

音楽的ビートが映像的ビートの知覚に及ぼす引き込み効果 Drawing-in Effect on Perception of Beats in Multimedia

長嶋洋一

Yoichi NAGASHIMA

ASL/SUAC

nagasm@computer.org

概要 聴覚的情報と視覚的情報を同時に視聴するマルチメディアコンテンツにおける「ビート」に注目し、音楽的ビートが映像的ビートの知覚に及ぼす局所的な引き込み効果を提唱した。これを解析・検証するための新しい心理学実験システムを制作し、被験者テストによる実験・評価を行った。あわせて、コンピュータを用いた心理学実験で重要となる、システムのレイテンシ(遅延)について検討・考察した。具体的には、従来から進めてきたComputer Music関係の研究を受け継ぎ、新しいモデルを提唱し実験的に検証するために、(1)新しい視点での「ビート」概念の定義、(2)実験に適した映像素材の検討と制作、(3)実験に適した音楽素材と被験者インターフェースの開発と映像系との同期制御、(4)被験者の音楽経験など個人的要因によるばらつきへのシステム対応(チューニング)、(5)システムのレイテンシの検討と心理学的影響の考察、(6)実験データの効果的な可視化、などの新しいアプローチを行った。

Abstract : This research is about "the interaction of beats" in multimedia. Like a music video, the case where the beat of music and images synchronizes completely, it is generally natural and comfortable for people. In this research the tempo between images and music is slightly different. I produced the experiment system to generate visual beats and musical beats with slightly different tempo in real-time, and tested with many human subjects.

キーワード : リズム、ビート、心理学実験、レイテンシ、引き込み効果

Keywords : Rhythm, Beat, Psychological Experiment, Latency, Drawing-in Effect

1. はじめに

人間はリズムとともに生存/生活している。心理学の領域では、音楽/音声などの聴覚情報、映画/ビデオなどの視覚情報、ダンス/身振りなどの身体運動情報に対して、時間的に繰り返すリズムを知覚・認知する人間の心理的メカニズムが探究されてきた。Computer Musicと総称される音楽情報科学の領域でも、リズム・ビート・拍子などの時間的要素を対象としたテーマにおいて、聴覚情報はもちろんのこと、視覚情報や身体運動情報と一体となったマルチメディア情報・マルチモーダル情報が研究の対象となってきた([1]-[65])。

本研究では、聴覚的情報と視覚的情報を同時に視聴するマルチメディアコンテンツの感覚間調和 intersensory harmony に関して、新たな視点でビートを「周期的に繰り返しリズムのノリが知覚されるアクセント部分」と再定義した。その上で、音楽的ビートが映像的ビートの知覚に及ぼす局所的な「引き込み効果」を提唱し、これを解析・検証するための新しい実験システムをデザインし、被験者テストによる実験・評価を行った。また音楽情報科学の視点から、コンピュータを用いた音楽心理学/知覚認知心理学の実験で重要となる、実験システムのレイテンシ(遅延)についても詳細に検討・考察した。

本研究での「引き込み効果」とは、非線形システムの同期現象である引き込み現象 entrainment とは異なる概念である。一言で言えば、等間隔(一定テンポ)の映像的ビートに対してスイッチによるタッ

ピングを行っている人間が、その映像的ビートと同じテンポの音楽的ビートを聴取している時に、両者のごくわずかなテンポの違いからビート位置がずれてきた時に、等間隔である筈の映像的ビートが音楽的ビートに引っ張られるように移動して、そのずれが大きくなると(無意識に気付いて)元に戻る、というような振るまいを起こす、という効果である。

リズムは音楽だけでなくすべての時間的事象の形態化において広く見られる心理現象である [1]。リズムは継起する事象の秩序立った特性であり、予測可能性により成立している人間の心理的な(心的構築によって生まれる)概念、「秩序の知覚」である [2][4]。これまで音楽心理学においては、リズムの知覚の実験的研究として、拍節のリズム、リズム形式、強弱の知覚、高低の知覚、長短の知覚、時間的構造の知覚、反応時間、リズム記憶と再生、主観的リズム、テンポ変化の知覚、タッピング(リズムと合わせて叩く身体動作)実験、リズムとテンポの同期・同調、リズムの群化(時間的パターン認識、時間的ゲシュタルトの知覚)、近接リズムの影響による錯覚現象、など多様なテーマの研究が行われてきた [1][2][8][11][13][34][39][43][46]。

身体運動と関係づけたリズムの特性としては、(1)リズムは人間に同期反応 synchronized response を誘発しやすい、(2)リズムは力の消費を最小化調整する要素として身体活動を支配する、(3)音響リズムは運動リズムをひきおこし、運動リズムは音響リズムをひきおこす(聴覚の筋運動感覚的側面)、な

どが報告されている [2][4][16][18][37][38][51]。視覚的メディアである映画/映像においては、ショットとショットとを編集montageする際に、刺激の強弱・断続・連続などの視覚的リズムを考慮して、画像対象とは別にフィルム/フレーム自体のリズムが研究されてきた [5][6][10]。

これら身体運動のリズムや視覚的リズムの領域においても、人間が身体運動や映像などの刺激を能動的に知覚することでリズムを心的に構築する、という音楽心理学と共通の報告がある[18][19][26][41]。また注目すべき研究として、音楽ビートの感じ方には Tago Beat、Toe Beat、Heel Beat、というまったく異なる3種類があると報告された[61][62]。本研究ではこれらの視点に注目して「ビート」の概念を再定義した。

従来心理学実験/研究による音楽的情報と視覚的情報の相互関係については、(1)音楽と映像のリズム/テンポが同期しているとダイナミックな印象や調和感を与える、(2)音楽と映像のリズム/テンポがばらばらであると違和感を感じる、(3)リズムが倍数関係であれば人間は容易に拍を分割して自然に同期する、(4)映像のリズムに自然なように音楽の速度を変える実験では整数倍のリズムに同調する、などの直感的に賛同できる結果が得られている [13][19][35][36]。

これらの心理学実験に用いられた視覚的情報(刺激)としては、画像/映像の持つ意味的な影響を排除する目的で、LEDランプなど発光体の点滅、モニター内の正方形や円環など単純図形上の輝点の運動など、結果として被験者の反射神経や運動神経に大きく影響されるものが多かった [12][29][30][53]。しかしこれらの単純・単調な視覚刺激リズムの繰り返しは、慣れにより飽きや眠気など被験者の注意力を奪うため、短時間の実験でのみ有効で、知覚認知の漸進的プロセス(2-3分間以上のオーダ)の実験には適さない [17][32][33]。さらに、テレビや映画などマルチメディアコンテンツの主目的である「エンターテインメント」「アート」等の性格が乏しい(面白くない・楽しめない・現実的でない)、という問題点も指摘された。これらの指摘は音楽的情報(刺激)についても同様に有効である。

また、マルチメディア心理学実験において、モニター画面内の映像/画像の瞬時変化や光体の点滅という視覚刺激を用いる場合、10-100msecのオーダである視覚的反応時間との関係、あるいは無意識に画面内を瞬時走査する眼球運動(サッカドsaccade)の影響を十分に検討する必要がある [3][7][9][13][14][19][22][30][54]。さらにコンピュータを心理学実験に使用する場合には、その実験システム(ハードウェア、ソフトウェア)のレイテンシ(遅延)と誤差について注意しなければ、心理学的時間オーダを越えるレイテンシや誤差により実験の意味を失う危険性もある [20][23][25][27][56][60]。

時間的・空間的に複数のリズムが存在する場合の相互作用については、テンポの同調モデルや追従trackingを実現する課題が多くのテーマで研究されており、心理学実験においては何を基準として何を変化させて何を調べるのか、という実験計画も重要

である [21][24][42][44][45][55]。また、リズムに関する聴覚的な現象を効果的に可視化する手法、心理学現象をニューラルネットワーク等でモデル化する手法についても多くの研究が行われてきた[40][47][48][49]。

これらの関連研究・先行研究について筆者の調べた限りにおいては、本研究で筆者が提唱するような「引き込み効果」についてのアプローチは見られなかった。その理由として、(1)従来研究における実験システムの技術的な制約から、本研究の対象となる「引き込み効果」のオーダの時間的な現象を正確に規定して心理学実験を行うことが困難であった、(2)視覚的情報と聴覚的情報の時間的同期についての研究そのものがまだ萌芽の段階にある、(3)時間的同期に関する研究の多くが、単一の刺激に対する時間的遅延などに注目されており、ほぼ等間隔で眷属するビートのような対象はほとんどアプローチされていない、等がある。

このような状況のもとで、筆者は従来から進めてきたComputer Music関係の研究を受け継ぎ、新しいモデルを提唱し実験的に検証するために、(1)新しい視点での「ビート」概念の定義、(2)実験に適した映像素材の検討と制作、(3)実験に適した音楽素材と被験者インターフェースの開発と映像系との同期制御、(4)被験者の音楽経験など個人的要因によるばらつきへのシステム対応(チューニング)、(5)システムのレイテンシの検討と心理学的影響の考察、(6)実験データの効果的な可視化、などの新しいアプローチを行った [28][50][57][58][59][60]。

2. ビートの引き込み効果モデル

ここでは本研究において新しく提唱した、音楽的ビートが映像的ビートの知覚に対して局所的に引き込み効果をもたらす、というモデルを解説する。

まず初めに、村尾の提唱する新しい音楽ビートの感じ方として Tago Beat、Toe Beat、Heel Beat について紹介する [61][62]。この視点と関連して、本研究における用語として、テンポ・拍子・アクセント・ノリ・ビート・引き込み現象/引き込み効果について、概念と定義を明確化する。そして、映像のテンポと音楽のテンポとがほとんど同じであるがわずかに異なっている、という状況において起きると予想した局所的引き込み効果について、そのメカニズム仮説を提示し、この仮説に基づいた場合にデータとして得られると予想される状況について述べる。

2-1. 村尾のビート分類

音楽における拍節構造、アクセント、グルーピングなどを研究してきた村尾は1988年に、音楽ビートの感じ方を区別する新しい用語として Tago Beat、Toe Beat、Heel Beat を提唱した [61][62]。本研究においてはビート、ノリ、アクセントなどの概念を対象にするために、ここで簡単に Tago Beat、Toe Beat、Heel Beat について紹介し、次項での用語や概念の説明において「村尾の音楽ビート分類」ということで、適宜参照して関連付けることにする。

Tago Beat

着地(打ち)止め、裏跳ねによるビート。着地(落下)点を拍子の表、跳ね上げ点を裏側とする。日本の農民の手拍子、演歌の合の手などで表拍から裏拍まで揉み手をするノリは日本人の高齢者には親しみがある(図1)。

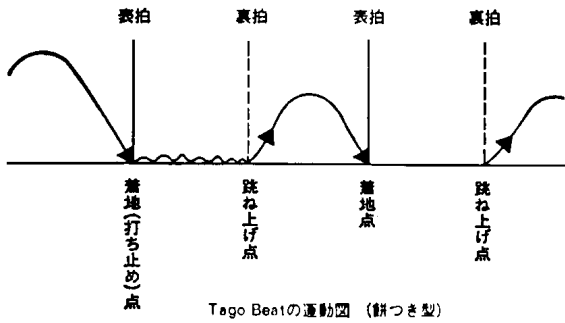


図1 Tago Beat

Toe Beat

弾力性があり、踵(かかと)から爪先(トゥ)へ跳ね上がるビート。基本型は落下点と跳ね上げ点が一致し、拍点に向かって加速し、その後は減速する。バレエのステップのようにビートがグループ化されたものもある(図2)。

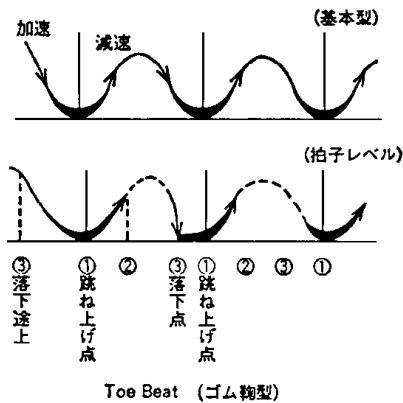


図2 Toe Beat

Heel Beat

爪先を支点として踵(ヒール)を打ち付けることによって拍点を得る。ジャズ・ロック・ポップスなどで一般的なビート。タゴビートのように打ち止めせず、トゥビートのように弾力で跳ね返りもせず、ハンマーのように叩き返す。グループ化された場合には弱拍に強勢のアクセントが置かれ、これは手拍子の位置となる(図3)。

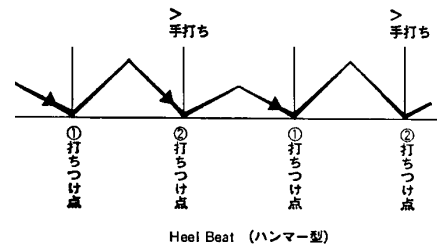


図3 Heel Beat

2-2. 関連用語と概念

これまで多くの心理学実験においては、人間のリズム感覚が時間的に変化したり同調・同期したりする現象や、複雑なリズムを知覚・認知する現象などから、「メロディー/和声のリズム」「テンポ」「拍節構造」「拍子の体制化」「グルーピング」「拍子の解釈」等の、音楽全般に関する高度で複雑な概念を対象としてきた [33]。リズムという用語自体もそうであるように、これらの概念の中には定義が研究ごとに微妙に異なり無用な誤解を生む可能性を持つ場合も少なくない。そこで本研究では、音楽的な対象を以下のように単純化し、登場する主要な用語についてその概念を再定義した。

テンポ

本研究においては音楽と映像それぞれのテンポは常に一定とする。本研究における実際の心理学実験の中では、「映像的ビートのテンポよりわずかだけ音楽的ビートのテンポが速い」という状況と、逆に「映像的ビートのテンポよりわずかだけ音楽的ビートのテンポが遅い」という状況を、数十秒間をかけて交互に行き来して両方の現象を実現しているが、想定しているそれぞれのローカルテンポについては一定である。つまり、クラシック音楽の「芸術的表現」のような局所的なテンポバート等の無い、シーケンサで機械的に生成したポップスやロックなどの、テンポ一定の音楽に対象を限定する。さらに、音楽アンサンブル演奏など人間が能動的にテンポを変えられる状況ではなく、例えばポップス音楽をウォークマンで聞くような、人間が自分で音楽のテンポを変えられない受動的な状況が対象である。映像についても同様で、映像的ビートを等間隔で繰り返すテンポも一定であり、変えられない。

拍子

本研究においては音楽の拍子も単純化して、いわゆるエイトビートのロック音楽、つまり「8分の8」だけとする。実際の心理学実験では、ドラムパートのクローズドハイハット(HH)音は8分音符でなく16分音符を刻んでいるが、バスターン(BD)とスネアドラム(SD)が生成するリズムの拍子としてはあくまでエイトビート(8/8)である。これによって、拍節構造やグルーピングといった複雑な要素を排除して、主眼である繰り返しリズムのアクセントとしてのビートだけを対象として明確にする。なお村尾の音楽ビート分類で言えばこの音楽は「ヒールビート」で

あり、弱拍で強勢のアクセント部分の手拍子にはスネアドラム(SD)が打たれることになる。

視覚的情報である映像については、「3-2-2. QuickTimeムービーの制作」で詳しく述べるように、アクセントから次のアクセントまで「滑らかに変化する」状態が持続するので、音楽的な拍子の言い方では「2拍子」のようなものと言える。ただし実験に用いる映像素材の映像的ビートについては、アクセント部分に意図的に時間幅を持たせているので、村尾の音楽ビート分類の「ヒールビート」ではない。

アクセント

リズムにおけるアクセントとは、周期的に繰り返すリズムの中で「強・弱」のうちの「強」に当たる部分である。これまでエイトビートの基本リズムは、1小節の繰り返し単位を4つの4分音符の並びとした「1・2・3・4」のうち「1・3」がアクセント(強拍:BD)で、「2・4」が裏拍(弱拍:SD)とされた。しかし、ここに同時に8分音符のHHを8つ入れて、1小節の繰り返し単位を8つの8分音符の並びの「1・2・3・4・5・6・7・8」としてみると、実は「1・5」のBDも「3・7」のSDも共にアクセントであり、裏拍というのはHHのみの「2・4・6・8」である事が分かる。このいずれが優勢と知覚されるかは音楽のテンポにも関係する。これは村尾の音楽ビート分類の「ヒールビート」の概念を導入することにより明確に理解でき、「弱拍で強勢のアクセント」というのは矛盾せず、ここに手拍子ないしSDが自然に入る。

本研究における心理学実験の基準テンポとしては、8分音符を110msecとした。これは音楽用語の速度記号の言い方では、およそ「4分音符 = 136」(アレグロ)という事になり、アクセントとしては上記の「1・5」のBDと「3・7」のSDの位置が該当する。なお、音楽において音量だけでなく音価durationが影響するように、リズムのアクセントとは時間的には「瞬間」ではなく、心理的にはある時間幅を持っている、という点を指摘しておく [1][2]。

ノリ

広辞苑の「リズム」には「繰り返して起こる事柄によって起こる快い印象」という定義もあるが、心理学では「心地よい」という価値観が捨象される事も多い[34]。しかしリズムの「精神の抑制からの解放という生命現象」「力の消費を最小化調整する要素として身体活動を支配」「運動性の刺激と情緒的反応を伴う」等の特性をふまえれば、これは重大な切り捨てとも言える [2][18]。

本研究においては「ノリ」という概念を提案し、従来の「リズムのアクセント部分」でなく、ビートの定義のために「ノリのあるリズムのアクセント部分」を考えることにする [31][63][64][65]。簡単に言えば、一例としてロックコンサートで聴衆が心地よく踊っている場合、その身体動作リズムのアクセント部分にはノリがある(乗れる)、という事である。ここ対象とする音楽のリズムは村尾の音楽ビー

ト分類における「ヒールビート」なので、ノリのあるビートとは、強拍のBDの部分と、弱拍で強勢のアクセントであるSDの部分ということになる。

上記のエイトビートで1小節の繰り返し単位を8つの8分音符「1・2・3・4・5・6・7・8」としてみると、「1・5」のBDと「3・7」のSDのアクセントのうち、ロックコンサートの聴衆が音楽に合わせて足踏みをしたり手を叩く「ノリ」の場所は、テンポが比較的遅ければ「1・3・5・7」の全てとなり、テンポが比較的速い場合にはBDの場所だけでなくSDの「3・7」(アフタービート)となる事が一般的である。

ただしこれは音楽を受動的に享受する聴衆の側の場合であり、そのコンサートで演奏している(ビートを生み出している)ドラマーは、テンポに関わらず(いくら速くなくても)「1・3・5・7」の全てをノリのあるビートと体感しているし、BDを単純な「1・5」でなく、「1・6」や「1・5・6」や「1・4・6」とキックしてドライブ感を増した場合でも、ノリは実際の音が出る「6」や「4」でなく「5」に感じている。なお、アクセントは瞬間でなく時間幅があるので、ノリにも当然ながら時間的な広がりがある。

ビート

上述のように、本研究における「ビート」は、たんに「リズムのアクセント部分」というだけでなく、「ノリのあるリズムのアクセント部分」と定義する。ノリがあるため、このビートの繰り返しを知覚する人間は「気持ちいい」心理状態(情緒的反応)になる、という部分が重要である。ロックコンサートでミュージシャンが提供する音楽のビートを享受する聴衆の場合、ビートに乗ることで自分のリズムのテンポを音楽のテンポと同調/同期させることが「快い」状態への最適解となる。

本研究で注目するのは、人間の知覚心理において「ビートにはかなりの時間的な幅がある」という点である。ここでは、音楽的ビートはヒールビートであるため、聴覚の時間的分解能から比較的狭い時間幅であると考えられる。これに対して身体的ビートは、マーチングバンドの行進ステップと演奏しているマーチ音楽のビートを調べた研究により、その時間的な幅が広いこと、経験者と初心者などの個人差が大きいことが報告された [31][63][64][65]。イメージ上の運動であれば質量ゼロで瞬間的に動き出したり停止したりできるが、実際には運動している肉体の重みによる時間的な幅と、トレーニングによる制御能力の差が出てくるのは当然である。これは村尾の分類で言えばヒールビートとは違うものであり、身体運動において自然なのはむしろタゴビートないしトゥビートである、という可能性も示唆している。

視覚的ビートについては、視覚の時間的分解能から推定すると、この中間の時間帯であると考えられる。また、視覚に特有の反応速度と視覚の残存時間(残像)の効果により、聴覚より広い心理的ビート時間幅を持つとも考えられる。これも村尾の分類で言えばヒールビートとは違う、タゴビートないしトゥビートの性格をより強く持っている可能性も示唆し

ている。

そこで本研究では、テンポが自分では変化できないという制約下で音楽情報と映像情報とを同時に視聴する場合、この視覚的ビートの時間幅の広さを無意識に活用してノリを得ている、というモデルを提案した。

「引き込み現象」と「引き込み効果」

非線形振動子(リミットサイクル)は相互作用によって引き込み現象entrainmentを示す。振動子集団が相互引き込みによってマクロなリズムを発生させる例は、心拍、脳波、概日性生理リズム、蛍の集団発光など、生命現象を中心として広く自然界に見られる。異なる周波数の時計を結合させると同期する「ホイヘンスの時計」や、二つの発振回路を隣接させると発振周波数が吸い込まれて同調するなど、物理現象としても引き込み現象はよく知られている。日本心電学会用語委員会は、entrainmentの訳語として「引き込み」や「乗り込み」でなく、能動的ニュアンスの「乗っ取り」や「令引(造語)」を提案している。

音楽アンサンブルのリズムにおいても、タッピング時の脳波測定実験とともに引き込み現象entrainmentによるリズム同調が研究されている[52]。これは日常的な音楽アンサンブルにおいて、指揮者やバンドマスターが指示しなくても全体のテンポやリズムが同期していくメカニズムとして自然なものであり、受動的ニュアンスの「引き込み」が適している。この場合、引き込み現象によって一方のリズム/テンポが次第に相手のリズム/テンポに同調していく。

これに対して本研究の実験では、音楽的ビートと映像的ビートは制約条件によって同調することなく、両者のテンポはわずかに異なったままそれぞれ続く。そして知覚的にこの折り合いをつけるための心理的メカニズムが、提唱するモデルである。

従って、引き込み現象という意味での同調は発生せず、あくまで知覚における心理的影響として引き込まれる「効果」がある、という点が大きく異なる。そこで敢えて本研究では「引き込み効果」という用語を定義し、混乱を避けるためにentrainmentという英語を用いずに一般的な英語でDrawing-in Effectと直訳した。ただし英語表記については、今後より適切な表現があれば躊躇なく改訂していく方針である。

2-3. 身体的ビートと音楽的ビートと映像的ビート

前項の「関連用語と概念」の定義を受けて、本研究が提唱する「引き込み効果」の説明のために、あらためて身体的ビートと音楽的ビートと映像的ビートについて整理する。ビートの定義を「ノリのあるリズムのアクセント部分」(そこに乗ることで気持ちいい部分)としたので、これらのビートを考える上では、ビートの時間的な幅の検討とあわせて、運動性の刺激や心理状態(情緒的反応)という視点が重要となる。

2-3-1. 身体的ビート

本研究のテーマは音楽的ビートと映像的ビートとの関係であり、身体的ビートは考察のための補助的な情報となるが、「音楽的ビートの筋運動感覚的側面」と「音映像的ビートの身体運動的側面」とを媒介している、という意味では非常に重要である。

聴覚による音楽的ビートや視覚による映像的ビートに比べて、身体的ビートの場合には、身体動作は自己の肉体という質量を伴う現象であるため、ビートの時間幅はかなり広いと考えられる。文献[64][65]のマーチングバンドの行進ステップのビートとマーチ音楽のビートとを調べた研究から検討すると、身体的ビートに対応するタイミングとしては、(1)足の裏が着地して地面を叩く瞬間、(2)体重がもっとも重くかかる(知覚される反作用が最大)時間帯、(3)地面を強く蹴る(力が入る)瞬間、(4)地面から離れる瞬間(つま先の蹴り出し)、(5)太ももが一番高い瞬間、などの候補が考えられる[31]。このような種類の身体動作のビートは、村尾の分類では「トゥビート」として考えることが適当であろう。

リズムの本質である「継起する事象の秩序立った特性」により、人間はビートをあらかじめ無意識に予測でき、心的構築によってそのアクセントの瞬間に向けてあらかじめ先行して身体動作を開始し、アクセントの瞬間にそのピーク(タッピング、ステップなど)を迎え、瞬時にこの一致を知覚フィードバックしてその同期を確認し満足することで快さを得ると考えられる。ここで慣性質量や筋力など身体の物理的条件をあらかじめ経験的/無意識的に学習していることは、指揮法メソッドに顕著のように非常に重要である。指揮法とは、視覚を通じて音楽的ビートを指揮者から演奏者に伝えるための技法であり、視覚を経由して演奏者が身体的ビートを確実に受け取るための運動軌跡のルール、そして指揮者自身が視覚を経由して音楽的ビートを効果的に表現/伝達するための身体動作の運動学的なルール、と言うことができる。

ビートという言葉から拍手(ハンドクラップ)のような打撃をイメージすると、ノリのあるビートの瞬間に身体(の一部)が高速で運動したり打撃の反発があったりする、と誤解されるが、身体的ビートはむしろ「運動の合間の静止」である、という視点があることもここで指摘しておきたい。もっとも典型的にはリズムカルなダンスの一種であるブレイクダンスを観察すると、ビートの瞬間は強制的に身体が静止して、その個々の静止点の間を高速で身体(の一部)を運動させている。このような種類の身体動作のビートは、村尾の分類ではまさに「タゴビート」として考えることが適当であろう。

2-3-2. 音楽的ビート

音楽的ビートにおいても、アクセントという事で音量の極大点、あるいは複数の楽器音が同時に鳴るポイントという誤解があるが、本当にノリのある音楽的ビートは身体的ビートと同じく、むしろ「点」でなく「幅を持つ」と考えられる。ソウルやレゲエなどの音楽はもちろんのこと、ロック音楽であって

もドラマーの定番テクニックとして常識となっているのは、スネアドラム(SD)を入れるタイミングを微妙に「溜める」こと、すなわち微小時間だけジャストのタイミングから遅らせることである。SDだけをタメて遅らせても、他の楽器を叩くリズム全体の枠組みやテンポ自体はつられて遅れてはいけない、というのが初心者ドラマーに対する基礎的なアドバイスである。これはノリの考え方を含めた音楽的ビートという視点からは、SDの入るビートにはかなりの時間幅がある、という事実を示している。クラシックで有名な「ウイナワルツ」において、2拍目のタイミングが早めにずれることでノリが出ているのもまったく同様である。

ただし、本研究においては対象とする音楽としては「ヒールビート」、つまりアクセント部分に時間幅がなく叩き付けて打ち返す、というリズムであり、ここでは音楽のビートの時間幅は十分に小さいものとして検討を進める。

2-3-3. 映像的ビート

映像的ビートとしては、映像の中に明らかな視覚的アクセントが周期的に存在し、その連なりが繰り返しによりリズムを作り出すタイミング、とまず定義する。一般には映像が瞬間的に激しく動く場所がアクセントであると思われがちであるが、突然の動き出しは予期できないので、ノリのあるビートとはなりえない。これとは逆に、ほぼ一定で動いている映像が周期的に静止し、しかもその繰り返しパターンが予測される関係性を持つ場合には、周期的に出現する静止画の部分の逆をビートとして容易に知覚できる。ブレイクダンスの静止姿勢のアクセント感はこの原理をそのまま実現した好例であり、踊っている本人の身体感覚としてだけでなく、これを見て人間にとってもビート感を与える。

本研究の心理学実験のための映像素材(視覚的刺激)については、知覚心理学から得られた研究結果を検討して、ムービーという動画の中で、周期的に動きと静止がなめらかに交互に繰り返す、というリズムを構成し、ここからノリのある映像的ビートを被験者に提供することにした。言い換えれば、この視覚刺激素材のムービーにおいては、ブレイクダンスのビート感を重視し、視覚的ビートとして「変化する映像が周期的に静止する」というタイミングがビートとなるように制作した。これは、ある意味ではタゴビートの静止点とトゥビートの加速感・減速感を折衷したようなものとも言えよう。

2-4. 局所的なビートの引き込み効果

本研究では、聴覚的情報と視覚的情報を同時に視聴するマルチメディアコンテンツにおいて、ビートを「周期的に繰り返しリズムのノリが知覚されるアクセント部分」と定義し、音楽的ビートが映像的ビートの知覚に及ぼす局所的な「引き込み効果」を提唱する。まず最初に、思考実験として2つの一般的な体験例を紹介し、その上で心理的な現象としての解釈を検討する。

思考実験1

ある一定の快適なテンポで運動している人間がいる。ウォーキングでも行進でもマラソンでも構わないが、その身体運動のビートとテンポはちょうど早く、これを断固として変えずに続けることがもっとも気持ちいい状態である。この人間が、同時にウォークマンで「ほとんど同じテンポの音楽」を聞いてその音楽ビートを楽しんでいるとする。この音楽は自分が演奏しているわけではないので、テンポは変えることはできない。運動のテンポも快適なので、無理に音楽には合わせない(変わらない)。

身体運動ビートというのはいろいろな種類があり、さらにかなり緩やかな時間幅がある。足の裏が着地して地面を叩く瞬間、体重がもっとも重くかかる(反作用も最大)の瞬間、地面を強く蹴る瞬間、地面から離れる瞬間(つま先の蹴り出し)、逆に太ももが一番高くなる瞬間、など色々ある。

すると、最初は音楽の表ビート(BDのビート)で音楽と運動が同期することで気持ちいい、と同期したビートを楽しんでいるのが、次第にズレた時に、この曖昧な時間幅の範囲でいつのまにか音楽の裏ビート(SDのビート)に変わること、再びこれを許容してビートにノれてしまっていた(以下、これが交互に続く)。これが「引き込み効果」である。

思考実験2

新幹線の車窓から風景を見ている。あるいは高速道路を走る高速バスの車窓から風景を見ている。スピードは一定であり、移動は長く続いた直線で変化がなく、そこに一定の間隔できちんと並んだ電柱や標識などの目印が、一定のリズムで視界に飛び込んで去って行く。これは視覚的情報のリズムであり、加速度的に近付いては消え去って行く躍動感の繰り返しにノリを感じるために、この風景が視覚的ビートとして心地よい。

ここでも、同時にウォークマンで「ほとんど同じテンポの音楽」を聞いてその音楽ビートを楽しんでいるとする。自分が演奏しているわけではないので、音楽のテンポは変えることはできない。

風景のビートは快適、音楽のビートも快適、しかしいずれも変えられないが、微妙に音楽のビートが風景のビートと違う場合にどうなるか。視覚的ビートにもいろいろな種類があり、さらに緩やかな時間幅がある。遠くから新しい目印が視界に入る瞬間、そこから近づくまでぐんと加速する瞬間、自分の真横を通過する瞬間、など色々ある。

すると、最初は音楽の表ビート(BDのビート)で音楽と風景が同期することで気持ちいい、と同期したビートを楽しんでいるのが次第にズレた時に、この曖昧な時間幅の範囲でいつのまにか音楽の裏ビート(SDのビート)と風景と同期することで、再びこれを許容してビートにノれてしまっていた(以下、これが交互に続く)。これが「引き込み効果」である。

引き込み効果

上記の2例はいずれも同一の現象である。ここでは「ほとんど同じテンポの音楽」が、微妙に運動の

テンポよりも速いとして(逆でも議論は同様である)、思考実験1の運動の例を、さらに時間的に細かく追跡する。

歩行運動の例では、「運動ビートの瞬間」が、さっきは地面から離れる瞬間(つま先の蹴り出し)が音楽ビートと一致していたのが、次第に面を強く蹴る瞬間、さらに足の裏が着地して地面を叩く瞬間、と少しずつ前倒しされていく。しかし、いずれも時間的に幅を持つ運動ビートの幅の中なので許容され、乗れる。

村尾の音楽ビート分類で言えば「トウビート」の拍子化された図式のように、アクセント付近にはいくつかのポイントがあり、音楽よりもアクセントの許容幅の広い身体運動ビートとしては、そのどこかにはまればノリとして許容できる。

そしてズレがいよいよ大きくなると、例えば音楽の表ビート(BDのビート)と同期していた運動ビートが、今度は自然に音楽の裏ビート(SDのビート)に「乗り換えて」またまた両者の同期したビートを楽しむ、という現象が起きる。細かく見れば、さっきまで足の裏が着地して地面を叩く瞬間と一致していた音楽のビートがずれたと知覚されると、音楽の裏の反対側のビートがこれから足の裏が着地して地面を叩く瞬間にほぼ近くなり、このずれは次第に縮まって一致してきて、さらに地面を強く蹴る瞬間へ、地面から離れる瞬間(つま先の蹴り出し)へ、と一致しながらノリが続く。

結果として、音楽のビートのテンポも運動のビートのテンポもともに変化しない(引き込み現象 entrainmentではない)にもかかわらず、運動ビートの許容幅の中で音楽ビートとの一致を自然に楽しむことで、音楽ビートに対して見てみると、しばらくは表ビートで乗り、やがて裏ビートで乗り、また次に表ビートでのノリに、と(無意識的に)乗り換えている。一致していると解釈することで気持ちいい、という瞬間に着目して運動のビートを見ると、そこでは少しだけ速いテンポの表ビートに引き寄せられて局所的に微小に加速し、やがて後から来る裏ビートに乗り換えてまた加速し、という速度の微小なギザギザ変化を繰り返すことになる。これが「引き込み効果」である。

2-5. 引き込み効果の定量化と可視化

この引き込み効果は少なくとも筆者の経験と内観においては明確なのであるが、これを心理学実験によって顕在化・定量化すること、さらに理解のために可視化することはできるだろうか。そのヒントは、上記の最後の部分、「一致していると解釈することで気持ちいい、という瞬間に着目して運動のビートを見ると、少しだけ速いテンポの表ビートに引き寄せられて局所的に微小に加速し、やがて後から来る裏ビートに乗り換えてまた加速し、という速度の微小なギザギザ変化を繰り返す」にあると考えられる。

ビートの定義としてノリ、すなわち気持ちいい、言い換えれば半ば無意識に人間がビートのノリを納得している瞬間を、音楽的ビートと映像的ビートと同時に被験者から獲得できれば、その時間情報を分

析することが可能である。ここでポイントとなるのは、リズムやビートの性質として最初から指摘してきた「身体動作(筋運動感覚)的側面」である。簡単に言えば、音楽を聞いて、映像を見て、そこにビートを感じている人間は自然に身体動作のビートを伴うのであるから、それをタッピング動作として取り出せばいいのでは、というのが本研究における心理学実験の基本方針である。

この視点から、引き込み効果の定量化のために心理学実験に必要な要素を整理すると、概念的には以下となる。その具体的な詳細については、制作のところで詳述する。

1. テンポが一定でビート感のある音楽的刺激。飽きないものが欲しい
2. テンポが一定でビート感のある映像的刺激。飽きないものが欲しい
3. 上記テンポの一方は、基本的には一定でありながら、実験の中では微妙に変化させたい
4. 被験者がビート感の反応を返すタッピングのセンサ情報
5. この「音楽」「映像」「タッピング」を同一の基準タイミングで正確に計測・記録するシステム

さらに、この心理学実験の結果データが得られたとして、引き込み効果の可視化のために必要となる要素を整理すると、概念的には以下となる。その具体的な詳細については、制作のところで詳述する。

1. 実験時間の経過をX軸方向、個々のデータの時間的位置をY軸方向にプロット
2. 音楽的ビートと映像的ビートをそれぞれプロット
3. この両者のうち一方を等速に固定すると、それがX軸上に等間隔で並ぶ
4. もう一方は少しずつずれるので、X軸と一定の角度の平行線上にプロットされる筈
5. この平面上に、さらに被験者がビート感の反応を返すタッピングをプロットして比較する
6. 時間のずれを定量化してX軸まで垂線で結び、その長さとする
7. 見やすさのために視点を俯瞰の位置に置いた3次元表示としたい

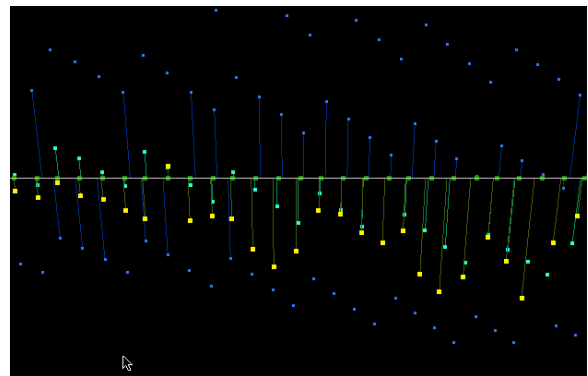


図4 3次元表示された「引き込み効果」
(緑:映像ビート、青:音楽ビート、黄:タッピング)
図4は、本研究の心理学実験において実際に被験者から得られたデータの一部分を3次元表示したものである。

3. 心理学実験システムの制作

ここでは本研究のために新しく制作した心理学実験システムの設計と実装の全般について、個別の要素ごとに具体的な検討事項と設計・開発・実装の実例について述べる。システムの制作と実験は、2003年3-4月の約2週間にわたって静岡文化芸術大学1106研究室で行われた。以下は、この心理学実験システムの全体図である(図5)。

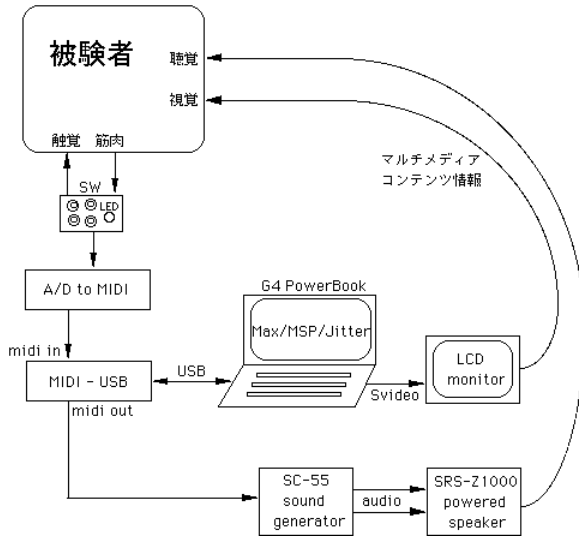


図5 心理学実験システムの全体図

3-1. プラットフォーム

心理学実験システムのプラットフォームと、使用したツール類の詳細は以下の通りである。コンピュータ(OS)としてUnix(Linux/IRIX)やWindowsでなくMacintoshを利用した理由は、中核となるソフトウェア"Max/MSP/Jitter"はUnixやWindowsに対応しておらず、本実験を行った時点ではMacintosh(OS9.2)でのみ使えたからである。なお本実験後、2003年4月になってMac OSX対応版が公開されたが処理能力はやや低かった。そして2003年8月21に、数年来の懸案であったWindows版がリリースされた。今後のMax/MSP/Jitterについては、Mac OSXとWindows XPのクロスプラットフォームとして進展するという情報であるが、これまでの状況から将来についてはまだ未確定であろう。

メインプラットフォーム

Apple Macintoshコンピュータ"G4PowerBook" : G4プロセッサ、クロック667MHz、メモリ1GB、MacOS9.2.2、QuickTime5.0.2、OMS2.3.8、S-video出力、RAMディスク100MB
ソフトウェア環境

Cycling'74 "Max/MSP/Jitter"(version4.1) 上での自作パッチ(実験ソフト/データ可視化ソフト)、他に "Jedit", "Graphic Converter", "Premiere", "Morphor", "Play It Cool", "Claris Works", "iMovie", "Fetch", "MacDraw Pro"等のソフトウェアを適宜使用

MIDI関係

USB-MIDIインターフェース : Roland "UM-2"、GM音源モジュール : Roland "SC-55mkII"、被験者MIDIタッピング装置 :

自作(CPU:AKI-H8)

オーディオ/ビデオ関係

ソニー パワードスピーカ"SRS-Z1000"、シャープ 液晶モニター"LC-10RV2-A"、リオン 騒音計"NL-05"

レイテンシ計測関係

東芝コンピュータ"DynaBookSS-R590"(Windows95)、Pico Technology デジタルサンプリングオシロスコープ"ADC-200"、日立 プローブ"AT-10AR"、ソフトウェア "Picoscope for Windows"

システム全体のプラットフォームとなり、実験ソフトウェア自身、データの3次元可視化ツール、レイテンシ計測ツールの全てを構築したベースは、Cycling74社のソフトウェア"Max/MSP/Jitter"である。なお、このツールは世界的に有名なので解説は省略する(文献[71]を参照のこと)。

このMax/MSP/Jitter(メモリ割当て:200MB)をプラットフォームとする事で、1msec分解能でのMIDI情報処理、44.1KHzサンプリングでの音響情報処理、ベストエフォート条件において同等の分解能での画像処理をシームレスに構築できる。これは心理学実験において重要となるシステムのレイテンシや人間の知覚スピードに対して、詳細が不明な複数の機材などの要素を接続することで発生する不確定要素を排除するという重要なメリットとなる。

3-2. 映像素材の制作

ここでは、本研究の心理学実験において使用した映像素材の制作について述べる。まず最初に、従来の心理学・認知科学の研究において心理学実験に用いられた視覚的刺激の問題点を検討する。次に実際の映像素材をQuickTimeムービーとして制作した過程を紹介し、さらに心理学実験システム内での再生環境を報告する。

3-2-1. 視覚的刺激の問題点の検討

従来の研究で視覚的刺激として使用されてきたものに、モニタ上の抽象的な図形の明滅やLEDランプ等を瞬間的に点灯させるフラッシュタイプがある。しかしこれは、音楽知覚との同期というよりは、視覚的刺激の変化に対する反射神経・運動神経のテストという性格が強くなる問題点が指摘された。また、以下のようにこれまでに知られている知覚心理学的な報告との関係にも注意する必要がある。(1)視覚の不連続の閾値は100msec、(2)光刺激を瞬間と意識する視覚の瞬間域は60-100msec、(3)断続する視覚的刺激を融合して連続と感じる視覚系臨界融合周波数CFFは低輝度の場合には2-3Hz、高輝度では60Hz、(4)視覚システムによって光パターンが保持される時間は150-200msec、(5)静止画などを注視する平均的な時間は200-300msec [7][9][13][15][19]。

従来の研究で用いられた他の視覚的刺激としては、指揮棒の先端が描く放物線や、単純な幾何学図形の軌道上を運動する輝点により、ビート点への到達を予測させる軌跡追跡タイプがあった。しかしこれは、単純な視覚的刺激の繰り返しで、短時間で飽きがきたり眠気を催しやすい。本研究では実験時間としておよそ2-3分間の間に起きる心理的知覚現象を計測するために、その間ずっと被験者の注意(興

味)を喚起するような、自然で一般的な映像が視覚刺激の素材として要求される。

従来の研究で一般的な映像を使用したものに、風景などの静止画や、映画のワンシーンなどの映像を強制的に切り替える瞬間をビート(アクセント)とする、映像スイッチングタイプがある。しかし人物写真とか風景写真のショットを一定時間ごとに切り替えた場合、新しいショットが提示されると、サッカーsaccadeという眼球運動(痙攣的にピクッと動く)が起きる。このサッカーの瞬間は視覚的に抑制されて目は見えない状態である(これにより視野のブレ感覚を抑制する) [19]。映像や画像を瞬間的に切り替えるこのような視覚的刺激の場合、被験者は短い注視時間で、かつ長いサッカー距離をもった画面走査を行って、画面全体のおおまかな情報を掴もうとする。被験者の視線はめまぐるしい移動を強制され、意味的な繋がりが無い(余計な先入観や予測や意味を排除)ために、集中を持続して画面を注視することは非常に困難となる。

映像において画面を切り替える手法としては、上記のように瞬間的にスイッチングするだけではない。より自然に画像を切り替えるための代表的な方法(エフェクト)としては、(1)ワイプ、(2)クロスズーム、(3)クロスディゾルブ(クロスフェード)、(4)モーフィング、等がある。(1)のワイプの場合、各種のワイプパターンが知られているが、いずれも画像そのものの情報に加えて、ワイプの運動パターンと運動方向という新しい情報が追加されるために、視覚的刺激としては適当ではない。(2)のクロスズームにおいても、画像そのものの情報に加えて、ズームインとズームアウトという運動の知覚情報が追加されるために、視覚的刺激としては適当ではない。(3)のクロスディゾルブ(クロスフェード)の場合には運動成分はなくなるものの、輝度変化や透明度変化により前後の画像が接続されるために、肝心の映像変化のアクセントである視覚的ビートが時間的に曖昧になる欠点がある。

ところで(4)のモーフィングについては、前後の静止画に似た性質の画像(例えば、ほぼ同じポーズとアングルの犬の画像同士)であれば、モーフィングの対応点(鼻から鼻、尻尾から尻尾など)を指定することで、スムーズに前後の画像が繋がる。これはほとんどサッカー眼球運動を起こさない、視点を固定しつつも注意を喚起する有効な素材である。また、モーフィングしている最中はいわば運動中の状態で、モーフィングのスタートとエンドの部分は静止していることで、これを繰り返すことはリズムを生み出し、うまく編集することでノリのある映像的ビートを構成できる可能性がある。そこで本研究の心理学実験においては、映像素材としてこのモーフィングを採用した。

3-2-2. QuickTimeムービーの制作

以上のような状況を受けて、本研究では以下に述べる手順で、被験者テスト(心理学実験)のためのサンプル映像をQuickTimeムービーとして制作した。

映像の素材は、幾何学図形などの作為的な図形でなく、人間にとって自然な静止画の風景写真をペー

スとするために、筆者がかつて映像作品の制作のために取材撮影した約500体のお地藏さんのデジカメ写真を用いた。取材地は静岡県奥浜名湖地方にある、奥山半僧坊・方広寺(五百羅漢で知られる名刹)である。

作業はまず、500枚を越えるデジカメ写真から、(1)画面中央に地藏が1体だけ写っている、(2)周囲は自然の風景(森林や土手など)、(3)個々の地藏は全て異なったもの、という条件で240枚を選んだ。芸能人とか動物の写真でなく地藏としたのは、(a)単なる風景でなく中心に地藏というテーマがあると被験者が集中できる、(b)人間や動物などでは個人的・個性的な嗜好で反応して実験誤差を生む、(c)地藏は全て石で出来ているために基本的なテクスチャが共通で、無個性の同等キャラクタと知覚しやすい、(d)ほぼ中央に無個性の地藏があると切り替え時に視線が画面全体を無駄に走査する必要がない、等の理由による。

次に、この240枚のVGA(640*480ピクセル)静止画データから作為なく撮影順に隣接した2枚を次々に取り出して、シェアウェア"Morpher"(M.Fujimiya氏制作)を利用して、以下の設定条件でモーフィングムービーを制作した。

1. モーフィング前後のそれぞれの地藏の17箇所(頭頂、両目、鼻、口、顎、両耳上、両耳下、両肩、胴体両脇、胴体両底部、底部中央)を個別にクリックして対応づける
2. モーフィングは等速でなく、最初と最後が減速され途中が加速されS字カーブ特性の"Natural Mode"を指定する
3. スタート画像と最終画像との間を30段階に分割してモーフィングを行い(全体としては31フレームのムービー)、レンダリング最高画質モードを指定する
4. フレーム速度を10fpsとし(ムービーの長さは約3秒)、画質は無圧縮で最高画質モードを指定する

これによって、240枚の静止画から隣接する2枚を選んで1本のムービーが作られ、次のムービーは最終画像をスタートとして次の1枚の画像と組み合わせるので、ムービーは全部で239本となった。画像無圧縮指定のために全てのムービーのサイズは同一(約19MB)である。モーフィング速度として単純なりニア特性でなく、最初と最後が減速され途中が加速されるS字カーブ特性の"Natural Mode"を指定したことは重要であり、結果として視覚的ビートのノリを生み出す「自然な静止時間」が得られた。

上記の設定条件では標準的なG4Macコンピュータ1台で1本のムービーのレンダリング処理に3-4時間かかり、239本のムービー制作の期間は45台のMacを使って約1週間であった。

239本のQuickTimeムービーが完成した後に、シェアウェア"Play it Cool!"(Bridge 1 Software社)を利用して、隣接する静止画同士が結合するように全て連結して、全体として1本のムービー(4552MB)となった。例えば静止画[A]から静止画[B]までのモーフィングムービーの最後は静止画[B]であり、ここに静止画[B]から静止画[C]までのモーフィングムービーを連結することで[A] [B] [C]という滑らかなムービーとなり、これを240枚の静止画が順に登場するようにした。

モーフィング速度をS字特性として静止画の付近

では減速している事と、静止画は結合した両方のムービーに共通で2フレーム連続している事により、ムービー全体を見た場合には、静止画部分を(ブレイクダンスと似た感覚で)映像的ビートと見た場合には、ビート部分の候補時間帯が広くなり、ピンポイントに限定されている反射神経テストのような従来の提示映像刺激に比べて自然な実験素材を実現できた。

このVGAサイズ無圧縮のQuickTimeムービーはDVD-R(4.3GB)にも入らないサイズとなったため、実際に実験で利用する素材ムービーとして、ソフトウェア"Premiere"(Adobe Systems)を使用して圧縮した。ムービー再圧縮のため設定した条件は以下である。

1. 画素数を320*240ピクセルと1/4に設定する
2. 画像圧縮プロトコルとして「フォトJPEG」形式(画質 : 50パーセント)を指定する
3. 速度(デュレーション)を4倍速(400パーセント)に指定する(フレームを間引く)
4. フレーム速度を10fpsとする
5. これにより静止画から静止画までの平均フレーム数は7.75フレーム(ムービー全体 : 1852フレーム [=31*239/4])となる

画素数を1/4に削減した理由は、「標準画像の解像度を1/2-1/4-1/8と劣化させても、コントラストを劣化させても、画面サイズを1/4-1/9-1/16と小さくしても、画像に対する認知度はほとんど変化しなかった」という過去の研究報告による[19]。以上により、完成したQuickTimeムービーのサイズは45.9MBとなった。

3-2-3. Max/MSP/Jitterシステム上でのQuickTimeムービーの再生

完成した映像素材QuickTimeムービーは、100MBの領域を確保したRAMディスク内に置くことで、HDDアクセス時の速度変動の無いスムーズなムービー再生を実現した。QuickTimeムービーは、初期値110msecごとにインクリメントされる数値をフレーム指定として与えることで、フレーム単位で駆動して再生した。図6は、この心理学実験システムMax/MSP/Jitterメインパッチである。なお、このメインパッチ中にあるサブパッチの詳細については、音楽素材および被験者インターフェイスに関するものなので、それぞれのところで詳述している。

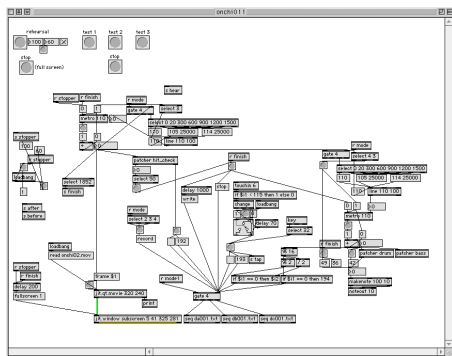


図6 心理学実験システムメインパッチ

3-2-4. 被験者への映像素材の提示

320*240ピクセルのムービー用スクリーンウインドウは、コンピュータのサブ画面内に移動して「フルスクリーン」指定によってVGA画面(640*480)にフルサイズ拡大され、S-video画像出力端子からSビデオケーブル(長さ1メートル)で接続されたSHARP LCDモニタに表示した。

LCDモニタは被験者の正面およそ50cmの距離に置き、画面の輝度は「明るい・標準・暗い」の3段階のうち、もっとも映像刺激として鮮明な「明るい」を設定した。モニタ背面のスタンドを被験者自身が調整して、もっとも画面を見やすい角度に調整したために、液晶画面に特有の角度による見にくさは回避できた。

3-3. 音楽素材の制作

次いで、本研究の心理学実験において使用した音楽素材の制作について述べる。まず最初に、従来の心理学・認知科学の研究において心理学実験に用いられた音楽的刺激の問題点を検討する。次に実際の音楽素材をMax/MSP/Jitterのパッチとして制作した過程を紹介し、さらに心理学実験システム内での再生環境を報告する。

3-3-1. 音楽的刺激の問題点の検討

従来の研究で音楽的刺激として使用されてきたものに、メトロノーム音やクリック音やスタッカートで同一音高のピアノ音などがある。これは、和音やメロディーによって調的なリズムやフレーズのリズムなどの別種の要素が入り込むことを排除するためであるが、クリックとクリックとの時間的間隔の心理学実験ならともかく、本研究で目指すリズムの知覚実験としては重大な欠陥がある。

その第一の理由は、単純なクリックの羅列は心理的なゲシュタルト知覚の働きによっているに(2拍子系/3拍子系等)体制化されてしまう、という問題である [2]。第二の欠点は、短時間でなくビート感の引き込み効果という知覚認知の漸進のプロセス実験の刺激とするには、単調さに起因する慣れ・飽き・眠気などによって素材として適さない[17][33]。そして第三の、というよりも最大の問題点は、単なるクリックの羅列にはノリがないので、本研究の視点からはビートが存在しないことになるのである。

ノリのある音楽的素材としては、ロックやポップスなどの既存のダンスミュージックを使用すればいいが、これは高次の音楽的要素が多く含まれているために、心理学実験の音楽的刺激としては適さない。そこで次の候補として、ドラムマシンのようなシンプルなりズムパターンを繰り返すシーケンスデータを作り、これをシーケンサを用いて再生することが容易に思い付く。実験においてMax/MSP/Jitterで映像的ビートをフレーム単位で制御すると同時に、音楽的ビートも音符単位で時間制御するためには、簡単には標準MIDIファイル(SMF)形式などのシーケンスデータをDTMの手法で制作(作曲)し、これをMax/MSP/Jitterのシーケンスオブジェクトで再生

するMIDIテンポクロックを制御すればよい。

しかしこの方法でも、以下のような問題点と課題がある。

まず、(1)無個性の地蔵の羅列である映像素材は記憶しにくい、ドラムとベースによる音楽素材はシーケンサで固定して繰り返し再生すると、楽曲の一部ということで簡単に覚えてしまう。これでは、それぞれ約3分間の素材で2パターンの実験を行いたい、という状況では、慣れてしまっ「飽き」「眠気」などの問題が出てくる。

また、(2)被験者ごとのレスポンス特性をリハール段階でのチューニングにより補正するためには、再生のMIDIクロックでなく個々の音楽素材の生成時にタイミングを発生して、より確実なデータを収集したい。これはシーケンサでは不可能であるが、アルゴリズム作曲の手法によりMax/MSP/Jitterでは容易である。

そして、(3)被験者だけでなく実験実施者(筆者)自身も、次に出てくる音(コードのルート等)を予測できない状況がより好ましい。明らかにこの方が完全に映像と音楽のビートという条件だけを抽象化するのに良好であるが、これは乱数要素を盛り込んだアルゴリズム作曲の手法により、Max/MSP/Jitterでは実現容易である。

そこで本システムでは、以下に述べるように固定的な音楽素材のシーケンサ再生でなく、乱数を用いたMax/MSP/Jitterによるリアルタイムのアルゴリズムミック音楽生成を用いた。

3-3-2. Maxによる音楽素材生成パッチの制作

まず音源について検討した。Max/MSPの環境ではQuickTime標準のソフトウェアGM音源にも対応しているが、この音源には発音抜けと遅延のばらつきという重大な欠点があるために使用できない[27][56]。そこで、OMS経由でUSB-MIDIインターフェース(Roland UM-2)により接続する外部音源として、先行研究により発音遅延とそのばらつきについて良好な実験結果データが得られているRoland社のSC-55mkIIを使用した。ソフトウェアの音楽情報科学実験システムにMIDIを利用する場合にはレイテンシと遅延のばらつきについて十分に検討する必要があるが、本実験の場合には、被験者からのタッピング入力にも同じMIDIインターフェースを経由したことで、共通の遅延要素として誤差が相殺されるメリットも生まれた。

具体的には、ムービーを1フレーム進める基準クロック(16分音符に対応)である110msecごとのタイミングでハイハット音を16ビート刻む時間的単位(テンポ：4分音符 = 136.36/分)のもとで、いわゆるエイトビートのシンプルなロックドラムとスラップベースによるリズム系を作成した。ドラムやベースは同一パターン(フレーズ)を繰り返すと被験者が慣れてきて集中力が低下するために、8分音符が4個という半小節ループを1単位として、ドラムは2単位ごと、ベースは3単位ごとに乱数によりそれぞれ4種類のパターンから1種類を選択するようにした。さらにベースのスケール(調性感)のルート音は同様に2単位ごとに平均率12音から乱数で選択するように

したために、ドラムとベースだけの単純な音楽でありながら刻々と進行していき、単純な繰り返しとして知覚しにくい、無個性なビートを純粋に提供する音楽素材を実現できた。

ここで、一般的なBGM(ボーカルの無い音楽)ではドラムとベースとともに定番である「コード」(ピアノ、カッティングギター等の複音楽器パート)が無いのは何故か、という質問が予想されるが、これには明確な理由がある。コード楽器により音楽パターン中の刻々の和声が決まづけられると、人間は2の累乗や3の倍数などの枠組みの大局的な音楽構造、そしてドミナントモーションに代表される和声進行による音楽構造を無意識に予想してしまう。これは本研究の目的からすれば余計な音楽認知的要素を持ち込むことであり、実験として避けたい事である。ベースラインは単音であるために、明確なコード(和声)を単一に予測させにくく、そのフレーズのある種のスケールから選ぶことで無調のような不自然さは避けながら、刻々とルート音を移動(転調)させることで音楽的構造が予測しにくい、それでいてノリのある音楽を生成する要素として、敢えてドラムとベースだけ、という構成を採用した。

以下、具体的なサブパッチにより内容を紹介する。

バスドラムとスネアドラム

ドラムスの楽器は、メインパッチ内で全ての16分音符ごとに鳴るクローズドハイハット以外には、バスドラムとスネアドラムの2種類だけとした。以下のパッチによって、パターンを2回ループすること、4種類のパターンのうちの1つがランダムに選択されて繰り返す(図7)。

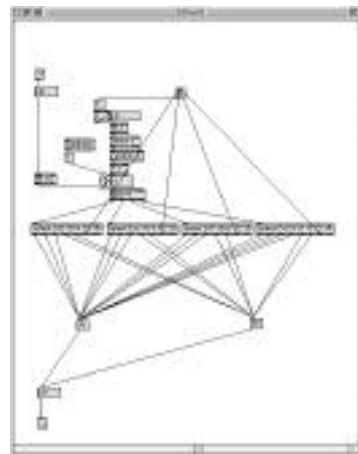


図7 バスドラムとスネアドラムのサブパッチ

ベース

ベースは以下のパッチによって、パターンを3回ループすること、4種類のパターンのうちの1つがランダムに選択されて繰り返す。これと同時に、パターンを2回ループすること、フレーズの調的中心となるルートの音を平均率12音中の隣接した9個の中からランダムに選択する。この組み合わせによ

て、ドラムとも異なる周期でベースのパターンとルートが変化するために、全体としてはシンプルでビート感がありながら、なんだか少しずつ違っていて気が抜けない、そこそこ注意を喚起しつつ飽きない、という音楽素材を実現できた(図8)。

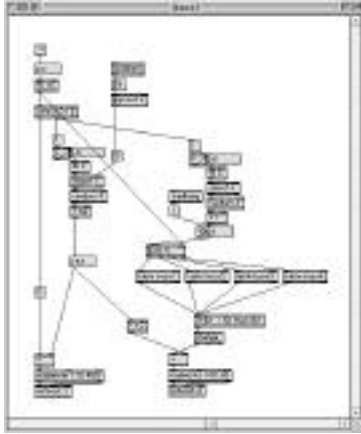


図8 ベースのサブパッチ

ベースのフレーズ

ベースの4種類のフレーズパターンは以下のテーブルパッチによって生成した。飽きが来ないこと、ビート感を出すために細部では8ビートでなく16ビートの要素が盛り込まれているが、それぞれのフレーズの冒頭は必ずルート音の長い音を置くことでエイトビートのノリを確保した(図9)。

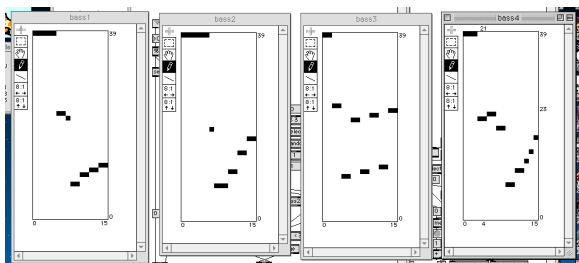


図9 ベースのフレーズのサブパッチ

3-3-3. 被験者への音楽素材の提示

このMaxからのMIDI音楽演奏情報は、Roland USB-MIDIインターフェース"UM-2"からMIDIケーブル(長さ3メートル)によって Roland GM音源モジュール"SC-55mkII"に与えられ、そのステレオサウンド出力(RCAライン:長さ1.5メートル)が ソニー パワードスピーカ"SRS-Z1000"によって音響となった。スピーカから被験者までの距離は約2メートル、左右スピーカの間隔は約1メートル、スピーカの高さは被験者の視線から上方へ約50cmである。

被験者の位置において、リオン騒音計"NL-05"にて計測した、実験室のバックグラウンドノイズの音圧レベルは45.5dbであった。本研究の心理学実験における、同じ被験者の位置におけるサウンドの音圧レベルは、10回の計測の平均値として以下のように

なった。なお、「実験1」「実験2」「実験3」という実験のモードについては、第5章で詳しく解説している。

「実験1」におけるヒット位置を示すピアノ音 56.6db
 「実験1」においてヒットした時のピアノ和音 62.7db
 「実験2」「実験3」においてBGMが鳴っている状態61.9db

3-4. 被験者インターフェース

ここでは、本研究の心理学実験において被験者のインターフェースとしてオリジナル開発したスイッチ(センサ)システムの制作について述べる。内容としては、(1)汎用アナログ-MIDIコンバータとして開発したセンサシステム、(2)このセンサに取り付けたマイクロスイッチとLED、(3)センサからのMIDI情報を映像情報と音楽情報と組み合わせて心理学実験データを得るためのMax/MSP/Jitterのパッチ、の3段階からなる。

なお、このセンサシステムに限らず、内部にマイクロコンピュータ(CPU)を内蔵した電子情報機器においては、ハードウェアやソフトウェアの内部処理を詳細に検討せずブラックボックスの状態のまま利用することは、データの信頼性や誤差・分解能などに重大な問題がある。

その好例は、Windows APIなどシステムソフトウェアのサービスルーチンをそのまま無批判に使用した心理実験であり、「マニュアルに1msecの分解能と書いてあるので実験の時間的分解能は1msecである」などというとんでもない間違いを平気で犯す危険性がある。ソフトウェアで「1msecの分解能」というのは「他に何のタスクも割り込みも発生しない場合の最大性能」というだけであり、実際には10msecとか20msecかかる場合も頻発し、基準時間という意味の分解能としては何も保証していない。たまたま最大瞬間風速のようなベストの値として、最良の場合には1msecで処理できるように設計しているつもりである、というだけである。

このような状況を受けて、ここではオリジナル開発した被験者インターフェースのシステムについては回路図からCPUソースコードまでを詳細に報告して、必要に応じて正確な追試による検証・批判に対応する。なお、「実験1」「実験2」「実験3」という実験のモードについては、第5章で詳しく解説している。

3-4-1. 汎用アナログ-MIDIコンバータ

本システムは情報処理学会から依頼されたチュートリアルのために1999年に制作したものである[66]。中核となるコンピュータシステムは、日立の32bit CPU H8/3048Fを搭載した、秋月電子のAKI-H8カードマイコンである。この基板にあるMAX232の11/12ピンをカットするだけでMIDI対応となる。全体の回路図は図10のようになる。

このシステムでは、CPUのクロックは16MHzであり、インストラクションサイクルは125nsecである。文献[71]で公開しているソースプログラムから判るように、A/Dコンバータの変換速度として134ステートの高速モードを指定し、単一モードで8チャ

ンネルの入力をソフトスキャンしているの、1チャンネルあたりの最大変換速度は134 μ secであり、前回のアナログ入力電圧と比較して変化があるとFIFOバッファに積んでMIDI出力する。データフォーマットとしては、「チャンネルプレッシャー」ステータスの1チャンネルから8チャンネルを用いた。

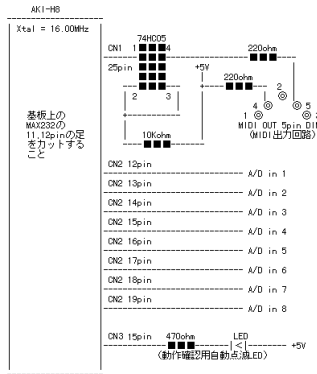


図10 汎用アナログ-MIDIコンバータ回路図

外部とのインターフェースとしては、8チャンネルの電圧入力(+5V-GND)端子から外部基板に延長して、4個のボリュームと4個の拡張端子を設けた。全体の外形は以下のようなものである(図11)。



図11 汎用アナログ-MIDIコンバータの外形

3-4-2. スイッチ(センサ)システム

この汎用アナログ-MIDIコンバータの外部電圧入力基板から+5Vの電源とGNDと1チャンネルの入力ラインとをケーブルで引き出して、図12のようなスイッチ基板をこの心理学実験のために制作した。

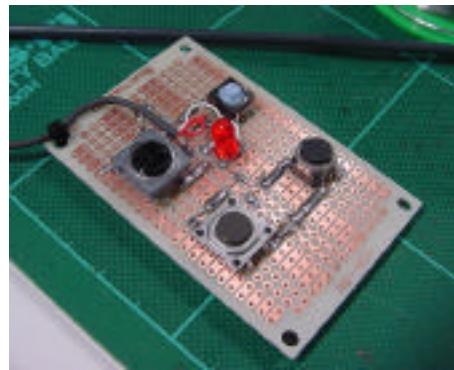


図12 タッピング用自作スイッチ

これは、タッチ感やストロークやクリック感の異なる4つのマイクロスイッチを並列に接続して、+5V電源により中央のLEDを点灯させる、という基板であり、そのLED端子電圧の変化を汎用アナログ-MIDIコンバータに出力した。

アルプス社製3種類、松下電気製1種類のこれらスイッチのストローク(接点間の移動距離)は、0.3mmが2種類、1.0mmが2種類であった。クリック感があるものと無いものの両方があったが、被験者の選択はまったくまちまちであり、どちらが多いという偏った傾向は見られなかった。

実際の心理学実験においては、被験者は「どれも同じ情報が出るので好きなスイッチを押して下さい」「押し方や持ち方も自由です」という指示により、思い思いの方法でスイッチ操作を行った。たいていの被験者はどれかのスイッチを押し続けたが、机の上で押していたのを途中で基板ごと持ち上げて押すようになった被験者と、両手で左右のスイッチを交互に押すという被験者がそれぞれ1人現れた。

3-4-3. 実験用パッチでのインターフェース処理

被験者のスイッチ操作によるMIDI情報は、Max/MSP/Jitterの走る同じコンピュータの同じUSB-MIDIインターフェース"UM-2"のMIDI入力にMIDIケーブル(長さ50cm)によって接続した。この情報は、心理学実験システムのメインパッチにある「touchin 6」オブジェクトにより入力され、通常はプルアップされて+5V付近である電圧が、スイッチによりGNDに接続されてLEDが点灯し低下した電圧として、バリュー115より小さい、というイベントとして検出される。この変化は逆方向(押していたスイッチを再び離す)のイベントと分離され、さらにチャタリング防止のために70msecのソフトウェアフィルタを経て、パッチ内部で「tap」というイベント情報となる。

で」を繰り返すデータ

なお、2バイトメッセージのステータス192-194を用いた理由は、不要な3バイトステータスを用いれば、イベントごとの情報量が増えることによる、無意味な誤差の可能性が増大するからである。安易にコントロールチェンジやピッチベンドなどの3バイトステータスメッセージを用いれば、単純に処理時間として1.5倍、さらにmakenoteオブジェクトによるノートイベントを使えば3倍の無駄な時間がかかることになるので、このような視点はリアルタイム心理学実験において重要であると考えられる。

3-5-2. 実験データの形式

以下は、テキストファイルとして得られたデータの一例である(図14)。このようにテキストファイル化されていることで、表計算ソフトに取り込んで分析処理することも容易である。データはスペースで区切られた数字の組み合わせであり、最初の数字が記録スタートからの経過時間(msec)、次の情報がMIDIメッセージである。それぞれのデータの最後にデリミタとしてコロンの付いているので、これを処理系でデータのまとまりの確認として利用すればよい。

```
13460 193 1;
13530 194 6;
13750 194 7;
13970 194 0;
14190 194 1;
14213 192 1;
14310 193 0;
14410 194 2;
14630 194 3;
14850 194 4;
15069 192 0;
15070 194 5;
15115 193 1;
```

図14 得られたデータの一例

この実験では記録される情報は全てステータス192-194の2バイトメッセージ、そしてこれに続くデータも0から7までの数字しか登場しないように統一している。一般のMIDI情報をそのままファイル化した場合には、2バイトメッセージと3バイトメッセージとが混在され、さらに「ランニングステータス」のルールによって、メッセージの先頭のステータスバイト(MSBが1、128以上のデータ)を省略することができる。これは多量の演奏情報の伝送効率のためによく使われているが、受け取り側でこの解釈処理をすることが要請されている。本研究の実験データではランニングステータスを使っていないので、この例外処理もなくデータの処理は簡単である。

時間情報を見ると、時間分解能としては1msec単位のデータとなっていることが判る。ただし、実際に計測処理の精度が本当に1msec分解能であるかどうか、というのはこのデータとは別の問題であり(この部分を誤解した心理学研究が少なくないのは残念である)、これは第4章のレイテンシのところ

検討する。

3-5-3. 3次元可視化パッチ

Max/MSPはそれ自体の段階で、QuickTimeムービーを単純に再生する機能を持っていたが、Jitterへと発展したことで、そのグラフィクス機能は飛躍的に向上した。その大きな特徴の一つに、Open-GLの標準対応があり、これまでUnix環境でテキストベースでプログラミングしていたOpen-GLによるリアルタイム3次元CGが、MaxのGUIで利用できるようになった。以下、簡単なパッチの解説(詳細は文献[71]を参照のこと)により、「3次元可視化プログラム」の処理アルゴリズムを紹介する。

メインパッチ

3次元可視化プログラムのメインパッチの実行モードでは、「read」オブジェクトにより、実験データを選択して読み込んで表示処理を行う。また、多数並んだカラーパレットオブジェクトにより、以下のそれぞれのカラーを任意に設定できる。

- スクリーンの背景色
- 映像ビートのドットの色
- 被験者のタッピングのドットの色
- 被験者のタッピングのドットからX軸までの垂線の色
- 音楽ビート(アクセント)のドットの色
- 音楽ビートのドットからX軸までの垂線の色
- 音楽裏ビート(弱拍)のドットの色
- 音楽裏ビート(弱拍)のドットからX軸までの垂線の色

3次元可視化プログラムのメインパッチの編集モードでは、通常は実行モードの画面内にある変数オブジェクトと反転スイッチ オブジェクトによって、以下のパラメータや設定を任意に切り替えてデータの3次元表示を制御できる。

- フルスクリーンモードのON/OFF
- 座標軸の表示ON/OFF
- 3次元カメラの焦点距離
- X, Y, Z軸それぞれのズーム
- X, Y, Z軸それぞれの方向への平行移動
- 視点(カメラ)からX, Y, Z軸それぞれの方向への回転

環境設定サブパッチ

メインパッチ中にある「初期化」サブパッチでは、スクリーン内を初期化としてブラックに塗りつぶし、データをプロットするためのグレーの直線をX軸と平行に描画している。また、Jitterの基本機能として、50msecごとにOpen-GLのレンダリング処理を行っている。

データ処理関係パッチ

メインパッチ中にある、「count1」サブパッチ内で「read」コマンドで「Text」オブジェクトに読み込んだ実験データは、大きく2段階のパスとして処理する。これは被験者ごとにデータ数が同一ではないため、まずこの「count1」では、データを次々に読み込んでその総数を獲得するとともに、初期状態の画面でX軸方向に0.0から3.0までの範囲で全データを3次元表示するために、3.0をこのデータ数で割ることで、次の段階で表示するためのX軸方向の増分値を得る。

メインパッチ中にある、「count2」サブパッチは「count1」の処理が終わると続けて行われる処理であり、「count1」で確定したデータ総数に対して、同じく計算されたX軸方向の増分値に対して、それぞれの描画サブパッチに実際のタイミングデータを分配する。ここで得られた値によって、JitterがOpen-GLの描画タイミングごとに新たに3次元表示データを計算して描画するので、その合間にリアルタイムにマウスによって視点を変更したりズームングや平行移動によって、自在にデータを可視化できることになる。

描画関係パッチ

メインパッチ中にある、「映像ビート」描画のサブパッチでは、この実験においては、常に「映像ビート」を基準として、初期化パッチでX軸と平行に描画した直線上に、等間隔でプロットする。ここは可視化において重要であるが、表示としては映像素材のテンポを固定して音楽素材のテンポを微小変化させた「実験2」で「映像ビート」が等間隔になるのは当然である。

しかし、音楽素材のテンポを固定して映像素材のテンポを微小変化させた「実験3」においても、物理的には伸び縮みしている「映像ビート」を等間隔となるように表示上は時間軸を局部的に伸長させて、相対的に音楽ビートとタッピングのビートとの時間差を表示することで、実験としては正反対のものを同一の視点から可視化することに成功した。

メインパッチ中にある、「タッピング」描画のサブパッチでは、上記のように「実験2」「実験3」それぞれで換算されて等間隔に描画されている映像ビートのタイミングに対して、まずはタッピングのタイミングの時間から相対時間を計算して、これをY軸方向の大きさとして描画している。さらに、前後でもっとも至近な映像ビートとの相対時間(ビートのずれ)を比較して、映像ビートよりも早い場合にはZ軸方向として下(マイナス)方向、映像ビートよりも遅い場合には上(プラス)方向となるので、これが見やすいように、それぞれの該当する時間地点(X軸上)まで細い垂線を描画するようにした。

メインパッチ中にある、「音楽ビート」描画のサブパッチでは、音楽ビートには、ノリのあるビートの部分と、裏拍でビートのない部分とがあるが、ここでは2種類のパッチによって、それぞれ色と大きさの異なるドットで表示して、音楽がテンポの変化とともに映像ビートと絡む様子、そしてタッピングのデータと絡む様子を可視化した。

4. システムのレイテンシの検討

ここでは、音楽情報科学の視点から、コンピュータを用いた音楽心理学/知覚認知心理学の実験で重要となる、実験システムのレイテンシ(遅延)について詳細に検討・考察する。図15は、本研究の心理学実験システムの全体図を、レイテンシの検討という視点からさらに詳細に示した図である。

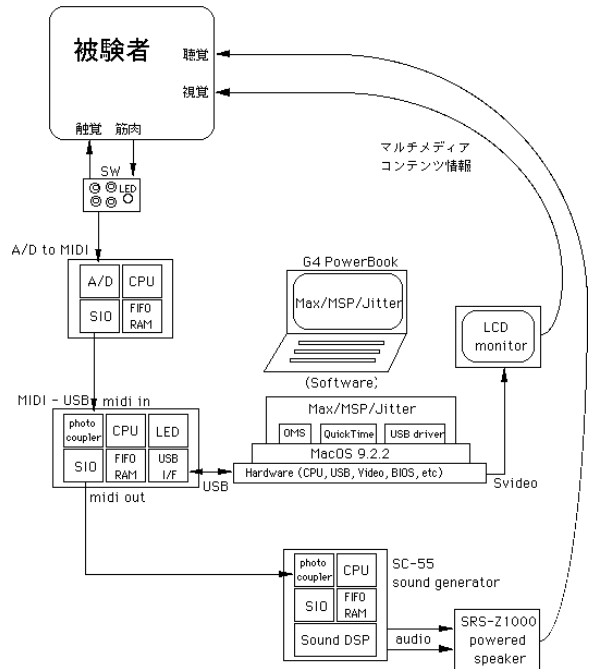


図15 心理学実験システムの詳細全体図

4-1. 心理学実験におけるMIDI/USB利用の注意点

音楽関係のコンピュータシステムではMIDIは1985年から定番のインターフェースであり、さらに最近では多くのコンピュータの外部インターフェースとしてUSBが用いられている。本研究の心理学実験でも、パソコンの外部インターフェースとしてはUSBしかないので、「USB-MIDIインターフェース」装置を介して情報交換を行っている。

ところで、パソコン自体もそうであるが、MIDI機器も、さらに「USB-MIDIインターフェース」装置も、内部にコンピュータ(CPU)というソフトウェア機構を持ったハードウェアであり、ここには多段の複雑なソフトウェア処理がイモズル式に連なっていることになる。この事実と問題点について検討する。

4-1-1. デジタル情報の流れ

あらためて上記の心理学実験システムの詳細全体図を眺めてみると、本研究の心理学実験における情報の流れにおいて、実験中のリアルタイム情報は簡単には以下の5つのCPUによるソフトウェア処理を経由していることが判る。これは最終的な出力としては音楽情報のルートであるが、映像情報のルートについても、パソコンからの映像出力信号を実際の「光」に変換するLCDディスプレイ内部に、CPUではないがデジタル信号処理システムがあって画像情報として出力されているので、少なくとも4つのCPU処理は経由していることになる [27][56]。

A/D to MIDIインターフェース

スイッチ操作による電圧変化をMIDIに変換するシステム。内部のCPUのソフトウェアによる処理であり、本質的に時間幅のある遅延を伴う。スイッチの変化からMIDI情報が送出されるまでの処理は以下のステップである。なお、このシステムは筆者が開発したオリジナルであるので、記述は確定している事実である。

1. タイマにより定期的に8チャンネルのアナログ電圧をスキャンしてA/D変換する
2. A/Dデータの前回データと比較して変化があると、データをMIDIフォーマットに変換して内部FIFOバッファに積む
3. メインルーチンを巡回中にFIFOバッファに変化があると、そのデータをシリアルポートからMIDIとして出力する

USB-MIDIインターフェース

ソフトウェアが、MIDI情報をパソコン側の専用ドライバソフトを介してやりとりする装置。MIDI受信はハンドシェイクやポーリングなど「暇な時に受信する」というシステムではデータ落ちが生じて、これは「音がONされてOFFを受け落とし鳴り続ける」というような致命的なトラブルとなるので、割り込みとFIFOバッファリングという手法で「なにがなんでもとりあえず受信する」というのが一般的である。ここにはCPUのソフトウェアを用いるのが一般的であり、ローランド UM-2 においても内部CPUが以下のような処理を行っていると考えられる。なお、メーカーからは詳細な技術情報は公開されていないので、これはあくまで推定である。

1. MIDI受信データはフォトカプラからCPUのシリアルポートに入力され、割り込みによってこれをFIFOバッファに積む
2. FIFOバッファに積まれたデータはMIDIメッセージとして意味を持つかが解釈され、メッセージとして完結しないなど無意味なデータは捨てられる
3. MIDIメッセージとして意味を持つと解釈されたメッセージはUSBバスにパケットとして送られる

Macintoshコンピュータ

上記の心理学実験システムの詳細全体図の中央のコンピュータの処理の部分は最小限に簡略化されているが、およそ以下のような処理を行っている。ここでは、心理学実験として肝心のMax/MSP/Jitterでのソフトウェア処理の部分は「(スイッチの)情報が入ったら(サウンドの)情報を出す」と簡略化しておく。なお、メーカー等からは詳細な内部技術情報は完全には公開されていないので、これはあくまで推定である。

1. USB-MIDIインターフェースUM-2から、このデバイスからのリクエストとしてUSBバスに情報が送られる
2. コンピュータのUSBデバイスはBIOSを経由してOSのUSB処理ルーチンをコール(ソフトウェア割り込み)
3. Max/MSP/Jitterの起動時にOMSが呼び出され、このOMSはUM-2の専用USBドライバを読み込んでいたために、OS USBドライバ OMS Maxという流れでこのMIDI入力情報をMax/MSP/Jitterが受け取る
4. (心理学実験システム内の処理)
5. Max/MSP/JitterからのMIDI情報出力は、上記と逆の流れで、Max OMS USBドライバ OS USBハードウェア、と経由してUSBバスからUM-2に送られる

USB-MIDIインターフェース

1. パソコンからのMIDIメッセージはUSBバスからUM-2に与えられる
2. CPUはUSBパケットを解析してMIDI送信すべきMIDIメッセージを抽出する
3. 送信すべきMIDIメッセージはステータスバイトとデータバイトそれぞれにシリアルポートに送られ、オープンドレイン/オープンコレクタのドライバを経由して電流出力される

MIDI音源SC-55

MIDI情報を受信して、GM/GS規約に従った音色の音源装置としてMIDI情報をサウンド電気信号に変換する一種の電子楽器。過去の研究[27][56]からは、MIDI受信は単にFIFOバッファリングで受信するだけでなく、時間的にほぼ連続して受信した発音イベントを「和音」として同時出力するために、最初のMIDI受信からすぐに発音処理を行わずに意図的な遅延を加えている報告があった。なお、メーカーからは詳細な技術情報は公開されていないので、これはあくまで推定である。

1. MIDI受信データはフォトカプラからCPUのシリアルポートに入力され、割り込みによってこれをFIFOバッファに積む
2. FIFOバッファに積まれたデータはMIDIメッセージとして意味を持つかが解釈され、メッセージとして完結しないなど無意味なデータは捨てられる
3. MIDIメッセージとして意味を持つと解釈されたメッセージは内部バッファに蓄積される
4. 入力された発音イベントの情報のうちMIDIチャンネル情報によって、そのMIDIチャンネルに既に設定されている(あるいは電源ON時/デフォルト時の初期値)音色番号が呼び出され、その音色番号に対応した楽音合成パラメータが音源LSI (DSP)にセットされる
5. 入力された発音イベントの情報のうちMIDIチャンネル情報とノートナンバ(音高)情報とベロシティ(音量とON/OFF)情報によって、過去に同一の発音イベントとの関係を検索し、さらにこの発音のための音源チャンネルの空き状態を検索し、「音源チャンネルが空いていて過去に同じ音は発音中ではない」場合には発音処理を開始する(アサイン処理)
6. 上記で過去に同一の発音が既にあったり音源チャンネルが同時発音数の限度で満杯の場合には、決められたルールによりしかるべき対応処理を行う(無視したり、どこか別の音を強制消音して音源チャンネルを空けたり、等)
7. 発音の場合には、音色パラメータ、ピッチ情報などを設定した上で、ここにベロシティ情報に対応した音量情報を乗算し、さらにそのMIDIチャンネルに設定されていたボリューム情報とエクスプレッション情報とを乗算して最終的な発音の音量情報とし、さらにそのMIDIチャンネルに設定されていたパンポット情報に応じてステレオの左右チャンネルにそれぞれ乗算してステレオの両チャンネルからのサウンド情報に累算する
8. 音源部は、上記のような処理を同時発音数の全ての音源チャンネルについて累算して、D/Aコンバータの左右それぞれの出力情報とする
9. 出力段のD/Aコンバータは、このデジタル電気信号を44.1kHzサンプリング、16ビットリニア、2チャンネルとしてアナログ電気信号に変換してライン出力する

ごく簡略しても以上のような情報処理の流れを経て、本研究の心理学実験での「スイッチを押す」「サウンドが鳴る」という動作が実現されている。従って、この所要時間、その誤差とばらつき、知覚心理学から得られている人間の反応時間や知覚処理時間、というような時間オーダでの比較と検討は心理学実験において本質的に重要である。

4-1-2. MIDI音源の遅延の計測実験の例

筆者の過去の研究において、心理学実験にも使われるようになったMIDI音源(ハードウェア製品、ソフトウェア音源)のMIDI入力から発音までの遅延と発音抜けの状況を調査した [27][56]。ここではその中から、遅延の分析、原因の検討/考察の部分を省略して、実験で得られたMIDI遅延データの概略についてのみ簡単に紹介する(文献[71]を参照)。

MIDIは既に15年以上の歴史を持つComputer Musicにおけるスタンダードであるが、これまで主に楽器メーカーがBlack Boxとしていたこともあり、その「遅延」に関しては多くの誤解が存在する。ここでは、市販されているMIDIシンセサイザやMIDI音源モジュール、さらにパソコン等のソフトウェアによるGM音源ドライバやGM音源モジュールについて、実際にMIDI信号を受けてから発音されるまでの発音の遅延とそのばらつきと確実性、さらにMIDI情報のトラフィックに対する振舞いについて実験的に検証し、特に音楽心理学のための実験環境として安易に利用されている現状に関して、その問題点を検討した。

分解能が4分音符単位で960とか謳われている市販のシーケンスソフトの時間管理にそのような精度が無いのは常識であり、実際のMIDI信号を処理する部分がブラックボックス内に入っているソフトウェアをこの計測に使用することは無意味である。そこでまず、必要な基準MIDIメッセージのみを正確に送出し、さらにその出力タイミングと厳密に同期したトリガ信号を発生する、計測専用ツールを開発した。

基準MIDIメッセージとして「MIDI1チャンネル、ノートナンバ60、ベロシティ100でON、その直後にOFF」という6バイトのMIDI情報を送出して、その最終バイトの送出と同期してデジタルポートに幅2.5m秒のトリガパルスを生成した。このMIDI信号を対象の音源に直接接続し、GMリセットによりそのサウンド出力に現れるピアノ音の立ち上がり点までの遅延をストレージスコープで20回計測し、平均遅延時間と最大値最小値をプロットした。これを「シングルトリガモード」と呼ぶ。

これとともに、多量のMIDIトラフィックの中からこの基準MIDIメッセージを受信して発音する「トラフィックモード」も設けて併せて計測した。ここでは、シングルトリガモードの情報と対応したトリガ信号を生成しながら、さらにバックグラウンドとして「他の二つのMIDIチャンネルで、ノートナンバ0から127までを順に、ベロシティ1でONして直後にOFFする多量のメッセージ」を流した。送出レートのチューニングにより、実際にMIDIケーブル上を流れるデジタル信号が約70%になるように調整した。これは実際の音楽の場では多数のセンサフュージョン情報をマージしたような非常に過酷な条件であるが、同時発音数は常に1音で情報としてはMIDI規格に従った条件であり、ベロシティが最小音量である多量のメッセージに埋没することなく、基準メッセージのサウンドだけが明確に区別され発音されることが求められる。

この実験で対象とした音源を図16に示す。説明の便宜上、<1>-<24>の記号を割り当てた。ハード音源としては、<1>SC-88STPro、<2>SC-55mark2、

<3>05R/W、<4>GMega、<23>JV30、<22>SCP-55、パソコンのソフト音源としてはMacintosh PowerBook 2400上のQuickTime3.0、WindowsではESSチップを用いたものとしてはFM音源、Microsoft GS Wavetableドライバ、VSC-55、OPLチップを用いたものとしてはOPL3-FM、OPL3-SA、SGMP driver、さらにSGI Indy上でのIRIX6.2のソフトGMシンセサイザ。WindowsパソコンのMIDIインターフェースについては、デスクトップ機がローランドシリアル(SC-88Pro)とSB16、ノート機がローランドシリアルとSCP-55を取り替えてつ使用した。

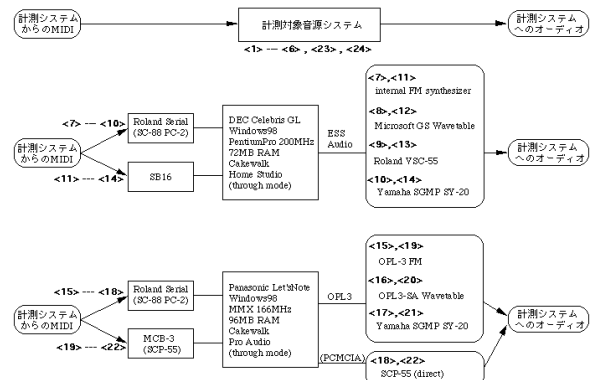


図16 実験で対象とした音源

図17-図20は、図16で測定対象とした24タイプの音源の計測結果である。なお、図の中に発音遅延の最大値と最小値を示しているものの中で、「たまに鳴った」「*回鳴らなかった」というような記述のある音源のデータは、実質的にはあまり意味をなさない事に注意すべきである。バックグラウンドに多量の「ベロシティ=1」のイベントがあるにせよ、個々の瞬間での同時ON数は常に1か2かゼロであり、MIDI情報としては必ず発音すべきであるのに、これはたとえば10回続けて鳴らない、というような現象である。その意味するのは、センサデータのストリームとマージしているような状況であるにせよ、演奏者が鍵盤に向かって渾身の一撃の演奏を行ったのに、これが空を切って音源が鳴らない「ことがある」、という音楽的には致命的な事実であり、私見としては「音源(楽器)としては失格」という印象を持たざるをえない。また、この測定に使用した音源装置やソフトウェアについては、製造ロットやバージョンによって、ユーザの見えないところで仕様が変更されていることも予想できるので、追試に関しては十分に注意する必要がある点を指摘しておく。

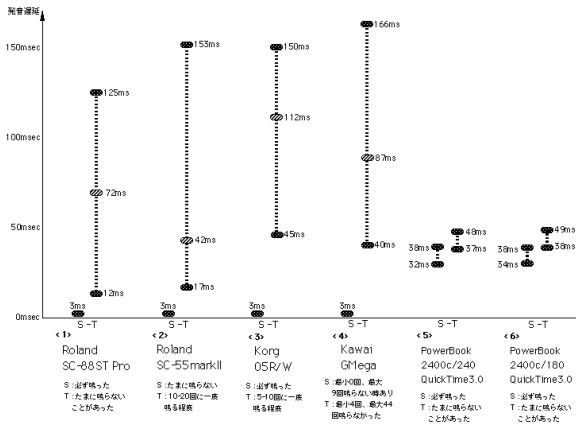


図17 計測結果 (音源<1>-<6>)

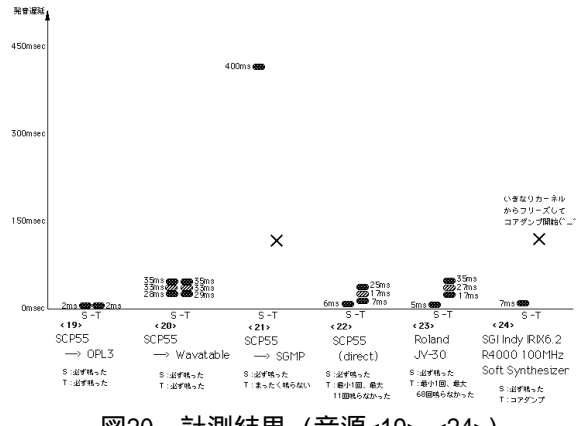


図20 計測結果 (音源<19>-<24>)

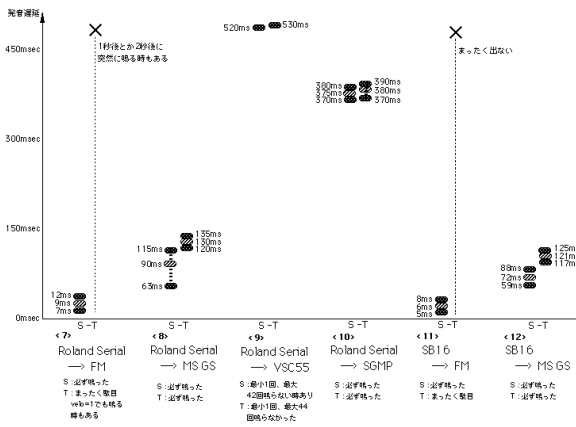


図18 計測結果 (音源<7>-<12>)

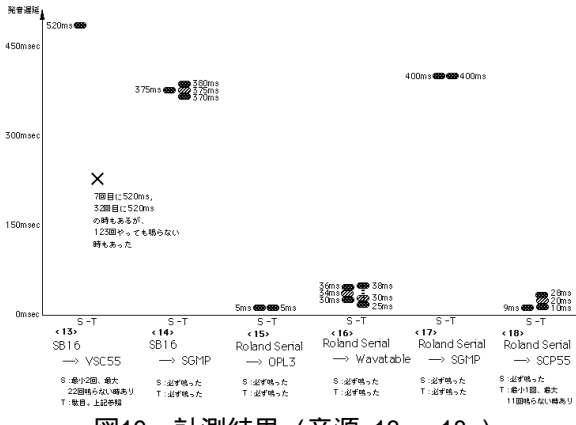


図19 計測結果 (音源<13>-<18>)

上記の実験結果のような特性を持つMIDI音源を音楽心理学の実験の道具として使用した場合、人間の聴覚と音楽心理学的な影響から、発音遅延やばらつきによって、実験の意味が消滅する危険性がある。つまり、もともと正確な「同時性」の概念の消滅しているMIDIの世界において、さらに実験データをブラックボックスのシーケンサとMIDI音源に委ねた場合には、音楽心理学的に意味を持つ(感性評価として有意な差がある)オーダの時間的誤差がある、というのは、実験の意味を喪失させることになりかねない。パソコンとシーケンサソフトとMIDI音源(楽器メーカーによるハード的な製品であっても、パソコンの内蔵音源チップであっても、ソフトシンセなどソフト音源であっても)、というお手頃な道具の普及は、音楽心理学実験をあまりコンピュータの得意でない人々にまで広げる貢献をしているものの、ブラックボックス故の上記の問題点によって研究の意義を脅かすとなれば大きな問題であろう。

北村・山田らの研究によれば、人間が「同時に鳴る」3和音を聴取した際に、5msecとか10msecとかのオーダの発音遅延を特定の音に作用させただけで、和音全体の感覚的な印象が異なる、とされている[67]。もし、ここに発音遅延の絶対値が20msecのオーダでばらつく音源を使用した場合、シーケンサのソフト上では「同時」と思いこんでいる実験データが、実験の意味を失わせるオーダで効いていることは明白であるが、シーケンサと市販音源を安易に使用している研究者は、この事実気付かない、という可能性がある。カラオケのBGMを自動再生している場合にはまったく問題ないことであるが、音楽心理学の実験にはそうそう気楽に使ってはいけな、という事実を認識すべきである。なお文献[67]の実験には、当然のことながらMIDIシーケンサ等は使用せず、正確なタイミングの和音の音響波形を編集して、DATで再生して聴取実験している。

4-1-3. レイテンシを許容できる実験条件

以上のようなシステム遅延の状況から、音楽心理学においてコンピュータシステム(ソフトウェアを含む)を利用する場合には、以下のようなガイドライン例を十分に検討する必要がある。

1. 実験における情報の流れをステップとして明確に区分する

2. ステップごとの遅延時間の絶対値を明確にする
3. ステップごとの遅延時間のばらつき範囲を明確にする
4. ばらつきが絶対値より大きい場合には時間の加算の意味がなくなる
5. 個々のステップの平均遅延時間を合計することが無意味にならないか検討する
6. 人間の知覚心理学から得られている時間と比較検討する
7. 知覚時間より十分に小さい遅れや誤差は許容できる
8. 知覚時間と同等、あるいはそれ以上のオーダの遅延やばらつきがあれば実験に使用できない

4-2. レイテンシ計測の実験とデータ

ここでは、心理学実験システムの詳細全体図のうち、どの部分を計測したデータであるか、を明確にしなが、11種類のレイテンシ計測を行った結果を詳細に報告し、次の検討の資料とするとともに追試のための条件を全て明確にする。

4-2-1. レイテンシ計測実験1

この実験では、図21のような情報の流れを計測した。図22はこのレイテンシ計測実験1の計測データの一例である。グラフ上段のトリガ信号としては、被験者のタッピング操作のスイッチの両端にかかる電圧そのものを計測して、プルアップしている電圧がGNDに落ちた瞬間を時間計測の起点とした。

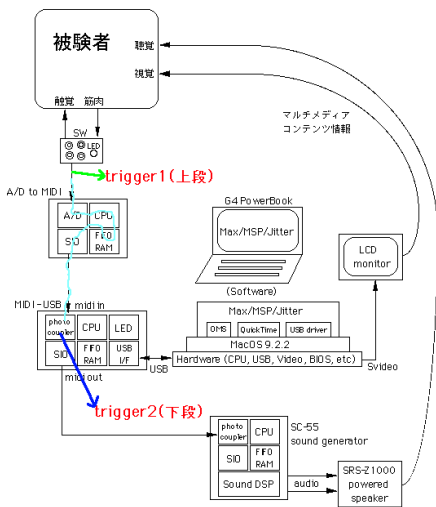


図21 レイテンシ計測実験1の情報の流れ

グラフ下段の計測信号としては、USB-MIDIインターフェースUM-2の内部のMIDI受信フォトカプラの出力信号とした。このフォトカプラ(Sharp PC410)の変換時間は、Typicalで50nsecと高速なので、MIDI受信に対しての出力はほぼ同時と考えてよい。

結局、この実験では、「被験者がスイッチを押してからUM-2にMIDI情報が届くまでの遅延」が計測できる。情報の遅延要因としては、汎用A/Dインターフェース内部のCPUによるA/D変換時間、MIDI情報化してFIFOバッファに積まれる時間、31.25KbpsのシリアルI/OであるMIDIにより伝送される伝送遅延の合計となる。

グラフ下段では2ブロックのデータ(2バイト)が見

られるが、これはスイッチONの情報を2バイトメッセージである「チャンネルプレッシャー」として送出しているからであり、最初の1バイトがステータス、続くバイトがデータである。

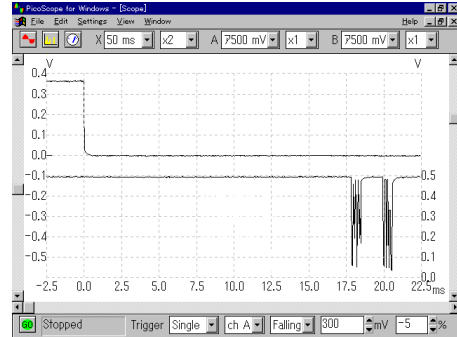


図22 レイテンシ計測実験1の計測データの一例

4-2-2. レイテンシ計測実験2

この実験では、図23のような情報の流れを計測した。図24はこのレイテンシ計測実験2の計測データの一例である。

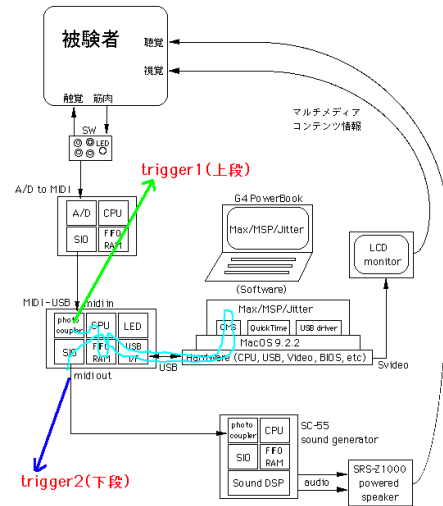


図23 レイテンシ計測実験2/6/7/8の情報の流れ

グラフ上段のトリガ信号としては、実験1の下段の情報である、USB-MIDIインターフェースUM-2の内部のMIDI受信フォトカプラの出力信号とした。

グラフ下段の計測信号としては、UM-2のMIDI出力オープンドレインゲートの入力信号とした。ゲートの遅延は5nsecから10nsecなので、これはUM-2からのMIDI送信と同時に考えてよい。

UM-2からG4PowerBookコンピュータのMax/MSPに入った情報は、「midiin」オブジェクトから直結された「midiout」オブジェクトによりそのままスルー出力されてUM-2に戻り、MIDI出力される。

結局、この実験では、「UM-2にMIDI情報が届いてからMax/MSP/Jitterを経由してUM-2から送信されるまでの最短時間の遅延」が計測できる。情報の遅延

要因としては、UM-2内部のCPUによるMIDI受信時間、USB情報化してMacintoshとやりとりする時間、USBドライバからOMSを経由してMax/MSP/Jitterと行き来する時間、USBを経由してUM-2からMIDIにより伝送される伝送遅延の合計となる。

グラフ上段では2ブロックのデータ(2バイト)が離れて見られるが、これは汎用A/Dインターフェースの方でMIDI出力にもFIFOバッファを介してデータを分散させているからである。これに対して、下段のUM-2の出力は2バイトが密着して一気に送信されている。

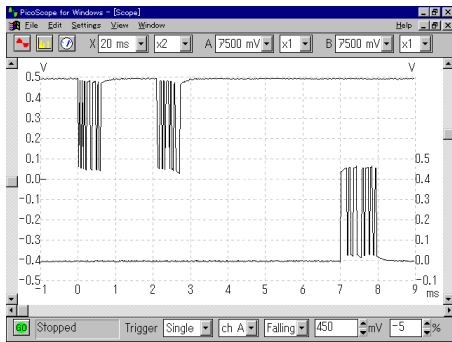


図24 レイテンシ計測実験2の計測データの一例

計測データとしては、上段の2バイトメッセージの先頭にトリガをかけてここからの時間としているが、上述の処理という意味では、上段の後のデータが届いたところでMIDIメッセージとして完結する(例えば後半のデータが欠落していれば無視しなければならない)ので、このメッセージの最後を受けて初めて処理がスタートすることに注意しなければならない。

4-2-3. レイテンシ計測実験3

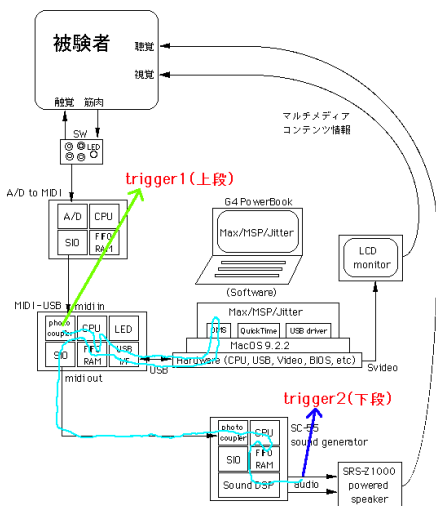


図25 レイテンシ計測実験3/4/5の情報の流れ
この実験では、図25のような情報の流れを計測した。図26はこのレイテンシ計測実験3の計測データ

の一例である。グラフ上段のトリガ信号としては、実験1の下段の情報である、USB-MIDIインターフェースUM-2の内部のMIDI受信フォトカプラの出力信号とした。

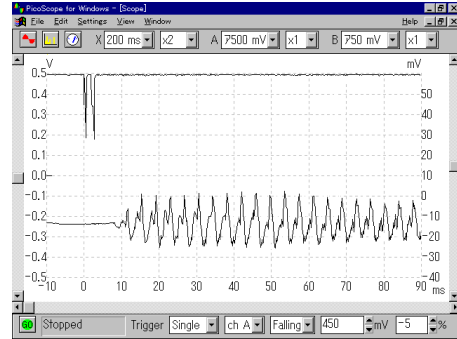


図26 レイテンシ計測実験3の計測データの一例

グラフ下段の計測信号としては、ピアノ音色を設定したGM音源SC-55mkIIのオーディオ出力端子とした。本当はSC-55のMIDI入力をトリガとしたかったが、ピアノ音などノートイベントについては、アクティブセンシング機能によって、UM-2から自動的に300msec以内ごとにメッセージ「0xFE」が送信されてしまい、オシロスコープのトリガを邪魔するので、スイッチ操作によりピアノ音を出す、という実験2の遅延時間が加わった条件として考えることにした。

UM-2からG4PowerBookコンピュータのMax/MSPに入った情報は、「midiin」オブジェクトから直結されてノートナンバ=60とした「makenote」-「noteout」オブジェクトにより出力されてUM-2に戻り、MIDI出力される。

結局、この実験では、「UM-2にMIDI情報が届いてからMax/MSP/Jitterを経由してUM-2から送信される、これを受信してGM音源SC-55mkIIでサウンドが発音されるまでの最短時間の遅延」が計測できる。情報の遅延要因としては、実験2の遅延に加えて、SC-55内部のCPUによるMIDI受信時間、SC-55内部の音源DSPシステムの楽音合成/発音処理の演算時間、SC-55からD/Aコンバータでアナログのオーディオ信号化する変換時間の合計となる。

グラフ上段では2ブロックのデータ、下段ではアナログのピアノ音の立ち上がりが見られる。

4-2-4. レイテンシ計測実験4

この実験では、計測実験3と同じく、図25のような情報の流れを計測した。

図27はこのレイテンシ計測実験4の計測データの一例である。計測実験3と同様に、グラフ上段のトリガ信号としては、実験1の下段の情報である、USB-MIDIインターフェースUM-2の内部のMIDI受信フォトカプラの出力信号とした。グラフ下段の計測信号としては、ピアノ音色を設定したGM音源SC-55mkIIのオーディオ出力端子とした。

違いとして、心理学実験2と3のようなBGMのある状況を想定して、計測に使うピアノ音とは別に、

MIDI2チャンネル2のベース音とMIDI10チャンネルのドラムパートを連続して再生し、この2つのチャンネルボリュームを最小の「1」とすることでほとんど聞こえない状態で鳴らしている。

結局、この実験では、計測実験3のシステムレスポンス(入力イベントがあったら最速で反応して発音する)に対して、BGMという余分なジョブが加わっている場合にどうなるか、という差分を見ていることになる。

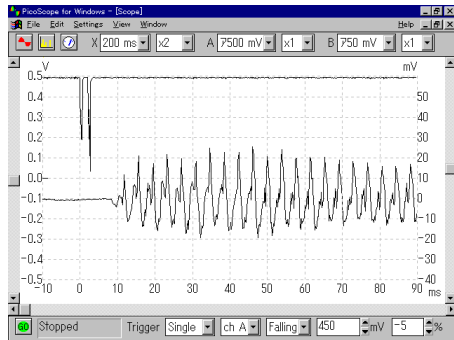


図27 レイテンシ計測実験4の計測データの一例

4-2-5. レイテンシ計測実験5

この実験では、計測実験2と同じく、図23のような情報の流れを計測した。

計測実験2と同様に、トリガ信号としては、実験1の下段の情報である、USB-MIDIインターフェースUM-2の内部のMIDI受信フォトカプラの出力信号とした。計測信号としては、UM-2のMIDI出力オープンドレインゲートの入力信号とした。

違いとして、Jitterを使って、320*240ドットのQuickTimeムービーを、S-Videoを経由して出力したサブスクリーンに最大化(640*480ドットのVGA)して同時に描画処理させている。

結局、この実験では、計測実験2のシステムレスポンスに対して、さらにリアルタイムのムービー再生出力、という余分なジョブが加わっている場合にどうなるか、という差分を見ていることになる。

4-2-6. レイテンシ計測実験6

この実験では、計測実験2と同じく、図23のような情報の流れを計測した。

図28はこのレイテンシ計測実験6の計測データの一例である。計測実験5と同様に、グラフ上段のトリガ信号としては、実験1の下段の情報である、USB-MIDIインターフェースUM-2の内部のMIDI受信フォトカプラの出力信号とした。グラフ下段の計測信号としては、UM-2のMIDI出力オープンドレインゲートの入力信号とした。

違いとして、計測実験5と同様にQuickTimeムービーのサブスクリーンでのVGA連続描画を行った上で、「midiin」から「midiout」に直結していたところに、msec単位の情報遅延オブジェクトである「pipe」を使って、意図的にMax/MSP/Jitter内での処理として「pipe 1」という「1msecの遅延」を明示的に加えた。

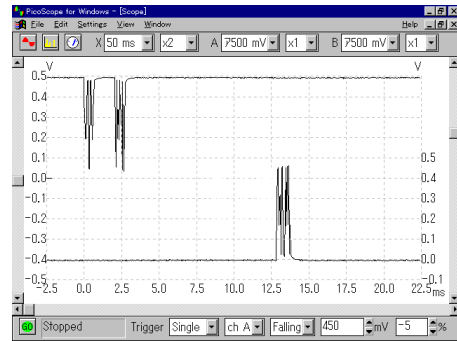


図28 レイテンシ計測実験6の計測データの一例

つまりこの実験は、計測実験5と比較した場合、ムービー再生を行っている中で、スペックと同様に「1msecだけ遅延が長い」結果が出たとすれば、Max/MSP/Jitterは相当に正確な時間管理をしているということになる。

4-2-7. レイテンシ計測実験7

この実験では、計測実験2と同じく、図23のような情報の流れを計測した。

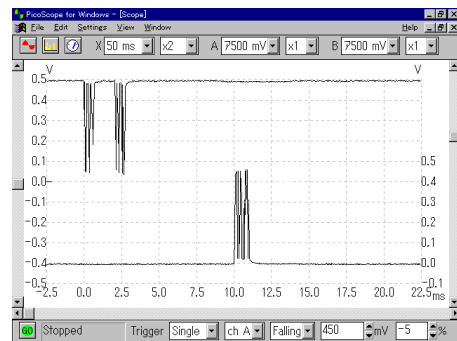


図29 レイテンシ計測実験7の計測データの一例

図29はこのレイテンシ計測実験7の計測データの一例である。計測実験5/6と同様に、グラフ上段のトリガ信号としては、実験1の下段の情報である、USB-MIDIインターフェースUM-2の内部のMIDI受信フォトカプラの出力信号とした。グラフ下段の計測信号としては、UM-2のMIDI出力オープンドレインゲートの入力信号とした。

違いとして、計測実験5/6と同様にQuickTimeムービーのサブスクリーンでのVGA連続描画を行った上で、「midiin」から「midiout」に直結していたところに、msec単位の情報遅延オブジェクトである「pipe」を使って、意図的にMax/MSP/Jitter内での処理として「pipe 5」という「5msecの遅延」を明示的に加えた。

つまりこの実験は、計測実験6と比較した場合、スペックと同様に「5msecだけ遅延が長い」結果が出たとすれば、Max/MSP/Jitterは相当に正確な時間管理をしているということになる。

4-2-8. レイテンシ計測実験8

この実験では、計測実験2と同じく、図23のような情報の流れを計測した。

図30はこのレイテンシ計測実験8の計測データの一例である。計測実験5/6/7と同様に、グラフ上段のトリガ信号としては、実験1の下段の情報である、USB-MIDIインターフェースUM-2の内部のMIDI受信フォトカプラの出力信号とした。グラフ下段の計測信号としては、UM-2のMIDI出力オープンドレインゲートの入力信号とした。

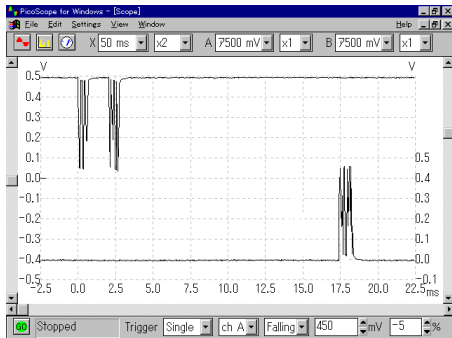


図30 レイテンシ計測実験8の計測データの一例

違いとして、計測実験5/6/7と同様にQuickTimeムービーのサブスクリーンでのVGA連続描画を行った上で、「midiin」から「midiout」に直結していたところに、msec単位の情報遅延オブジェクトである「pipe」を使って、意図的にMax/MSP/Jitter内での処理として「pipe 10」という「10msecの遅延」を明示的に加えた。

つまりこの実験は、計測実験6/7と比較した場合、スペックと同様に「10msecだけ遅延が長い」結果が出たとすれば、Max/MSP/Jitterは相当に正確な時間管理をしているということになる。

4-2-9. レイテンシ計測実験9

この実験では、図31のような情報の流れを計測した。図32はこのレイテンシ計測実験9の計測データの一例である。グラフ上段のトリガ信号としては、実験1の下段の情報である、USB-MIDIインターフェースUM-2の内部のMIDI受信フォトカプラの出力信号とした。

グラフ下段の計測信号としては、G4PowerBookのS-Video出力信号を接続したシャープ 液晶モニター LC-10RV2-Aの画面上に密着して貼り付けたフォトセンサ(Sharp 1S474)の電圧出力を計測した。Max/MSPの処理としては、MIDI入力があると、それまで初期化によりブラックだった画面内に明るい部分を持つQuickTimeムービーを出力する、というものであり、ムービーの中でほぼ真っ白という高輝度の部分にセンサを貼り付けた。LCDから約1メートル離れた被験者にはこの光の変化が遅れて届くが、光速がこの距離を伝わる遅延時間は約3億分の1秒なので、ほぼ無視できる。

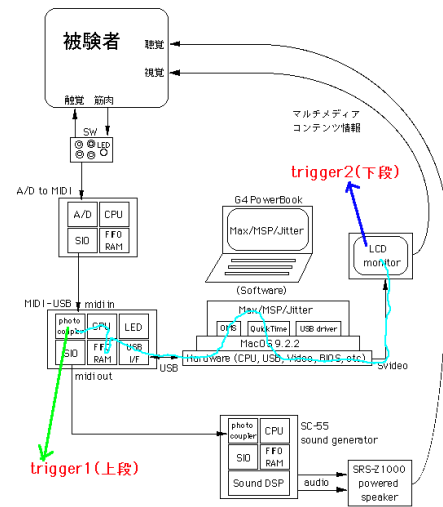


図31 レイテンシ計測実験9の情報の流れ

この実験は、MIDI入力に対して「なるべく即時にグラフィクスを出す」という実験条件を、実際に被験者が対峙しているLCDモニターによって再現しているものとなる。多くの心理学実験において、被験者に提示する視覚的刺激的提示の時間的遅延は、情報を送るパソコンの内部情報として定義しているが、VDTにしるCRTにしるプロジェクタにしるLCDにしる、いずれも情報機器として内部で信号処理/演算処理を行っているための遅延があり、さらに液晶の反応時間はかなり長いと言われている。本計測実験のような条件測定を行わずに視覚刺激の時間を定義した場合には、大きな問題となる可能性があるために設定した計測実験である。

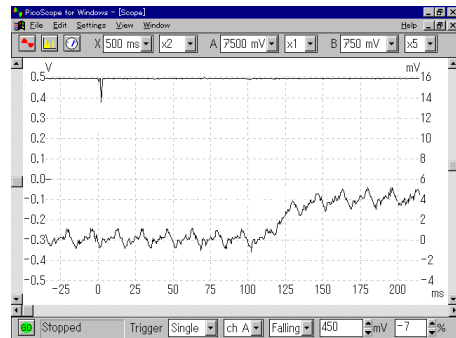


図32 レイテンシ計測実験9の計測データの一例

4-2-10. レイテンシ計測実験10

この実験では、図33のような情報の流れを計測した。図34はこのレイテンシ計測実験10の計測データの一例である。グラフ上段のトリガ信号としては、実験1の下段の情報である、USB-MIDIインターフェースUM-2の内部のMIDI受信フォトカプラの出力信号とした。

グラフ下段の計測信号としては、USB-MIDIインターフェースUM-2のパネルにある動作表示LEDのうち、「MIDI送信LED(green)」に密着して貼り付けた

フォトセンサ(Sharp1S474)の電圧出力を計測した。
UM-2からG4PowerBookコンピュータのMax/MSP/Jitterに入った情報は、「midiin」オブジェクトから直結された「midiout」オブジェクトによりそのままスルー出力されてUM-2に戻り、MIDI出力される。この実験は、CPU内蔵機器であるUM-2の内部動作遅延を検証する目的で行った。

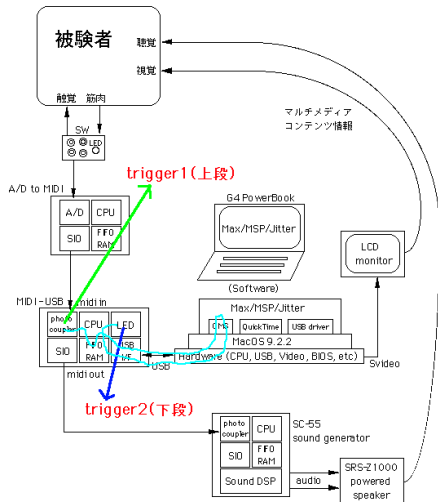


図33 レイテンシ計測実験10の情報の流れ

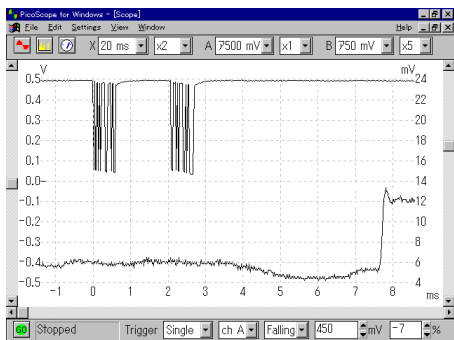


図34 レイテンシ計測実験10の計測データの一部

4-2-11. レイテンシ計測実験11

この実験では、図35のような情報の流れを計測した。図36はこのレイテンシ計測実験11の計測データの一部である。グラフ上段のトリガ信号としては、実験1の下段の情報である、USB-MIDIインターフェースUM-2の内部のMIDI受信フォトカブラの出力信号とした。

グラフ下段の計測信号としては、USB-MIDIインターフェースUM-2のパネルにある動作表示LEDのうち、「MIDI受信LED(red)」に密着して貼り付けたフォトセンサ(Sharp 1S474)の電圧出力を計測した。一般に、MIDIインターフェースの「MIDI受信モニタ」LEDは、MIDIラインそのものを表示するとはほんの一瞬で見えにくいために、CPUのソフトウェアで一定時間のタイマで点灯させることが多い。また、MIDIを受信した段階でこのLEDを点灯するシステム

であれば、計測実験10のようにコンピュータに行き帰ってくる遅延が不要となるので、より早く反応する筈である。この実験は、UM-2の場合にどうかを検証する目的で行った。

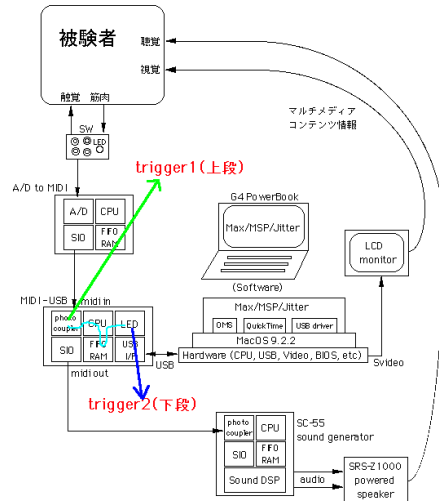


図35 レイテンシ計測実験11の情報の流れ

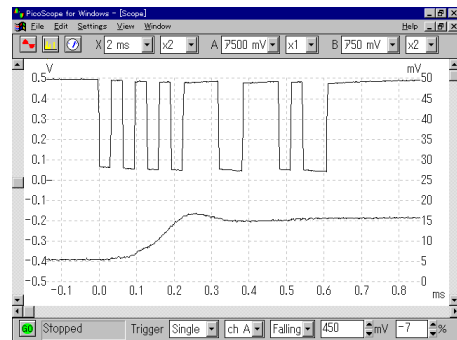


図36 レイテンシ計測実験11の計測データの一部

4-3. レイテンシ計測結果の検討

以上の11種類の遅延計測実験はそれぞれ20個のデータを取り、平均や標準偏差を求めた。図37がその実験結果データである。

4-3-1. 知覚心理学から得られている人間の感覚の時間データ

視覚的情報、聴覚的情報の知覚の時間的定量化などに関して、これまでに知られている知覚心理学的な報告を文献から整理してみると、以下のようになる。

聴覚刺激に対する最小反応時間 minimum reaction timeは100-150msecである [3]。

聴覚の不連続の閾値は10msec、視覚の不連続の閾値は100msecである [9]。

光刺激を瞬間と意識する視覚の瞬間域は60-100msecで、聴覚や触覚の瞬間域は10-50msecである [13]。

断続する視覚的の刺激が融合して連続と感ずる視覚系臨

界融合周波数critical flicker frequencyは、低輝度の場合には2-3Hzであるが、高輝度になると60Hz近くに達する [15]。
 視覚システムによって光パターンが保持される時間は150-200msecである [7]。
 静止画などを注視する平均的な時間は200-300msecである [19]。
 視覚的錯覚現象である「オッペル・クントの錯視」は聴覚の絡んだ時間間隔の知覚においても報告されており、ある時間的インターバルの中に、それを分節化する要素が含まれている場合には、含まれない場合よりも時間的インターバルは長く感じられる [14]。
 標準画像の解像度を1/2-1/4-1/8と劣化させても、コントラストを劣化させても、画面サイズを1/4-1/9-1/16と小さくしても、画像に対する認知度はほとんど変化しなかった [19]。

て2桁(20倍程度)ということで、この遅延の影響は無視できると判断できる。

計測実験2の結果検討

計測実験2で得られた遅延は、システムのプラットフォームとなったMax/MSP/JitterとMIDIとの行き来、という基本的なものであり、ある意味で本研究の根幹に関わるものである。計測実験1の誤差を共通するものとして切り離して、以降の計測実験では全てこの実験2と同様に、UM-2のMIDI入力のリニア出力を基準とした。
 その結果であるが、平均5msecで誤差も小さくプラスマイナス2msec以内、という予想外に良好な成績であった。これはMax/MSPでの「midiin」から「midiout」への直結ということで、MIDIメッセージとして完結するかどうかという評価なしにそのままスルーしているとしても、非常に優秀であると言える。この結果は、十分に注意してシステムを設計すれば、心理学実験のツールとして活用できる可能性を示していると言える。

実験データ(実測値): 単位 = msec

	実験1	実験2	実験3	実験4	実験5	実験6	実験7	実験8	実験9	実験10	実験11
data 1	18.0	6.0	12.0	10.0	3.5	9.0	13.0	16.0	7.5	5.0	0.15
2	11.0	6.0	10.0	9.0	4.0	11.0	8.0	13.0	30	4.0	0.15
3	10.0	4.5	10.0	11.0	6.0	8.0	10.0	14.0	35	5.0	0.15
4	8.0	4.0	8.0	8.0	4.5	12.0	11.0	18.0	30	3.0	0.15
5	14.0	7.0	8.0	10.0	5.5	12.0	10.0	13.0	40	7.0	0.15
6	10.0	6.0	9.0	9.0	6.0	8.0	12.0	15.0	55	5.0	0.15
7	3.0	3.5	11.0	8.0	7.0	12.0	11.0	13.0	35	4.5	0.15
8	16.0	5.0	8.0	11.0	4.0	8.0	9.0	15.0	30	5.5	0.15
9	13.0	4.0	10.0	10.0	5.0	13.0	12.0	16.0	50	7.5	0.15
10	18.0	5.5	8.0	11.0	4.0	10.0	9.0	15.0	40	5.5	0.15
11	6.0	5.5	11.0	11.0	7.0	11.0	10.0	17.0	45	7.0	0.15
12	10.0	3.0	11.0	8.0	4.0	13.0	8.0	17.0	25	7.5	0.15
13	7.0	5.0	9.0	8.0	8.0	11.0	9.0	16.0	100	6.5	0.15
14	5.0	5.0	12.0	10.0	7.0	12.0	9.0	15.0	40	3.5	0.15
15	16.0	8.0	10.0	12.0	4.0	12.0	12.0	14.0	25	7.0	0.15
16	13.0	4.0	9.0	9.0	4.5	9.0	12.0	17.0	120	7.0	0.15
17	6.0	5.5	10.0	11.0	6.5	12.0	8.0	13.0	40	8.0	0.15
18	9.0	5.0	10.0	9.0	5.0	8.0	10.0	13.0	70	5.0	0.15
19	15.0	4.5	11.0	9.0	8.0	9.0	12.0	14.0	25	7.0	0.15
data 20	7.0	7.5	9.0	10.0	7.0	11.0	11.0	16.0	35	3.5	0.15
平均	10.8	5.1	9.8	9.7	5.5	10.6	10.3	15.0	47.2	5.7	0.15
最大値	18.0	7.5	12.0	12.0	8.0	13.0	13.0	18.0	120	8.0	0.15
最小値	3.0	3.0	8.0	8.0	3.5	8.0	8.0	13.0	25	3.0	0.15
標準偏差	4.459	1.115	1.381	1.218	1.464	1.761	1.569	1.589	25.62	1.625	0.00

図37 レイテンシ計測実験の結果データ

4-3-2. レイテンシ計測実験の結果検討

計測実験1の結果検討

計測実験1で得られた遅延は、他の全ての実験結果に加算される性質のデータである。ここでは被験者のスイッチ操作の情報をUSB-MIDIインターフェイスに送るまでの遅延であり、平均11msecほどの遅延時間とともに、標準偏差4.46と比較的大きなばらつきがあって、20回の実験の最大値が18msec、最小値が3msecであった。

このばらつきの原因は明確であり、汎用A/DインターフェイスのA/D変換とMIDI処理におけるバッファリング方式により、たまたまサンプリングとFIFOバッファ処理のループのどこに位置したか、という偶然性により、最悪値としてはループの長さの遅延誤差が生じることになる。

心理学実験の時間誤差としては、プラスマイナス7-8msecということになるが、これは上記の人間の知覚心理学的時間データに比べると1桁ほど小さいものであり、特に本研究においては、さらに1桁ほど大きな視覚的刺激に対する被験者インターフェイスとして利用しているスイッチなので、オーダとし

計測実験3の結果検討

計測実験3で得られた遅延は、実験2のMIDI情報をさらにGM音源SC-55mkIIに入れて発音させる、という楽音合成処理の遅延が加わったものであり、簡単に言えば計測実験3の時間から計測実験2の時間を引き算すれば、それが音源SC-55の発音遅延である、という事になる。

その結果であるが、平均10msecで誤差も小さくプラスマイナス2msec以内、実験2のデータを差し引くと「ほぼ確実に5msecで発音」という、これまた良好な成績であった。過去の遅延計測実験[27][56]でも約3msecという結果が出ていたが、この実験では、Max/MSP内での処理として「makenote」が加わっていることを勘案すると、この情報の流れは十分に実験として小さな誤差で実現できていることが確認できた。

計測実験4の結果検討

ここから計測実験8までは、初めて遅延計測を行うこととなった、ソフトウェアMax/MSP/Jitterの内部処理というブラックボックスに対する、一種の「いじわるテスト」である。しかし、このような評価検討を行わずにソフトウェアを過信して利用するというのは、定量的な時間計測を伴う心理学実験としては致命的でもあり、本研究での報告が今後の心理学研究に警鐘を鳴らすことができれば幸いである。

この計測実験4では、単音でピアノ音を鳴らした計測実験3のバックグラウンド処理として、本研究の心理学実験2/3で用いたものと同じ、ベースとドラムによるBGM自動生成を走らせてみた。ノイズとなるトリガを起こさないようにMIDIボリュームメッセージを最小データの「1」としているものの、MIDI処理は普通に行われるので、過去の非力なパソコンではかなりの差が出た実験である。

しかし驚くべきことに、結果は計測実験3との違

いが無い、というものであった。BGMのイベント密度はそれほど高いものではなかったが、msecオーダのMIDIイベント処理に対して、新しいパソコンのソフトウェア処理は十分に余裕がある、という状況を立証する結果となった。

計測実験5の結果検討

この計測実験5では、Max/MSP/JitterでMIDI入力をそのままMIDI出力する、という計測実験2のバックグラウンド処理として、本研究の心理学実験1/2/3で用いたものと同じ、QuickTimeムービーをJitterにより連続再生してみた。

その結果、平均遅延は5.1msecに対して5.5msec、最大値と最小値もほぼ同様ということで、ほとんど映像処理の遅延の影響を感じさせないという良好な結果となった。これは従来のQuickTimeを使ったMax/MSPに比べて体感としても改善された印象があるが、Jitterの開発者であるKit Clayton氏がチューニングを図ったJitterの処理能力を賞讃すべきであると考えられる。この結果は、視覚的刺激を提供する心理学実験システムとしての有効性も示している。

計測実験6の結果検討

この計測実験6では、バックグラウンドでQuickTimeムービーをJitterにより連続再生した計測実験5と同様の条件において、平均遅延5msecで誤差も2msec以内という良好な成績であった計測実験2の「midiin」と「midiout」の間に、msec単位の情報遅延オブジェクトである「pipe」を使って、意図的に「pipe 1」という「1msecの遅延」を明示的に加えたものである。つまりこの実験は、計測実験5と比較して、スペックと同様に「1msecだけ遅延が長い」結果が出たとすれば、Max/MSP/Jitterは相当に正確な時間管理をしているということになる。

この結果は逆の意味であるが、また予想外のものであった。1msecの遅延を単純に入れただけなのに、平均遅延は10msecを越えて、1msecどころか5msecほど余計にかかった処理となっている結果であった。これはMax/MSP/Jitterのスケジューリングとして、MSP/Jitterのオーディオ系は44.1kHzのサンプリングであるのに対して、Maxのデータは最小分解能として1msecが設定できるのみであり、MIDIイベントの時間管理としての「msec」というのはあまり意味がない、という事実を明示した結果となった。

ただし、本研究の心理学実験における遅延と誤差のオーダは十分に知覚心理学で得られた範囲より小さいので、実験としての誤差については問題ないと考えられる。

計測実験7の結果検討

この実験は上記の計測実験6のやや意外な結果を受けて追加したものである。この計測実験7では、バックグラウンドでQuickTimeムービーをJitterにより連続再生した計測実験5と同様の条件において、「midiin」と「midiout」の間に、意図的に「pipe 5」という「5msecの遅延」を明示的に加え

たものである。

その結果は、これまた意外なことであるが、計測実験6とほぼ同様、精密に比較すればむしろやや良好(平均遅延がわずかに少なく、さらにばらつきがわずかに小さい)、という結果となった。この原因としては、最小限度である計測実験6の「pipe 1」においては、ソフトウェアのスケジューリングにやや歪んだバランスが悪影響を与えるが、ここでの「pipe 5」は全体としてはスケジューリングのバランスが良くなる、というソフトウェア科学から想定される動作が考えられる。

この5msecという遅延はBGM再生処理や映像再生処理と同等のオーダであるので、心理学実験ツールとしては、システム内部に5msec程度の遅延要因が平均的に存在する、という前提でシステム設計に当たるべきであるのかもしれない。

計測実験8の結果検討

この実験は上記の計測実験6/7の結果を受けて追加したものである。この計測実験8では、バックグラウンドでQuickTimeムービーをJitterにより連続再生した計測実験5と同様の条件において、「midiin」と「midiout」の間に、意図的に「pipe 10」という「10msecの遅延」を明示的に加えたものである。

その結果は、計測実験5の結果に綺麗に10msecを加算したようなデータであった。これを受けて計測実験7の結果を見てみると、「pipe 5」によって、計測実験5の結果に綺麗に5msecを加算したようなデータであると見てとれる。つまり、スペック的に元々無理である「pipe 1」という計測実験6を無視して、「pipe 5」と「pipe 10」を入れた結果を見てみると、かなり正確に5msecと10msecの時間加算が成立していたのである。

この結果を教訓とするなら、心理学実験ツールとしてMax/MSP/Jitterを使う場合には、MIDI処理の時間単位としては、1msecや2msecという指定は無意味であるが、5-10msec以上についてはシステムのスケジューリングのマージンに入ること、かなり正確にタイミングが実現されている、という事のようにある。

計測実験9の結果検討

この実験以降では、オシロスコープのプロープで回路信号を検出するのではなく、光速光センサを用いて、視覚刺激である「光」そのもののタイミングを計測した。具体的には、Max/MSP/Jitterの走るG4PowerBookのS-Video出力信号を接続したシャープ液晶モニタLC-10RV2-Aの画面上に密着して貼り付けたフォトセンサ(Sharp 1S474)の電圧出力である。

Max/MSP/Jitterの処理としては、トリガとなるMIDI入力があると、それまで初期化によりブラックだった画面内に明るい部分を持つQuickTimeムービーを出力する、というものであり、ムービーの中でほぼ真っ白という高輝度の部分にセンサを貼り付けた。

その結果は、平均遅延47msec、最大遅延

120msec、最小遅延25msec、という、今回の計測実験において他とオーダが違うような大きなものとなった。これは、知覚心理学的にも反応時間の短い聴覚に比べて時間オーダの大きな視覚であっても、実験によっては重要な問題となるレベルであると言える。特に、多くの心理学実験において、被験者に提示する視覚的刺激的の提示の時間的遅延を、情報を送るパソコンの内部情報(時計)として定義しているが、VDTにしるCRTにしるプロジェクタにしるLCDにしる、いずれも情報機器として内部で信号処理/演算処理を行っているためにこのような遅延があるとすれば重大な問題である。本計測実験のような条件測定を行わずに視覚刺激の時間を定義した場合には、大きな問題となる可能性があることを指摘しておきたい。

本研究における影響については、「映像ビート」のスレシールド時間として50-100msecという時間設定があるために、なんとかデータのオーダと誤差のオーダが拮抗する程度に収まるので、平均的データとして(相対的に)実験結果を評価する範囲においては問題はないと考えられるが、個々の瞬間のデータを絶対的(定量的に)に見ることは難しいと言える。

計測実験10の結果検討

この実験は、計測実験2と比較して検討する性格のものである。情報としてはMIDI入力を受けていたMacintosh内のMax/MSP/Jitterまで行き、そのままソフトウェアスルーでMIDI出力をしている、そのMIDI出力状態をUM-2のパネルあるグリーンLEDで表示している。このLEDに光速光センサ(Sharp 1S474)を密着させて計測したデータである。

結果は、計測実験2とほとんど同等のレスポンスとなり、予想したようにUM-2内部のCPUの処理として、MIDI情報を実際に出力している部分でグリーンLEDを点灯させ、その後はソフトウェアタイマによってしばらく連続点灯させている事が分かった。

計測実験11の結果検討

この実験は、計測実験10と比較して検討する性格のものである。送信モニタと同様に、「MIDI受信モニタ」LEDは、MIDIラインそのものを表示するとほんの一瞬見えにくいために、CPUのソフトウェアで一定時間のタイマで点灯させることが多い。この赤色LEDに光速光センサ(Sharp 1S474)を密着させて計測したデータである。MIDIを受信した段階でこのLEDを点灯するシステムであれば、計測実験10のようにコンピュータに行き帰ってくる遅延が不要となるので、より早く反応する筈である。

結果は、他の計測実験とはオーダの異なる高速な対応であった。20回の実験でまったくばらつきのない、150 μ secほどのレスポンスで、実際のMIDIメッセージが完結していないうちから点灯を開始し、さらにソフトウェアタイマによって連続点灯している。これはUM-2内部の処理として、MIDIのシリアル割込みルーチン内でMIDI受信のデータFIFOバッファリングと同時に、とりあえず「MIDIが来た」というモニタ赤色LEDを点灯させていると推定される。

4-3-3. 心理学実験としての評価検討

以上のレイテンシ計測実験により、本研究の心理学実験システムとしての遅延とばらつきの状況が明確に定量化できた。以前に行った実験に比べるとコンピュータの性能向上によりシステムのレスポンスは良好になっており、さらにSC-55を用いたサウンド系の遅延については心理実験として問題ないオーダの誤差とばらつきである事を確認できた。

唯一の問題として注意すべき遅延は、視覚的刺激である映像情報が、被験者に提示した液晶モニタの画面上の輝度変化という最終的な状態までで、計測実験9にあったような数十から100msecオーダの遅延時間を持った。これは本研究の心理学実験においては、「映像ビート」の平均的位置のばらつきが同様のオーダであるために問題はないが、単一の映像刺激を瞬間的に提示してこれに対する反応時間を計測する、等の知覚心理学実験を行うとすれば、その絶対値が問題となって実験の意味を失う可能性がある事も示した。

5. 被験者テストの3つのモード

ここでは、2003年3月-4月に、静岡文化芸術大学1106研究室にて行った、本研究の心理学実験の具体的な内容について詳細に報告解説する。概要としては以下のようなものである。

研究室中央の机に向かい、背もたれのある椅子に楽に腰掛け
てもらう
被験者は目の前のLCDモニタを見て、BGMとして流れる音楽を
聞き、スイッチを操作する
LCDモニタは被験者の正面およそ50cm。視線は水平よりやや
下向き
その手前にスイッチ基板があり、被験者は「どれも同じ情報
が出るので、好きなスイッチを押して下さい」「押し
方や持ち方も自由です」という指示により、思い思い
の方法でスイッチ操作を行う
パワードスピーカーから被験者までの距離は約2メートル、左
右スピーカーの間隔は約1メートル、スピーカーの高さは
被験者の視線から上方へ約50cm
被験者の位置における実験室のバックグラウンドノイズの音
圧レベルは45.5db、「実験1」におけるヒット位置を
示すピアノ音の鳴った時の音圧は56.6db、「実験1」
においてヒットした時のピアノ音の音圧は62.7db、
「実験2」「実験3」においてBGMが鳴っている状態の
音圧は61.9db
被験者はリハーサル、実験1、実験2、実験3を連続して行
う。所要時間は全部で約10分間程度
その後アンケートを無記名で記入して提出

5-1. 被験者とアンケート調査

被験者は、静岡県在住の43人であり、全員が視覚および聴覚およびスイッチ操作の運動能力において正常である。女性が34人、男性が9人であり、年齢は18歳から40歳までであるが18-22歳が大多数である。職業は学生および事務職員であるが学生が大多数である。各自の音楽経験、運動経験、テレビゲーム等の嗜好、ダンスなど音楽と運動の組み合わせへの嗜好などはアンケートにて回答を集めたが、これは実験終了直後に行ったので被験者の実験に対する予備知識とはならなかった。ここでは個別のアンケート回答については触れず、第9章の実験結果の考

察のところで、必要に応じて個人差の解釈などに利用することにした。

5-2. リハーサル

実際の心理学実験(「実験1」「実験2」「実験3」という3種類からなる本実験モード)そのものに入る前に、あらかじめ被験者がシステムや行動に慣れるための「リハーサルモード」を用意した。リハーサルモードにおいては、映像ビートを提示する映像素材としては本番の実験1/2/3とまったく同じものを液晶ディスプレイから提示した。被験者にすれば、いきなり突然に「お地蔵さんからお地蔵さんにモーフィング」という奇妙な映像に直面するので、このリハーサルにおいて心理的違和感を解消して慣れるための位置付けも大きいと考える。

この映像素材と同時に、本番の実験ではまったく用いないサウンドとして、システムの映像ビートと同期したタイミングでピアノ音「中央C」(MIDIノートナンバ=60)をスタカートで鳴らした。これは、「スイッチを押すべき映像ビートというのはここ」という実験の指示を明確に示すためのもので、被験者は4つのうちの自分に合ったスイッチを探しながら、まずは等間隔のこのタイミングでスイッチを押す。そして、映像ビートとスイッチONが一定のスレシールド時間帯の範囲内であれば、さらに和音で「G」(67)と「Bb」(70)の2音も鳴るようにした。これにより、被験者はスイッチを押して和音が鳴る(ヒットした)か単音だけである(外した)か、を容易に理解できる。

これを30秒ほど続けたところで実験者が必ず被験者に指摘したのは、「今は映像の静止点の部分でピアノの音がしますが、これは本番の実験では鳴りません。あくまで、目で見えた映像の静止点でスイッチを押して下さい」という事である。大部分の被験者は、等間隔に聞こえるピアノ音と同時にスイッチを押す傾向にあったが、この指示により視覚に注意を向けることになった。

また、一部の被験者は映像の制止点を予測してヒットするようにスイッチ操作を行うことが難航した。これには視覚的知覚と運動神経的なものと両方の原因が考えられるが、この場合には実験者(筆者)が、映像ビートに対して前後にどれだけの時間幅のスイッチ操作までをヒットと看做すか、という2つのパラメータを(裏で)マニュアル操作した。

ヒットの許容幅を被験者によって変えることについては、本研究における実験の主眼は「実験2」と「実験3」にあり、ここではこの調整したパラメータはまったく使用していない(「ヒット」という概念は実験2/3には存在せず、単にスイッチ操作のタイミングを記録する)。リハーサルおよび「実験1」というのは、いろいろな被験者を同じような心理的条件とスイッチ操作条件でメインの実験に誘導するためのものでしかないので、本質的な問題は起きないと思う。むしろ、なかなかヒット音が出ないために心理的に焦ったり、自分の運動神経や反射神経に不安を持ち始める被験者にとって、実験者がこのパラメータを調整してヒットするようになって、結果として余計な精神的圧迫を解消して実験に進ん

だ、という意味で、この機能は有効であった。

5-3. 実験1

上記のリハーサルモードを受けて、本実験モードの最初の実験1として以下のような内容を設定した。まず、「映像の停止点の部分で等間隔で鳴っていたピアノの音はもう鳴りません。あくまで、目で見えた映像の静止点でスイッチを押して下さい。静止点をつまみ捕まえると和音が鳴ります。慌てて押し直したりしないで、毎回押すようにして下さい。50回成功すると終わります」と告げた。

実験1においては、映像ビートを提示する映像素材としては、リハーサルモードやメインの実験2/3とまったく同じものを液晶ディスプレイから提示した。そして、映像ビートと被験者のスイッチONが、リハーサル時に設定した一定のスレシールド時間帯の範囲内であれば、和音で「G」(67)と「Bb」(70)の2音が鳴るようにした。これにより、被験者はスイッチを押して和音が鳴る(ヒットした)か鳴らない(外した)か、を容易に理解できる。

ヒットした回数は記録され、50回のヒットで実験は終了し(映像がブラックに変わる)、このデータを実験者(筆者)がテキストデータとしてファイル名を付けて保存する10数秒間が被験者の軽い休憩時間となった。50回のヒットにより実験1が終了するまでの時間は被験者によって大きなばらつきがあり、早い場合には60-70回程度で終了するが、難航して200回以上の映像ビートに挑戦するケースもあった。映像素材は240枚の静止画しかないが、その最後まで行っても計50回のヒットに達しない、というケースは一人もなく、全員がそれ以内に成功・終了した。

5-4. 実験2

上記の実験1を受けて、メインの実験2として以下のような内容を設定した。まず、「いよいよ本番の実験です。今の実験と同じように目で見えた映像の静止点でスイッチを押して下さい。ただし今度はピアノの和音(ヒット)は鳴らないので、視覚を頼りにして下さい。また、今度はBGMが鳴ってきますので、これをさりげに聞きながら行って下さい。全体で3分近い長さがありますので頑張ってください」と告げた。

実験2においても、映像ビートを提示する映像素材としては、他の実験とまったく同じものを液晶ディスプレイから提示したが、実験3とともにここでは用意した映像素材(ムービー)の全体を走らせた。この実験では「ヒット」の概念はなくなり、映像ビート、スイッチ操作のタイミング、そして新たに音楽ビートが全て記録された。

実験2においては、映像ビートのテンポは、1フレームごとに110msecというペースで完全に一定である。その一方で、音楽ビートのテンポは、次第に速くなり、そこでいったん一定となり、今度は次第に遅くなり、そこでいったん一定となり、再び第に速くなり、・・・という動きを繰り返した。心理学実験システムのメインパッチにあるように、初期値は110msec間隔の音楽ビートが、

実験開始後2秒の時点から25秒かけて105msec間隔までリニアに加速
 実験開始後30秒の時点から25秒かけて114msec間隔までリニアに減速
 実験開始後60秒の時点から25秒かけて105msec間隔までリニアに加速
 実験開始後90秒の時点から25秒かけて114msec間隔までリニアに減速
 実験開始後120秒の時点から25秒かけて105msec間隔までリニアに加速
 実験開始後150秒の時点から25秒かけて114msec間隔までリニアに減速

という変化をもった。被験者が(意識したにせよしないにせよ)この音楽ビートの速度変化のパターンを経験するのはこれ一度きりであり、パターンを覚えるという点は考慮する必要がないので、この変化は全ての被験者に同一の条件とした。

5-5. 実験3

上記の実験2を受けて、引き続きメインの実験3として以下のような内容を設定した。まず、「あと1回、今の実験とほぼ同様のことを行います。全体で3分近い長さがありますが、これで最後ですので頑張って下さい」と告げた。

実験3においても、映像ビートを提示する映像素材としては、他の実験とまったく同じものを液晶ディスプレイから提示し、実験2とともに用意した映像素材(ムービー)の全体を走らせた。この実験でも「ヒット」の概念は不要であり、映像ビート、スイッチ操作のタイミング、そして音楽ビートが全て記録された。

実験3においては、実験2とは反対に、音楽ビートのテンポが1フレームごとに110msecというペースで完全に一定である。その一方で、映像ビートのテンポは、次第に速くなり、そこでいったん一定となり、今度は次第に遅くなり、そこでいったん一定となり、再び第に速くなり、・・・という動きを繰り返した。心理学実験システムのメインパッチにあるように、初期値は110msec間隔の映像ビートが、

実験開始後2秒の時点から25秒かけて105msec間隔までリニアに加速
 実験開始後30秒の時点から25秒かけて114msec間隔までリニアに減速
 実験開始後60秒の時点から25秒かけて105msec間隔までリニアに加速
 実験開始後90秒の時点から25秒かけて114msec間隔までリニアに減速
 実験開始後120秒の時点から25秒かけて105msec間隔までリニアに加速
 実験開始後150秒の時点から25秒かけて114msec間隔までリニアに減速

という変化をもった。この違いに気付いた被験者もいれば、まったく気付かない被験者もいた。

6. 実験結果

実験結果について、本研究で提唱した引き込み効果の検討を含めた分析は第9章で検討することにして、ここでは得られたデータそのもの一部を、実際の数値データ、全体を描画した静止画データと

して紹介する。3次元可視化されたデータの見方については、第7章で詳しく解説している。全てのデータ、さらにムービーとして見やすくしたデータは全て文献[71]にあるので、参照されたい。

6-1. 実験で得られた生データの一部

ここではまず最初に、文献[71]の全データを分析する追試のために、実験で得られたデータのごく一部の実例をそのまま紹介する。データは43人の被験者ごとに、以下の6つである。

実験1で記録されたテキストデータ(図38)
 実験2で記録されたテキストデータ
 実験3で記録されたテキストデータ
 実験1の結果データ全体の3次元可視化データ(図39)
 実験2の結果データ全体の3次元可視化データ(図40)
 実験3の結果データ全体の3次元可視化データ(図41)

```
0 192 0;
545 193 0;
770 192 1;
1285 193 1;
1650 192 0;
2285 193 0;
2420 192 1;
3255 193 1;
3300 192 0;
4040 193 0;
4180 192 1;
4930 193 1;
5060 192 0;
5830 192 1;
5830 193 0;
6515 193 1;
6710 192 0;
7370 193 0;
7590 192 1;
8070 193 1;
8470 192 0;
9155 193 0;
9240 192 1;
9960 193 1;
10120 192 0;
10915 193 0;
11000 192 1;
11730 193 1;
11880 192 0;
12520 193 0;
12650 192 1;
13505 193 1;
13530 192 0;
14290 193 0;
14410 192 1;
15145 193 1;
15290 192 0;
15995 193 0;
16060 192 1;
16800 193 1;
```

図38 被験者2の実験1の結果テキストデータ

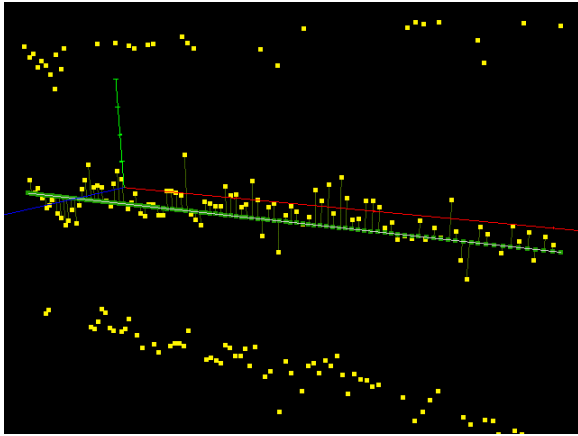


図39 被験者3の実験1の結果の3次元可視化データ

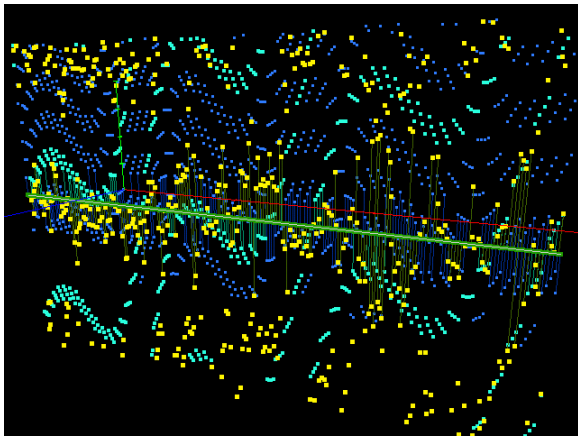


図40 被験者4の実験2の結果の3次元可視化データ

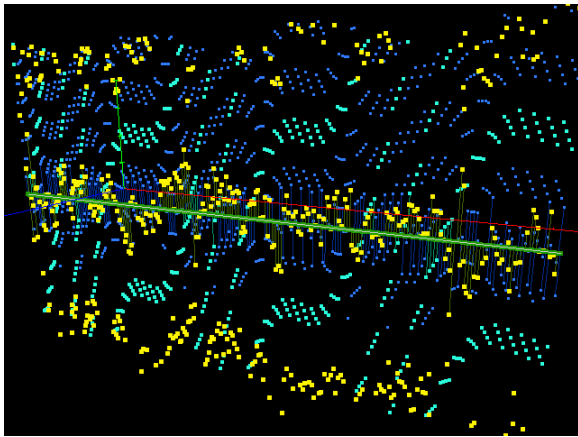


図41 被験者5の実験3の結果の3次元可視化データ

6-2. 3次元可視化データのムービー

本研究の実験システムの一部としてオリジナル開発した、結果データの3次元可視化ツールの画面(図39-図41など)については、この静止画だけを見ると、3次元可視化の効果として、以下のような疑問が提起されうる。すなわち、(1)3次元CG技術の原理からいって、面を描画してこそ3次元特有の立体感

や奥行き感が表現できるのであり、本ツールのような点と線による描画で3次元CGの恩恵が得られるのか疑わしい、(2)むしろ複数の実験結果の重ね合わせ表示は、結果を読みにくくする可能性さえある、(3)図39-図41などの一連の可視化結果は、斜めの視点から表示しているように見えるが、これも2次元的に真正面から表示したほうが読みやすい、等である。

これは本論文のように紙面に固定された静止画だけを眺めているための誤解であり、本システムの3次元可視化ツールの最大の効果は、(1)大量のデータを任意の軸方向にズームできること、(2)3次元空間を任意に回転/平行など視点を移動できること、というリアルタイム性・インタラクティブ性にある。本研究で説明を目指した「引き込み効果」は非常に微妙な時間オーダで起きる現象であり、大量のデータのごく一部にのみ特徴的に出現する。そこで、各被験者のデータを本ツールによってマウスやトラックボールで自在に動的に扱うことで、特徴的な場所を適確に発見し、拡大して現象を検討できる、という意義を指摘しておきたい。

この意味で、文献[71]においては、43人の被験者の実験2および実験3で得られたデータを、上述の3次元可視化ソフトウェアに一部自動パラメータプレイヤーを組み込んで、以下のような流れで全体を俯瞰したムービーを自動生成して公開している。このムービーを見てみれば、上記の「リアルタイム性・インタラクティブ性」と、データから現象を見い出すためのツールの有効性が理解されよう。

1. まず、全体のデータを描画する(この描画が終了した状態が、ここに示した結果データの静止画である)
2. 時間軸(データ軸方向)を拡大して個々のイベントが見えるようにズームアップする
3. 時間軸(データ軸方向)にシフト(平行移動)して、実験の経過をシミュレーションする
4. 実験の最後のところまでいったら元の原点まで戻してズームも戻す
5. データ全体の様子を把握するために視点を3次元空間で回転させる
6. 最後に、いちばん最初の視点に戻す(この状態も、ここに示した結果データの静止画である)

7. 実験データの解析

ここでは、実験結果データに対する解析検討の第一段階として、定性的な結果データの傾向について考察した。データは全て公開しているので、実験の追試、ないしこのデータに対する新しいアプローチでの分析検討にも対応している。本研究では第一ステップとして、実験1によって被験者が大きく2つのグループに分かれた点に注目して、アンケートによる回答とあわせて検討した。

7-1. 実験1の解析

実験1の結果データで予想以上に顕著だったのは、データを一瞥するだけで簡単に指摘できるほど、被験者がかなり明確に2群に分かれたことである。一例として、図42は被験者6の、また図43は被験者7の、実験1の結果の3次元可視化データである。次節において、図42の被験者6は「タッ

ピング初心者グループ」に、図43の被験者7は「タッピング上級者グループ」と分類した、それぞれの典型例である。

図42の被験者6の3次元可視化データは、スイッチのタッピングを示すイエローの矩形が、映像ビートを示すグリーンデータの軸に対して上下のあちこちに広範に分布している。これに対して図43の被験者7の3次元可視化データは、スイッチのタッピングを示すイエローの矩形が、グリーンデータの軸のほとんど近辺にだけ分布している。この実験1は、映像ビートに対してタッピングが「ヒット」するとそのサウンドが鳴り、50回のヒットで実験が終了するので、ほとんど毎回「ヒット」した被験者7のグラフの方が時間的に早く完了しており、データ軸上の映像ビートのグリーン矩形が離散的に見えている。これに対してなかなか「ヒット」しない被験者6の場合には、全体の実験時間として長い時間を必要としているために、画面全体を実験時間と正規化した関係で、データ軸上の映像ビートを示すグリーンの矩形は密接に並んでいる。

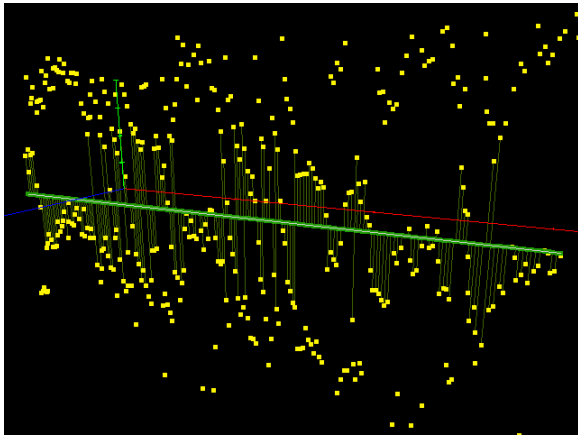


図42 被験者6の実験1の結果の3次元可視化データ

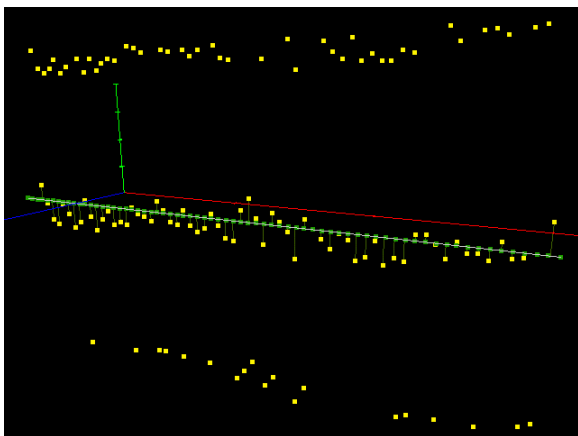


図43 被験者7の実験1の結果の3次元可視化データ

この明確な違いは、被験者の全体にわたって顕著であり、実験1の結果である3次元可視化データを一瞥するだけで、被験者は容易に2群に分類できた。

そこで以下では、この2群それぞれに分類した被験者の実験1の結果の中から無作為に選んだ5人ずつを並べて、あわせてそれぞれの被験者が無記名のアンケートとして回答した内容から特徴的な傾向が見られないか検討した。被験者全員のデータとアンケート回答については、文献[71]で全て公開しているので、アンケートについては本稿では省略する。

ここで、以降の検討のための便宜上の呼び方として、実験1の結果の3次元可視化データとして、スイッチのタッピングを示すイエローの矩形が映像ビートの乗ったデータ軸の近辺に分布しているグループを、「タッピング上級者グループ」と名付けた。この意味するところは、(1)本実験で使用した映像素材に対して、モーフィングの静止点を映像ビートとするという定義に良好に対応してこの映像ビートを正確に捕えていること、(2)その映像ビートに対してスイッチを正確にタッピングすることで高い「ヒット」率を実現していること、の2条件を満たしている、ということである。

また、同様に便宜上の呼び方として、実験1の結果の3次元可視化データとして、スイッチのタッピングを示すイエローの矩形が映像ビートの乗ったデータ軸に対して上下のあちこちに広範に分布しているグループを、「タッピング初心者グループ」と名付けた。この意味するところは、(1)本実験で使用した映像素材に対して、モーフィングの静止点を映像ビートとするという要請にうまく応えられていない(映像ビートにピンとこない)、(2)映像ビートに対してスイッチを正確にタッピングする運動神経の部分が苦手で「ヒット」率が低い、のいずれかの条件が想定される、ということである。

7-1-1. タッピング上級者グループ

図44は被験者9の、図45は被験者10の、図46は被験者12の、図47は被験者14の、図48は被験者16の、実験1の結果の3次元可視化データである。

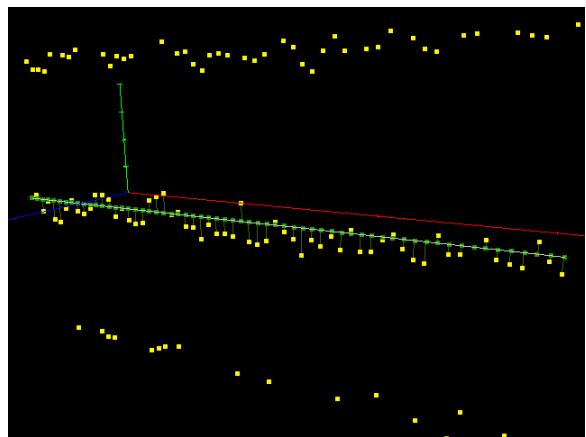


図44 被験者9の実験1の結果の3次元可視化データ

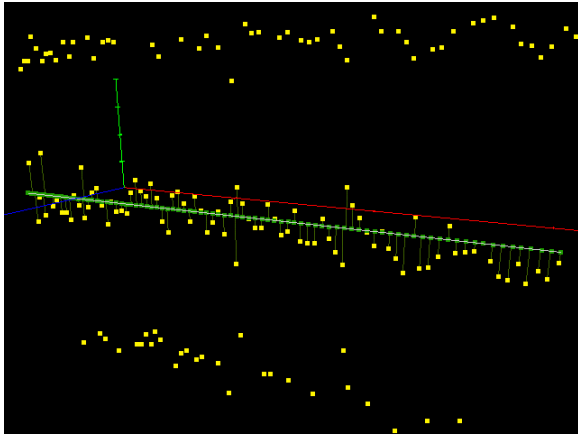


図45 被験者10の実験1の結果の3次元可視化データ

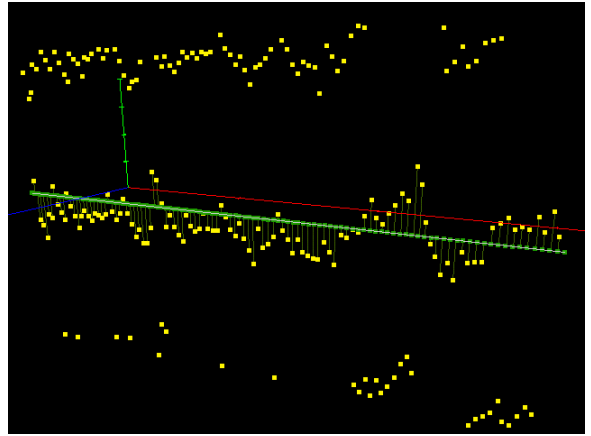


図48 被験者16の実験1の結果の3次元可視化データ

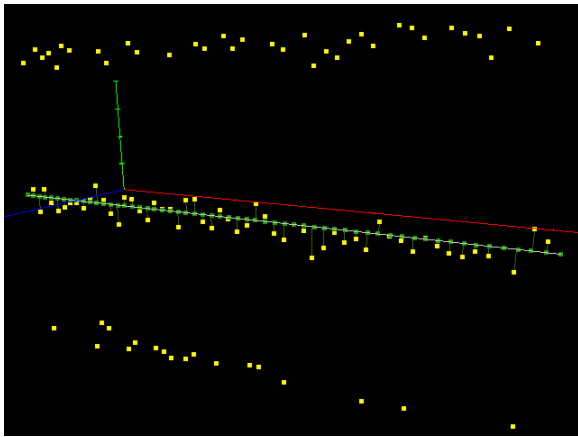


図46 被験者12の実験1の結果の3次元可視化データ

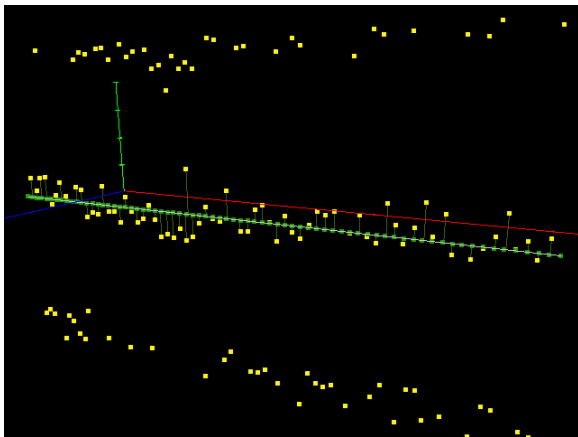


図47 被験者14の実験1の結果の3次元可視化データ

7-1-2. タッピング初心者グループ

図49は被験者8の、図50は被験者11の、図51は被験者13の、図52は被験者15の、図53は被験者19の、実験1の結果の3次元可視化データである。

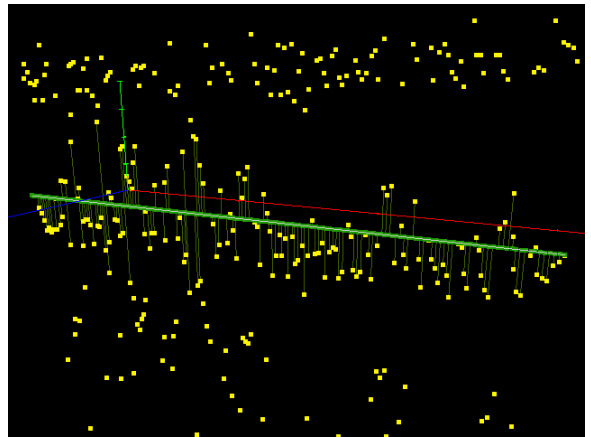


図49 被験者8の実験1の結果の3次元可視化データ

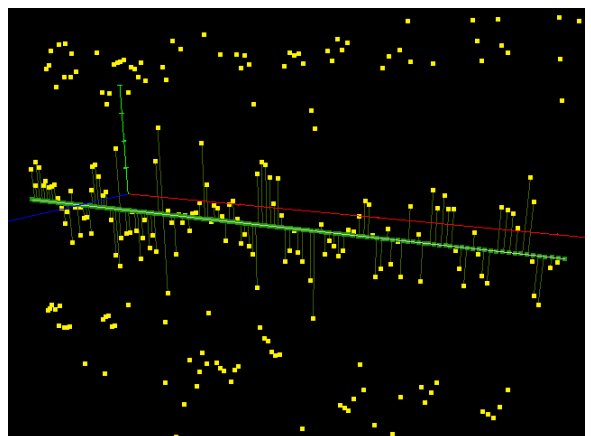


図50 被験者11の実験1の結果の3次元可視化データ

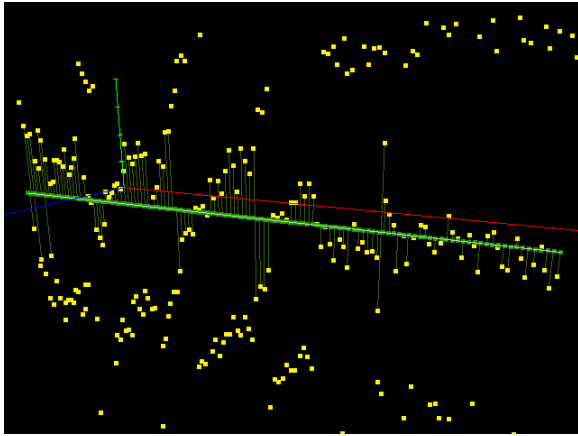


図51 被験者13の実験1の結果の3次元可視化データ

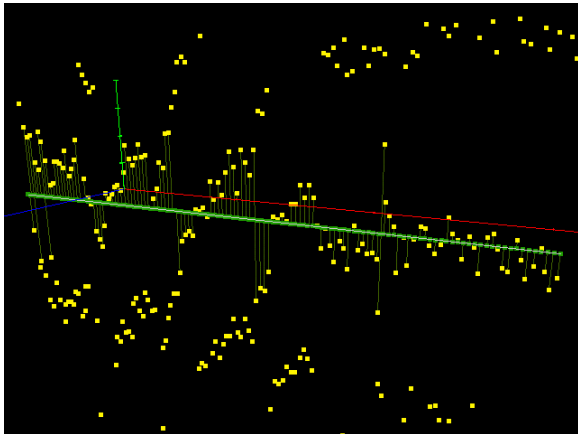


図52 被験者15の実験1の結果の3次元可視化データ

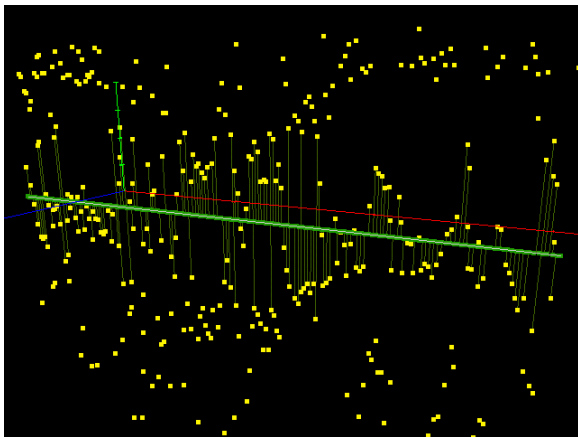


図53 被験者19の実験1の結果の3次元可視化データ

7-1-3. 実験1の結果とアンケート回答

上記のような結果となったが、筆者はこのデータの明確なグループ化と、アンケート結果から得られた傾向との関係性を抽出することに成功していない。「読書や勉強などで集中している時に音楽が聞こえていると集中できない」という質問項目は、もしかすると音楽聴取のチャンネルを無意識に閉じる

ことができる被験者がいて、その場合には音楽ビートによる引き込み効果が出て来ないか少ないのでは、と設定した。しかしこの質問項目も、43人の被験者の中で21人がYES、「タッピング上級者」グループ32人中15人がYES、「タッピング初心者」グループ11人中6人がYES、とほぼ半分であり、有意な傾向は得られなかった。

この他のアンケート項目についても、とりたてて明確なグループ化の理由づけになるようなものは得られていない。この点は、今後の検討を待つことにする。

7-2. 実験2の解析

実験1において「タッピング初心者グループ」に分類された11人については、もともとBGMの無いところで映像ビートをタッピングするだけで、3次元可視化データのタッピング情報が映像ビートの乗ったデータ軸に対して上下のあちこちに広範囲に分布した。この原因はまだ明確になっていないが、(1)本実験で使用した映像素材に対して、モーフィングの静止点を映像ビートとするという要請にうまく応えられていない(映像ビートにピンとこない)、(2)映像ビートに対してスイッチを正確にタッピングする運動神経の部分が苦手な「ヒット」率が低い、のいずれかの条件が原因として有力である。

このようにBGMナシの状況で大きくばらつく被験者のデータであれば、本研究においてより細かい時間的オーダとして注目したい「音楽ビートによる引き込み効果」は、実験結果として十分な精度で得られる可能性が乏しい。そこで本研究では第一段階として、「実験2の解析」としては、実験1の結果

「タッピング上級者グループ」に分類された32人のデータだけを検討することにした。なお32人全てのデータについての検討は文献[71]で公開しているが、ここでは一例として「被験者2」のデータを例として、3次元可視化されたデータの「読み方」のみを紹介した。

図54は被験者2の、実験2の結果の3次元可視化データである。

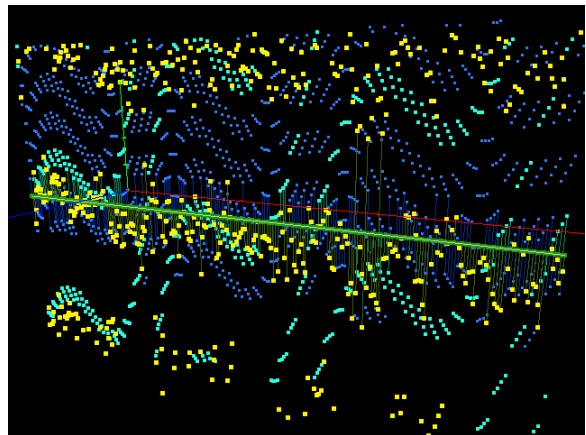


図54 被験者2の実験2の結果の3次元可視化データ

最初に、「3-5. 実験データの記録と3次元可視

化」で開発の部分を紹介した、自作の3次元可視化ソフトウェアによって得られる、実験結果データの図の見方についてまとめて紹介しておく。

まず全体の見方としては、画面左端にある座標軸として、垂直方向の緑色のZ軸は「個々のビートからの時間のずれの大きさ」(相対時間)を示す。右方向に伸びた赤色のX軸は、心理実験が進行する時間軸を示す。この全体画面では緑色の太い直線のように見えている部分は、時間軸方向に拡大すると、等間隔に並んだ「映像ビート」の点を示す緑色の矩形となる。この中心には緑色の直線を描画している。これを以降は「データ軸」と呼ぶことにする。左方向に伸びた青色のY軸は、便宜的にデータ軸の描画をX-Z平面およびX-Y平面から一定距離に離すために用いた次元であり、複数の実験データ(X-Z平面に平行な平面上にデータが並ぶ)をこの軸方向にいくつも並べて3次元的に比較する機能を想定したものである。この画面だけでは分かりにくいのが、画面はX,Y,Z軸方向それぞれに任意に拡大縮小でき、さらにX,Y,Z軸方向それぞれに任意に平行移動でき、さらにX,Y,Z軸に対してそれぞれに任意に視点回転することができる。

画面内にあるイエローの矩形は、それぞれのタイミングで被験者がスイッチをタップした瞬間を示す。基準となる等間隔の映像ビート(緑色の矩形)の乗ったデータ軸に対して、映像ビートのタイミングよりもタップが時間的に早い場合にはデータ軸に対してZ軸の負方向(下側)に、時間的に遅れた場合にはデータ軸に対してZ軸の正方向(上側)にイエロー矩形を置き、そこからデータ軸に垂線を下ろしている。この画面でデータ軸付近に多数のタップ点(イエローの矩形群)があり、さらにZ軸の下に離れた部分に同じデータ群が見えているのは、この画面のZ軸方向のズームングにおいては、隣の映像ビートまでの時間間隔(約880msec)だけ、Z軸方向でこの両群が離れていることを意味している。Open-GLでの描画範囲は限定しているので、もう少しだけZ軸方向のズームングを拡大すると、この下にあるデータ群は画面の下に消え去って見えなくなる。

画面内にあるライトブルー(水色)の矩形は、音楽ビートのタイミングを示す。さらに小さな青の点は、8ビートごとにライトブルーの音楽ビートが来る、その合間にハイハットで刻んだ音楽タイミングを示している。全体図で明らかかなように、音楽ビートのデータは時間的にめらかなに上昇・下降を繰り返しており、これは等間隔の映像ビートに対して音楽ビートのテンポが速くなったり遅くなったりしている事を示している。これは実験2のデータなので、X軸(データ軸)方向はこの画面全体で、実験時間の全体、すなわちおよそ3分間(約180秒)を示している。このように、X軸方向の時間とZ軸方向の時間とでは、この画面の場合200倍程度のスケールの違いがある(検討するケースごとにさらに変化すること)に注意されたい。

7-3. 実験3の解析

「実験2の解析」に続いて、「実験3の解析」につ

いても、実験1の結果「タッピング上級者グループ」に分類された32人のデータだけを検討した。個別のコメントは省略するが、以下に3件の代表データのみを紹介する。全被験者のデータについては文献[71]を参照されたい。

図55は被験者2の、図56は被験者3の、図57は被験者4の、実験3の結果の3次元可視化データである。いずれも実験結果データを、データ軸方向(X軸方向)に一部拡大表示して抜き出した。それぞれいくつかの区間で、音楽ビートを示すライトブルーとブルーのドット列の傾き(微妙なテンポの変動)に、イエローのタッピングのドット列が同じ傾きで追従している場所がわかる。そして、この傾きによって緑色のドット列の乗っている映像ビートのデータ軸から大きく離れると、音楽ビートへの追従をやめて新しいビート点から追従を開始している。これが、筆者が本研究で検証したかった「引き込み効果」の現れの例である。

第9章で全体のデータについての考察を述べるが、実験3では音楽ビートの方のテンポを一定にして、映像ビートのテンポを相対的にわずかに増減させている。しかし3次元可視化ツールの表示においては、厳密には微妙に間隔が疎密になっている映像ビートを等間隔になるようにnormalizeして、対応する音楽ビートの時間的なずれを表示して、実験2と同様の見え方にしている(一見すると変化しているように見える音楽ビートの変化の傾向が実験2とは違っていることで検証できる)。そして実験3で得られた結果としては、どの被験者についても実験2とほとんど同じ傾向が得られた、という点に注目したい。これは、映像的ビートと音楽的ビートでは、本実験で行ったような環境/条件においては、役割が対等ではなく、「音楽ビートが映像ビートに影響を与える」という関係が、どちらのテンポが微妙に変動していても得られた、という事を意味する。この点については、今後の研究の重要なテーマであると考えている。

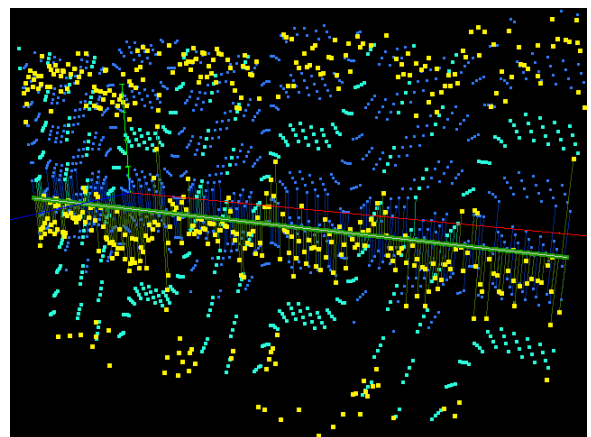


図55 被験者2の実験3の結果の3次元可視化データ

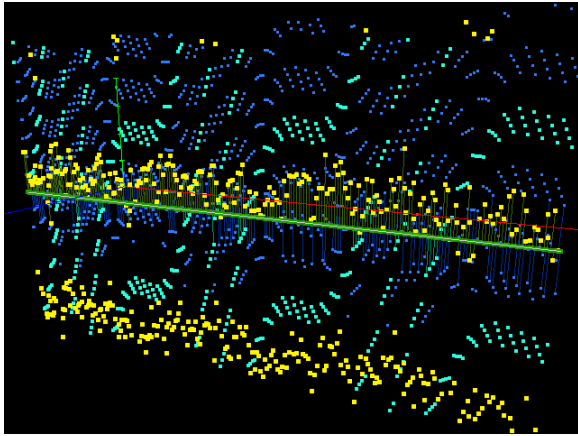


図56 被験者3の実験3の結果の3次元可視化データ

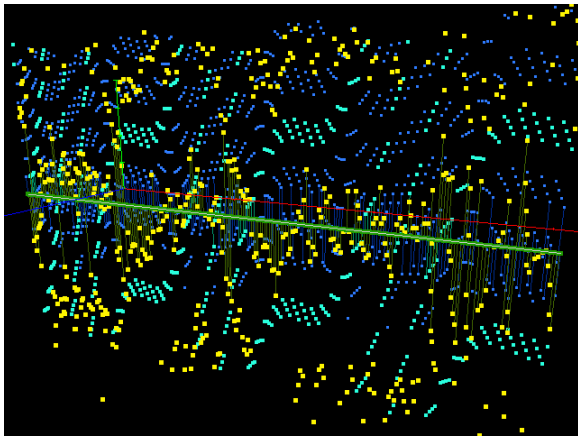


図57 被験者4の実験3の結果の3次元可視化データ

8. 瞬時変化映像による補足実験

メインの心理学実験が全て終了したあとで、実験の様様を記録したビデオで被験者の視線を解析していたところ、ほぼ同じ地点に地蔵の顔があるモーフィング映像、という素材そのままに、ほとんど被験者の視線が動いていないという事実を確認できた。これは予備的な検討の際に考慮した、瞬間的な映像刺激の提示で起きるサッカードをうまく避けた事実と思われるが、実際に本研究のデータとしてより明確化することを目指して、補足となる実験を行った。

8-1. 補足実験の概要

補足実験[A]では、映像刺激の違いだけを鮮明にして比較検討したいので、心理学実験システムも、システムのメインパッチも、まったく実験2-3と同一である。唯一の違いは、以下に述べるように映像素材であるQuickTimeムービーを別のものに差し換えただけであり、このムービーについても、画質や毎秒フレーム数、全体の長さなどは全て同一条件とした。被験者については、メインの心理学実験を受けた被験者には予備知識なり先入観があるので、メインの心理学実験の被験者とは異なる学生1名(男

性、20歳、音楽経験あり[SUAC軽音の部長])を用いた。実験は、上述のメイン実験2(モーフィング映像)に相当するものを行い、続いて以下のように映像を切り替えた新しい実験2を行った。従って、実験3に相当する、「音楽ビートのテンポが一定で映像ビートのテンポだけが変化する」というパターンは行っていない。

同じ被験者で引き続き行った補足実験[B]で用いたムービーでは、モーフィングでなく全体を240枚の静止画として等分割した。つまり、メインの実験では地蔵A、地蔵B、・・・とモーフィングでつながっている動画であるのに対して、この補足実験の映像素材はモーフィング1本分の時間だけ連続した静止画があり、次のモーフィング映像に対応する新しい静止画に瞬間的に切り替わる。被験者には、モーフィング映像では静止画の瞬間をビートとして定義してスイッチ操作してもらい、静止画の瞬時切り替え映像においては切り替えの瞬間をビートとするように指示した。

静止画素材については、市販の著作権フリーの素材写真集からジャンルをアトランダムに選択し、さらに手作業で隣接する静止画が似たようなものにならないように配置した。個々の静止画はメイン実験のモーフィングの1ブロックの時間と同一の条件の枚数で並び、音楽ビートに対する位置関係(モーフィングにおける静止画の場所と瞬間的に切り替わる場所)も同一にした。

実験[B]中の被験者の行動として顕著だったのは、メインの心理学実験1/2/3においては、どの被験者もほとんど全身の身体動作を大きく伴った「ビートにのる」という行動は見られなかったのに対して、この補足実験の被験者の場合、それも瞬間的に切り替わる映像素材の場合のみ、かなり激しい全身の揺れによって映像ビートを捕らえようと(無意識に行動)した、という点である。この補足実験の被験者も、他の被験者と同じ「地蔵モーフィング」の映像素材での実験[A]においては、身体はリラックスして椅子に腰掛けてスイッチ操作をしていたのに対して、静止画の瞬時切り替えになった途端に映像ビートを追い掛けるのが困難になったためか(モーフィングの場合には運動の予測が付きやすい)、上体を前後に激しく揺すって、(音楽ビートとずれている)映像ビートとの同期を図っている行動となった。この行動はこれ自体が非常に興味ある現象であり、本研究から派生するものとして研究を行う余地がある、と報告しておきたい。

8-2. 補足実験のデータの検討

図58は、補足実験の実験結果データを解析する3次元可視化ツールの画面例である。

以下の補足実験データの解析においては、この画面内にある3次元表示スクリーンの全体表示をまず示し、続いてそれぞれ5枚ずつの、時間軸方向に拡大表示して時間軸方向に移動させた詳細データを示す。

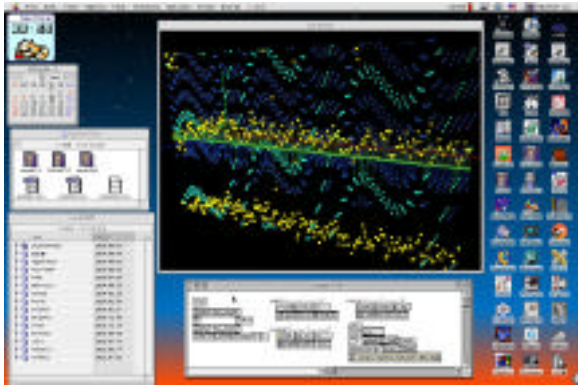


図58 3次元可視化ツールの画面例

8-2-1. モーフィングによる映像素材の実験

[A]の結果

図59はこの実験[A]の結果データ全体のいちばん最初の部分を、データ軸方向(X軸方向)に拡大表示した例である。実験の冒頭はいきなり映像も音楽も始まるので、この被験者のデータでもそうであるように、最初のあたりは映像ビートとタッピングデータとがばらばらに離れている。これは他の被験者でも一般的なものである。

この被験者は比較的节奏感が良く、映像ビートに対するタッピングの時間的偏差の絶対値が非常に小さい。また、このズーム画面の左半分、つまり前半では、音楽ビートが右下がり、つまり音楽のテンポは次第に遅くなっているが、この時点では被験者は実験冒頭ということで映像ビートを探してタッピングすることに集中していて、音楽ビートの変化にあまり影響されていない事がわかる。

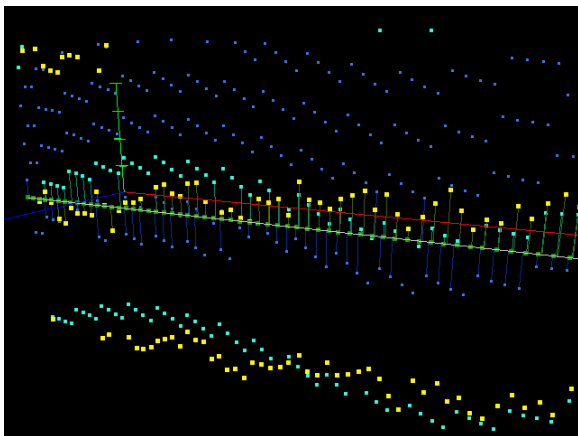


図59 補足実験[A](モーフィング)の拡大データ

図60は全体図をデータ軸方向(X軸方向)に拡大表示して、図59よりも時間的に後の状況まで平行移動したところである。この画面の左半分(前半)では、映像ビート(グリーン)の乗ったデータ軸と音楽ビート(ライトブルー)がほとんど重なっていて、さらに音楽ビートのテンポがほぼ一定(遅くなってやがて速くなっていく折り返し点)であることもあり、

タッピングは映像ビートおよび音楽ビートの地点に非常によく乗っていることがわかる。

そしてこの図の右半分(後半)では、テンポが次第に速くなったために映像ビートと音楽ビートが大きく離れてきて、タッピングは音楽ビートを分割するブルーの点をなぞりながら、時間的には右下がりの「のこぎり波形」のような動きをしている。これは本研究で提唱する「引き込み効果」の見事な実例である。

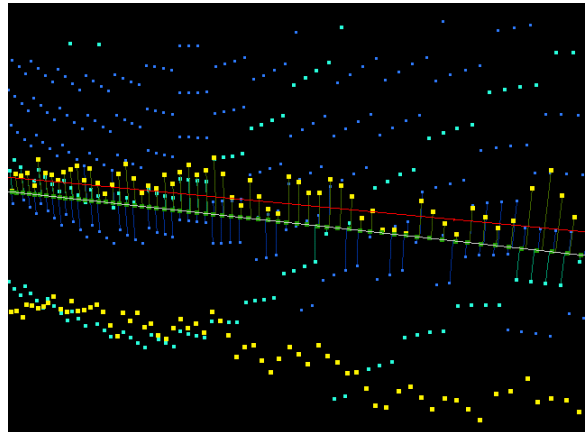


図60 補足実験[A](モーフィング)の拡大データ

図61は全体図をデータ軸方向(X軸方向)に拡大表示して、図60よりもさらに時間的に後の状況まで平行移動したところである。この画面では、音楽ビートのテンポが速くなったあとで反転して、後半では遅くなってきている。被験者のタッピングはあいかわらず正確であるが、後半の減速区間(音楽ビートの列が右下がり)において、黄色のタッピング矩形の連なりが、右下がりの「のこぎり波形」となっていることがわかる。これも筆者の提唱する「ビートの引き込み効果」と考えることができる。

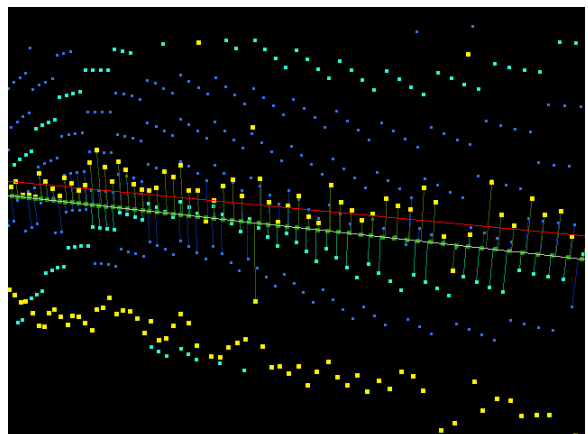


図61 補足実験[A](モーフィング)の拡大データ

8-2-2. 静止画の瞬時変化による映像素材の実験[B]の結果

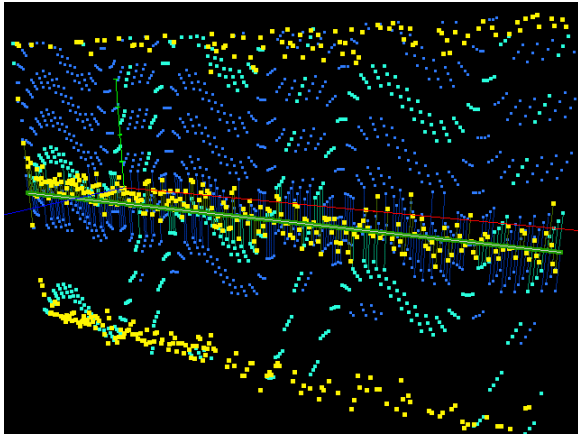


図62 補足実験[B] (静止画瞬時変化)の全体データ

図62は、この補足実験のために特別に制作した、「静止画の瞬時変化による映像素材」を用いた実験[B]の結果であり、同等の条件で全体を表示したものである。まずこの全体画面をモーフィング素材の実験[A]の結果と比較して特徴的なのは、被験者のタッピングのデータであるイエローの矩形が、画面の中央のデータ軸の周囲とともに、表示エリアの上限の部分と下限の部分にもほぼ同等に出ている、という点である。これはモーフィングの映像素材の時には見られなかった傾向であるだけでなく、実はメインの心理学実験の被験者からは、ほとんどこのようなデータは得られていないという特殊なものである。

モーフィングによる映像素材では、タッピングのデータはデータ軸のわずかに上方にあるために、映像ビートの単位だけずれた同じデータ群は画面下方にだけ出現した。これは、この被験者が映像ビートに対して、ほぼ一律にわずかに遅れた時点をタッピングの「ヒット点」として認識していることの証しである。これに対して、図62のように上下にほぼ同等に出現するのは、タッピングのデータが完全にデータ軸の中央に平均を持っているからであるが、モーフィング素材ではこれはほとんど見られない現象なのである。

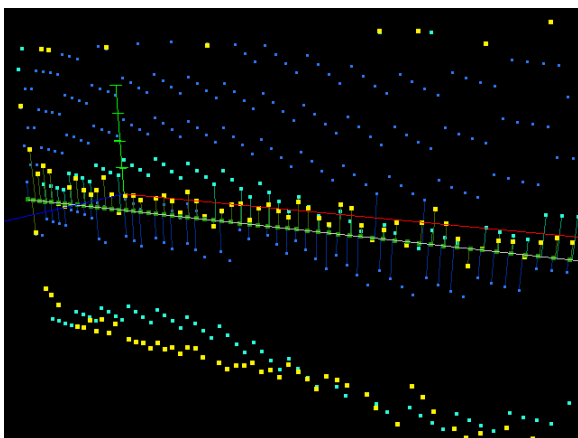


図63 補足実験[B] (静止画瞬時変化)の拡大データ

図63は、この全体図のいちばん最初の部分を、データ軸方向(X軸方向)に拡大表示した例である。実験の冒頭はいきなり映像も音楽も始まるので、最初のあたりは映像ビートとタッピングデータとがばらばらに離れている。

そして、図64と図65の2つの図は全体図をデータ軸方向(X軸方向)に拡大表示して、図63よりも時間的に後の状況まで平行移動したところである。この特徴というのは、モーフィング映像素材に比べて、引き込み効果のような音楽ビートによる影響がほとんど見られないこと、そして映像ビートに対する時間的誤差が非常に小さい、という事である。これは、被験者は瞬間的に変化する映像の連続に対して、激しく上体と脚を揺すって「映像ビートに同期したタイミング」をなかば無意識に作り出していたためであり、音楽などを聞いている余裕はない、という外見上の印象であった。その結果、一定の傾向で映像ビートから遅れたり早まるというのではなく、身体動作というメトロノームによって、平均値がゼロである(データ軸上に並ぶ)ようなタッピングを実現したものであると思われる。

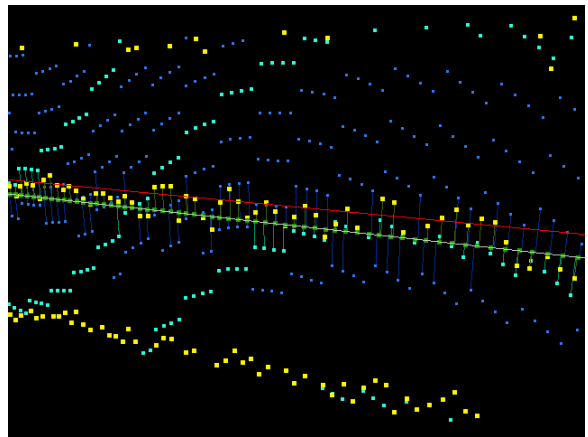


図64 補足実験[B] (静止画瞬時変化)の拡大データ

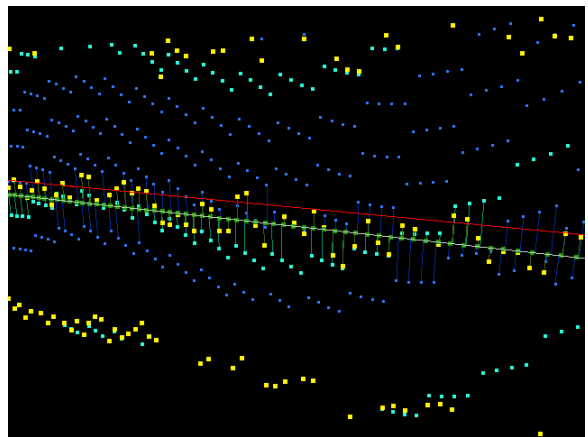


図65 補足実験[B] (静止画瞬時変化)の拡大データ

この結果は、静止画の瞬時切り替え、という映像

素材では、本研究で求める「映像ビート」を自然に実現できていない、という条件の傍証であると思われる。今後、リズムやビートの知覚については、さらに映像刺激の性格を検討する必要があるであろう。

9. 実験データの考察

以上の実験結果および補足実験の結果を受けて、本研究で提唱した「ビートの引き込み効果」について考察する。

9-1. 引き込み効果と個人差の検討

実験1において「タッピング初心者グループ」に分類された11人については、もともとBGMの無いところで映像ビートをタッピングするだけで、3次元可視化データのタッピング情報が映像ビートの乗ったデータ軸に対して上下のあちこちに広範囲に分布した。「タッピング上級者グループ」では本研究で提唱する「音楽ビートによる引き込み効果」を確認することができたが、この「タッピング初心者グループ」のデータから何か特徴的な傾向を検討・抽出することができれば、今後の新たな研究課題としての可能性がある。

主観的な申告によるアンケートの結果は、この両グループの区別と明確に対応づけることができなかったが、実験2および実験3で得られたデータは両グループとも全て完備しているため、以下のような視点で今後解析することで、新しい関係性を見出す可能性は残されている。ここではいくつかの仮説を提唱しておく。

- * 「タッピング初心者グループ」のばらついているデータに対して、ある種のソフトウェア的なフィルタ操作を実現することで一種のノイズ除去を実現し、主観的なリズム知覚に対応したビートのタッピング特性を抽出できれば、実験2および実験3での影響を見出すことができるかもしれない
- * 「タッピング初心者グループ」であっても音楽的ビートに対しては大きく影響される、という可能性は少なくないので、データのうち大きく離れたものを消去して有効なデータだけに限定しても意味ある結果が得られるかもしれない
- * 集団として43人の被験者というのは十分に大きいので、個性を捨象して全体のデータを全てまとめて統計処理してしまっても、意味のあるデータの傾向が得られるかもしれない
- * 本研究の実験では、実験2と実験3との結果については、定性的にはほぼ同等の傾向が得られたが、「映像が等速度で音楽がわずかにテンポ変動」と「音楽が等速度で映像がわずかにテンポ変動」とでは現象の出方が異なる可能性も十分にある。この点に注目してさらに分析することには意味がある
- * 予備実験とシステムパラメータの対応のチューニングの方法をより進展させて、予備実験の段階で個人差を有効に補正して、あとは得られたデータを全て単純に統計処理できるようにする工夫の余地がある

9-2. 結論

本研究の冒頭で、「思考実験」として説明した「引き込み効果」の例は、「テンポが一定の音楽を聞いてノリを感じながら、これとほんのわずかなだけテンポの異なるビートで歩行する人の、ステップのビートが音楽ビートに影響される」という状況であった。本研究の心理実験では、この音楽の部分はまったく同一なものとして、歩行のステップの代わりに、等間隔で繰返される映像ビートに合わせたタッピング、という身体動作に置き換えた。基本的には、等間隔にやってくる映像ビートに合わせてスイッチを叩く、という動作も一種のノリとして気持ち良さを返すであろう、というアナロジーである。そこでまず、本研究で提唱する「引き込み効果」を、この思考実験とおなじ形式で以下のように表現しておくことにする。

ある一定のテンポで等間隔に静止点を持つ映像があり、この静止点に合わせてスイッチをタッピングする、というタスクを実行している。そのタッピングのビートとテンポはちょうど快く、これを変えずにノリをもって続けることが気持ちいい状態である。そして同時に、ドラムとベースによる「ほとんど同じテンポの音楽」も聞こえていて、その音楽ビートも楽しんでいる。この音楽は自分が演奏しているわけではないので、テンポは変えることはできない。タッピングのテンポも快適なので、変化しない。

この実験では、映像ビートに相当する静止画部分にはかなり時間幅がある。モーフィングで動いていた映像がほぼ静止した瞬間、静止画とはっきり認識した瞬間、静止画がまだ続いているという瞬間、ゆるやかにモーフィングがスタートした瞬間。

そして映像のビートと音楽のビートはほんのわずかなテンポの違いから次第にずれてくると、ノリを失わないように、映像ビートの解釈点がわずかに移動しつつ、被験者は両者を同時に合わせる、という気持ち良さを維持しようとする。映像ビートに合わせる、というタスクのタッピングにおいて、さっきはモーフィング終了の瞬間が音楽ビートと一致していたのが、次第に静止画の中間地点、モーフィングの始まる瞬間、と少しずつ移動していく。しかし、いずれも時間的に幅を持つ映像ビートの幅の中なので許容され、乗れる。

そしてズレがいよいよ大きくなると、例えば音楽の表ビート(BDのビート)と同期していた映像ビートが、今度は自然に音楽の裏ビート(SDのビート)に「乗り換えて」またまた両者の同期したビートを楽しむ、という現象が起きる。細かく見れば、さっきまで静止画としてタッピングする瞬間と一致していた音楽のビートがずれたと知覚されると、音楽の裏表の反対側のビートがこれから静止画としてタッピングす瞬間にほぼ近くなり、このずれは次第に縮まって一致してきて、さらに移動しながらノリが続く。

結果として、音楽のビートのテンポも映像のビートのテンポもともに変化しない(引き込み現象entrainmentではない)にもかかわらず、映像ビートの許容幅の中で音楽ビートとの一致を自然に楽しむことで、音楽ビートに対して見てみると、しばらくは表ビートで乗り、やがて裏ビートで乗り、また次に表ビートでのノリに、と(無意識的に)乗り換えている。一致していると解釈することで気持ちいい、という瞬間に着目して映像タッピングのビートを見ると、そこでは少しだけ速いテンポの表ビートに引き寄せられて局所的に微小に加速し、やがて後から来る裏ビートに乗り換えてまた加速し、という速度の微小なギザギザ変化(ノコギリの歯のような形状)を繰り返すことになる。これが

「引き込み効果」である。

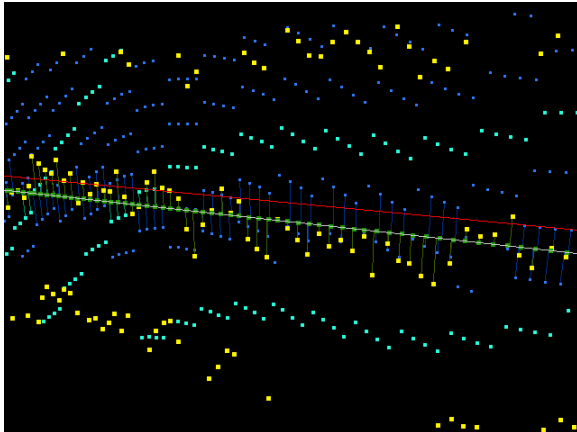


図91 被験者24の実験3の結果データの一部

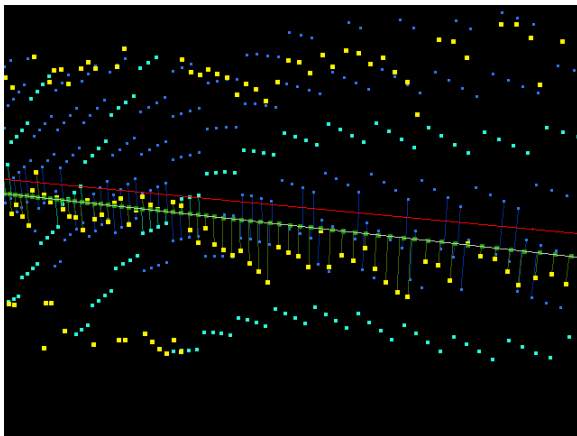


図92 被験者33の実験3の結果データの一部

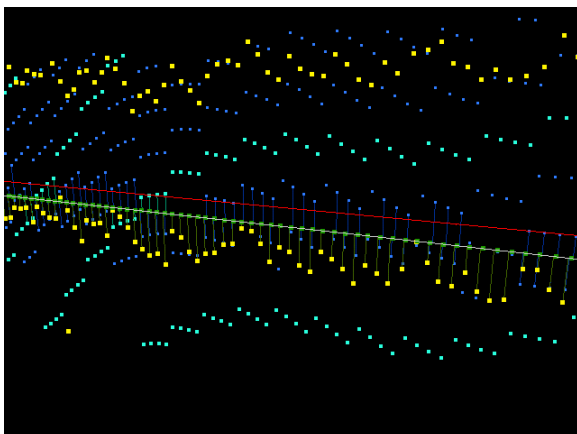


図93 被験者38の実験3の結果データの一部

以上のようなモデルを確認した上で、上に実験3の結果データの中から、この現象の典型的な3人の被験者のデータを示す。図91は被験者24の実験3の結果データの一部、図92は被験者33の実験3の結果データの一部、図93は被験者38の実験3の結果データの一部である。

タの一部である。上記の「引き込み効果」仮説の解説とともに、そのギザギザ(ノコギリ歯の形状)の行き来する様子を抽出できたことで、本実験でも十分に仮説が説明されていることが判る。定量的でなく定性的な現象の抽出であるが、「引き込み効果」はこのように確認できたと考える。これが本研究の結論である。

9-3. 今後の課題

本研究の実験結果データは全てを公開しているので、今後、新たな解析の指針と手法により新しい知見が得られる可能性がある。これが今後の第一の課題である。

本研究は、「ビート」「ノリ」というような視点から音楽的ビートと映像的ビートについて検証した初めての実験である。より新しい切り口から新しい心理学実験をデザインして、新しい仮説・現象を求めて行くこと、これが今後の課題の第2である。

本研究の心理学実験では、映像的ビートという刺激として、モーフィング映像を用いて、その静止画となる瞬間をビートと定義して被験者にタッピングを指示した。しかし一部の被験者にとって、これは不自然な要請であった。より効果的、有効であり、他の知覚心理学的な問題のない映像刺激についての検討が、今後の課題の第3である。

本研究では、音楽的ビートと映像的ビートの関係について調べて、付随的にタッピングという身体動作が関係した。しかしこれとは別に、フットタッピングとかパッドを叩く、というようなより積極的な身体動作ビート、あるいはダンスとの関係という要素も興味ある分野である。ここでは「音楽」「視覚」「身体動作」と要素が3つになるので、より関係性を洗練させる必要もある。この分野が今後の課題の第4である。

10. おわりに

人間はリズムとともに生存/生活している。本研究では、聴覚的情報と視覚的情報を同時に視聴するマルチメディアコンテンツの感覚間調和 intersensory harmonyに関して、新たな視点でビートを「周期的に繰り返しリズムのノリが知覚されるアクセント部分」と再定義した。その上で、音楽的ビートが映像的ビートの知覚に及ぼす局所的な「引き込み効果」を提唱し、これを解析・検証するための新しい実験システムをデザインし、被験者テストによる実験・評価を行った。また音楽情報科学の視点から、コンピュータを用いた音楽心理学/知覚認知心理学の実験で重要となる、実験システムのレイテンシ(遅延)についても詳細に検討・考察した。

リズムやビートというのは、音楽においてだけでなく、生命活動においても根源的な要素であり、あらゆる意味でのダイナミクスの源泉である。そしてマルチメディアというのは、人間が外界と、あるいは相互にコミュニケーションする重要な要素である。本研究のテーマはこの両者の融合領域に挑戦したものであり、このような実験の一つや二つではとうてい片鱗も見えない深遠なものであるが、逆にコ

ンピュータ支援によって、より明確にものごとの本質が抽出できる実験の可能性も大きくなってきた。過度に妄信せず心理学実験システムをデザインしながら、慎重に実験を進めて新たな知見を得られるよう、今後も精進していきたい。

謝辞

本研究にあたって、重要な視点と意見を提供いただいた村尾忠廣氏に感謝します。また、本研究にあたって、有意義な議論をしていただいた、新山王政和、山田真司、照岡正樹、中村文隆の各氏に感謝します。また、本研究に関して有意義な意見・討論をいただいた、日本音楽知覚認知学会、情報処理学会音楽情報科学研究会の皆様感謝いたします。また、本研究にあたって、被験者として協力いただいた、静岡文化芸術大学の学生・職員の皆さんに感謝いたします。

参考文献

- [1] 梅本亮夫：音楽心理学，p.594，誠信書房，1966.
- [2] ダイアナ・ドイチユ/寺西立年，大串健吾，宮崎謙一訳：音楽の心理学，p.663，西村書店，1987.
- [3] リタ・アイエロ/大串健吾訳：音楽の認知心理学，p.371，誠信書房，1998.
- [4] 梅本亮夫：音楽心理学の研究，p.375，ナカニシヤ出版，1996.
- [5] 岡田晋：映像学・序説，p.260，九州大学出版会，1996.
- [6] ジェイムズ・モナコ/岩本憲児他訳：映画の教科書，p.438，フィルムアート社，1983.
- [7] P.H.リンゼイ，D.A.ノーマン/中溝幸夫，箱田裕司，近藤倫明訳：感覚と知覚，p.319，サイエンス社，1983.
- [8] 森敏昭，吉田寿夫：心理学のためのデータ解析テクニカルブック，p.349，北大路書房，1990.
- [9] R.H.ディ/島津一夫，立野有文訳：知覚的解決：知覚心理学，p.249，誠信書房，1972.
- [10] 国際電気通信基礎技術研究所：視聴覚情報科学，p.271，オーム社，1994.
- [11] 淀川英司，東倉洋一，中根一成/電子情報通信学会編：視聴覚の認知科学，p.192，電子情報通信学会/コロナ社，1998.
- [12] 武藤真介：計量心理学，p.155，朝倉書店，1982.
- [13] 松田隆夫：知覚心理学の基礎，p.294，培風館，2000.
- [14] 相場覚，鳥居修晃：知覚心理学，p.210，放送大学教育振興会，1997.
- [15] 福田忠彦：生体情報システム論，p.242，産業図書，1995.
- [16] 森本正昭：情報処理心理学，p.226，誠信書房，1979.
- [17] 波多野誼余夫：音楽と認知，p.155，東京大学出版会，1987.
- [18] L.クラークス/杉浦実訳：リズムの本質，p.142，みすず書房，1994.
- [19] 中島義明：映像の心理学-マルチメディアの基礎，p.270，サイエンス社，1996.
- [20] 西井雄一郎，栗本育三郎：ドラムパッドを利用した実演音とクリック音との聴覚的ずれ検出装置について，情報処理学会研究報告，Vol.93，No.32 (93-MUS-1)，pp.17-23，1993.
- [21] 後藤真孝，村岡洋一：音楽音響信号に対するビートトラッキングシステム，情報処理学会研究報告，Vol.94，No.71 (94-MUS-7)，pp.49-56，1994.
- [22] 山田真司，井村和孝，新井裕子，小田満理子，西村英樹：音楽演奏者の時間的制御能力について，情報処理学会研究報告，Vol.95，No.46 (95-MUS-10)，pp.21-28，1995.
- [23] 堀内靖雄，三井卓，井宮淳，市川薫：二人の間による演奏の収録と分析，情報処理学会研究報告，Vol.96，No.53 (96-MUS-15)，pp.21-26，1996.
- [24] 後藤真孝，村岡洋一：音楽音響信号を対象としたビートトラッキングシステム ---小節線の検出と打楽器音の有無に応じた音楽的知識の選択---，情報処理学会研究報告，Vol.97，No.67 (97-MUS-21)，pp.45-52，1997.
- [25] 堀内靖雄，三井卓，財津茜，市川薫：二人の間による演奏の協調動作について，情報処理学会研究報告，Vol.98，No.74 (98-MUS-26)，pp.103-108，1998.
- [26] 堀内靖雄，財津茜，市川薫：人間の演奏制御モデルの推定，情報処理学会研究報告，Vol.99，No.51 (99-MUS-30)，pp.59-64，1999.
- [27] 長嶋洋一：MIDI音源の発音遅延と音源アルゴリズムに関する検討，情報処理学会研究報告，Vol.99，No.68 (99-MUS-31)，pp.31-38，1999.
- [28] 長嶋洋一，中村文隆：メディアアートにおける画像系の制御について，情報処理学会研究報告，Vol.2000，No.76 (2000-MUS-36)，pp.31-36，2000.
- [29] 浜中雅俊，後藤真孝，麻生英樹，大津展之：学習に基づくクオンタイズ-発音時刻の楽譜上での位置の推定，情報処理学会研究報告，Vol.2001，No.45 (2001-MUS-40)，pp.21-28，2001.
- [30] 難波精一郎，林勇気：画像中の円の落下と音の変化の共鳴現象，情報処理学会研究報告，Vol.2002，No.40 (2002-MUS-45)，pp.9-12，2002.
- [31] 新山王政和，村尾忠廣，南曜子，小川容子：音楽ビートと運動ビートのタイムラグについて---マーチングステップの熟達者と未経験者の相違について---，情報処理学会研究報告，Vol.2002，No.40 (2002-MUS-45)，pp.79-84，2002.
- [32] 下迫晴加，石田時敬，菊地正：音系列における時間間隔の変化の検出，情報処理学会研究報告，Vol.2002，No.40 (2002-MUS-45)，pp.91-96，2002.
- [33] 後藤靖宏，阿部純一：拍子解釈の基本的嗜好性と漸進的確立，音楽知覚認知研究，Vol.2，日本音楽知覚認知学会，pp.38-47，1996.
- [34] 末富大剛，中島祥好：リズム知覚研究の動向，音楽知覚認知研究，Vol.4，No.1，日本音楽知覚認知学会，pp.26-41，1998.
- [35] 菅野禎盛，岩宮眞一郎：音楽のリズムと映像の動きの同期が音楽と映像の調和に及ぼす効果，音楽知覚認知研究，Vol.5，No.1，日本音楽知覚認知学会，pp.1-10，1999.
- [36] 岩宮眞一郎，上月裕，菅野禎盛，高田正幸：音楽の調性及びテンポと映像の速度及び密度が映像作品の印象に及ぼす影響，音楽知覚認知研究，Vol.8，No.2，日本音楽知覚認知学会，pp.53-64，2002.
- [37] Manfred Clynes：Time, Timeconsciousness and Music, Proceedings of 1st International Conference on Music Perception and Cognition, pp.249-254, Kyoto, Japan, 1989.
- [38] Yoshitaka Nakajima, Takashi Nomura, Takashi Tsumura：Physically Wrong Rhythms can be Subjectively Correct, Proceedings of

- International Symposium on Musical Acoustics, pp.187-190, Tokyo, Japan, 1992.
- [39] David Rosenthal : A Model of the Process of Listening to Simple Rhythms, Proceedings of ICMC1988, pp.189-197, International Computer Music Association, 1988.
- [40] Gregory Kramer, Stephen Ellison : Audification ---The Use of Sound to Display Multivariate Data---, Proceedings of ICMC1991, pp.214-221, International Computer Music Association, 1991.
- [41] David Rosenthal : Intelligent Rhythm Tracking, Proceedings of ICMC1992, pp.227-230, International Computer Music Association, 1992.
- [42] Jeff A.Bilmes : Techniques to Foster Drum Machine Expressivity, Proceedings of ICMC1993, pp.276-283, International Computer Music Association, 1993.
- [43] Peter Design, Henkjan Honing : Advanced Issues in Beat Induction Modeling ---Syncopation, tempo and Timing---, Proceedings of ICMC1994, pp.92-94, International Computer Music Association, 1994.
- [44] Simon C.Roberts, Michael Greenhough : Rhythmic Pattern Processing using a Self-Organising Neural Network, Proceedings of ICMC1995, pp.412-419, International Computer Music Association, 1995.
- [45] Leigh M.Smith : Modeling Rhythm Perception by Continuous Time-Frequency Analysis, Proceedings of ICMC1996, pp.392-395, International Computer Music Association, 1996.
- [46] Vijay Iyer, Jeff Bilmes, Matt Wright, David Wessel : A Novel Representation for Rhythmic Structure, Proceedings of ICMC1997, pp.97-100, International Computer Music Association, 1997.
- [47] Eric Gottesman : The Psycho-Adaptive Listening Machine ---An Application of Perceptual Control Theory to Computer Music---, Proceedings of ICMC1998, pp.195-198, International Computer Music Association, 1998.
- [48] Tommi Ilmonen, Tapio Takala : Conductor Following with Artificial Neural Networks, Proceedings of ICMC1999, pp.367-370, International Computer Music Association, 1999.
- [49] Ali Taylan Cemgil, Bert Kappen, Peter Design, Henkjan Honing : On Tempo Tracking ---Tempogram Representation and Kalman Filtering---, Proceedings of ICMC2000, pp.352-355, International Computer Music Association, 2000.
- [50] Yoichi Nagashima : Improvisation-II ---A Performing/Composing System for Improvisational Sessions with Networks---, Entertainment Computing, pp.241-248, IFIP / Kluwer Academic Publishers, 2003.
- [51] Hiromi Hashimoto, Masashi Yamada : Temporal Fluctuation in Equal Interval Tapping using Various Muscle Groups, Proceedings of 17th Congress of the International Association of Empirical Aesthetics, pp.525-528, Takarazuka, Japan, 2002.
- [52] 吉田友敬, 山本佐代子, 廣瀬百合子, 武田昌一 : 音楽リズムの同調と引き込み現象(1), 日本音楽知覚認知学会平成14年度秋季研究発表会予稿集, 日本音楽知覚認知学会, pp.79-86, 2002.
- [53] 難波精一郎, 林勇気 : 画像中の円盤の動きと音の変化の同期---枠組みの影響---, 日本音楽知覚認知学会平成14年度秋季研究発表会予稿集, 日本音楽知覚認知学会, pp.115-120, 2002.
- [54] 刈谷亜未太, 青野正二, 桑野園子, 難波精一郎 : 規則的なパルス列のタッピング作業におよぼす挿入刺激の効果, 日本音楽知覚認知学会平成14年度秋季研究発表会予稿集, 日本音楽知覚認知学会, pp.143-150, 2002.
- [55] 吉田友敬, 山本佐代子, 廣瀬百合子, 武田昌一 : アンサンブルにおけるリズムの同調との分析, 日本音響学会全国大会講演論文集, 日本音響学会, pp.863-864, 2003.
- [56] 長嶋洋一 : MIDI音源の発音遅延と音楽心理学実験への影響, 日本音響学会音楽音響研究会資料 Vol.18, No.5, 日本音響学会, pp.47-54, 1999.
- [57] 長嶋洋一, 橋本周司, 平賀謙, 平田圭二 : コンピュータと音楽の世界, p.484, 共立出版, 1999.
- [58] 長嶋洋一 : コンピュータサウンドの世界, p.180, CQ出版, 1999.
- [59] 長嶋洋一 : Java & AKI-80, p.176, CQ出版, 1997.
- [60] 長嶋洋一 : 作るサウンドエレクトロニクス, <http://nagasm.suac.net/ASL/mse/index.html>, 1999.
- [61] 村尾忠廣 : タゴリズムからの発見, 季刊音楽教育研究56, 音楽の友社, pp.177-190, 1988.
- [62] 村尾忠廣 : <拍ノリ>の裏・表 タゴビートの裏・表, 音楽教育学 18-1, 日本音楽教育学会, pp.31-36, 1988.
- [63] 村尾忠廣(代表) : わらべ唄・自由斉唱におけるピッチの統一化と運動ビートの同期化プロセスの研究, 研究成果報告書 平成10-12年度文部科学省研究費補助金基盤研究(B)(2) 研究課題番号10480045, p.118, 2000.
- [64] 新山王政和 : フットタッピングによるテンポ同期の実験研究, 音楽教育学 27-1, pp.53-68, 日本音楽教育学会, 1997.
- [65] 新山王政和 : 音楽ビートと運動ビートの知覚に関する分析的研究 ---「テンポ同期」から「ステップ動作」までの一連の検証実験のまとめ---, 音楽教育学第33-2号別冊 日本音楽教育学会第33回プログラム, pp.15, 日本音楽教育学会, 2003.
- [66] 長嶋洋一 : オリジナルMIDIセンサを作ろう, <http://nagasm.suac.net/ASL/original/index.html>, 1999.
- [67] 梅田利彦, 山田真司, 北村音一 : ピアノの長3和音の各音の発生時間による変化, 日本音響学会講演論文集, 日本音響学会, pp.447-448, 1989.
- [68] 音楽的ビートが映像的ビートの知覚に及ぼす引き込み効果(1) --- その第1報と実験計画 ---, <http://1106.suac.net/news/docs/onchi-1.pdf>, 2003.
- [69] 音楽的ビートが映像的ビートの知覚に及ぼす引き込み効果(2) --- 心理学実験システムの開発とレイテンシの計測 ---, <http://1106.suac.net/news/docs/sigmus0308.pdf>, 2003.
- [70] 音楽的ビートが映像的ビートの知覚に及ぼす引き込み効果(3) --- 被験者インターフェースとデータの3次元可視化 ---, <http://1106.suac.net/news/docs/FIT2003.pdf>, 2003.
- [71] <http://nagasm.suac.net/ASL/beat/index.html>, 2003. (本論文の公表後に公開予定)