

画像データベースと視考支援環境

笠原 久嗣 岸本 登美夫

NTT ヒューマンインタフェース研究所

視覚的思考（視考）支援環境としての画像データベース利用環境の高度化に関して述べる。まず、人間と画像情報記憶機械としてのコンピュータの相互協調の内容について考察し、人間の思考過程とその各過程を支援する画像情報へのアクセス形態から、視考支援のステージとして Browsing, Hunting, Processingの3段階を抽出しそれぞれの研究課題を明らかにした。続いて、Browsing ステージを扱った我々の研究例として画像情報群ブラウジング支援のためのナビゲーションインタフェースの研究、Hunting ステージの研究例として3次元形状に関わる対話や思考を支援する3次元インタフェースの研究をそれぞれ紹介し、画像データベース利用技術を考察した。

Image Database as Visual Thinking Augmentation Environment

Hisashi Kasahara Tomio Kishimoto

NTT Human Interface Laboratories

1-2356, Take, Yokosuka-shi, Kanagawa 238-03, Japan

We refer to the cooperative environment between a user and a computer which enhances the value of the image database system up to the level where a powerful visual thinking space is created. The human thinking stages and the ways of mnemonic access in the course of them are studied, as "browsing", "hunting" and "processing" stages, to clarify the scope of our research. Our recent research results are introduced. They include the navigational cooperation which augments the user's mnemonic browsing process and the 3-dimensional interaction which operates 3-D shape as a supplement to their mnemonic hunting process.

1. はじめに

視覚的思考（視考）支援環境としての画像データベース利用環境の高度化に関する我々の考え方と具体的研究内容のいくつかを紹介する。

本稿では、まず、視考支援環境における人間と画像情報記憶機械としてのコンピュータの相互協調の内容について考察を加え、その研究課題を明らかにする。続いて、視覚情報を生かした画像データベース利用環境の研究の中から、画像情報群ブラウジング支援のためのナビゲーションインタフェース、3次元形状に関わる対話や思考を支援する3次元インタフェースの研究をそれぞれ紹介し、その視考支援系としての位置づけを示す。

2. 視考支援環境の研究

我々は、コンピュータをベースとした画像情報記憶機械を“Mnemonic Agent”と呼ぶこととし、画像データベース技術をよりヒューマンインタフェース的視点からとらえてきた。ここでは、人間が優れる能力と Mnemonic Agent が優れる能力とが互いに協調し補い合う環境が研究テーマとなる。

我々は、

(a) コンピュータの蓄積パワー：画像情報をありのまま正確に大量にかつインデックスとの関連を維持しつつ長期間保持できる能力、

(b) コンピュータの通信パワー：地理的に離れたところの画像情報も居ながらにして得られる能力、そして、

(c) コンピュータの表示パワー：3次元的に動きも含めて色つきで高精細にそして臨場感をもって画像情報が再現できる能力、

のような、本来コンピュータが得意としており人間の能力をはるかに超えているこれらのすばらしいパワーを人間の Mnemonic Power として最大限まで活用したい。そして、一方人間の得意分野である創造的思考の世界を刺激し強的に

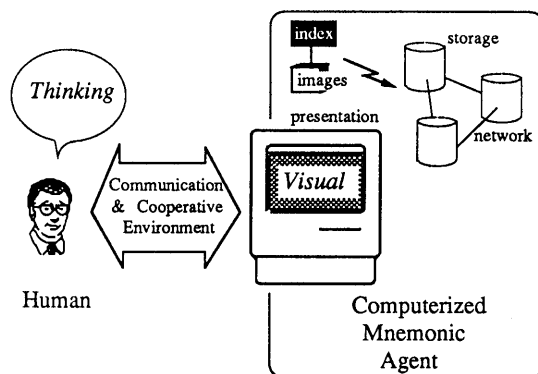


図1. Visual Thinking Environment

増強したい、知的生産能力を向上させたい、と考えてきた。従って、図1に示すように Storage（蓄積）、Network（通信）、Presentation（表示）の各パワーをGivenとした時、人間とMnemonic Agentとの間のヒューマンインタフェースとしての協調環境がどうあるべきか、コンピュータの3大パワーに支援された人間の視覚的思考（以下、“視考”と略称）環境、というところに研究興味がある。

ここでの重要な課題は、人間の思考過程と、その各過程を支援する画像情報へのアクセス形態である。我々は、Mnemonic Agentによる視考支援のステージとして図2に示す3つの段階を考えている。

(1) Mnemonic Browsing：何か発想／思考の手がかりを記憶の中から見つけ出そうとする人間の思考過程を支援

ここでは、このステージを、ターゲット画像の特定の内容イメージを利用者である人間が持っていない、あるいは言葉や絵などのメディアで表現しきれないステージとして定義する。このステージでは、Mnemonic Agentから画像的手がかりを与えてもらうことが必要でブラウジング技術が重要である。このステージでの視考支援のポイントとしては、

a. 高速に一覧できること（1つずつでも何個ずつ同時にでもよい）、

b. 様々な軸で一覧できること（ブラウジング方向の選択性），
 c. 連想的アクセス，
 cについては、たとえば、ブラウジング中の画像情報と関連のある画像情報が常に同時に表示されるようにすれば、これを手がかりとして連想的により有効な手がかりに近づいていく、あるいは思いもよらなかった有効な情報を発見する方向に思考の軌道修正をする、といったことを支援できる。このステージにおいては、人間は全ての画像情報に対し意識しながら（言葉を変えれば、発想をめぐらしながら）1つ1つ見ていくので、極力むだな情報の無意味な提示は避ける必要がある。

（2） Mnemonic Hunting：人間にとってある特定の意味をもつ画像を記憶から直接引き出す／あるいは記憶中にあったかどうか確認しようとする人間の思考過程を支援

このステージではBrowsingと違って手がかりはすでに人間が持っているとして定義する。このステージでは、探索の手がかりを Mnemonic Agent に伝える手段の利便性が要求される。つまり人間から Mnemonic Agent へのコミュニケーション能力の向上が重要である。このステージでの視考支援のポイントとしては、手がかりを提示するメディアに制約がないことがあげられ、手がかりは言葉であったり、絵やスケッチであったり、実物であったり、あるいは身ぶり／手

ぶりであったりするだろうが、いずれにしても手がかりの意図を正確に Agent が理解し、あるいは理解しがたい場合は対話により助けになる情報を人間から引き出すための手続きを実行することが必要である。

（3） Mnemonic Processing：手がかりとなる画像情報素材をベースとして新しい発想／思考を具体的な形に組み上げようとする人間の思考過程を支援

このステージでは、アイデアの視覚化やそのアイデアの有効性のシミュレーションによる確認が重要で、この過程により更に新しい手がかりを得て再び上記2つの過程を繰り返すことにより、発想をふくらませていく。従ってこのステージでの視考支援のポイントとしては、

- a. プロセスしたアイデアを視覚的に短期記憶としてとどめ、比較評価に供すること、
 - b. アイデアを目的に合わせ視覚的にシミュレーションし、評価すること、
 - c. あとあと使えるようなアイデアを長期記憶の中に他の情報と関連づけて蓄積保存すること、
- があげられる。

以上の3つの視考支援段階は単独ではなく、視考の過程に応じて適切に組み合わせられて用いられることが必要となる。

以下、視覚情報を生かした画像データベース利用環境としてのナビゲーションインタフェース及び3次元インタフェースの研究を紹介しつつ、視考支援環境としての画像データベース利用技術を具体的に考察していく。

3. 画像情報群ブラウジング支援

3. 1. 研究の位置づけ

画像素材のデータベースでは、画像の有する情報量の大きさ／多様性に対して検索キーあるいは書誌インデクスとして利用者が記述できる情報量が極めて限られ、一般的なデータベースの手法の応用として書誌インデクスを作成して

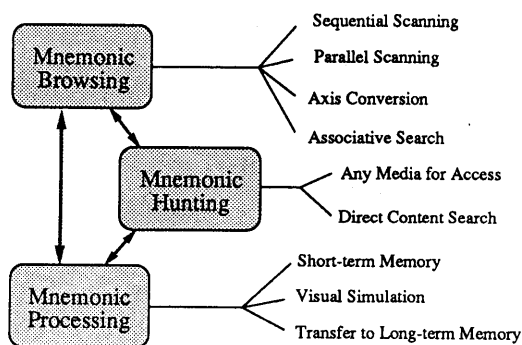


図2. Visual Thinking Augmentation

データベース化し、それをキーワードマッチングで検索するという形態だけでは利用環境として不十分なケースが多い。その主な理由としては次の2点が考えられる。

①画像検索の手がかりである検索キーの表現手段と内容が制約される

② Mnemonic Browsing のステージのように手がかりを持たない場合にデータベースを利用する術がない

①については4節にその考察を譲り、ここでは②への対処を試みた我々の研究を紹介する。

我々がこの研究において想定している画像データベース利用ステージは、Mnemonic Browsing のステージであり(2節(1)参照)、画像の手がかりを得ようとする利用者がコンピュータの記憶保持している画像情報群の中をブラブラと歩き回る(ブラウジング)際、いかにこれを的確な方向に誘導するか(ナビゲートするか)、という課題を扱った研究である。一般的にブラウジングでは、情報間の関連を見定めながらの進行方向制御が課題となる。数値・文字データベースや知識ベースに対しては、従来種々のブラウザの研究が行なわれてきているが、画像データベースに関しては、意味的関連に加えて視覚的関連の表現が必要で特殊な工夫を要する。本研究のポイントは2つあって、1つは方向誘導のための視覚的インタフェース、そしてもう1つは画像情報群をいかにブラウジング空間としてモデル化するかという点である。

3. 2. 視覚的インタフェース

我々のシステムでは、図3に示すような視覚的インタフェースを構成している。ここでの基本的なコンセプトは、

a. Mnemonic Agent は、情報記憶空間をシーンとして可視化する、

b. シーンは利用者の現在位置といくつかの Visual なナビゲーター群(ブラウジング誘導子)で構成する、

c. ナビゲータは画像オブジェクトとそれらの

記憶中での関連を反映して作成する、

d. 利用者によるナビゲータの選択によりブラウジング方向を逐次選択していく、
というものである。

図の中央の領域が利用者の記憶空間中の現在位置を示す”コアイメージ”を表示する領域で、現在参照中の画像がそのまま表示される。その回りの小さなウィンドウ群がナビゲータを表示する領域でそれぞれ特定の関連情報群へ利用者を逐次誘導する。

ナビゲータは対応する画像オブジェクトとコアイメージとの関連内容に基づいて生成され、ナビゲータの選択によりデータベース空間内の注視点を徐々に望みの方向に移動させていく。関連内容はブラウジング対象となるデータベースの構造に依存するが、

(a) 画像的(or 撮像対象)属性

ex. 存在/非存在, 数(多少, 増減), 形状, 姿勢(方向角差: 傾く), 大きさ(伸長/収縮方向, 面積差/体積差), 位置(特定座標軸への位置射影: 上下左右), 配置・包含関係(主体-客体間位置関係: 近づく/集まる/上がる, 対象間距離/方向角差), 部品の接続関係(主体-客体間接続関係: 接続位置, 方向), 構成

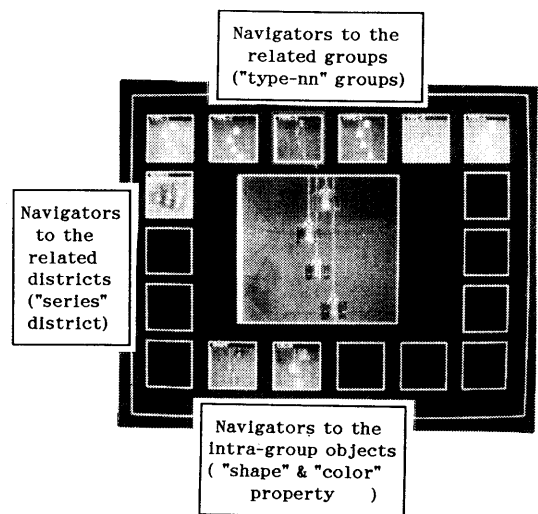


図3. Visual Scene for Navigation

要素の付加／削除，色彩（色相Hue／彩度Chroma），明るさ（明度Value）→ 例えば，I H S 表色系の3座標

(b) 視点的属性

ex. 撮像方向／位置（方向角差，位置差）

(c) 地理的属性

ex. 地理上の位置（方向角差，位置差）

(d) 時間的属性

ex. 時刻（順方向（or 逆方向）時間差），時間（長／短）

(e) 意味的属性

ex. カテゴリ，参照関係，部分－全体関係，部分－部分関係

など多数考えられる。実際の実施においては、これらの情報属性のいくつかについて、それぞれ遠近測度が応用毎に規定され、コアイメージに対する属性方向とその遠近測度の評価尺度に基づき、ナビゲータが生成される。従って、普遍的に情報の遠近が規定されるものではない。たとえば、図3の照明器具のカタログブラウジングでは、形状－色彩－大きさ－数の各属性をこの順の優先度で情報間の関連評価に使用した結果が反映されており、たとえば形状が同じ情報についてのみ色彩の差分の評価が行なわれている。

ここでは、Mnemonic Agent が提示するこのシ

ーンが画像データベース利用のための画像的手がかりとなっており、これにより、

① キーワードレスで直接データベース世界に入り込み、

② 利用者のコアイメージに対する満足度に応じて、

③ 再度検索式を投入することなく、

④ Mnemonic Agent から提示された属性の関連方向やその量に対する利用者の視覚的認知メジャーに応じて、

⑤ ナビゲータの選択という簡便な操作で、画像データベースが利用できる。

3. 3. 画像管理モデル

画像情報群空間の管理モデルとしては、“群化と関連”（Grouping and Relation）という独自のモデルを適用しており、これによりナビゲーションを効果的に駆動する環境を構築している。群化は画像情報の特定の意味的まとまりを反映し、関連は画像情報間あるいは情報群間の同値性や差分関係を反映する。図4の左側が群化の階層を、右側が群内の情報間関連と本モデルの特徴の1つである、1つの画像情報の複数グループへの多重帰属の例を示している。我々のシステムでは、このモデルとの対応で言うと、図3の上端の複数のウィンドウを画像情報の各

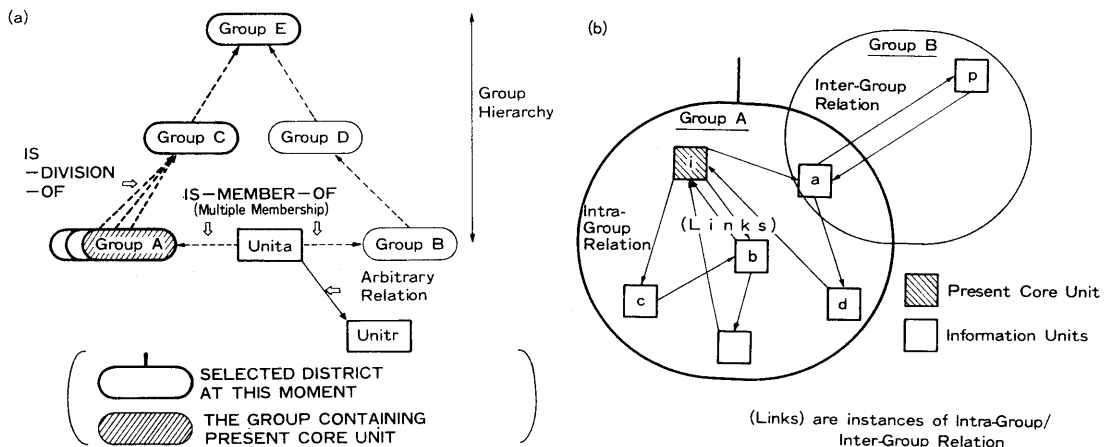


図4. Image Space Modeling for Navigation

グループ化階層へのナビゲータ，下部の各ウィンドウをグループ内の各属性に対応した関連方向の画像情報群へのナビゲータ，左側の各ウィンドウを他のグループ階層—これをディストリクトと呼んでいる—へのナビゲータとしてそれぞれのナビゲータ領域を使用している。このように、群化した画像情報群を代表する画像を選定してナビゲータとすることにより、画像群間の大局的関連参照が効率的に誘導できる。

この研究では、前節で述べた Mnemonic Browsing 支援の要素をすべて包含しており、ブラウジング軸方向毎の Sequential Scanning, 複数のブラウジング軸の並列的一覧, ブラウジング軸の2次元空間への写像, そして画像そのものをナビゲータとして用いることによる視覚的連想サーチの手がかりの提供などを実現している。

このようにして、何か発想／思考の手がかりを記憶の中から見つけ出そうとする人間の不確定な思考過程を支援することを試みている。

4. 3次元形状ハンティング支援

4. 1. 研究の位置づけ

最近のコンピュータによる3次元画像情報処理技術の進展は著しいものがある。ところが、

人間が言葉や2次元の画像から3次元的視点の画像をいともたやすくイメージし認識／理解するのに対し、現在のコンピュータはこの能力においてはまだまだ極めて低いレベルにあると言わざるを得ない。一方、3次元的画像をリアルに正確に視覚化し操作できるコンピュータの能力は一般の人間をはるかにしのいでいる。

我々は、コンピュータの記憶する画像情報群の中から3次元の形状手がかりによって、該当する画像を引き出そうとする状況において、人間が提示する手がかりである形状表現（言葉であったり、絵やスケッチであったり、あるいは実物であったりする）と記憶中の3次元画像とを結びつける、Mnemonic Agentが処理可能な中間メディアの可能性を検討しているが、現在は超二次関数表現の適用性について分析・検討を行なっている。そして、3次元形状を検索キーとする3次元画像データベースへの適用が試みられている。

この研究において前提とした視考のステージは Mnemonic Hunting (2節(2)参照)である。ここでは、利用者から Mnemonic Agent への手がかりの正確な伝達が肝要である。伝達メディアの各々について考えると、3次元の形状手がかりを伝えるには、

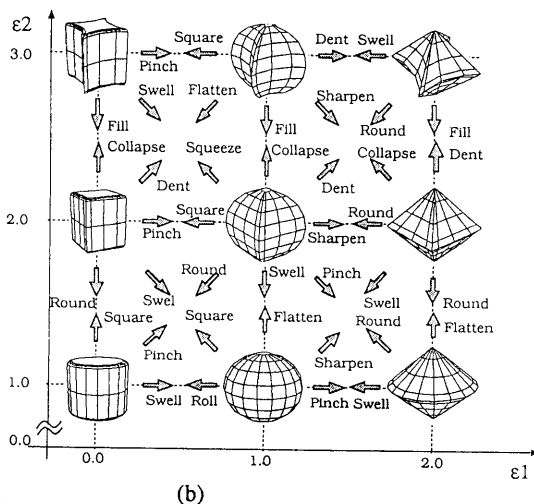
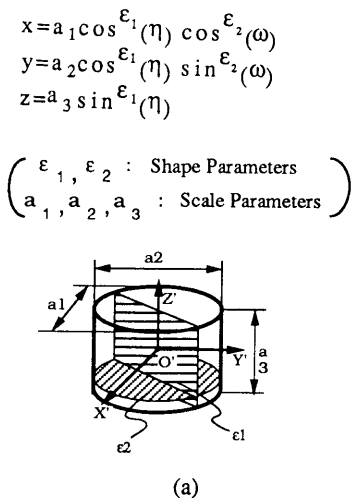


図5. Superquadrics Representation

(a)言葉では、表現能力が小さすぎる、
 (b)スケッチや図面は、絵ごころが必要で思いどおりの描画は困難である、
 (c)3Dグラフィックスは、操作が専門的でむずかしくまた造形ごころが必要、
 (d)実物は、手に入れにくい、
 などの障害がある。そこで本研究では、人間の得意な形状理解と Mnemonic Agent の得意な形状の視覚化／生成機能の協調により、人間からコンピュータへの記憶探索手がかりの正確な伝達を支援する環境を構築しようとしてきた。そして、ここでは、言語メディアと画像メディアの相互補完，サンプルをベースとした変形による対話などへの適用性から超二次関数を中間表現メディアとして導入した。

4. 2. 超二次関数表現

この関数は、図5(a)に示す式のように、立体の垂直及び水平断面形状を表わす形状パラメータ (ϵ_1, ϵ_2) とスケールパラメータ (a_1, a_2, a_3) により表現され、形状を表現する言葉—丸い，角張っている，細長い等—と関数のパラメータ群との対応がとりやすいという特性を持っている。また、形状パラメータの変化に追従して表現形状が連続的に変形することか

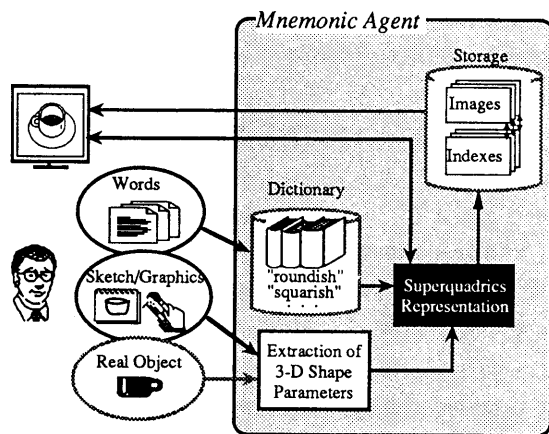


図6. 3-D Visual Thinking Augmentation Environment

ら、変形の制御にも、言葉による、程度（“もっと”，“もう少し”など）や、方向（“丸まる”，“角張る”など）の指示が可能であるという特長がある。

4. 3. 3次元視考支援系

検討しているシステムの機能構成を図6に示す。また、実験的に構築した3次元オブジェクトデータベースのインタフェース例を図7に示す。ここでは、超二次関数による各データベー

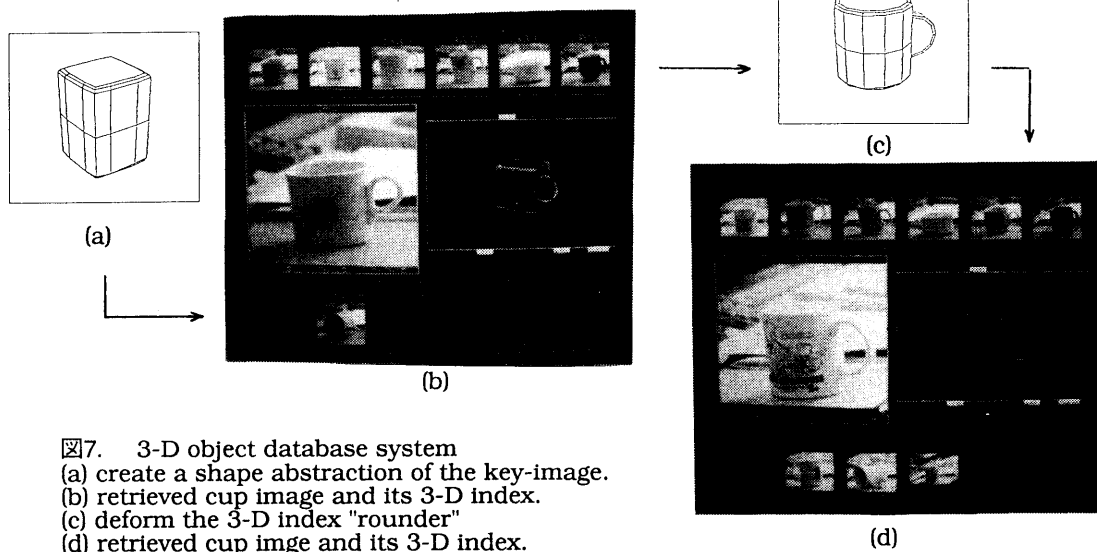


図7. 3-D object database system
 (a) create a shape abstraction of the key-image.
 (b) retrieved cup image and its 3-D index.
 (c) deform the 3-D index "rounder"
 (d) retrieved cup image and its 3-D index.

ス化画像に対する形状インデックス付与と変形操作とを応用した画像検索対話システムが構成されている。

このシステムでは、利用者が入力手段として言葉やスケッチ等を用いて3次元情報を入力することにより、システムは超二次関数を介在として3次元の意味を認識し、これとインデックス（これも超二次関数で表現されている）との比較（関数パラメータのみの比較になるため容易）により、検索処理を行なって結果を利用者に返送する。検索結果は、候補画像とその画像に対応するインデックスを示す画像の並びとして示される。利用者はこの後は、新たに得られたインデックス画像に所望の変形を加える（サンプルをベースとした変形）ことにより、検索の意図を更に細部まで絞り込んでいく手続きをくり返し利用することができる。

このようにして、利用者は、手がかりの伝達メディアの制約から解放された環境でコンピュータがどのように手がかりを解釈したかを視覚的フィードバックで確認しつつ、画像データベースが利用できる。

5. おわりに

以上、視考支援環境としての画像データベース利用環境の高度化について、我々の関連研究を紹介した。視考支援環境は、利用者による連想、類似／相違判断を支援し、新しいアイデアやデザインの創出へと利用者の思考を誘導していく環境である。従って、現在のコンピュータの画像思考能力（コンピュータの連想、類似／相違判断）を期待あるいは開発して、コンピュータに創造的発想をさせようとする研究ではない。あくまでも、利用者とコンピュータそれぞれの得意分野の能力の相互協働系に主眼をおいた研究であるところにその特徴がある。

今後のビジュアル情報社会を展望すると、コンピュータやインフラとしての通信ネットワークの機能強化という従来から精力的に取り組ま

れてきた研究開発に加えて、その利用側の技術一本稿で述べたようなヒューマンインタフェース的視点からの使いこなしの技術一へもより目が向けられていくことが重要で、特に専門家以外に広範な直接利用層が想定される画像情報システムに関しては、その視点からの研究をベースとしてより優れた一使える一道具の提案をコンピュータ研究者の側から積極的に行なっていくことが望まれる。“いいセンス”を持ったコンピュータシステムとの“話せる”関係を具現化していくことがブレークスルーにつながる。

[文献]

- [1] Kasahara, H., Y. Tonomura & T. Kishimoto "Image Mnemonics by a Computer -Visual Thinking Augmentation-", International Symposium Computer World '89, (Sep.27-29, Osaka), 1989, pp63-70.
- [2] 岸本, 笠原 "言語としての画像小特集: 5. データベース検索への応用", 画像電子学会誌 Vol.18, No. 4, pp191-196, 1989.
- [3] Kasahara, H., T. Kishimoto "Navigation in the Visual Information Database World", International Symposium on Database Systems for Advanced Applications (Apr.10-12, Seoul), 1989, pp25-32.
- [4] 笠原, 岸本 "画像データベースナビゲーション", 電子情報通信学会 技報IE88-46, 1988.
- [5] 堀越, 笠原 "画像データベースにおける三次元インデクシング", 電子情報通信学会 技報IE88-111, 1989.
- [6] 堀越, 笠原 "画像記憶術-3Dビジュアルコミュニケーション", 電子情報通信学会秋季全国大会 SD-3-2, 1989.