

顔画像データ検索時の検索特性について

尾 田 政 臣

A T R 視 聴 覚 機 構 研 究 所

画像検索システムでは、検索者が求める検索対象をどのようにシステムに伝え検索させるかが検索の容易性を確保するための重要な課題である。この課題に対して種々の試みが行われている。ところが、それらの方法では検索イメージが検索前にしっかり固まっていることを前提としていた。筆者は、検索前には漠然としたイメージしか無い場合についての画像検索方法を提案した。本稿では、提案するシステムに対して検索実験を行いその有効性を明らかにする。検索実験では、検索回数、検索されたデータの品質評価、検索時間について調べた。約6万個の顔データの中から望ましい10個の顔データを検索するのに200~300個の顔データを評価することで十分なことが分かった。

Retrieval Properties of Humans as to Facial Image Retrieval System

Masaomi ODA

ATR Auditory and Visual Perception Research Laboratories

2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02, Japan

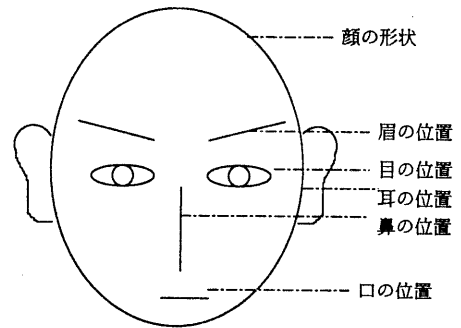
How to transmit the retriever's intention to the system will be the most important problem in constructing an easy-to-use and efficient interface for an image retrieval system. Many studies have focused on ways to improve the transmission; however, all of them assumed a fixed target image in the human brain before retrieval. The previous paper proposed a new retrieval mechanism for an image database. A system with this mechanism is effective for a retriever who has no concrete target image before retrieval. This paper investigates the efficiency of the new mechanism. The evaluation of only 200~300 faces is sufficient for finding the 10 preferable faces from 60-thousand-face data.

1. まえがき

画像データベースの検索は、テキスト情報の検索と比べ言語的な対応が取りにくいいため、その検索には困難が伴う。それを解決すべく、言語にあいまいさを許したり、略画を用いるなど種々の検索法が提案されている [1-9]。それ等に共通することは検索者の頭の中には検索すべき対象のイメージが検索前に存在し、それ等と照合しながら検索を進めるモデルを想定していた。

しかし、将来的には種々の画像データベースが構築されて行くと、上記モデルに当てはまらない検索法が求められたり、当初からそのような検索を目的としない画像データベースの構築が求められる可能性がある。例えば、カタログデータベースやデザインに関するデータベースの場合を考えて見る。我々が物を購入する場合の意思決定過程を思い浮かべると、明確に購買する対象が決まっている場合と、ウインドウショッピングや通信販売カタログを見ているうちに購買対象のイメージが固まる場合があることに気が付く。また、デザインに関するデータベースでは、利用開始当初は新たにデザインした意匠が既に登録されていないか確認するための手段として利用されていても、データの量、質が向上するに従って、新たなデザインをする場合の支援ツールとして利用される可能性が高まることが予想される。このような利用を行う場合には、検索前に検索者の求めるイメージが固まっていることを前提とする検索方式では、検索が困難になる。

そこで筆者は、検索対象が明確でない場合にも検索が容易な画像検索システムを提案した [10]。それは、人間が複数の画像から概念を形成する過程が次の概念形成過程に影響することを示した心理実験結果 [11] を利用したものである。線画による顔画像検索システムを例として、検索当初において検索すべきイメージが明確でない場合にも容易に好みの顔画像を検索可能であることを示した。本稿では、さらに検索回数、検索時間、検索されたデータの品質等について調べた新たな実験結果をもとに、検索システムの有効性について検証する。



特徴パラメータ値

(顔の形状,眉の位置,眉の形状,目の位置,目の形状,
鼻の位置,鼻の形状,口の位置,口の形状,耳の位置)

(0, 0, 1, 0, 0, -1, 1, -1, -1, 0)

図1. 顔画像の例

2. 検索システムの概要

本検索システムは、検索前に検索すべき画像のイメージがはっきりしない場合や、言語的に表現しにくい画像を検索する場合にも効率的な検索が可能か否かを検証することを目的に構成されている。その目的に合うデータとして線画の顔画像を用いた (図1)。線画を用いる利点は、データベースの構築時にデータの収集及びシステム作成の点で容易になる他、顔の類似性や特徴点など実験条件の統制及び結果の解析が容易になることである。また、顔画像は、好みの顔のイメージを問われたときに、必ずしも明確なイメージがなく言語的にも表現しにくいデータであることから実験の目的に合うと判断し採用した。欠点は、実用データベースでは写真等が用いられる場合が想定され、線画による実験結果が利用できるかどうか検証が必要になることであろう。しかし、線画による実験も原理的な面で十分有効な結果をもたらすものと考えられる。実用システムについては種々の要因が加算されるので、その上でさらに個々に対処する必要があるだろう。

本システムでは、顔の形状、眉の位置、眉の形状、目の位置、目の形状、鼻の位置、鼻の形状、口の位

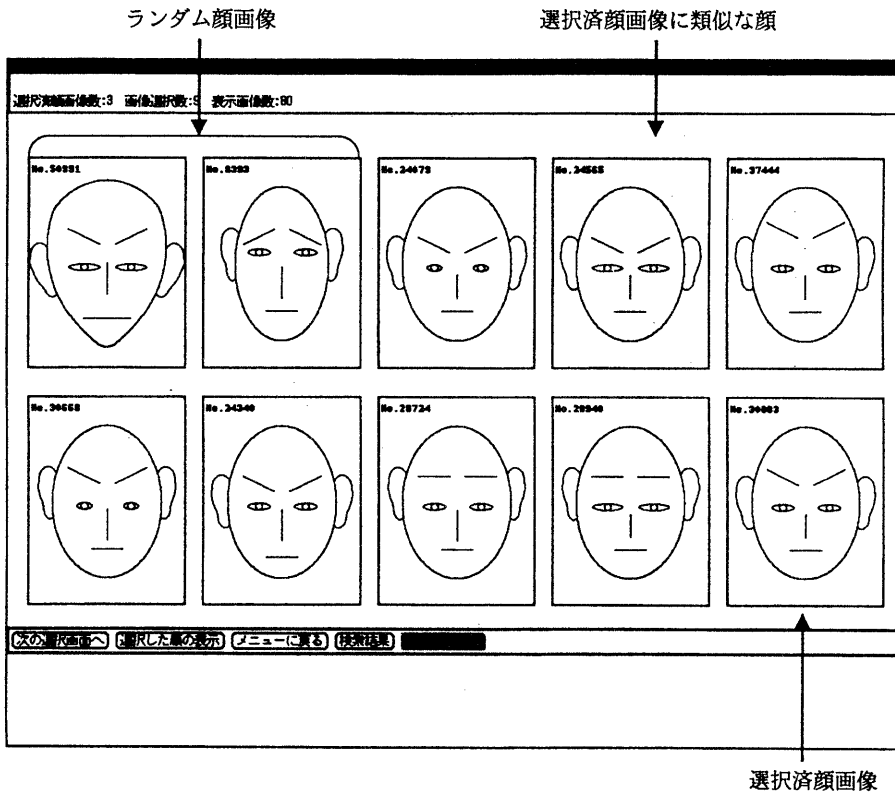


図2 選択候補データの表示方法

置、口の形状、耳の位置と10個の特徴をパラメータとした。各パラメータは、口の形の大、中、小、や目の位置の下、中、上のように3値(-1、0、1)のうちいずれかを取る。各パラメータの3値の全ての組み合わせを取り、約6万個の顔画像を生成し、データとしてシステムに格納してある。

本システムは、10個のデータを随時ディスプレイに提示し、その中から検索者の求めるデータに近いものを選んでもらい、漸次求めるデータのイメージを固めつつ、最終的に求めるデータを決めるように検索を進める。その過程を効率的に行うために、システムは次に示す特徴を有する。

- (1) 検索者は10個の顔画像を同時に見比べ評価する。
- 相対評価により、望ましいデータの選択を容易化できる。
- (2) 選択済の画像に類似の顔が7個表示される。

それまでに検索者が選んだ好みの顔画像に類似の顔をシステムが代行検索し、検索者に提示する。この過程自身が望ましい顔画像のイメージを形成する上での文脈をなしており、この文脈に沿った形で望ましい顔画像のイメージが作られて行く。

- (3) 選択済の顔画像の中から1個が再評価のため表示される。

望ましい画像が次々に選択されて行った場合、選択された画像について再評価する必要が出てくる。そこで、選択済顔画像の中から、ひとつをランダムに選び10個ずつ提示する顔画像の一つとして提示し、再評価させる。再評価の結果、再び選択されなければ、その時点では望ましいデータと判断されなかったものとして選択済画像から除く。

- (4) 検索途中での好みの変化に対応できるよう、ランダムに選んだ2個の顔画像が表示される。

選択済顔画像に類似の顔画像だけを提示していた

表1 検索実験結果

| 実験項目 | 検索数 | | 選択数 | | 検索時間 | | F 値 | | 最良の顔と平均顔との比較 | |
|--------------------|-----|-------|------|------|--------|-------|------|------|--------------|------|
| | 平均値 | 標準偏差 | 平均値 | 標準偏差 | 平均値 | 標準偏差 | 平均値 | 標準偏差 | 自乗誤差 | 標準偏差 |
| 文脈依存メカニズム 有り1回目 | 205 | 114.9 | 40.4 | 19.5 | 6'31" | 3'41" | 15.2 | 5.1 | 1.24 | 0.57 |
| 文脈依存メカニズム 有り2回目 | 436 | 125.5 | 73.2 | 10.9 | 13'16" | 4'10" | 16.1 | 5.1 | 1.57 | 0.60 |
| 文脈依存メカニズム 無し1回目 | 377 | 292.5 | 14.7 | 2.4 | 5'39" | 1'54" | 44.8 | 7.1 | 4.50 | 1.15 |
| 文脈依存メカニズム 無し2回目 | 793 | 573.2 | 25.5 | 5.1 | 10'29" | 5'23" | 43.4 | 5.7 | 4.12 | 0.86 |

のでは、文脈に依存して一定の方向に求める顔画像の絞り込みが進む危険がある。即ち途中でイメージの変更を行おうとしても、ある段階まで絞り込みが進むと他の特徴への変更は非常に困難になる。このため、システムがデータの中からランダムに2個の画像を選び出し、10個の提示画像のうちの2個として提示する。

(5) 任意の時期に、選択済の顔画像を再表示し、再評価できる。

選択済顔画像は最大10個を限度に保存する。10個以上選択された場合は、古いものから順に捨てる。検索が進んだ段階で選択済顔画像の再評価が必要になる。一度に、全ての選択済顔画像の表示を可能とし、相対的な評価に基づいて容易に望ましくないデータを削除可能とする。

検索の開始は、検索者がマウスでボタンを押下することにより始まる。システムは、それを契機にランダムに選び出した10個の顔画像を提示する。検索者はこれ等の画像を評価し、望ましいものをマウスで任意の個数選択する。システムは、この選択された顔の特徴の平均を求め、さらにこの平均的顔画像に類似の顔画像をデータベースから7個選ぶ。また、データベースから2個の顔画像をランダムに選び、さらにそれまでに選んだ顔画像のうちから1個をランダムに選びそれ等を合わせて、次の選択候補データとして図2に示すように表示する。検索者は表示された10個の中から望ましい顔を選択する。このようにし

て、検索が続けられる。

システムは、選択済顔画像に類似する画像をデータベースから求めるに当っては以下の式に基づいて計算を行う。

$$S_j = \sum_{i=1}^N \alpha_i (\bar{x}_i - x_{ij})^2$$

S_j : 画像データベース中の任意の画像jと選択済顔画像の平均値との類似度

α_i : 重み係数

\bar{x}_i : 選択済み顔画像の特徴パラメータ値の平均値

x_{ij} : 任意の画像jの特徴パラメータiの値

N: 特徴パラメータの数 (10)。

S_j の値が最も小さい顔画像データが、平均値に最も近いデータと見なす。ただし、データの検索過程では計算量の低減を図るため次の方法をとる。

特徴の平均値を求め、その中で小さいものから順に3個のパラメータを求める。それを検索者が最も注目しているパラメータと考え、全データのうち3個の特徴を満足するデータを類似性計算の対象とする。なお、重み係数 α_i は今回の実験では1に固定した。

本システムは、SUN - sparc2上に構築され、10個の顔画像の表示は、1.4秒程度で完了する。したがって、検索者は殆ど待ち時間なしで検索を実行できる。

3. 検索実験方法

学生被験者10名を、検索者として実験を行った。

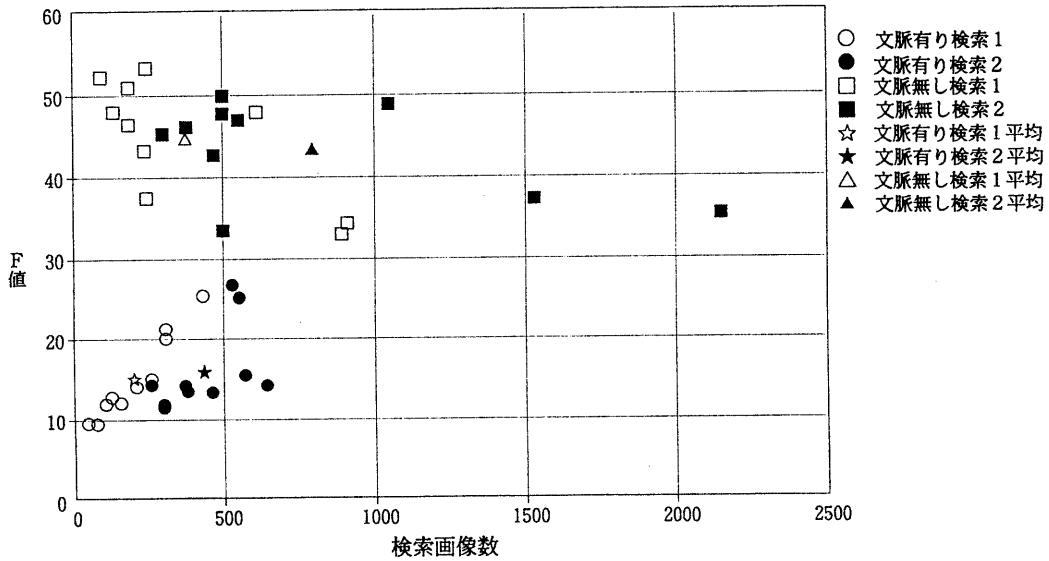


図3 顔画像検索における検索画像数とF値の関係

この中にはマウス操作の未経験の者もいるので、画像の検索操作の説明の後、2~3分間程度練習を行わせて操作に支障がなくなった段階で実験を開始した。被験者には、システムの中に約6万個の顔画像があり、その中から自分の好みの顔データを10個選び出し、満足の行く10個のデータが揃った段階で検索終了を告げるよう教示した。その時点で一度検索を中断し、それまでに要した検索画像回数、選択した画像数、検索時間、選択した顔データを集計した。さらに、選んだ10個の中で最も好ましいデータを1個選んでもらった。その後、選択顔画像の品質をもっと望ましいものにするよう検索を続け、望ましいデータの出現が期待できないか、またはデータの改善が進まないと判断した時点で検索終了を告げるよう教示し、検索を再開した。この2回目の検索結果についても、第1回目と同様のデータの集計並びに最も好ましいと思うデータの選択を行ってもらった。

文脈依存メカニズムを利用したシステムとの比較評価のために、常に10個のデータをランダムに表示し被験者に選ばせる、文脈依存メカニズム無しの実験を行った。文脈依存メカニズム無しの実験は文脈依存メカニズム有りの場合と操作上は全く同一である。

実験は、同一被験者に文脈依存メカニズム有りと無しを1回ずつ行わせた。ただし、システム操作のなれ等による偏りを除くため、被験者を文脈依存メカニズム有りを先に行うグループと文脈依存メカニズム無しを先に行うグループの2群に分けた。

4. 検索実験結果

文脈依存メカニズム無しの場合と有りの場合の検索結果を表1に示す。また、各被験者についての検索画像数とF値の関係について図3に示す。表に於て、1回目、2回目は各々第一回目の検索結果と引き続き行われた第2回目の集計結果を示す。F値とは、最終的に選ばれた10個の顔画像の品質を評価するために導入した値である。

$$F = \sum_{i=1}^N \sum_{j=1}^M (\bar{x}_i - x_{ij})^2$$

N: 特徴パラメータの数 (10)

M: 選択済みデータの数 (10)。

F値が小さいほど類似の顔が求められたことになる。F値は常に10個の選択済み顔画像で評価しているが、最も類似した顔画像が求められた場合のF値は8.1となる。また、全くランダムに画像を選んだとすると、F値の予測値は66.7である。

4.1 検索画像数

表1に於て「検索数」は何個の画像を検索し眺めたかを示している。したがって、検索数が少ないほど検索が容易なことを示す。また、一度に表示する画像数が10であるから、「検索数」を10で割った数が画面が変わって新しい顔画像の選択候補が表示された回数である。

文脈依存メカニズム有りの1回目の検索数は平均で205回で文脈依存メカニズム無しの1回目は377回である。文脈依存メカニズム有りは無しに比べ約1/2程度の検索数となった。

4.2 選択画像数

表1に於て「選択数」は、検索者が表示された画像から好みの画像として何個選んだかを示す。第1回目の検索結果の比較では文脈依存メカニズム有りの場合が40.4、文脈依存メカニズム無しの場合が14.4と文脈依存メカニズム有りの場合が無しに比べ約3倍になっている。この理由は、第一に文脈依存メカニズム有りでは再評価用の画像が必ず1個表示されており、これが望ましいと思う時は必ず選択しなければならないため、選択数が文脈依存メカニズム無しより増加する。第一回目の集計で比較すると検索数の差が、約170であるから毎回選択動作を行うと仮定すると文脈依存メカニズム有りの場合が無しの場合に比べ選択数は17回多くなるはずである。しかし、実験結果ではその差は25.7となっており、この理由だけでは説明しきれず他の原因が求められる。実際の検索状況を眺めると文脈依存メカニズム無しではなかなか望ましい顔画像が出現しないため、選択の回数が少なくなっているように思われる。これに比べ文脈依存メカニズム有りでは、それまでに選んだ顔画像より望ましい画像が次々出現するため、選択回数が上がる。これ等の理由により文脈依存メカニズムの選択数が多くなったと考えられる。

これを裏付けるように文脈依存メカニズム有りの場合は205個の画像に対し40.4個の画像が選ばれている。したがって、5.1個の顔画像につき1回の選択率である。一方、文脈依存メカニズム無しの場合、

377個の画像に対し14.7個の画像が選ばれたのであるから、25.6個につき1回の選択率となっている。これ等の結果は、検索者は文脈依存メカニズム有りでは好ましい顔を次々に選び望ましい顔のイメージを頻りに改良して行くが、文脈依存メカニズム無しではなかなか好ましい顔が出現しないため、好ましいと思うものだけを選択して行くことを示している。

4.3 選択画像の品質

上記の選択方法の違いは、選択された顔の質として現れる。文脈依存メカニズム有りの第1回目の結果では、F値が15程度と非常に小さく、類似性が高いことを示している。これに対し文脈依存メカニズム無しでは、F値が44程度で非常に悪い値にとどまっている。

類似性が高いとしても満足できるデータであるとは限らないが、①被験者が好ましいデータとして選り②文脈依存メカニズム有りと無しを選択した顔データの特徴は類似した傾向にあること③選択された10個の顔画像の特徴が似ていること、から好ましいデータが選ばれていると推定される。選択した10個の顔の平均値とその中から「最良の顔」として選んだ顔との類似性を各パラメータの平均自乗誤差で評価した(表1)。第1回目の結果では、文脈依存メカニズム有りは無しの場合に比べ1/4程度の誤差である。即ち、最良の顔も他の選択された顔と類似性が高く、選択された顔が望ましいものであったことが分かる。

4.4 検索時間

検索時間について見ると、文脈依存メカニズム有りと文脈依存メカニズム無しで若干文脈依存メカニズム無しの検索時間が短い、ほぼ同程と見てよい。この理由は、類似性が高いほど顔画像の良否を決めるのに時間がかかるためと考えられる。文脈依存メカニズム有りでは類似性の高いデータが次々出てくるため、評価に時間がかかるのであろう。検索数当りの検索時間は、文脈依存メカニズム有りで0.52個/sec、文脈依存メカニズム無しで1.11個/secと、ほぼ1個のデータ当り2倍の時間がかかっている。ただし、検索時間がほぼ同一としてみると文脈依存メ

カニズム有りは、検索数及び検索データの品質において、文脈依存メカニズム無しに比べ、はるかに有利になっている。

5. 考察

上記の結果から、検索時間が同一の条件下では文脈依存メカニズム有りでは少ない検索数で望ましい画像が選べることが明らかになった。

それでは、検索をどの程度続ければ良いのであろうか。データベース内のデータを総て評価できれば良いが、一般的にはデータ数が多くなれば不可能になる。そのことは本実験において文脈依存メカニズム有り及び無しでも検索時間がほぼ同一時間になったことからもうかがえる。文脈依存メカニズム無しの場合は文脈依存メカニズム有りには比べ、はるかに悪い評価値で終了している。その理由として今回の実験では最長の検索時間が30分程度であった。検索時の疲労感が推測される。勿論、検索に対する動機付け等に依存するが、今回の実験ではそれが限度のようである。

文脈依存メカニズム有り無しと無しの各々の実験について、第1回目の集計と第2回目の集計を比較する。

文脈依存メカニズム無しでは、第1回目の結果では検索数が377回から、第2回目の終了時点では793回と416個増加した。その結果、F値は44.8から43.4へと1.4低下し改善の傾向にある。これに比べ文脈依存メカニズム有りでは、第1回目の結果で検索数が205から、第2回目で436と235個増加した。しかし、F値は15.2から16.1と0.9増加し、品質の低下が見られる。個々のケースでは改善がある場合と劣化する場面があるものの、検索数を増加させても品質の改善は望めない。この様子は、図3からも読み取れる。文脈依存メカニズム無しでは検索数を増やすに従って、F値は改善されるが、上記のようにデータ数や時間的な制約から限度がある。一方、文脈依存メカニズム有りでは、200~300個の検索画像数でほぼ満足できる結果となった。

次に、イメージを形成して行く過程とそれを支援するメカニズムとの関係について考える。好みの顔

を選び出すという条件では、必ずしも一種類の好みの顔に限らず種々の顔が選び出す可能性がある。しかし、実験結果から文脈依存メカニズム有りでも無しでも一貫した傾向を示している。これはF値の標準偏差が小さいことで裏付けられている。実際の検索過程を観察しても途中から、それまで選んだ顔と著しく異なる顔に切り換える検索者はいない。即ち、検索初期の段階で選んだ顔には好みの顔の特徴がかなり含まれていると考えられる。

それでは、選択済み顔画像を表示し、再評価させるメカニズムが不要であろうか。実際の使い方を見ると再評価のメカニズムは必ず使用されている。選ばれた顔を一覧にしてみると好みの良否の判断が容易になる。途中段階で選択済み顔画像を表示し再評価を行い好みの顔の絞り込みを行った方がシステムのメカニズムがうまく働き速く収束することになる。

一方、選択された好みの顔画像に類似の顔の他に2個のランダムに選んだ顔画像を表示しているが、その必要性はあるだろうか。実際の検索過程に於ても使用頻度はそれほど高くはない。しかし、途中で好みの顔を変更する可能性があることと、類似の顔に収束した場合に、一部の特徴を微調整する手段として必須な機能である。この機能はそれ以上に顔の好し悪しを判断するのに役立っているのではないかと思われる。なぜなら、顔の良さの評価を行うときに類似の顔だけが表示されたのでは、その表示された顔の中の相对比较は容易でも、それ等が他の顔に比べてどの程度良いのかの評価が下しにくい。異なった顔との比較をすることにより絞り込まれつつある顔の評価が容易になるものと考えられる。

この点から考えると、現状のように全くランダムに選んだ顔画像を表示するよりも、それまでに選ばれた顔にない特徴を持つ顔画像を表示した方がもっと良いシステムとなる。また2個のランダムな顔画像についてもお互いに異なる特徴を持つものを表示したほうが良いであろう。これ等のロジックとランダムに表示する画像の個数などは実用システムの設計時にシステム構築の容易性と効果を考え合わせながら決める必要がある。

6.まとめ

本稿では、文脈に依存させながら次の候補を選ばせると言うように、変更の可能性を取り入れ再評価できるようにしたシステムの有効性を、人間の検索特性に基づいて検証した。その結果、予め検索すべきイメージかはっきりしていない場合でも、全画像を対象とすることなく20ないし30回の非常に少ない検索回数で目的の画像を絞り込めることを示した。また、検索回数をこれ以上増やしても品質の改善が望めないところまで絞り込みが図れることが分かった。

本方式は従来の検索方式と融合することも可能である。例えば検索意図がある程度しっかりしている場合、候補の絞り込みの段階までは言語的キーまたはメニューにより候補を絞り込ませ、そのあとで質の高い理想イメージに到達させるのに本方式を組み合わせる。これにより検索候補の絞り込みがすみやかに進められ検索時間の削減が図られるであろう。

今回の実験では簡単な類似度計算式を品質評価式として用いたがそれでも満足行く結果を得た。実用システムでは状況がもっと複雑になるので工夫が必要になるだろう。現実場面での人間の概念形成過程の行動パターンや画像パターンに対する心理尺度を工夫すればもっと使い勝手の良いシステムが構成できると思われる。これらは今後の課題である。

謝辞

本研究を行う機会と有益なコメントを頂いたATR視聴覚機構研究所葉原耕平会長、淀川英司社長並びに上野圭一室長に謝意を表します。

参考文献

- [1] 加藤、下垣、藤村；画像対話型商標、意匠データベース TRADEMARK、1988, Vol. 14, 681-688.
- [2] S.K.CHANG 他；An Intelligent Image Database System, IEEE, Trans. on SOFTWARE ENGINEERING, 1988, Vol. 14, No. 5, 681-688.
- [3] 木見尻、安居院、中嶋；SD法を用いた画像検

索システムのための一検討、電子情報通信学会春季全国大会、1989, SD-7-3, 6-353~354.

[4] 栗田、下坂、加藤；主観的類似度に適応した画像検索、情報処理学会論文誌、1990, Vol. 131, No. 2, 227-236

[5] 柴田、井上；画像データベースの連想検索方式、電子情報通信学論文誌D-II, 1990, Vol. J73-D-II, No. 4, 526-536.

[6] 柴山他；知識を用いた図形の類似検索法、IE90-8, 55-60.

[7] 美濃、岡崎、坂井；対象物の属性特徴による画像検索法、1991, Vol. 132, No. 4, 513-522.

[8] 印象言語に基づく顔画像の検索、HCI91-44, 17-24.

[9] 近藤他；曖昧な記憶からの高速画像システム、第23回画像工学コンファレンス、1992, 39-42.

[10] 尾田；人間の認知特性を利用した画像検索システム、情報処理学会ヒューマンインタフェース研究会、1992, 44-19, 141-148.

[11] 尾田；図形概念形成過程の文脈依存性とその応用、情報処理学会ヒューマンインタフェース研究会、1991, 35-2, 9-16.