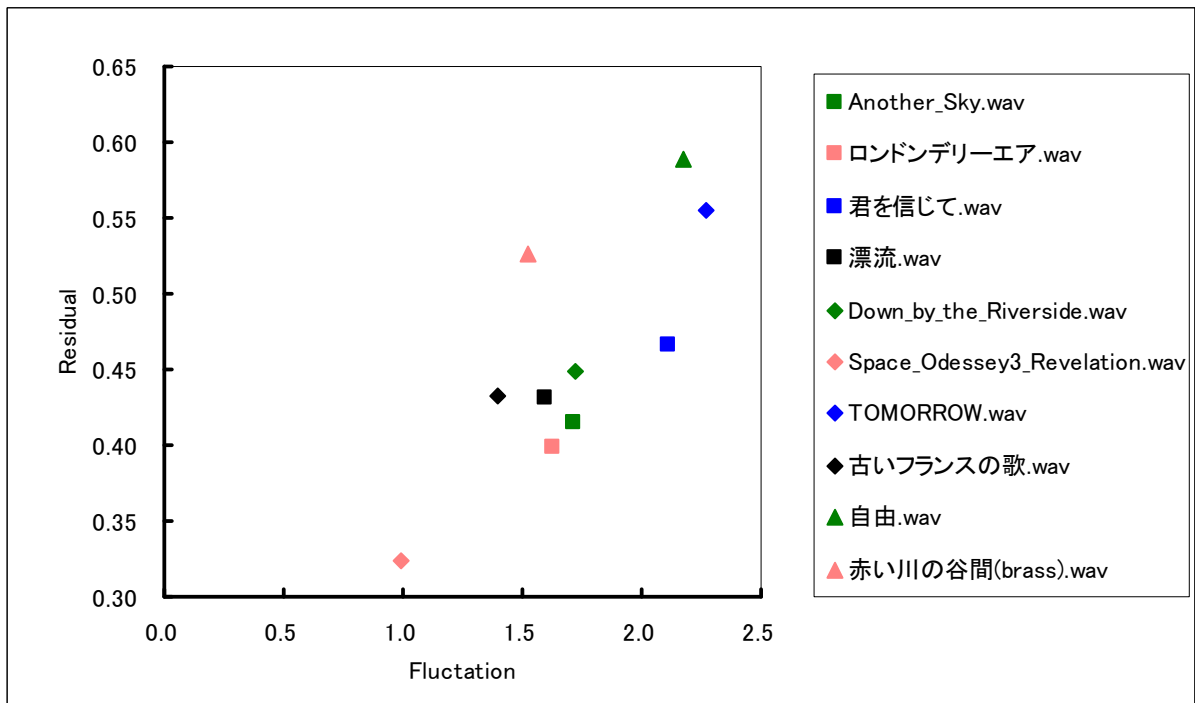


図 3.4 ゆらぎ値随伴量の相関性 (ゆらぎ値と残差二乗和)

### 3.1.3. ゆらぎ値随伴量と音声の変化

続いて、楽曲のゆらぎ値及び残差二乗和を人為的に変更した際の音声の変化を調査した。尚、切片は原曲の収録音量と対応している為[15]、変更前と変更後の音圧が一定となるように切片を調整している。調査に用いた楽曲は表 3.2 に示す Another\_Sky である。

図 3.5～3.7 に、ゆらぎ値を 1, 2, 3, 更に±10%に変化させた結果を示す。同様に、残差の倍率を 0.4 倍, 0.8 倍, 1.2 倍に変化させた結果を図 3.8～3.9 に示す。表 3.3, 3.4 に 3 種の随伴量の実測値を示す。

表 3.3 ゆらぎ値随伴量の実測値（ゆらぎ値の変更）

Title	Fluctation setting	Fluctation	Intercept	Residual
Another_Sky	Default	1.715343	7.704218	0.414398
	1	1.000000	3.681235	0.414398
	2	2.000000	8.623341	0.414398
	3	3.000000	10.976942	0.414398
	-10%	1.546430	6.676341	0.414398
	+10%	1.890081	8.293712	0.414398

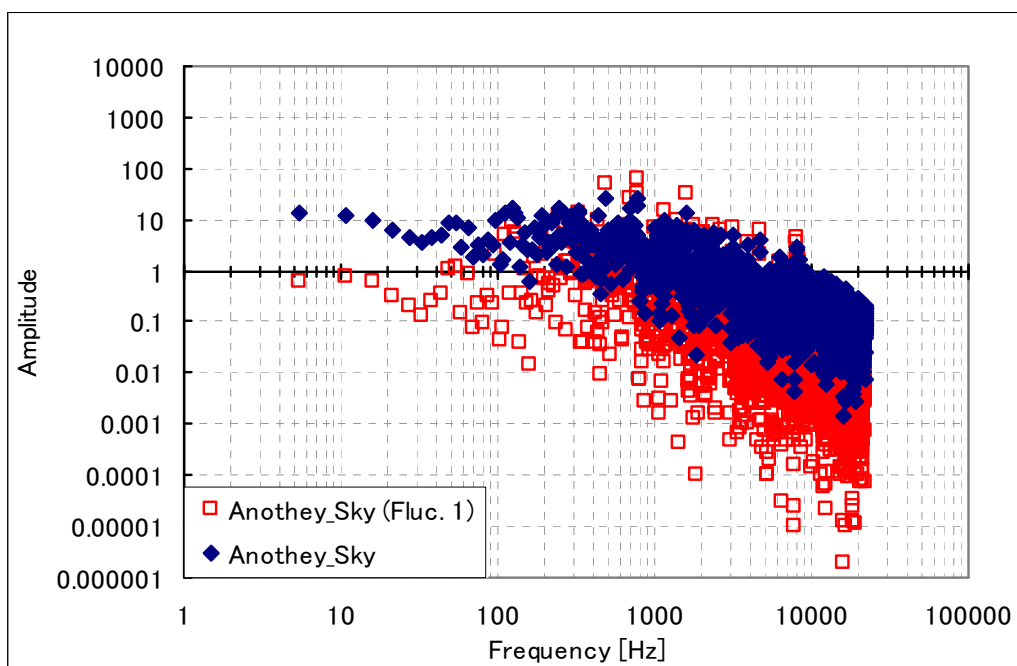
表 3.4 ゆらぎ値随伴量の実測値（残差の変更）

Title	Residual setting	Fluctation	Intercept	Residual
Another_Sky	Default	1.715343	7.704218	0.414398
	×0.4	1.715343	7.315150	0.165759
	×0.8	1.715343	7.625076	0.331318
	×1.2	1.715343	7.557243	0.497278

図 3.5 及び図 3.6 より、ゆらぎ値が 3 に近づくにつれて高周波成分の振幅の減衰が目立つことが分かる。つまり、高周波成分の音量は原曲に比べて低下したと考えられる。更に図 3.7 より、ゆらぎ値を+10%変化させると原曲より高周波成分がやや減衰していることが分かる。また、ゆらぎ値を-10%変化させると、高周波成分がやや増加することが分かる。

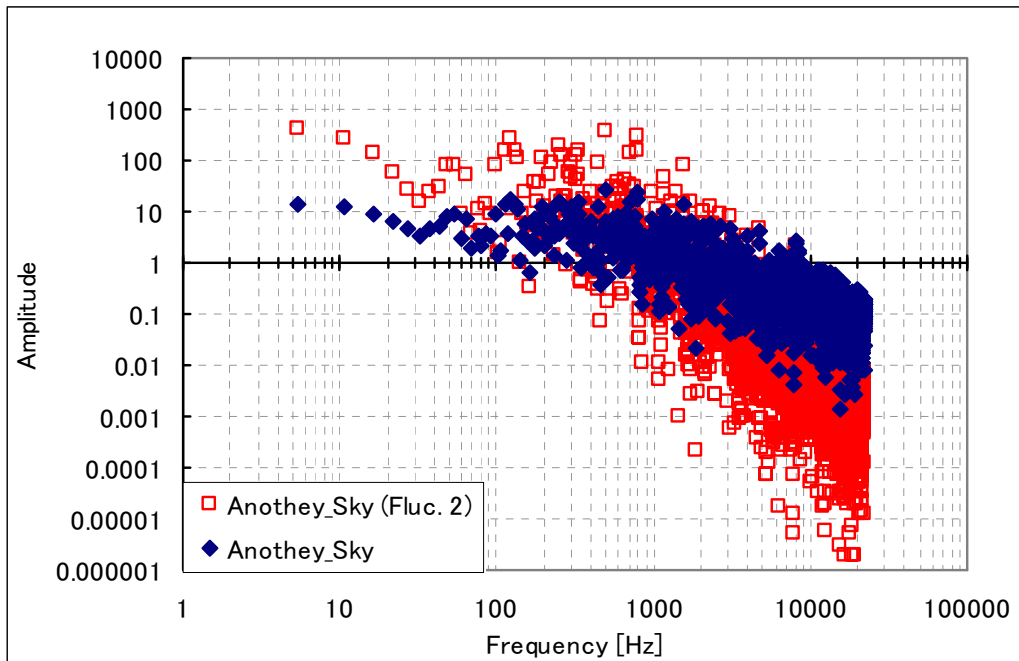
図 3.8 及び図 3.9 より，残差を変化させた場合は，残差に乗算する値を増加させるほど残差が増加していることが分かる．

以上の結果から，音声信号における『ゆらぎ値』の変更は，低周波成分及び高周波成分を多く増減させる事が可能と考えられる．しかしながら，ゆらぎ値の変更割合が大きすぎれば，原曲の持つ感性的印象とは大きく異なってしまうと推定される．また，『残差二乗和』は，近似直線への当てはまりを示す値である．残差を変更した楽曲を聴いたところ，高周波成分が強調されている印象を受けた．この傾向は表 3.2 に示す全ての楽曲で同様であった．

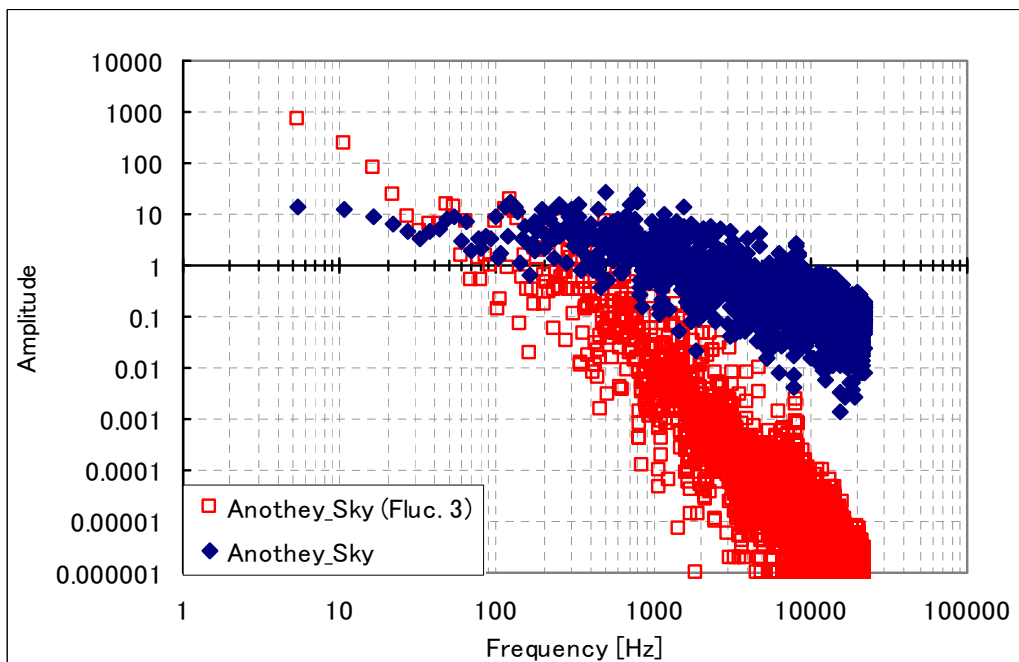


ゆらぎ値 1 (実測値：表 3.3(a))

図 3.5 ゆらぎ値の変更

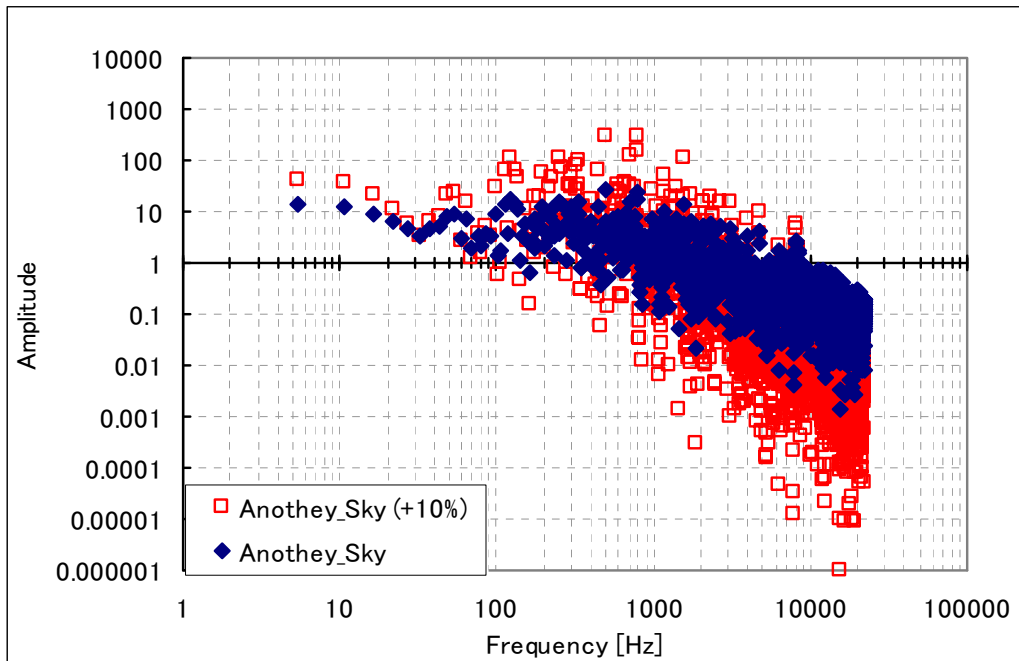


(a) ゆらぎ値 2 (実測値：表 3.3(a))

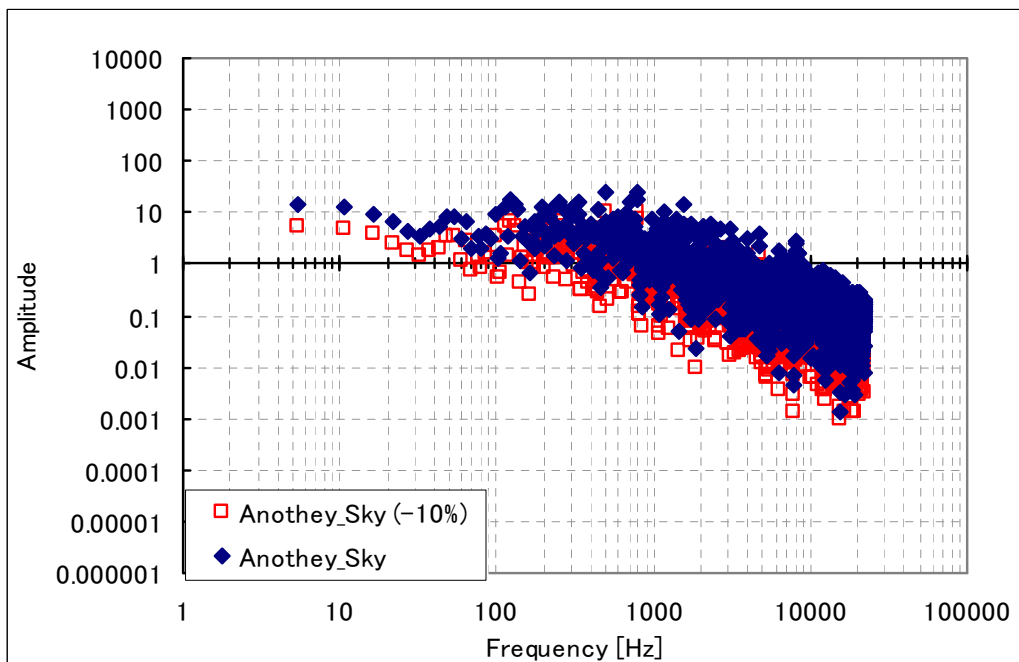


(b) ゆらぎ値 3 (実測値：表 3.3(a))

図 3.6 ゆらぎ値の変更

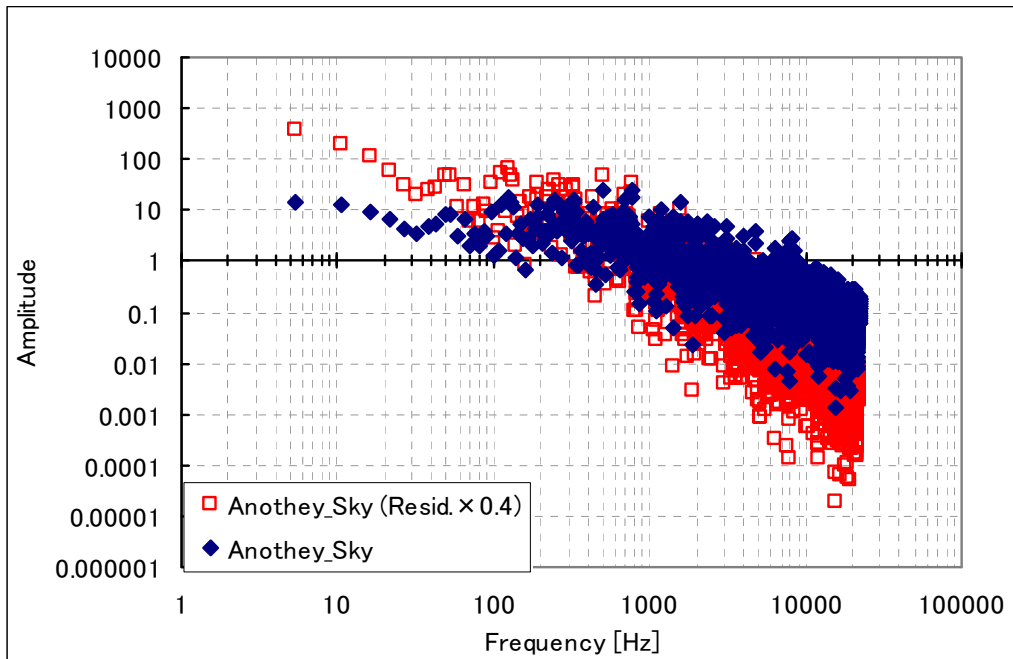


(a) ゆらぎ値+10% (実測値：表 3.3(a))

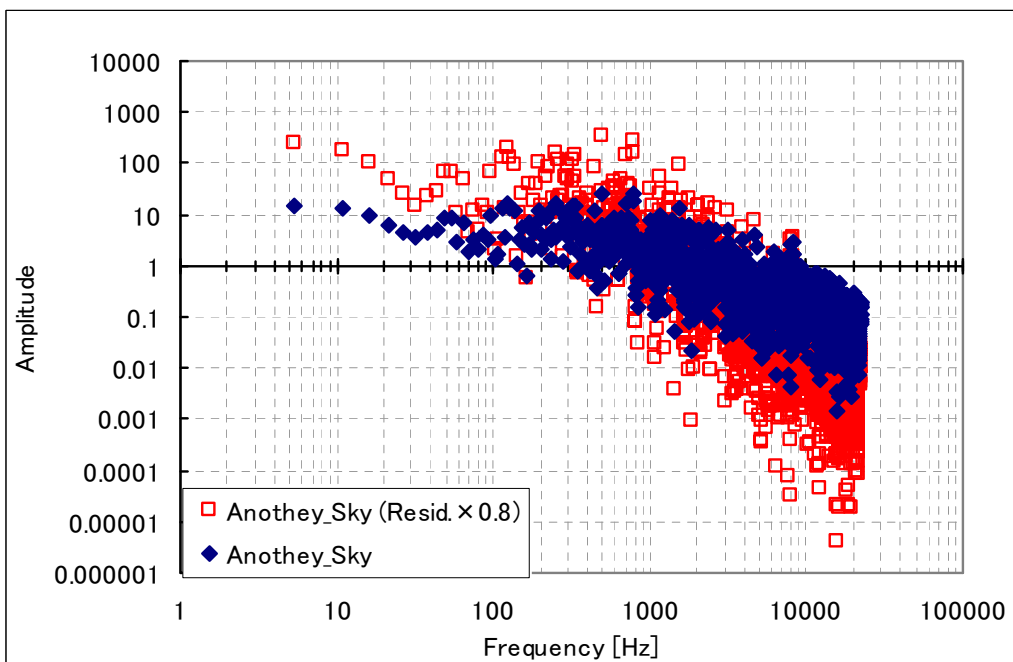


(b) ゆらぎ値-10% (実測値：表 3.3(a))

図 3.7 ゆらぎ値の変更

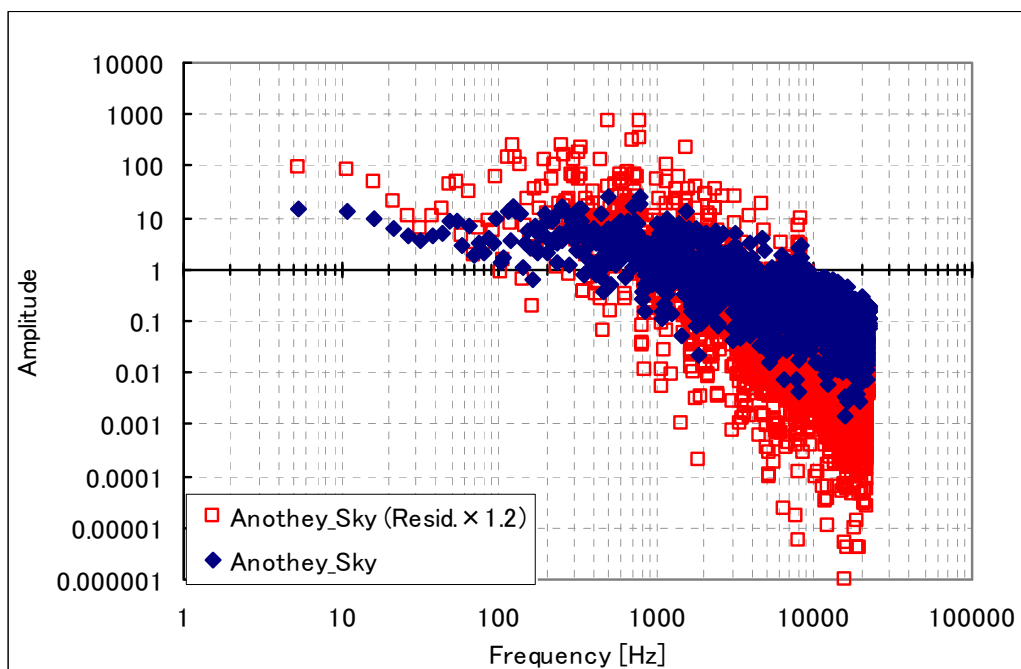


(a) 残差 0.4 倍 (実測値 : 表 3.3(b))



(b) 残差 0.8 倍 (実測値 : 表 3.3(b))

図 3.8 残差の変更



残差 1.2 倍 (実測値 : 表 3.3(b))

図 3.9 残差の変更

### 3.2. 本調査概要

随伴パラメータと楽曲の感性的印象の関係性を探るため、質問紙調査を行った。被験者は 20 代前半の学生 34 名である。質問紙調査は 2 種類行った。

調査 1 は、楽曲の感性的印象を問う質問である。調査 2 は各楽曲のゆらぎ値、残差二乗和を変化させたときの感性的印象を問う質問である。これらの質問紙調査で使用した楽曲は、表 3.2[11]に示したものである。調査に用いた質問紙を図 3.10~3.12 に示す。尚、本章における調査結果及び考察は第 4 章にて述べる。



### 3.3. ゆらぎ値随伴量と楽曲の感性的印象（調査1）

楽曲の感性的印象を問う調査1では、被験者に対して全10曲（1曲あたりの再生時間20秒）を最後まで聴かせ、毎曲終わるごとに表3.5[18]に示す凡例について1点～4点で評価させた。更に楽曲の総合的な印象（総合評価）を10点満点で評価させた。

表3.5 調査1に使用する凡例[18]

項目1	ゆったりした	1 ⇔ 4	せかせかした
項目2	重厚	1 ⇔ 4	軽薄
項目3	自然的	1 ⇔ 4	人工的
項目4	後ろ向きな	1 ⇔ 4	前向きな

項目1～4の評点及び総合評価は最頻値とし、最頻値が複数個存在する場合は、これらの値の平均を取った。総合評価に関しては平均値と標準偏差も示している。また各評価項目（項目1～4及び総合評価）の評点とゆらぎ値随伴量との相関係数を算出し、ゆらぎ値随伴量による楽曲の感性的印象を分析した。分析結果については、有意水準は1%にて平均値の差の検定を実施した[18]。

更に、(3.1)式に示す重回帰式を用いて、各被験者における楽曲の各評価項目の評点と、ゆらぎ値及び残差二乗和との関連性をより詳細に分析した。また、(3.1)式において、 $y_i$ は或る被験者の任意の評価項目における評点である。 $x_{i1}$ 及び $x_{i2}$ は、ゆらぎ値及び残差二乗和である。又、添字 $i$ は表3.2[11]に示す各楽曲の番号と対応する。尚、係数 $\alpha_0$ 、 $\alpha_1$ 、 $\alpha_2$ は未知数である。

即ち(3.1)式を用いれば、或る被験者の楽曲10曲に対する項目1の評点は、 $y_1, \dots, y_{10}$ の10組の重回帰式で表される事が分かる。

$$y_i = \alpha_0 + \alpha_1 x_{i1} + \alpha_2 x_{i2} + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma) \quad (i=1, \dots, 10) \quad (3.1)$$

各楽曲の各評価項目について、被験者 34 枚分の重回帰式を得た後、最小二乗法 [19, 20] を用い、個々の重回帰式における係数  $\alpha_0$ ,  $\alpha_1$ ,  $\alpha_2$  を求める。但し、(3.1) 式を用いて相関係数を求めた際、 $\alpha_0 = 0$  の場合が最も相関係数が高くなっていた為、 $\alpha_0$  は 0 に設定した。

続いて、被験者の感性的印象の分類を行う為、各被験者における項目 1~4 及び総合評価の重回帰式の各々におけるゆらぎ値係数  $\alpha_1$  及び残差二乗和係数  $\alpha_2$  について、Ward 法 [16] によるクラスター分析を行った。クラスター化が行われる都度、有意水準を 1% とした Wilks' Lambda 検定 [16, 19, 21] を行い、有意水準 1% にてクラスター間の有意差が無くなるまで上記の操作を繰り返した。

次に、次元の異なるゆらぎ値係数と残差二乗和係数を比較する為、基準化ゆらぎ値係数  $\alpha_1^*$  (Normalized Fluctuation Coefficient; NFC) 及び基準化残差二乗和係数  $\alpha_2^*$  (Normalized Residual Coefficient; NRC) を (3.2) 式, (3.3) 式を用いて定義した [13]。 (3.2) 式, (3.3) 式において、 $x_1^*$ ,  $x_2^*$  はゆらぎ値及び残差二乗和の基準化説明変数を表し、 $\bar{x}_1$  及び  $\bar{x}_2$  はゆらぎ値及び残差二乗和の説明変数の平均を表す。又、 $\sigma_n$  は標準偏差である。

$$x_n^* = \frac{x_n - \bar{x}_n}{\sigma_n}, \quad \therefore x_n = x_n^* \sigma_n + \bar{x}_n, \quad (n = 1, 2) \quad (3.2)$$

$$\begin{aligned} y &= \alpha_1 x_1 + \alpha_2 x_2 \\ &= \alpha_1 (x_1^* \sigma_1 + \bar{x}_1) + \alpha_2 (x_2^* \sigma_2 + \bar{x}_2) \\ &= (\alpha_1 \sigma_1) x_1^* + (\alpha_2 \sigma_2) x_2^* + (\alpha_1 \bar{x}_1 + \alpha_2 \bar{x}_2) \\ &= \alpha_1^* x_1^* + \alpha_2^* x_2^* + \alpha_3 \end{aligned} \quad (3.3)$$

### 3.4. ゆらぎ値随伴量的人為的変更と楽曲の感性的印象（調査2）

調査2では、楽曲のゆらぎ値又は残差二乗和を変更した Wave ファイル，及び変更していない Wave ファイルを被験者に聴かせ，最も好きだと感じた Wave ファイルを，各楽曲について1つずつ選択させた．各楽曲にて用意した Wave ファイルは次に示す（ア）～（オ）の5通りである．

- （ア） 原曲と同じもの
- （イ） ゆらぎ値を10%減少したもの
- （ウ） ゆらぎ値を10%加算したもの
- （エ） 残差を0.8倍したもの
- （オ） 残差を1.2倍したもの

本調査では直感的に感性的印象を判断させる為，各 Wave ファイルの初めの10秒間を聴かせ，どれが最も良いかを被験者に評価させた．但し，パラメータを変更した Wave ファイルの提示順序はランダムである．そして，この投票数で楽曲の感性的印象の向上の度合いを評価する．

本調査では被験者数が34人と小標本であるため，二項検定[19, 21]によって各楽曲の Wave ファイルにおける有意差を判断する，即ち，各楽曲において（ア）～（オ）の5通りに対する投票であるから，平均確率  $p_0=1/5$  であり，ゆらぎ値随伴量を変更した Wave ファイルの投票比率  $p$ （=投票数/被験者数）が有意差を持って  $p_0$  より大きいか否かを見る．つまり，本検定における帰無仮説  $H_0$  と対立仮説  $H_1$  は以下の通りである．

$H_0$  : ゆらぎ値随伴量を変更した Wave ファイルの投票比率  $p$  = 平均確率  $p_0$

$H_1$  : ゆらぎ値随伴量を変更した Wave ファイルの投票比率  $p >$  平均確率  $p_0$

本検定には以下の(3.4)式を用い、投票数  $b$  に対する確率  $P(k \geq b)$  を有意水準  $\alpha$  と比較する。(3.4)式が成立すれば帰無仮説  $H_0$  を棄却し、対立仮説  $H_1$  を採択する。そうでない場合は帰無仮説  $H_0$  を採択する。

$$P(k \geq b) = \sum_{k=b}^n {}_n C_k p_0^k (1-p_0)^{n-k} \leq \alpha \quad (3.4)$$

更に、重回帰分析によって投票数に対するゆらぎ値随伴量の関係を調査する。3.1節の予備調査より、切片はゆらぎ値及び残差二乗和との相関関係が強いこと明らかである。そこで多重共線性を回避する為、切片を説明変数から除去し、以下の(3.5)式を用いて重回帰分析を行う[11]。(3.5)式において、目的変数  $y_i$  は被験者の楽曲への投票数であり、説明変数  $x_{i1}$  及び  $x_{i2}$  は、ゆらぎ値及び残差二乗和である。尚、誤差  $\varepsilon_i$  は互いに独立に  $N(0, \sigma^2)$  に従うと仮定する。

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2) \quad (3.5)$$

楽曲の感性的印象に関するアンケート(1)

<はじめにお読みください>

(1) 各設問の音声は1回のみ流れます。  
 (2) このアンケートは統計的に処理をしますので、個人が特定されることはありません。  
 (3) **総合評価は1点(嫌い)～10点(好き)の10段階で記入してください。**  
 以上3点に注意の上、アンケートに回答してください。

✚ 1 番目の音声について、当てはまるものに○をつけてください。

ゆったりした ←	1	2	3	4		→ せかせかした
重厚 ←	1	2	3	4		→ 軽薄
自然的 ←	1	2	3	4		→ 人工的
後ろ向きな ←	1	2	3	4		→ 前向きな
総合評価 : _____						

✚ 2 番目の音声について、当てはまるものに○をつけてください。

ゆったりした ←	1	2	3	4		→ せかせかした
重厚 ←	1	2	3	4		→ 軽薄
自然的 ←	1	2	3	4		→ 人工的
後ろ向きな ←	1	2	3	4		→ 前向きな
総合評価 : _____						

✚ 3 番目の音声について、当てはまるものに○をつけてください。

ゆったりした ←	1	2	3	4		→ せかせかした
重厚 ←	1	2	3	4		→ 軽薄
自然的 ←	1	2	3	4		→ 人工的
後ろ向きな ←	1	2	3	4		→ 前向きな
総合評価 : _____						

✚ 4 番目の音声について、当てはまるものに○をつけてください。

ゆったりした ←	1	2	3	4		→ せかせかした
重厚 ←	1	2	3	4		→ 軽薄
自然的 ←	1	2	3	4		→ 人工的
後ろ向きな ←	1	2	3	4		→ 前向きな
総合評価 : _____						

✚ 5 番目の音声について、当てはまるものに○をつけてください。

ゆったりした ←	1	2	3	4		→ せかせかした
重厚 ←	1	2	3	4		→ 軽薄
自然的 ←	1	2	3	4		→ 人工的
後ろ向きな ←	1	2	3	4		→ 前向きな
総合評価 : _____						

<裏面に続きます>

図 3.10 質問紙 (調査 1, 1/2 頁)

6 番目の音声について、当てはまるものに○をつけてください。

ゆったりした ←	1	2	3	4	→ せかせかした
重厚 ←	1	2	3	4	→ 軽薄
自然的 ←	1	2	3	4	→ 人工的
後ろ向きな ←	1	2	3	4	→ 前向きな

総合評価： \_\_\_\_\_

7 番目の音声について、当てはまるものに○をつけてください。

ゆったりした ←	1	2	3	4	→ せかせかした
重厚 ←	1	2	3	4	→ 軽薄
自然的 ←	1	2	3	4	→ 人工的
後ろ向きな ←	1	2	3	4	→ 前向きな

総合評価： \_\_\_\_\_

8 番目の音声について、当てはまるものに○をつけてください。

ゆったりした ←	1	2	3	4	→ せかせかした
重厚 ←	1	2	3	4	→ 軽薄
自然的 ←	1	2	3	4	→ 人工的
後ろ向きな ←	1	2	3	4	→ 前向きな

総合評価： \_\_\_\_\_

9 番目の音声について、当てはまるものに○をつけてください。

ゆったりした ←	1	2	3	4	→ せかせかした
重厚 ←	1	2	3	4	→ 軽薄
自然的 ←	1	2	3	4	→ 人工的
後ろ向きな ←	1	2	3	4	→ 前向きな

総合評価： \_\_\_\_\_

10 番目の音声について、当てはまるものに○をつけてください。

ゆったりした ←	1	2	3	4	→ せかせかした
重厚 ←	1	2	3	4	→ 軽薄
自然的 ←	1	2	3	4	→ 人工的
後ろ向きな ←	1	2	3	4	→ 前向きな

総合評価： \_\_\_\_\_

以上です。ご協力、ありがとうございました。

図 3.11 質問紙（調査 1, 2/2 頁）

楽曲の感性的印象に関するアンケート(2)

<はじめにお読みください>

- (1) 楽曲は1回しか流れません。
- (2) このアンケートは統計的に処理をしますので、個人が特定されることはありません。
- (3) 今から、曲目番号1~10について、処理を施したものを1曲につき①~⑤の5パターン聞いて頂きます。各曲目につき、最も良いと思う番号の欄に、1つだけ✓をつけてください。

以上3点に注意の上、アンケートに回答してください。

曲目	①	②	③	④	⑤
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					

以上です。ご協力、ありがとうございました。

図 3.12 質問紙 (調査 2)

## 第4章

### 調査結果と考察

#### 4.1. ゆらぎ値随伴量による楽曲の感性的印象（調査1）

楽曲の感性的印象を問う調査1では、被験者に対して全10曲（1曲あたりの再生時間20秒）を最後まで聴かせ、毎曲終わるごとに4.2節の表4.1[18]に示す凡例について1点～4点で評価させた。更に楽曲の総合的な印象（総合評価）を10点満点で評価させた。結果を表4.1に示す。表4.1の内、項目1～4の値は最頻値である。最頻値が複数個存在する場合は、それらの値の平均を取った。総合評価においては最頻値と、平均値及び標準偏差を示している。

図4.1は項目1と残差二乗和の関係を表した散布図である。表4.1の残差二乗和との相関係数は0.62であり、中程度の相関が見られる。項目1の評価値に注目すれば、残差二乗和 $\geq 0.50$ の楽曲で評価値が3以上の楽曲は10曲中3曲であり、残差二乗和 $< 0.50$ の楽曲で評価値が3未満の楽曲は10曲中6曲である。

即ち、残差二乗和は『せかせかした』印象を与えていると考えられる。このことから、残差二乗和 $\geq 0.50$ である楽曲は『せかせかした』印象が強くなり、残差二乗和 $< 0.50$ である楽曲は『ゆったりした』印象が強くなると考えられる。

図4.2は項目4と切片の関係を示した散布図である。表4.1の残差二乗和との相関係数を見ると0.58であり、中程度の相関が見られる。項目4の評価値に注目すれば、切片 $\geq 7.7$ の楽曲で評価値が3以上の楽曲は10曲中5曲であり、切片 $< 7.7$ の楽曲で



評価値が3未満の楽曲は10曲中4曲である。即ち、切片は『前向きな』印象を与えていると考えられる。この事から、切片 $\geq 7.7$ である楽曲は『前向きな』印象が強くなり、切片 $< 7.7$ である楽曲は『後ろ向きな』印象が強くなると考えられる。

表 4.1 楽曲の感性的印象に関するアンケート（調査1）[18]

番号	ファイル名	ゆらぎ値	切片	残差二乗和	項目1	項目2	項目3	項目4	総合評価		
									最頻値	平均値	標準偏差
1	Another_Sky.wav	1.715343	7.704218	0.414398	2.0	3.0	3.0	3.0	6.0	5.9	2.54
2	ロンドンデリーエア.wav	1.627244	7.080130	0.398816	1.0	2.0	2.0	2.0	7.0	6.2	3.04
3	君を信じて.wav	2.110503	9.850914	0.466082	2.0	2.0	2.0	4.0	6.0	6.4	2.01
4	漂流.wav	1.592919	6.873790	0.430658	1.0	1.0	2.0	1.0	7.0	6.0	1.43
5	Down_by_the_Riverside.wav	1.724475	8.052250	0.448193	3.0	4.0	4.0	4.0	6.0	5.3	3.10
6	Space_Odessey3_Revelation.wav	0.990190	3.543654	0.323250	1.0	2.0	2.0	1.0	6.0	5.6	2.73
7	TOMORROW.wav	2.267830	11.134446	0.555204	3.0	3.0	4.0	4.0	7.0	6.6	3.41
8	古いフランスの歌.wav	1.398202	5.805316	0.433054	1.0	2.0	2.0	1.0	4.0	5.4	2.58
9	自由.wav	2.172833	11.011055	0.588947	4.0	4.0	4.0	4.0	5.0	5.8	1.80
10	赤い川の谷間(brass).wav	1.520269	7.336829	0.526494	3.0	4.0	3.0	4.0	7.0	5.9	2.29
相関係数	ゆらぎ値	—			0.44	0.03	0.25	0.54	0.11		
	切片	—			0.51	0.09	0.30	0.58	0.10		
	残差二乗和	—			0.62	0.20	0.38	0.51	0.06		

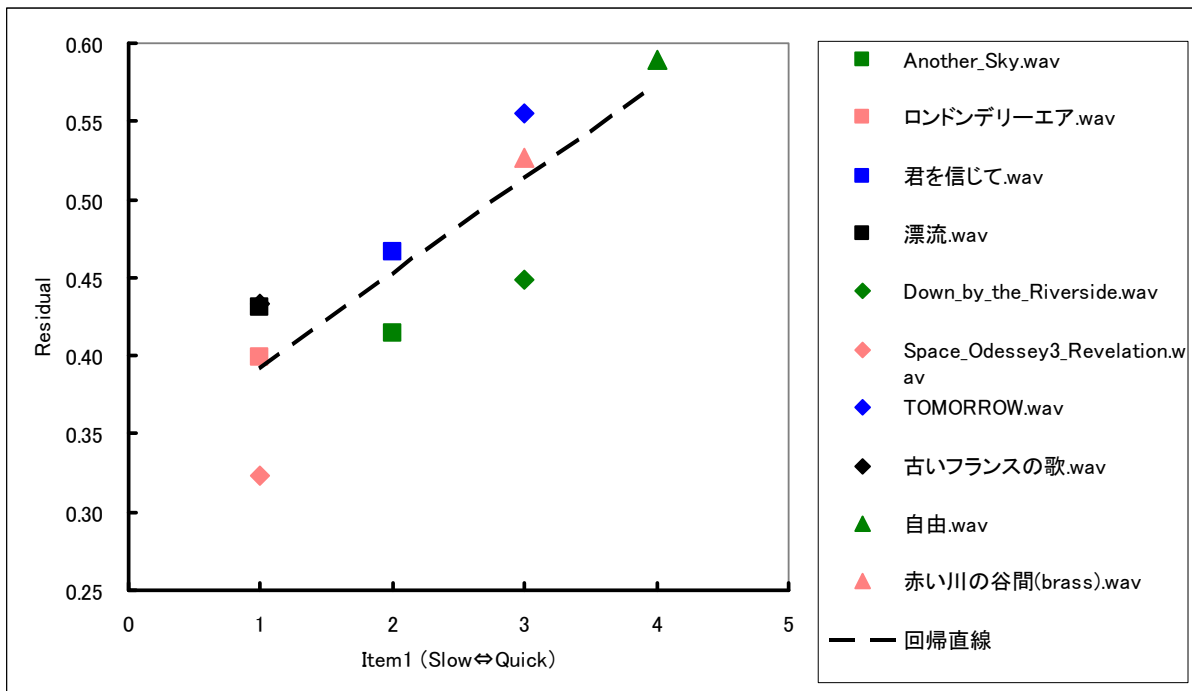


図 4.1 項目1と残差二乗和の関連性[18]

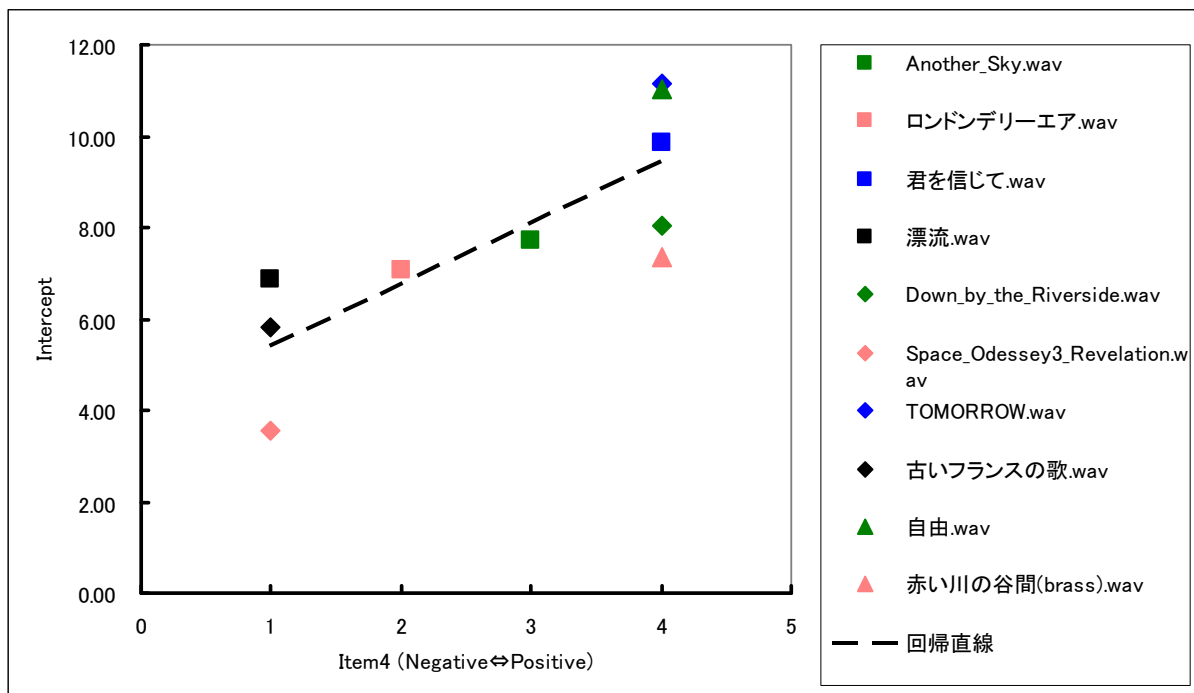


図 4.2 項目 4 と切片との関連性[18]

続いて、項目の評価値と各帯域におけるゆらぎ値随伴量について平均値の検定を実施した。検定内容は次の(A)及び(B)である。

(A) 項目 1 :

残差二乗和 $\geq 0.50$ である楽曲は『せかせかした』印象を与え、残差二乗和 $< 0.50$ である楽曲は『ゆったりした』印象を与える

(B) 項目 4 :

切片 $\geq 7.7$ の楽曲は『前向きな』印象を与え、切片 $< 7.7$ の楽曲は『後ろ向きな』印象を与える

以上について項目の評価値とゆらぎ値随伴量について平均値の検定を行ったところ、1%の有意水準で有意であった。この検定は、例えば残差二乗和 $\geq 0.50$ の楽曲に

対する 34 名に評価の平均と、残差二乗和 $<0.50$ のそれらの平均についての差の検定を意味する[11]. よって、今回調査した楽曲においては、残差二乗和は『せかせかした』印象と相関性が高く、切片は『前向きな』印象と相関性が高い事が確認された.

更に、各被験者における楽曲の項目 1~4 及び総合評価（以下、各評価項目）の評点と、ゆらぎ値及び残差二乗和との関連性をより明らかにする為、(4.1)式に示す重回帰式を用いて分析を行った. この分析は、音声信号の全域 (All-Frequency domain ; 以下 AF, 0~22050Hz), 低域 (Low-Frequency domain ; 以下 LF, 0~300Hz), 中域 (Middle-Frequency domain ; 以下 MF, 300~1000Hz), 高域 (High-Frequency domain ; 以下 HF, 1000~22050Hz) の 4 つの周波数帯域について行った[15].

各被験者の重回帰式におけるゆらぎ値係数 $\alpha_1$ 及び残差二乗和係数 $\alpha_2$ の組を Ward 法にてクラスター分析すると、概ね 3 つのグループに分けられた. 以下に、周波数帯域ごとの結果及び考察を述べる.

## I. AF (全域)

図 4.3~図 4.7 の散布図に、(4.1)式にてゆらぎ値係数 $\alpha_1$ 及び残差二乗和係数 $\alpha_2$ を算出し重回帰分析を行った結果を示す. これらの図より、Ward 法によるクラスター分析の結果、同じ感性的印象を有するグループは、Group1, Group2, Group3 の 3 つに分類された事が分かる. また、表 4.2 に各グループの基準化ゆらぎ値係数 (Normalized Fluctuation Coefficient; 以下 NFC)  $\alpha_1^*$  及び基準化残差二乗和係数 (Normalized Residual Coefficient; 以下 NRC)  $\alpha_2^*$  の平均値を示す.

表 4.2 にて Group1 の  $\text{NFC}\alpha_1^*$  と  $\text{NRC}\alpha_2^*$  を比較すれば、各評価項目において  $\text{NRC}\alpha_2^*$  の方が  $\text{NFC}\alpha_1^*$  よりも高いことが分かる. 更に、 $\text{NRC}\alpha_2^*$  は正の値であり  $\text{NFC}\alpha_1^*$  は負の値である事も分かる. ここで(4.1)式を考慮すると、ゆらぎ値  $x_1$  が増加すれば、各評価項目の評点は低下し、ゆらぎ値  $x_1$  が減少すれば、各評価項目の評点はやや向上すると考えられる. また、残差二乗和  $x_2$  の増減によって、

各評価項目の評点はゆらぎ値  $x_1$  を変化させた時と比較して大きく増減すると考えられる。

Group2 では項目 1~3 における  $NFC\alpha_1^*$  は負の値であり、その絶対値は  $NRC\alpha_2^*$  よりも低いことが分かる。つまり、Group1 と同様にゆらぎ値  $x_1$  の減少・増加によって項目 1~3 の評価が増加・減少し、残差二乗和  $x_2$  の増加・減少によって項目 1~3 の評価が増加・減少すると考えられ、影響度はゆらぎ値  $x_1$  より残差二乗和  $x_2$  の方が大きいと考えられる。また、項目 4 では  $NFC\alpha_1^*$  と  $NRC\alpha_2^*$  の双方が正の値であるが、 $NRC\alpha_2^*$  の方が高いため、残差二乗和  $x_2$  の影響度の方がゆらぎ値  $x_1$  の影響度より大きいと考えられる。一方、総合評価に注目してみれば、 $NFC\alpha_1^*$  は  $NRC\alpha_2^*$  よりも高く、且つ  $NRC\alpha_2^*$  は負の値である為、ゆらぎ値  $x_1$  の増減の影響度が大きいと考えられる。

Group3 では項目 1, 2 及び総合評価の  $NRC\alpha_2^*$  の方が  $NFC\alpha_1^*$  よりも高い事が分かる。つまり項目 1, 2 及び総合評価の評点の増減に寄与するのは、ゆらぎ値  $x_1$  の増減よりも残差二乗和  $x_2$  の増減の方であると考えられる。また項目 3, 4 では  $NFC\alpha_1^*$  が  $NRC\alpha_2^*$  の 3 倍以上である。よって項目 3, 4 では、ゆらぎ値  $x_1$  は残差二乗和  $x_2$  よりも評点の増減に寄与すると考えられる。

更に、各グループにおける相関係数を調査した結果、0.949 以上であった。即ち、(4.1)式の重回帰式を用いて、少なくとも結果の 94.9%を説明できることが判明した。

以上を総括すると、AF では残差二乗和  $x_2$  の方がゆらぎ値  $x_1$  よりも感性的印象に対する、より強いファクターであると考えられる。

## II. LF (低域)

表 4.3 に、LF における各グループの  $NFC\alpha_1^*$  及び  $NRC\alpha_2^*$  の平均値、及び重回帰係数を示し、図 4.8~図 4.12 にクラスター分析の結果を示す。

表 4.3 の  $NFC\alpha_1^*$  及び  $NRC\alpha_2^*$  の値に注目すれば、項目 1~項目 4 において  $NRC\alpha_2^*$

の方が  $NFC\alpha_1^*$  よりも高くなる傾向が見られる。しかしながら項目3及び総合評価の重相関係数は0.316及び0.021と低い為、この2つの評価項目に関しては被験者の好みが大きく反映されたと考えられる。

従って、LFにおける項目1, 2, 4では残差二乗和  $x_2$  の方がゆらぎ値  $x_1$  よりも感性的印象に対する、より強いファクターであると考えられるが、項目3及び総合評価では個人差が大きい為ばらつきが生じたと考えられる。

### III. MF (中域)

表4.4に、MFにおける各グループの  $NFC\alpha_1^*$  及び  $NRC\alpha_2^*$  の平均値、及び重回帰係数を示し、図4.13～図4.17にMFのクラスター分析の結果を示す。

表4.4より、Group1とGroup2において、各評価項目の  $NRC\alpha_2^*$  が  $NFC\alpha_1^*$  よりも高くなっていることが分かる。従って双方のグループにおける感性的印象は、残差二乗和  $x_2$  の影響の方がゆらぎ値  $x_1$  による影響よりも大きいと考えられる。

更に、MFにおける全体的な傾向として、残差二乗和  $x_2$  の影響の方がゆらぎ値  $x_1$  による影響よりも大きくなっていることが分かる。

### IV. HF (高域)

表4.5に、HFにおける各グループの  $NFC\alpha_1^*$  及び  $NRC\alpha_2^*$  の平均値、及び重回帰係数を示し、図4.18～図4.22にHFのクラスター分析の結果を示す。

表4.5より、Group2とGroup3において項目1と項目4の  $NFC\alpha_1^*$  が  $NRC\alpha_2^*$  よりも高いことが分かる。これ以外については、 $NRC\alpha_2^*$  が  $NFC\alpha_1^*$  よりも高いことが分かる。

従って、Group2とGroup3における項目1と項目4の感性的印象は残差二乗和  $x_2$  の影響よりもゆらぎ値  $x_1$  の影響が大きいと考えられ、それ以外については、残差二乗和  $x_2$  の影響の方がゆらぎ値  $x_1$  の影響よりも大きいと考えられる。

以上，I～IVにて述べた各帯域について考慮すれば，表4.2～4.5に示す分析結果より， $NRC\alpha_2^*$ は $NFC\alpha_1^*$ よりも概ね高くなる傾向にある事が分かった。即ち，ゆらぎ値 $x_1$ を変動させるよりも残差二乗和 $x_2$ を変動させた方が感性的印象への影響が大きいと考えられる。更に，表4.3～4.5より，高周波になるほど各評価項目の重相関係数が高くなっている事が分かる。従って，高周波ほど残差二乗和 $x_2$ の影響が強くなると考えられる。

表4.2 AF（全域）の重相関係数及びNFC ( $\alpha_1^*$ )，NRC ( $\alpha_2^*$ ) 一覧[15]

評価項目	重相関係数	有意水準	回帰係数の平均値					
			Group 1		Group 2		Group 3	
			$\alpha_1^*$	$\alpha_2^*$	$\alpha_1^*$	$\alpha_2^*$	$\alpha_1^*$	$\alpha_2^*$
項目1 (ゆったりした⇄せかせかした)	0.991	1%	-0.949	1.082	-0.295	0.625	0.132	0.252
項目2 (重厚⇄軽薄)	0.980	1%	-0.941	1.154	-0.404	0.772	0.024	0.408
項目3 (自然的⇄人工的)	0.980	1%	-0.657	0.968	-0.097	0.528	0.395	0.112
項目4 (後ろ向き⇄前向き)	0.990	1%	-0.812	1.103	0.199	0.317	0.620	0.014
総合評価	0.949	1%	-0.939	1.925	1.478	-0.186	0.150	0.831

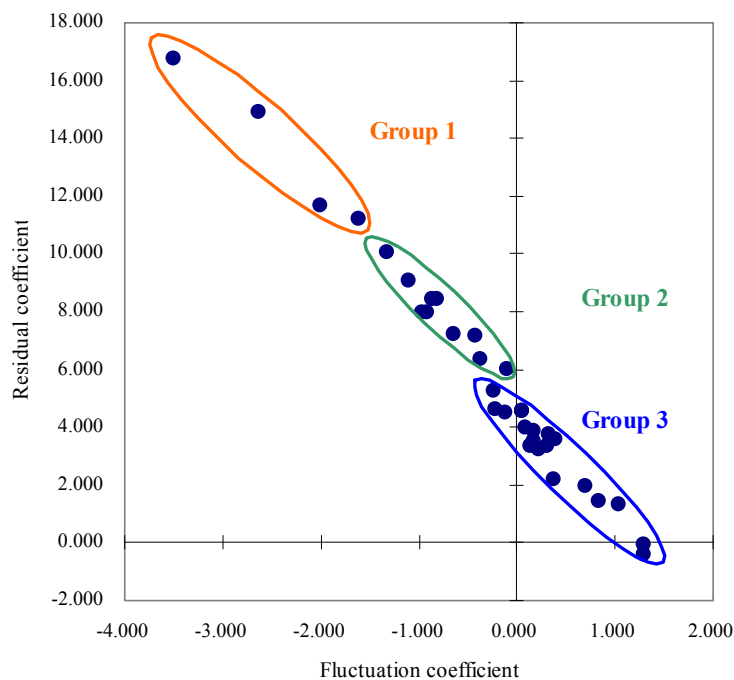


図4.3 項目1の回帰分析結果（AF）[15]

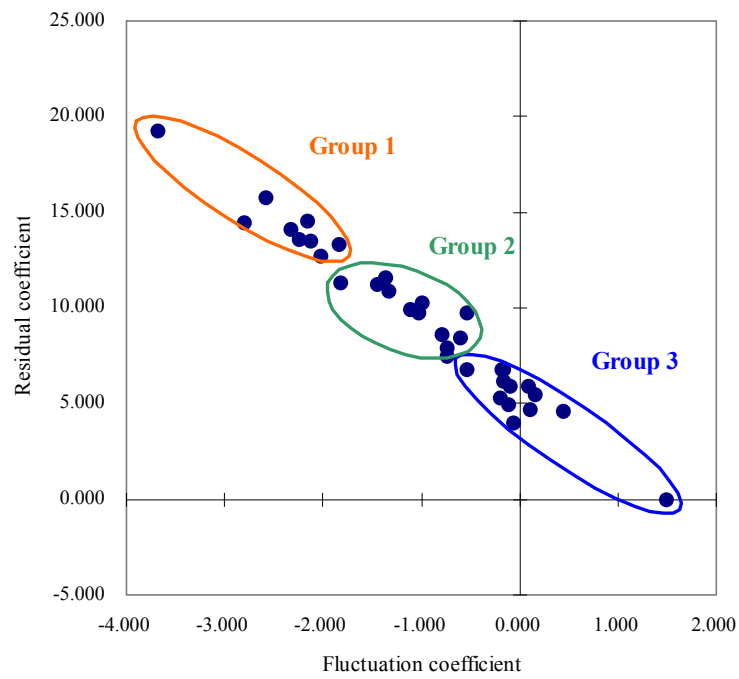


図 4.4 項目 2 の回帰分析結果 (AF) [15]

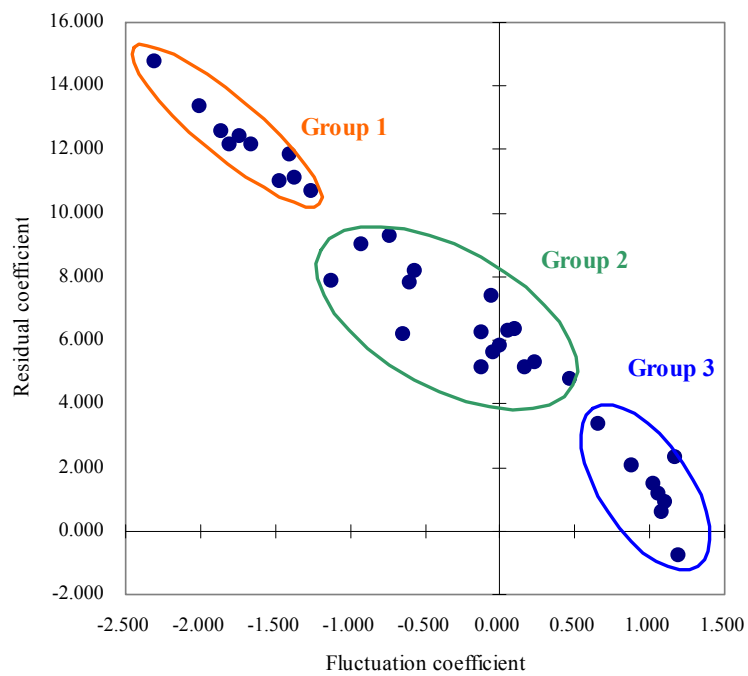


図 4.5 項目 3 の回帰分析結果 (AF) [15]

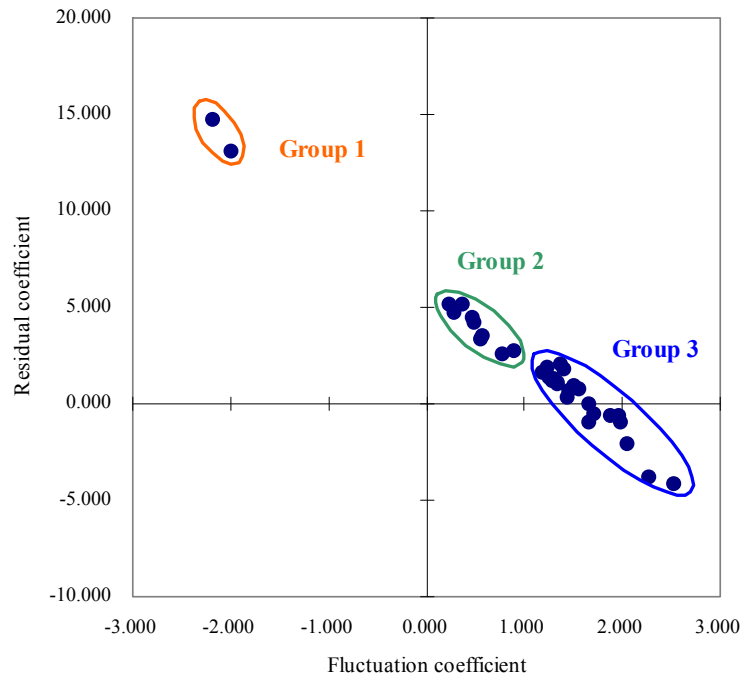


図 4.6 項目4の回帰分析結果 (AF) [15]

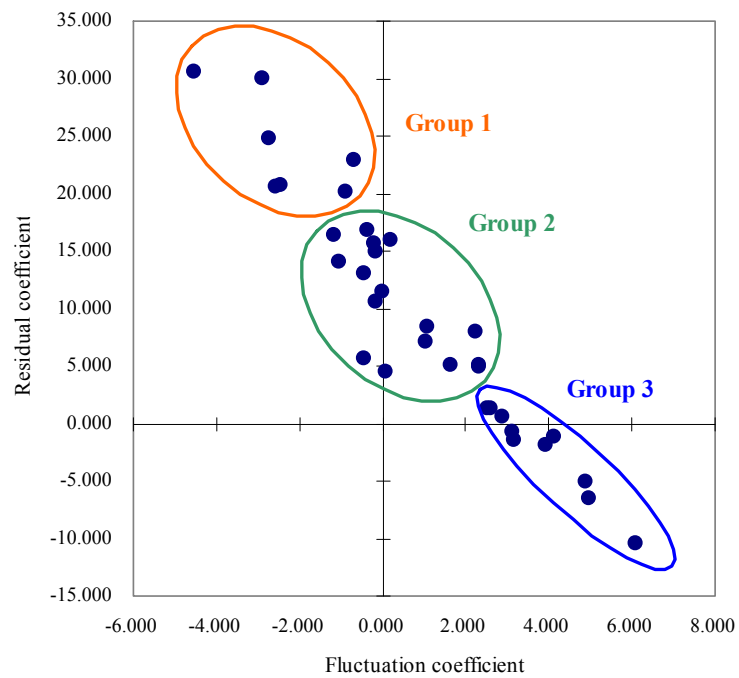


図 4.7 総合評価の回帰分析結果 (AF) [15]



表 4.3 LF（低域）の重相関係数及び NFC ( $\alpha_1^*$ ), NRC ( $\alpha_2^*$ ) 一覧[15]

評価項目	重相関係数	有意水準	回帰係数の平均値					
			Group 1		Group 2		Group 3	
			$\alpha_1^*$	$\alpha_2^*$	$\alpha_1^*$	$\alpha_2^*$	$\alpha_1^*$	$\alpha_2^*$
項目1 (ゆったりした⇄せかせかした)	0.763	1%	-0.182	0.662	0.403	0.448	-	-
項目2 (重厚⇄軽薄)	0.669	1%	0.099	0.757	0.563	0.520	-	-
項目3 (自然的⇄人工的)	0.316	1%	0.085	0.788	0.258	0.634	-	-
項目4 (後ろ向き⇄前向き)	0.829	1%	-0.143	0.862	0.449	0.629	-	-
総合評価	0.021	1%	0.904	1.684	1.177	1.228	0.830	0.317

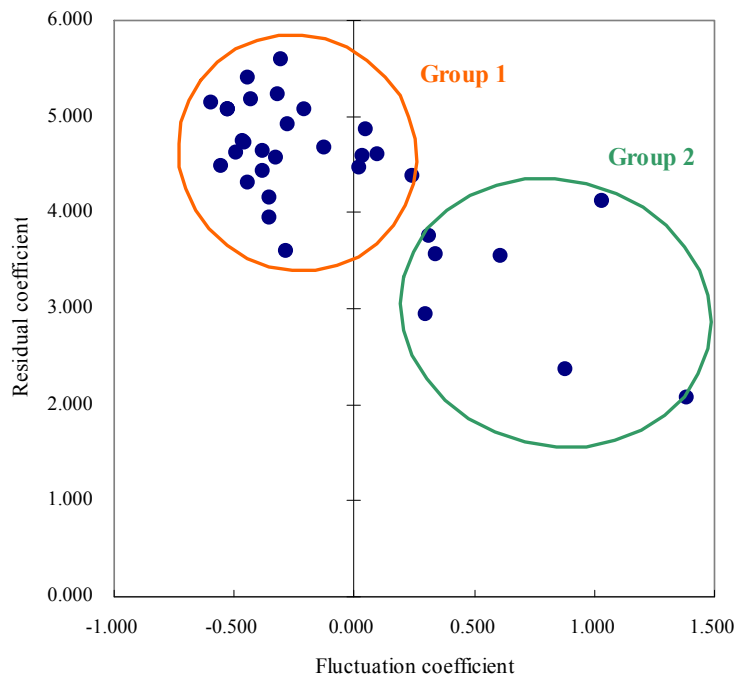


図 4.8 項目 1 の回帰分析結果 (LF) [15]

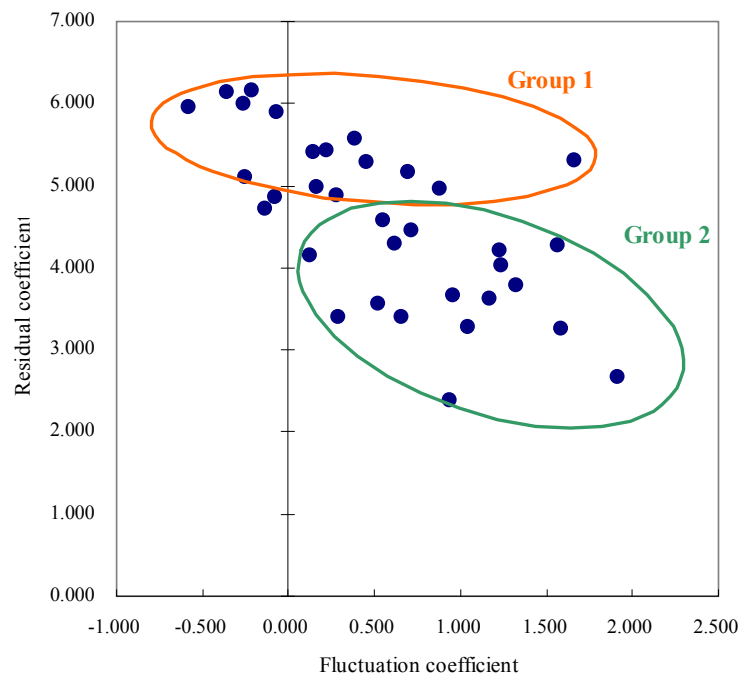


図 4.9 項目 2 の回帰分析結果 (LF) [15]

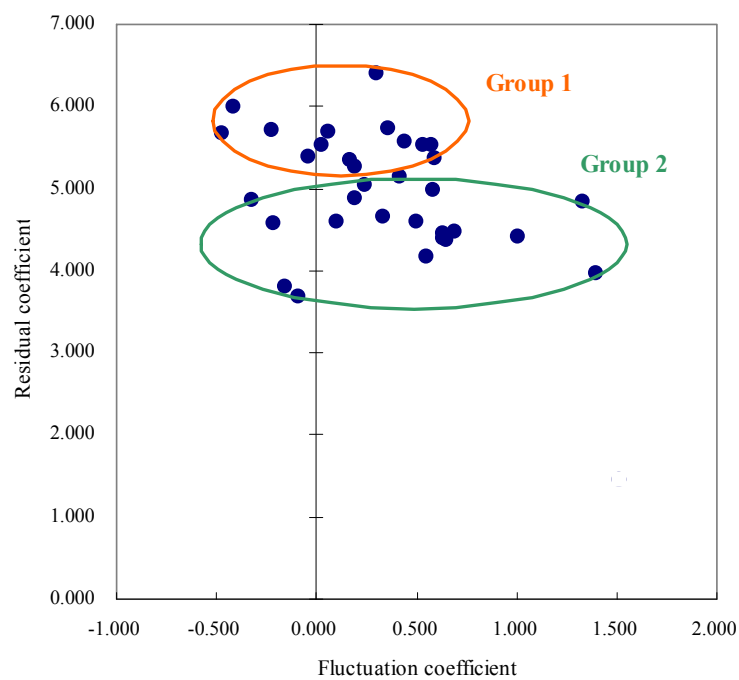


図 4.10 項目 3 の回帰分析結果 (LF) [15]

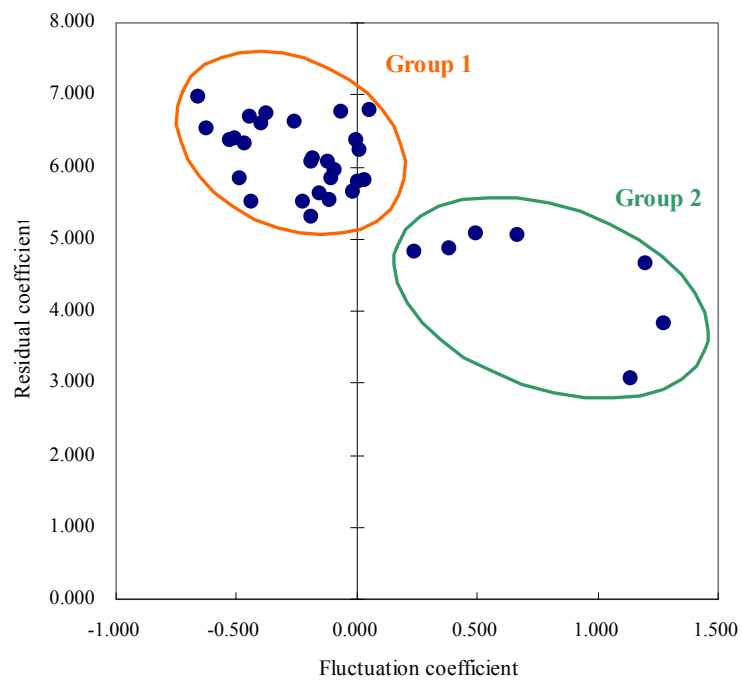


図 4.11 項目4の回帰分析結果 (LF) [15]

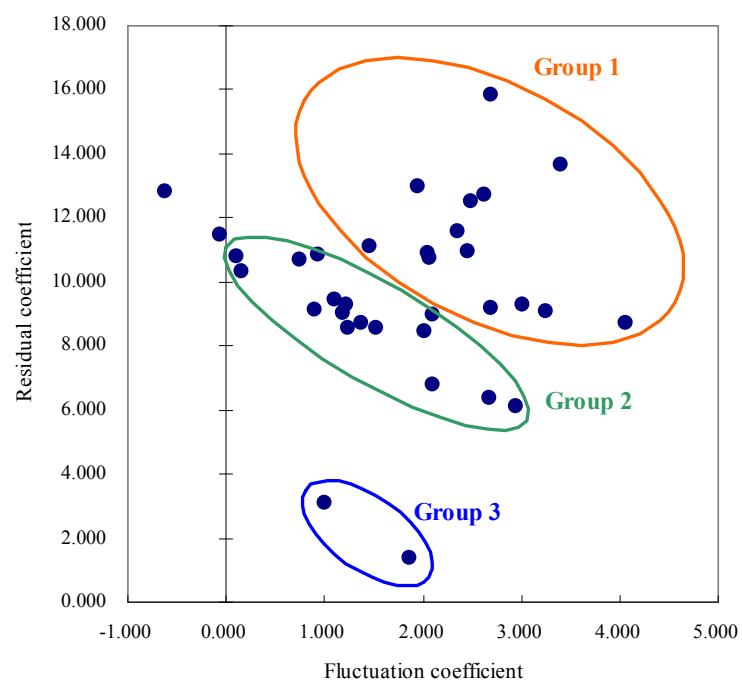


図 4.12 総合評価の回帰分析結果 (LF) [15]

表 4.4 MF（中域）の重相関係数及び NFC ( $\alpha_1^*$ ), NRC ( $\alpha_2^*$ ) 一覧[15]

評価項目	重相関係数	有意水準	回帰係数の平均値					
			Group 1		Group 2		Group 3	
			$\alpha_1^*$	$\alpha_2^*$	$\alpha_1^*$	$\alpha_2^*$	$\alpha_1^*$	$\alpha_2^*$
項目1 (ゆったりした⇄せかせかした)	0.831	1%	-0.671	0.887	-0.174	0.670	-	-
項目2 (重厚⇄軽薄)	0.816	1%	-0.293	0.880	0.112	0.629	0.699	0.349
項目3 (自然的⇄人工的)	0.879	1%	-0.444	0.983	-0.050	0.778	0.303	0.422
項目4 (後ろ向き⇄前向き)	0.911	1%	-0.556	1.048	-0.218	0.842	0.289	0.583
総合評価	0.613	1%	-0.173	2.070	0.397	1.365	0.841	0.620

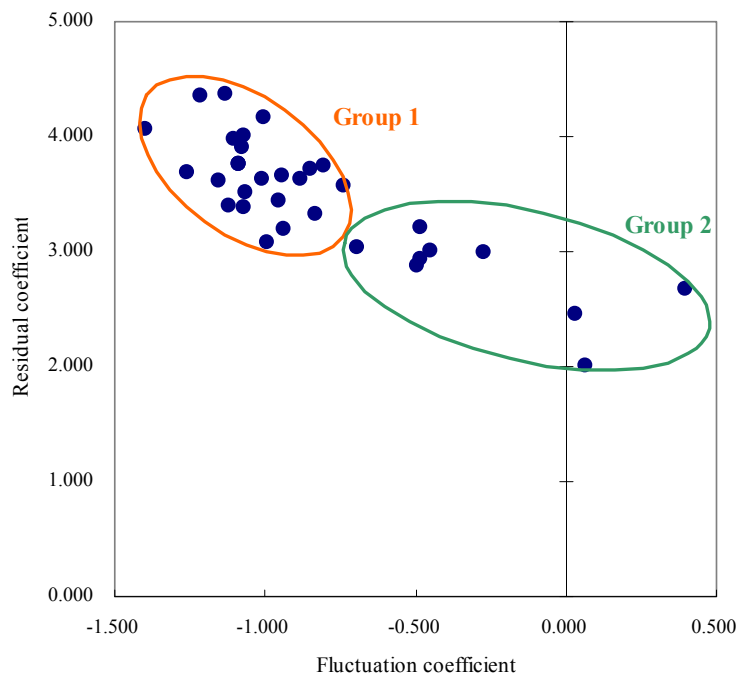


図 4.13 項目1の回帰分析結果 (MF) [15]

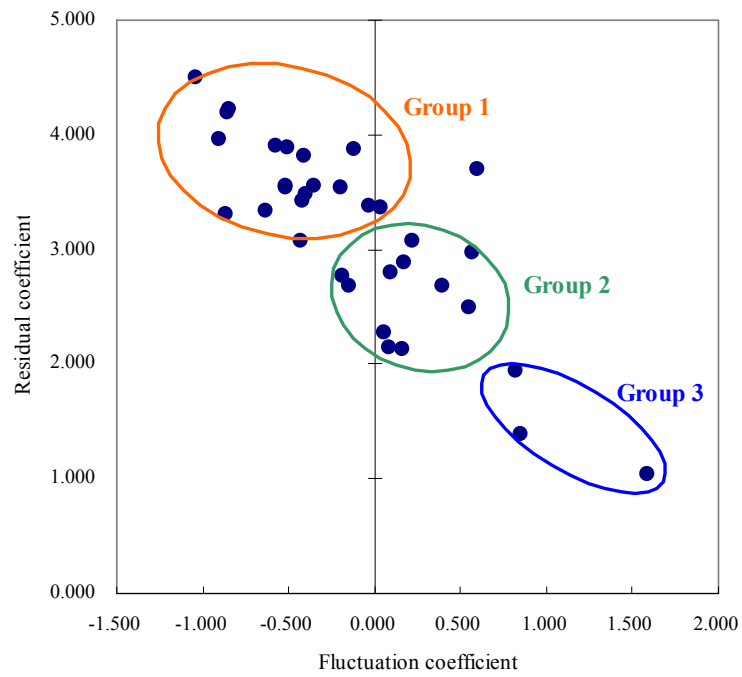


図 4.14 項目 2 の回帰分析結果 (MF) [15]

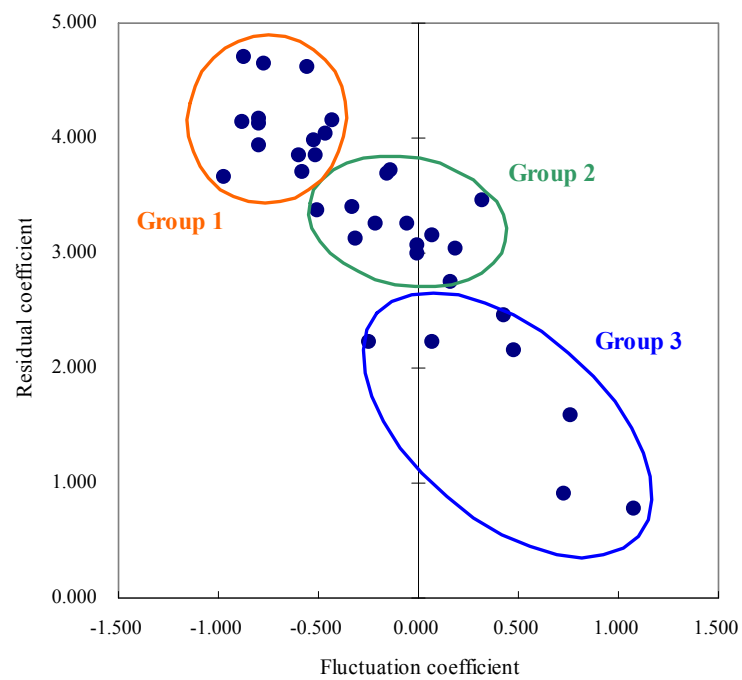


図 4.15 項目 3 の回帰分析結果 (MF) [15]

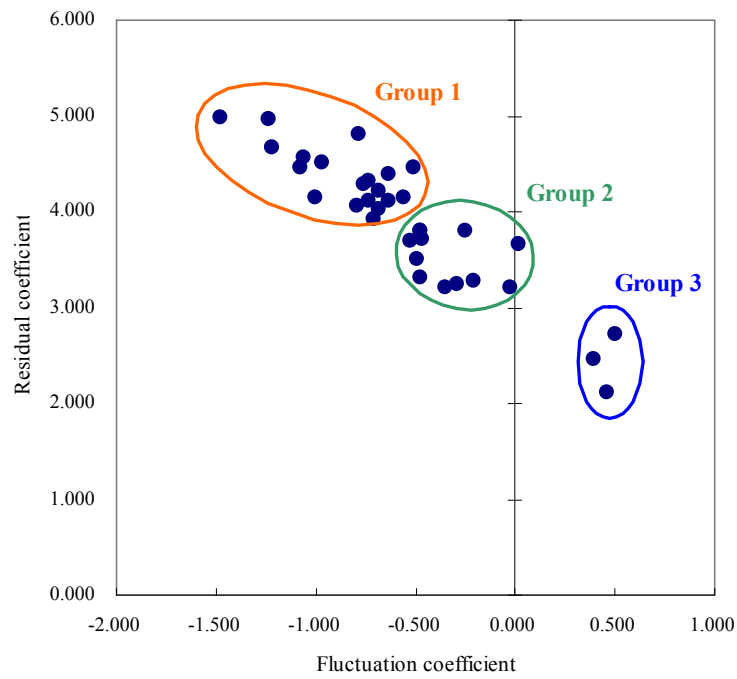


図 4.16 項目4の回帰分析結果 (MF) [15]

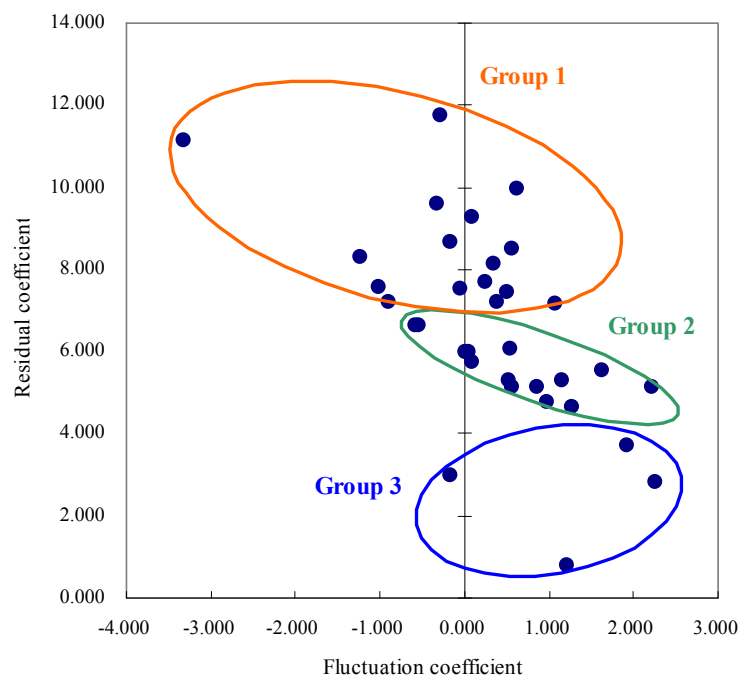


図 4.17 総合評価の回帰分析結果 (MF) [15]

表 4.5 HF（高域）の重相関係数及び NFC ( $\alpha_1^*$ ), NRC ( $\alpha_2^*$ ) 一覧[15]

評価項目	重相関係数	有意水準	回帰係数の平均値					
			Group 1		Group 2		Group 3	
			$\alpha_1^*$	$\alpha_2^*$	$\alpha_1^*$	$\alpha_2^*$	$\alpha_1^*$	$\alpha_2^*$
項目1 (ゆったりした⇄せかせかした)	0.984	1%	-0.361	0.502	0.359	0.143	0.649	-0.031
項目2 (重厚⇄軽薄)	0.973	1%	-0.846	0.840	-0.293	0.503	0.280	0.240
項目3 (自然的⇄人工的)	0.947	1%	-0.423	0.585	0.046	0.398	0.367	0.178
項目4 (後ろ向きな⇄前向きな)	0.984	1%	-0.859	0.841	0.365	0.227	0.709	0.061
総合評価	0.880	1%	-0.620	1.299	0.260	0.661	1.053	0.193

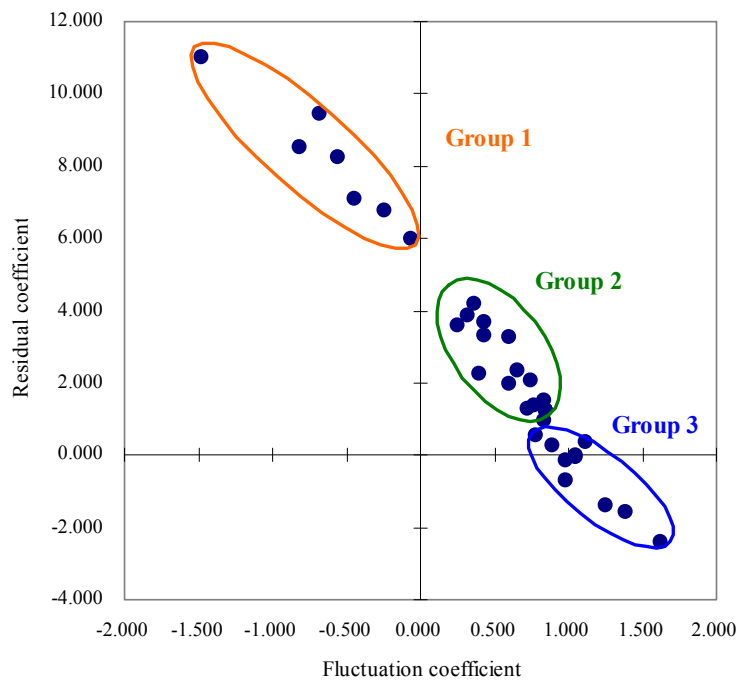


図 4.18 項目1の回帰分析結果 (HF) [15]

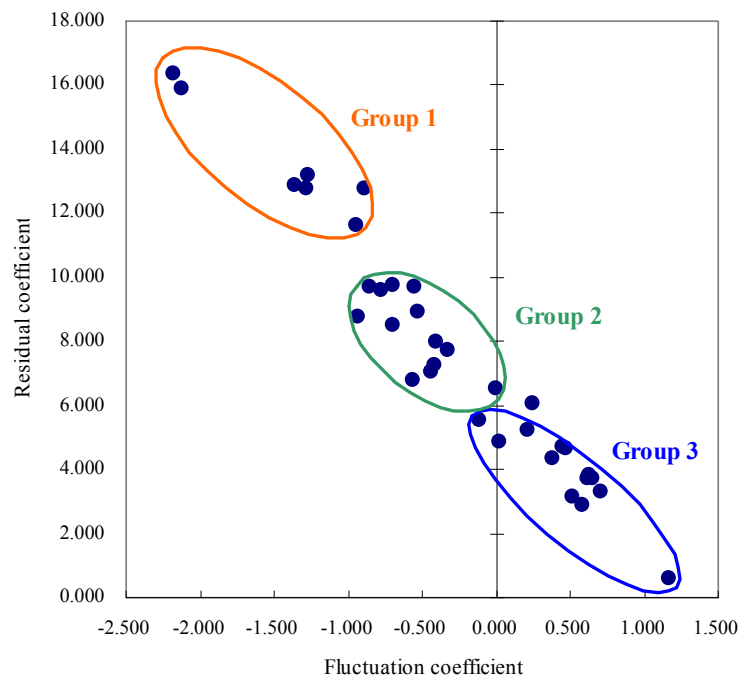


図 4.19 項目 2 の回帰分析結果 (HF) [15]

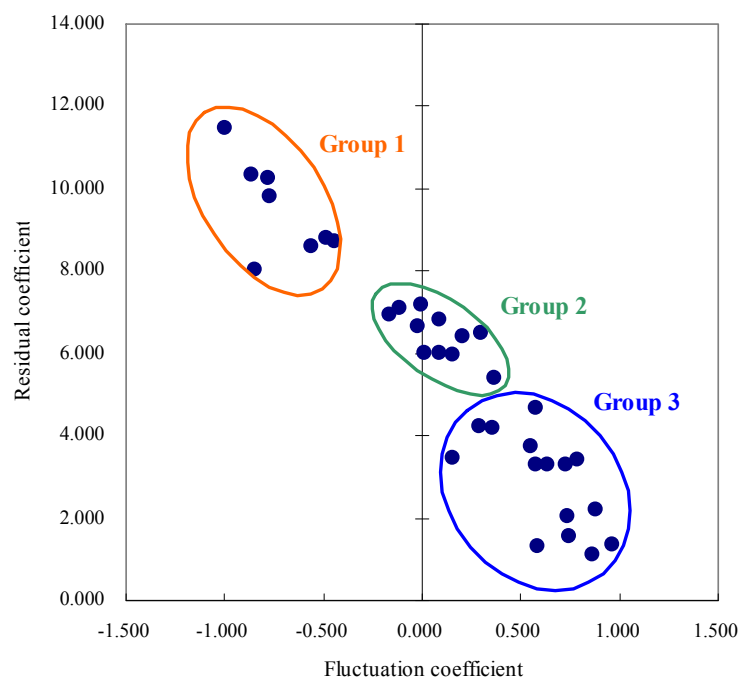


図 4.20 項目 3 の回帰分析結果 (HF) [15]



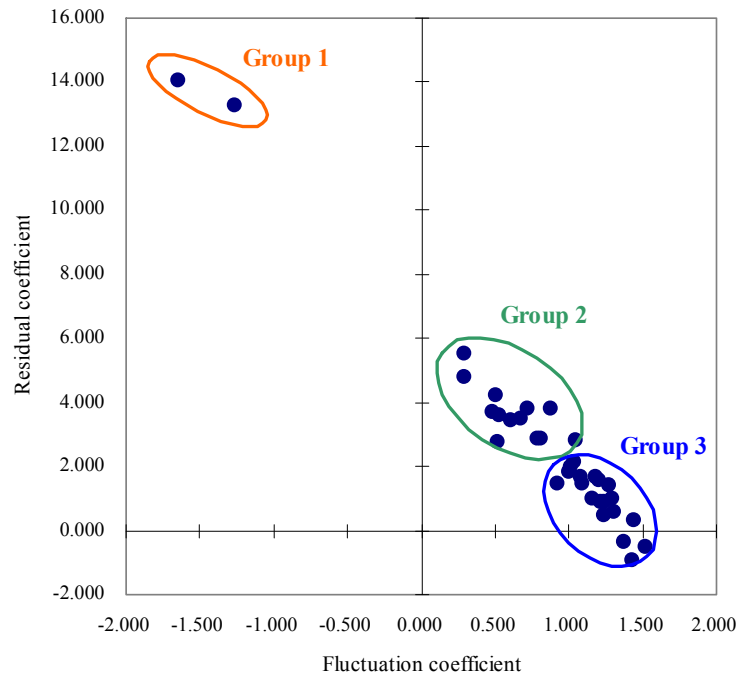


図 4.21 項目4の回帰分析結果 (HF) [15]

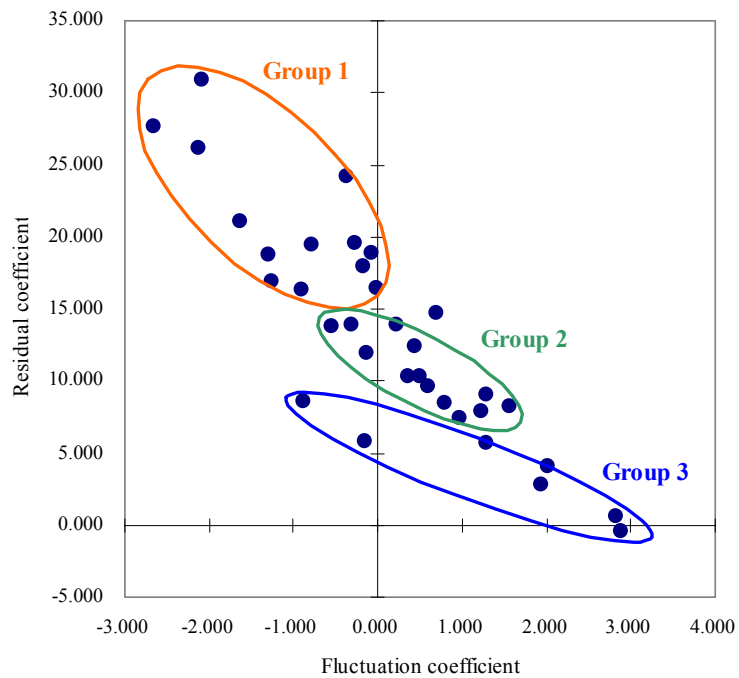


図 4.22 総合評価の回帰分析結果 (HF) [15]

## 4.2. ゆらぎ値随伴量的人為的変更と楽曲の感性的印象（調査2）

楽曲のゆらぎ値随伴量を変更した Wave ファイル, 及び変更していない Wave ファイルを被験者に聴かせ, 最も好きだと感じた Wave ファイルを, 各楽曲について1つずつ選択させた. 各楽曲に各楽曲にて用意した Wave ファイルは次に示す (ア) ~ (オ) の5通りである. 結果を表 4.6 に示す. 単位は[人]である. 図 4.23 は表 4.7 を図示したものである.

- (ア) 原曲と同じもの
- (イ) ゆらぎ値を 10%減少したもの
- (ウ) ゆらぎ値を 10%加算したもの
- (エ) 残差を 0.8 倍したもの
- (オ) 残差を 1.2 倍したもの

図 4.23 より, ゆらぎ値随伴量を変更した Wave ファイルが最も高い得票数を得た楽曲は, 10 曲中 8 曲 (楽曲番号 2, 3, 4, 5, 6, 7, 8, 10) であることが分かる. これら 8 曲の各最高投票数に対して, 二項検定[19, 21]によって統計的に意味を持つか否かを検証した. 帰無仮説  $H_0$  及び対立仮説  $H_1$  を以下に示す.

$H_0$ : ゆらぎ値随伴量を変更した Wave ファイルの投票比率  $p$  = 平均確率  $p_0$

$H_1$ : ゆらぎ値随伴量を変更した Wave ファイルの投票比率  $p >$  平均確率  $p_0$

4.3 節に示した (4.4) 式を実際に計算すれば, 投票数  $b$  が 12 以上ならば有意水準 5% で対立仮説  $H_1$  が採択され, 投票数  $b$  が 11 以上ならば有意水準 10% で対立仮説  $H_1$  が採択されることが分かる.

従って図 4.23 を見れば, 楽曲番号 4 の最高投票数は有意水準 5% で平均より高く, 楽曲番号 6, 8 では有意水準 10% で平均より高いことが分かる. しかし, 楽曲番号 2, 3,

5, 7, 10 については平均程度となる。一方, 楽曲番号 2~8, 10 の原曲への投票数は平均程度, 或いはそれ以下となる。

表 4.6 ゆらぎ値随伴量を変更した際の感性的印象に関するアンケート(調査2) [18]

番号	ファイル名	ゆらぎ値	切片	残差二乗和	原曲	ゆらぎ値-10%	ゆらぎ値+10%	残差×0.8	残差×1.2
1	Another_Sky.wav	1.715343	7.704218	0.414398	10	9	5	4	6
2	ロンドンデリーエア.wav	1.627244	7.080130	0.398816	6	5	4	9	10
3	君を信じて.wav	2.110503	9.850914	0.466082	2	8	7	7	10
4	漂流.wav	1.592919	6.873790	0.430658	6	10	3	12	4
5	Down by the Riverside.wav	1.724475	8.052250	0.448193	6	6	4	9	8
6	Space_Odessey3_Revelation.wav	0.990190	3.543654	0.323250	8	5	2	11	8
7	TOMORROW.wav	2.267830	11.134446	0.555204	6	6	3	9	10
8	古いフランスの歌.wav	1.398202	5.805316	0.433054	2	5	11	10	6
9	自由.wav	2.172833	11.011055	0.588947	12	7	1	6	8
10	赤い川の谷間(brass).wav	1.520269	7.336829	0.526494	8	4	9	7	5

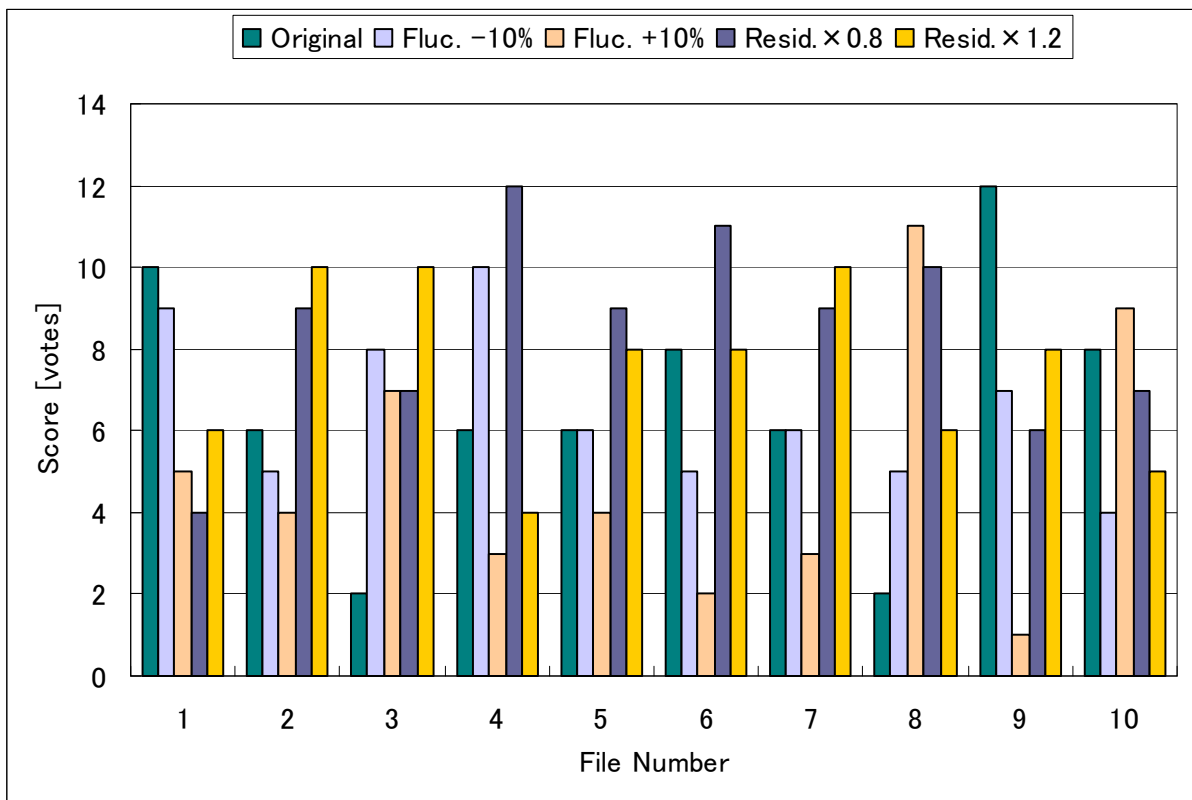


図 4.23 随伴パラメータの変更と投票数 (Fluc. : ゆらぎ値, Resid. : 残差二乗和) [18]

即ち、10 曲中 8 曲（楽曲番号：2～8, 10）において最高投票数を得た曲のバリエーションは、ゆらぎ値随伴量を変更したものであった。この内、平均との有意差があるのは 3 曲（楽曲番号：4, 6, 8）であり、これらの原曲とも差があることが分かる。

続いて、4.3 節に示す(4.5)式の重回帰式を用いて投票数に対するゆらぎ値随伴量の関係を調査した結果、重相関係数が 0.918 となり、この重回帰式で評価の約 91.8% を説明できることが分かった。次に重回帰式の優位性を確かめるために分散分析を行い、5%の有意水準で有意であるという結果になった。重回帰式の各係数に着目すれば、ゆらぎ値の係数 $\beta_1=3.20$ 、残差二乗和の係数 $\beta_2=2.94$  となり、総合評価に対するゆらぎ値と残差二乗和の影響は、ほぼ同程度であることが分かった。

## 第 5 章

# HMGD 処理の試験的適用

### 5.1. ヒストグラムマッチングの原理

HMGD (Histogram Matching based on Gaussian Distribution) 処理[13, 14, 20]はヒストグラムマッチング[22-28]を利用した感性的印象効果を高める自動補正処理の一つであり, ヒストグラムをガウス分布にマッチングさせる処理である. 我々は, この処理が音声信号及び画像信号の両方に有効ではないかと考え, HMGD 処理を音声信号の振幅値, 及び画像信号の輝度軸に対して試験的に適用し, 処理前と HMGD 処理後の感性的印象効果を比較調査した.

HMGD 処理はヒストグラムの伸縮変換によって行う事が出来る. 始めに, 音声信号の振幅値及び画像信号の輝度値  $x, y$  に対する確率密度関数を  $f(x), g(y)$  と定義する. 更に,  $x$  と  $y$  からなる連続かつ単調増加な関数  $y=\phi(x)$  を定義すれば,  $f(x)$  及び  $g(y)$  の関係は図 5.1(a) で示される.

図 5.1(a) より,  $x$  の取り得る値の範囲が  $0 \sim L$  である為,  $y$  の取り得る値の範囲は  $0 \sim \phi(L)$  である事が分かる. ここで確率を  $P$  と定義し,  $f(x), g(y)$  の累積確率密度関数を (5.1) 式及び (5.2) 式で定義する. 更に,  $y_0 = \phi(x_0), y_0 + dx = \phi(x_0 + dx)$  と置けば (5.3) 式が導かれる.

$$P(0 \leq x \leq L) = \int_{x=0}^{x=L} f(x) dx = 1 \quad (5.1)$$

$$P(0 \leq y \leq \phi(L)) = \int_{y=0}^{y=\phi(L)} g(y)dy = 1 \quad (5.2)$$

$$\begin{aligned} f(x)dx &= P(x_0 \leq x \leq x_0 + dx) \\ &= P(\phi(x_0) \leq y \leq \phi(x_0 + dx)) \\ &= P(y_0 \leq y \leq y_0 + dy) \\ &= g(y)dy \end{aligned} \quad (5.3)$$

よって、以下に示す (5.4) 式が成り立ち、(5.5) 式が導かれる。

$$f(x)dx = g(y)dy = g(y)\phi'(x)dx \quad (5.4)$$

$$f(x) = g(y)\phi'(x) \quad (5.5)$$

(5.5) 式より、 $\phi(x)$  と  $g(y)$  が既知であるならば、 $f(x)$  は一意に定まる事が分かる。即ち、 $\phi(x)$  はヒストグラムの変換関数である。

例えば、ヒストグラムイコライゼーションでは  $\phi(x)$  を (5.6) 式として設定する。

$$\phi(x) = L \int_0^x f(x)dx \quad (5.6)$$

(5.6) 式の両辺を微分すれば、

$$\phi'(x) = Lf(x) \quad (5.7)$$

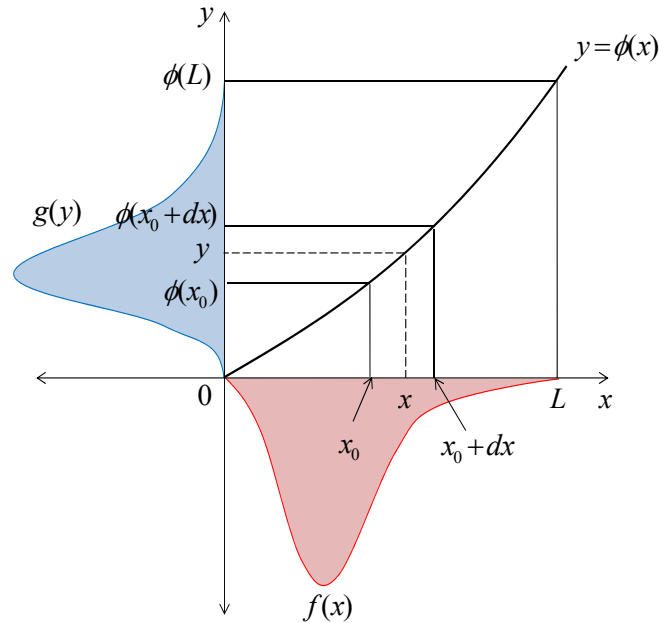
である。従って (5.7) 式を (5.5) 式に代入すれば、以下の (5.8) 式が得られる。

$$f(x) = g(y)\phi'(x) = g(y)Lf(x) \quad (5.8)$$

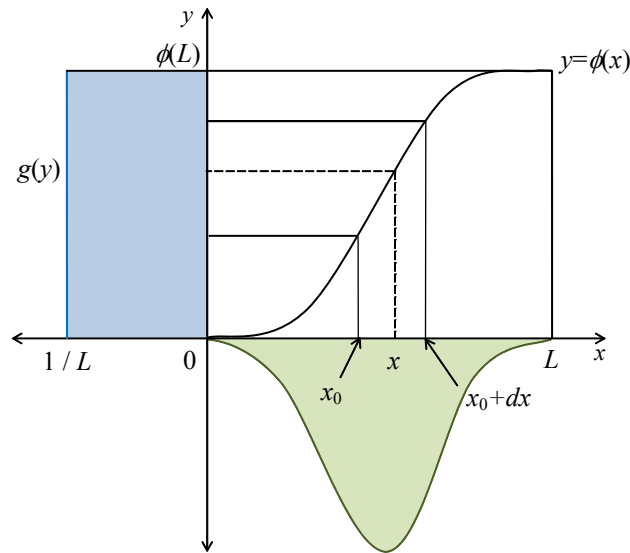
よって、(5.8) 式より以下の (5.9) 式が求められ、 $g(y)$  は図 5.1 (b) に示す様に一様分布となる。つまり、 $g(y)$  は  $f(x)$  にヒストグラムイコライゼーション処理を施した結果となる事が分かる。

$$g(y)L = 1, \quad g(y) = 1/L \quad (5.9)$$

即ち，変換関数  $\phi(x)$  を，平均  $\mu$ ，分散  $\sigma^2$  のガウス分布の累積確率密度関数とすれば，HMGD 処理を行える事が分かる．



(a)  $y=\phi(x)$ ,  $f(x)$  及び  $g(y)$  の関係図



(b) ヒストグラムイコライゼーション

図 5.1 ヒストグラムマッチングの概念図 [13, 14, 20]

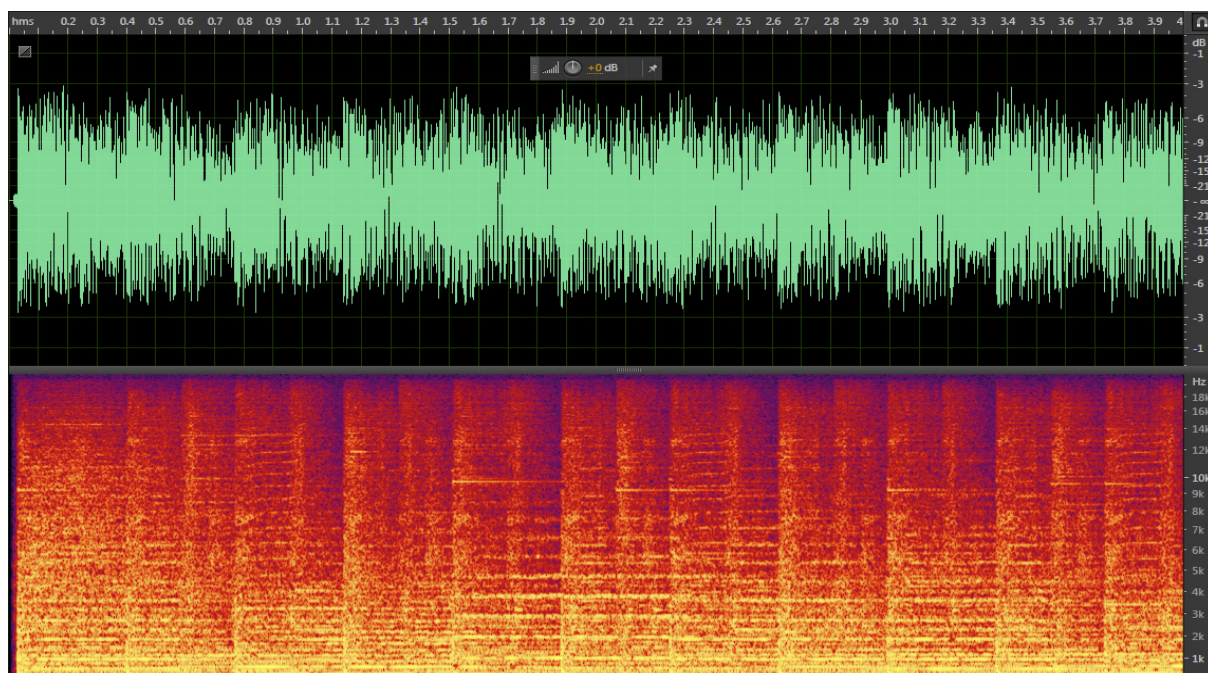
## 5.2. 音声信号への適用

音声信号の振幅値のヒストグラムから累積ヒストグラムを求め、HMGD 処理を行った。音声信号のフォーマットはサンプリング周波数 44.1kHz, 8bit である。HMGD 処理でマッチングさせるガウス分布は文献[29]に基づき、 $\sigma = 50$ ,  $\mu = 128$  とし試験的に処理を施したが、音割れを抑える為  $\sigma = 40$  に変更して実験を行った。

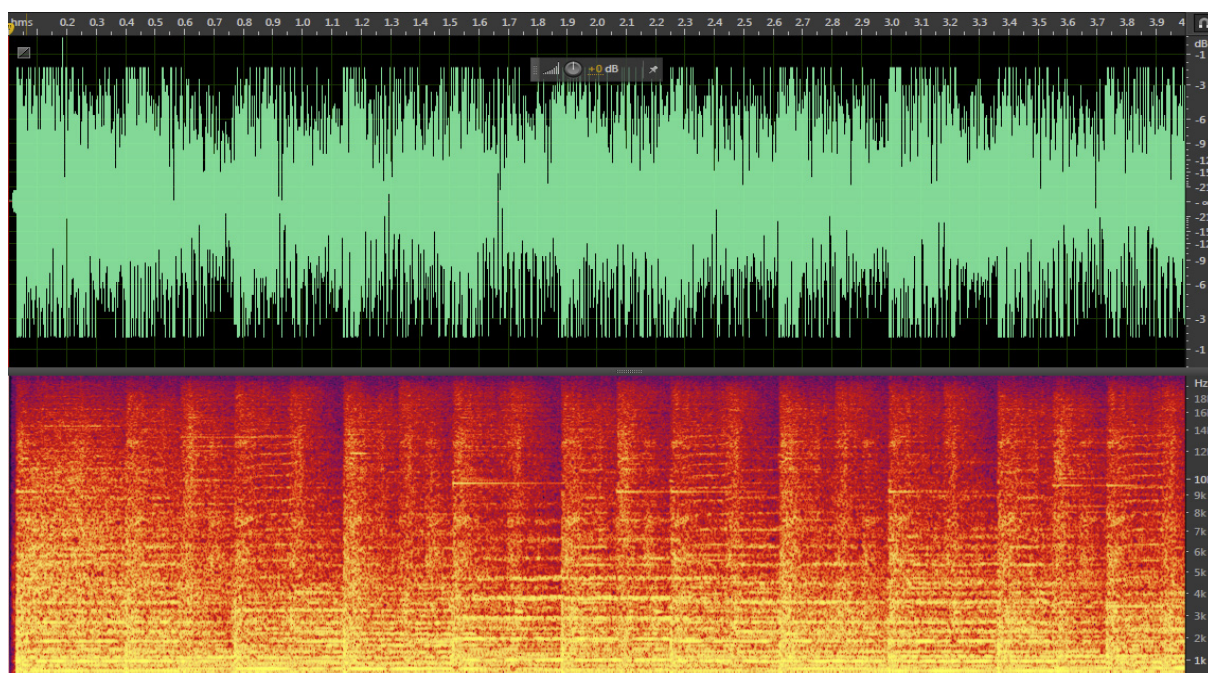
図 5.2 に元音声、及び HMGD 処理音声の波形及びスペクトログラムを示し、図 5.3 に元音声および HMGD 音声のヒストグラムを示す。音声波形及びスペクトログラムの観測には Adobe Audition CS6 を用いている。尚、スペクトログラムは、周波数成分の強度を視覚化したものである。図 5.2 では黒色に近づく個所の周波数成分が弱く、黄色に近づく個所の周波数成分が強い事を表している。

図 5.2 の音声波形を比較すれば、図 5.2(b) の HMGD 処理音声波形は図 5.2(a) の元音声波形より振幅が増加している事が分かる。また、図 5.2 のスペクトログラムを比較すれば、図 5.2(b) の HMGD 処理音声のスペクトログラムは、図 5.2(a) の元音声のスペクトログラムよりも全ての周波数において黄色い部分が増加している事が分かる。更に図 5.3 のヒストグラムを比較すれば、図 5.3(b) の HMGD 音声のヒストグラムは図 5.3(a) の元音声のヒストグラムと比較して、振幅値のダイナミックレンジが広がっている事が分かる。即ち、音声信号における HMGD は振幅の増加のみならず、全ての周波数における成分を強調する効果があると考えられる。また、ゆらぎ値及び残差二乗和も変化している為、音声信号の HMGD 処理によって感性的印象も変化する事が推測される。





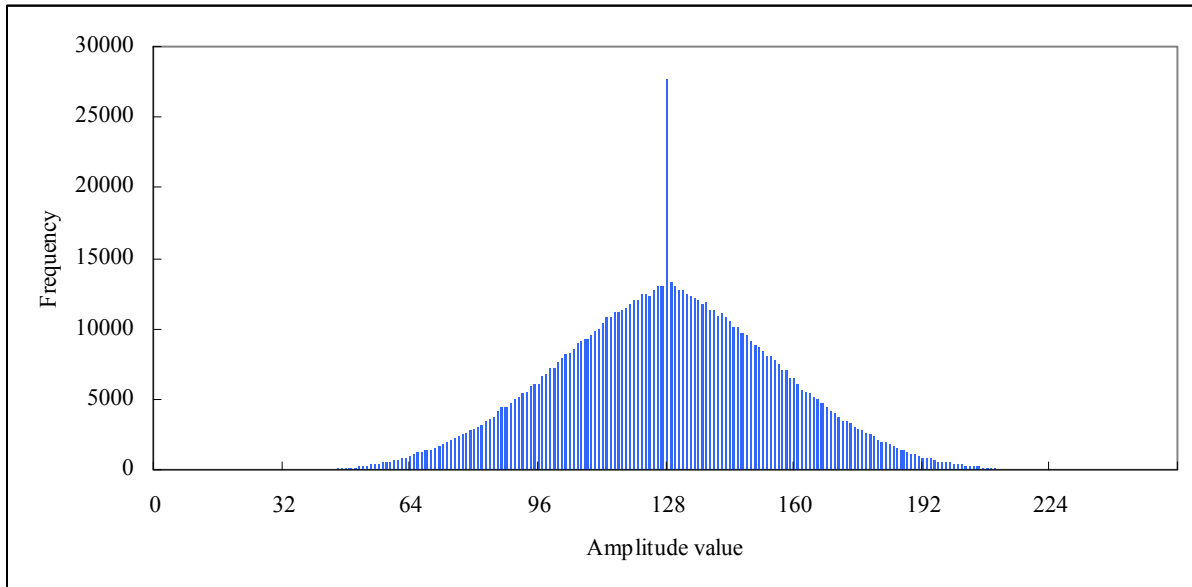
(a) 元音声 (ゆらぎ値 : 2.132, 残差二乗和 : 0.673)



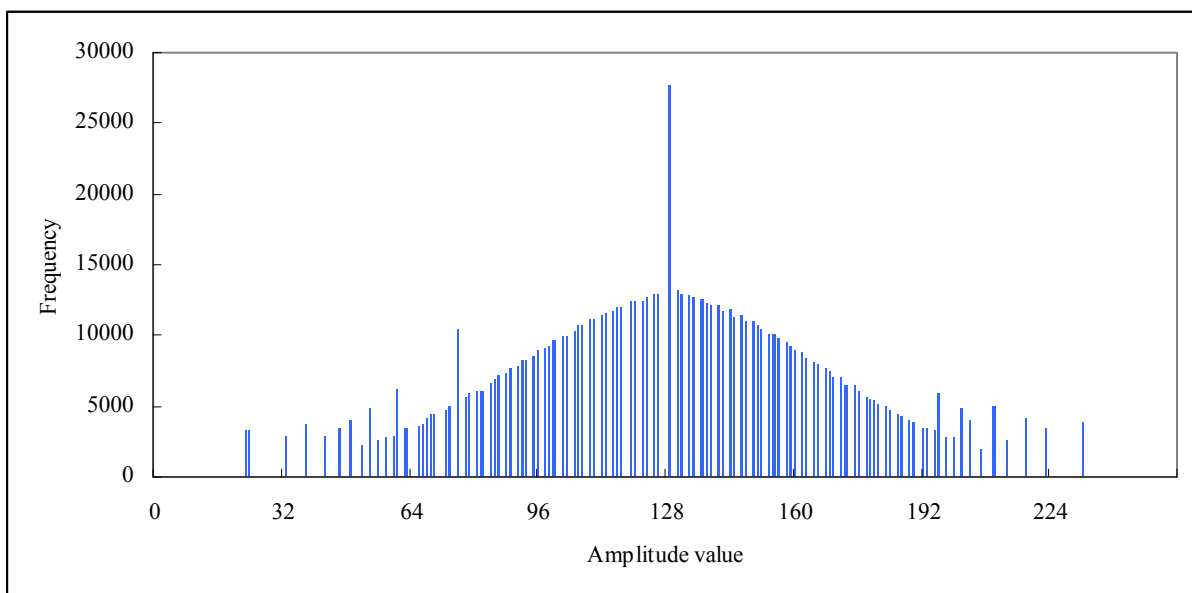
(b) HMGD 処理音声 ( $\sigma=40$ ,  $\mu=128$ , ゆらぎ値 : 2.081, 残差二乗和 : 0.614)

図 5.2 音声波形 (各図上段) 及びスペクトログラム (各図下段)

(お願いヴァレンティヌ.wav)



(a) 元音声 (ゆらぎ値 : 2.132, 残差二乗和 : 0.673)



(b) HMGD 処理音声 ( $\sigma=40$ ,  $\mu=128$ , ゆらぎ値 : 2.081, 残差二乗和 : 0.614)

図 5.3 元音声及び HMGD 処理音声のヒストグラム (お願いヴァレンティヌ.wav)

### 5.3. カラー画像への適用

一方、カラー画像に対する HMGD 処理では画像の輝度のヒストグラムに対して音声信号の場合と同様に HMGD 処理を行った。但し、HMGD 処理でマッチングさせるガウス分布は文献[29]に基づき、 $\sigma = 50$ 、 $\mu = 128$ としている。

図 5.4~5.6 に処理前の画像（元画像）、HMGD 処理画像、及び元画像のヒストグラムを示す。尚、ヒストグラム（図 5.4(c)、図 5.5(c)、図 5.6(c)）の赤線は、ヒストグラムのピーク部分を示す。

図 5.4 では、HMGD 処理によって陰部の輝度を上げ、視認性を高めている事が分かる。更に図 5.4(c)より、この画像のピークは単一であることも分かる。

図 5.5 及び図 5.6 では、HMGD 処理によって立体感が薄れたり、色味が不自然になったりする事が分かり、HMGD 処理が効果的に機能していないと考えられる。図 5.5(c) 及び図 5.6(c)より、これらの画像のヒストグラムのピークは複数である事が分かる。

即ち、カラー画像の輝度における HMGD 処理では、HMGD 処理前にヒストグラムのピークが単一か否かを識別する処理方法を開発する必要がある、HMGD 処理はヒストグラムのピークが単一である画像に対して適用する事が望ましいと考えられる。

更にカラー画像における HMGD 処理を行った際の感性的印象の変化を客観的に分析する為、質問紙調査を実施した。調査方法は ITU-T P. 910[30]に準拠している。但し本調査の評価対象は静止画像である為、評価項目を静止画像に関する以下の 5 項目に限定している。

- (A) 明るさ
- (B) コントラスト
- (C) 色
- (D) 輪郭
- (E) 主印象

被験者は 20 代の大学生 36 名であり，3 章及び 4 章における音声信号の調査で用いた被験者と重複していない。

被験者には元画像及び HMGD 処理画像を同時に提示し，各々の画像における各評価項目を 1.0～5.0 点の範囲で評価させた。

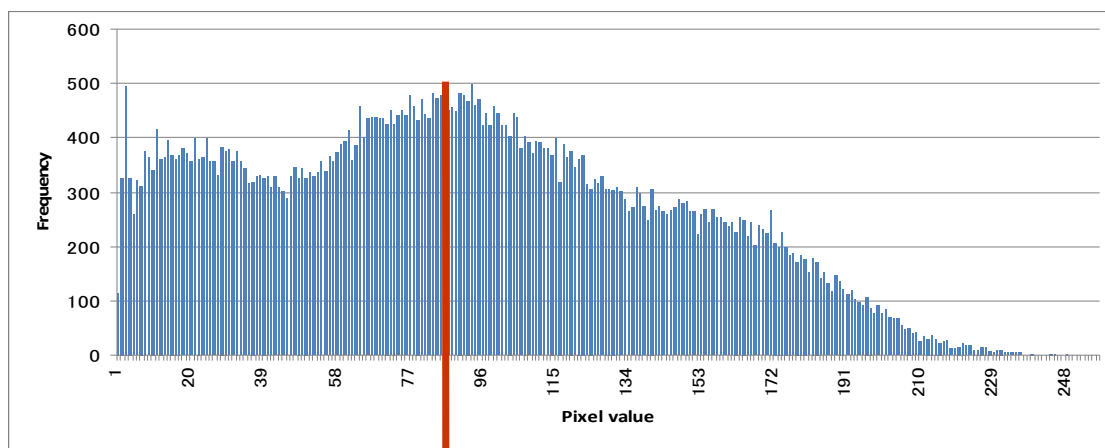
調査に用いた質問紙を図 5.7～5.11 に示し，調査に用いた元画像及び HMGD 処理画像を図 5.12～5.15 に示す。また，図 5.16 に単峰性の画像 (図 5.12(b), (c), 図 5.13(a), 図 5.14(a)) に対する調査結果を示し，図 5.17 に多峰性の画像 (図 5.12(a), 図 5.13(b), (c), 図 5.14(a), 図 5.15) に対する調査結果を示す。図 5.16～5.17 に示す調査結果は，被験者 36 人が付けた評点の平均値である。また，図中において赤丸を付けた個所は，元画像と HMGD 処理画像の平均値の有意差を有意水準 5% の t 検定 [19] を行い，有意差が認められた評価項目を示す。

図 5.16 より，単峰性の元画像に対して HMGD 処理を行うと，色及び主印象の評点が向上していることが分かる。また，明るさ及びコントラストの評点も向上する傾向が見られる。一方，図 5.17 より多峰性の元画像に対して HMGD 処理を行うと，色，輪郭，主印象の評点が低下する事が分かる。更に明るさとコントラストの評点も低下する傾向が見られる。従って，HMGD 処理は単峰性の画像の感性的印象を向上させる効果があると考えられる。



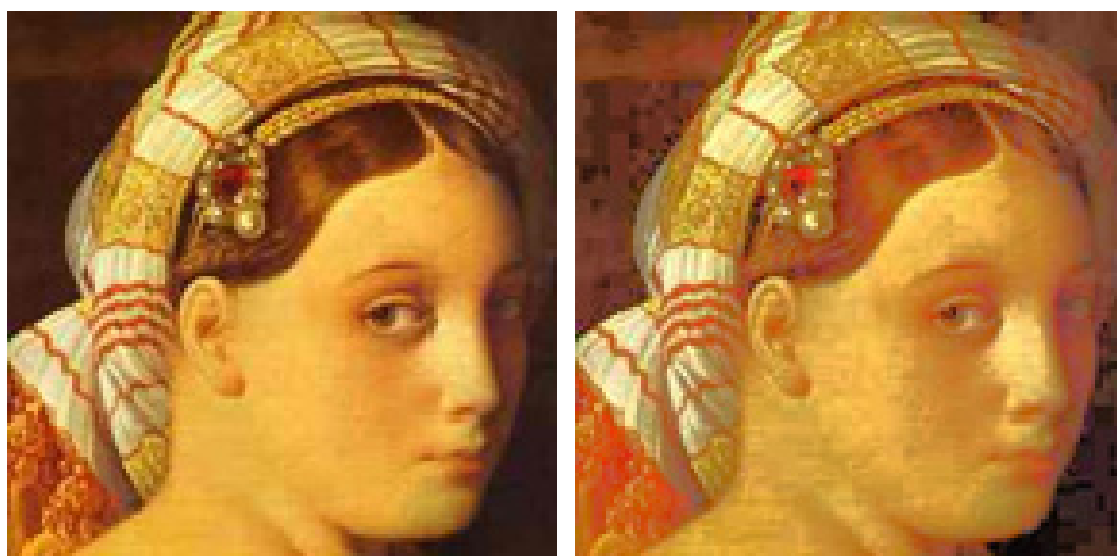
(a) 元画像

(b) HMGD 処理画像



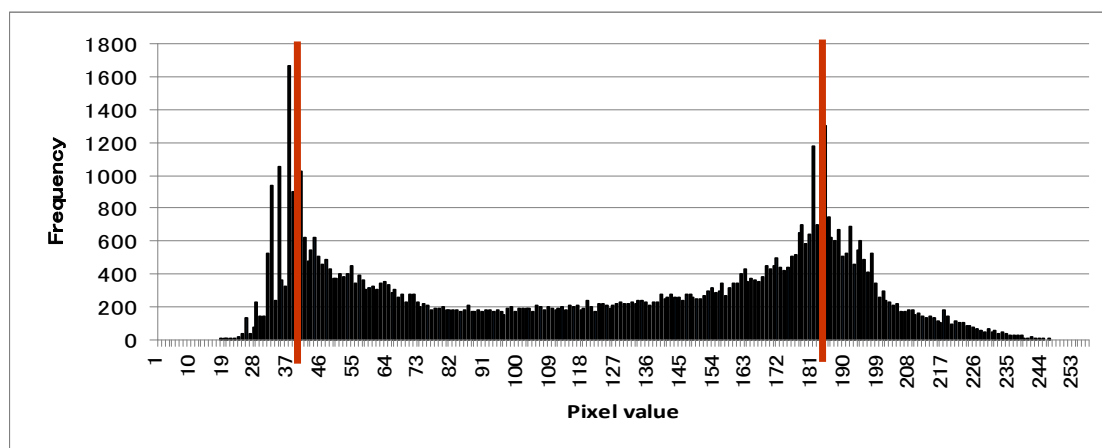
(c) 元画像のヒストグラム

図 5.4 カラー画像に対する HMGD 処理(1) [13, 20]



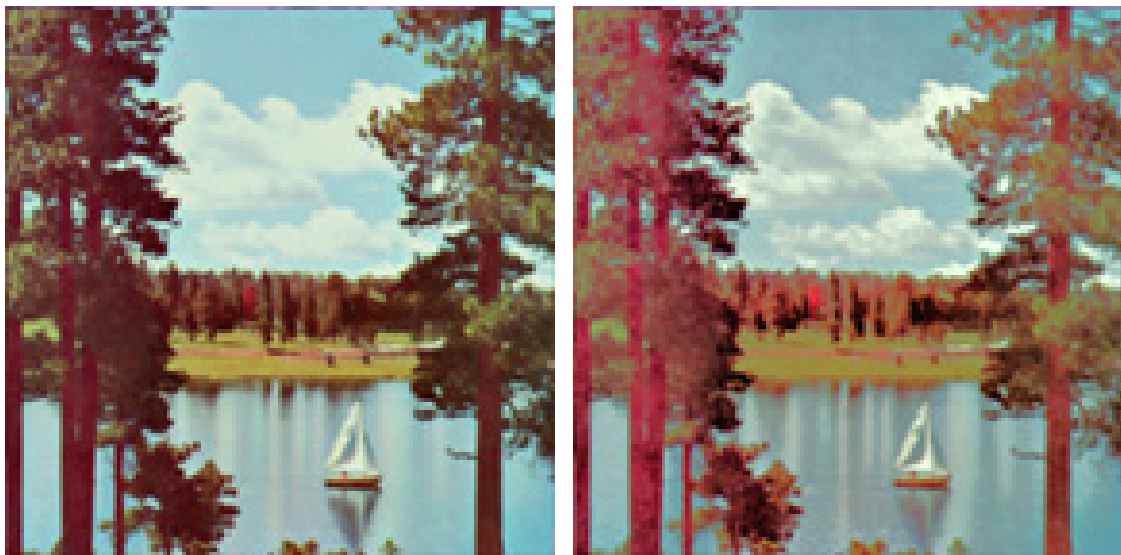
(a) 元画像

(b) HMGD 処理画像



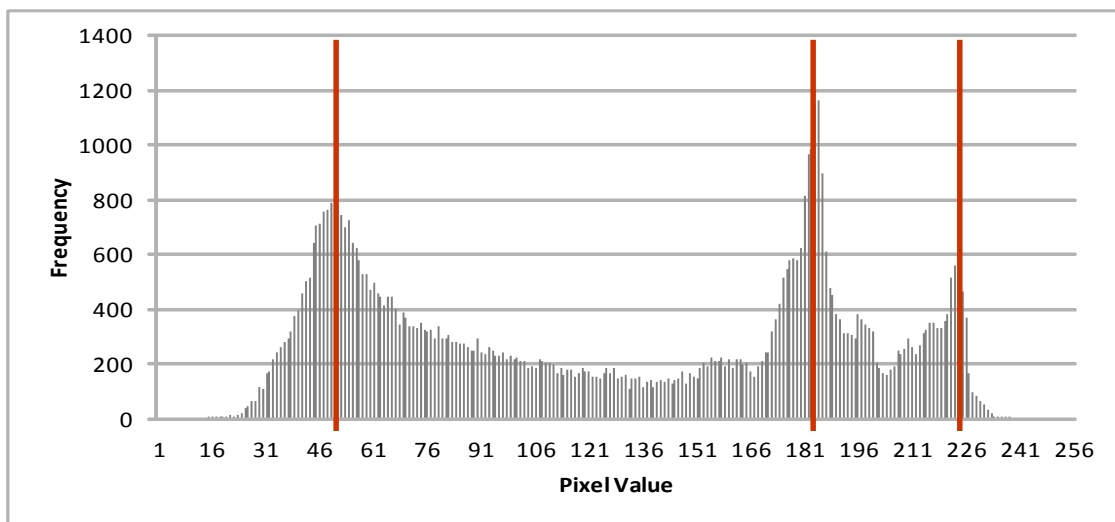
(c) 元画像のヒストグラム

図 5.5 カラー画像に対する HMGD 処理(2) [13]



(a) 元画像

(b) HMGD 処理画像



(c) 元画像のヒストグラム

図 5.6 カラー画像に対する HMGD 処理(3) [13]

画像の評価についてのアンケート（説明）

このアンケートは、画像の印象を感性的に評価するためのアンケートです。Web サイトにて 10 組の画像を提示しますので、以下の要領でアンケート回答用紙に記入してください。

※アンケート用紙ご記入にあたってのお願い

- ・あまり考え込まず、直感で回答してください。
- ・画像の好き嫌いでなく、画質を評価してください。
- ・評価は 1.0～5.0 まであります。以下の表が点数の目安です。

Quality	
5	Excellent
4	Good
3	Fair
2	Poor
1	Bad

<記入例>

・明るさ    1            2            3            4            5

暗い |-----| 明るい

・コントラスト

おだやか |-----| 鮮やか

A	B
2.3	3.2
3.5	1.1

<手順>

1. こちらから指定する Web サイトに Internet Explorer を用いてアクセスしてください。
2. 指示がありましたら、『開始』ボタンを押して始めてください。
3. A、B の 2 つの画像が表示されますので、それぞれの画像についての画質を解答用紙に記入してください。
4. 3.が終わったら、『次へ』ボタンをクリックしてください。
5. 3.～4.を繰り返してください。10 組目の画像の評価が終わったら、『終了』ボタンをクリックしてください。

図 5.7 画像に関する感性的印象評価の質問紙（1/5 頁）



画像の評価についてのアンケート

学籍番号 (                      ) 名前 (                      )

<1 組目>

	1	2	3	4	5			
・明るさ	----- ----- ----- ----- -----					暗い		
・コントラスト	----- ----- ----- ----- -----					おだやか		
・色	----- ----- ----- ----- -----					不自然な		
・輪郭	----- ----- ----- ----- -----					あいまい		
・主印象	----- ----- ----- ----- -----					悪い		
						明るい		
						鮮やか		
						自然な		
						はっきり		
						良い		

<2 組目>

	1	2	3	4	5			
・明るさ	----- ----- ----- ----- -----					暗い		
・コントラスト	----- ----- ----- ----- -----					おだやか		
・色	----- ----- ----- ----- -----					不自然な		
・輪郭	----- ----- ----- ----- -----					あいまい		
・主印象	----- ----- ----- ----- -----					悪い		
						明るい		
						鮮やか		
						自然な		
						はっきり		
						良い		

<3 組目>

	1	2	3	4	5			
・明るさ	----- ----- ----- ----- -----					暗い		
・コントラスト	----- ----- ----- ----- -----					おだやか		
・色	----- ----- ----- ----- -----					不自然な		
・輪郭	----- ----- ----- ----- -----					あいまい		
・主印象	----- ----- ----- ----- -----					悪い		
						明るい		
						鮮やか		
						自然な		
						はっきり		
						良い		

図 5.8 画像に関する感性的印象評価の質問紙 (2/5 頁)

<4 組目>

	1	2	3	4	5		
・明るさ	----- ----- ----- ----- -----					暗い	明るい
・コントラスト	----- ----- ----- ----- -----					おだやか	鮮やか
・色	----- ----- ----- ----- -----					不自然な	自然な
・輪郭	----- ----- ----- ----- -----					あいまい	はっきり
・主印象	----- ----- ----- ----- -----					悪い	良い

	A	B

<5 組目>

	1	2	3	4	5		
・明るさ	----- ----- ----- ----- -----					暗い	明るい
・コントラスト	----- ----- ----- ----- -----					おだやか	鮮やか
・色	----- ----- ----- ----- -----					不自然な	自然な
・輪郭	----- ----- ----- ----- -----					あいまい	はっきり
・主印象	----- ----- ----- ----- -----					悪い	良い

	A	B

<6 組目>

	1	2	3	4	5		
・明るさ	----- ----- ----- ----- -----					暗い	明るい
・コントラスト	----- ----- ----- ----- -----					おだやか	鮮やか
・色	----- ----- ----- ----- -----					不自然な	自然な
・輪郭	----- ----- ----- ----- -----					あいまい	はっきり
・主印象	----- ----- ----- ----- -----					悪い	良い

	A	B

図 5.9 画像に関する感性的印象評価の質問紙 (3/5 頁)

<7 組目>

	1	2	3	4	5	
・明るさ	----- ----- ----- ----- -----					
	暗い				明るい	
・コントラスト	----- ----- ----- ----- -----					
	おだやか				鮮やか	
・色	----- ----- ----- ----- -----					
	不自然な				自然な	
・輪郭	----- ----- ----- ----- -----					
	あいまい				はっきり	
・主印象	----- ----- ----- ----- -----					
	悪い				良い	

A	B

<8 組目>

	1	2	3	4	5	
・明るさ	----- ----- ----- ----- -----					
	暗い				明るい	
・コントラスト	----- ----- ----- ----- -----					
	おだやか				鮮やか	
・色	----- ----- ----- ----- -----					
	不自然な				自然な	
・輪郭	----- ----- ----- ----- -----					
	あいまい				はっきり	
・主印象	----- ----- ----- ----- -----					
	悪い				良い	

A	B

<9 組目>

	1	2	3	4	5	
・明るさ	----- ----- ----- ----- -----					
	暗い				明るい	
・コントラスト	----- ----- ----- ----- -----					
	おだやか				鮮やか	
・色	----- ----- ----- ----- -----					
	不自然な				自然な	
・輪郭	----- ----- ----- ----- -----					
	あいまい				はっきり	
・主印象	----- ----- ----- ----- -----					
	悪い				良い	

A	B

図 5.10 画像に関する感性的印象評価の質問紙 (4/5 頁)

<10 組目>

・明るさ	1	2	3	4	5	
	暗い					明るい
・コントラスト						
	おだやか					鮮やか
・色						
	不自然な					自然な
・輪郭						
	あいまい					はっきり
・主印象						
	悪い					良い

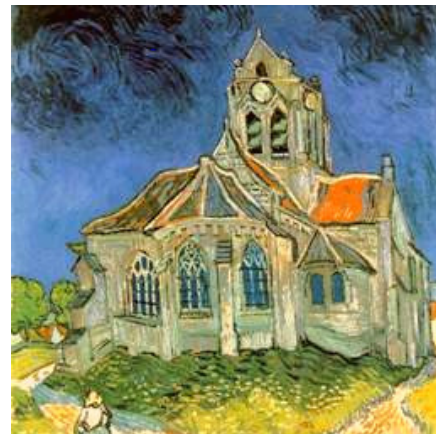
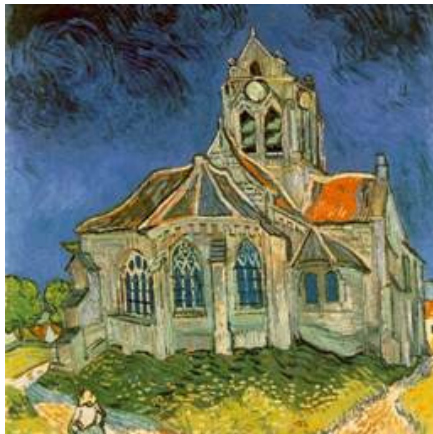
A	B

アンケートは以上です。ご協力、ありがとうございました。

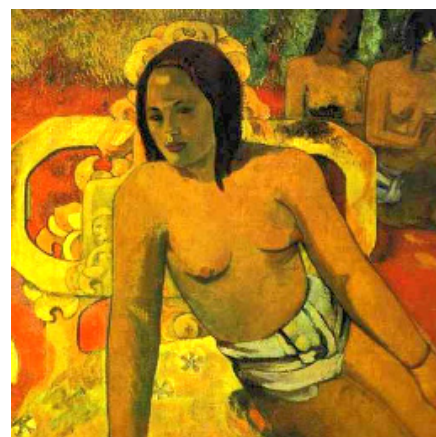
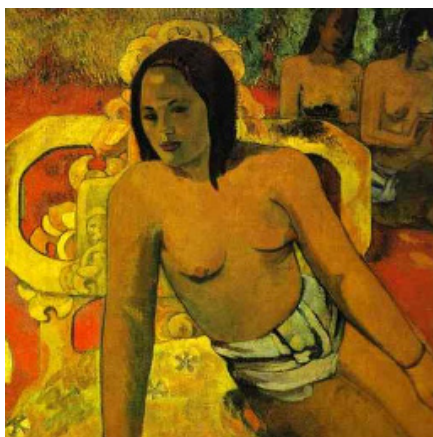
図 5.11 画像に関する感性的印象評価の質問紙 (5/5 頁)



(a) 画像 1 (左 : 元画像, 右 : HMGD 処理画像)



(b) 画像 2 (左 : 元画像, 右 : HMGD 処理画像)



(c) 画像 3 (左 : 元画像, 右 : HMGD 処理画像)

図 5.12 感性的印象評価に関する調査に用いた画像 (1/4)





(a) 画像 4 (左：元画像, 右：HMGD 処理画像)

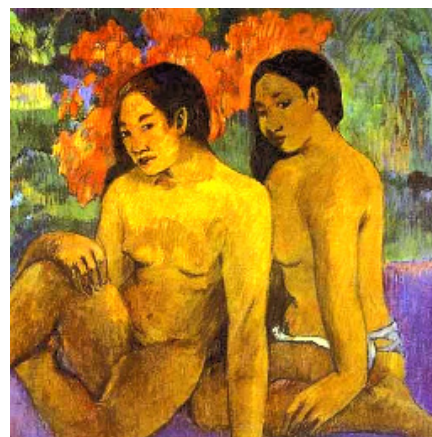
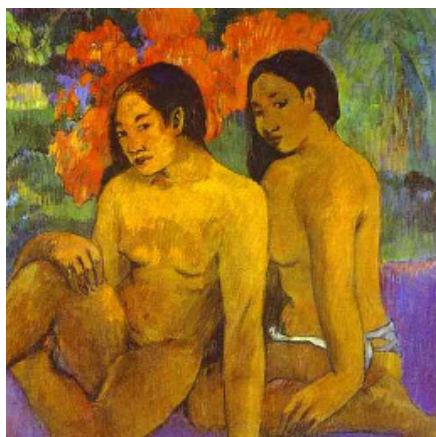


(b) 画像 5 (左：元画像, 右：HMGD 処理画像)



(c) 画像 6 (左：元画像, 右：HMGD 処理画像)

図 5.13 感性的印象評価に関する調査に用いた画像 (2/4)



(a) 画像 7 (左：元画像, 右：HMGD 処理画像)



(b) 画像 8 (左：元画像, 右：HMGD 処理画像)



(c) 画像 9 (左：元画像, 右：HMGD 処理画像)

図 5.14 感性的印象評価に関する調査に用いた画像 (3/4)



図 5.15 感性的印象評価に関する調査に用いた画像 (4/4)  
 画像 10 (左：元画像, 右：HMGD 処理画像)

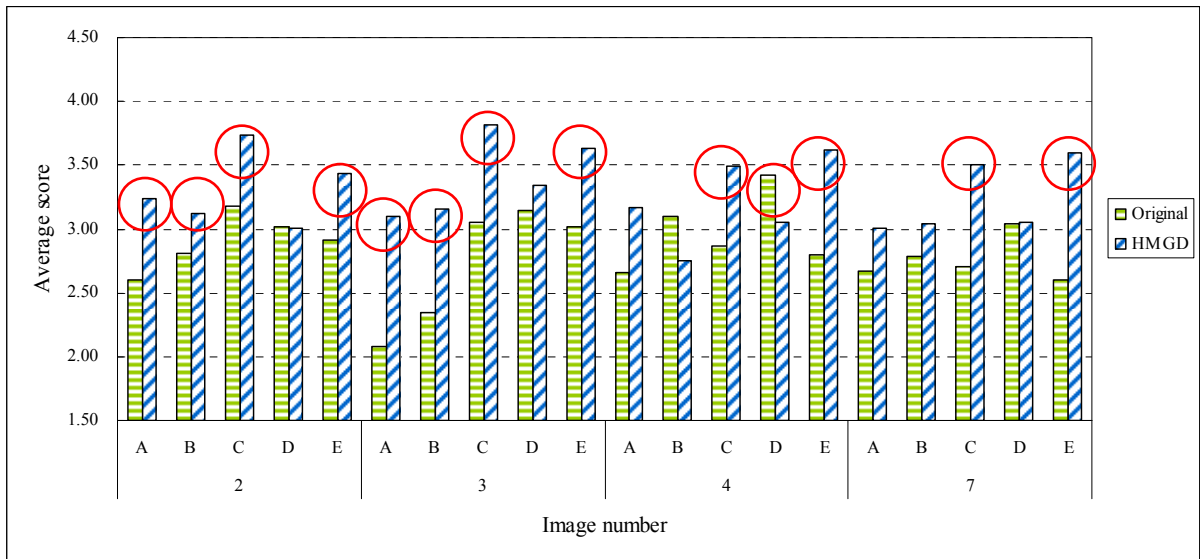
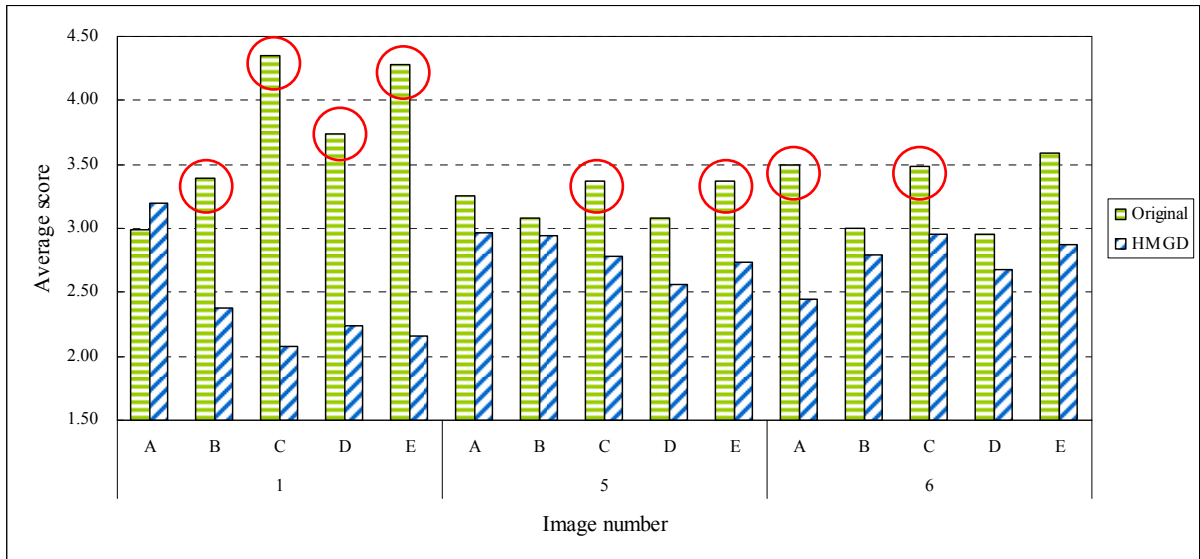
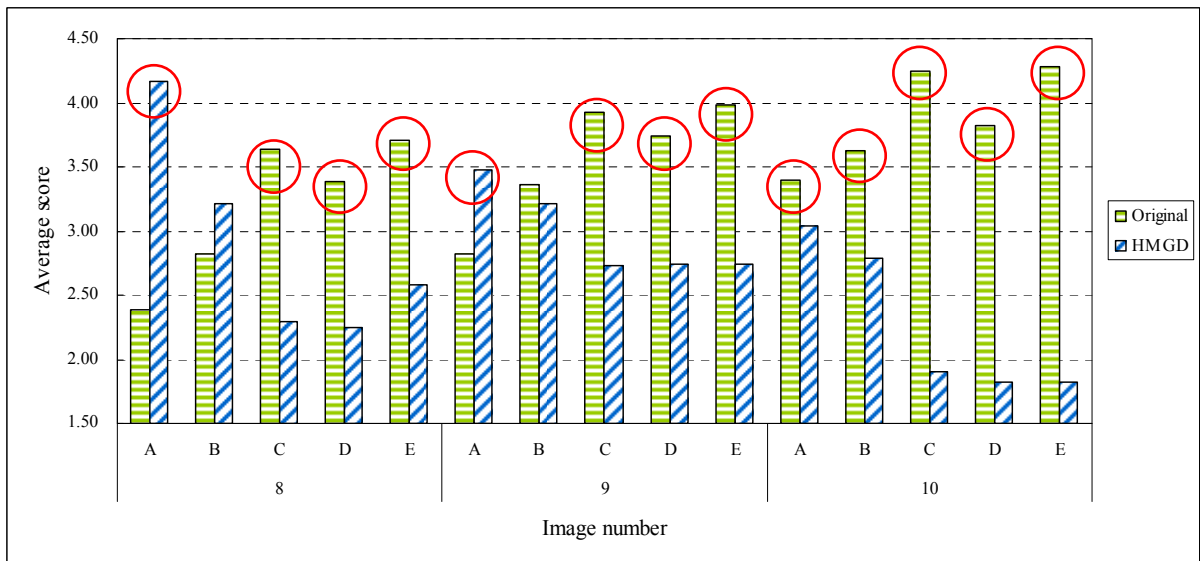


図 5.16 画像の感性的印象評価に関する調査結果 (単峰性)  
 A：明るさ, B：コントラスト, C：色, D：輪郭, E：主印象





(a)



(b)

図 5.17 画像の感性的印象評価に関する調査結果（単峰性）

A：明るさ，B：コントラスト，C：色，D：輪郭，E：主印象

#### 5.4. 曲率によるヒストグラムピークの自動検出

5.2 節及び 5.3 節で述べたように、音声及び画像信号に対する HMGD 処理は、ヒストグラムが単峰性であれば感性的印象の向上が見込まれる事が分かった。そこで本節は、ヒストグラムのピーク数を曲率の近似解を求める事で自動検出する手法を述べる。

始めに、 $y$  を  $x$  の関数とするとき、曲率  $R$  は以下の (5.10) 式で定義される [14, 20]。従って、近似解を求めるには関数  $y$  の微分及び 2 階微分を求めればよい事が分かる。

$$R = \frac{\left(\frac{d^2y}{dx^2}\right)}{\left\{1 + \left(\frac{dy}{dx}\right)^2\right\}^{\frac{3}{2}}} \quad (5.10)$$

続いて、 $g(x)$  を (5.11) 式に示すように、分散  $\sigma^2$ 、平均  $a$  におけるガウス分布の確率密度関数として定義する。尚  $g(x)$  は音声信号のヒストグラムにおける振幅値、或いはカラー画像の輝度値  $0 \sim L$  と対応する。また、定数  $K$  は (5.12) 式を満たすとする。

$$g(x) = \frac{K}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-a)^2}{2\sigma^2}\right\} \quad (5.11)$$

$$\frac{K}{\sigma\sqrt{2\pi}} \int_0^L \exp\left\{-\frac{(x-a)^2}{2\sigma^2}\right\} du = 1 \quad (5.12)$$

ここで  $y = f(x)$  を累積ヒストグラムとすれば、 $f(x)$  は以下のように表す事が出来る。

$$y = f(x) = \int_0^x g(u) du = \frac{K}{\sigma\sqrt{2\pi}} \int_0^x \exp\left\{-\frac{(u-a)^2}{2\sigma^2}\right\} du \quad (5.13)$$

従って、(5.13) 式の  $x$  における微分は (5.14) 式となる。

$$\frac{dy}{dx} = \frac{K}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-a)^2}{2\sigma^2}\right\} = g(x) \quad (5.14)$$

更に (5.13) 式の 2 階微分を計算すれば、

$$\begin{aligned}\frac{d^2y}{dx^2} &= \frac{d}{dx} g(x) = \frac{K}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-a)^2}{2\sigma^2}\right\} \cdot \left(-\frac{1}{2\sigma^2}\right) \cdot 2(x-a) \\ &= \frac{K}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-a)^2}{2\sigma^2}\right\} \cdot \left\{-\frac{(x-a)}{\sigma^2}\right\}\end{aligned}\quad (5.15)$$

である。(5.15)式には(5.11)式で定義した  $g(x)$  が含まれているため、整理すると以下の(5.16)式が得られる。

$$\frac{d^2y}{dx^2} = \frac{d}{dx} g(x) = \frac{K}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-a)^2}{2\sigma^2}\right\} \cdot \left\{-\frac{(x-a)}{\sigma^2}\right\} = g(x) \left(\frac{a-x}{\sigma^2}\right) \quad (5.16)$$

(5.14)式は  $y$  の微分であり、(5.16)式は  $y$  の 2 階微分であることは明らかである。従って、これらの式を(5.10)式に代入すれば、曲率の近似解として(5.17)式が得られる。

$$R = \frac{g(x) \left(\frac{a-x}{\sigma^2}\right)}{\left\{1 + g(x)^2\right\}^{\frac{3}{2}}} \approx \frac{a-x}{\sigma^2} g(x) \quad (5.17)$$

(5.17)式より、曲率  $R$  の正負は変数  $x$  によって決定される事が分かる。 $R$  の値としては以下の 3 通りの場合が考えられる。

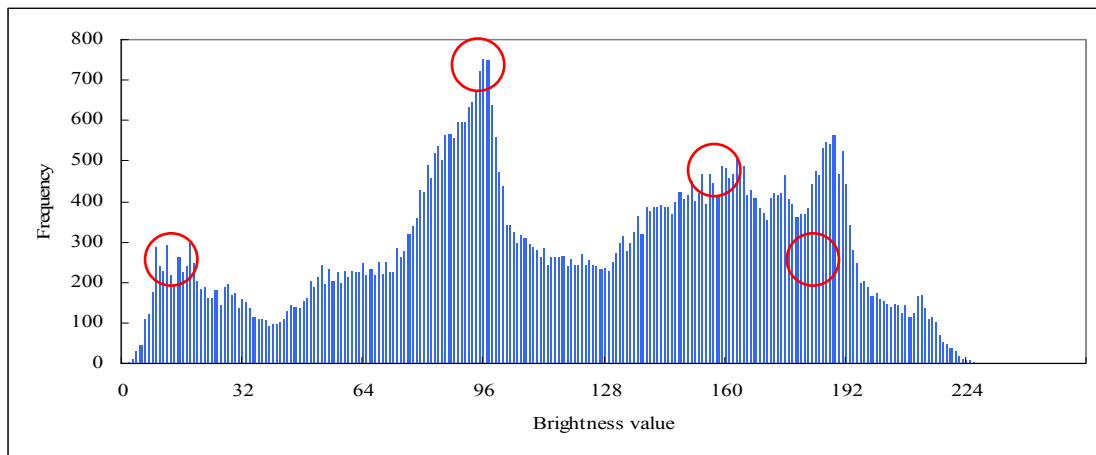
- (i)  $x < a \quad \dots \quad R > 0$
- (ii)  $x = a \quad \dots \quad R = 0$
- (iii)  $x > a \quad \dots \quad R < 0$

即ち、ヒストグラムのピークとなる個所では曲率が 0 である事が分かる。よって、曲率が正から負になる回数を検出するとヒストグラムのピーク数が求められる。

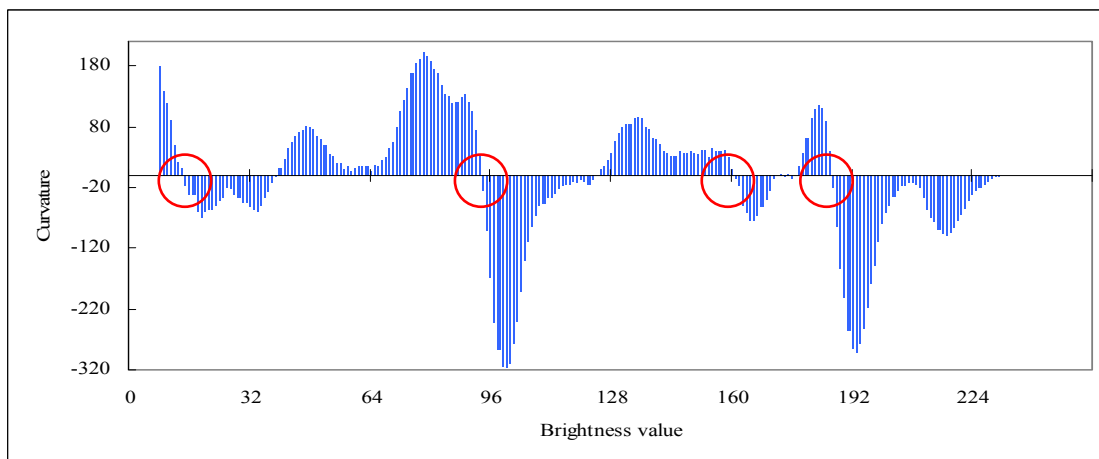
図 5.18 にカラー画像のヒストグラムに対し曲率計算を用い、ピーク数を検出した例を示す。図 5.18(c) の曲率を示したヒストグラムを見れば、図 5.18(b) に示すヒストグラムのピーク部分 (○) で曲率が正から負へと大きく変化している事が分かる。



(a) 元画像



(b) 元画像のヒストグラム



(c) 元画像の曲率ヒストグラム (ピーク数: 4)

図 5.18 カラー画像におけるヒストグラムの曲率算出例

## 第 6 章

### 結論

本論文では、音声信号からのゆらぎ値を求める際に同時算出されるゆらぎ値、切片、残差二乗和のゆらぎ値随伴量を感性的特徴量として捉え、それらと感性的印象との対応関係について調査した。調査方法としては質問紙を用いた試聴実験を行い、ゆらぎ値随伴量と楽曲の感性的印象との関係及び、ゆらぎ値、残差二乗和を各々独立且つ人為的に変更した場合の感性的印象の変化について分析を行った。また、音声信号に対する感性的印象を詳細に分析するため、音声信号に対する被験者の感性的印象の傾向を Ward 法によってクラスタ化し、グループ毎に音声信号のゆらぎ値算出随伴量と感性的印象についても分析した。

更に、メディア信号の感性的印象が見込まれる自動処理として提案されている HMGD 処理を音声信号に対して適用し、その感性的印象の変化について調査した。その応用として、種々の画像に対して HMGD 処理を適用し、元の画像と HMGD 処理画像とを比較し、感性的印象の変化に対しても検討及び考察を行った。

音声の回帰分析を行い、ゆらぎ値随伴量として、ゆらぎ値、切片、残差二乗和を求めた結果、ゆらぎ値と切片、切片と残差二乗和の相関関係が比較的強くなっていることが分かった。従って、音声信号における感性的印象の分類は、ゆらぎ値と残差二乗和とを用いる事が望ましいと考えられる。

続いて、ゆらぎ値、残差二乗和を変更したところ、再生される PS が拡張され、特定の周波数帯域の PS が増幅或いは減衰した。ゆらぎ値の変更については、元の楽曲が有するゆらぎ値を考慮しない場合、元の楽曲と比較して高周波成分の分布が大きく異なる事が分かった。従って、ゆらぎ値の変更範囲は、元の楽曲の有するゆらぎ値を考慮し、±10%程度が望ましいと思われる。

更に、残差二乗和の変更（残差をスケールリングして回帰直線への当てはまり度合いを変更すること）では、残差二乗和の値が増加する程、高周波成分が強調されている印象を受けた。従って、残差の増減によって高周波成分が強調或いは減衰され得る可能性があると思われる。

楽曲の感性的印象を問う調査1では、被験者34名に対して全10曲の楽曲を聴取させ、4.2節に示す表4.1に示す凡例、及び総合的な印象について質問紙による評価を行った。その結果を用いて、ゆらぎ値随伴量と感性的印象の相関性を調査分析した。

結果、残差二乗和 $\geq 0.50$ である楽曲は『せかせかした』印象が強くなり、残差二乗和 $< 0.50$ である楽曲は『ゆったりした』印象が強くなる傾向が見られた。また、切片 $\geq 7.7$ の楽曲は『前向きな』印象が強くなり、切片 $< 7.7$ の楽曲は『後ろ向きな』印象が強くなる傾向が見られた。これらの結果に対して平均値の検定を実施したところ、有意水準1%で有意であった。従って、今回調査した楽曲においては、切片は『前向きな』印象と相関性が強くなる傾向であり、残差二乗和は『せかせかした』印象と相関性が強くなる傾向であると考えられる。

更に、楽曲の感性的印象とゆらぎ値及び残差二乗和との関連性をより明らかにする為、重回帰式を用いた分析を行った。この分析は、音声信号の全域 (All-Frequency domain; 以下 AF, 0~22050Hz), 低域 (Low-Frequency domain; 以下 LF, 0~300Hz), 中域 (Middle-Frequency domain; 以下 MF, 300~1000Hz), 高域 (High-Frequency domain; 以下 HF, 1000~22050Hz) の4つの周波数帯域について行った。

結果、AFでは残差二乗和の方がゆらぎ値よりも感性的印象に対する影響が大きくなる事が分かった。また、LF, MF, HFのいずれの周波数帯域においても残差二乗和の方

がゆらぎ値よりも感性的印象に対する影響が大きいことが分かった。即ち、本論文で用いた楽曲の感性的印象に大きな影響を与えるのはゆらぎ値よりも残差二乗和であり、残差二乗和を人為的に変更すれば、感性的印象をより向上できるのではないかと考えられる。また、高周波なほど重相関係数が高くなっていることから、感性的印象には高周波成分が寄与すると考えられる。

また、楽曲のゆらぎ値随伴量を人為的に変更した場合の調査2では、10曲中8曲が、ゆらぎ値及び残差二乗和を変更した曲で最も高い得票数を得ており、その内3曲は平均との有意差が見られた。従って、ゆらぎ値随伴量の人為的変更によるゆらぎ値随伴量の変換は、感性的印象の向上に寄与し得ると考えられる。

そして、音声信号のヒストグラムに対して HMGD 処理を行った結果、スペクトログラムの比較より、得られた音声は振幅の増加のみならず、全ての周波数における成分を強調する傾向にある事が分かった。一方カラー画像では、ヒストグラムのピークが単一（単峰性）な画像には HMGD 処理によって陰部の輝度を上げ視認性を高める事が分かったが、ヒストグラムのピークが複数（多峰性）である画像には HMGD 処理によって立体感が失われる、或いは色味が不自然になる等の現象が確認された。従って、HMGD 処理は単峰性の画像に対して適用するのが望ましいと考えられる。

更に質問紙調査を行った結果、単峰性の画像に対して HMGD 処理を行うと、明るさ及びコントラストが改善する傾向が見られた。その上、色味及び主印象が向上し、感性的印象が改善された事が分かった。また、多峰性の画像に対して HMGD 処理を行うと、色味、輪郭及び主印象が低下し、感性的印象は向上しないことが分かった。

今回用いた楽曲のゆらぎ値随伴量では、ゆらぎ値と切片、切片と残差二乗和、ゆらぎ値と残差二乗和の何れも高い相関を示した。従って、ゆらぎ値随伴量を一般的な指標とするために、今後の調査で用いる楽曲及び音声は、ゆらぎ値随伴量が高い相関を示さないように留意の上、選択する必要があると考えられる。また、感性的印象を更に詳細に分類するため、より多くの感性形容詞を用いて各被験者の選択傾向を分析し、

統計的な調査を行う必要があると考えられる。

試聴実験には、デジタルオーディオプレイヤー、アンプ、スピーカー等の機材を用いた。しかしながら、各種音響機器、実験に使用する室内の音響特性等、ゆらぎ値随伴量以外の要素が変化する場合でも、それらが感性的印象に大きく影響を及ぼす可能性があり、検討の余地がある。また、本論文では、AF, LF, MF, HF の4つの周波数帯域における感性的印象とゆらぎ値随伴量を分析したが、各帯域におけるゆらぎ値随伴量を人為的に変更した音声信号を用い、感性的印象の変化が理論通りとなるか否かを検討する余地が残されている。

カラー画像の HMGD 処理においては、多峰性の画像において感性的印象が向上しなかった。従って、ヒストグラムの形状を分析し、平均及び分散を自動的に変更したり、多峰性の画像に対して複数のヒストグラムのピークに対してマッチングを行える様にしたりするアルゴリズムを開発することで、感性的印象が改善される可能性が残されている。

また、本章で述べたように、感性的印象には高周波成分が大きく寄与している傾向が見られている。そこで、ゆらぎ値随伴量を『感性的質感情報』として捉えれば、信号圧縮の過程において失われる質感情報をゆらぎ値及び残差として記録する事により、非可逆圧縮された信号の品質向上が期待出来ると考えられる。

具体的には、送信側において圧縮前の信号における高周波成分のゆらぎ値及び残差を別途記録し、非可逆圧縮を行う。その後、受信側へ非可逆圧縮信号と共にゆらぎ値及び残差を伝送する。受信側では、始めに非可逆圧縮信号とゆらぎ値及び残差の分離を行う。次に非可逆圧縮信号の伸張を行う。更にゆらぎ値及び残差を用いて、伸長された信号の高周波成分を推定し復元する。

非可逆圧縮はその原理上、圧縮された信号を伸張及び復元しても圧縮前の信号とは必ずしも一致するとは限らない事が知られている。しかしながら、上述の方式を用いる事により感性的質感情報を保持しつつ、品質の良い信号圧縮を実現できる可能性が見込まれる。



## 謝辞

本研究に当たり，暖かいご指導とご助言を賜りました，香川大学工学部教授・服部哲郎先生に，この場をお借りしまして厚く御礼申し上げます。

また，研究やプログラムの製作に関しまして，ご協力を頂きました，香川大学工学部服部研究室の皆様，試聴実験にご協力頂きました被験者の皆様に御礼申し上げます。

## 参考文献

- [1] 武者 利光 著：人が快・不快を感じる理由 - 音・映像・触感・・・“心地よさ”の不思議に迫る (KAWADE 夢新書), 河出書房新社 (1999).
- [2] 武者 利光 著：ゆらぎの世界 - 自然界の  $1/f$  揺らぎの不思議, 講談社ブルーバックス, 講談社 (1990).
- [3] 赤池 弘次, 和田 孝雄 著：生態の揺らぎとリズム, 講談社サイエンティフィク, (2001).
- [4] 堀内 征二 著：ゆらぎの不思議, 信濃毎日新聞社 (1997).
- [5] 武者 利光 著：ゆらぎの発想 -  $1/f$  ゆらぎの謎にせまる, 日本放送出版協会 (1998).
- [6] 武者 利光 著：ゆらぎの科学 <1>～<10>, 森北出版 (1991～1999).
- [7] 野村 俊昭, 石井 六哉 著：スペクトル解析による快音・不快音の特徴抽出, 電子情報通信学会技術研究報告, DSP, デジタル信号処理, pp. 39-46 (1995).
- [8] 菅井 桂子, 斎藤 兆古, 堀井 清之 著：音楽に伴う  $1/f$  ゆらぎ周波数成分の抽出とその人間生理への応用, 法政大学情報メディア教育研究センター研究報告, Vol. 23 (2010).
- [9] 泉 哲也, 服部 哲郎, 藤田 笑美子, 杉本 峻一, 川野 弘道 著：カラー画像におけるゆらぎ値算出随伴量と感性的印象, 日本感性工学会論文誌, Vol. 9, No. 2, pp. 243-250 (2010).
- [10] 大類 重範 著：デジタル信号処理, 日本理工出版会 (2001)
- [11] 川上 裕介, 服部 哲郎, 山松 達矢, 泉 哲也, 川野 弘道 著：音声信号のゆらぎ値算出随伴量と感性的印象との関係に関する試験的調査, 日本感性工学会論文誌, Vol. 10, No. 3, pp. 365-374 (2011).

- 
- [12] Yusuke Kawakami, Tetsuo Hattori, Hiromichi Kawano, Tetsuya Izumi : Experimental Investigation on Correlation between Kansei Impression and Feature Quantity of Sound Signal in Fourier Domain, Proceedings of 2012 International Symposium on Nonlinear Theory and its Applications, IEICE (電子情報通信学会), A1L-A, pp.1-4 (2012).
- [13] Yusuke Kawakami, Tetsuo Hattori, Haruna Matsushita, Yoshiro Imai, Hiromichi Kawano, Janaka Rajapakse : Arrangement Effect of Histogram Matching for Image and Sound Signal, Proceedings of International Conference on Humanized Systems 2013, IEEJ (電気学会), EISS (電子・情報・システム部門), ISBN 978-4-9905392-1-4, pp.87-91 (2013).
- [14] Yusuke Kawakami, Tetsuo Hattori, Daisuke Kutsuna, Haruna Matsushita, Yoshiro Imai, Hiromichi Kawano, Janaka Rajapakse : An Automated Color Image Arrangement Method Based on Histogram Matching, Proceedings of ICBAKE 2013 (2013 International Conference on Biometrics and Kansei Engineering), ISBN978-0-7695-5019-0, pp.31-34 (2013).
- [15] Yusuke Kawakami, Tetsuo Hattori, Hiromichi Kawano, Tetsuya Izumi : Statistical Investigation on Relation between Feeling Impression and Feature Parameters of Sound Signal, International Journal of Affective Engineering(日本感性工学会・英文論文誌), Vol.13, No.1, pp.71-80 (2014).
- [16] 佐藤 義治 著：多変量データの分類－判別分析・クラスター分析－，朝倉書店 (2009).
- [17] wav ファイルフォーマット (<http://www.kk.iiij4u.or.jp/~kondo/wave/>)
- [18] 川上 裕介：音声信号の感性的印象とフーリエ領域における特徴量との相関性に関する研究，香川大学大学院工学研究科博士前期課程信頼性情報システム工学専攻修士論文 (2011).
- [19] 武藤 真介 著：統計解析ハンドブック，朝倉書店 (1995).
- [20] Yusuke Kawakami, Tetsuo Hattori, Yoshiro Imai, Haruna Matsushita,

- 
- Hiromichi Kawano, Janaka Rajapakse: “Kansei Impression and Automated Color Image Arrangement Methods”, *Journal of Robotics, Networks and Artificial Life*, Vol.1, No.1, pp.61-68 (2014).
- [21] B. S. Everitt: “The Cambridge Dictionary of Statistics”, Cambridge University Press 2nd edition (2002).
- [22] Rafael C. Gonzalez, Richard E. Woods: “Digital Image Processing”, Addison-Wesley Publishing Company (1993).
- [23] Bernd Jahne: “Digital Image Processing --Concepts, Algorithms, and Scientific Applications --”, 4th Edition, Springer (1995).
- [24] Scott E. Umbaugh: “Computer Vision and Image Processing: A Practical Approach Using CVIPtools”, Prentice Hall PTR (1998).
- [25] Wilhelm Burger, Mark J. Burge: “Principles of Digital Image Processing: Fundamental Techniques”, Springer (2009).
- [26] Richard Baldock, Jim Graham: “Image Processing and Analysis: A Practical Approach”, Oxford University Press (2000).
- [27] Richard Szeliski: “Computer Vision: Algorithms and Applications”, Springer (2011).
- [28] 忽那 大輔: ヒストグラムの特徴に基づく伸縮変換によるカラー画像補正, 香川大学大学院工学研究科博士前期課程信頼性情報システム工学専攻修士論文 (2013).
- [29] International Telecommunication Union: “ITU-T Recommendation P.910: Subjective video quality assessment methods for multimedia applications”, ITU-T (2008).

## 業績

### <学会論文誌論文>

- [1] 川上裕介, 服部哲郎, 山松達矢, 泉哲也, 川野弘道: “音声信号のゆらぎ値算出随伴量と感性的印象との関係に関する試験的調査”, 日本感性工学会論文誌, Vol.10, No.3, pp.365-374, 2011.
- [2] Yusuke Kawakami, Tetsuo Hattori, Hiromichi Kawano, Tetsuya Izumi: ” Statistical Investigation on Relation between Feeling Impression and Feature Parameters of Sound Signal”, International Journal of Affective Engineering, Vol.13, No.1, pp.71-80, Special Issue on ISAE 2013, 2014.
- [3] Yusuke Kawakami, Tetsuo Hattori, Yoshiro Imai, Haruna Matsushita, Hiromichi Kawano, Janaka Rajapakse: ” Kansei Impression and Automated Color Image Arrangement Methods”, Journal of Robotics, Networks and Artificial Life, Vol.1, No.1, pp.61-68, 2014.

### <国際会議論文（査読付き）>

- [1] Yusuke Kawakami, Tetsuo Hattori, Hiromichi Kawano and Tetsuya Izumi: “Correlation between Feeling Impression and Quantities Accompanying Fluctuation Computation in Sound Signal”, Proceedings of 8th International Conference on Humanized Systems (ICHS 2012), IEEE RA Korea Chapter, pp.124-129, Daejeon, Korea, August 16-18, 2012.
- [2] Yusuke Kawakami, Tetsuo Hattori, Hiromichi Kawano, Tetsuya Izumi: ” Experimental Investigation on Correlation between Kansei

- 
- Impression and Feature Quantity of Sound Signal in Fourier Domain” ,  
Proceedings of 2012 NOLTA (International Symposium on Nonlinear Theory and  
its Applications), A1L-A, pp. 1-4, Majorca, Spain, Oct. 22-26, 2012.
- [3] Yusuke Kawakami, Tetsuo Hattori, Tetsuya Izumi, Hiromichi Kawano, Yoshiro  
Imai : ” On a relation between feeling Impression and 3PACF of sound signal” ,  
Proceedings of 2013 AROB (18th International Symposium on Artificial Life  
and Robotics) (CD), ISBN978-4-9902880-7-5, pp.276-279, Daejeon, Korea,  
Jan. 30-Feb. 1, 2013.
- [4] Daisuke Kutsuna, Tetsuo Hattori, Yusuke Kawakami, Haruna Matsushita,  
Yoshiro Imai, Tetsuya Izumi : “Color image arrangement by elastic transform  
based on histogram matching”, Proceedings of 2013 AROB ( 18th International  
Symposium on Artificial Life and Robotics) (CD), ISBN978-4-9902880-7-5,  
pp.280-283, Daejeon, Korea, Jan.30-Feb. 1, 2013.
- [5] Yusuke KAWAKAMI, Tetsuo HATTORI, Hiromichi KAWANO, Tetsuya IZUMI :  
“Statistical Investigation on Correlation between Feeling Impression and  
Feature Quantity of Sound Signal”, JSKE, Proceedings of ISAE 2013 (First  
International Symposium on Affective Engineering 2013) (CD), ISSN2187-669X,  
pp.297-302, Kitakyushu, Japan, Mar. 6-8, 2013.
- [6] Yusuke Kawakami, Tetsuo Hattori, Daisuke Kutsuna, Haruna Matsushita,  
Yoshiro Imai, Hiromichi Kawano, Janaka Rajapakse : ” An Automated Color  
Image Arrangement Method Based on Histogram Matching” , Proceedings of  
ICBAKE 2013 (2013 International Conference on Biometrics and Kansei  
Engineering) (CD), ISBN978-0-7695-5019-0, pp.31-34, Tokyo Japan, July 5-7,  
2013.
- [7] Yusuke Kawakami, Tetsuo Hattori, Haruna Matsushita, Yoshiro Imai,  
Hiromichi Kawano, Janaka Rajapakse : ” Arrangement Effect of Histogram  
Matching for Image and Sound Signal ” , Proceedings of ICHS 2013

- (International Conference on Humanized Systems 2013) (CD), ISBN978-4-9905392-1-4, pp.87-91, Kagawa Japan, September. 16-18, 2013.
- [8] Yusuke Kawakami, Tetsuo Hattori, Yoshiro Imai, Haruna Matsushita, Hiromichi Kawano, Janaka Rajapakse: "Automated Color Image Arrangement Method and Kansei Impression", Proceedings of ICAROB 2014 (International Conference on Artificial Life and Robotics 2014) (CD), ISBN978-4-9902880-8-2, pp.260-263, Oita Japan, January. 11-13, 2014.

#### <口頭発表>

- [1] 川上裕介, 服部哲郎, 山松達矢, 泉哲也, 川野弘道: "音声信号のゆらぎ値算出随伴量と感性的印象との関係に関する試験的調査", 第12回日本感性工学会大会—予稿集—(CD), 大会査読セッション, 3A1-5, pp.1-9, 9月, 2010.
- [2] 川上裕介, 服部哲郎, 川野弘道, 泉哲也: "フーリエ領域における各周波数帯域の特徴量と楽曲の感性的印象との関係に関する調査", 第7回日本感性工学会春季大会—予稿集—(CD), 企画セッション「感性情報処理のためのパターン認識」, pp.1-4, 3月, 2012.
- [3] 川上裕介, 服部哲郎, 川野弘道, 泉哲也: "フーリエ領域における音声信号特徴量と感性的印象の試験的調査", 2012電気関係学会四国支部連合大会—予稿集—(CD), 一般セッション, 16-36, pp.1, 9月, 2012.
- [4] 川上裕介, 服部哲郎, 川野弘道, 泉哲也: "ゆらぎ値を含む音声信号の特徴量による感性的影響の調査 (Investigation of Feeling Effect by Feature Quantity Including Fluctuation in Sound Signal)", 2013電気学会電子・情報・システム部門大会—予稿集—(CD), 一般セッション, GS7-1, pp.1490-1495, 9月, 2013.
- [5] 川上裕介, 服部哲郎, 松下春奈, 今井慈郎, 泉哲也, 川野弘道: "画像・音声信号に対するHMGDの自動補正効果", 2013電気関係学会関西支部連合大会—予稿集—(CD), 一般セッション, G13-3, pp.403-404, 11月, 2013.