

「Enra-Enra」 (類似画像検索のためのユーザインタフェイス)

広池 敦 武者 義則

情報ベース機能日立研究室, RWCP

東京都国分寺市東恋ヶ窪 1-280

(株)日立製作所 中央研究所

{he, sha}@crl.hitachi.co.jp

要約

類似画像検索における検索結果を、画像の特徴量から構成される仮想的な3次元空間に可視化するシステムを開発した。本システムは、数10万件規模の静止画像コンテンツを対象とする画像特徴量に基づいた検索を行う。検索によって得られる約2000件の検索結果は、サムネイル画像により構成される時空間パターンとしてユーザに提示される。本稿では、このシステムのユーザインタフェイス上の諸機能、上記の空間表現に適したウォークスルー方法について紹介する。

“Enra-Enra” A User Interface for Similarity-Based Image Retrieval

Atsushi Hiroike Yoshinori Musha

Information-Base Functions Hitachi Laboratory, RWCP

c/o Central Research Laboratory, Hitachi Ltd.,

1-280, Higashi-koigakubo, Kokubunji-shi, Tokyo 185-8601, JAPAN.

{he, sha}@crl.hitachi.co.jp

Abstract

We have developed a user interface for similarity-based image retrieval, where the distribution of retrieved data in a high-dimensional feature space is represented as a *dynamical* scatter diagram of thumbnail images in a 3-dimensional visualization space constructed from the feature space. In this paper, we explain the functions of the user interface, and a new walk-through method suitable for browsing the visualization space.

1 はじめに

画像自体の持つ類似性に基づいて画像検索を行うシステムが、近年、数多く発表されている。我々は、特にユーザインタフェイスの問題に焦点を当てて、類似画像検索を行なうプロトタイプシステムの開発を行なってきた[2], [3], [4], [6], [7], [8], [9], [10], [12], [11]。我々のシステムの基本的な特徴は、画像検索の結果を、画像特徴量から構成される2次元ないしは3次元の空間上に可視化する点にある。このような

ユーザインタフェイス上での検索結果の表現については、[1], [5], [13]等が同様のモデルを論じている。今回報告する我々のモデルでは、検索結果の分布は、画像特徴量から構成される3次元可視化空間中の時空間パターンとして表現される。この基本的な表現方法自体は、過去に、[6], [8], [11]等で報告した。本報告では、システム構成、ユーザインタフェイスの種々の機能、空間内のウォークスルー機能について、より詳細に論じたい。

まず、本論に入る前に、システムを開発するに際

しての我々の基本的なコンセプトについて述べる。

1.1 「感覚器管の延長」としてのシステム

類似画像検索が通常の検索と大きく異なるのは、キーワード等のアルゴリズム的に厳密に定義できる属性ではなく、「画像間の類似性」という本来人間が主観的に判断する情報に基づいて検索する点である。このような主観的な判断を計算機上のアルゴリズムとして実現するのが、いわゆるパターン認識技術である。ただし、特に、限定のない多様な画像を対象とした場合、現状では、その実現は非常に困難と言える。

通常、パターン認識技術を応用する際には、対象となるパターンの性質を適切に見極めることにより認識精度の向上を目指す。勿論、これ自体は正しいアプローチと言える。しかし、例えば、WWW から収集される画像、あるいは、個人がデジタル映像機器で撮影した画像等を想定した場合、「限定が少ないデータ集合に対して応用できる技術」もまた、非常に重要といえる。大量のパターン情報が容易に入手できるようになった現在では、このような技術の重要性が今後増してくるであろう。

一方、「類似性」が、主観的な判断である以上、その基準は、個人間のもとより個人内でも変動するものと考えられる。従って、たとえ計算機のパターン認識能力が人間並みに進歩したとしても、ユーザが、(計算機上のシステムを含む) 他者の判断に対して満足する保証はない。ユーザは、あくまで自らの主観の方を重視するかも知れない。

以上のような考察から、我々は、類似画像検索を行なうシステムは、以下のような発想で設計すべきであると考えた。通常の場合、検索システムでは、なるべく絞り込んだ精度のよいデータを見つけ、ユーザに知らせることが目標となる。しかしながら、ユーザがシステムの判断に満足する保証はない以上、このようなアプローチには限界がある。そこで、我々は、「なるべく多くのデータをユーザに見せること」を目標として設定する。計算機は、パターン認識能力では人間に劣るが、大量のデータを扱うことに関しては人間より優れている。我々は、人間のパターン認識能力、判断能力と、計算機を持つ大規模データ処理能力とを適切に「接続する」ことを目指している。この場合の「接続する」とは、人間が直接は見るできない膨大なデータが張る情報空間を、ユーザにとって直感的に理解可能な表現へと変換す

ること、および、その情報空間に対する直感的な操作方法をユーザに提供することを意味する。通常、パターン認識のモデルは「判断の主体」のモデル化を目標としている。これに対して、我々は、人間の「感覚器管の延長」となるようなシステムの構築を目標としている。

1.2 パターン情報空間の可視化

システム内部の情報空間を人間にとって直観的に理解可能なものへと変換した例としては、現在、ユーザインタフェイスとして一般的となっているデスクトップメタファがある。デスクトップメタファでは、ファイルシステム上の論理的な関係が可視化され、それまでのコマンドライン入力より、より直感的なファイル操作が可能となっている。そのアナロジーで言えば、我々のシステムは、「パターン情報間の持つ計量的関係を可視化するもの」である。この可視化、および、可視化された表現上での操作が適切にモデル化されたなら、パターン情報を扱うのに適したユーザインタフェイスが実現可能となろう。

パターン情報間の計量的関係を可視化することは、古くから、多変量解析の分野で研究されてきたことであり、我々も、これらの技術を応用している。我々は、多変量解析の手法を、静的なデータの解析ではなく、ユーザの要求に応じて動的に変化するデータ集合(検索結果の集合)の表現のために用いている。

2 データの内部表現

ここでは、ユーザインタフェイス上の機能の前提となる、画像情報に関するシステム内部での表現、および、処理について述べる。

2.1 特徴量

画像特徴量としては、主として、以下の2種類を用いている。

色特徴量 画像の色ヒストグラムに基づく特徴量。ヒストグラムは、RGB色空間を $N_r \times N_g \times N_b$ のセルに分割することによって計算する。構造的な情報が必要な場合は、画像を $N_x \times N_y$ の矩形領域に分割してから、各領域ごとにヒストグラムを計算する。この場合、得られる特徴量の次元は、 $N_r \times N_g \times N_b \times N_x \times N_y$ となる。

微分特徴量 画像中の輝度勾配ベクトルの方向分布に基づく特徴量。ベクトルの方向を N_θ レベルに量子化する (ただし、角度方向の範囲は $-\pi/2 \leq \theta < \pi/2$ とし、明暗の反転の違いは無視する)。その上で、 N_θ 次元の特徴量ベクトル $v = \{v_k | k = 0, \dots, N_\theta - 1\}$ を以下のように計算する。まず、各画素について、輝度勾配ベクトル (f_x, f_y) の量子化された方向 k を決める。次に v_k の値を $v_k \rightarrow v_k + f_x^2 + f_y^2$ と更新する。全画素について、この処理が終了したら、最後に、加算された v_k の各値を画素数 S で次式のように正規化する。

$$v_k \rightarrow \sqrt{v_k/S}, \quad k = 0, \dots, N_\theta - 1$$

色特徴量の場合と同様、構図的情報が必要な場合、画像を $N_x \times N_y$ に分割してから上記の計算を行なう。また、この種の特徴量の場合、解像度の違いに応じて異なる情報が抽出されることはよく知られている。 N_l レベルの解像度ピラミッドを用いた場合、最終的に得られる特徴量の次元数は、 $N_\theta \times N_x \times N_y \times N_l$ となる。

通常の場合、検索時には、データベース全体での主成分分析を行なうことにより、上記の特徴量ベクトルを適切な次元数に圧縮したものを特徴量として用いる。実際の例で言えば、例えば、1024 次元の色特徴量 ($N_r = N_g = N_b = N_x = N_y = 4$) と 512 次元の微分特徴量 ($N_\theta = 8, N_x = N_y = 4, N_l = 4$) を、それぞれ、100 次元のベクトルに変換した上で、検索、あるいは、可視化表現のための座標値計算等を行なう。

2.2 類似度

画像間類似度は 2 乗距離に基づいて定義する。 N_f を特徴量ベクトルの種類の数、 $\{x_i\}$ と $\{y_i\}$ ($i = 0, \dots, N_f - 1$) をそれぞれ画像 X と画像 Y の特徴量ベクトルの集合とする。この時、 X と Y との類似度は、

$$s(X, Y) = \exp \left(- \sum_i^{N_f-1} w_i \|x_i - y_i\|^2 \right) \quad (1)$$

ただし、各 2 乗距離は、特徴量の種類ごとに、 $\langle \|x_i - y_i\|^2 \rangle = 1$ となるよう正規化されているものとする。また、 $w = \{w_i\}$ は非負の重みベクトルである。

我々のシステムでは、検索のキーとなる画像の集合をグループ化して管理する。ある画像 X と、あるキー画像のグループ C との類似度は、

$$s(X, C) = \min_{Y_i \in C} s(X, Y_i) \quad (2)$$

として定義する。各キー画像グループは、“正・負”の属性を持つ。“負”のキー画像グループは、それに類似した画像を検索結果から除外したい場合に用いる。システムは、データベース中の各画像 X について、まず、最も類似度が高いキー画像グループ C を求め、 C が“正”の画像グループなら $s(X, C)$ を、“負”の画像グループなら $-s(X, C)$ を、そのデータの類似度とする。最終的な検索結果の件数は、類似度に関する閾値と件数の最大値の両者によって決まる。

なお、重みベクトル w は、キー画像グループごとに定義される。 w の値の和が大きくなるとそのグループ内の類似度が一様に小さくなる。これによって、どのグループとの類似性を重視するかが決まる。ただし、「 w の値が大きくなるとそのグループが重視されなくなる」というのは、人間の直観とは一致しない。そこで、ユーザレベルで指定する重みベクトル W は、次のように定義されている。

$$w_i = N_f \left(\sum_i W_i \right)^{-2} W_i$$

この場合、 W の平均値が大きくなると、そのグループの類似度が増大する。

2.3 検索結果に対する統計処理

我々のシステムでは、検索結果の統計量を以下のように定義する。 $\{X_i\}$ を最も類似度が高いキー画像グループが C である検索結果の集合とし、 $\{x_{ij}\}$ ($j = 0, \dots, N_f$) を X_i の特徴量ベクトルの集合とする。キー画像グループ C における j 番目の特徴量の統計量は、

$$\begin{aligned} \mathcal{N} &= \sum_i s(X_i, C)^p \\ \mu_j &= \mathcal{N}^{-1} \sum_i s(X_i, C)^p x_{ij} \\ \Sigma_j &= \mathcal{N}^{-1} \sum_i s(X_i, C)^p (x_{ij} - \mu_j)(x_{ij} - \mu_j)^t \end{aligned}$$

p は、低類似度のデータの統計量計算への効果を調整するためのパラメータで、 $p = 0$ の場合、通常の設定通りの統計量となる。逆に、 $p \rightarrow \infty$ の場合は、 C に含まれるキー画像 (すなわち類似度 1 のデータ)

のみの統計量となる。後述する可視化表現を用いたユーザインタフェース上では、2000件規模のデータを検索結果として扱うが、実際には、高類似度のデータが強調されるような表現法がとられている。パラメータ p は、このような表現法と統計量との整合性を持たせるために導入した。

3 システム構成

システムは、C++で記述された2つのサーバプログラムとJavaで記述されたクライアントプログラムからなる (fig.1)。検索サーバ (M-Server) は、起動時に、データベース中の全件データの特徴量を圧縮された形式 (例えば、ベクトルの各要素を8ビット整数に変換する) でメモリ内に格納する。その後、クライアントプログラムの要求に応じて、検索処理、統計量の計算等を行なう。一方、可視化サーバ (V-Server) は、最大2000件程度までの検索結果を、「可視化空間」中のサムネール画像の3次元的な散布図として表示する。V-Serverが表示する画像を直接マウスクリックする場合等を除き、通常のユーザ入力、クライアントプログラムが受け付ける。

M-Serverは、本稿で述べるシステムのみに対応したものではなく、要求が異なる複数種類のクライアントへのインタフェースを持つ構成となっている。各クライアントは、クライアントタイプ (クライアントの種類) とクライアントIDによって管理され、固有のメモリリソースをM-Server内に確保する。実際のメモリリソースの内容はクライアントタイプごとに異なるが、本稿で述べるシステムの場合、検索結果の集合、検索結果に関する統計量等が保持される。

検索、あるいは、あるいは、ユーザインタフェース上での可視化に、実際にどのような特徴量を用いるかは、クライアントが、簡易言語 (「特徴量記述言語」) によってM-Serverに指定する。「特徴量記述言語」は、C++で記述された特徴抽出アルゴリズムとのインタフェースを提供することを主な目的として、我々が開発した言語である。画像特徴量の選択、再構成等が柔軟に行なえるような仕様となっており、内部的には、実行時におけるシステムの柔軟性と、大規模データの処理を想定したオーバーヘッドの最小化とを同時に満すように設計されている。言語の解釈処理は、基本的には記述を指定した時点でのみに行なわ、実際の特徴抽出等を行なう時点では、コンパイル済みのコードが直接呼ばれる。

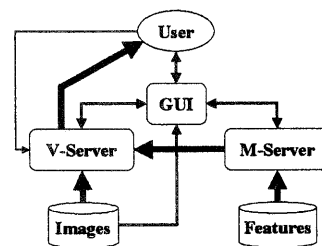
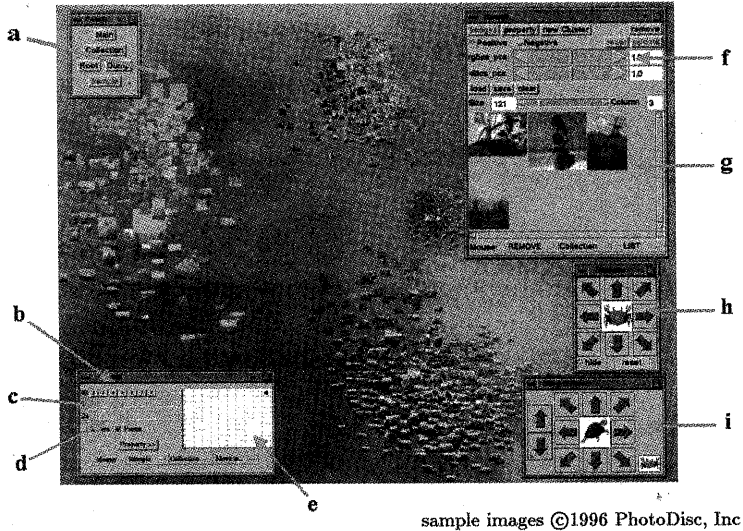


図1: システム構成。矢印の太さは、移動するデータ量を表す。

本稿のシステムの場合、M-Serverは、クライアントが通知するキー画像の指定等の検索条件を受理すると、まず検索処理を行い、次に、その検索結果を対象とした統計処理 (分散共分散行列の計算、主成分分析等) を行なう。これらの処理が終了すると、V-Serverに対して、類似度でソートされた検索結果を、可視化のための座標値ベクトルと共に通知する。V-Serverは、M-Serverからの通知に基づき、新たな表示画像のロード、不要なデータの破棄等を行なう。M-Serverから送られた座標値ベクトルは、V-Serverでは、3次元座標値の系列として解釈され、これに基づいて各画像の運動軌道が設定される。例えば、ある画像に対して12次元の座標値ベクトルが付与されていた場合、V-Serverは、そのデータを、4ステップの運動軌道を持つ3次元オブジェクトとして扱う。実際の運動生成では、相前後する2つのステップ間での座標値は、線型補間によって定める。

このシステムの処理で、実際に処理上最も負荷が大きいのは、V-Serverでの新規表示画像の登録、すなわち、画像のロードである。ユーザ操作へのリアルタイム応答を損わないために、表示内容の更新は、画像全件の登録を待たずに、各時点での登録済みの画像について、順次行なわれる。また、クライアントによる検索条件の変更が起きた場合は、画像の登録は随時キャンセルされる (すなわち、以前にM-Serverから通知された検索結果の内、低類似度のものは破棄される)。なお、ハードウェア的には、V-Serverが、大量のテキストチャータデータを扱う関係上、十分はテキストチャータマッピングメモリを持ったグラフィックボードが必要となる。



sample images ©1996 PhotoDisc, Inc

図 2: ユーザインタフェイスの表示例.

4 ユーザインタフェイス

4.1 可視化表現

図 2は、我々のシステムの表示例である。右上のパネル(図 2-g)に表示されているのがキー画像、背景に拡がるのが検索結果が表示される「可視化空間」(図 2-a)である。この例では、約 10 万件のコンテンツから、類似度が高い上位 2000 件の検索結果が、最も近いキー画像の周りに島を作るような形式で表示されている。

この場合の検索結果の配置は、以下のように行なわれている。システム内では、色特徴量、微分特徴量それぞれについて検索結果の統計量、および、それらの分散共分散行列の固有空間が求められる。その上で、まず、各キー画像を、微分特徴量の固有空間

上に配置する。この場合の原点は、可視化空間の原点(すなわち表示系の原点)である。一方、各検索結果については、最も近いキー画像の位置を原点とした上で、色特徴量の固有空間上に配置する(ただし、実際には、キー画像と検索結果を区別した処理は行なっていないため、配置を決める処理の流れ自体は、この記述とはやや異なる)。

3で述べたように、本システムの表示系(V-Server)は、各画像を独立な運動軌道を持つオブジェクトとして扱う。画像特徴量を運動軌道の座標値として用いた場合、各画像は、基本的には独自の動き方をするが、類似性が高い画像どうしは、互いに引き付けあつような形で、やはり類似した動き方をする。結果として、画像の分布全体は、「有機的な運動性を印象として与える時空間パターン」として表

現される(図 2-j)。図 2 の上図の例では、実際には、キー画像は不動で、検索結果のみがステップ数 4 の運動起動を持つように設定されている。すなわち、M-Server(検索サーバ)は、V-Server に対して、色特徴量の 12 次元までの固有空間上での座標値(ただし、原点をキー画像の位置に移動したものを)、各検索結果と共に送付している。

図 2 の左下にあるパネルは、可視化表示のコントロールを行なうためのものである。図 2-b のボタンの列には、運動軌道が現在どのステップにあるかが示される。ユーザは、このボタンを押すことによって、運動軌道を任意のステップにジャンプさせることができる。そのすぐ下のスライダ(図 2-c)は、運動の速度を決めるコントローラである。そのまた下のスライダ(図 2-d)は、運動を一時的に停止させている場合にのみ使えるもので、相前後するステップ間の任意の時点に表示を移動させる。任意時点での各画像の表示位置は、3 で述べたように、それらの 2 つのステップでの座標系上での位置を線型補間して得られたものである。

これら運動軌道に関連するコントローラの右にある矩形の GUI 部品(図 2-e)は、類似度と画像表示の大きさとの関係を決めるためのコントローラである。このシステムでは、高類似度の検索結果は、3 次元オブジェクトとして大きく、低類似度の結果は、小さく表示する。類似度が 0 の検索結果(これは、実際には表示系には通知されないが)は、大きさが 0 のオブジェクトとなる。すなわち、本システムでは、検索結果・非検索結果の関係が、連続的な印象となるようユーザに提示される。類似度 s と表示の大きさ S の関係は、 $S = \beta s^\alpha$ という形で与えられ、図 2-e の矩形領域の横軸が α (類似度効果の強さ)、縦軸が β (全体的なスケール)に対応している。ユーザは、矩形領域中の点を移動させることによって、例えば、表示を絞り込みたいなら類似度効果を強める、といったように、類似度と表示の大きさの関係を自由に設定することができる。

新たなキー画像の登録は、通常は可視化空間内の画像をクリックすることによって行なう。ただし、場合によっては、Java のパネルの上で、直接データベース全体の画像リストから選択する、あるいは、予めセーブしておいた画像リストから選択する、といった方法もとられる。キー画像の削除は Java のパネルの上で行なう。特徴量重みの指定についても、Java のパネルの上で行なう(図 2-f)。

実際に、可視化空間中に浮ぶ画像をキー画像としてクリックすると、その画像は、次第に大きくなりながら、類似した画像を従えて、それまで属していた画像分布の島を離れていき、新たな島を形成する。逆に、あるキー画像をキーから削除すると、そのキー画像に対応する島に属する各画像は、次第に小さくなりながら、以前属していた島に戻っていくか、ないしは、どのキー画像にも類似していない画像については、可視化空間から消滅する。特徴量重み指定にともなう検索条件の変更についても同様で、その結果は、リアルタイムの運動としてユーザに提示される。この場合、画像の出現・消滅に加え、最も類似したキー画像が変化することにより、島間での画像の移動が生じる。

2.2 で述べたように、我々のシステムでは、キー画像は、グループ化されて管理される。各キー画像のグループは、GUI 上では、それに属するキー画像が表示された 1 つのパネルとしてユーザに提示される。図 2 の例では、右上のパネルに対応するもの以外に、実際には、他に 3 つのキー画像グループが登録されている(対応するパネルは、Java 上で非表示となっている)。これらのキー画像グループには、いずれも“負”の属性が指定されており、それらと類似度が高い画像は、検索結果から排除されている。このキー画像グループの登録・削除は、Java のパネル上で行なう。ユーザは、これによって、キー画像をカテゴリカルに管理することができる。また、検索および可視化に用いる特徴量の指定は、グループごとに独立に行なえる。例えば、あるグループだけを(キー画像ごとに島を形成するのではなく)グループ全体が 1 つの固まりとなるような表示に設定することも可能である。キー画像グループ全体は、階層的に管理されている。同時に複数のキー画像グループが可視化空間に表示される場合は、上位のキー画像グループの指定に従って、下位のグループの表示上の原点位置が決められる。

4.2 ウォークスルー機能

このシステムで表現される可視化空間は、通常の VR システムで扱われるような、現実中存在する空間(例えば、建造物の中、街並、等)を模したのではなく、画像特徴量が張る情報空間という元々抽象的な世界を表現したものである。実際に、この空間中を、VRML 等が提供する通常のウォークスルー機能のみで移動していると、すぐに迷子になってし

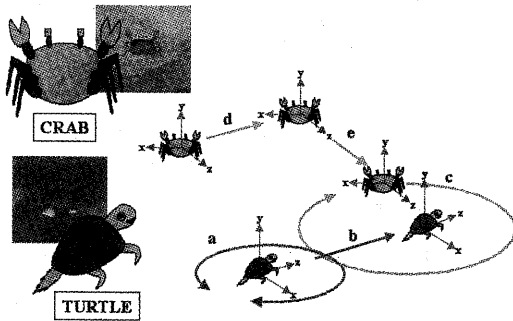


図 3: “カニ”と“カメ”(“Crab” and “Turtle”).

まう。これは、我々が日常的に用いている自己中心的回転(首振り)に基づく視点移動が、この可視化空間に適合しないことを意味する。実際、空間中の画像分布を全体的に眺める際には、このような自己中心的回転ではなく、画像分布の周りを公転するような形式で視点を移動できた方が便利である。可視化空間の性質から要求される、このような特殊なウォークスルー機能を、直感的に理解できるような形でユーザに提供するために、我々は、以下のようなユーザインタフェイスを開発した。

可視化空間中には、“カニ”と“カメ”と名付けられた2つのオブジェクトが存在する(図3)。これらは、異なる視点およびウォークスルー機能を持っている(ちなみに、図2の上の図は、カニから見た可視化空間である)。カメは、通常のウォークスルー機能を表現したオブジェクトで、自己中心的首振り回転(図3-a)を行うことによって、任意の方向に移動できる(図3-b)。一方、カニの回転は、常に、カメを中心として行なわれる(図3-c)。また、カニは、常にカメに面し、カメとの距離を一定に保ちながら、カメの移動を追従する(図3-d)。カニ自体が自発的に行なえる並進移動は、カメへの接近、あるいは、カメからの逃避のみである(図3-e)。

図2のhとiが、これらのオブジェクトを操作するためのパネルである。図2-hのパネル中央にあるボタン(図中ではカニのマーク表示されている)は、視点切替用のトグルボタンで、可視化空間の表示をカニの見え、あるいは、カメの見えに切り替える。その周囲にある矢印が付いた8つのボタンは、現在の視点における (x, y) 平面上の並進移動のためボタンである。一方、図2-iは、オブジェクトの回転および z 方向の移動のためのパネルで、中央にあるボタン(図中ではカメのマーク表示されている)は、制御対

象となるオブジェクトを切り替えるためのトグルボタンである。その周囲にある8つのボタンは、左右、上下、および、斜め方向の回転のためのボタンで、制御対象に応じた回転を行なう。その左の縦に並んだ2つのボタンは、 z 方向移動(前進・後退)用のボタンである。すべての移動および回転運動は、対応する矢印ボタンを押すと開始し、離すと止まる。また、運動速度は、マウス側のボタンに応じて切り替わる。

視点の切替と制御対象の切替は全く独立である。ユーザは、カニの視点からカメの動きを眺めつつ、カメの動きを制御することができ、これによって、カニの回転の中心を、画像分布を眺めるのに適切な可視化空間中の位置へと、容易に設定することができる。

図2-hの (x, y) 平面上の並進移動については、カメ視点の場合は、一種のショートカット的な働きとなる。まず、カメ自身が、現在の (x, y) 平面上のある方向に回転した後、前方(回転後の z 方向)へと進む。一方、カニ視点の場合は、カニは、前述した定義より自由な方向転換ができない。従って、「カメが動くことによって結果として視点が移動する」という形式で動作する。例えば、ユーザが、カニ視点で $+x$ 方向のボタンを押すと、まず、カメがカニの $+x$ 方向と一致する方向に回転する。その後、カメの方が前方方向に進み、カニはその動きを追従する。結果として、視点はカニの $+x$ 方向に移動する。

実際に、カニ・カメを操作してウォークスルーを行なっていくと、互いの位置および姿勢の相対的な関係は当然変化していく。カメの視点では、通常の場合カニをすぐに見失う。一方、カニの視点でも、カメは常に見えてはいるが、姿勢は必ずしも一致しない。カメの制御をする場合、これだと不都合な場合もある。また、現在のカメが見ている視点でカニで見たい、という場合も出て来る。このような場面对応するために用意されているのが、図2-iのパネルの右下にある小さなボタン(図中ではカニが表示されている)である。これは、制御対象の回転についてのショートカット機能を与える。制御対象がカメの場合に、このショートカットボタンを押すと、カメは、カニの姿勢と z 方向が反転した姿勢(カニに面した姿勢)へと、まず回転する。以降、ショートカットボタンを連続して押すと z 方向の反転に伴う回転が順次起きる。一方、制御対象がカニの場合は、カニは、カメの姿勢と全く一致した姿勢へと回転する。カニは、公転運動をするように回転するので、実際には、これは、カメの後側に回り込むような

運動となる。連続してショートカットボタンを押すと、カメの場合と同様、 z 方向の反転が起き、カメの前(後)に回り込むような動作を行なう。なお、これらの回転運動は、すべて、1軸を中心とした連続的な回転運動として実現されている。その際の回転軸および回転角の導出は、クライアントプログラムの側で行なっている。

カメのみが持つ機能としては、画像へのジャンプ機能がある。ユーザは、可視化空間中の画像をクリックすることによって、好きな画像の位置へとカメを飛ばすことができる。

5 まとめと今後の課題

本稿では、類似画像検索のためのユーザインタフェイスとして我々が開発した、検索結果を3次元の時空間パターンとして表現するシステムについて紹介した。現状の我々が用いている可視化空間の軸は、検索結果の分布の主成分分析に基づくもので、「特徴量空間から可視化空間へ次元を圧縮した際の分布の再現性を2次のオーダで最適化する」という意味しか持たない。また、運動の生成に関しても、単に主成分分析で得られた固有ベクトルを切り替えているだけである。ユーザにとってより意味のある可視化表現を得るためには、可視化空間の軸の構成法、時間発展の仕方等に関して、さらに検討していく必要がある。

ユーザインタフェイス上に関しては、今回のシステムでは、マウス操作を基本としている。これについては、現在、我々は、ゲームパッドのような外部入力装置を用いることによって、より直感的な操作性が実現できないかと検討している。これによって、例えば、立体視表示を用いた場合でも、問題なく操作できるシステムが実現できるであろう。

参考文献

- [1] Gupta, A., Santini, S. and Jain R., "In search of information in visual media." *Communication of the ACM*, 40-12, 35-42, 1997.
- [2] 武者, 森, 広池: 大量画像を対象とする特徴量空間の可視化, 第3回知能メディアシンポジウム, 301-308, 1997.
- [3] 武者, 広池, 森, 杉本: Visualizing feature space for massive image contents, RWC Technical Report, TR-98001, 233-238, 1998.
- [4] 広池, 武者, 杉本: 画像情報の計量化と可視化—特徴量を用いた画像データベースのブラウジングシステム—, 第4回関西情報関連学会連合大会, 33-36, 1998.
- [5] Rubner Y., Tomasi, C. and Gauibas, L. J., "Adaptive Color-Image Embeddings for Database Navigation," *IEEE ACCV98*, 104-111, 1998.
- [6] 広池, 武者, 杉本: VR空間を用いた画像特徴量空間の可視化—画像データベースの検索・ブラウジングのためのユーザインタフェイス, 信学技報, PRMU98-86, 17-24, 1998.
- [7] Musha, Y., Mori, Y., Hiroike, A. and Sugimoto, A., "An Interface for Visualizing Feature Space in Image Retrieval," *Machine Vision and Applications*, 447-450, 1998.
- [8] Hiroike A., Musha, Y., "Visualization of information spaces to retrieve and browse image data," *Third International Conference on Visual Information Systems*, Springer-Verlag, 155-162, 1999.
- [9] 武者, 広池, 杉本: 類似画像検索における検索結果の可視化インタフェース—可視化軸として意味軸を用いる方法—, 信学技報, PRMU99-57, 59-64, 1999.
- [10] 武者, 広池, 杉本: 類似画像検索における特徴量空間の可視化インタフェース, 信学論, J82-D-II, 10, 1626-1633, 1999.
- [11] 広池, 武者: 類似画像検索のための可視化とウォークスルー, インタラクシオン 2000, 39-40, 2000.
- [12] 武者, 広池: 意味を表現する空間上に, 検索結果を表示する可視化インタフェース, インタラクシオン 2000, 205-212, 2000.
- [13] 木本: 3次元探索空間を用いるインタラクティブな画像検索システムとその検索特性, 情報処理学会論文誌, 41, 3, 638-657, 2000.