

## マルチメディアへの心理学的接近

柳川 友香 大谷 康子 桑木 景子 奥 正廣 伊吹 公夫

東京工科大学

視聴覚マルチメディアシステムの設計、及びその発展には人間の認知機構の理解が不可欠である。そのためには人間の認知機構への理解を深めることが必要であり、心理学的検討の積み重ねが要求される。特に、人間の情報認知には様々な要素が関わってくることから、その相互関係が認知にどのように関わってくるかを検討していく必要がある。

そこで、分散システムの一例である計算機オーケストラと音響認識による動画制御システムを試作事例として、昨年に引き続き約100人ほどの被験者を対象とした2つの認知心理学的実験調査を行った。これらの実験、及びその検討結果の報告を今回行う。

PSYCHOLOGICAL APPROACH TO  
MULTIMEDIA SYSTEM

Tomoko Yanagawa Yasuko Ohtani  
Keiko Kuwaki Masohiro Oku Kimio Ibuki  
Tokyo Engineering University  
1404-1 Katakura, Hachioji-shi, Tokyo 192, Japan

Thorough understanding of the human cognitive mechanism is useful for design and improvement of the audio-visual multimedia system. For the understanding, it is important to pile up many psychological investigations of human cognition. Especially, it needs to investigate and examine how each factor and their mutual relations affect human cognitive mechanism, because many factors are concerned in the mechanism.

Therefore, we designed two of the trial systems, the computer controlled orchestra, and the audio controlled animation system, and made two cognitive experiments on about 100 people. These experiments and the results of their examination will be reported in this paper.

1. はじめに ～ 知覚メディアに対する心理学的検討の意義 ～

コンピュータ・システム、特に知覚メディアシステムのさらなる発展を目指すためには人間自身の機能を把握する必要がある。そのなかでも、特に視聴覚メディアシステムでは、人間の視聴覚認知機能との調和のとれたシステム設計が要求されるのではないだろうか。

人間は外界から知覚によって情報を取り入れ認識する。この際、対象からの情報をそのまま認識するのではなく、むしろ外界、あるいは感覚相互間での影響、本人の感性、経験などによって影響を受けた情報を認識しているのではないだろうか。人間の認知機構を知るためには、まず情報を認識する際に何が影響するのか、そして情報をどのように処理しているかを知る必要がある。(図1)

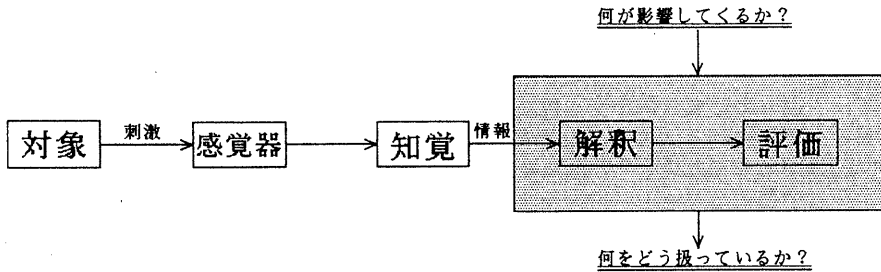


図1 知覚情報処理モデル

そこで今回、コンピュータ・オーケストラ、及び、かえるの手旗信号の2つの試作システムを取り上げ、聴覚、そして視覚から取り込まれる情報を人間が認識する際、どのようにその情報を処理しているのか、そして何が認識に影響を与えているのかを検討していくことにした。

2. 音楽における聴覚情報処理

人間が何かの音を認識する、つまり”聞く”という現象は図2の様になっていると考えられる。この時、音源からの”音”は直接的にそのまま認識されるのではなく、何らかの処理が行われる。その結果、認識される音は音源からの”音”そのものではなく、聞いている環境や、聞いている人の感性、経験など様々な影響を受けた情報として処理されたものではないかと思われる。

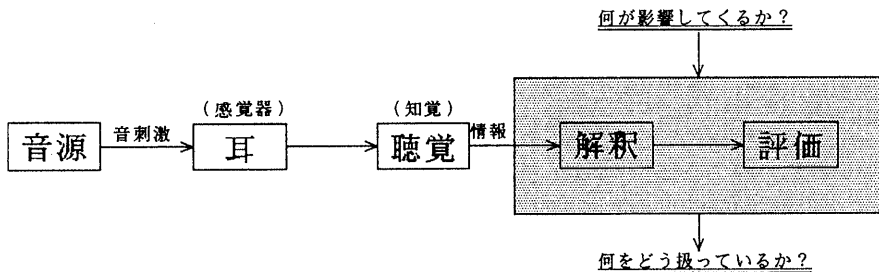


図2 聴覚情報処理モデル

聴覚情報処理は視覚情報処理にくらべ主観的な要素に影響されやすい。特に音楽に対する評価は、曲の親密度、被聴者の音楽経験度など、様々な要素がからみあって、その反応には個人差があると考えられる。しかし、実際には個人差はあってもある程度似たような反応を示す場合もある。

今回、人が音楽を聴く場合、これを個人差はあるにしろ図3の様な経路をたどって情報として処理しているのではないかと考え、このモデルの検証を行うことにした。

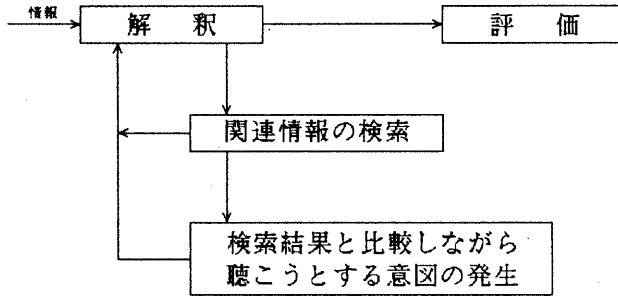


図3 音楽情報処理モデル

まず、左右のスピーカーより、演奏のタイミングはずれているが、その他の条件（曲名、テンポ、音色、……等）は全て同じ状態の演奏を流し、中央で被験者に聴かせて、その演奏をどう感じるか、演奏のずれに対する反応を中心に調査した。そして、その結果よりある程度傾向が現れている反応について、なぜそのような反応が現れたのかを検討することから図3のモデルの検証を行うことにした。詳しくは第5章で述べる。

### 3. 複数メディアに対する知覚情報処理

聴覚と同様に視覚の場合にも図4の様なモデルが考えられる。この場合も様々な要素が情報を処理する際に影響を与えているのではないと思われる。

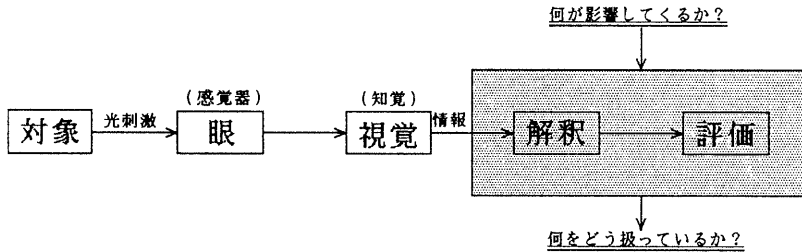
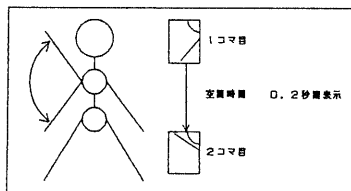


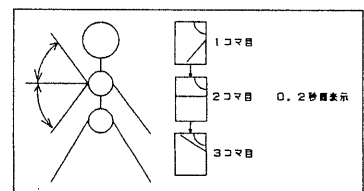
図4 視覚情報処理モデル

昨年度報告したように<sup>1)</sup>、図5のそれぞれの動画の自然性の比較を図6 (a) (b) の各システムにおいて行ったところ被験者の反応は異なった。しかし、何故反応が異なったのかは結果に影響を与える様々な要因が、実験環境自体に含まれてしまっているので原因を特定する事が出来ていなかった。考えられる要因としては、聴こえてくる音の違い、動画を見ている条件（演奏するか否か）の違い、動画の動きからくる予測、などが考えられる。



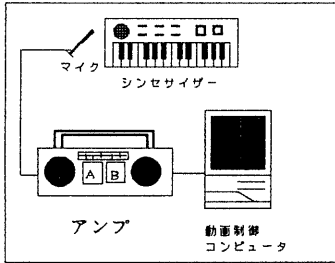
(a) 2コマ

(注)  
動画は、聴こえてくる音の周波数解析により音程を判断し、対応する部所を動かす仕組みになっている。



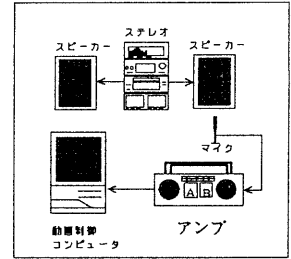
(b) 3コマ

図5 昨年度実験 ディスプレイ画面



(a) 演奏しながら見る場合

(注) 昨年度実験結果  
自然に見えるとの回答  
システム (a)  
2 コマ …… 42 %  
3 コマ …… 79 %  
システム (b)  
2 コマ …… 67 %  
3 コマ …… 88 %  
(複数回答可)



(b) 演奏を聴きながら見る場合

図6 昨年度実験 実験システム

そこで、今回の実験では、(b)のシステムで同一の曲のテンポを3段階に変化させて流し、その各々において2つの動画を被験者に同時に比較してもらうことにした。動画の一つ一つの動き自体は変化しないが、曲のテンポが速くなると単位時間あたりの音数が多くなるため結果的にその動く回数が頻繁になる。曲のテンポの変化と、動画の動きに対して被験者がどのような反応を示すかを調査し、さらなる検討を行った。詳細は第6章で述べる。

#### 4. 試作システム

今回の心理実験に用いた試作システムを図7に示す。

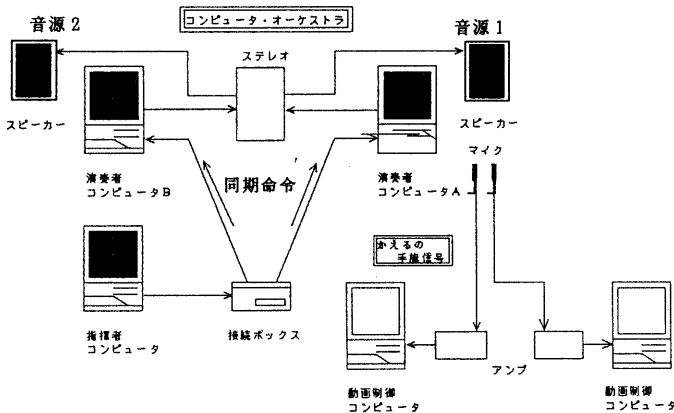


図7 実験システム

このシステムは、コンピュータ・ネットワークによる最小規模の分散制御システム(コンピュータ・オーケストラ)と、このオーケストラによって演奏される連続音を音源として認識し、各音程に対応して動画の動きを制御する音響認識による動画制御システム(かえるの手旗信号)の複合システムである。コンピュータ・オーケストラでは指揮者コンピュータから各パートに当たる演奏者コンピュータを制御する方式をとる。

第2章で述べた聴覚に関する実験はコンピュータ・オーケストラで、第3章で述べた動画の自然性に関する実験はかえるの手旗信号で行った。詳しくは第5章、第6章でそれぞれ述べる。

#### 5. コンピュータ・オーケストラによる音響認知実験

今回の実験では、片方の音源の演奏開始時間を約0.5秒遅らせ、演奏がずれている状態で、その印象を質問紙(Semantic Differential method: SD法)で尋ねた。実験には表1の4曲を採用し、被験者には、図8のように2つの音源と正三角形をなすような位置に立ってもらい、1曲を担当者が選択して3段階の速さについて、それぞれの演奏の印象を表2の5項目について5段階評価で

回答してもらった。

(なお、今回選んだ4曲はどれも印象に大差はなく、印象の類似した曲であるとして検討してよいことを予備実験で確認してある。)

そして、この回答をもとに、

A 曲の違いは、演奏に対して被験者の感じる印象に影響を与えるかどうか

B テンポの違いは、演奏に対して被験者の感じる印象に影響を与えるかどうか

の各点について検討した。

|   |
|---|
| (a) 指の歌<br>(b) おうま<br>(c) もうもう うしさん<br>(d) かえるのうた |
|---|

表1 実験曲

|   |
|---|
| 違和感のある／ない<br>余韻のある／ない<br>エコーのある／ない<br>ずれのある／ない<br>好きな／嫌いな |
|---|

表2 質問項目

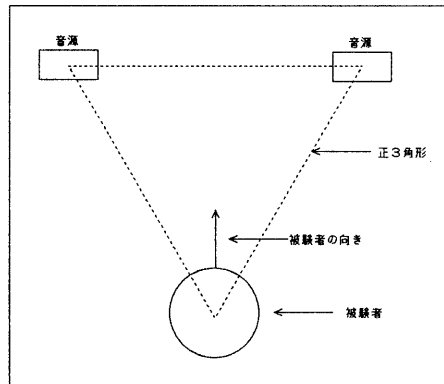


図8 被験者の立ち位置

回答結果をもとに、 $\chi^2$ 検定を中心に行った検討結果は次の通りである。

まず、曲の違いと演奏に対する印象との関係だが、以下の2点について検討した。

「曲の違いはずれの感じ方に影響するかどうか」については、5%の危険率で検定により、曲の違いによるずれの感じ方への影響が確認された。特に「かえるのうた」については他の曲に比べてずれを強く感じる傾向がでていた。

「曲の違いは余韻の感じ方に影響するかどうか」についても、同じく5%の危険率で検定により、曲の違いからくる余韻への影響が確認された。特に「おうま」については他の曲に比べて余韻を感じないという傾向がでていた。

次に、テンポの違いと演奏に対する印象との関係だが、これも以下の2点について検討した。

「曲の速さの違いは好き嫌いの感じ方に影響しない」については同様に5%の危険率で、速さの違いからくる好き嫌いへの影響が確認された。この場合、「速い」ほど“嫌い”と感じる比率が低くなる傾向が表れた。

「速さの違いは余韻の感じ方に影響しない」についても5%の危険率で、速さの違いからくる余韻の感じ方への影響が確認された。特に遅いほど余韻を強く感じる傾向がでていた。

このような各傾向が現れた要因として、図9の各図のような経路で被験者が演奏を聴いているためではないかと考えられる。

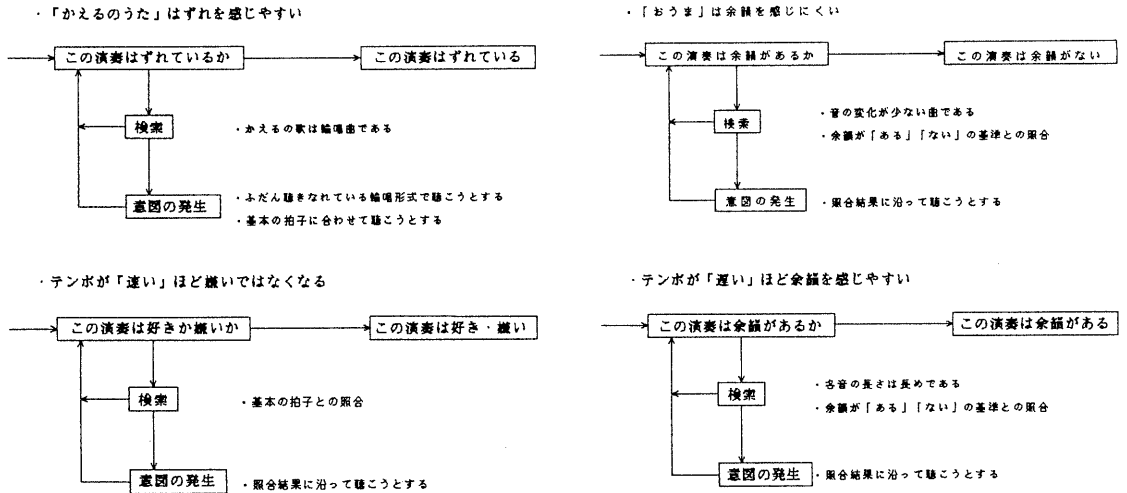


図9 実験結果に基づく音楽処理モデル

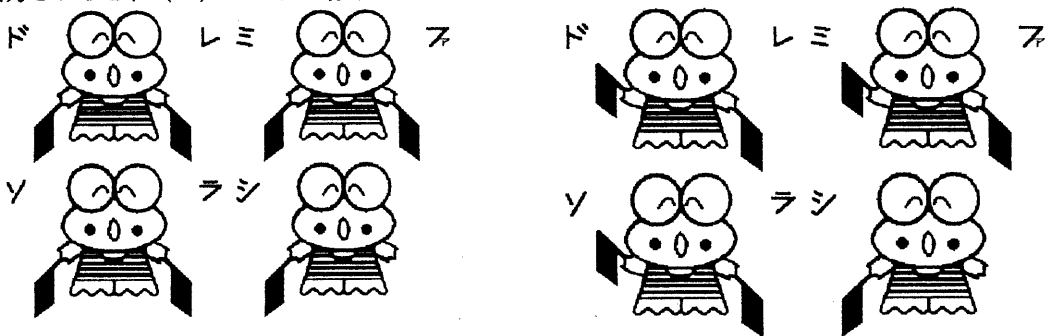
図9の各図は図3の音楽情報処理モデルに基づいたものである。よって今回の実験においてはこのモデルが成立するのではないと思われる。

## 6. 動画の自然性に関する実験

今回の実験条件は以下の通りである。

|          |       |       |       |
|----------|-------|-------|-------|
| 一分当たりの音数 | 48回   | 60回   | 120回  |
| 2音の時間間隔  | 1.25秒 | 1.00秒 | 0.50秒 |

ディスプレイ上には、図10(a)のような動画を表示する。これは、例えばド・ミ・ソの音が入力されると、(b)のように動くようになっている。



(a) 静止状態

(b) ド・ミ・ソが入力された場合

図10 ディスプレイ表示画面

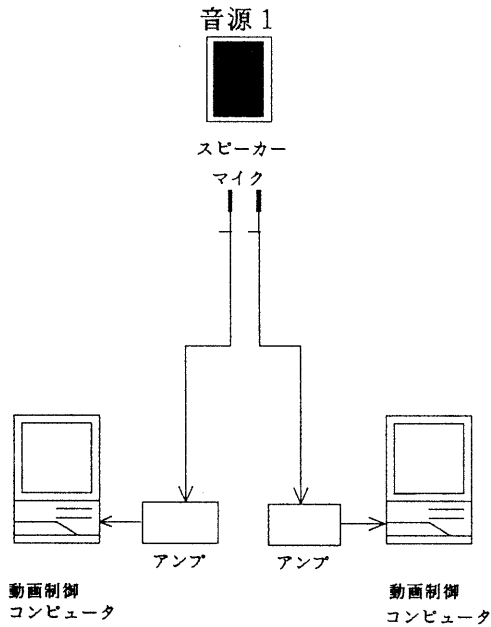
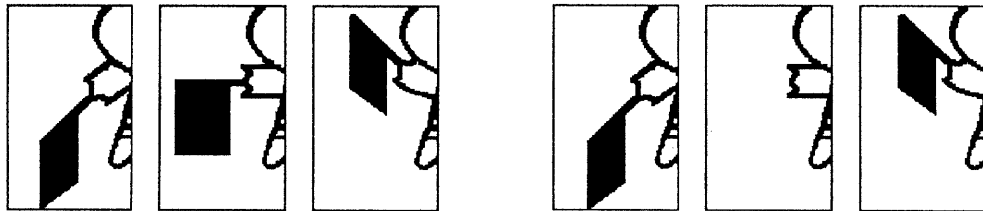


図 1.1 実験システム概略図

そして図 1.1 のシステムにおいて、1 台のコンピュータには図 1.2 (a) 3コマアニメーションを、もう 1 台には (b) 2コマアニメーションを表示した。被験者には、2 台のコンピュータを同時に見てもらい、どちらのアニメーションの方が動きが自然に見えるか、ということで、その見やすさと音楽とのタイミングの 2 つの点で比較してもらった。そしてその回答から

- A 単位時間当りに入力される音数の違いによって、動画が見やすいと思えるコマ数は変化するかどうか
- B 単位時間当りに入力される音数が、動画と音楽のタイミングに影響を及ぼすかどうか

の 2 点について検討した。



(a) 3コマアニメーション

(b) 2コマアニメーション

図 1.2 実験アニメーションの構成

被験者の回答をもとに  $\chi^2$  検定を中心として行った検討結果は次の通りである。

まず、仮説「単位時間当りに入力される音数の違いによって、動画が見やすいと思えるコマ数は変化しない。」つまり、「入力される音数が多くても、少なくても、動画が見やすいと思えるコマ数は、2コマになったり、3コマになったりしない。」についてであるが、動画の見やすさについての回答結果は図 1.3 のようになった。この結果は、動画が見やすいと思えるコマ数を 3コマとした場合も 2コマとした場合も、それぞれの値は仮説より導出した期待値 48.3、29.0 から離れている。実際の検定結果でも、この仮説は (5% の危険率で) 棄却された。すなわち「単位時間当りにマイクから入力される音数が変化すると、動画が見やすいと思えるコマ数が変化する。」ということになる。

また、仮説「単位時間当りに入力される音数の違いが、動画と音楽とのタイミングに影響を及ぼさない。」についてだが、動画とタイミングについての回答結果は図 1.4 のようになり、各値は、仮説より導出した期待値から離れている。実際の検定結果でも、この仮説は (5% の危険率で) 棄却された。つまり「単位時間当りに入力される音数が変化すると、音楽とタイミングよく動いて見える動画のコマ数は変化する。」ということになる。

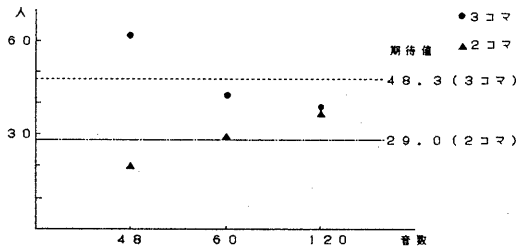


図13 音数と見やすさ

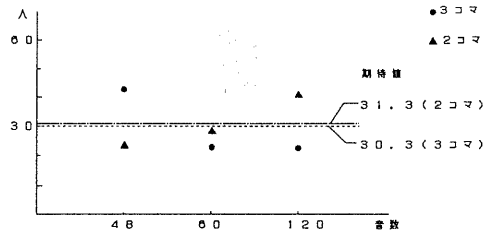


図14 音数とタイミング

以上2つの結果より、音数が少ない場合には3コマ、多いときには2コマアニメーションが見やすい傾向があるといえる。これは、単位時間当りに入力される音数が多くなると、動画が次々に制御されなければならないとなり、3コマアニメーションの2コマ目がなくても、動画の動きが雑には見えず、むしろコマ数が少ない方が余分なコマが表示されず動画として動きが見やすいと感じられ、逆に、単位時間当りに入力される音数が少ないと、動画と音1つ当りの対応時間が長くなるため、コマ数を減らしてしまうと、動きが雑に見えてしまい、自然性を損なう(1コマ1コマの区切りが見えてしまう)ので、コマ数を多くした方が、動きが自然に感じられるためではないか、と考えられる。

#### 7. おわりに ～ 今後への課題 ～

これまでの研究から、人間の視聴覚情報処理においては様々な要素が関わってくることが予想されてきた。一つ一つの要素に対してさらに検討を深めていくとともに、今後は各要素間の相互関係についても検討を進めて行きたいと思う。特に、

- 1) 採用した刺激(曲、図形など)の違いによる影響
- 2) 実験方式、及びその検討手段の改良
- 3) 反応の傾向と個人差について
- 4) 感覚間の相互作用とその影響
- 5) 人間の視聴覚情報処理と行動との関係

などに関して、様々な実験により、今回提唱したモデル検証を重ねるとともに、具体的なシステム設計のための資料を豊富にすることによって、人間の認知機構についての研究を深めて行きたいと思う。

#### 謝 辞

大学祭において、被験者として本実験に御協力頂いた方々に深く感謝する。

#### 参 考 文 献

- 1) 桑木景子、斎藤正彦、金子文昭、尾垣晴仁、伊吹公夫 :  
 "マルチメディア分散制御の視聴覚芸術への応用"  
 情処研報 Vol. 93, No 23 p. 53~p. 62 (1993)