平成１４年度

修士論文

音の印象の操作を目指した基礎的研究

電気通信大学大学院　情報システム学研究科

情報ネットワーク学専攻

学籍番号：0151020　　橋本 和哉

指導教官

阪口 豊

出澤 正徳

伊藤 秀一

平成１５年２月３日提出

目次

1.　はじめに

1.1　序論 ・・ 3
1.2　研究の目的 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 5
1.3　本論文の構成 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 5

2.　予備知識

2.1　音の心理的側面
2.1.1 音の３要素 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 6
2.1.2 音色について・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 6
2.1.3 音色の主観尺度 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 9
2.2　音の物理的側面
2.2.1 音の波形の物理構造 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 10
2.2.2 音の３要素と波形・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 13
2.3　心理量と物理量の対応
2.3.1 表現語の定量化への試み ・・・・・・・・・・・・・・・・・・・・・・・・・ 16
2.3.2 その他の報告 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 17

3.　実験の準備

3.1　問題解決への基本方針・・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 18
3.2　心理次元と物理次元
3.2.1 心理次元の構成 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 19
3.2.2 音の操作方法 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 21
3.4　聴こえの大きさの問題・・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 27
3.5　実験環境について・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 28

４．　実験と考察

4.1　予備実験 [共通して知覚される心理量]
4.1.1　実験の目的 ………………………………… 29
4.1.2　刺激音 ………………………………… 29
4.1.3　実験方法 ………………………………… 32
4.1.4　実験結果と考察 ………………………………… 33
4.2　心理実験 1 [各物理操作が音の印象に与える影響の差]
4.2.1　実験の目的 ………………………………… 35
4.2.2　刺激音 ………………………………… 35
4.2.3　実験方法 ………………………………… 38
4.2.4　実験結果と考察 ………………………………… 39
4.3　心理実験 2 [音の操作に対する心理次元の変化量]
4.3.1　実験の目的 ………………………………… 44
4.3.2　刺激音 ………………………………… 44
4.3.3　実験方法 ………………………………… 52
4.3.4　実験結果と考察 ………………………………… 53

5.　シンセサイザーの実装

5.1　シンセサイザーの仕様 ………………………………… 60
5.2　音の制御方式
5.2.1　段階的なパラメータの寄与 ………………………………… 62
5.2.2　心理量と物理量の対応行列 ………………………………… 63
5.2.3　データの処理 ………………………………… 63

6.　まとめ

6.1　全体的な考察 ………………………………… 68
6.2　結論 ………………………………… 71

謝辞　72
参考文献　73

1. はじめに

1.1 序論

音に対して抱く印象とはどんなものか、という疑問がこの研究を始めた理由である。人が音の印象を表す際には「美しい」、「力強い」、「明るい」などの形容詞を用いて表現する。ではこの形容詞に例えられた音は一体何をもってそのような印象を人に与えているのであろうか。明るい音とは一体何をもって明るいとされているのであろうか。このように「音の物理的側面と心理的側面の対応関係を調べる」ということを本研究の主題に取り上げてみた。

音といっても様々な音がある。音楽は当然のこと、道行く車の音、明け方によく聞く鳥の鳴き声、誰かの足音、隣の人の話し声、と我々の周辺には様々な音が満ち溢れている。そのような中、「何か特定の音」というように人間に意味的情報を与える音は除外して、純音のように「何の音だか分からない音」を研究の題材にとった。音の印象を測る際に、その対象とする音が何かしら特有の意味をもってしまうと、話が心理次元において複雑化するからである。あくまで、「何の音だか分からない短い提示音」に対して人間が受けとる印象を問題にする。

色聴という音と色との共感覚現象がある。どうやらこの色聴という能力を持った人は、音を聴くと色が見えるらしい。共感覚現象とは視覚と聴覚というように情報の入る器官や脳の感覚受容器が違っていても、その感覚に共通なものがあったり、相互の感覚が影響しあったりする知覚現象のことである。青色にコーディネートされた部屋に入ると涼しく感じるという視覚と体感の共感覚現象は有名な話である。昔からよく女性の甲高い声を「黄色い声」と例えたりすることからも、我々も少なかれこのような能力をもっていると思われるが、果して人間一人一人に対してどこまで共通する感覚なのであろうか。

２年くらい前か、PC 上で楽曲の旋律を担う音色を自分で作ることを趣味としていた時期がある。といっても、とあるシンセサイザーソフトに少しばかりパラメータを与えてやるとすぐに音ができるという非常に簡単なもので、実際にそのソフトが何をして音を作っていたのか詳しくは分からない。（確か FM 音源だった）今考えれば、この曲にはこの音色などと格好つけて音を作っていたが、作った音のほとんどが自分の好みに合う似たりよったりの音色であった。それでも結構たくさんの音色を作ったと思う。そして作った音を保存する際にはいつも、その音色から感じる印象を代表する言葉でもって名前をつけていた。その中でも一番多用したのが「黄色い音」、「青白い音」などの色表現であった。これと前述の「色聴」とは何か関係あるのであろうか。このことは、私が音と感覚という分野に興味をもった一番大きな理由である。音を聴いて色が見えることはないが、その音色を一言で表現するならやはり「黄色」という言葉が一番しっくりくるのである。自分でない他の人ならばどう表現するのか、気になるところである。いずれにせよ、一般的に誰に対しても同じような印象を与える音は必ず存在すると筆者は考えている。

このような背景でもって、一般的に人が「何の音だか分からない短い提示音」に対して共通に感じることができる印象と、それに対応する物理量の対応について調べていく。その方針として、「基準となる音に、特定の操作を行った時に音の印象がどのように変化するのか」ここに問題を絞って話を進めていく。音波の物理量変化と、印象の変化の対応関係が分かれば、音の印象を操作することが可能になるはずである。本研究では 1000Hz,500msec の純音を基準音として「いくつかの音の物理操作に対する音の印象の変化」を調べた。この対応関係を利用すれば、原点の周辺（基準となる音）という極めて限られた範囲でしか適応できないが、「音に明るい感じを与えたい」・「カタイ音を作りたい」など、心理次元から音を操作することができるかもしれない。これを目指し、本研究では、音の印象を心理量から直接操作することを試みるシンセサイザーを作製した。このようにして音色研究という分野の一考察を独自の手法で行った。実験において、筆者の主観に依存した何の確証もない値、独自のデータ処理方法を使うなど、現物合わせが多く、非常に粗っぽい研究であるが、筆者はこの研究においては、問題意識に対する「細部」よりも「全体的な一連の流れ」を何より大切にしたかった。

1.2 研究の目的

マニピュレータとは、シンセサイザーを使って音を作る仕事をする人達の総称である。彼らが音作りをする際には、多数の音響エフェクトのパラメータ値を変化させ、自分の耳で音を聴き分けて調節することにより、目的の音に徐々に近づけていくという方式がとられている。これに対して、前述の感覚パラメータの入力により、目的とする音に近い状態を作り、細部は自分の耳で調節するという方式を取り入れられないかと検討する。

本研究ではこれを実現するための第一歩として、音の心理量と物理量の対応関係を心理実験により割出すことを主な目的とする。そして、実験より得られたデータを用いて、実際に音の印象を心理次元から直接操作することを試みる。

1.3 本論文の構成

ここで本論文のおおまかな流れを簡単に説明する。２章では予備知識として、これまでに調べられている音の性質を心理的側面と物理的側面に分けて紹介する。３章では実験準備として、問題解決への基本的な方針と、本研究で扱う音の物理次元からの操作方法と音の印象の尺度となる心理次元を確立する。これらを前提に、４章で心理実験を行い、それぞれの実験結果について考察する。そして、実験から得られた結果を用いて、本研究の目的に従う簡単なシンセサイザーを作製した。５章にその詳細を示す。最後に６章で、それぞれの結果を踏まえた本研究全体のまとめをする。

2. 予備知識

2.1 音の心理的側面

2.1.1 音の３要素

我々が日常生活において音として認識しているのは、音の波動（音波）によって生じた空気の圧力の変動である。音とは音波によって引き起こされる聴覚的現象であり、その性質は主に「音量」、「音程」、「音色」と大きく３つに分類されている。ここで「音量」は音の大きさ、「音程」は音の高さ、「音色」は音の質にそれぞれ対応している。

人間が音を聞き分ける際には、その音の大きさや高さを判断の基準にしているのではなく、むしろその音の音色に着目している場合が多い。音色は、音の大きさ、高さに並ぶ聴覚上で最も基本的な性質の１つとされている。しかし、大きさや高さの感覚量がほぼ一次元的に変化する量であるのに対し、音色は心理的に多次元的性質を持っており、さらに音色を規定する物理的要因も多次元的であるために物理量に対応した尺度の構成は極めて難しいと考えられている。

2.1.2 音色について

＜音色の定義＞

音色とは聴覚上の性質の１つで、２音の大きさ、及び高さが共に等しくてもその２音が異なった感じを与えるとき、その相違に対応する性質と定義されている。（JIS，ASA 規格）（図 2.1）

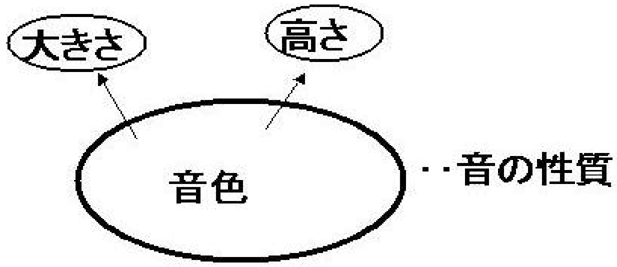


図 2.1　音色の定義（JIS,ASA）

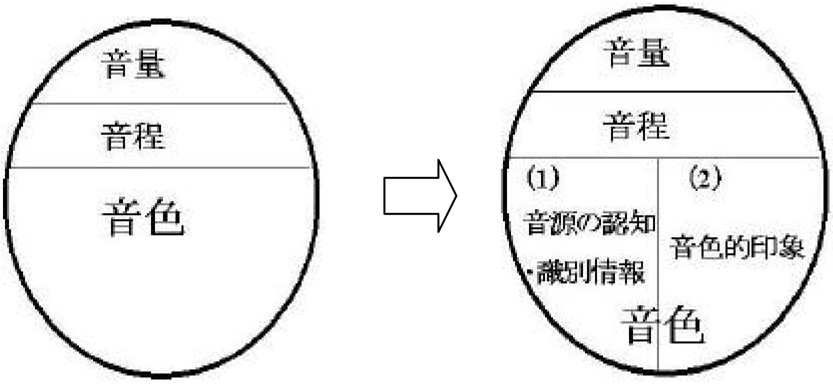
しかし、この定義では音から受ける聴覚上の印象から、心理物理学的に比較的よく知られている音の大きさ、高さという性質を取り除いた残りを音色という性質に押し当てているに他ならない。

＜音色の定義の提案＞

Helmholtz は音源の認知・識別（聴こえた音が何の音であったかということ）の手掛かりが、音色の相違にあることに注目し、音色を音源の認知に関与する属性として位置づけ、一方で我々は音を聞いて単に音源を認知するだけでなく、その音色特有の様々な感じ（音色的印象）を受けることがあると示した。このように、音色には「認知・識別」の側面と、「音色的印象」の側面があることを考慮して音色を再定義すると次のようになる。（図 2.2）[2]

(1) 音源が何であるか認知するための手掛かりとなる性質

(2) 音を聴いた主体が音から受ける印象のことで、澄んだ・濁った・明るい・暗い等の共感覚的表現で示される性質



(JIS，ASA)　(Helmholtz)

図 2.2　Helmholtz による音色の定義

このように音色は定義そのものがまだ十分確立されていない性質であり、議論の余地を残しているといえる。その他の音色の定義に関する意見として、「音の持続感」と「音源定位」も音を区別しうる性質の１つであり、これらを音色の性質と分けるという考え方も提案されている。また別の意見として、JIS などの定義を狭義の音色とし、さらに広義の音色を「音の大きさ、高さを総合した印象」として定義している考え方もある[1]。

これに対して、筆者が考える音色とは音そのものから感じる印象のことである。聴取した音事象の中に含まれる規則性のある物理特性それぞれの印象の総和が、音色という一つの音事象※の印象を形成すると考えている（図 2.3）。言い換えれば、Helmholtz の提案する「音色的印象」を「音の大きさ、高さ等の、音の性質全体を総合した印象」と解釈し、音の印象そのものを音色という言葉で表してしまう考え方である。ただし、その一つ一つの物理特性から生じる印象が互いにどのように干渉し合うかは分からない。（非常に粗っぽいやり方であるが）本研究において、その相互干渉は無視できるものと仮定して話を進めている。

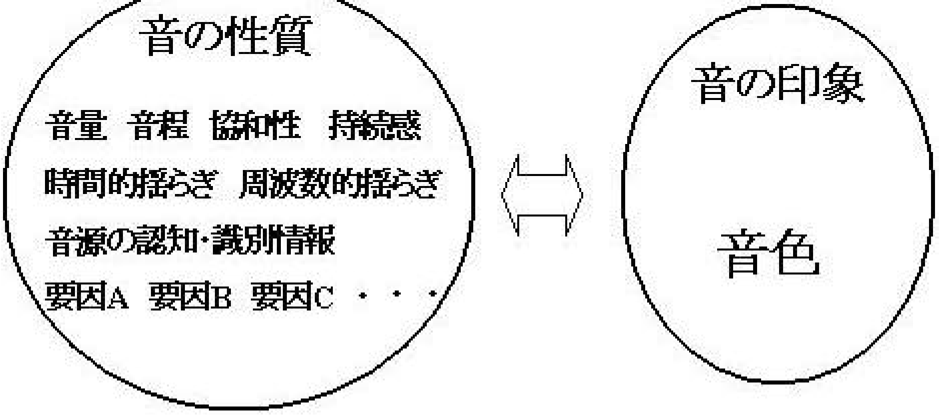


図 2.3　筆者の提案する音色の定義

[音事象]

- **日常生活において、一つの音として知覚される事象は、立ち上がり,継続部,立ち下がりという３つの要素からなる時間的な構造をもっている。この３つの要素からなる音の始まりから終わりまでの単位を「音事象」と呼ぶ[11]。**

2.1.3 音色の主観尺度

＜音色表現に使われる言葉＞

人が音の印象を表現する方法として、一般的に表現語という定性的な言葉でもって表す場合が多い。この「音を表現する言葉」を整理・分類し、共通な因子を見出すことにより、音色の主観度構成を行う試みがなされている。ここで得られた結果では、人間が音を聴取した際に音色から受ける感覚は、主に次に挙げられる３つの主観的な因子をもつ感覚と関係していることが示唆されている[1][3][8]。

Ⅰ, 美的・叙述的因子

・(美しい － 汚い)・(快 － 不快)
・(澄んだ － 濁った)・(鮮やか)
・(潤いある － 乾いた)　　　　等

Ⅱ, 量的・空間的因子

・(豊かな － 乏しい) ・(迫力ある － 物足りない)
・(力強い － 弱々しい)
・(広がりある － まとまった)　等

Ⅲ, 明るさ・金属性を示す因子

・(明るい － 暗い)
・(鋭い － 鈍い)・(金属的な)
・(軽い － 重い)・(丸い)　　　　等

(Ⅳ), 柔らかさを示す因子

・(柔らかい － 硬い)

このⅠ~Ⅲ(Ⅳ)に分類された各同一因子内の言葉は非常に相関性が高く、異なる２音において、その感覚の優劣を比べた時には同等の評価を得る場合が多い。

2.2 音の物理的側面

2.2.1 音の波形の物理構造

ここで以下の議論の準備として、いくつかの専門用語について説明する。[1]

[純音]

・瞬時音圧が時間の正弦関数である音。すなわち正弦関数一つで示すことができる音波のこと。一般に次式で示される。

$$P(t) = A \cdot \sin(2\pi f t + \theta)$$

$P(t)$：瞬時音圧
A：振幅
f：振動数
t：時間
θ：位相角

[複合音]

・純音以外の音の総称である。

[部分音]

・周期性のある複合音は、波形をフーリエ級数展開することにより、いつくかの純音の足し算として表すことができる。このとき、元の波形の構成要素となる一つ一つの正弦波（純音）を、その複合音の「部分音」と呼ぶ。

[倍音]

・複合音を構成する部分音の周波数が整数倍関係にあるとき、その周波数構造を「倍音構造」という。このとき、最も低い周波数をもつ部分音を「基本音」といい、その基本音の２倍の周波数をもつ部分音を「第２次倍音」、以下第３

次、第４次倍音という。高さを感じる音は全てこの倍音構造をもっていて、その音の高さは基本音の高さと一致する。

[楽音]
・楽器音の総称である。打楽器の音を除いて、基本的に全て倍音構造をもっている。

ここで、以上に説明した用語を図示する。（図 2.4）

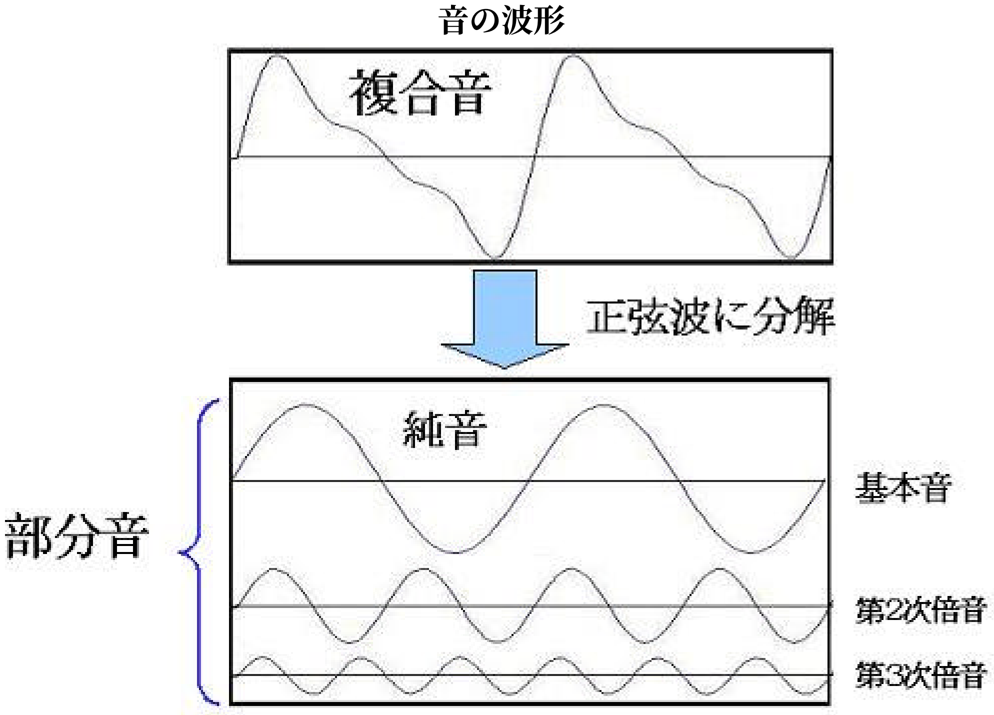


図 2.4　音の波形の物理構造

[振幅変調音]（AM 音：amplitude modulated sound）
・振幅を時間的に変調させた音のこと。（図 2.5a）以下に純音を正弦波で変調した例を示す。

$P(t) = A\sin(2\pi ft)\cdot(1+\sin(2\pi gt))$

f：元の純音の周波数（搬送周波数）
g：変調する正弦波の周波数（変調周波数）

また、ここで

$$P(t) = A\sin(2\pi ft)\cdot(1+\sin(2\pi gt))$$
$$= A\sin(2\pi ft) + \frac{A}{2}\cos(2\pi(f-g)t) - \frac{A}{2}\cos(2\pi(f+g)t)$$

より、振幅変調することによって $f+g$ と $f-g$ の「側帯波」と呼ばれる周波数成分が新たに生じる。

[周波数変調音]（FM 音：frequency modulated sound）
・周波数を時間的に変調させた音。振幅は一定で、周波数が時間的に変化する。
（図 2.5b）以下に純音を正弦波で変調した例を示す。
周波数変調音では、$f\pm g$, $f\pm 2g$, $f\pm 3g$,…という多くの側帯波が生じる。

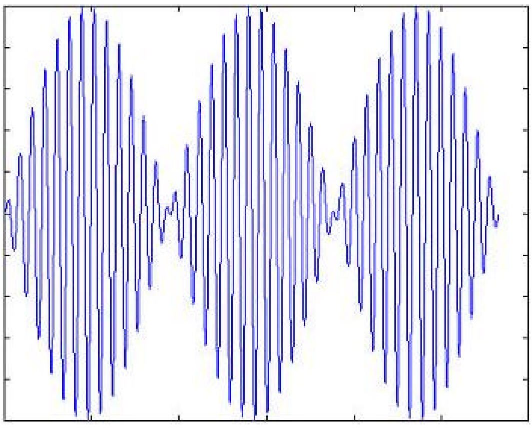


図 2.5a　AM 変調音の時間波形

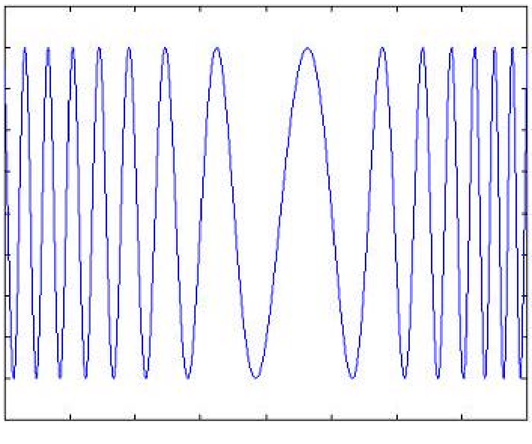


図 2.5b　FM 変調音の時間波形

2.2.2 音の３要素と波形

JIS,ASA 規格で定義されている「音量」、「音程」、「音色」と音の波形の対応関係を以下に記す。

＜音量＞
音の大きさに相当する音量は、音波の振幅と非常に高い正の相関をもつ。すなわち波形の振幅の増加によって、音量がほぼ一次元的に大きくなる。しかし、人間の聴覚特性では、知覚する音の高さや成分周波数帯域によっても、聴こえの大きさが異なるという性質があるため、音の大きさの知覚は振幅のみに依存するというわけではない。

＜音程＞
音の高さに相当する音程は、音波の成分周波数（特に基本周波数）と非常に高い正の相関をもつ。こちらも音量と同様に、音程が波形の周波数の増加によって、ほぼ一次元的に高くなる。しかし全ての音が音程感をもつわけではない。連続スペクトル周波数構造をもつホワイトノイズ等の雑音が代表的な例である。このように音の高さに相当する感覚には、音程感のあるものと、ないものに分けられる。前者に対する音程感は、前述のように基本周波数の高さである。しかし、例外として、440Hz , 880Hz , 1320Hz の倍音構造をもつ音から 440Hz の成分周波数を取り除いたにもかかわらず、音全体の高さが 440Hz に聴こえるという「missing fundamental」という現象が報告されている。

音の高さと近い関係をもっている「音調性」という言葉がある。音の高さを次第に上げていくと、当初の音と非常に類似性のある音がやってくる。これはオクターブと呼ばれ、２倍の周波数になるごとに循環的に類似性が戻ってくるという性質である。これを音調性という。しかし、必ずしも周波数が２倍になった時に、その類似性が戻ってくるとは限らない。この関係は 200～1000Hz の基本周波数をもつ音に対しては、ほぼ成立するが、これ以上、及びこれ以下の周波数の音に対しては、少し周波数の幅を広げなければならない。例として、2000Hz の１オクターブ高い音は 4000.21Hz と少し周波数の幅が広がっている。このように、心理オクターブは音響的オクターブよりも広がっていることが報告されている。[8]

＜音色＞

音色を規定する物理的要因は、音のスペクトル、波形、音圧、およびそれらの時間的変化に関係していると言われている。特に楽音に限定して考えれば、音色を規定する物理的要因は、音の倍音構造とスペクトルに大きく依存している。しかし、これは対象とする楽音が周期性のある定常状態である場合に限定して考えた結果であり、実際の音を考える場合には、音の立ち上がりと立ち下り部の過渡部分に現れる非定常部分から生じる、音の持続感・時間的変化（エンベロープ）というパラメータも音色を規定する物理的要因として考慮する必要がある。

よって、対象とする音をスペクトル構造が時間的に変化しない「定常音」と、スペクトル構造、音量、音程が時間的に変化する「変動音」に分けて考慮すると、それぞれの音色は以下のような物理パラメータを周波数領域に持つと考えられる[1]。

・ 定常音の場合（図 2.6a）
（ｉ）スペクトル包絡
（ii）各部分音の相対振幅の大きさ
（iii）成分周波数の相対関係
（iv）成分間の位相関係

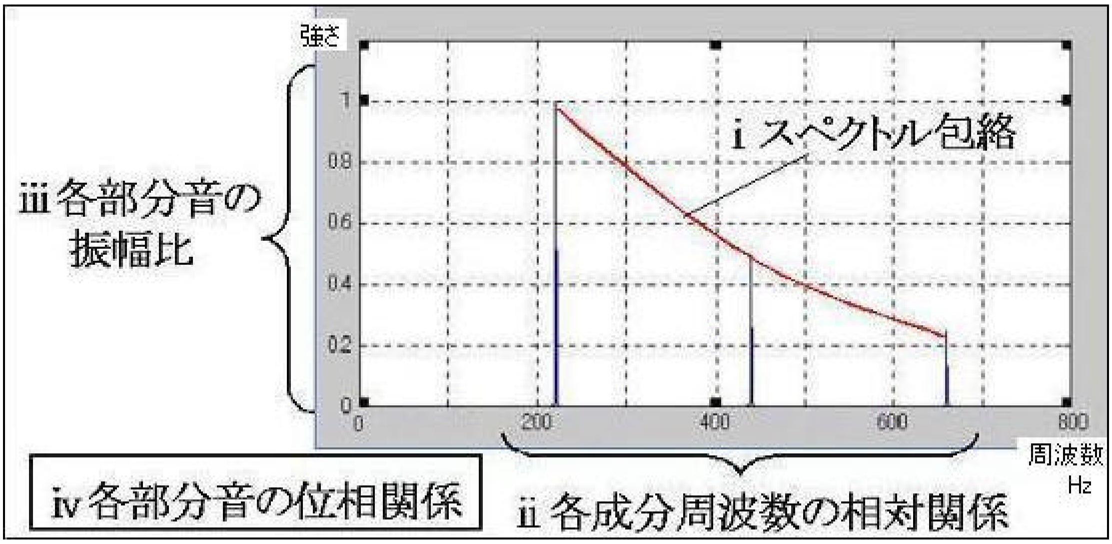


図 2.6a　定常音における周波数領域の波形

・変動音の場合（図 2.6b）

（ⅰ）スペクトル包絡の時間変化

（ⅱ）各部分音の振幅の時間変化

（ⅲ）各部分音の周波数の時間変化

（ⅳ）時間領域における波形の時間包絡（エンベロープ）

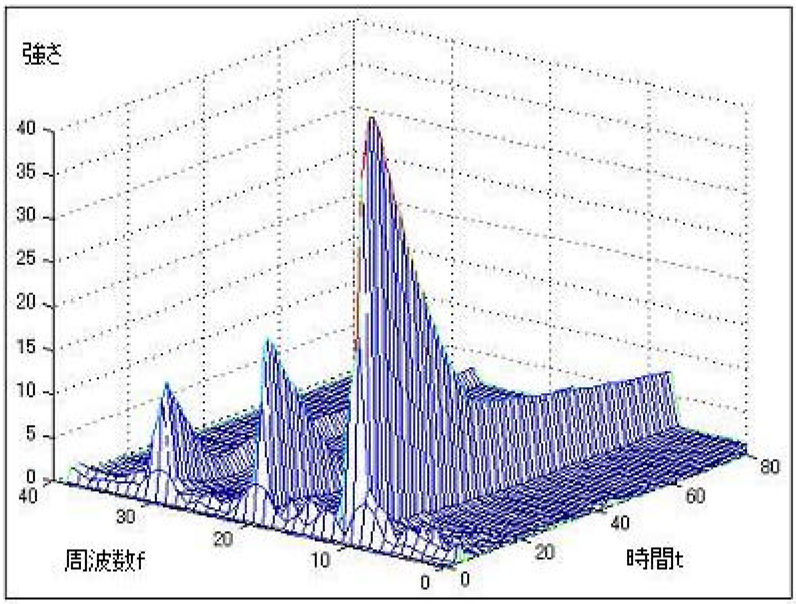


図 2.6b　変動音のスペクトログラム

＜広義と狭義の音色＞

音色を規定する物理要因は、部分音のスペクトル構造であるが、スペクトル構造が時間的に変化しない定常音から感じられる印象を”timbre”と呼び、その timbre や音の大きさ、高さが時間的に変化する変動音から感じられる印象を”sonance”と呼ぶことが提案されている。[2]

2.2.3 心理量と物理量の対応

2.3.1 表現後の定量化への試み

2.1.3 に挙げた音を表現する言葉の主要な主観的因子が具体的にどのような物理量によって規定されているかが明らかになれば、音色を定量的に扱うことが可能になると考えられる。しかしながら、筆者らの知るところでは、現在のところでは主観的因子と物理量との関係を明らかにした報告は少ない[1][6]。以下、表現語と物理量との対応を計ることで、表現語の定量化を試みた研究を簡潔に紹介する。

＜先行研究１・音の鋭さ＞[1][7]

Bismarck は、帯域幅が 200Hz〜1kHz、スペクトル包絡の傾斜が 0 dB/oct である帯域雑音より感じられる鋭さの感覚 S_{on} を音の鋭さの基準量に定めた。そして、下限周波数を 200Hz に固定し、上限周波数 f_{ul} と、スペクトル包絡の傾斜 k を変化させることより、いつくかの帯域雑音と調波複合音の定常音における相対的な鋭さの感覚量 S/S_{on} を測定した。

この実験の結果、帯域雑音と調波複合音のどちらの場合においても、音の鋭さは、ほぼ同様に知覚されているとうことがわかった。これは、連続的、離散的というスペクトル構造の違いは、音の鋭さの感覚量にそれほど影響しないことを示唆している。また、上限周波数 f_{ul} の増加やスペクトル包絡 k の高域の増加につれて、鋭さは単調に増加した。さらに、帯域雑音においては、上限周波数が 1〜2 臨界帯域分上昇することによる鋭さの増分が、f_{ul} を固定した状態で傾斜を +6dB/oct したときに相当していることが導出された。
このようにして、Bismarck は鋭さという表現語を定量的に測定して、その対応する物理量を探し出す研究を行った。

＜先行研究２・音の粗さ＞[1]

Terhardt は、正弦波振幅変調音を対象にして音の粗さを測定した。これによると、粗さの感覚量は、主に変調度 m に強く依存する。ここで、振幅変調音の

粗さ r と変調度 m の間には、 $r \approx k \cdot m^2$ という関係がある。k は搬送周波数 f_{car} と、変調周波数 f_{mod} に依存する定数である。

粗さに関する実験には、振幅及び周波数変調音が用いられることが多い。よって、一般的な音の粗さを、求めるためには、物理量が定まった特定の変調音を基準音として規格化し、それに対して相対的な粗さを求めることが必要とされる。

2.3.2 その他の報告

果たして、対象とする音のどのような物理量が人間の音色の知覚に対して影響しているのであろうか。以下に、2.3.1 以外に報告されたいくつかの表現語と物理量との対応を示す[7]。

・純音は部分音が１つという最も単純な音の構造を持っているが、これは一般にとても澄んだ感じで知覚されることに対し、部分音の数が多くなると音色は豊かな感じが増してくる。

・多くの楽音のように部分音が基音と整数倍にある音に比較して、非整数倍の周波数成分を含む音は、濁った感じを与える。

・高次倍音の強さが増すと音は明るく、あるいは荒々しくなる。

・奇数次倍音だけを含む音は空虚な感じを、偶数倍音だけを含む音は充実した感じを与える。

・部分音の相対振幅の側面から、低音領域の周波数成分が強い音は重厚な印象を与え、およそ 1000Hz 以下の中音域の成分が強い音は力強く、高音域の成分が強いと音は硬く、時には荒々しく感じる。

楽音の定常音部分に限定して考えれば、このような部分音の構造がそれぞれの楽器の持つ音色の特徴を大きく支配しているといえる。

3．実験の準備

3.1 問題解決への基本方針

本研究では、音を印象という心理次元から直接操作することを目指している。これを実現するために、ある基準となる音を変化させ、その変化前と変化後の音を比較して、音の印象の相対的な差を測定することにより、音の印象という心理量と、音波の変化量という物理量の対応関係を調べていきたい。図 3.1 に、基準音にある操作を施したとき、「明るさ」という心理量が増加した例を図示する。

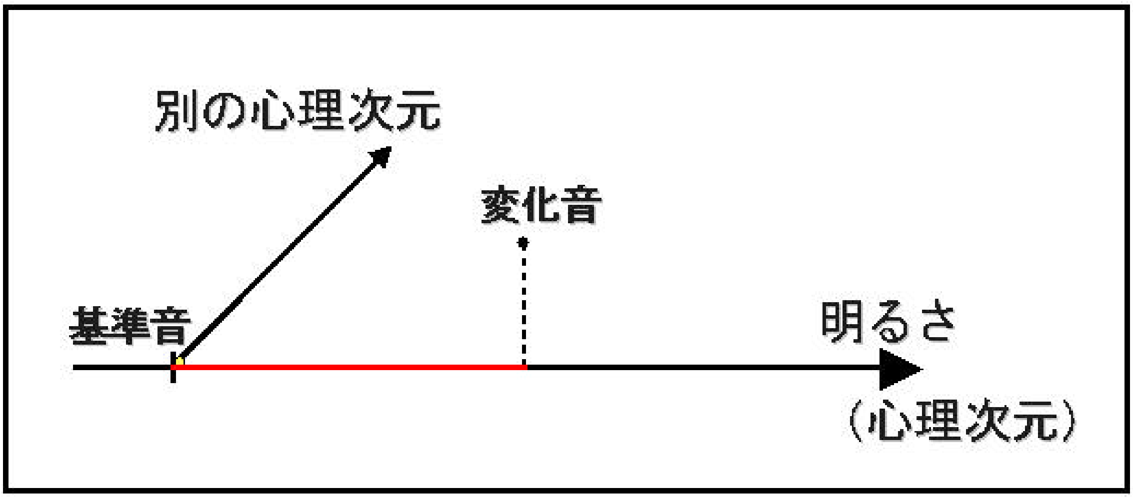


図 3.1 物理量の変化によって生じる心理次元の変化例

本研究では、基準となる音を、最も単純な正弦波関数で表すことができる純音に設定した。詳細は以下の通りである。

$$S(t) = A \cdot \sin(2\pi f t) \qquad (70\mathrm{dB}(\mathrm{SPL}))$$

S(t)：瞬時音圧 (dB)
A：振幅係数
f：周波数 1000 (Hz)
t：時間 $0 \leqq t \leqq 500$ (msec)

70dB(SPL)とは、おおよそ我々の日常会話レベルの音の大きさである。

3.2 心理次元と物理次元

ここでは、本研究で取り扱う音の「心理量」と「物理量」について説明する。音の心理量とは「人間が音を聴取した際に抱く印象」のことである。これは主に、2.1.3 に示すような言葉で表現される感覚のことである。一方で、音の物理量とは「音波の状態・特性、及びこれらを構成するもの」のことである。音の基本周波数、振幅、周波数構造、音源などがこれにあたる。そして、心理量が広がる空間を「心理次元」、物理量が広がる空間を「物理次元」と定義する。本論文においては、言葉の意味を以上のように解釈してもらいたい。図 3.2 にこの定義の概略を図示する。

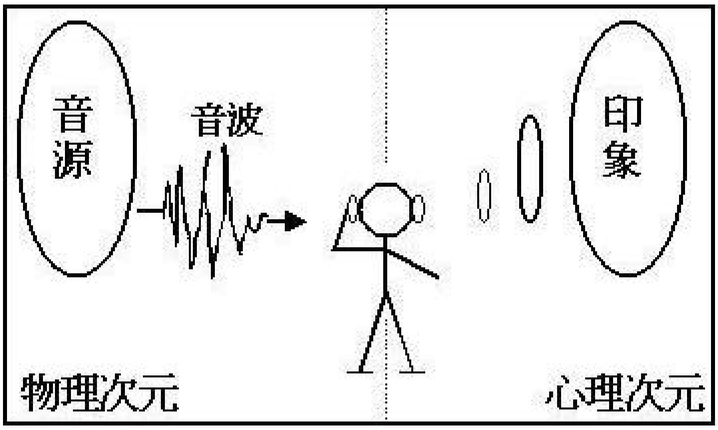


図 3.2　音の物理次元と心理次元

3.2.1 心理次元の構成

ここで、本研究で主に着目する心理次元を構成する。

2.1.3 に示すＩの因子群における、美しい・快という言葉は、その言葉自体が連動して感じる、極めて個人差のある心理的・感性的な感覚量であるといえる。つまり、音を聴取した時に感じる印象として、この２つの言葉は相関性が高い。美しければ快であるし、逆に汚ければ不快であり、その評定尺度は人それぞれであると考えられる。これより（快 ― 不快）の言葉は（美しい ― 汚い）の言葉で代表される感覚の中にまとめてしまうことにする。また一方で、力強い・迫力あるという言葉にも同じような関係が成り立つといえる。

このような方式で、音の印象を表現する言葉を整理すると、2.1.3 に示す音表現に使われる言葉は以下の９つの対になる。ここで、対になっている反対の意味を持つ言葉を、その一方のマイナス量と考えると、言葉の意味で示される空間（心理次元）が９つ出来上がる。本研究では、主にこの９つの心理次元を対象として議論する。

Ⅰ

１．美しい　－　汚い
２．澄んだ　－　濁った
３．潤った　－　乾いた

Ⅱ

４．力強い　－　弱々しい
５．まとまった　－　広がった

Ⅲ

６．明るい　－　暗い
７．鋭い　－　鈍い
８．軽い　－　重い

Ⅳ

９．硬い　－　柔らかい

これらの表現語を眺めると、聴覚は主に視覚と触覚に共通した感覚反応を引き起こす場合が多いことに気づく。確かに「おいしい音」、「辛い音」、「苦い音」、「臭い音」などと味覚と嗅覚に共通する感覚表現はない。ここでは人間の深層心理について追求する気はないが、少なくとも音の印象について考える際には、聴覚以外の感覚から得られた体験について考えることが重要であり、おそらくそれは日常生活の広い体験に根ざすものであると推測できる。

3.2.2 音の操作方法

音を印象という心理次元から直接操作するためには、音の心理量に対応する物理量を介して音を操作する必要がある。音波は物理量である。本研究では主に、以下に示す６つの物理次元からの操作（パラメータ操作）によって音を変化させる。

＜振幅操作＞

2.2.2 に示すように、音波の振幅は音量とほぼ一次元的な対応関係がある。この操作は、3.1 の基準音の振幅係数 A を変化させることである。

＜基本周波数操作＞

2.2.2 に示すように、音波の基本周波数は音程とほぼ一次元的な対応関係がある。この操作は、3.1 の基準音の周波数 f を変化させることである。

＜エンベロープ操作＞

エンベロープとは、巨視的にみたパワー時間変化のことで、日本語では包絡線という。波形を例にとって考えれば、音波の振幅の頂点を滑らかな曲線でつないだものがこれに相当する。また、スペクトルの包絡線をスペクトルエンベロープという。

一つの音事象に対するエンベロープパターンは、ほぼ無限に存在すると考えられる。本研究においては「音の立ち上がり」と「音の立ち下り」という２つのパラメータのみに注目することによって、音の持続感を決定することにした。ここで、音の持続感を形成するエンベロープ関数 $En(t)$を以下のように定義する（図 3.3）。

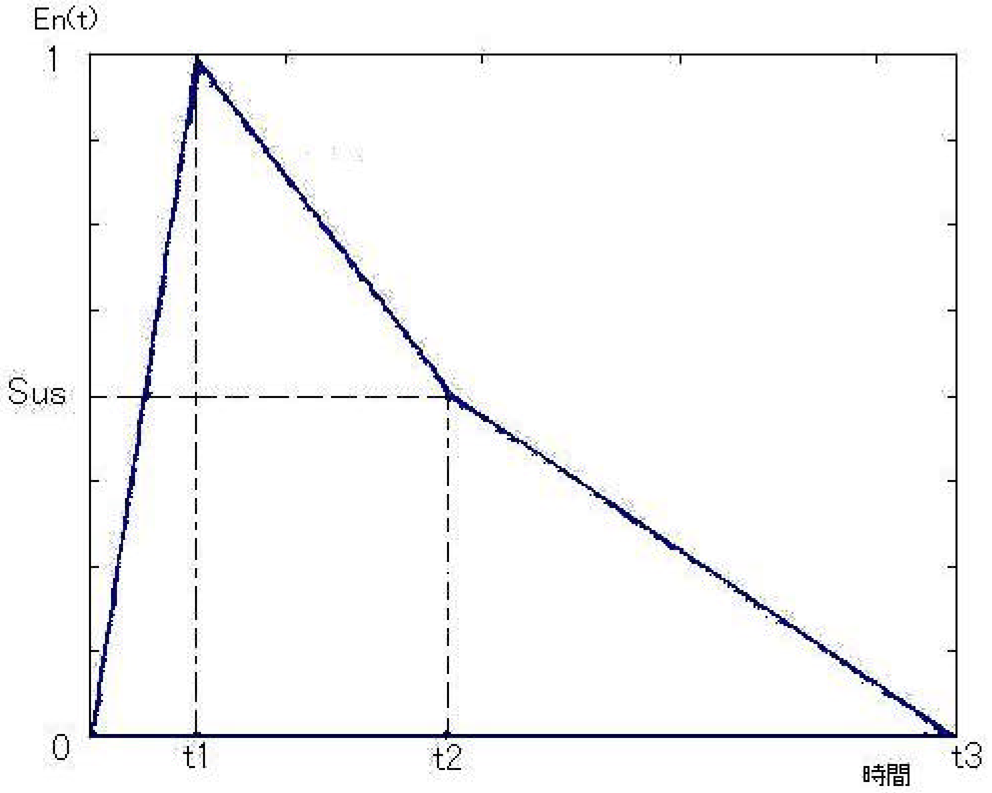


図 3.3　**エンベロープ関数** En(t)

図 3.3 **に示すように、最大値を１とするこれらの包絡線を形成するエンベロープ関数** *En(t)***は以下の式で示される。**

$$
\mathrm{En}(t) = \begin{cases} \dfrac{1}{AT}t & (0 \leqq t \leqq t_1) \\ -\dfrac{1-Sus}{t2-t1}\mathrm{t} + \dfrac{t2-Sus \cdot t1}{t2-t1} & (t_1 < t \leqq t_2) \\ -\dfrac{Sus}{t3-t2}\mathrm{t} + \dfrac{Sus \cdot t3}{t3-t2} & (t_2 < t \leqq t_3) \end{cases}
$$

(　$t_1 = 1/AT$ ，$t_2 = t_1 + (1-Sus)/DC$ ，t_3 =**任意（音の終わり時刻）**　)

ここで扱う各パラメータの詳細を以下に記す。

[*AT* (アタックレート)]
　$0 \leqq t \leqq t_1$ 区間における包絡線の傾き。

[*DC* (ディケイレート)]
　$t_1 < t \leqq t_2$ 区間における包絡線の傾き。

[*Sus* (サスティーンレベル)]
　t_2 の値を決定する。*DC* で減衰する区間において、最大振幅からどこまで減衰するのか決定する

　この En(t)と 3.1 の S(t)を用いて

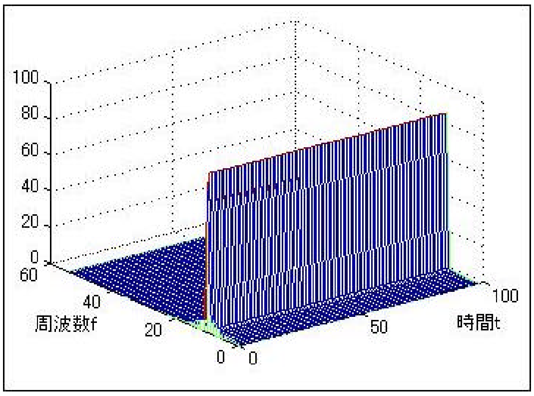
$$S_{en}(t) = En(t) \cdot S(t)$$

$S_{en}(t)$：瞬時音圧 (dB)

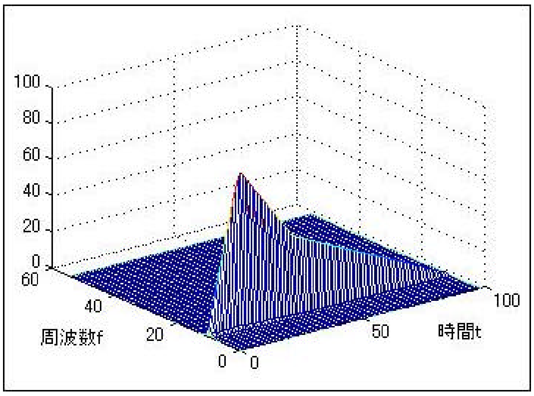
（ $0 \leqq E(t) \leqq 1$ より $S_{en}(t)$は S(t)の振幅を最大振幅にとる）

となる。

　以下にエンベロープ関数を与える前と後のスペクトログラムを示す。(図 3.4)



S(t)



S_{en}(t)

図 3.4　S(t)と $S_{en}(t)$のスペクトログラム

このように、一つの音事象 $S_e(t)$は「最大振幅に到達するまでの時間 t_1」と「最大振幅から *Sus* によって決定される音圧に減衰するまでの時間 t_2」と「t_2から音の鳴り終わる時間 t_3」の３つの時間幅によって包絡線の傾きが変化する。

＜振幅変調＞

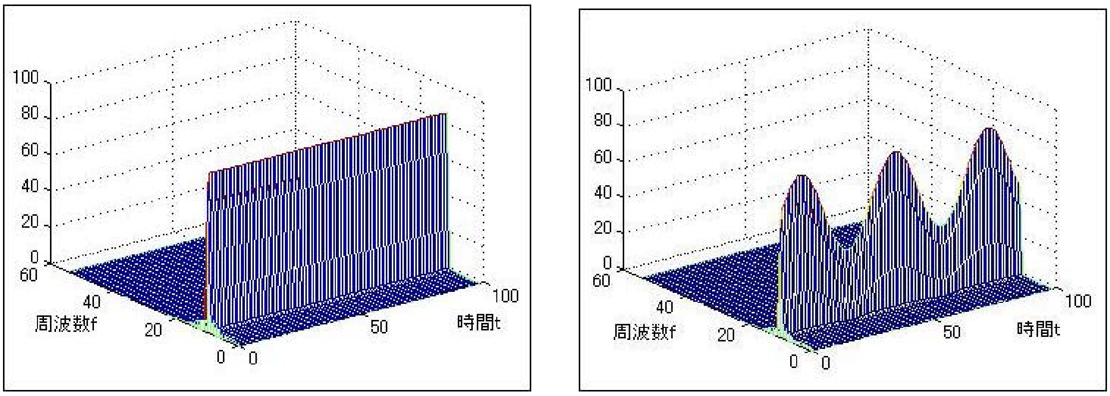
振幅変調音とは 2.2.1 で示したように、振幅を時間的に変調させた音のことである。振幅変調は「元の波形に対しての時間軸上における係数の掛け算」ということに関して、エンベロープ関数と同じ物理操作と考えられるが、本研究においては、エンベロープを「音の持続感」、振幅変調を「時間的な揺らぎ」と考えているために、別の物理操作として取り扱っている。

基準音 $S(t)$を正弦波で振幅変調した時の式を以下に記す。

$$S_{am}(t) = S(t)\cdot\{\ (1 - D_{am}) + D_{am}\cdot\sin(2\pi f_{am}t)\ \}$$

$S_{am}(t)$：瞬時音圧 (dB)
D_{am}：変調する正弦波の振幅係数 ($0 \leqq D_{am} \leqq 1$)
f_{am}：変調周波数 (Hz)

D_{am}は言い換えれば「変調の深さ」である。元の波形の振幅をどれくらいの振幅で揺るがすか、時間波形の揺らぎの深さを決定するパラメータである。振幅変調前と後のスペクトログラムを以下に示す。（図 3.5）



$S(t)$　　　　$S_{am}(t)$

図 3.5　$S(t)$と $S_{am}(t)$のスペクトログラム

＜周波数変調＞

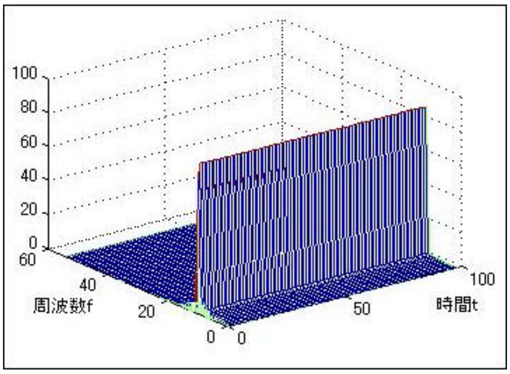
周波数変調音とは 2.2.1 で示したように、周波数を時間的に変調させた音のことである。

基準音 S(t)を正弦波で周波数変調した時の式を以下に記す。

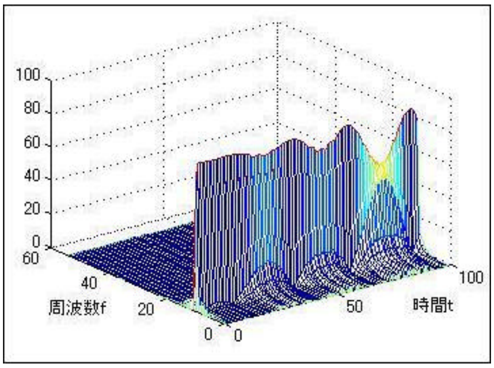
$$S_{fm}(t) = \sin(2\pi t \cdot (f + D_{fm} \cdot \sin(2\pi f_{fm} t)))$$

$S_{fm}(t)$：瞬時音圧 (dB)
D_{fm}：変調する正弦波の振幅係数
f_{fm}：変調周波数（ Hz ）

周波数変調前と後のスペクトログラムを以下に示す。（図 3.6）この２つのスペクトルの時間変化を見比べてみると、周波数変調音 $S_{fm}(t)$のスペクトログラムは側帯波による影響を強く受けていることが見られる。



S (t)



$S_{fm}(t)$

図 3.6　S(t)と $S_{fm}(t)$のスペクトログラム

＜倍音成分の追加＞

2.2.2 で挙げた音色を規定する物理的要因に

（iii）成分周波数の相対関係

という要因がある。これは音の印象を大きく変化させる要因であり、別の言葉で「協和性」と呼ばれている。協和・不協和という概念は、調和構造を有する楽音に対して使われ、一般的に部分音の成分周波数の相対的な比が単純な整数比の時に協和性がよくなることが示されている[1]。

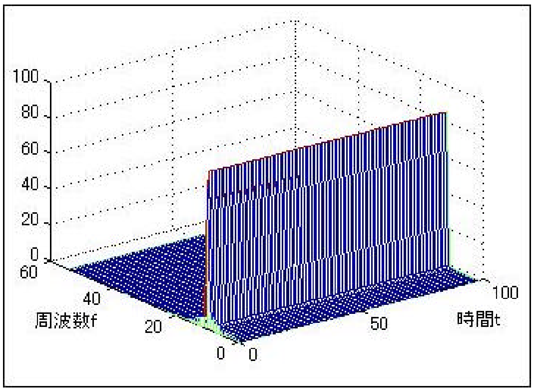
一つの音事象に対する協和・不協和パターンは、エンベロープパターン同様、ほぼ無限に存在すると考えられる。よって、本研究では、基準音の基本周波数の倍音成分に、いくらかの係数掛けをして、元の波形式に追加した時の音の印象の変化を考えることにした。こうして作製された音は、基本音と第２次倍音しかもたない単純な音であるが、倍音構造にある協和性が音の印象に対して及ぼす影響を、大いに含んでいると筆者は考えている。

基準音 $S(t)$に、第２次倍音を追加した時の式を以下に記す。

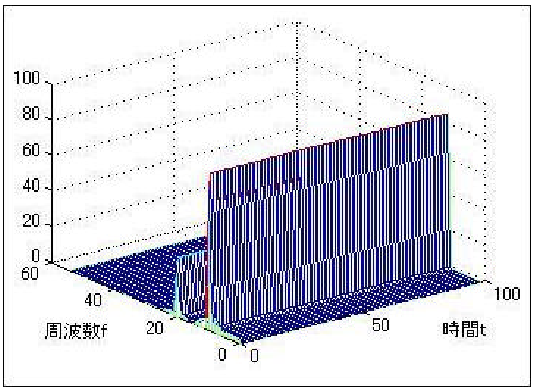
$$S_{ha}(t) = S(t) + A_{ha} \cdot \sin(2\pi f_2 t)$$

$S_{ha}(t)$：瞬時音圧 (dB)
f_2：倍音成分（ Hz ）
A_{ha}：倍音成分の振幅

倍音成分追加前と後のスペクトログラムを以下に示す。（図 3.7）



S (t)



$S_{ha}(t)$

図 3.7　$S(t)$と $S_{ha}(t)$のスペクトログラム

3.4　聴こえの大きさの調節

3.3 にて作製する音は基準と比べると、若干大きさが異なって聴こえる。よって本研究においては、振幅操作音以外の全ての刺激音を、基準音となる 1000Hz・70dB・500msec の純音から聴こえる音の大きさと同等の評価を得ることができる音の大きさに調節した。その際に用いた手法を以下に簡潔に記す。

＜音エネルギー一定＞

エンベロープ生成、AM 変調、FM 変調の操作により作製した刺激音に関しては、音のエネルギーの総和を統一することにより、聴こえの大きさを統一した。これには、刺激音の時間波形の総区間をフーリエ級数展開して求めたスペクトルの 2 乗和を、基準音のものと等しくする手法を用いた。

＜音の大きさの等感曲線から算出＞

2.2.2 に示したように、聴こえの大きさは知覚する音の高さによって、同じ音圧であっても異なる。よって、高さの異なる刺激音の大きさにおいては、純音の音の大きさの等感曲線（ロビンソン・ダットソン曲線）から 1000Hz・70dB の純音と同等の評価を得ることができる音圧レベルを求め、その値から計算された係数を波形の時間式に掛けることにより、聴こえの大きさを統一した。

＜複合音の大きさの算出・フレッチャーの方法＞[1]

聴こえの大きさは、知覚する音の周波数成分の帯域幅によっても変化する。一般的に、知覚される複合音の周波数帯域幅が広がると、同じ音圧であっても聴こえる音の大きさが大きくなる。この帯域幅はいくつかの区間に分けられていて「臨界帯域」と呼ばれている。[1]音の周波数成分帯域がこの臨界帯域をいくら跨ぐかで、音の大きさが異なる。

倍音成分を追加した音は、同じ音圧であっても基準音より大きく聴こえる。本研究では、この聴こえの大きさを揃えるために「フレッチャーの複合音の大きさの算出法」を用いて、複合音の音の大きさを算出し、これを基準音の大きさに等しくするように調整した。

＜急激な立ち上がり・立ち下がりの調節＞

本研究で扱う全ての刺激音は、音の始まりと終わりの 2msec にフェードイン・アウトの効果をつけることにより、音とび問題に見られるクリック雑音のようなノイズの影響を回避した。

3.5 実験環境について

実験では、刺激音を以下に示すような PC－AT 互換機から出力し、出力音声をオーディオアンプで増幅した後、被験者にヘッドフォンを通じて提示した（図 3.8)。尚、実験は全て正常な聴力を持つ２０代の男女に対して、防音設備のない実験室で灯りを点けた状態で行った。個々の実験装置の仕様を表 3.1 に示す。

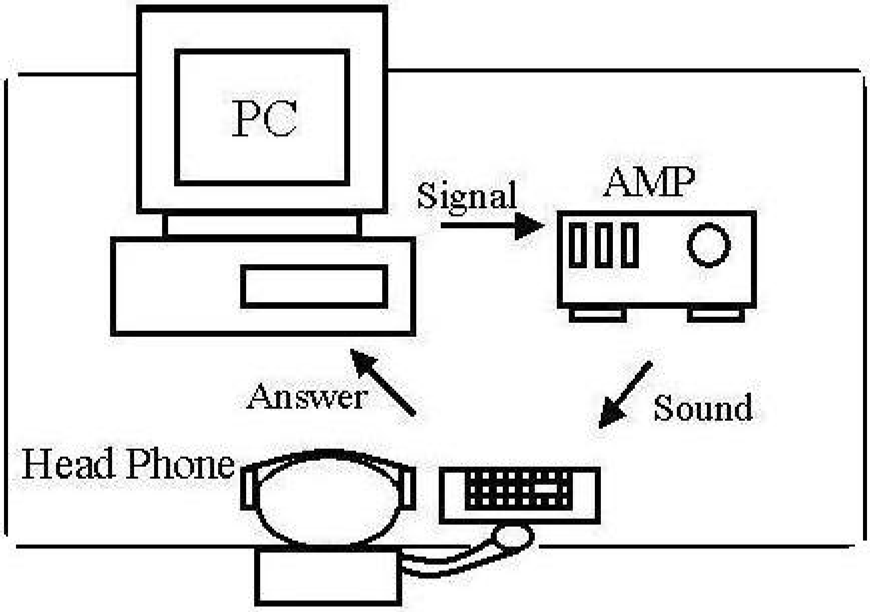


図 3.8 実験装置の構成

装置	名称
Computer	自作機 (CPU:1000Hz,Memory:128Mbyte)
Audio Amplifier	STAX SRM-T1S
Headphone	STAX Lambda Nova Signature

表 3.1 実験装置の仕様

４．実験と考察

4.1 予備実験 [共通して知覚される印象の確認]

4.1.1 実験の目的

人間の感性とは千差万別であるが、その様な中でも人間が音から共通して感じることのできる感覚は存在するのではないかと筆者は考えている。果たして、一体どの様な音がどの様な感覚を人間に共通して感じさせるのか。これを検証するために簡単な実験を行った。ここでは４つの刺激音を用いて、人がどのような音をどのように感じるのか、大まかな対応関係の確認する。また、この実験より先行研究で報告されていることがどこまで信用できるのか、自分が考えている音と印象の関係が、他の人にどこまで一致するのか確認したい。

4.1.2 刺激音

実験に用いた刺激音の、加工前の詳細を表 4.1 に示す。これらの音の立ち上がり・立ち下がりを除去し、ほぼ大きさ一定の持続状態として聴取できる刺激音を作り出した。各刺激音のサンプリング周波数は 44.1kHz、提示時間は 800msec に統一した。４つの刺激音は全て、多少の時間的な揺らぎがあるものの、ほとんどが過渡的な立ち上がり・立ち下がりのない定常状態の音として聴こえるものである。しかし、音圧についての厳密な測定はしておらず、筆者が同じ大きさの音に聴こえるように調節する程度に留まった。（およそ７０ｄB(SPL)である。）各刺激音のスペクトログラムは以下の通りである（図 4.1a〜4.1d）。

YAMAHA-MIDI-XG-Sound	
Stimulus1.1	Program No49 -String Ensemble1-
Stimulus1.2	Program No17 -Drawbar Organ-
Stimulus1.3	Program No74 -Flute-
Stimulus1.4	Program No41 -Violin-

表 4.1 加工前の刺激音の詳細

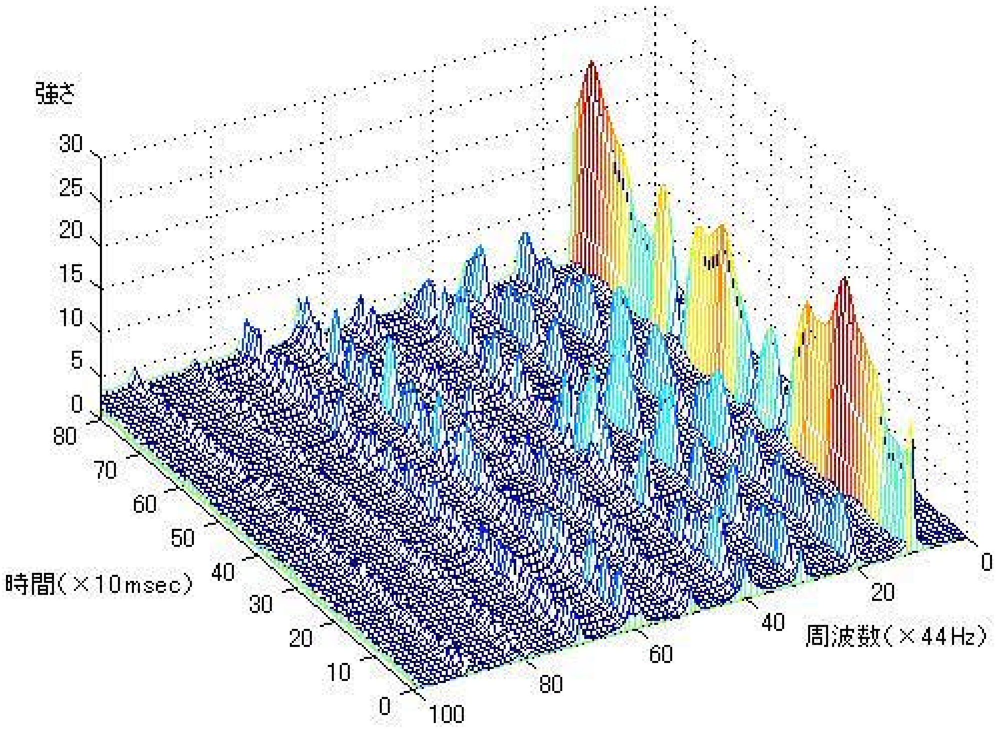


図 4.1a Stimulus1.1 のスペクトログラム

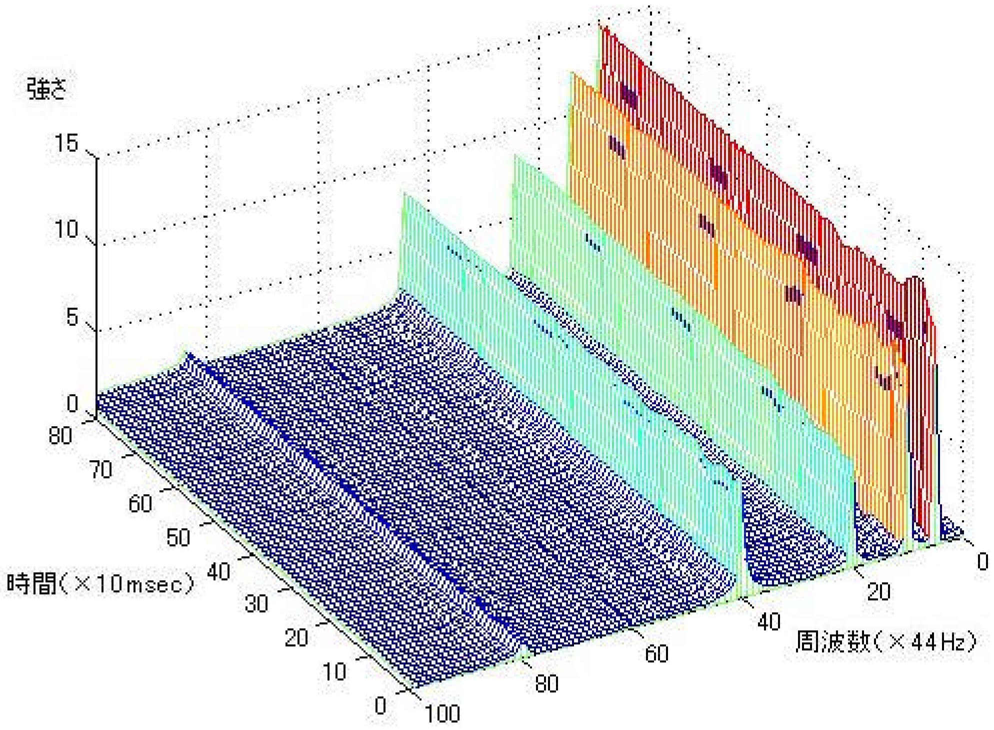


図 4.1b Stimulus1.2 のスペクトログラム

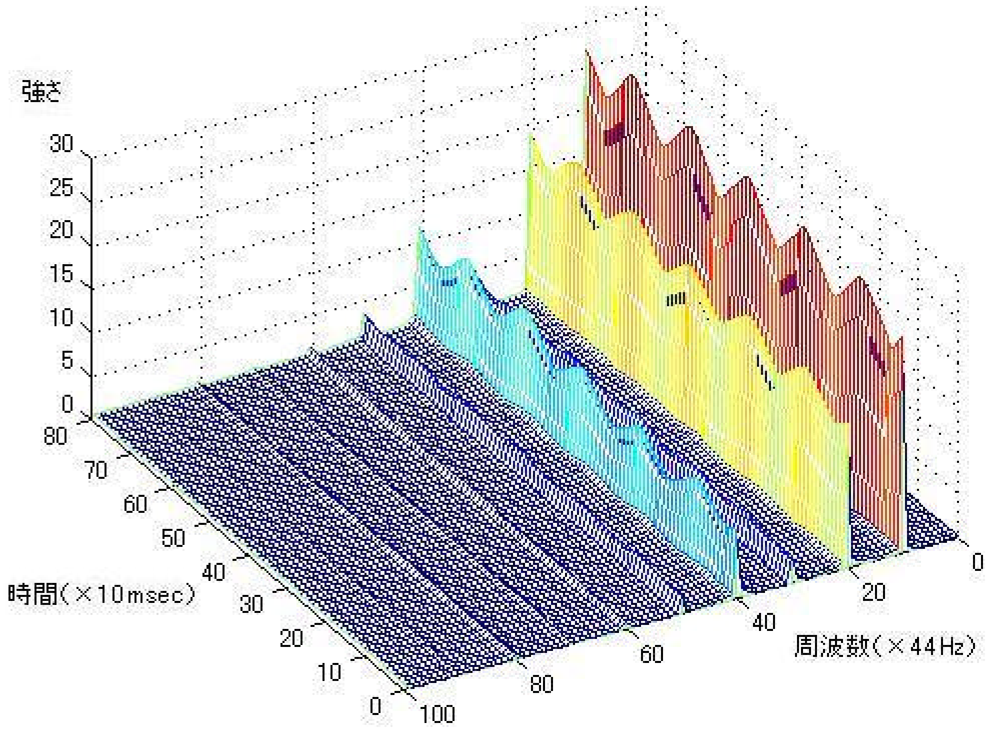


図 4.1c Stimulus1.3 **のスペクトログラム**

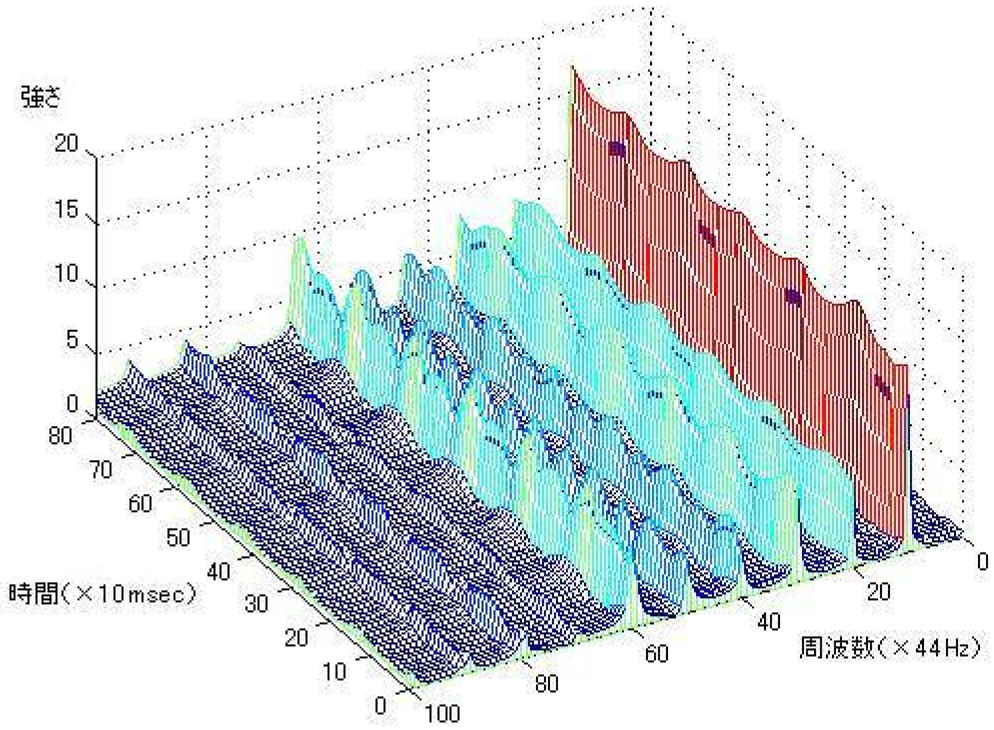


図 4.1d Stimulus1.4 **のスペクトログラム**

4.1.3 実験方法

実験にはサーストンの一対比較法を用いた。一対比較法とは被験者に対して、複数個の刺激の中から、２つ選択する組み合わせとなる全通りの刺激対を提示し、その後で被験者に特定の判断基準に基づいて、どちらかの刺激を選択してもらうという心理実験方法である。この結果をサーストンの比較判断の法則で処理することにより、刺激間においての「特定の判断基準」に対する間隔尺度を構成することができる。

被験者には、４つの刺激音からなる刺激対（順番も考慮してここでは１２通りの対）をランダムに提示して、どちらの音が判断基準となる感覚を強く感じるか聴き比べて、強制選択してもらう。判断基準となる尺度には、3.2.1 に示す９つの表現語を用いた。これより、被験者には全部で１０８の刺激対について、それぞれの判断基準に基づき評定してもらうことになる。以下に、実験のイメージ図を示す（図 4.2）。刺激音は「Stimulus1.x」→「500msec の無音状態」→「Stimulus1.y」の順で 2 度繰り返し提示される。

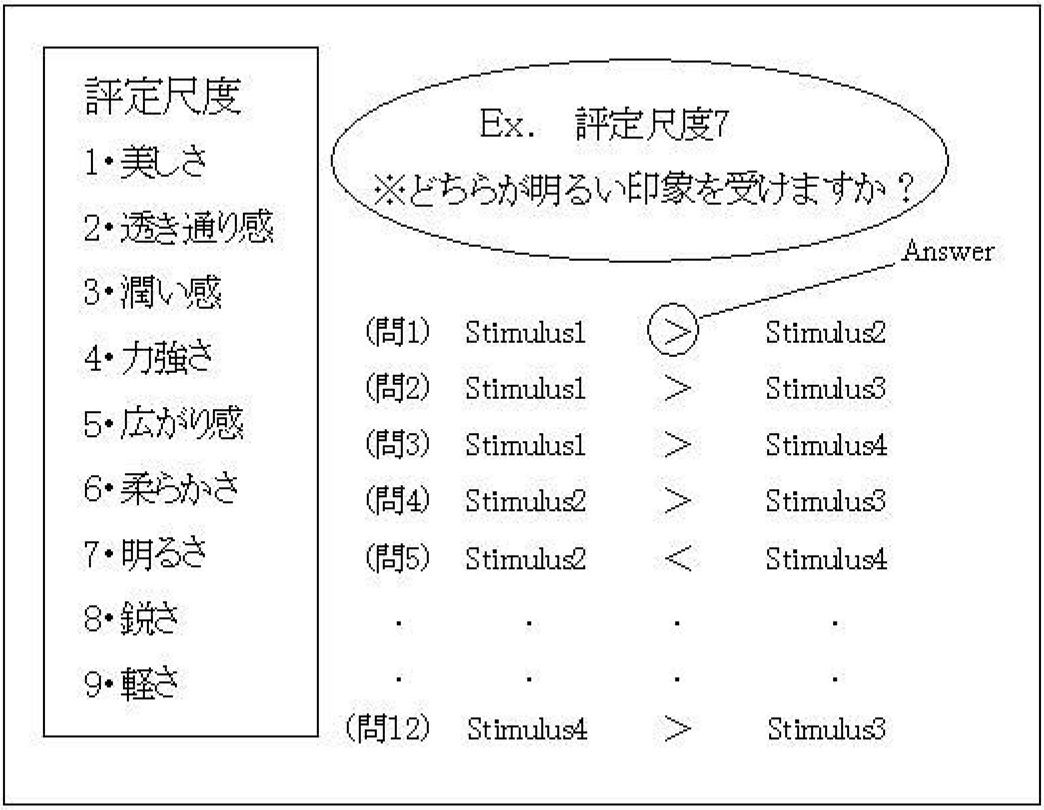


図 4.2　実験１のイメージ

4.1.4 実験結果と考察

＜結果＞

　実験を５人の被験者に対して行った結果を図 4.3 に示す。

1

Si ＞ Si				
i＼j	1	2	3	4
1		6	8	9
2	4		8	10
3	2	2		8
4	1	0	2	
「美しさ」				

2

Si ＞ Si				
i＼j	1	2	3	4
1		8	8	7
2	2		4	6
3	2	6		4
4	3	4	6	
「透き通り感」				

3

Si ＞ Si				
i＼j	1	2	3	4
1		2	4	2
2	8		5	7
3	6	5		8
4	8	3	2	
「潤い感」				

4

Si ＞ Si				
i＼j	1	2	3	4
1		5	5	3
2	5		5	2
3	5	5		5
4	7	8	5	
「力強さ」				

5

Si ＞ Si				
i＼j	1	2	3	4
1		10	10	10
2	0		7	6
3	0	3		3
4	0	4	7	
「広がり感」				

6

Si ＞ Si				
i＼j	1	2	3	4
1		2	2	5
2	8		3	8
3	8	7		10
4	5	2	0	
「柔らかさ」				

7

Si ＞ Si				
i＼j	1	2	3	4
1		9	10	10
2	1		2	2
3	0	8		3
4	0	8	7	
「明るさ」				

8

Si ＞ Si				
i＼j	1	2	3	4
1		8	9	8
2	2		3	1
3	1	7		1
4	2	9	9	
「軽さ」				

9

Si ＞ Si				
i＼j	1	2	3	4
1		9	8	7
2	1		5	2
3	2	5		1
4	3	8	9	
「鋭さ」				

図 4.3　実験結果

（刺激音 S_i と S_j の比較結果：図は (刺激音 S_i) > (刺激音 S_j) と判断された選択度数を表す。S_1：Stimulus1.1 , S_2：Stimulus1.2 , S_3：Stimulus1.3 , S_4：Stimulus1.4）

＜考察＞

　この実験結果から、以下の２つの考察をした。

・明るさ因子群の高い相関性について

　判断基準７～９である「明るさ」・「鋭さ」・「軽さ」に関しては、多少の差があるものの、各刺激音から受ける印象の順位差が、以下の通りに全て同じであった。

Stimulus1.1　＞　Stimulus1.4　＞　Stimulus1.3　＞　Stimulus1.2

　この結果から、音から受ける印象の中で「明るい」・「鋭い」・「軽い」という感覚は、同じ物理要因によって引き起こされている可能性が大きいと推測できる。しかしながら、ここではその物理要因の推測はできるものの、断定には到らない。各刺激音のスペクトログラム及び単一時刻における周波数構造を見る限りでは、やはり先行研究から報告されるように、高周波成分の強さが影響していると思われる。また、stimulus1.1 と stimulus1.4 の周波数波形に見られる側帯波と、スペクトログラムに見られる成分周波数の時間的な揺らぎから、明るさ因子に代表される感覚は、振幅変調の物理要因に関しても強く影響されるのではないかと考えられる。

・音に広がり感を与える物理要因について

　この実験において、Stimulus1.1 の評価は全ての判断基準に対して大きかった。その中でも、特に「広がり感」という感覚尺度に対しては、５人の被験者全員が例外なく「一番広がりある感じがする」と答えている。この点に注目して、各刺激音のスペクトログラムを見比べてみると、Stimulus1.1 と他の刺激音の決定的な違いは、振幅変調による変調の深さではないかと推測できる。これより、筆者は「音の広がり感」という印象に関しては、変調周波数との兼ね合いもあるが振幅変調の深さが非常に強く影響しているのではないかと考える。

＜反省点＞

　この実験方法では、既存の音に対しての大まかな音と感覚との分類は可能だが、どのような物理量が音の印象に直に関係しているのか絞り込むことができない。すなわち、この実験結果でもって、刺激間の相対距離をサーストンの比較判断の方法で求めたとしても、距離に対応する物理量を一意に絞り込むことができない。

　よって、実験１では既存の音響信号に特定の操作を加えることにより音を変化させ、その操作に対する心理量の変化を調べることによって、音波の物理量と音の印象を対応させせることにする。

4.2 心理実験１ [各物理操作が音の印象に与える影響の差]

4.2.1 実験の目的

音波の物理量と音の印象の対応関係を調べる本研究では、既存の音響信号に特定の操作を加えることにより音を変化させ、その操作に対する心理量の変化を調べる。ある特定の物理操作による変化量が音の印象に対してどのように影響するか、この対応関係を求めたい。そのために、まずは 3.2.2 で示した４つの物理操作がどのような印象を強く引き起こす要因になるか、大まかにこれを特定することにした。

4.2.2 刺激音

刺激音には 3.1 に示す基準音と、これを 3.2.2 で示した４つの物理操作で変化させた音、計５つの音を用いた。それぞれの物理操作に用いたパラメータの値と、各刺激音のスペクトログラムを以下に示す（表 4.2，図 4.4a～4.4e)。刺激音のサンプリングレートは 44.1kHz で、提示時間は 500msec、高さは 1000Hz、聴こえの大きさは、基準音（70dB(SPL)）と同等の評価が得ることができる大きさに統一した。

刺激音	物理操作	各パラメータと詳細
2.1	無	基準音
2.2	振幅変調	D_{am}=0.2 , f_{am}=6
2.3	周波数変調	D_{fm}=1 , f_{fm}=6
2.4	エンベロープ	AT=1000 , DC=10 , Sus=0.5
2.5	倍音追加	A_{ha}=0.1

表 4.2 刺激音の詳細

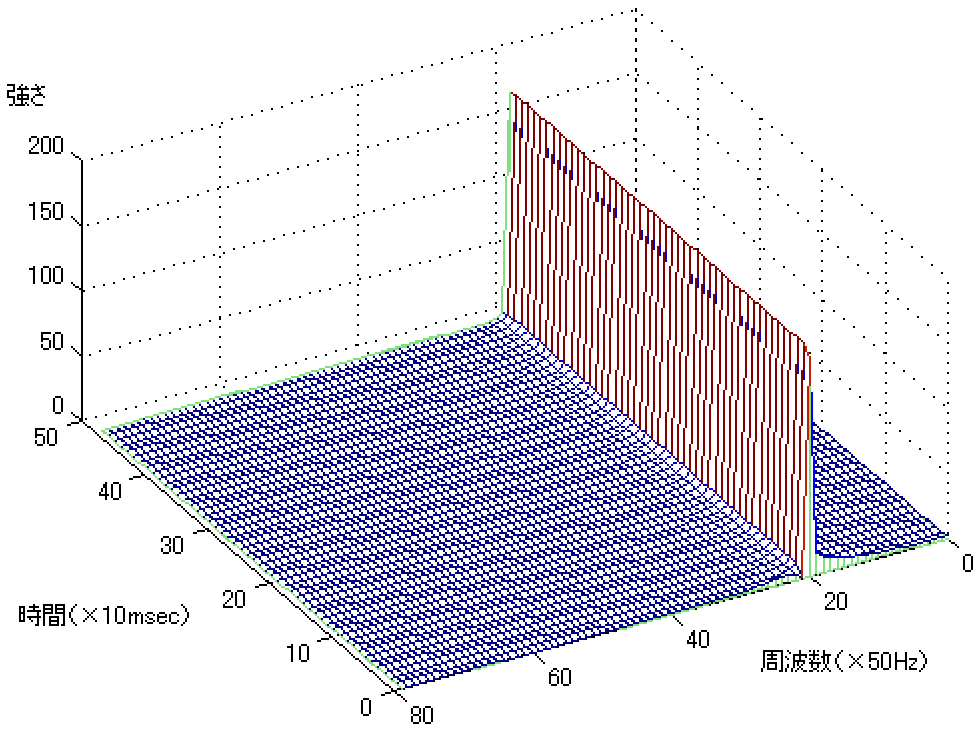


図 4.4a　Stimulus2.1 のスペクトログラム

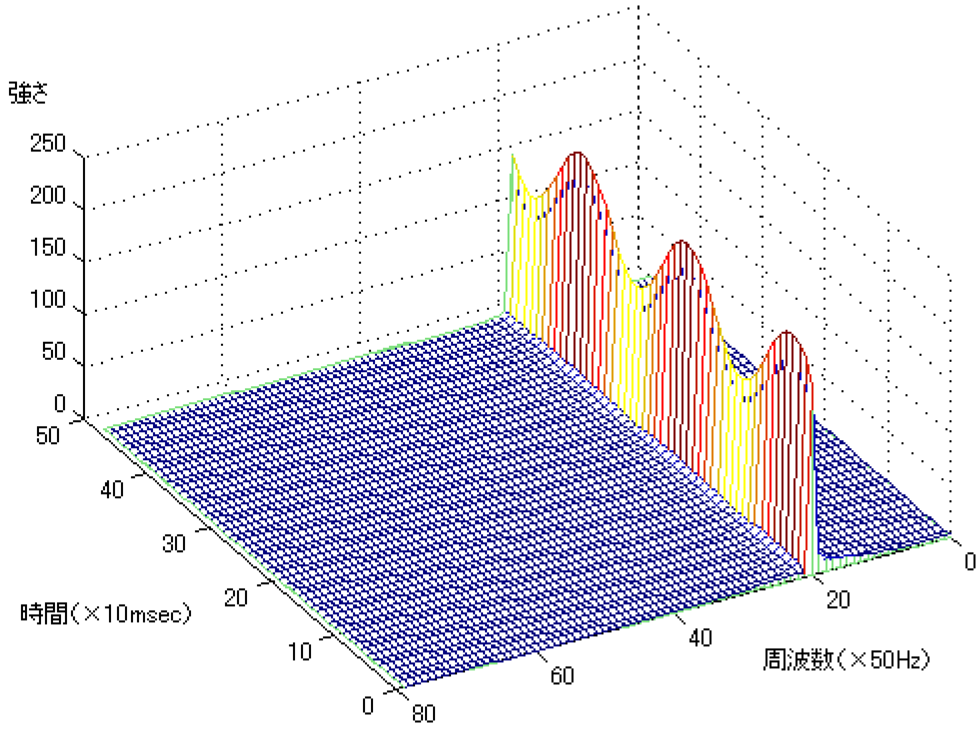


図 4.4b　Stimulus2.2 のスペクトログラム

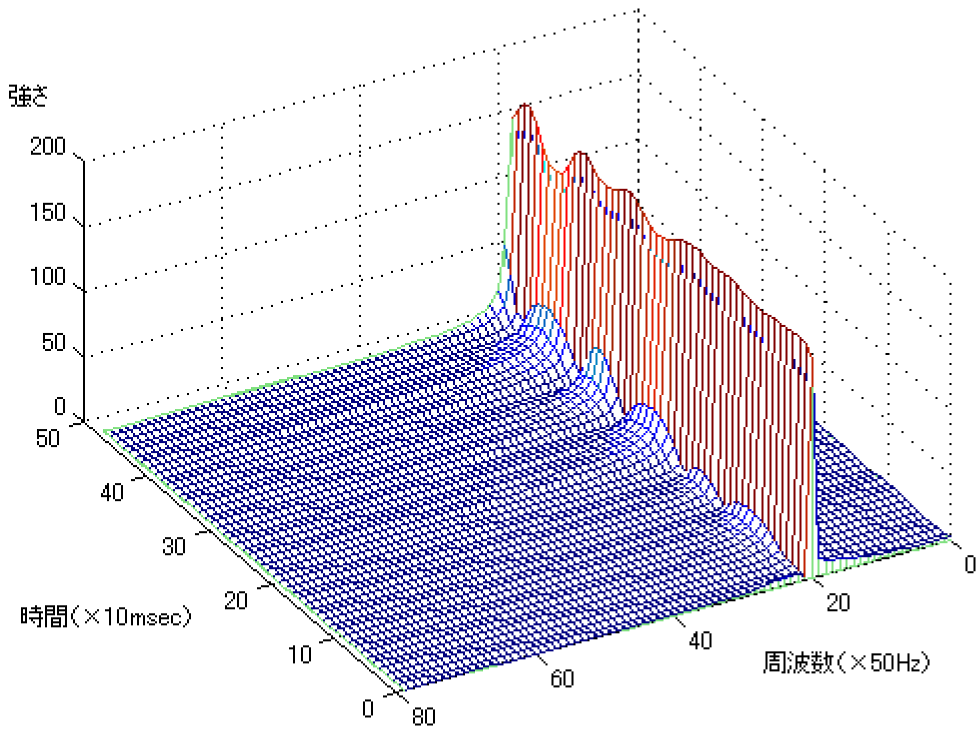


図 4.4c　Stimulus2.3 のスペクトログラム

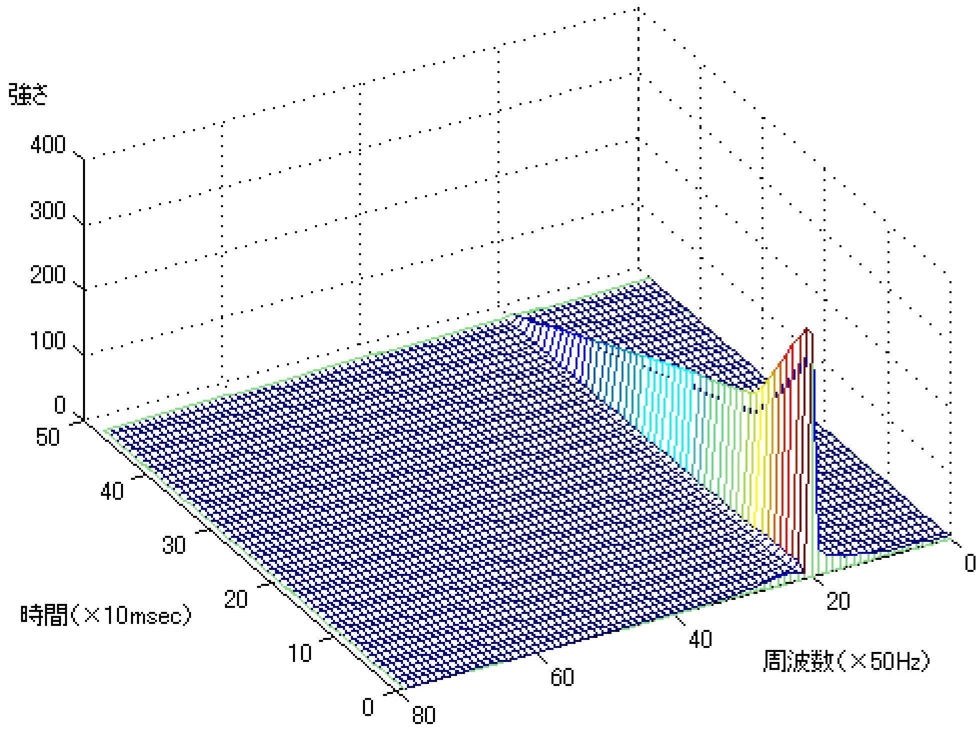


図 4.4c　Stimulus2.4 のスペクトログラム

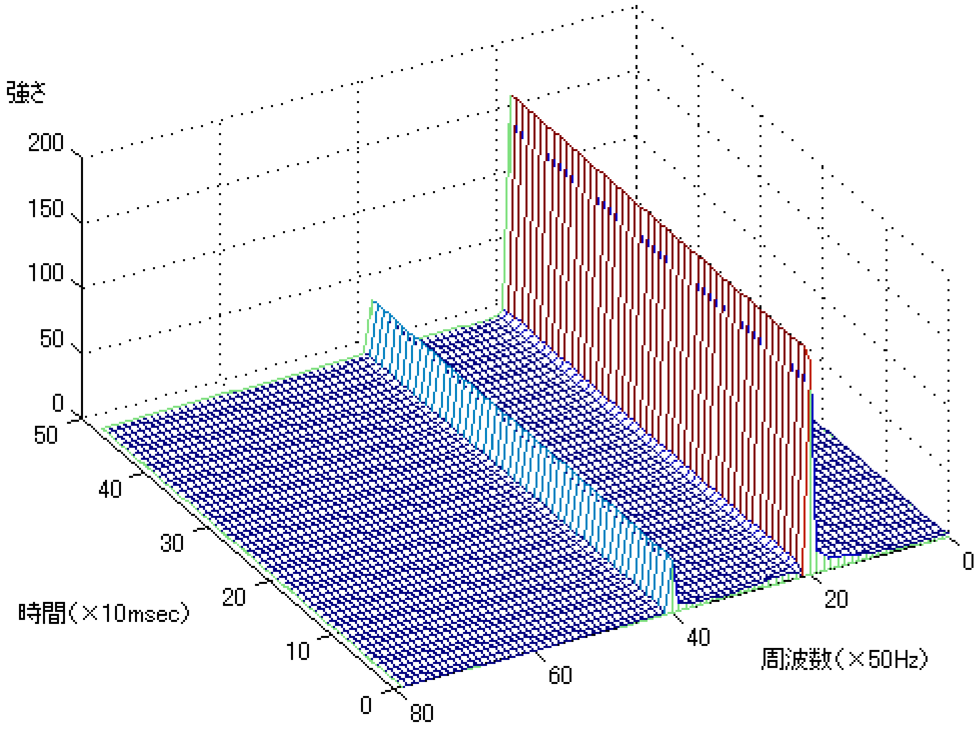


図 4.4e　Stimulus2.5 のスペクトログラム

これらの刺激音は、3.2.2 の計算を MATLAB 上で実行することにより作製した。また、全ての刺激音は基準音と比べて少し異なって聴こえる程度に調節した。比較対象を、「基準音から少しずらした範囲」と限定して議論しなければ、心理量と物理量の対応関係が複雑化するからである。本研究では物理量と心理量の対応関係が、この限定された範囲では少なくとも一次元的な対応をすると仮定している。しかしその対応が、線形的か、対数的か、どのような関係でもって成立するかは分からない。

4.2.3 実験方法

実験には、サーストンの一対比較法を用いた。被験者には、５つの刺激音からなる刺激対（２０通りの対）をランダムに提示して、どちらの音が判断基準となる感覚を強く感じるか聴き比べて、強制選択してもらう。判断基準となる尺度には、3.2.1 に示す９つの表現語の中から「美しさ」・「広がり感」・「明るさ」・「硬さ」の４つの表現語を、それぞれの因子群を代表する尺度として用いた。これより、被験者には全部で８０の刺激対について、それぞれの判断基準に基づき評定してもらうことになる。実験１に用いた手法は、比較する刺激音が異なること以外は、予備実験と全く同じである。

4.2.4 実験結果と考察

＜結果＞

実験を１０人の被験者に対して行った結果を図 4.5 に示す。

Ⅰ「美しさ」（Si＞Si）					
i＼j	1	2	3	4	5
1	＼	3	6	3	13
2	17	＼	15	11	16
3	14	5	＼	7	11
4	17	9	13	＼	16
5	7	4	9	4	＼

Ⅱ「広がり感」（Si＞Sj）					
i＼j	1	2	3	4	5
1	＼	3	9	11	16
2	17	＼	15	12	17
3	11	5	＼	13	17
4	9	8	7	＼	12
5	4	3	3	8	＼

Ⅲ「明るさ」（Si＞Sj）					
i＼j	1	2	3	4	5
1	＼	10	11	1	9
2	10	＼	13	2	9
3	9	7	＼	2	6
4	19	18	18	＼	16
5	11	11	14	4	＼

Ⅳ「硬さ」（Si＞Sj）					
i＼j	1	2	3	4	5
1	＼	11	12	3	6
2	9	＼	10	2	8
3	8	10	＼	3	7
4	17	18	17	＼	15
5	12	12	13	5	＼

図 4.5　実験２の結果

（刺激音 S_i と S_j の比較結果：図は (刺激音 S_i) ＞ (刺激音 S_j) と判断された選択度数を表す。S_1：Stimulus2.1 , S_2：Stimulus2.2 , S_3：Stimulus2.3 , S_4：Stimulus2.4 , S_5：Stimulus2.5）

＜考察＞

サーストンの一対比較法で求められる尺度は順序尺度であり、判断が正規分布をする仮定のもとに、間隔尺度に変換する手続きを行う。そのために被験者の数は５０名以上が好ましいとされているが、規模の大きさを考慮した結果、被験者の数は１０人に留まった。しかしながら、この結果から求めた間隔尺度においても十分、音の印象と各物理操作の大まかな対応関係がつかめると考えている。

図 4.5 にある結果から、サーストンの比較判断の法則に基づき、４つの尺度に対して、各刺激間の相対距離を求める。図 4.5 の選択度数を被験者の数で割った(刺激音 S_i) > (刺激音 S_j)の選択比率を図 4.6 に、そしてこの確率を正規変換表からｚ変換し（標準偏差・６８％の確率をｚ＝１とする)、ｚの値を平均することより刺激間の相対距離を求めた結果を図 4.7a～4.7d に示す。

Ⅰ「美しさ」（Si＞Si）					
i＼j	1	2	3	4	5
1		15	30	15	65
2	85		75	55	80
3	70	25		35	55
4	85	45	65		80
5	35	20	45	20	

Ⅱ「広がり感」（Si＞Sj）					
i＼j	1	2	3	4	5
1		15	45	55	80
2	85		75	60	85
3	55	25		65	85
4	45	40	35		60
5	20	15	15	40	

Ⅲ「明るさ」（Si＞Sj）					
i＼j	1	2	3	4	5
1		50	55	5	45
2	50		65	10	45
3	45	35		10	30
4	95	90	90		80
5	55	55	70	20	

Ⅳ「硬さ」（Si＞Sj）					
i＼j	1	2	3	4	5
1		55	60	15	30
2	45		50	10	40
3	40	50		15	35
4	85	90	85		15
5	60	60	65	25	

図 4.6　$S_i > S_j$ の選択比率

Ⅰ「美しさ」（Si＞Si）					
i＼j	1	2	3	4	5
1	0	1.04	0.52	1.04	−0.39
2	−1.04	0	−0.68	−0.12	−0.84
3	−0.52	0.68	0	0.39	−0.12
4	−1.04	0.12	−0.39	0	−0.84
5	0.39	0.84	0.12	0.84	0
合計	−2.21	2.68	−0.43	2.15	−2.19
平均	−0.553	0.67	−0.108	0.538	−0.548
差	0	1.223	0.446	1.091	0.006

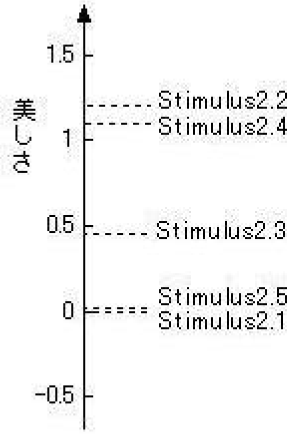


図 4.7a　「美しさ」の尺度における刺激間の相対距離

Ⅱ「広がり感」（Si > Si）					
i＼j	1	2	3	4	5
1	0	1.04	0.12	−0.12	−0.84
2	−1.04	0	−0.68	−0.25	−1.04
3	−0.12	0.68	0	−0.39	−1.04
4	0.12	0.25	0.39	0	−0.25
5	0.84	1.04	1.04	0.25	0
合計	−0.2	3.01	0.87	−0.51	−3.17
平均	−0.05	0.753	0.218	−0.128	−0.793
差	0	0.803	0.268	−0.078	−0.743

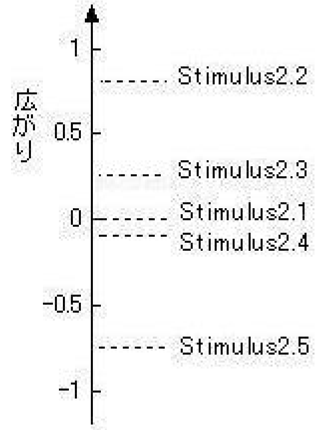


図 4.7b　「広がり感」の尺度における刺激間の相対距離

Ⅲ「明るさ」（Si > Si）					
i＼j	1	2	3	4	5
1	0	0	−0.12	1.65	0.12
2	0	0	−0.39	1.28	0.12
3	0.12	0.39	0	1.28	0.52
4	−1.65	−1.28	−1.28	0	−0.84
5	−0.12	−0.12	−0.52	0.84	0
合計	−1.65	−1.01	−2.31	5.05	−0.08
平均	−0.413	−0.253	−0.578	1.263	−0.02
差	0	0.16	−0.165	1.676	0.393

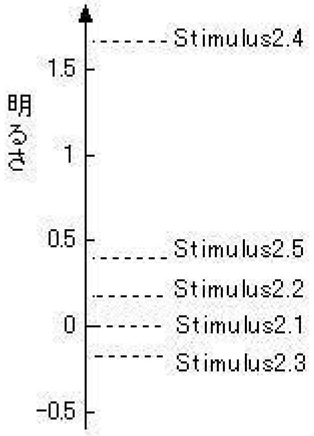


図 4.7c　「明るさ」の尺度における刺激間の相対距離

Ⅳ「硬さ」（Si > Si）					
i＼j	1	2	3	4	5
1	0	−0.12	−0.25	1.04	0.25
2	0.12	0	0	1.28	0.25
3	0.25	0	0	1.04	−0.12
4	−1.04	−1.28	−1.04	0	−0.84
5	−0.25	−0.25	0.12	0.84	0
合計	−0.92	−1.65	−1.17	4.2	−0.46
平均	−0.23	−0.413	−0.293	1.05	−0.115
差	0	−0.183	−0.063	1.28	0.115

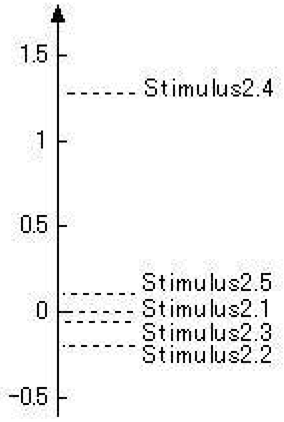


図 4.7d　「硬さ」の尺度における刺激間の相対距離

これらの結果から、基準音との差が大きかった物理操作について考察する。

・「美しさ」について

美しさの度合いは、全ての物理操作によって上昇した。特に振幅変調とエンベロープ操作についてはその差が大きく見られた。前述の通りに、振幅変調とは周期的な音の大きさの変化である。人間が音楽鑑賞にあたって、多少揺らぎがある音を好むのはよく知られたことである。おそらく、この結果も同じ理由によるものであると思われる。

また、エンベロープ操作を施した音が一番「美しい」とされた理由としては、人間は何の音だか分からない音に対して、一定の大きさを持続する音よりも、音の鳴り始めから終わりまでが時々刻々と分かる減衰音のほうが、自然な音と解釈できるためであると筆者は考える。人間が自然の音を好むことは、川のせせらぎや森林に鳴り響く鳥の声などの自然界に広がる音を聴くと落ち着くという事実からも明らかである。

・「広がり感」について

広がり感に一番大きく作用した物理操作は、予備実験での考察の通りに振幅変調であった。これより、振幅変調は音に広がり感を与える操作であることがいえる。しかし、どのようなパラメータでもって振幅変調すれば音により強い広がり感を持たすことができるかは、現状況では分からない。ただ、振幅変調と周波数変調による影響を強いということから、何かしら周期的な揺らぎがこの印象を引き起こす要因になっていると推測できる。

また、倍音成分の周波数を追加した音は、広がる感じに対する負の評価が大きかった。これは倍音成分の追加という操作が、「広がり感」のマイナス量と考えられる「まとまり感」に強く作用したためと考えられる。これは恐らく、周波数が整数倍にある成分を加えたことが、音にまとまり感を与えたのではないかと筆者は考える。2.2.1 で言及したように、振幅変調という物理操作は、数式側から見れば変調周波数に依存する側帯波成分を追加したと同じことである。成分周波数を加えるという同じ物理操作にもかかわらず、与える印象が全く逆になったということは、少なくとも成分周波数と加えることによる変化する「協和性」が音の「広がり感」－「まとまり感」という尺度の心理量に強く影響する性質であることは間違いないと思われる。

・「明るさ」について

先行研究からの報告によれば、音の「明るさ」は高次倍音成分の強さに強く影響するとされている。しかし、実験結果ではエンベロープ操作を施した音が跳び抜けて一番明るいという評価を受けた。Stimulus2.4 は他の刺激音と違い、立ち上がり・立ち下がりのはっきりした減衰音である。現時点においては、この物理要因が「明るさ」を引き起こすのではないかと推測される。

Stimulus2.4 以外の刺激音に関しては、ほぼ予想通りの結果となった。わずかな差であるが、高次倍音を含む Stimulus2.5 と、予備実験で考察した通りに振幅変調音である Stimulus2.2 が基準音よりも明るいと評価された。

・「硬さ」について

「明るさ」の次元と同様に、エンベロープ操作を施した音が跳び抜けて一番明るいという評価を受けた。その他の刺激間においては、あまり差が見られないが振幅変調と周波数変調による評価が低いことから、周期的な揺らぎは、音に「柔らかい」という印象を与えるのではないかと思われる。

＜反省点＞

ここでは 3.2.2 で定めた音の物理操作が、音の印象にどのように影響するのか、対象とする心理次元において、その差を求めることにより考察した。その結果、いくつかの表現語について対応する物理量を絞り込むことができた。しかし、この方法でもって 3.2.2 に示す６つの音の物理操作から、いくつかのパラメータ値を用いて各操作ごとに２～３つづの刺激音を作製し、3.2.1 で定めた６つの心理次元に対して同様の実験を行ったとすると、被験者には最低でも、1000 回以上の判断をしてもらうことになり、負担が大きすぎる。

よって、この方法を少し変えて、比較対象を全て基準音とすることで、刺激音の数を増やし、「音の印象の度合い」を被験者に、各心理次元において７段階評価的してもらうことで、被験者の負担を減らすことにした。

実験２では、この手法を用いて各物理操作による変化量とその心理次元の変化量の対応関係を調べる。

4.3 心理実験２ [音の操作に対する心理次元の変化量]

4.3.1 実験の目的

実験２では、「音の印象の度合い」を各心理次元において７段階に定め、各物理操作による変化量とその心理次元の変化量の対応を測ることにした。物理操作により変化した音は、基準音と比べてどれくらい印象に差が生じるか、これを被験者に７段階で評価してもらうことで、心理量の変化を定量的に測定することを試みる。

4.3.2 刺激音

音の印象は、音量や音程によっても異なって知覚される。これより、この実験では音の大きさと高さの変化も、音の物理操作の一つと考えて、計６つの物理操作によって基準音を変化させて刺激音を作製した。それぞれの物理操作に用いたパラメータの値と、各刺激音のスペクトログラムを以下に示す（表 4.3, 図 4.8a～4.8e）。刺激音のサンプリングレートは 44.1kHz で、提示時間は 500msec に統一した。また、音程操作以外の物理操作における刺激音の高さは 1000Hz に、音量操作以外の物理操作における刺激音の聴こえの大きさは、基準音（70dB(SPL)）と同等の評価が得ることができる大きさに統一した。

刺激音	物理操作	各パラメータと詳細
3.1	無	基準音
3.2	音量変化	振幅 2 倍
3.3	音量変化	振幅 1/2 倍
3.4	音程変化	周波数 2 倍
3.5	音程変化	周波数 1/2 倍
3.6	エンベロープ	AT=500 , DC=20 , Sus=0.5
3.7	エンベロープ	AT=10 , DC=0.01 , Sus=0.5
3.8	振幅変調	D_{am}=0.5 , f_{am}=6
3.9	振幅変調	D_{am}=1 , f_{am}=12
3.10	周波数変調	D_{fm}=1 , f_{fm}=10
3.11	周波数変調	D_{fm}=10 , $\mathrm{f_{fm}}$=1
3.12	周波数変調	D_{fm}=10 , f_{fm}=10
3.13	倍音追加	A_{ha}=0.1
3.14	倍音追加	A_{ha}=1

表 4.3 刺激音の詳細

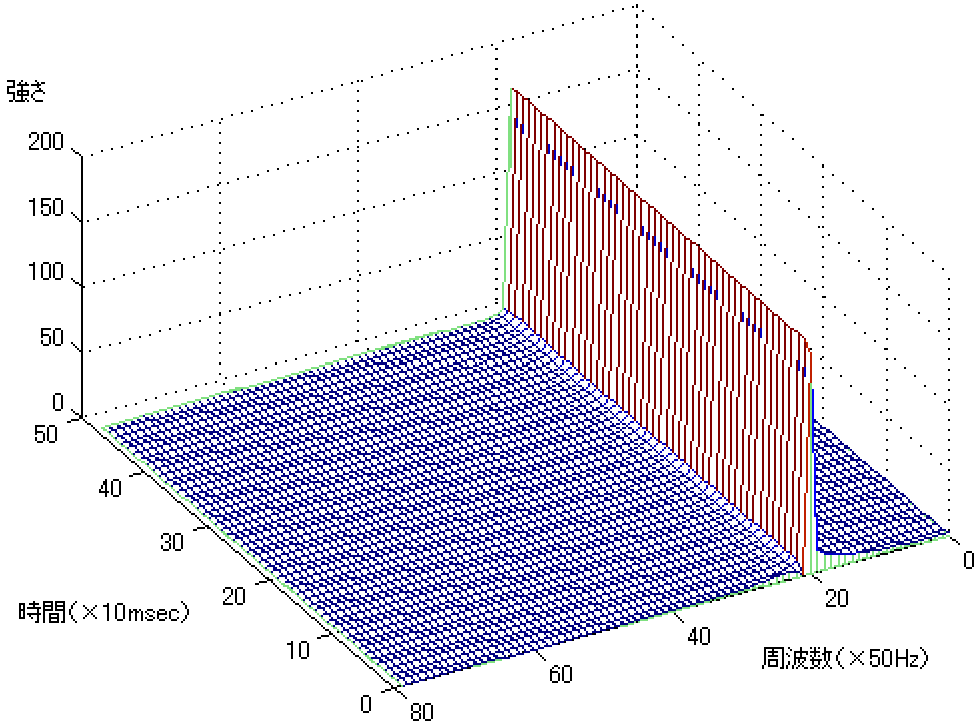


図 4.8a　Stimulus3.1 のスペクトログラム

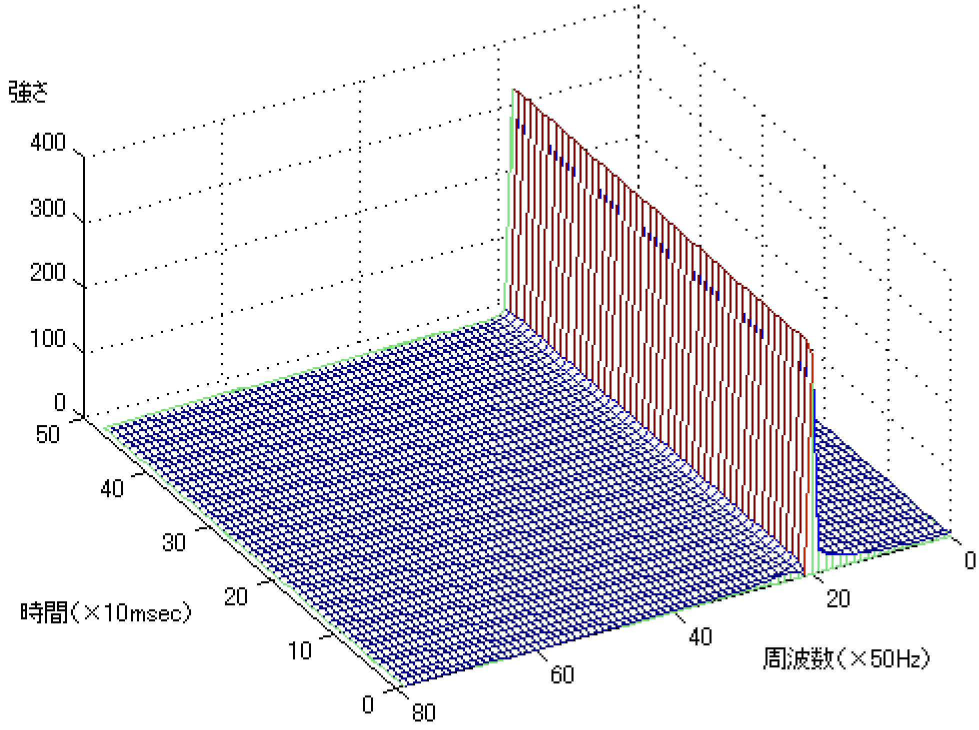


図 4.8b　Stimulus3.2 のスペクトログラム

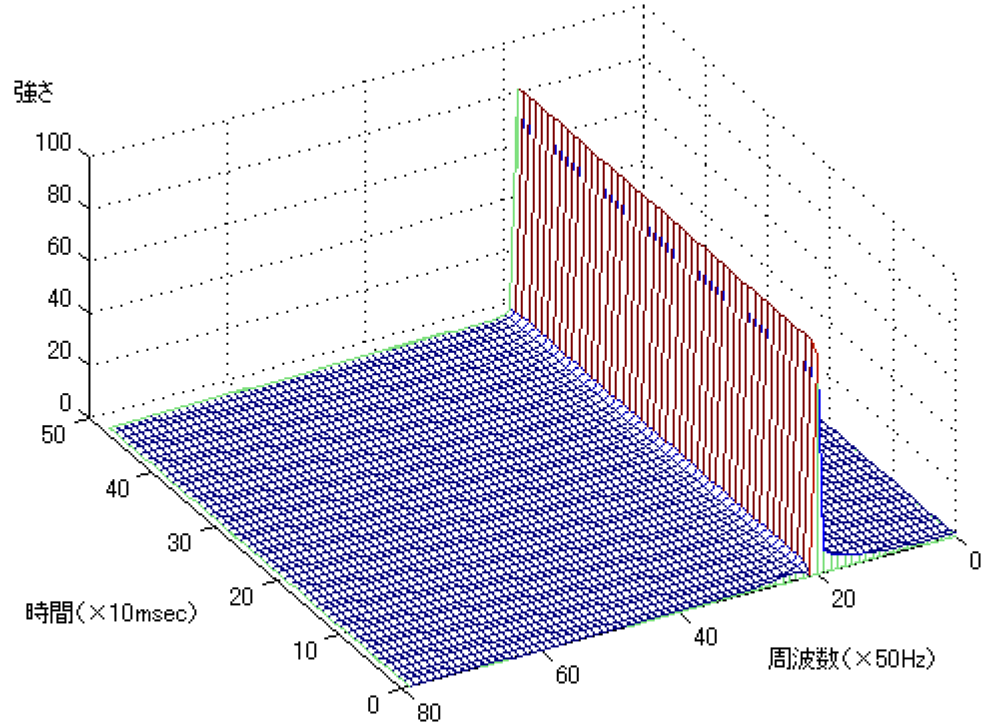


図 4.8c Stimulus3.3 のスペクトログラム

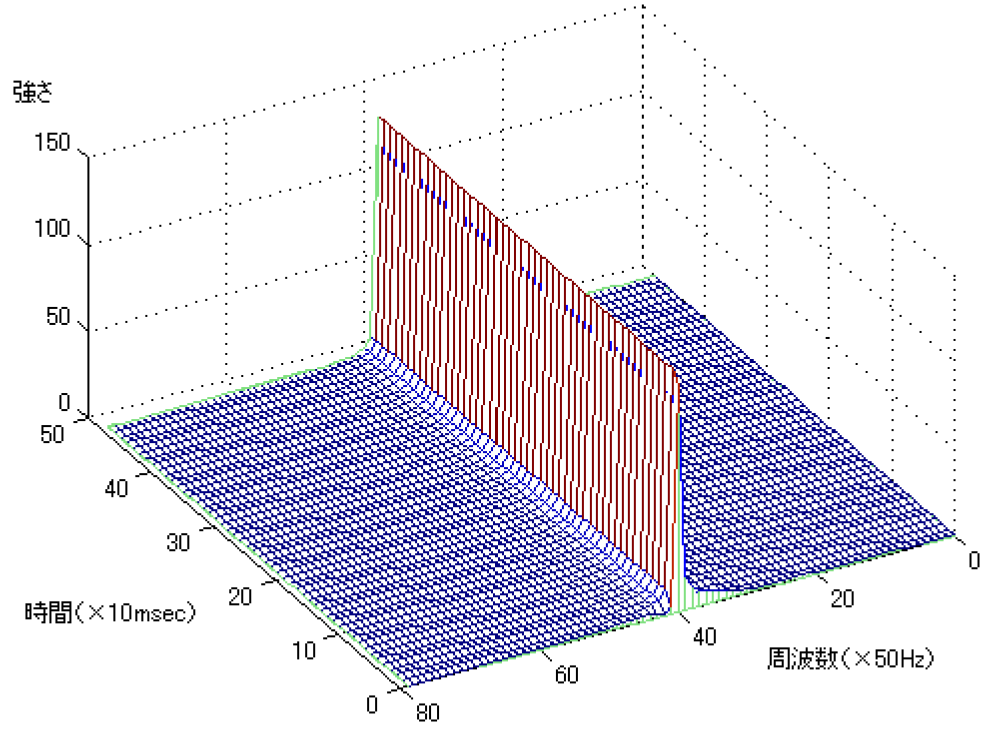


図 4.8d Stimulus3.4 のスペクトログラム

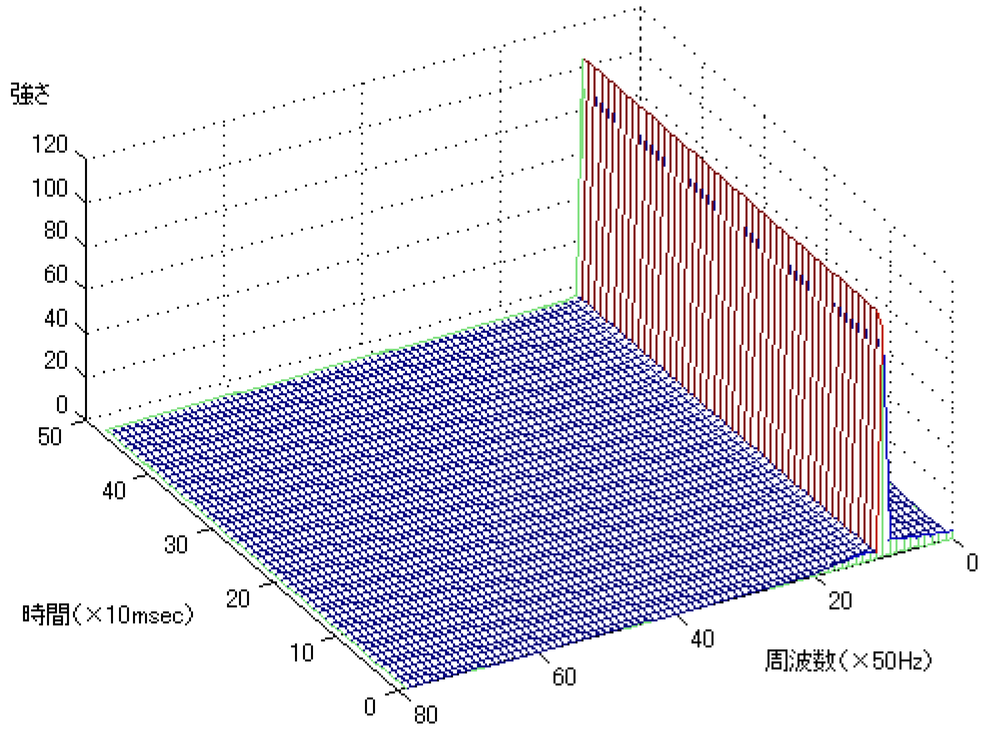


図 4.8e　Stimulus3.5 のスペクトログラム

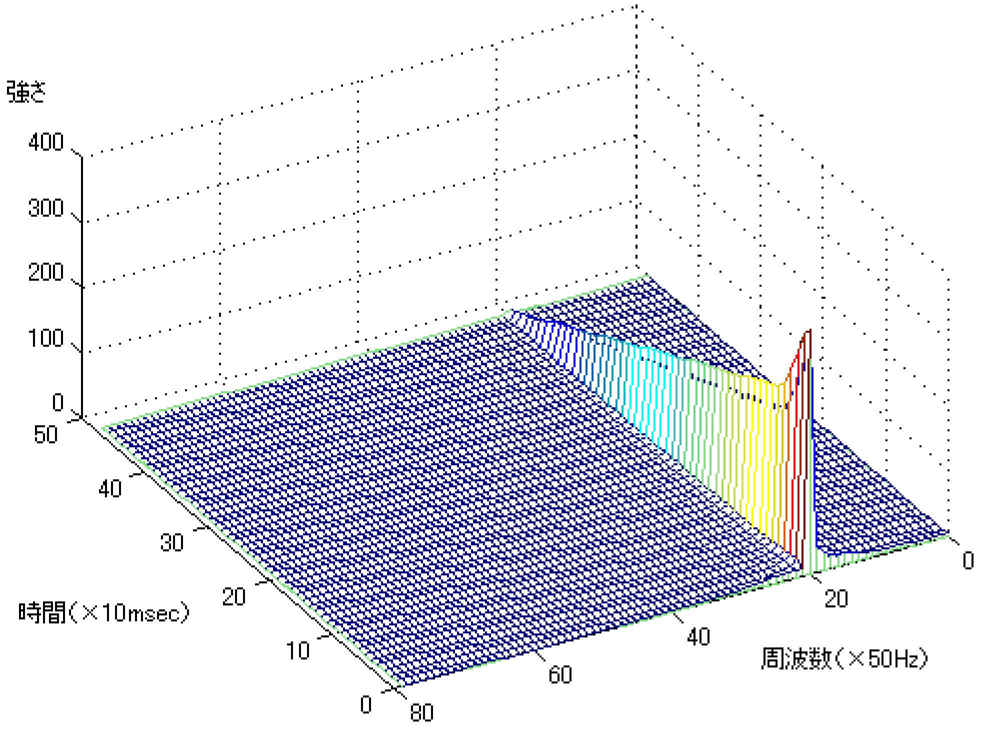


図 4.8f　Stimulus3.6 のスペクトログラム

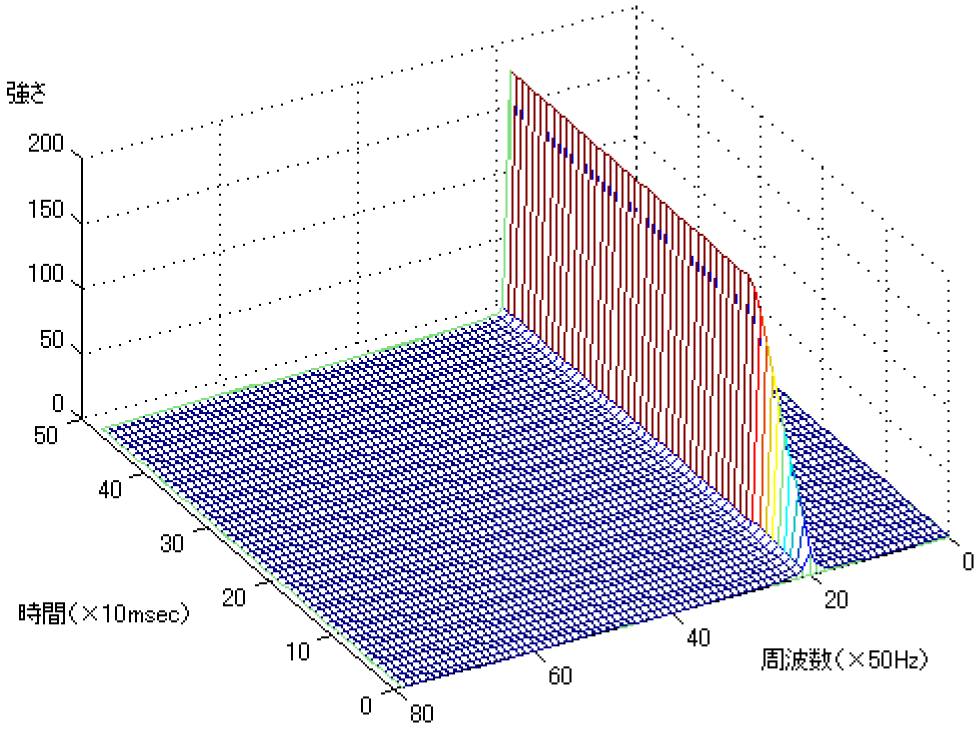


図 4.8g　Stimulus3.7 のスペクトログラム

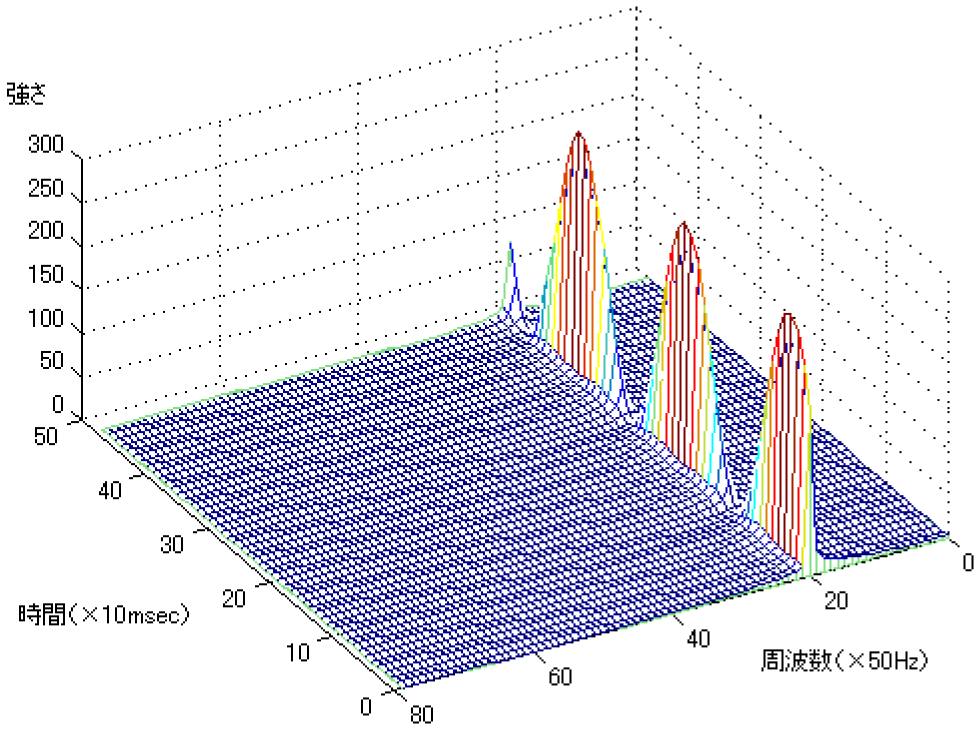


図 4.8h　Stimulus3.8 のスペクトログラム

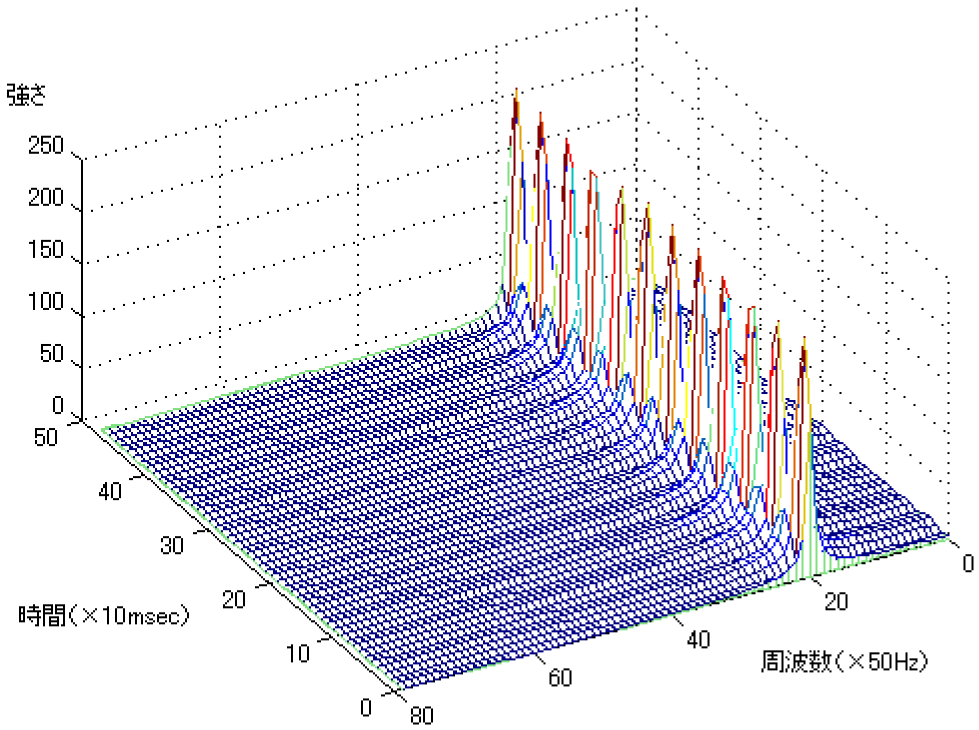


図 4.10i　Stimulus3.9 のスペクトログラム

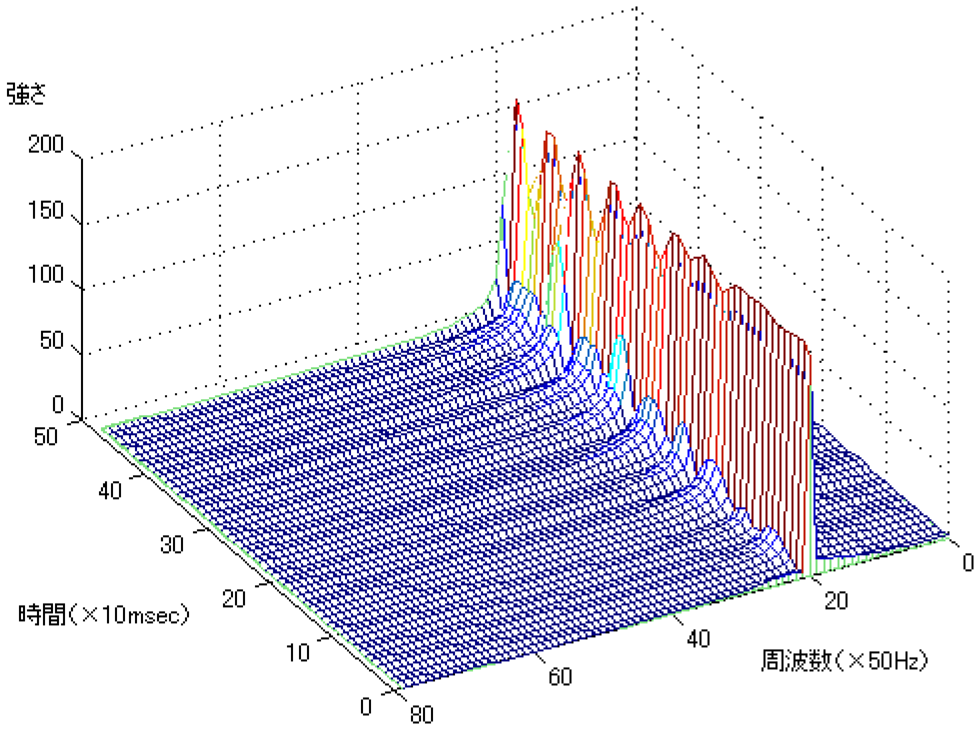


図 4.10j　Stimulus3.10 のスペクトログラム

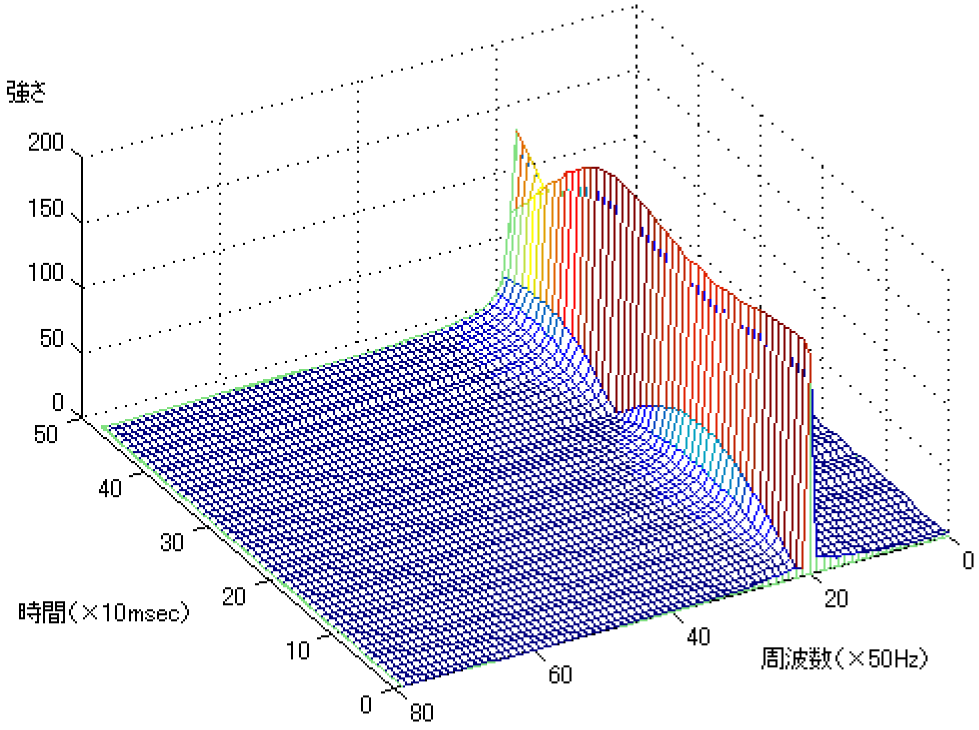


図 4.10k　Stimulus3.11 のスペクトログラム

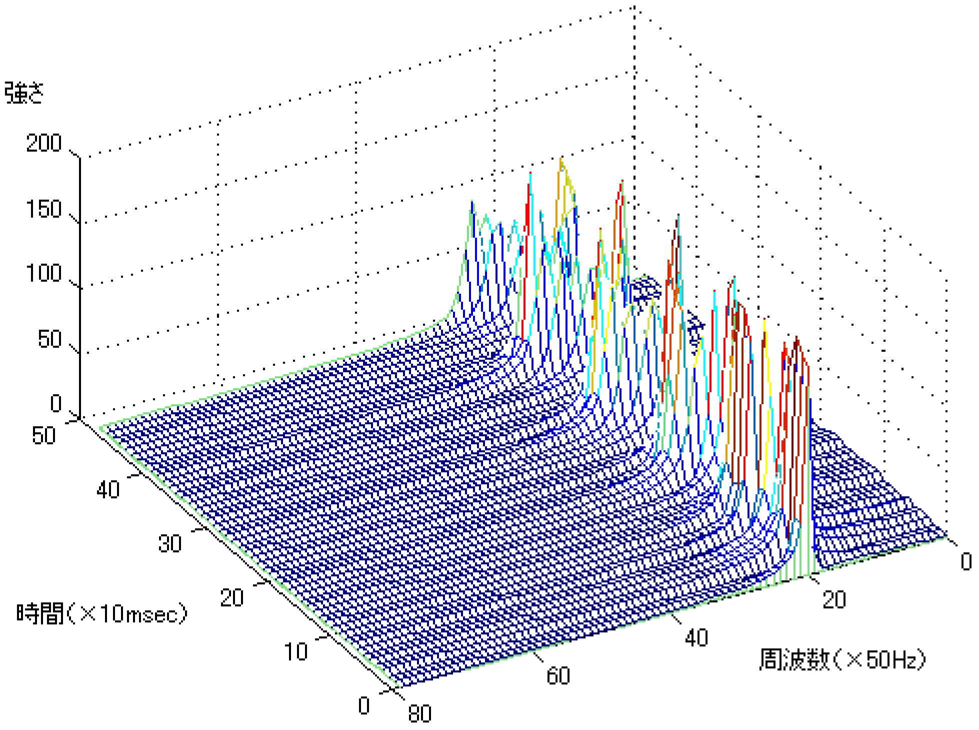


図 4.10l　Stimulus3.12 のスペクトログラム

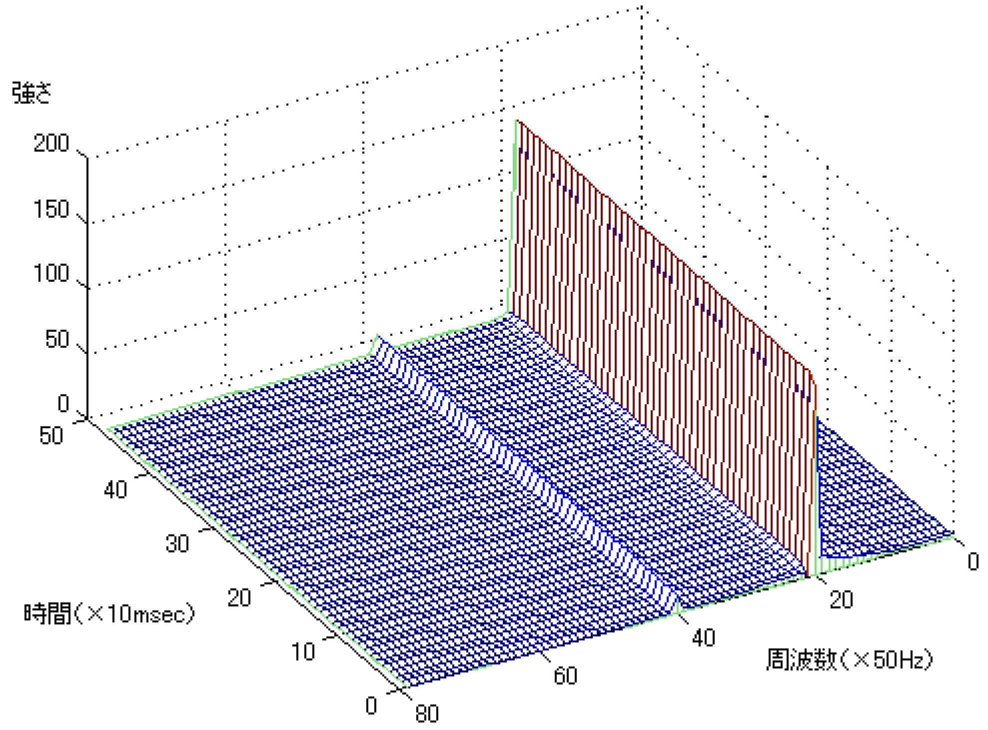


図 4.10m　Stimulus3.13 のスペクトログラム

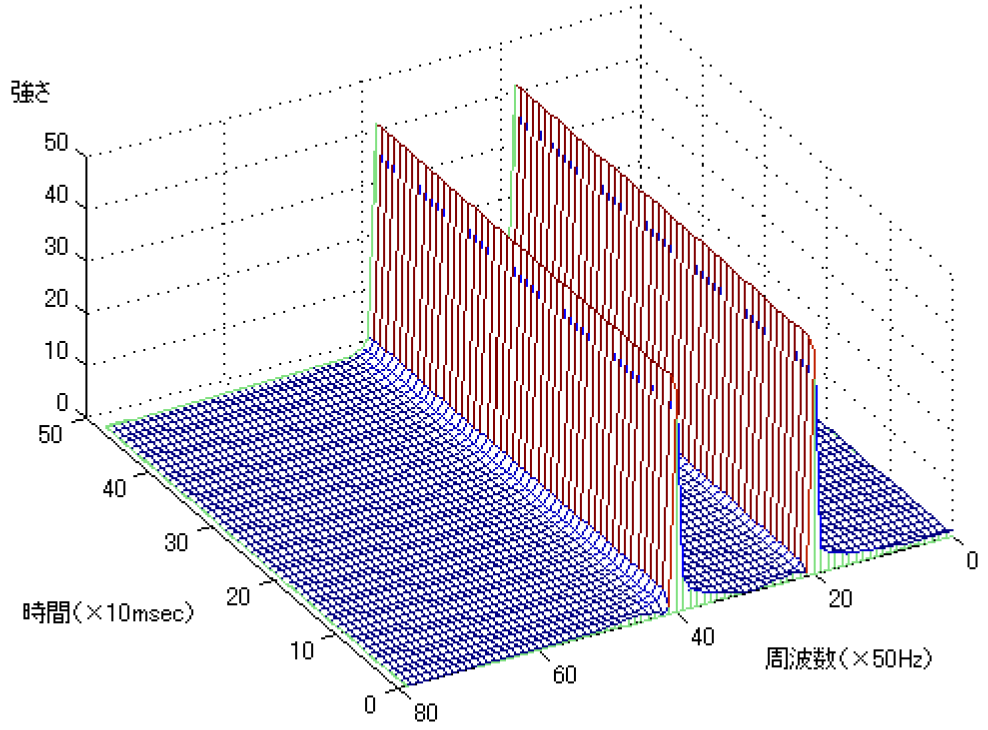


図 4.10n　Stimulus3.14 のスペクトログラム

4.3.3 実験方法

実験には、基準音と各物理操作を施した音の印象の差を７段階評定する手法を用いた。被験者には、基準音と各１３種の刺激音を聴き比べてもらい、2.1.3 に示した９つの感覚次元それぞれについて７段階の評価点をつけてもらう。刺激音は、「基準音」→「500msec の無音」→「Stimulus3.x」の順で 2 度繰り返し提示される。被験者は、一つ一つの心理次元である判断基準に注目し、１３組の刺激対に対して全部で１１７回の評価をすることになる。以下に、実験のイメージ図を示す（図 4.9）。

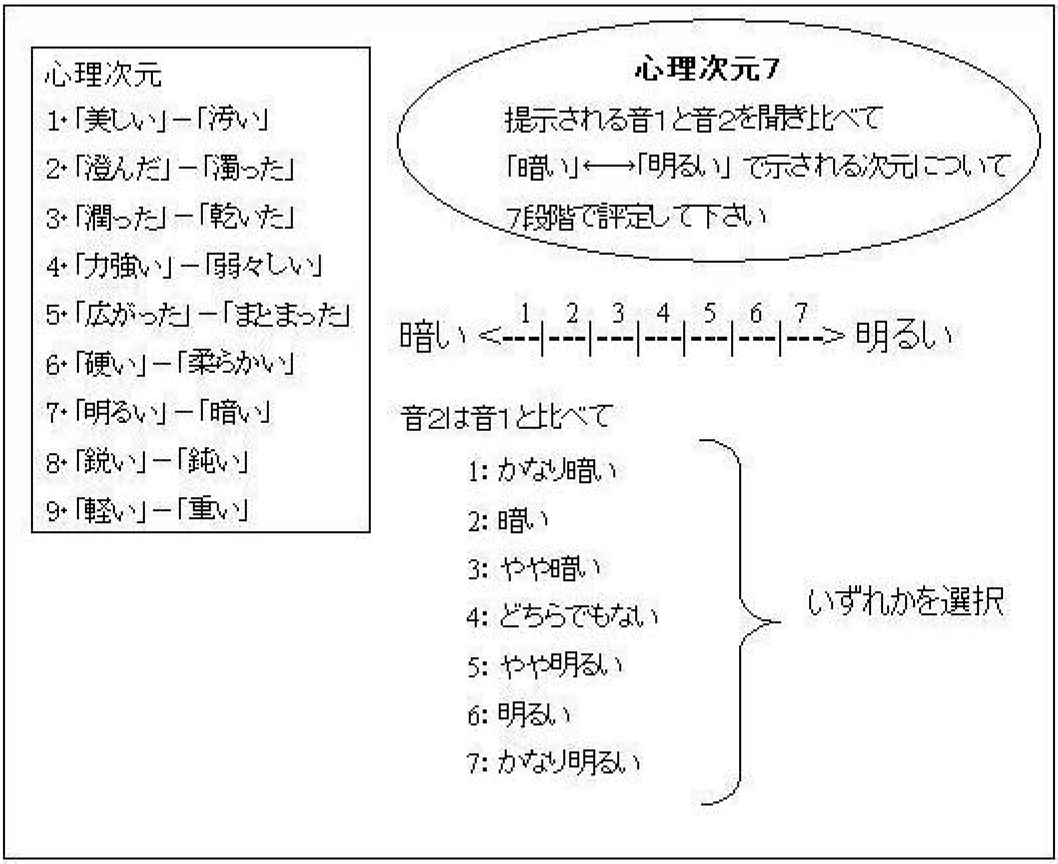


図 4.9　実験２のイメージ

4.3.4 実験結果と考察

＜結果＞

実験２は５人の被験者に対して行った。この実験において、それぞれの心理次元を７段階の序数尺度に分けた。本研究では（非常に大雑把だが）その序数尺度を心理量の程度を表す連続量として扱うことにする。評定尺度法は一対比較法とは異なり、被験者の主観が強く影響するため、ばらつきも多いと思われる。よって実験２の結果をデータの平均と標準偏差を対比して、各心理次元別に以下に示す（表 4.4a～4.4i）。

刺激音	物理操作	評価の対象となる感覚次元 -3　-2　-1　0　1　2　3 汚い＜－｜－｜－｜－｜－｜－｜－＞美しい	
		平均	標準偏差
Stimulus3.2	音量変化	0.4	0.49
Stimulus3.3	音量変化	-0.2	0.75
Stimulus3.4	音程変化	0	0.89
Stimulus3.5	音程変化	-0.4	0.49
Stimulus3.6	エンベロープ	**1.0**	1.10
Stimulus3.7	エンベロープ	0.4	0.80
Stimulus3.8	振幅変調	**1.8**	0.75
Stimulus3.9	振幅変調	-0.4	1.02
Stimulus3.10	周波数変調	-0.2	1.17
Stimulus3.11	周波数変調	-1.0	0.89
Stimulus3.12	周波数変調	**-1.4**	1.36
Stimulus3.13	倍音追加	-0.6	1.50
Stimulus3.14	倍音追加	-0.4	1.20

表 4.4a　「美しい」－「汚い」次元における評価点の平均と標準偏差

刺激音	物理操作	評価の対象となる感覚次元 -3　-2　-1　0　1　2　3 濁った＜－｜－｜－｜－｜－｜－｜－＞澄んだ	
		平均	標準偏差
Stimulus3.2	音量変化	0.8	0.75
Stimulus3.3	音量変化	-0.2	0.40
Stimulus3.4	音程変化	1.8	0.75
Stimulus3.5	音程変化	-0.6	0.49
Stimulus3.6	エンベロープ	0.2	1.33
Stimulus3.7	エンベロープ	0.4	0.80
Stimulus3.8	振幅変調	2.4	0.49
Stimulus3.9	振幅変調	-0.6	1.36
Stimulus3.10	周波数変調	-0.6	1.02
Stimulus3.11	周波数変調	-0.4	0.49
Stimulus3.12	周波数変調	-0.6	1.74
Stimulus3.13	倍音追加	-1.2	1.17
Stimulus3.14	倍音追加	-0.2	0.75

表 4.4b　「澄んだ」－「濁った」次元における評価点の平均と標準偏差

刺激音	物理操作	評価の対象となる感覚次元 -3　-2　-1　0　1　2　3 乾いた＜－｜－｜－｜－｜－｜－｜－＞潤った	
		平均	標準偏差
Stimulus3.2	音量変化	0	0.63
Stimulus3.3	音量変化	0.2	0.40
Stimulus3.4	音程変化	-1.4	0.80
Stimulus3.5	音程変化	-0.2	0.40
Stimulus3.6	エンベロープ	0.2	1.60
Stimulus3.7	エンベロープ	0	1.10
Stimulus3.8	振幅変調	1.0	1.67
Stimulus3.9	振幅変調	1.0	1.10
Stimulus3.10	周波数変調	0.4	1.02
Stimulus3.11	周波数変調	0	0.63
Stimulus3.12	周波数変調	1.0	0.89
Stimulus3.13	倍音追加	-0.6	1.20
Stimulus3.14	倍音追加	-0.6	1.02

表 4.4c　「潤った」－「乾いた」次元における評価点の平均と標準偏差

刺激音	物理操作	評価の対象となる感覚次元 -3　-2　-1　0　1　2　3 弱々しい＜─｜─｜─｜─｜─｜─｜─＞力強い	
		平均	標準偏差
Stimulus3.2	音量変化	2.4	0.49
Stimulus3.3	音量変化	-2.0	0.63
Stimulus3.4	音程変化	1.0	1.41
Stimulus3.5	音程変化	-1.2	1.17
Stimulus3.6	エンベロープ	0.6	0.49
Stimulus3.7	エンベロープ	-2.0	0.89
Stimulus3.8	振幅変調	0	0.89
Stimulus3.9	振幅変調	0.2	0.75
Stimulus3.10	周波数変調	-0.8	0.75
Stimulus3.11	周波数変調	-1.4	1.50
Stimulus3.12	周波数変調	-1.0	0.63
Stimulus3.13	倍音追加	-0.4	1.50
Stimulus3.14	倍音追加	0.4	1.02

表 4.4d 「力強い」─「弱々しい」次元における評価点の平均と標準偏差

刺激音	物理操作	評価の対象となる感覚次元 -3　-2　-1　0　1　2　3 広がった＜─｜─｜─｜─｜─｜─｜─＞まとまった	
		平均	標準偏差
Stimulus3.2	音量変化	-0.4	0.49
Stimulus3.3	音量変化	0.8	0.75
Stimulus3.4	音程変化	1.0	1.67
Stimulus3.5	音程変化	0.2	1.33
Stimulus3.6	エンベロープ	0	1.67
Stimulus3.7	エンベロープ	0	1.41
Stimulus3.8	振幅変調	-1.6	1.02
Stimulus3.9	振幅変調	-1.0	0.89
Stimulus3.10	周波数変調	-1.6	0.49
Stimulus3.11	周波数変調	0	0.63
Stimulus3.12	周波数変調	-1.6	0.49
Stimulus3.13	倍音追加	0.8	1.17
Stimulus3.14	倍音追加	0.6	1.20

表 4.4e 「まとまった」─「広がった」次元における評価点の平均と標準偏差

刺激音	物理操作	評価の対象となる感覚次元 -3　-2　-1　0　1　2　3 柔らかい<－｜－｜－｜－｜－｜－｜－>硬い	
		平均	標準偏差
Stimulus3.2	音量変化	1.0	1.41
Stimulus3.3	音量変化	-0.6	1.02
Stimulus3.4	音程変化	1.4	1.36
Stimulus3.5	音程変化	**-1.6**	1.02
Stimulus3.6	エンベロープ	**2.0**	1.10
Stimulus3.7	エンベロープ	**-1.8**	0.40
Stimulus3.8	振幅変調	-0.2	2.32
Stimulus3.9	振幅変調	0.4	1.02
Stimulus3.10	周波数変調	-1.0	1.10
Stimulus3.11	周波数変調	-0.8	0.75
Stimulus3.12	周波数変調	**-1.8**	0.98
Stimulus3.13	倍音追加	1.2	1.17
Stimulus3.14	倍音追加	1.0	1.10

表 4.4f 「硬い」－「柔らかい」次元における評価点の平均と標準偏差

刺激音	物理操作	評価の対象となる感覚次元 -3　-2　-1　0　1　2　3 暗い<－｜－｜－｜－｜－｜－｜－>明るい	
		平均	標準偏差
Stimulus3.2	音量変化	**1.6**	0.80
Stimulus3.3	音量変化	**-1.0**	0
Stimulus3.4	音程変化	**2.4**	0.49
Stimulus3.5	音程変化	**-2.0**	0.89
Stimulus3.6	エンベロープ	**1.8**	0.75
Stimulus3.7	エンベロープ	-0.6	0.80
Stimulus3.8	振幅変調	1.2	1.47
Stimulus3.9	振幅変調	**1.0**	0.89
Stimulus3.10	周波数変調	-0.2	0.75
Stimulus3.11	周波数変調	-0.2	0.75
Stimulus3.12	周波数変調	0.4	1.02
Stimulus3.13	倍音追加	0.4	1.62
Stimulus3.14	倍音追加	1.2	1.32

表 4.4g 「明るい」－「暗い」次元における評価点の平均と標準偏差

刺激音	物理操作	評価の対象となる感覚次元 -3 -2 -1 0 1 2 3 鈍い＜－｜－｜－｜－｜－｜－｜－＞鋭い	
		平均	標準偏差
Stimulus3.2	音量変化	**1.4**	1.36
Stimulus3.3	音量変化	-0.8	0.40
Stimulus3.4	音程変化	1.4	1.85
Stimulus3.5	音程変化	**-1.6**	0.80
Stimulus3.6	エンベロープ	**1.6**	1.02
Stimulus3.7	エンベロープ	-0.6	0.80
Stimulus3.8	振幅変調	-0.2	0.40
Stimulus3.9	振幅変調	-0.4	1.36
Stimulus3.10	周波数変調	**-1.0**	0.89
Stimulus3.11	周波数変調	-1.2	1.72
Stimulus3.12	周波数変調	-1.0	1.26
Stimulus3.13	倍音追加	0.2	1.72
Stimulus3.14	倍音追加	1.0	1.20

表 4.4h 「鋭い」－「鈍い」次元における評価点の平均と標準偏差

刺激音	物理操作	評価の対象となる感覚次元 -3 -2 -1 0 1 2 3 重い＜－｜－｜－｜－｜－｜－｜－＞軽い	
		平均	標準偏差
Stimulus3.2	音量変化	0	1.10
Stimulus3.3	音量変化	0.6	1.50
Stimulus3.4	音程変化	**1.8**	1.17
Stimulus3.5	音程変化	**-1.2**	0.40
Stimulus3.6	エンベロープ	0.8	1.17
Stimulus3.7	エンベロープ	**1.0**	0
Stimulus3.8	振幅変調	**1.2**	0.75
Stimulus3.9	振幅変調	0.4	1.02
Stimulus3.10	周波数変調	0	0.63
Stimulus3.11	周波数変調	0	0
Stimulus3.12	周波数変調	0.8	0.75
Stimulus3.13	倍音追加	0.4	1.36
Stimulus3.14	倍音追加	0.4	1.02

表 4.4i 「軽い」－「重い」次元における評価点の平均と標準偏差

＜考察＞

・音量と音程について

実験２の結果の中で、音量と音程による印象の測定結果を切り出したものを、以下の表 4.5 に示す。従来の音色の定義からすれば、音量と音程という特性は音色を定める物理要因とは別とされているが、これらの物理特性に関しては音の印象との対応関係がはっきりと見られた。

感覚次元	音量を操作した音		音程を操作した音	
	Stimulus3.2	Stimulus3.3	Stimulus3.4	Stimulus3.5
美しさ	0.4 (0.49)	-0.2 (0.75)	0 (0.89)	-0.4 (0.49)
透明感	0.8 (0.75)	-0.2 (0.40)	1.8 (0.75)	-0.6 (0.49)
潤い感	0 (0.63)	0.2 (0.40)	-1.4 (0.80)	-0.2 (0.40)
力強さ	2.4 (0.49)	-2.0 (0.63)	1.0 (1.41)	-1.2 (1.17)
まとまり感	-0.4 (0.49)	0.8 (0.75)	1.0 (1.67)	0.2 (1.33)
硬さ	1.0 (1.41)	-0.6 (1.02)	1.4 (1.36)	-1.6 (1.02)
明るさ	1.6 (0.80)	-1.0 (0)	2.4 (0.49)	-2.0 (0.89)
鋭さ	1.4 (1.36)	-0.8 (0.40)	1.4 (1.85)	-1.6 (0.80)
軽さ	0 (1.41)	0.6 (1.50)	1.8 (1.17)	-1.2 (0.40)

表 4.5　音量と音程操作による物理量変化と感覚量の対応
(（）内は標準偏差を表す)

表を見る限りでは、音の大きさ・高さの増加とともに各９つの感覚次元もほぼ一次元的に変化していることがわかる。このことは、音量 64～76dB(SPL)、音程 500～2000Hz 内にある純音は、音の振幅と周波数を操作することによって、音の印象が多少ながらも一次元的に操作できることを示唆している。特に、標準偏差が小さく評価点が大きい音量操作に対する「力強さ」・「明るさ」、音程操作に対する「透明感」・「明るさ」・「軽さ」については、感覚量との良い対応が期待できると考えられる。我々が生活の中でよく耳にする、音量の大きい音に対する「強い音」、音程の低い音に対する「重低音」という言葉は、この結果に基づいているものだと考えられる。

・その他、定量化できそうな心理次元

基準音との差を比べた際に、評価点が大きく、かつ標準偏差が小さかった物理操作に関しては、被験者全員に共通した印象を与える物理操作であったと期待できる。(上記の表の中では太字で示してある)ここに、これまでの実験で考察した通りに、「明るさ」に対する周波数と振幅揺らぎ、「広がり感」にたいする周期的な揺らぎ、「硬さ」に対するエンベロープ具合の対応関係が見られる。これらの結果に関しては標準偏差も小さく、人に共通した印象をもたらしているといえる。

＜反省点＞

この実験では、被験者に音から感じる印象を基準音と比較させ、７段階の評価付けをしてもらった。これより、どの物理操作が効果的に目的とする音の印象に働きかけることができるか分かったが、果たしてこの序数尺度である評価点の平均を、連続的な心理量の値として扱ってよいか問題である。やはり、心理量と物理量の対応を定量的にとるには、振幅変調の変調周波数を固定して、変調の深さのみを一次元的に変化させて、多数の被験者を対象に測定するなどの、一つ一つの物理操作に対するパラメータについて地道に測定する方法しか現時点においてはないのかもしれない。

(本研究全体にいえることだが)この実験において一番反省しなければならないことは、被験者の数の少なさである。評定尺度法という実験方法では５０人以上の被験者が必要とされていて、得られた評価点が正規分布に従うものということを前提に、序数尺度を間隔尺度に変換し、尺度構成を行ことができる。しかし、実験２では被験者５人と数的に乏しいのもであったが、データの信頼性は低いが、大よその対応関係をつかんでいると思われる。実験２から得られた結果の、評価点が比較的高く標準偏差の小さい値を用いて、次節で簡単なシンセサイザーを作製した。これは音の世界に対して、非常に制限された範囲の音しか作ることができないが、物理次元と心理次元の両方から音の操作ができるように設計されている。

５．シンセサイザーの実装

実験結果を踏まえ、研究目的に従順するシンセサイザーを作製する。

3.2.2 で示した６つの物理操作の重ね合わせにより基準音を変化させるシンセサイザーを作る。このシンセサイザーの意義は、「明るい」・「硬い」などの心理次元からの入力を可能とし、目的の印象を段階的に自動的に作ってくれるところにある。シンセサイザーは、６つの物理操作を重ねて、目的の印象を与えてくれる音を作り出す。このシンセサイザーを動かすためには、実験により求めた物理操作と心理量の対応を表す行列が、音の制御過程において必要になる。ここでは実験２の結果(表 4.4a～4.4i)から､評価点が高く標準偏差が小さい値、または、評価点の高い値だけを取り出し、その他の良く分からない値に関しては考えないことにより、音を制御する行列を作った。シンセサイザーの視覚的イメージを図 4.10 に示す。

5.1 シンセサイザーの仕様

シンセサイザーは 1000H,500msec の純音（サンプリングレート 44.1kHz）を 3.2 で紹介した４つの物理操作と振幅係数・基本周波数の操作によって変化させて音を作る。各物理操作のパラメータは図 4.12 の下側に位置するスライダーによって操作できる。「再生」ボタンにより音が提示され、作った音のスペクトログラムが上側に表示される「Reset」ボタンは入力された値を全て初期化する。（基準音に戻す）中央部に位置する９つの心理次元の「＋」と「－」ボタンにより、自動的にスライダーにパラメータが渡され、目的の印象を与えてくれる音を作り出す。その際に、最下部に位置する「固定」のチェックボックスにより、変化させたくないパラメータを固定することができる。ユーザーは、大まかな音の感じを心理次元から決定し、細かいところはスライダーを自分で操作することにより音を作る。各パラメータの範囲を表 5.1 に示す。

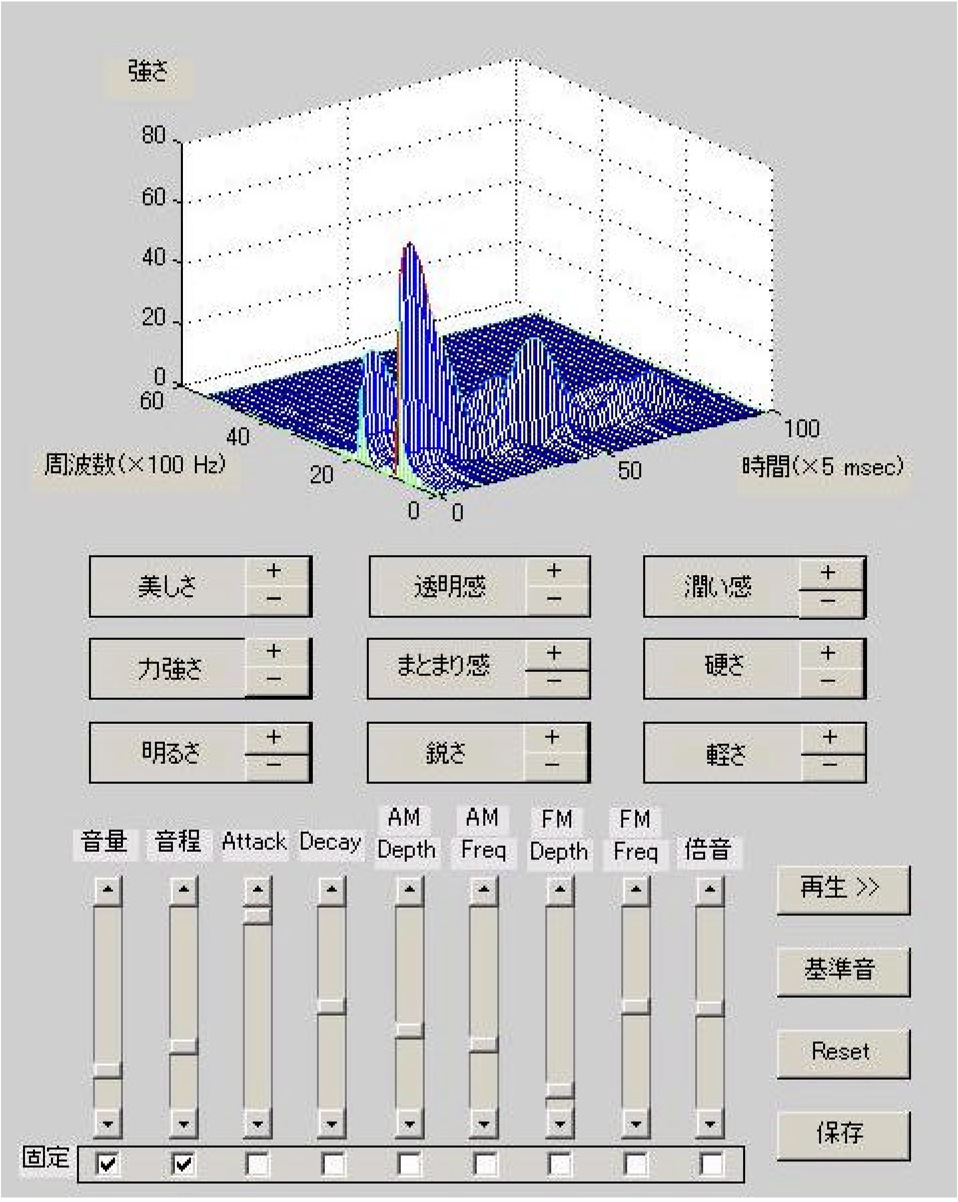


図 4.10　心理次元から音を操作するシンセサイザーの視覚的イメージ

物理量	スライダー表示	変数	上限	下限
振幅係数	“音量”	A	1	0
基本周波数	“音程”	f	2500	250
エンベロープ関数パラメータ	“Attack”	AT	500	10
	“Decay”	DC	20	0.001
		Sus	0.5（固定）	
振幅変調パラメータ	“AM Depth”	D_{am}	1	0
	“AM Freq”	f_{am}	20	0
周波数変調パラメータ	“FM Depth”	D_{fm}	20	0
	“FM Freq”	f_{fm}	20	0
倍音の振幅係数	“倍音”	A_{ha}	1	0

表 5.1　各操作による変数の範囲 (周波数の単位は Hz)

5.2 音の制御方式

このシンセサイザーは、心理次元からの入力を以下の物理操作によって制御する。各操作に対して上記範囲内にある 6 段階のパラメータを用意し、対応関係のある心理量の「＋」と「－」の入力により、それぞれの値を変化させる。(振幅係数と基本周波数については１１段階）ただし、この値は目的の心理量を増加及び減少させるだけであって、心理次元における等間隔性の保証は全くない。実際に連続的な物理量と心理量との対応関係を調べた訳ではないので、あくまで、段階的に与えられるパラメータの範囲内においては、その物理操作が目的の心理量との一次元的な対応をするであろうと仮定して音を制御しているにすぎない。

5.2.1 段階的なパラメータの寄与

＜振幅操作＞

$$A = A_0 \cdot 10^{(1.2 \cdot n_a/20)} \qquad (n_a = -5,-4,-3,-2,-1, 0, 1, 2, 3, 4, 5)$$

A　：振幅係数
A_0　：振幅係数の初期値

n_aが±１変化するごとに音圧は 1.2dB(SPL)変化する。n_a=±5 の時、音圧は

元の音よりも±6ｄB(SPL)変化する。心理実験２における Stimulus3.2 と Stimulus3.3 に対して行った振幅操作を、段階的に分けたものがこれにあたる。また、A=A_0 の時に音圧が 70dB(SPL)に近い値となるように設定しない場合は、このパラメータは固定することが望ましい。初期値は n_a=0 の状態である。

＜基本周波数操作＞

$$f = f_0 \cdot (\sqrt[5]{2})^{n_f} \qquad (n_f = -5,-4,-3,-2,-1, 0, 1, 2, 3, 4, 5)$$

f　：基本周波数(Hz)
f_0　：基本周波数の初期値(= 1000Hz)

n_fが±１変化するごとに基本周波数は、$(\sqrt[5]{2})^{n_f}$倍になる。（n_f=5 の時に 2 倍に、n_f=-5 の時に 1/2 倍に変化）心理実験２において Stimulus3.4 と Stimulus3.5 に対して行った基本周波数操作を、段階的に分けたものがこれにあたる。基本周波数の初期値は 3.1 の基準音と同じく 1000Hz（n_f=0 の状態）である。

＜エンベロープ操作＞

3.2.2 に示したエンベロープ関数の各パラメータを、以下のように段階的に与える。

$$AT = 110 - 20n_{at} \qquad (n_{at} = 0, 1, 2, 3, 4, 5)$$

$$DC = 2n_{dc} \qquad (n_{dc} = 0, 1, 2, 3, 4, 5)$$

（ただし、n_{dc}=0 の時は DC=0.001 とする）

Sus=0.5

初期値を n_{at}=0, n_{dc}=0 の状態として、AT は 110～10 に、DC は 20～0.001 の範囲内でそれぞれ段階的に変化する。Sus の値は常に一定である。AT は n_{at} により単調減少、DC は n_{dc} により単調増加する。Stimulus3.6 は DC=20、Stimulus3.7 は AT=10 のパラメータを持つことから、作製する音は、n_{at} の増加により最大振幅までの立ち上がり方は Stimulus3.7 に、n_{dc} の増加により最大振幅から Sus までの減衰の様子は Stimulus3.6 にそれぞれ近づいていく。

＜振幅変調＞

3.2.2 に示した振幅変調の各パラメータを以下のように段階的に与える。

$D_{am} = 0.1 n_{am} \qquad (n_{am} = 0, 1, 2, 3, 4, 5)$

$f_{am} = 6$ (Hz)

f_{am} を固定して D_{am} のみを 0～0.5 の範囲で段階的に変化させる。$n_{am} = 5$ の時に D_{am}=0.5 となり Stimulus3.8 と同じ振幅変調となる。初期値を n_{am}=0 の状態とする。

＜周波数変調＞

3.2.2 に示した周波数変調の各パラメータを以下のように段階的に与える。

$D_{fm} = 2 n_{fm} \qquad (n_{fm} = 0, 1, 2, 3, 4, 5)$

$f_{fm} = 10$ (Hz)

f_{fm} を固定して D_{fm} のみを 0～10 の範囲で段階的に変化させる。$n_{fm} = 5$ の時に D_{fm}=0.5 となり Stimulus3.12 と同じ周波数変調となる。初期値を n_{fm}=0 の状態とする。

＜倍音成分の追加＞

3.2.2 に示した倍音成分の振幅係数を以下のように段階的に与える。

$A_{ha} = 0.2 n_{ha} \qquad (n_{ha} = 0, 1, 2, 3, 4, 5)$

A_{ha} を 0～1.0 の範囲で段階的に変化させる。$n_{ha} = 5$ の時に A_{ha}=1.0 となり Stimulus3.14 と同じ基本周波数に対する倍音成分の比となる。初期値を n_{ha}=0 の状態とする。

このように、ここで使われる各物理操作は、心理実験２の刺激音作製に使われた各物理操作のパラメータを変化の上限として、段階的にパラメータが与えられる。

5.2.2 心理量と物理量の対応行列

実験２の結果から、5.2.1 に示す各操作の変化上限に等しいパラメータにより作製した刺激音の、心理量と物理量の対応結果を以下にまとめる（表 5.2）。

心理次元	音量		音程		エンベロープ		振幅変調	周波数変調	倍音成分追加
	St3.2	St3.3	St3.4	St3.5	St3.6	St3.7	St3.8	St3.12	St3.14
美しさ	0.4 (0.49)	-0.2 (0.75)	0 (0.89)	-0.4 (0.49)	1.0 (1.10)	0.4 (0.80)	1.8 (0.75)	-1.4 (1.36)	-0.4 (1.20)
透明感	0.8 (0.75)	-0.2 (0.40)	1.8 (0.75)	-0.6 (0.49)	0.2 (1.33)	0.4 (0.80)	2.4 (0.49)	-0.6 (1.74)	-0.2 (0.75)
潤い感	0 (0.63)	0.2 (0.40)	-1.4 (0.80)	-0.2 (0.40)	0.2 (1.60)	0 (1.10)	1.0 (1.67)	1.0 (0.89)	-0.6 (1.02)
力強さ	2.4 (0.49)	-2.0 (0.63)	1.0 (1.41)	-1.2 (1.17)	0.6 (0.49)	-2.0 (0.89)	0 (0.89)	-1.0 (0.63)	0.4 (1.02)
まとまり感	-0.4 (0.49)	0.8 (0.75)	1.0 (1.67)	0.2 (1.33)	0 (1.67)	0 (1.41)	-1.6 (1.02)	-1.6 (0.49)	0.6 (1.20)
硬さ	1.0 (1.41)	-0.6 (1.02)	1.4 (1.36)	-1.6 (1.02)	2.0 (1.10)	-1.8 (0.40)	-0.2 (2.32)	-1.8 (0.98)	1.0 (1.10)
明るさ	1.6 (1.36)	-1.0 (0)	2.4 (0.49)	-2.0 (0.89)	1.8 (0.75)	-0.6 (0.80)	1.2 (1.47)	0.4 (1.02)	1.2 (1.32)
鋭さ	1.4 (1.10)	-0.8 (0.40)	1.4 (1.85)	-1.6 (0.80)	1.6 (1.02)	-0.6 (0.80)	-0.2 (0.40)	-1.0 (1.26)	1.0 (1.20)
軽さ	0 (1.36)	0.6 (1.50)	1.8 (1.17)	-1.2 (0.40)	0.8 (1.17)	1.0 (0)	1.2 (0.75)	0.8 (0.75)	0.4 (1.02)

表 5.2　実験２の結果・心理次元に対する評価点の平均
（表中にある St は Stimulus の略, （）内は標準偏差を表す）

表 5.2 の評価点の平均と標準偏差から、心理次元に影響を及ぼす物理量と思われるものに ±１と数字を割り当て、特に影響が見られないものは 0 と割り当てる。これにより心理量と物理量の対応関係を表す行列を作る。この行列は、0 と±１の値だけで構成される。表 5.2 に示されるの心理次元に対し、正の対応には +１、負の対応には-１をつけ、対応関係がないと見なされるものには 0 をつける。対応関係の有無の決定には、次の２つの方法をとった。

（1）｜(評価点平均)｜ － (標準偏差) ＞ 0　を満たすもの

ばらつきを考慮して、被験者によっては様々な応えがあるものは評価点が高くても対応関係がないと見なす。一方で、評価点が低くても被験者全員が同じように応えたものに関しては対応があると見なす。例として、表 5.2 における「振幅変調」に対する「潤い感」と「明るさ」は、評価点の平均が高いが、標準偏差が大きく、ばらつき具合が大きい。被験者によっては、全く「潤い感」や「明るさ」を感じなかった人がいることを示唆する。よって、振幅変調は「潤い感」と「明るさ」に対応しないと判断する。この方法により作った心理量と物理量

の対応行列を表 5,3a に示す。

（2）｜(評価点平均)｜ ＞1　を満たすもの

（1）に付け加え、ばらつきを考慮せず、評価点の平均が比較的高いものに対しても全て対応があると見なす。データの信用性は低いが、シンセサイザーの音の制御過程において、心理次元に対応する物理操作が増えるので、音の変化の度合いが分かり易い利点がある。この方法により作った物理量と心理量の対応行列を表 5.3b に示す。

心理次元	振幅	基本周波数	エンベロープ		振幅変調	周波数変調	倍音成分追加
	A	f	AT	DC	D_{am}	D_{fm}	A_{ha}
美しさ	0	0	0	0	1	-1	0
透明感	0	1	0	0	1	0	0
潤い感	0	-1	0	0	0	1	0
力強さ	1	1	-1	1	0	-1	1
まとまり感	0	0	0	0	-1	-1	0
硬さ	0	1	-1	1	0	-1	0
明るさ	1	1	0	1	0	0	0
鋭さ	1	1	0	1	0	0	0
軽さ	0	1	1	0	1	1	0

表 5.3a　心理量と物理量の対応行列（1）

心理次元	振幅	基本周波数	エンベロープ		振幅変調	周波数変調	倍音成分追加
	A	f	AT	DC	D_{am}	D_{fm}	A_{ha}
美しさ	0	0	0	1	1	-1	0
透明感	0	1	0	0	1	0	0
潤い感	0	-1	0	0	1	1	0
力強さ	1	1	-1	1	0	-1	1
まとまり感	0	0	0	0	-1	-1	0
硬さ	0	1	-1	1	0	-1	1
明るさ	1	1	0	1	1	0	1
鋭さ	1	1	0	1	0	-1	1
軽さ	0	1	1	0	1	1	0

表 5.3a　心理量と物理量の対応行列（2）

「力強さ」に対する音量と倍音追加の操作に関しては、シンセサイザーの制御過程において、3.4 に挙げる聴こえの大きさの制御が十分にできなかったため、聴こえの大きさが大きく聴こえてしまう。そのため無条件で正の対応をつけた。また、表 5.2 の音量変化と音程変化の結果に関しては、増加・減少の２つの結果が、共に（１）または（２）の条件を満たす場合のみ、心理次元との対応があると見なし、表 5.3a,5,3b において「音量」と「音程」という一つの物理操作にまとめた。

シンセサイザーは、これら２パターンのどちらか一方を用いて、目的の印象を 5.2.1 に示した段階的なパラメータ寄与による物理操作でもって作り出す。

5.2.3　データの処理

このシンセサイザーの使い方と共に、5.21 の段階的な物理操作と 5.22 の対応行列を用いた音の制御を簡単に説明する。(２)の行列を用いて、音を制御した場合について考える。まず、初期状態で「美しさ」の「＋」ボタンを押した場合には、「美しさ」の心理次元と、正の対応関係のある DC と D_{am} のパラメータである n_{dc} と n_{am} に＋１が加算され、エンベロープ操作と振幅変調を決定するパラメータに DC=2 , D_{am}=0.1 という値が与えられる。これにより音が少し「美しい」とされる音に変化する。この場合は、振幅変調と立ち下がり操作の各操作とも、n は５段階に定められているので、この操作が５回以上行なわれると、n=5 の状態が保たれて音が変化しなくなる。この状態で「－」のボタンを押した場合に、n_{dc} と n_{am} に－1 が加算され、「－」ボタンを押すごとに 5 段階で初期状態の音に戻る。元の初期状態の音に戻ってからも、「美しさ」の「－」ボタンを押し続けると、今度は「美しさ」と負の対応関係にある D_{fm} のパラメータである n_{fm} に＋１が加算され、これにより周波数変調の操作が施され、音が少し「汚い」とされる音に変化する。「美しさ」以外の心理次元からの入力も同様に、対応する物理操作段階的な操作により目的の感覚の音を作る。音の世界に対して、非常に狭い範囲でしか機能しないが、このような仕様でシンセサイザーは心理次元からの音の操作を可能にしている。

＜考察＞

実際に、このシンセサイザーを筆者が使ってみたところ（１）の行列を用いた音の制御は、データの信頼性が高い一方で、音の物理操作が少ないという欠点があるため、本研究のシンセサイザーとしては多少物足りないと思った。(先入観からそう思ったのかもしれないが）多少大雑把にデータを扱った（２）の行列を用いた音の制御の方が、心理次元の入力による音の変化の幅が大きいので、目的のシンセサイザーとして適しているという印象を受けた。

6. まとめ

6.1 全体的な考察

実験結果から、本研究から得られたことを報告する。本研究では基準音を一つ定め、その音に特定の物理操作を加えることにより音を変化させて、元の基準音と聴こえた感じの印象を比較することにより、加えた物理操作により生じた印象がどのようなものか測定した。音の操作方法は簡単な数式表せるくらい簡単なものであったが、いくつかの心理次元に対しては非常に良い対応が見られた。本研究に用いた６つの物理操作と、９つの心理次元との対応を以下に考察する。

＜振幅変化による音の印象の変化＞

心理実験２の結果から、64～70ｄB の範囲でしか適応できないかもしれないが、波形の振幅という音の大きさとほぼ一次元的な対応関係がある物理量は、音の大きさだけでなく、音の「力強さ」・「明るさ」・「鋭さ」に影響する物理量であることが分かった。音の「力強さ」と「大きさ」は言葉的に非常に相関性の高いものであるかもしれないが、心理実験２の結果においては、この音量と「力強さ」に対する対応が、ばらつきが少なく、かつ評価点が高いという一番良い対応関係にあった。しかし、音の大きさの変化による音の変化は、他の５つの物理操作とくらべて一番乏しいものであり、音を作るにあたりこのパラメータはあまり重要でないことは避けられない事実である。

＜基本周波数変化による音の印象の変化＞

心理実験２の結果より、この物理量に対しても 500～2000Hz の範囲内では、波形の振幅変化と同様に心理次元との一次元的な対応関係が見られた。多少のばらつきが見られたが、基本周波数の変化は本研究で対象とした音の心理次元の「美しさ」以外の全ての次元に影響した。この結果から、一般に人は高い音に対しては「澄んだ」・「乾いた」・「力強い」・「まとまった」・「硬い」・「明るい」・「鋭い」・「軽い」という印象を抱き、低い音に関しては全く逆の印象を抱く傾向にあることが分かった。特に音の「明るさ」については、音の振幅変化に対

する「力強さ」のように、ばらつきが少なく、かつ評価点の高い結果であった。これは、人に共通して同じ印象を抱かせている音に他ならない。先行研究によれば音の「明るさ」は高次倍音成分の強さによると言われているが、周波数の高い音の成分が「明るさ」に影響する物理量であることに間違いないと思われる。しかしながら、少なくとも音楽に使われる音を作るためには、基本周波数は固定して話を考えなければ意味がない。この研究の実用性を考えるならば、他の物理変化に注目して話を進めなければならない。

＜エンベロープ操作による音の印象の変化＞

本研究で対象としたエンベロープは、音の世界に対して非常に単純なものでしかなかったが、少なくとも、音の立ち上がり・立ち下がりという２つの大きな物理要因の音の印象に対する影響を調べることが出来た。その結果、心理実験１と２の結果から立ち上がり・立ち下がりの方の強い（*AT* と *DC* が大きい）音からは「美しい」・「明るい」・「鋭い」・「硬い」という印象が感じられ、逆に立ち上がり方が弱い（*AT* が小さい）音からは、「弱々しい」・「軽い」・「柔らかい」印象を受けることが分かった。ここで、立ち上がり係数 *AT* に関しては、「硬い」－「柔らかい」という次元における一次元的な対応関係が見られる。これに関しては、我々が普段生活する中で硬い物体を叩いた時に生じる音が、立ち上がりの強い音に、逆に柔らかい物体を叩いた時には立ち上がりの弱い音が生じるという生活的な背景が要因なのかもしれない。

＜振幅変調による音の印象の変化＞

本研究では、パラメータの与え方が非常に狭い範囲内であったことが理由か、振幅変調による音は「美しい」・「澄んだ」・「潤いある」・「弱々しい」・「広がりある」・「明るい」・「軽い」という、人間的にはプラスイメージとして感じる印象に対する評価が大きかった。前述のように、人間が多少揺らぎのある音を好む傾向にあり、この音が人間に好ましい印象を抱かせたのかもしれない。筆者もこの操作による音が非常に「美しい音」であると感じる。しかし、この操作は２つのパラメータの与え方によって、「美しい音」にも「汚い音」にも聴こえる。本研究では、シンセサイザーの音の制御過程おいて、変調周波数を固定して特に変調の深さを狭い範囲内で段階的に変化させることにより、音の印象を変化させた。これに関しては特に実験から、一次元的に印象が変化するという傾向を割り出した訳ではなく、あくまで筆者の主観的な感覚のみで決定したもので対処してしまった。これに関しては、今後の課題として一度実験してみなければならない。

＜周波数変調による音の印象の変化＞

この物理操作をかけた音に対しては、振幅変調とは逆に「汚い」・「濁った」・「潤った」・「広がりある」・「弱々しい」・「鈍い」・「柔らかい」と、やや人間にはマイナスイメージとなるような感覚の評価が高かった。周波数変調は周波数を時間的に変化させる操作である。音の周波数という物理量は、基本周波数の変化による結果から分かるように、音の印象に対して敏感に効いてくるものである。より、音の印象の問題に関して周波数変調による印象変化を測る場合には、変化のさせ方を小さくすべきであったと反省する。

・倍音成分を追加した音の印象の変化

先行研究によれば、高次倍音成分の強さは音の「明るさ」と「鋭さ」に非常に影響する物理量であるとされている。しかし、心理実験２の結果を見ると確かに「明るさ」と「鋭さ」に対する評価が高いものであったが、標準偏差が大きく判断が被験者によってばらついた結果となった。これの理由として、第２次倍音成分しか追加しなかった、被験者の数が少なかったという要因が考えられる。また、音の「まとまり感」に関して、実験２では倍音成分の追加によって対応関係があると思われたが、心理実験２ではさほどの評価を受けなかった。標準偏差が高いことからも、人によっては「まとまっている」と感じるかもしれないが、一般的に人に共通して感じられる対応関係でなければ、議論の意味がない。
「まとまり感」－「広がり感」で示される次元は、周期的な揺らぎのある音に関しては「広がり感」との対応関係がはっきりしていたが、倍音成分の追加によって「まとまり感」が増えるということは必ずしも言えないようである。これについては第 3 次倍音、第 4 次倍音を加えた音に対しても測定してみる必要がある。また、実験では倍音成分を追加した音が少なからず「汚い」と判定される結果になった。余談だが、筆者は全ての実験結果をから、人間は高次倍音成分を含まない純音のような「澄んだ音」に減衰と時間揺らぎがかかったものを好む傾向にあるのではないかと推測する。ちょうど風鈴の音や鈴虫の鳴き声に非常に似ている音である。

6.2 結論

本研究の目的は音の印象を操作することである。この問題に対して、ある特定の心理量に影響を及ぼす物理操作を見つけることができれば、その物理操作を用いて音の印象の操作が可能になるのではないかと考えた。本研究では、振幅操作、基本周波数操作、エンベロープ操作、振幅変調、周波数変調、倍音成分追加という音の物理次元からの操作によって、基準となる音を変化させて、それに応じて変化した心理量に着目した。実験２では、基準音と変化音の９つの心理次元における印象の差を、評価点をつけることにより測定した。その結果、被験者全員が共通して大きな評価点をつけた心理量と物理量の対応に着目する。そして、実験に用いた刺激音に与えたパラメータの値を上限として、この物理操作のパラメータ与え方を５段階に分ける。このパラメータを段階的に与えることにより、その操作に反応する心理量は、等間隔的に変化する訳ではないが一次元的に増加・減少する。本研究ではこのような方式によって、心理次元から音を操作することを試みた。

この方式は、音の変化が極めて狭い範囲の条件内では成立するだろう。しかし、これについては、後の実験により検証せねばならない。本研究ではこの検証よりも先に、研究目的に従順するシンセサイザーの作製を優先した。本研究の意図を明確に人に理解してもらいたかったことを理由に、内容よりも結果を先に出すことにした。これに限らず、本研究では全体的な研究目的の達成の流れを急いでしまったため、少なからず細部に渡って注意が行き届いていない。反省すべき点である。

このシンセサイザーは現時点では単純な音しか作製することができない仕様になっているが、基準音となる音を複数に増やす、高次倍音成分を追加するなどの物理操作を増やす、時間の概念を取り入れる、現在のシンセサイザーの方式として用いられている既存の音の重ね合わせによる音の作製を可能にし、これを心理次元から制御するなどの新しい機能を複数追加することが出来れば、実用化できるかもしれない。

謝辞

本研究を遂行するにあたり非常に多くの人の力をお借りしました。被験者として協力してくれた方々、ゼミ発表にて実験における詰めの甘さを次々と指摘して下さった先生方、退任されてからもいろいろな分野で面倒を見続けて下さった、同研究室元助手の関根先生、そして様々な相談をしながら、共に学業に励んだヒューマンインターフェース講座の皆様方に心から感謝します。

最後に、音の研究がしたいと半ば強引に研究室に配属希望した、何の知識なかった自分を受け入れてくれ、最後まで面倒を見続けてくれた阪口先生に、心より御礼申し上げます。

参考文献

[1] 難波精一郎 編：聴覚ハンドブック, ナカニシヤ出版 (1984)

[2] 難波精一郎 :“ 音色の定義を巡って ”, 日本音響学会誌, 49, 823-832 (1993)

[3] 曽根敏夫, 城戸健一, 二村忠元 :“ 音の評価に使われる言葉の分析 ”, 日本音響学会誌, 18, 320-326 (1962)

[4] 安倍幸治, 小澤賢司, 鈴木陽一, 曽根敏夫 :“ 音色表現語、感情表現語及び音情報関連語による環境音評価 ”, 日本音響学会誌, 54, 343-350 (1998)

[5] 植松尚, 小澤賢司, 鈴木陽一, 曽根敏夫 :“ 高次高調波成分のみからなる複合音の音色知覚過程とそのモデルに関する検討 ”, 日本音響学会誌, 52, 948-956 (1995)

[6] 金子英二, 小松包治 :“ 楽音の評価についての一考察 ”, 日本音響学会誌, 51, 613-615 (1995)

[7] G. von Bismarck :“ Sharpness as an attribute of the timbre of steady sound ” , Acustica, 30 , 146-159 (1974)

[8] 境久雄：聴覚と音響心理, (境久雄 編) , コロナ社 (1995)

[9] 難波精一朗・桑野園子：音の評価のための心理学的測定法, 3 章, 45-144, コロナ社

[10] 安部素嗣 :“ 楽器音の学習認識システム ”, 東京大学工学部 計数工学科 卒業論文, (1991)

[11] 鈴木基 :“ 音の立ち上がりのずれが音源分離に与える影響 ”, 電気通信大学大学院 情報システム学研究科 修士論文, (1999)