

視聴覚情報を用いた言語・概念の獲得

中 西 宏 文

Hirobumi NAKANISHI

(情報科学選修)

Language and Concept Acquisition using visual and auditory information

Hirobumi NAKANISHI

(Department of Information and computer Sciences)

1. は じ め に

人間が外界から知識を得る手段は、視覚と聴覚によるものが大部分を占める。つまり、目で見たり耳で聞いたものを一種の「パターン」として捉え、それを理解し知識として蓄積したり行動のために利用する。

今日盛んに行われるようになった人工知能の研究は、こういった人間にのみ可能であった知識の保有・処理等を情報処理技術の立場から計算機処理の対象とし計算機に演繹・推論能力をも持たせようとするもので、将来的には人間の知的機能を模倣し、あるいはその代行をすることをも目指している。しかしながら、従来の人工知能の研究は、対象に関する概念・意味体系をあらかじめ分析し知識として記憶しておき、その上で推論操作等を行ういわば静的システムの研究にとどまっている。つまり、与えられた知識によって外的刺激に反応するが、外的刺激に対してシステムの知識を変更することはなかった。このため、環境に適応して能力・効率を高めていく事ができないばかりでなく、まったく未知の外的刺激に対してなんら合理的な反応を行うことができない。しかも、外的刺激には様々な種類があり相互に関連しているにもかかわらず、ただ一種類の外的刺激を扱うだけのものがほとんどであった。

これらの問題を克服するには、新しい概念や意味体系を外的刺激から獲得する能力、および未知の複数の外的刺激に対する意味付けを行うメカニズムの解明が必要である。このような概念獲得メカニズムが明らかになれば、人間の情報処理方式の解明ばかりでなく、人工的なシステムに未知の環境での適応能力を与えたり、大規模な知能システム（例えば音声言語理解、マルチメディア・インターフェース）のデータベースを自動作成する場合にも有用である。そのためには、人工知能の重要な基礎研究である「知覚」、「認知」、「学習」、「推論」、「記憶」、「知識の獲得・表現・利用」、「概念の形成」といった人間の情報処理メカニズムの解明が不可欠である。

本研究では、従来の研究では一度も試みられていない視覚と聴覚という2つの外的刺激

を結びつけて未知の入力に対する概念の獲得方式を定式化する。具体的には、イ) 視聴覚情報による子供の概念獲得の計算機シミュレーションを行うこと、ロ) 機械的に効率のよい帰納的・確率的概念獲得アルゴリズムを確立すること、ハ) 学習システムによって得られた情報を音声言語理解システムや画像理解・生成へフィードバックする方法の検討、及び、ニ) 音声一言語一概念一画像の相互関連を整理し、概念獲得という立場から融合化することを目的とする。今回はこれらの内、イ) とロ) の部分に関し基礎的な実験を行ない良好な結果を得たのでこれについて報告する。

なお本研究は入力として文字列とその意味を与えて文法と意味解析規則を帰納的に学習するシステムを開発したことが出発点となっている。^{1),2),3),4)}

2. 幼児期における言語および概念の獲得^{5),6)}

人間の幼児が、言葉を使えるようになるのは、個人差があるが、およそ1歳前後といわれている。そして、このころの言葉は「ブーブー（自動車）」とか、「カチャ（傘）」などのような一語であるが、「自動車が通っている」「ここに傘がある」「傘を見つけた」と言うような文の機能を持っているので、一語文と言われている。

この時期には、その他意味の無い音のつながり、イントネーションは大人の文的な長い「お話」がみられる。また、似た音のものに同じ語を用いる般（汎）化の現象もみられ、幼児なりの言葉の世界を作り始める。

次の段階としては、1歳半前後において、これまで使用している一語文の一語が、述語部分になったり、述語の対象（目的語）の部分になったりして、二語結合文が使用されるようになる。たとえば、「ない」に「ブーブー」が結合して「ブーブー ナイ（おもちゃの自動車がここにはない）」とか、「ハナ（花）トッテ」とか言うようになる。場面がなくても言葉を聞いただけで、要求であるか叙述であるかが聞き手に判断できるようになる始まりである。しかし、これも最初は二語の羅列で、2歳前後になって、関係をあらわす助詞「が」「も」「に」などが使えるようになり、三語文・四語文も出てくる。

さらに、2歳近くになると「コレ、ナニ？」という二語文をやたらと使い、物の名前を知りたがるという時期がどの子にも現れ、命名期とも言われる。この時期には付属語の使用も始まり、格助詞「が」「に」「の」が使えたり、係助詞の「は」「も」を使えるようになる。その他、「タイ」「ナイ」等の助動詞や「テイル」「チャウ」等の補助動詞、および動詞の活用「行く」を「イコウ」、「言う」を「イワナイ」と言うことができるようになる。つまり、過去のこと、自分の意志、否定、願望などの表現ができるようになる。二歳後期になると三語文以上を使えるようになる。また、「ドウシテ」を使った質問文も使え、接続詞を使った長い文を使うようになり、800から1000語近くを使うようになるが構文能力は十分はでない。

こういった言葉やそれに対応する概念の獲得がどのようにして行われているのか次に考える。

言葉や概念の獲得の仕方は、その概念の種類や個人により様々であるが、およそ以下の4つの方法のいずれか、あるいはそのいくつかの組合せを通じて獲得してきたと考えられている。⁵⁾

(1) 定義による概念獲得

辞書や教科書や父母や教師などに直接概念の定義を与えられて獲得する。

もちろん、定義を与えられても録音テープのようにそれをそのまま暗記しただけでは、定義検査には合格できても、概念を獲得したとはいえない。定義文中に用いられている概念をすでに獲得しており定義文を理解できることが必要である。

(2) 事例による概念獲得

「かわいいワンワンだね」、「これはワンワンじゃないの、シカさんていうんだよ」などと、「ワンワン」の概念は個々の事例を通じて獲得される。個々のワンワンをワンワンの正事例 (positive instance)、シカさんのようなワンワンでない事例を負事例 (negative instance) という。

(3) 言語文脈による概念獲得

「ヒューマニズムに貫かれた生涯」とか「ヒューマニズムあふるる感動巨編!」などのようなその概念の用いられる言語文脈から理解したり誤解したりして持っている概念。

(4) 概念の名による複合概念の推察

複合概念を表わす語は、その複合概念を構成する概念の名を組み合わせで作られることが多い。このような概念については、直接その複合概念を確かめることをせず、「プロダクションシステム」だとか「複合概念」だとか「自由民主党」などというその名から推察して理解したり誤解したりすることも多い。

本研究で対象とする概念は主に (2) の事例による概念獲得である。

3. 事例による概念獲得⁷⁾

正事例と負事例とに分ける基準の学習の過程及びそれに作用する変数の研究は、弁別学習実験の手続きを用いて行なわれてきている。ここでは事例の構成のしかたや、試行ごとの分類反応を求めるかどうかなど実験手続きには色々のヴァリエーションがあり得るが、この手続きのいくつかの基本的な特徴を指摘しておく。

(1) 正事例と負事例を分類できることは概念の獲得の重要な指標であり、共通反応を要求する事例と他の事例とを弁別させる手続きは、概念獲得の問題であるといえる。いったん概念を達成すれば、新しい事例に対しても適切な分類反応が出来るという点でもこの課題は概念行動の要請に合致している。

(2) この手続きは、個々の事例についての情報から分類の基準を学習することを要請する。その意味で、この手続きは経験による概念獲得の課題であり、とくに概念達成であるといえる。

(3) 事例の構成のしかたには色々のものが考えられるが、新しい事例にも適用可能な概念を達成させるために、事例のなんらかの性質に基づいて分類できるような概念が設定される。

(4) 事例の組の構成を工夫することによって、人工的ではあるが被験者に対して一種

の世界（環境）を提供し、その構造を各様に変化させることができる。

（５）事例を継時的に与えることによって、思考過程を時間的に引き延ばし、それぞれの時点における被験者の反応を分析して、情報処理過程を明らかにする可能性を開く。

（６）弁別学習実験の手続きを採用することによって、学習心理学における弁別学習研究の成果を利用することが出来る。弁別学習実験は、タコやネズミなどの動物から幼児、成人まで色々の種や色々の発達段階にある被験体に、基本的には同じ手続きで実施できるので、概念獲得の機制やその能率の系統発生的または個体発生的比較に都合がよい。とくに、弁別学習の転移の問題については、弁別の移調や弁別移行学習及び学習セットの研究など、動物の弁別学習実験から始まった問題で概念の獲得に直接かかわる問題が多い。

（７）言語理解などに関しては、正事例だけからなされるとされている。負事例を与えることにより、効率的な学習が達成され得るが、多くの負事例から有効な事例を選択することは難しい。

４．本システムの概要

本システムでは、具体的に人間の幼児がどのような情報によって概念形成を行っているのか考察し同様の過程を計算機上で実現することを試みた。

人間の場合、いくつかの感覚器を単独であるいは組み合わせて使用し、外部からの刺激を感じてそれらの情報が脳に伝えられ様々な概念を獲得していると考えられる。その中でも特に、事物の名前等を学習するためには視覚と聴覚が最も重要な役割を果たしているであろうことは容易に想像がつく。そこでこの視覚と聴覚によって得られる情報、つまり音声と画像の情報をを用いて計算機に物の名前や位置等の概念を学習させるシステムを作成した。つまり、ある物を表現する画像があったとするとその画像を説明する文を音声によって与えることにより、逐次画像上の形状・色・大きさ・位置といった概念に対応する音声言語を獲得することが本研究の目標である。但し、物の名前や位置等の概念を単語として与えるのではなく、簡単な文の音声データとそれに対応する画像データとを用いて、形状・大きさ・位置・色等の概念を形状することとした。このことから、画像同士の類似性の自動判定・音声同士の類似性の自動判定・画像と音声の対応付け等の機能が基本操作となる。

５．画像からの図形データの抽出

カメラで取り込んだ図形画像から、直接概念を形成することは、難しい。そこで、概念形成に必要とされるパラメータを図形画像から抽出しなければならない。

今回の実験で形成する概念は、以下の５つの概念グループに分類される。

- (1)図形の存在
- (2)図形の位置
- (3)図形の大きさ
- (4)図形の色
- (5)図形の形

以下では、これらの概念、およびその概念を形成するために求めるパラメータについて

述べる。

なお、今回の実験では入力画像として 256×195 画素からなる256階調のモノクロ画像を用いる。

(1) 図形の存在性

図形の存在性の概念には、入力画像中に図形がある、図形を画像中に出す、あるいは図形を取り去る、といった概念が含まれる。これらの概念とマッチングを取るためのパラメータとして、画像中の図形の数を求める。

画像中に図形が存在するかどうかを判断する方法として、その画像に、2値化に適切な閾値が存在するかどうかを調べる。具体的には、判別分析法で求めた閾値の良さ（クラス分離度）[1]が、ある値より大きければ、図形が存在するものとする。図形が存在すれば、求めた閾値により2値化処理を行った後、ラスタースキャンにより図形を探す。

(2) 図形の位置

この概念グループには、図形の絶対的な位置の概念、すなわち、画像中の右、左、あるいは上や下と、相対的な位置の概念、つまりある図形の右だとかいう概念が含まれる。この概念とマッチングを取るために、2値化した図形画像より、図形の重心座標、すなわち図形を形成する画素の座標の平均(μ_x , μ_y)と、その画素の分散(σ_x , σ_y)をパラメータとして求める。

二つの図形間の位置に関するマッチングは、各図形の画素が、x方向に(μ_x , σ_x)、y方向に(μ_y , σ_y)で正規分布しているとみなし、それぞれの方向での、二つの図形の正規分布の重なりを求めることより行う。ある程度以上重なっていれば同じ位置にあるとし、同じ位置にない場合の相対的な位置関係は、二つの図形の重心座標の差により求める。

(3) 図形の大きさ

このグループには、図形が大きい、普通、小さいという概念が含まれる。そこで、パラメータとして、図形の絶対的な大きさ、つまり、2値化画像での図形の占める画素数（面積）を用いる。

図形間のマッチングは、2図形の面積の比率をとることにより行う。

しかし、例えば人間には四角形と三角形が視覚的にはほぼ同じ大きさに思えても、絶対的な面積ではかなり異なるような場合が考えられるが、今回の実験では考慮していない。

(4) 図形の色（グレースケール）

図形が黒い、または、灰色だ、という概念グループである。このグループに対しては、図形のグレースケールの平均値と、その分散をパラメータとする。

これらの値は、判別分析法で、2値化閾値を決定するときに求めた値を流用する。

図形間の色情報のマッチングは、位置のマッチングと同様に、正規分布の重なりを求めることにより行う。

(5) 図形の形状

図形の形を表す概念グループであり、今回は丸、三角形、四角形が含まれる。図形の形状を表すパラメータとして最適なものはまだ決定していない。図形の形状を捕らえる方法

として、図形の輪郭線を用いるものと、図形の面積的な要素から形状を表すものの2つが考えられる。

輪郭線を用いるものとしては、輪郭線をフーリエ変換したり、輪郭線をスプライン曲線で表したり、あるいは、輪郭線を直線要素と曲線要素に分解して表すものなどが考えられる。

図形の面積的な要素を用いるものには、図形から最大内接円を次々と取り除いてゆき、その図形を最大内接円のリストとして記述するものがある[2]。

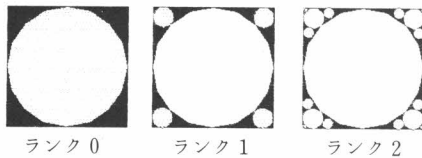
今回は図形を面積的に捕らえ、図形を形態学的に分割することにより表す方法を選んだ。この方法の利点は、図形の移動、相似変換や回転に対して強いデータが得られることである。ここでは、図形間の形状比較のために、それぞれの最大内接円にランクをつけ、そのランクを最大内接円リストに加える。

ランク付けは、まず、最初に求めた最大内接円のランクを0とし、以後は、求めた最大内接円が、より大きい(すでに取り除いた)最大内接円に接しているとき、その大きい内接円のランク+1を求めた最大内接円のランクとする。複数の円と接している場合は、その円の中の最大のランク+1とする。これにより、図形の大まかな構造を求める。

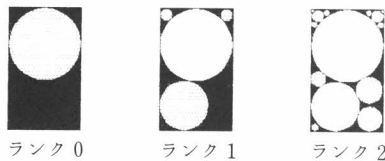
このように、ランクとは、その最大内接円の、構造的な重要度を示し、ランクが大きくなるのに従い、基本構造から細部構造を表す最大内接円となる。



(a) 正三角形



(b) 正方形



(c) 長方形

図1. 最大内接円による図形分割

図形間で形状のマッチングを取るには、まずランク0の最大内接円の数を比較し、次いでランク1の最大内接円の数、そして、ランク2について比較する。これで、図形の構造が比較できる。そして構造が同じであれば、個々の最大内接円について比較を行い、最終的な同型度を求める。基本構造が異なれば、図形の形が違うものとみなす。

今回は、丸、三角形、四角形といった単純な図形を対象としたので、ランクを設けただけで形状特徴が表せそうだが、より複雑な図形を対象とする場合には、ランクだけでなく、最大内接円間の隣接関係を考慮する必要がある。

また、長方形のような場合には、単純な円盤ではなく、図2のように円盤の軌跡で表現すべきなのだが、今回はそこまで行っていない。

・画像データ

量子化された4次元特徴量

(Y_1, Y_2, Y_3, Y_4)

$Y_1 \dots$ 形状に関するデータ

屈曲点数+1

y_2 ・・・大きさに関するデータ

1 (中), 2 (大), 3 (小)

y_3 ・・・色に関するデータ

1 (白), 2 (黒), 3 (灰色)

y_4 ・・・絶対位置に関するデータ

1 (中心), 2 (右), 3 (左), 4 (上), 5

(下)



現状



望ましい図形分割

図2. 望ましい図形分割

6. 音声データからの音声情報の抽出⁸⁾

前章で述べた画像データに対し、その画像に対応する音声から対応する音声情報を抽出する必要がある。このためには二つの画像間に類似性があると、それに対応する二つの音声間にも類似性があるとの反応に基づき、ここではDPマッチングによる二つの音声同士の類似区間の抽出を行った。

方法としては、二つの音声(文あるいは文の一部区間)の時系列データ同士のDPマッチングを行い、それによって算出された最適照合パスおよび照合距離により、類似区間を抽出する。

ここで、一番目の音声の時系列データを

$A = a_1 a_2 \dots a_i$, 二番目の音声の時系列データを $b_1 b_2 \dots b_j$, a_i と b_j のフレーム間距離を $d(i, j)$ とし、対称形のDPパスを用いると A と B の累積照合距離 $D(I, J)$ は次の漸化式で求められる。

$$D(i, j) = \min \begin{cases} D(i-1, j) + d(i, j) \\ D(i-1, j-1) + 2 * d(i, j) \\ D(i, j-1) + d(i, j) \end{cases}$$

但し今回は発声全体に対するパスの傾斜制限は用いなかった。

この方法で最適照合パスが求められると、音声の共通区間同士のパターンは区間長もほぼ同じで時間軸の伸縮があまりなく、しかも照合距離が比較的小さいと期待できるので、以下の3種類の基準により共通区間の抽出を行った。

(1) DPパスの線形性

(2) 平均距離

(3) 距離の最大値

ここで、「線形性」については、DPパスが i 方向、 j 方向に連続して4フレーム以上進めば共通部分でないとした。「平均距離」および「距離の最大値」はこれらが一定のいき値を越えるものは共通部分でないとした。

このようにして得られた共通区間を以下のアルゴリズムにおける共通区間として用いる。

7. 概念の獲得アルゴリズム

ここでは、3., 4. で述べた音声データと画像データから概念を獲得していくアルゴリズムについて述べる。また、画像と音声の対応づけができたものを辞書部として登録していくものとする。

以下、2つの文音声をマッチングし差異と共通部分を抽出した場合、それぞれの部分を特徴部、共通部と呼ぶ。また、画像と音声の対応づけができたものを辞書部として登録していくものとする。

・記号の定義

- R_{sen} ・ ・ 参考文・画像の番号
- I_{sen} ・ ・ 入力文・画像の番号
- R_{num} ・ ・ 辞書に登録されているデータ数
- $Re(L)$ ・ ・ L番目のデータ出現度数
- Word ・ ・ 辞書中の参照データ番号

ステップ1.

初期設定：参照文 $I_{sen} = 2$, $R_{num} = 0$ とする。

ステップ2.

$R_{sen} = I_{sen} - 1$ とし、 R_{sen} と I_{sen} つまり、ある文とその一つ前の文の音声データをマッチングし特徴部・共通部を抽出する。また、画像データを各概念を表す次元毎に、5. で述べた画像の特徴を量子化したもの同士をマッチングし音声の特徴部対応する次元(y_i)を求め、それ以外の次元($y_j, j \neq i$)を-1にクリアしたものを、その音声の特徴部と対応するデータとして割りあて、残りの次元($y_j, j \neq i$)を共通部に対応するものとする。もし共通部が全く抽出されない場合、ステップ4へ。

特徴部・共通部をそれぞれ対応する画像データと共に登録し、 $R_{num}, Re(word)$ をインクリメントする。

ステップ3.

$I_{sen} = I_{sen} + 1$ とし、辞書に登録された音声部と、画像パラメータの有効次元数の少ないものから順に、 I_{sen} の音声部とマッチングする。マッチングが取れたら、取れたものの画像パラメータにおいて、画像パラメータの対応する次元をクリアし、 I_{sen} の入力音声からマッチングが取れた音声データを削除する。但し、この際、画像パラメータの有効次元のデータ内容が一致せず削除できないときは、辞書中の登録データを、一致しない次元のパラメータは冗長なものとして、-1にクリアし訂正した上で削除する。これを全ての辞書と照合するか入力音声データがそれ以上削除できなくなるまで続ける。一定時間長以上の音声区間が一つでも削除されなかった場合は、ステップ2へ。

残った音声部・画像パラメータの次元を対応させて、辞書中に登録し、 R_{num} をインクリメントし、出現度数をカウントする。

ステップ4.

$I_{sen} < I_{max}$ ならばステップ3へ

上記のアルゴリズムに従って処理される過程を例示する。但し、入力文は以下の通りとする。

1. 丸があります。

2. 三角形があります。
3. 四角形があります。
4. 大きい丸があります。
5. 大きい三角形がある。
6. 小さい三角形がある。
7. 右に丸があります。
8. 左に三角形がある。
9. 上に四角形があります。
10. 下に丸があります。
11. 黒い丸がある。
12. 白い丸がある。
13. 左に白い丸があります。
14. 右に小さな三角形があります。
15. 下に黒い四角形がある。
16. 下に大きい丸があります。

ステップ1. 初期設定

ステップ2.

文1と文2の音声データをマッチングし特徴部・共通部を抽出する。

丸があります。 (1 1 1 1)

三角形があります。 (4 1 1 1)

特徴部として、音声データの「丸」・「三角形」が、共通部として「があります」に相当する部分が得られたとする。また、画像データとの対応づけで

「丸」 (1 -1 -1 -1)

「三角形」 (4 -1 -1 -1)

「があります」 (-1 1 1 1)

となる。この3個のデータが辞書に登録される。

ステップ3. 処理無し

ステップ4. ステップ3へ

ステップ3.

$I_{sen} = 3$ で、既に辞書中に3個のデータが登録されているので、これらの音声部と文3の音声をマッチングすると、「があります」に相当する部分で照合がとれる。この部分を削除すると

「四角形」 (4 -1 -1 -1)

が残りとして、得られるのでこれを登録する。

ステップ4. ステップ3へ

ステップ3.

$I_{sen}=4$ で、「大きい丸があります」のデータ(1 2 1 1)から、辞書部の音声と「丸」の部分が照合され、「大きい・・・があります」のデータとして、

(-1 2 1 1)

が残る。次に、登録データの「があります」の部分と残りの音声の一部と照合がとれるが、登録データの画像パラメータ

(-1 1 1 1)

と比較すると y_2 が一致しないので「があります」の登録データからこの次元を冗長なものとしてクリアし、

「があります」 (-1 1 1 1)と修正した上で削除すると

「大きい」 (-1 2 -1 -1)が残り、辞書に登録される。

このように繰り返していくと、1から12までの文と画像パラメータを上記のアルゴリズムで処理する事により、音声のマッチングが正しく行われた場合、以下の13種類の音と対応する画像パラメータが得られる。

「丸」 (1 -1 -1 -1)

「三角形」 (4 -1 -1 -1)

「四角形」 (5 -1 -1 -1)

「大きい」 (-1 2 -1 -1)

「小さい」 (-1 3 -1 -1)

「右に」 (-1 -1 -1 2)

「左に」 (-1 -1 -1 3)

「上に」 (-1 -1 -1 4)

「下に」 (-1 -1 -1 5)

「黒い」 (-1 -1 2 -1)

「白い」 (-1 -1 1 -1)

「があります」 (-1 -1 -1 -1)

「がある」 (-1 -1 -1 -1)

なお、本アルゴリズムは画像データの抽出ミスに対しては考慮していないため、画像データ抽出に当たってはいき値を最適化し希望する画像の特徴抽出データが得られるようにした。また音声区間の類似区間の抽出ミスに対しては、全ての概念が複数回出現するようにしたため、一度辞書に正しく登録された音声部に対しては、たとえそれ以降、本来あるべき共通区間が検出されなくても、問題が生じないようにした。しかし、学習の初期の段階で検出ミスが生じると希望する概念が獲得できないため、更にアルゴリズムを改良する必要がある。

8. 概念獲得実験

8.1 実験方法

実験は成人男性話者3名が7章で述べた16文を各1回ずつ発声したものをを用いた。特徴量としてはLPCメルケプストラム係数を用いた。

実験では、アルゴリズムのステップ2ならびにステップ3において、理論上期待される概念が獲得されるか否かをデータベースの評価を兼ねて調べ、最終的に辞書に登録されたものを確認する方法を用いた。画像の特徴パラメータ抽出に関しては前節で述べたようにしきい値を最適化したのですべて正しく抽出できた。また、音声データに関しても、共通区間検出の際のしきい値を最適化したため、1名の話者の2単語を除いて共通区間を正しく検出できた。

8.2 実験結果

今回用いた3名分の音声に対しては、1から16までの文によって、期待した理論通り、それぞれ13種類の概念が獲得できた。しかし、個々の獲得過程を調べると、1名の話者のデータにおいて、文5の「三角形」と13の「白い」が、辞書のものと音声部のマッチングが取れていなかった。しかし、本アルゴリズムでは入力音声に対して、共通部分が検出されなかった場合概念の獲得を行わない、つまり負事例による学習は行わないため問題は生じなかった。ここではこれらの単語は他の入力音声により適切な学習が行われ、最終的に目的とする概念が獲得できた。

8.3 問題点と今後の検討課題

今回は画像のデータをパラメータ化する際に、形状を表すパラメータとして屈曲点の個数という情報を使ったが、これには「屈曲点」という概念を人間が既に獲得している事を前提としてしまっているために、0からの概念獲得という点で問題が残った。この点で、画像の形状をパラメータ化する何らかの別の方法が必要であると思われる。

また、今回は学習させる文の順序をあらかじめ注意深く決定したために、最終的には期待通りの概念が獲得できたが、別の予備実験から初めに文2と3をもってきて実験を始めると、「三角形」「四角形」の概念が「三」「四」「角形があります」となり、獲得できない事がわかっておりこの点への対処も必要となるであろう。さらに、途中で期待通りの概念が獲得されなかった場合の事を考え、それぞれ最低2回以上学習するようにしたがこのような事は必須条件であると思われる。

9. ま と め

本研究では、視覚と聴覚という2つの外的刺激を結びつけて未知の入力に対する概念の獲得方法について検討し実験を行った。

その結果、逐次画像と音声を対応付けながら学習を行う事により、ある概念に対応する音と画像上の特徴とが対応付けられ、人間の概念獲得手法の初期段階が計算機上で実現できた。

但し、今回の実験では図形が1つに限られる点や、図形の概念を屈曲点の数といった形で扱ったためにあらかじめ「屈曲点」という概念を何も知識を持っていない幼児がイメージとして形状の区別に使えるかどうか疑問が残る。

今後、図形から抽出する特徴パラメータ、および複数図形の取り扱い、またその際の相対位置の問題等の解決をしていく必要がある。

参 考 文 献

- [1] 大津展之
“判別および最小2乗基準に基づく自動しきい値選定法”
電子通信学会論文誌
'80/4 Vol. J63-D No.4 p349-356
- [2] Ioannis Pitas & Anastasios N. Venetsanopoulos,
“Morphological Shape Decomposition”,
IEEE Trans. Pattern Anal.
Machine Intell.,
Vol 12, No.1, pp38-45, 1990.
- [3] 大須賀節雄, 佐伯 胖
“知識の獲得と学習”
オーム社, 知識工学講座 3
- [4] 大須賀節雄
“知識情報処理”
オーム社, 知識工学講座 1
- [5] 大久保愛
“幼児言語の研究”
あゆみ出版
- [6] 村田孝次
“言語発達研究”
培風館
- [7] 坂本昂
“現代基礎心理学 7”
東京大学出版会
- [8] 兵後・中川
“連続発生された二発話文間のDPマッチングによる共通部分の検出”
電子情報通信学会全国大会講論集(1989.3)