新しいフラクタル特徴とロバストな構図情報を用いた画像検索

Image Retrieval Using A New Fractal Feature and Robust Structural Information

正会員 望月 貴裕[†], 正会員 藤井 真人[†], 正会員 伊藤 崇之[†]

Takahiro Mochizuki[†], Mahito Fujii[†] and Takayuki Ito[†]

Abstract The rapid increase in the volume of electronically archived images and video materials has given rise to a need for new methods of image retrieval. Therefore, various approaches have been proposed. Most conventional image retrieval methods are based on the content of objects or on features describing only the partial properties of images. But the diverse images we must deal with have induced a necessity to retrieve images more flexibly and effectively, focusing on the multiple properties of images. In this paper, we describe a new method of image retrieval based on two new ideas. One is adopting a fractal sequence designed to include multiple information of texture and to relate one's first impression of an image. The other is structural information of images by allowing rectangular regions, which are set to extract features in multiple points on the image, to be located at slightly varied positions dependent on the image features. Additionally, we show with the results of some experiments that image retrieval that widely perceives the information of images more variously can be achieved because of our applying the new method.

キーワード：画像検索，テクスチャ特徴，構図情報，フラクタル，Active Net

1. ま え が き

画像を電子化して扱うことが容易な時代となり，様々なメディアを通して大量の画像にアクセス可能な環境が整いつつあるが，目的の画像を探すためには，目的にあった画像検索手法が必須となる．現状の画像検索は，キーワード検索，オブジェクト検索等，被写体の内容（「～が映っている」等）を検索のための手がかりとする手法が主流である．しかし，上記のような画像情報電子化の流れに伴い，再利用可能な画像が多様化しており，内容による検索だけでなく，「画像の構成が似ている」「画像の全体的な印象が似ている」等の，画像の性質を大まかに捉え，かつ人間の感覚により広く適応した画像特徴を手がかりとした「柔軟な」画像検索が有効な手段となることが予想される．

このような柔軟な画像検索の手法としては，画像全体の色，あるいはテクスチャに関する特徴分析に基づくもの^{1)～6)8)9)}や，さらに，特徴の画像内での大まかな分布，

いわゆる構図情報を利用するもの¹⁰⁾¹¹⁾などがある．これらの画像特徴は，言語化は難しいが全体的に見て似た印象のある画像を関連づける上で重要な情報であると考えられる．本稿では，白黒画像を処理対象とする場合を考慮し，テクスチャ特徴と構図情報による画像検索について考える．

テクスチャ特徴による画像分類および画像検索手法としては，画像全体から得られるエッジ方向特徴量を用いる手法¹⁾，注目画素対の角度，距離の変化および輝度差により特徴ベクトルを構成する手法²⁾，画像の離散コサイン変換(DCT)係数を用いる手法³⁾，DCT係数配列の分解により，疑似的に生成した多重解像度画像を用いる手法⁴⁾，Daubechies Wavelet変換係数を解像度毎に比較し画像を検索する手法⁵⁾，Gabor Wavelet変換係数から構成される特徴ベクトルにより検索する手法⁶⁾等が提案されている．しかし，これらの手法で用いられるテクスチャ特徴は，エッジ量とその方向，あるいは周波数とその方向等の種類の少ない情報から成るため，シンボルマーク等の単純な画像の検索や，テクスチャのみの分類等には適しているが，より多種の画像を対象として，柔軟かつ効率的な検索をするためには，さらに，テクスチャ全体の粗密，分布，複雑さ等の多面的な情報を含む画像特徴を導入する必要がある．そのような画像特徴として，フラクタルモデルに基づくテク

2001年1月31日，合同研究会(ヒューマンインフォメーション他)で発表
2002年1月17日受付，2002年12月27日最終受付，2003年2月7日採録
[†]NHK放送技術研究所

(〒157-8510 世田谷区砧 1-10-11, TEL 03-5494-3229)

[†]NHK (Japan Broadcasting Corporation) Science and Technical Research Laboratories

(1-10-11, Kinuta, Setagaya-ku, Tokyo 157-8510 Japan)

スチャ特徴が挙げられる。フラクタルモデルは、自然の造形物のモデルとして優れていることから、自然風景画像の特徴をとらえるための画像特徴として優れており⁷⁾、上述の画像特徴に比べ、より多種の実画像検索に適しているといえる。その代表例として、画像全体の輝度変化に関するフラクタル次元による画像検索手法⁸⁾⁹⁾が提案されている。ただし、検索対象画像が自然風景、またはそれに近いテクスチャのみで構成されるとは限らない場合には、人間の類似性の感覚と検索結果との差異が大きくなる場合があり、改善の必要がある。

構図情報に重きを置いた画像特徴抽出法は、画像を分割したサブブロックごとの正規化した濃淡情報を特徴として画像を検索する手法¹⁰⁾、色情報の空間分布に基づき画像を縦横方向へ分割し、各矩形領域から抽出したカラーヒストグラムを特徴とする手法¹¹⁾等、画像を固定されたブロック状に粗く分割する手法が主流である。しかし、これらの手法は、人間の感覚にとっては類似性が高いと思われる微妙な特徴分布の違いに対し、類似性が低くなるという問題がある。

本稿では、従来のフラクタル次元、および固定的にブロック分割する手法におけるこのような問題点を指摘するとともに、それらに対する改善策を提案する。具体的には、テクスチャの複雑性に関し、フラクタル次元よりさらに多面的な情報を持つと考えられるフラクタルシーケンスを提案し、これを検索で用いる画像特徴とする。さらに、構図情報検出のために、画像上に複数設置する特徴抽出サブブロックを、画像特徴に応じて位置可変にする手法を提案する。これらの手法を組合せることにより、従来手法に比べ「柔軟」な面をより強く持つ画像検索手法を実現できた。各改善策の効果をいくつかの検索実験により考察し、本提案手法が従来手法より優れていることを示す。

2. 画像のテクスチャに関する情報の表現

画像のテクスチャに関する情報の表現として、すでに述べたように様々な画像特徴が提案されている。その中でも、テクスチャ全体の粗密、分布、複雑さなどの情報を複合的に含む画像特徴として、フラクタル次元分析による手法が注目されている⁸⁾⁹⁾。以下、フラクタル次元による画像特徴分析の概要とその問題点について述べた後、本稿で新規に提案するフラクタルシーケンスとその有効性について述べる。

2.1 フラクタル次元による画像特徴分析

輝度変化のフラクタル次元は、画像の自己相似性を表す画像特徴の一つであり、複雑な画像ほど次元は大きく、単純な画像では次元が小さくなるという性質を持つ。したがって、画像のテクスチャの複雑性をうまく表現でき、人間が画像の初見から受ける印象に強く影響する特徴とも考えられている¹²⁾。フラクタル次元の計算方法として、掛村らは次のような輝度変化に関するフラクタル次元を提案し

図1 仮想体積
The virtual volume

ている⁸⁾。

- (1) 粗視化尺度系列 $r_0, r_1, \dots, r_N$ を準備する。
- (2) $r = r_0$ とする。
- (3) 画像を一辺 r 画素の正方形ブロック $N(r) \times M(r)$ に分割
- (4) 四角の画素の輝度値から構成される立体の「仮想体積」 $v(r, i, j)$ (図1) を計算する。 ($i = 0, \dots, N(r) - 1, j = 0, \dots, M(r) - 1$)
- (5) 全 $v(r, i, j)$ の正方形ブロックあたりの平均 $V(r)$ を計算する。
- (6) r を粗視化尺度系列に従って増加させ 3へ戻る。
 $r = r_N$ まで終了したら (7) へ
- (7) 横軸 $\log r$, 縦軸 $\log V(r)$ でプロットしたグラフの傾きがフラクタル次元であり、切片をその画像の複雑度と呼ぶ。フラクタル次元および複雑度の二つの値を画像特徴とする。

このようにして求めたフラクタル次元は、プロットしたグラフがほぼ直線になることを前提としている。実際、天体の物体表面、山肌、葉の集まり、空、海面の画像などの自然界に存在するテクスチャについては、プロットしたグラフがほぼ直線となり、それぞれが独自のフラクタル次元を持つことが知られている。しかし、その前提は自然界に存在するテクスチャから構成される画像以外では成立するとは限らない。よって、対象画像の種類を限定できない場合には、精度の良い分類および検索を行う事は困難であると思われる。例として、図2に大きさ 256×256 の二枚の PPM 画像を、粗視化尺度系列 $r = \{6, 8, 11, 16, 23, 32, 45, 64, 91, 128\}$ で上記のアルゴリズムにより、画像特徴の計算を行った時のグラフを示す。ここで、計算可能な尺度の最小値 $r = 2$ ではなく $r = 6$ から始めているのは、画像中の細かいノイズの影響を受けにくくするためである。画像 (a) と画像 (b) では、明らかに異なる印象を与えるテクスチャであり、導出されるグラフの形も大きく異なる。しかし、画像特徴、つまりグラフにフィットさせた直線の傾き、および切片を比較するとほぼ同じとなり、その結果、フラクタル次元を尺度とする類似度が高くなるという問題がある。

図 2 フラクタル次元による画像特徴の問題点
A problem with a image feature using the fractal dimension

図 3 テクスチャ画像例
Examples of texture images

2.2 フラクタルシーケンスによる画像特徴分析

2.1節にフラクタル次元による画像特徴を用いたテクスチャ画像検索の問題点を挙げたが、これは、計算時に導出されるグラフの形状を考慮せずに直線で近似することに主な原因があると考えられる。筆者らは、グラフの形状にも画像の重要なテクスチャ情報が含まれると考え、2.1 節に記述した輝度変化フラクタル次元の際に導出される $\log V(r)$ の数値列全体を、画像のテクスチャ的な情報を表す画像特徴として用いることにした。以下、この画像特徴をフラクタルシーケンスと呼ぶこととする¹³⁾。

2.3 フラクタルシーケンスの評価実験

画像特徴として、フラクタルシーケンスを用いる効果を、簡単なテクスチャ画像の検索実験によって示す。検索対象は、自然テクスチャに加え人工物表面のテクスチャも含む 256×256 の PGM 画像 322 枚である。画像例 10 枚を図 3 に示す。

<実験のための各設定>

粗視化尺度系列

$$r = \{6, 8, 11, 16, 23, 32, 45, 64, 91, 128\}$$

フラクタル次元特徴

$$\vec{F}_1 = [a, b]$$

求め方は 2.1 節参照。 a : フラクタル次元, b : 複雑度

フラクタルシーケンス特徴

$$\begin{aligned} \vec{F}_2 &= [f_0, f_1, \dots, f_{k-2}, f_{k-1}] \\ &= [\log V(2), \log V(3), \dots, \log V(64), \log V(128)]. \end{aligned}$$

k は尺度系列の要素数である。 $V(r)$ についても 2.1 節参照。

参照画像 R から計算したフラクタル次元特徴を $\vec{F}_1(R) = [a(R), b(R)]$ 、フラクタルシーケンス特徴を $\vec{F}_2(R) =$

図 4 フラクタル次元特徴による検索実験 1
The retrieval test 1 using the fractal dimension feature

図 5 フラクタル次元特徴による検索実験 2
The retrieval test 2 using the fractal dimension feature

$[f_0(R), f_1(R), \dots, f_{k-1}(R)]$ とし、検索対象画像 T からのフラクタル次元特徴を $\vec{F}_1(T) = [a(T), b(T)]$ 、フラクタルシーケンス特徴を $\vec{F}_2(T) = [f_0(T), f_1(T), \dots, f_{k-1}(T)]$ と表す。まず、フラクタル次元特徴を用いて検索実験を行った場合の $\vec{F}_1(R)$ と $\vec{F}_1(T)$ との画像特徴距離 $d_1(T)$ を式 (1) のように定める。この値が小さい程、類似度が高いものとする。

$$d_1(T) = \sqrt{(b(T) - b(R))^2 + (a(T) - a(R))^2} \quad (1)$$

また、フラクタルシーケンス特徴を用いて検索を行った場合の $\vec{F}_2(R)$ と $\vec{F}_2(T)$ の画像特徴距離 $d_2(T)$ は式 (2) のようにして計算する。

$$d_2(T) = \sqrt{(\bar{f}(T) - \bar{f}(R))^2 + \sum_{i=0}^{k-1} \{(f_i(T) - \bar{f}(T)) - (f_i(R) - \bar{f}(R))\}^2} \quad (2)$$

ここで、 $\bar{f}(T) = (\sum_{i=0}^{k-1} f_i(T))/k$ 、 $\bar{f}(R) = (\sum_{i=0}^{k-1} f_i(R))/k$ である。式 (2) の根号の中の第 1 項は、2つのフラクタルシーケンスをグラフにした時の「グラフの高さの差」であり、直線近似でいえば切片の差に相当するものである。この項は式 (1) の第 1 項に相当する。また第 2 項は、ベクトルの

図 6 フラクタル次元特徴による検索実験 3

The retrieval test 3 using the fractal dimension feature

各要素から要素の平均値を減じて正規化したベクトル同士の平方距離であり，グラフの形状の類似性を表す距離とみなすことができる．

<比較実験>

フラクタル次元特徴によるテクスチャ画像検索実験の結果を図 4，図 5 および図 6 に示す．各図とも，上段最左の画像が参照（問い合わせ）画像であり，その右側は，参照画像と類似度の高い上位 3 枚の画像である．また，下段左のグラフは，参照画像および，結果最上位画像についてのフラクタル次元計算の際に導出されるグラフ，つまりフラクタルシーケンスを示す．そして，下段右は，左のグラフに直線をあてはめたもので，傾きがフラクタル次元，切片が複雑度を表す．

各実験とも，主観評価ではあるが，参照画像と類似した画像が結果上位に来ているとはいえない．下段の二つのグラフを見てわかる通り，2.1 節で指摘した問題，つまり，参照画像と結果上位の画像では，グラフとしては異なる形をしているにもかかわらず，その違いを反映できない直線でグラフを近似してしまうことに起因した問題が生じている．これに対し，フラクタルシーケンス特徴で検索を行った結果を図 7，図 8 および図 9 に示す．フラクタル次元の場合と同様，上段最左の画像が参照（問い合わせ）画像であり，その右側は，参照画像と類似度の高い上位 3 枚の画像である．また，下段には，参照画像および，結果最上位画像のフラクタルシーケンスのグラフを下段に示した．

フラクタル次元特徴の場合に比べ，参照画像と類似した特徴を持つ画像が上位となっている．これは，フラクタルシーケンス特徴が，フラクタル次元特徴では表現できないテクスチャ情報を含むグラフ形状の異同を考慮しているためであり，実際に下段のグラフをみると，類似度上位の画像は参照画像とよく似たグラフ形状を持っていることがわかる．以上の実験により，画像全体のテクスチャ特性を手がかりとする検索を行うための画像特徴としては，フラクタル次元による画像特徴よりもフラクタルシーケンスが適していることがわかる．

図 7 フラクタルシーケンス特徴による検索実験 1

The retrieval test 1 using the fractal sequence feature

図 8 フラクタルシーケンス特徴による検索実験 2

The retrieval test 2 using the fractal sequence feature

図 9 フラクタルシーケンス特徴による検索実験 3

The retrieval test 3 using the fractal sequence feature

ここで，当然フラクタルシーケンスの方が特徴ベクトルの次元数が大きくなるため，類似度計算時間の問題を考える必要がある．しかし，CPU が Pentium 2GHz 程度の市販計算機を使用し，本比較実験（画像 322 枚）の合計類似度計算時間を測定したところ，フラクタル次元，フラクタルシーケンスどちらの場合も 0.01 秒未満であった．比較対

象画像数が数百枚～数千枚程度であっても、実用上問題ない程度の計算時間で済むと予想される。

3. 画像の構図に関する情報の表現

画像特徴の画像内での分布を大まかに表現した情報、いわゆる構図情報も、人間の類似感覚を左右する非常に重要な情報である。構図情報を利用した画像検索手法としては、画像をサブブロックに分割し、各ブロック単位で抽出した特徴を対応するブロック同士で比較する手法が一般的である。この章では、ブロック分割する際に、画像を固定的に分割する従来手法の問題点について述べ、その解決策として、Active Net を利用した位置可変なサブブロック配置手法、およびその評価実験について述べる。

3.1 位置固定な特徴サブブロック配置

従来の分割は、被写体の位置や画像の構成に関わらず、固定的に画像を分割する方法がとられてきた。しかし、固定的な分割によるサブブロック設定法では、背景の画像特徴がほぼ同じでも、物体領域の位置が若干異なる二つの画像において、その類似度が人間の感覚に比べて低くなるという問題がある。図 10 に、この問題が発生する例を示す。右の画像では、カメラのパンニング、チルトニングあるいは被写体である人物の移動により、左の画像に比べ被写体の位置が若干異なっている例である。被写体の位置ずれが小さいものであっても、固定的にブロックを配置すると、ブロック単位での画像特徴へ与える影響は小さくない。特に被写体輪郭近くのブロック（図 10 右側の灰色のブロック）では、左の画像との類似度の低下が人間の感覚以上に大きくなってしまふ。この類似度の低下を抑制するためには、被写体の輪郭をある程度「追従」して、特徴ブロックを設定する方法が有効であると考えられる。

3.2 Active Net を利用した位置可変な特徴サブブロック配置

本稿では、物体に追従した特徴サブブロックの設定法として、Active Net と呼ばれる動的網モデル¹⁴⁾をある画像特徴に依存して変形させ、変形後の網格子点上にサブブロックを設定する手法を提案する¹⁵⁾。

Active Net の変形は、網の持つエネルギーを小さくする方向に行われる。網の持つエネルギー E_{net} は式 (3) で表される。

$$E_{net} = \sum_i \sum_j \{E_{int}(\vec{v}(i, j)) + E_{img}(\vec{v}(i, j))\} \quad (3)$$

ここで、 (i, j) は網格子点を識別するインデックスで ($i = 0, 1, \dots, N-1, j = 0, 1, \dots, M-1 : N, M$ はそれぞれ縦横の網格子点数)、 $\vec{v}(i, j) = (x(i, j), y(i, j))$ は網格子点の xy 座標、 E_{int} は、近傍の網格子点との距離および移動速度、方向等に関する網の内部エネルギー、 E_{img} は Net を適応させる対象画像の画像特徴に依存したエネルギー（画像適応エネルギー）を表す。画像適応エネルギーに関与

図 10 分割による特徴サブブロックの問題点
A problem with feature sub-blocks by the dividing image

する画像情報によって、各網格子点がどこへ移動する傾向を持つかを制御することができる。画像適応エネルギーとしては、

- ・輝度値に比例したエネルギー関数
- ・エッジ量に比例したエネルギー関数

がよく用いられるが、前者は網格子点が画像の明部または暗部へ移動しやすくなり、後者はより強いエッジ部分へ移動する傾向を示す。本手法における目的は、網格子点が物体を追従することであるので、網格子点の位置が概ね背景と被写体の境界位置付近に定まることが望ましい。そこで本手法では、

- ・フラクタルシーケンスの空間変化に比例したエネルギー関数

を画像適応エネルギーとした。このようにすると、網格子点はフラクタルシーケンスの変化の激しい部分に移動する傾向を示すようになる。一般に、被写体と背景では、視距離や表面の物理的性質の違いによって、テクスチャ的画像特徴に差が出ることが多いと思われるので、網格子点が被写体と背景の境界近くへ移動することが期待できる。シンプルな画像に対するサブブロック設定例を図 11 に示す。同図 (a) は Active Net の初期状態であり、(b) は画像のフラクタルシーケンスの空間変化の大きさを白黒濃淡画像で表したものである。このフラクタルシーケンスの空間変化の具体的な求め方は、以下の通りである。

- (1) まず画像上に N_{fp} pixel 毎にサンプル点を設定。図 11 の例は $N_{fp} = 3$ とした。
- (2) 各サンプル点を中心とした一辺 N_{sq} pixel の正方形についてフラクタルシーケンスを計算。図 11 の例は、 $N_{sp} = 20$ 、粗視化尺度系列 $r = \{2, 4, 6, 8\}$ とした。
- (3) 隣接 4 サンプル点とのフラクタルシーケンス特徴ベクトル距離 (式 (2) 参照) の差分を計算し、その値を 0 - 255 の輝度値に正規化。差分が大きい程高い輝度値を割り当てる。
- (4) 各サンプル点を中心とした一辺 N_{fp} の正方形領域の画素値をすべて (3) で求めた輝度値とする。

図 11(c) は、同図 (b) を画像適応エネルギーとして変形させた後の Active Net である。そして、変形後の Active Net の各網格子点の位置を中心として、特徴サブブロックを設定した図が図 11(d) である。被写体である円を追従

図 11 Active Net による位置可変な特徴サブブロック
Feature sub-blocks which positions are varied using Active Net

してブロックが配置されていることがわかる。ここで、式(3)を減少させる反復回数を大きな値に設定すると、網が激しく変形し画像全体を覆う形状を保持することができず、画像の狭い領域へ特徴サブブロックが集中して設置されてしまうことになる。よって、反復回数は小さい方が好ましい。種々の画像による予備実験から、10~15 回程度が上記条件を満たす反復回数であると判断され、本実験では、反復回数を 10 回とした。

3.3 位置可変な特徴サブブロック設定の評価実験

Active Net を利用して、特徴サブブロックを被写体に追従して設定した場合の有効性を、固定的な画像分割との比較実験により示す。

実験で用いた画像は、図 12 の 4 枚の画像で、それぞれ、640×480 を 240×180 のサイズへ縮小したものである。(b),(c),(d) の画像は、撮影したカメラのチルトアップに相当する平行移動より、(a) の画像の地平線がそれぞれ、10pixel, 15pixel, 20pixel 下方へ移動したものである。固定的な分割による特徴サブブロック設定で画像特徴抽出を行った場合と、Active Net の変形による特徴サブブロック設定で画像特徴抽出を行った場合とで、参照画像 (a) の画像特徴と、カメラチルトアップ画像 (b),(c),(d) それぞれの画像特徴との特徴距離を比較した。

<実験のための各設定>

特徴サブブロックの大きさ 60 × 40

(Active Net の網格子点数は 4 × 4)

粗視化尺度系列 $r = \{2, 3, 4, 5, 6, 7, 8, 16, 32\}$

画像の縮小によりノイズが抑えられていると考え、 $r = 2$ からとした。

フラクタルシーケンス特徴ベクトルの求め方

- (1) 固定的な分割および Active Net を使用した分割のいずれの場合にも、画像上に $N_x \times N_y$ 個の特徴サブブロックを設定。本実験では $N_x = N_y = 4$
- (2) 各サブブロック (i, j) からフラクタルシーケンス $\vec{F}'(i, j)$ を求める。
($i = 0, 1, \dots, N_x - 1, j = 0, 1, \dots, N_y - 1$)

図 12 位置可変特徴サブブロックの効果検証用実験画像
Images to test effects of feature sub-blocks which positions are variable

表 1 比較実験結果
Result of comparative test

	固定分割	Active Net 使用
F_a と F_b との特徴距離	5.675260	4.649536
F_a と F_c との特徴距離	7.064926	4.784359
F_a と F_d との特徴距離	8.311714	5.618629

図 13 図 12 の画像への位置可変特徴サブブロック
Feature sub-blocks which positions are variable on images in Fig.12

- (3) それら $N_x \times N_y$ 個のフラクタルシーケンスの各要素を並べた

$\vec{F} = [\vec{F}'(0, 0), \dots, \vec{F}'(i, j), \dots, \vec{F}'(N_x - 1, N_y - 1)]$
をこの画像のフラクタルシーケンス特徴ベクトルとする。

また、二つの画像 R, T それぞれから計算したフラクタルシーケンス特徴ベクトル $\vec{F}(R) = [f_0(R), \dots, f_{N_d}(R)]$, $\vec{F}(T) = [f_0(T), \dots, f_{N_d}(T)]$ の特徴距離は式 (2) を用いて計算する。ここで、 $N_d = 9N_x N_y$ (粗視化尺度系列の要素数 × サブブロック数) である。

<比較実験>

実験結果を表 1 に示す。表の左から 2 列目が固定的な分割で特徴サブブロックを設定した場合、3 列目が Active Net の変形を利用して特徴サブブロックを設定した場合の実験結果である。また、表中の F_a は図 12 参照画像 (a) から 3.3 節の方法で抽出したフラクタルシーケンス特徴ベクトルを表し、 F_b, F_c, F_d も同様にそれぞれ画像 (b),(c),(d) から抽出したフラクタルシーケンス特徴ベクトルを表す。

結果を見ると、すべての画像 (b),(c),(d) とも、Active Net の変形を利用して特徴サブブロックを設定する手法の

図 14 DB1 画像例
Example of images in DB1

図 15 DB2 画像例
Example of images in DB2

図 16 DB3 画像例
Example of images in DB3

図 17 DB1 による実験結果
The result of retrieval test for DB1

方が画像特徴距離が小さい、つまり画像 (a) → (b) → (c) → (d) と地平線の位置が若干シフトしているにもかかわらず、類似度の低下を抑えることができています。これは、図 13 を見てわかる通り、すべての画像において地平線を追従して特徴サブブロックを配置していることによる効果である。このように、Active Net の変形を特徴ブロックの設定に用いることで、構図に関して柔軟に対応した画像特徴の抽出が可能となっている。

4. 実画像を用いた画像検索実験

2章および3章で述べたフラクタルシーケンス、および Active Net による特徴サブブロック設定手法を用いて、実画像による検索実験を行い、本手法の検索能力を評価した。本実験で用いる画像 DB は以下の3種類である。

DB1 プロ野球の放送番組からキャプチャした 240×180 の画像 500 枚

DB2 あるドラマ番組からキャプチャした 640×480 の画像 1000 枚

DB3 640×480 の風景画像 500 枚

各画像に対する特徴ベクトルは、3.3節の手順で計算したフラクタルシーケンス特徴ベクトルを用いた。その際、特徴サブブロック設定はすべて Active Net を使用し、

$N_x = N_y = 4$ とした。また、粗視化尺度系列は、解像度がそれほど高くなく、画像中のノイズの影響が少ないと考えられる DB1 については $r = \{2, 3, 4, 5, 6, 7, 8, 16, 32\}$ とし、比較的解像度が高い DB2 および DB3 については $r = \{6, 8, 11, 16, 23, 32, 45, 64, 91\}$ を用いた。各 DB の画像例を図 14、図 15、図 16 に示す。

4.1 検索処理

参照画像 R の特徴ベクトルを $\vec{F}(R)$ 、DB 中の画像 T の特徴ベクトルを $\vec{F}(T)$ ($T = 0, 1, \dots, N_{DB}-1$; $N_{DB} = 500$ または 100) とした場合、 R と T との類似度 $S_F(T)$ を以下の式に従って求める。

$$S_F(T) = 1 - \{d_F(T) / \max_t d_F(t)\} \quad (4)$$

ただし、 $d_F(T)$ はフラクタルシーケンス特徴ベクトル $\vec{F}(R)$ と $\vec{F}(T)$ との画像特徴距離 (式 (2) 参照) を表す。すべての画像 T について $S_F(T)$ を計算後、 $S_F(T)$ の値が大きい順に順位づけし、上位 N_{HIT} 枚を検索結果とする。

4.2 実験結果

DB1、DB2 および DB3 による検索実験例の結果画像をそれぞれ、図 17、図 18 および図 19 に示す。各図とも Ref. {1, 2, 3} が参照画像であり、1 位～5 位が式 (4) により

図 18 DB2 による実験結果
The result of retrieval test for DB2

図 19 DB3 による実験
The result of retrieval test for DB3

計算した参照画像との類似度の高い上位 5 枚の画像である。各実験とも非常に類似した画像が検索結果上位として出力されていることがわかる。

4.3 検索精度評価

本章では、フラクタルシーケンスによる画像特徴を導入し、かつ特徴サブブロックを Active Net により位置可変にした効果を客観的に評価するために、いくつかの手法と

比較実験を行った。DB1 に対する検索実験 (図 17) の各参照画像に対する検索精度を計算するために、それぞれ下記のような画像を「正解」とした。ここで、DB1 を使用した理由は、カメラワークのバリエーションがそれ程多くない野球中継の画像であるため、「正解」を定めやすいからである。

図 17 Ref.1 背景の大部分がテクスチャの粗い観客席。被写体はバストショットからウエストショットの大きさの選手 (76 枚)。

図 17 Ref.2 画像の上部がテクスチャの細かい観客席。残りの大部分がフィールド (55 枚)。

図 17 Ref.3 背景の大部分がフィールド。被写体はバストショットからウエストショットの大きさの選手 (78 枚) ここで、本手法との比較手法を考える。

まずサブブロックを固定的に設定した場合との比較手法として、

- ・サブブロックを固定的に設定しフラクタルシーケンスを画像特徴とする手法

を適用し、特徴ベクトルの計算は、3.3 節と同様に行う。

また、フラクタル次元特徴による手法との比較は、

- ・Active Net を使用しフラクタル次元による画像特徴を用いる手法

を用いることとし、特徴ベクトルは以下のように計算する。

<フラクタル次元特徴ベクトル (サブブロック可変)>

- (1) Active Net 使用により、画像上に $N_x \times N_y$ ($N_x = N_y = 4$) 個の特徴ブロックを設定。ただし、Active Net 変形のための画像適応エネルギーは、3.2 節のフラクタルシーケンスの空間変化の求め方とほぼ同じ手順で、フラクタル次元ベクトル (2.1 節参照) の空間変化を白黒濃淡画像で表したものをを用いる。

- (2) 各特徴ブロックからフラクタル次元ベクトルを求める。

- (3) それら $N_x N_y$ 個のフラクタル次元ベクトルの各要素を並べた $2N_x N_y$ 次元ベクトルを、この画像のフラクタル次元特徴ベクトルとする。

さらに、フラクタル以外の主な比較手法として、1 章で挙げたような、DCT 係数や Wavelet 変換係数を利用する手法が考えられるが、どちらも周波数とその方向に重きを置いた画像特徴であり、ほぼ類似した検索精度になると考えられる⁴⁾。さらに、後者は本手法に比べ特徴計算時間を要するため、本稿では、

- ・Active Net を使用し DCT 係数による画像特徴を用いる手法

を採用し、特徴ベクトル計算を以下のようにして行う。

<DCT 係数特徴ベクトル (サブブロック可変)>

- (1) Active Net 使用により、画像上に $N_x \times N_y$ ($N_x = N_y = 4$) 個の特徴ブロックを設定。ただし、Active Net 変形のための画像適応エネルギーは、3.2 節のフラクタルシーケンスの空間変化の求め方とほぼ同

じ手順で、 N_{dct} 次元 DCT 係数ベクトル⁴⁾の空間変化の大きさを白黒濃淡画像で表したものをを用いる。本実験では $N_{dct} = 9$ とする。

- (2) 各特徴ブロックから N_{dct} 次元 DCT 係数ベクトル⁴⁾を求める。
- (3) それら $N_x N_y$ 個の DCT 係数ベクトルの各要素を並べた $N_{dct} N_x N_y$ 次元ベクトルを、この画像の DCT 特徴ベクトルとする。

本手法および各比較手法による検索精度を「適合率・再現率グラフ」にしたものが図 20, 図 21, 図 22 であり、各図とも、横軸が全正解数に対しての正解検出数の割合を表す「再現率」、縦軸が上位から採用する枚数 N_{HIT} に対する正解検出数の割合を表す「適合率」である。ここで「適合率・再現率グラフ」は、 N_{HIT} の増変により構成することができ、一般的には右下がりの曲線となる。高再現率に対し高適合率を維持している場合に、検索精度が高いと評価される。

各グラフは、太実線が本手法、短破線が固定ブロックによる手法、細実線がフラクタル次元特徴ベクトルによる手法、長破線が DCT 係数ベクトルによる手法による実験結果を表す。

まず、本手法の結果に注目すると、図 20 では、再現率約 90% で適合率約 90%、図 21 では、再現率約 70% で適合率約 85% となっている。これらのデータは、本提案手法により高精度の検索ができることを示している。図 22 では、再現率約 60% で適合率約 70% とやや低めになっているが、これは、参照画像の大部分の輝度変化が平坦であり、テクスチャ領域が少ないためである。

固定サブブロック手法との比較では、図 22 で本提案手法の有効性が顕著に現れている。本手法を用いて、図 17Ref.3 に設定した特徴サブブロックを図 23 に示す。人物の輪郭を追跡してサブブロックが配置されていることがわかる。図 22 に比べ、図 20, 図 21 では Active Net を用いる本手法の効果があまり見られないが、これは、Net 変形のエネルギーとして用いたフラクタルシーケンスの空間変化画像において、背景と被写体領域の分離精度が低く、Net の変形が理想的に行われていないことが主な理由と考えられる。この点は、今後改善して行く必要がある。

フラクタルシーケンスを用いないテクスチャ特徴による手法との比較においても、本提案手法が優れていることがわかる。フラクタル次元特徴ベクトルを用いた手法では、図 17Ref.2 のような、固有のフラクタル次元を持つテクスチャ(非常に細かい歓客席、芝生)から構成される画像の検索は高精度で行えるが、中距離撮影の歓客席のような、人工的で粗いテクスチャを広く含む図 17Ref.1 の場合、2.1 節で言及したような問題が生じ、検索精度が低下してしまう。DCT 係数特徴ベクトルを用いた検索の場合については、各実験とも提案手法より精度が低い。これは特徴ベクトルがテクスチャ情報の一部(周波数、その方向)しか

図 20 図 17Ref.1 による検索精度比較

The comparison of retrieval accuracies for Fig.17Ref.1

図 21 図 17Ref.2 による検索精度比較

The comparison of retrieval accuracies for Fig.17 Ref.2

図 22 図 17Ref.3 による検索精度比較

The comparison of retrieval accuracies for Fig.17 Ref.3

図 23 図 17Ref.3 への本手法による特徴サブブロック
Feature sub-blocks on Fig.17 Ref.3 using our method

表現できていないことに起因していると考えられる。

以上の実験により、種類を限定しない一般的な画像に対し高精度の検索を実現するために、テクスチャの粗密、分布および複雑性等のグローバルな情報を含むフラクタル

シーケンスおよび、Active Net を用いた位置可変な特徴サブブロック設定が有効であることがわかった。

5. ま と め

本稿では、画像の内容による従来の画像検索でなく、画像の性質を複合的に捉えた情報による、人間の感性に広く適応するであろう「柔軟な」画像検索を実現するための画像特徴抽出、および検索の新しい手法を提案した。まず第2章で、従来の限定的な画像特徴とは異なり、画像のテクスチャ情報を広い意味で捉えていると考えられる画像特徴フラクタルシーケンスについて述べ、比較実験等でその効果を示した。また第3章では、従来の固定的な画像分割ではなく、画像特徴に応じて位置可変に特徴ブロックを設定することで、位置ずれにロバストな構図情報を取得する手法について述べ、比較実験によりその有効性を示した。さらに4章で、実画像に対して提案手法による画像検索実験を行い、様々な画像に対し高精度で検索が可能なることを示した。

今後の展開としてはまず、色に関する特徴との組合せが挙げられる。今回、輝度変化に関するフラクタルシーケンスのみを画像特徴としたため白黒画像を扱ったが、メディアに存在する画像の大部分はカラー画像であり、かつ色も画像に対する人間の印象を大きく左右する重要な情報である。したがって、人間の色感覚に近い色特徴を抽出できれば、色特徴の類似度とフラクタルシーケンスの類似度とを重み付けして足し合せることにより、利用者の検索意図により適した検索が可能となる。さらに、利用者の意図に基づいて、この重み係数の配分を変化させることで、より柔軟な検索が可能になると考えられる。また、色の類似度で絞り込んだ後に、フラクタルシーケンスの類似度で順位を入れ換える等の、階層的な類似度計算処理も、検索精度や速度の向上を図る方法として有望である。

今後の課題としてはまず、Active Net 変形のエネルギーとして用いるフラクタルシーケンス空間変化による、背景と被写体領域の分離精度を良くすることが挙げられる。これにより、Net の格子点が被写体領域の輪郭付近へ安定して移動することを可能とし、検索精度の向上が期待できる。また、フラクタルシーケンスの定量的評価も必要である。フラクタルシーケンスは、多面的なテクスチャ情報を複合的に含むと考えられるが、それらの情報の具体的な性質を分析し明確にすることで、特徴抽出を制御可能とし、検索をより効率的に行える手法としたい。さらに、人間の感覚に関する研究も重要な課題として挙げられる。脳が強く影響を受ける刺激にはどのような画像特徴が影響しているのか、画像のどのような領域が注視されるのか等の、視覚と類似感覚の関係についてより詳細に解析することは、本手法を改良するために有用であると考えている。

本研究を進めるにあたり、研究の機会を与えてくださったNHK放送技術研究所山田宰前所長、三宅誠所長、有益

な助言を頂いた、同所ヒューマンサイエンス磯野春雄前部長、河合直樹部長ならびに各位に感謝する。

〔文 献〕

- 1) 水野, 小林, 吉田, 酒井: "エッジ方向特徴量による画像検索", 信学技報, NIM98-103 (1998)
- 2) Dmitry Chetverikov: "Texture analysis using feature-based pairwise interaction maps", Pattern Recognition 32, pp.487-502 (1999)
- 3) Sang-Mi Lee, Hee-Jung Bae and Sung-Hwan Jung: "Efficient Content-based Image Retrieval Methods Using Color and Texture", ETRI Journal 20, 3 (1998.9)
- 4) Yu-Len Huang and Ruery-Feng Chang: "Texture features for DCT-coded image retrieval and classification", Proc IEEE Int Conf Acoust Speech Signal Process, 6, pp.3013-3016 (1999)
- 5) Ce Chen and Ross Wilkinson: "Image Retrieval using Multiresolution Wavelet Decomposition", International Conference on Computational Intelligence and Multimedia Applications, pp.824-829 (1998)
- 6) B.S.Manjunath and W.Y.Ma: "Texture Features for Browsing and Retrieval of Image Data", IEEE Transactions on Pattern Analysis and Machine Intelligence, 18, 8, pp.837-842 (1996.8)
- 7) Pentland, A.P.: "Fractal-based Description of Natural Scenes", IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-6, 6, pp.661-674 (1984)
- 8) 掛村, 東, 入江: "仮想体積を用いたテクスチャ変数とその応用", 信学論, J80-D-II, 9, pp.2411-2420 (1997.9)
- 9) 林, 田中: "混合フラクタルモデルに基づく画像特徴量とその画像検索応用における性能評価", 情処学論, 38, 10 (1997)
- 10) 中嶋, クティチ, 中村, 富永: "濃淡レイアウトに基づく類似画像検索方式", 画電学誌, 27, 5, pp.540-547 (1998)
- 11) 山本, 岩佐, 竹村, 横矢: "色情報の空間分布を考慮した類似画像検索" 信学技報, NIM99-46 (1999)
- 12) シュプリングー・フェアラーク東京: "フラクタルイメージ ~理論とプログラミング~"
- 13) 望月, 伊藤: "グローバルな特徴を利用した画像検索", 第59回情処学全大, 特1-229 (1999)
- 14) 坂上, 山本: "動的な網のモデル Active Net とその領域抽出への応用", テレビ誌, 45, 10, pp.1155-1163 (1991)
- 15) 望月, 伊藤: "Active Net による画像分割法", 信学会総全大, D-12-85 (2000.3)

もちづき たかひろ
望月 貴裕 1994年、東京工業大学工学部情報工学科卒業。1996年、同大学院修士過程修了。同年NHK入局。同放送技術局報道技術センター中継制作部を経て、1998年より、放送技術研究所にて画像認識の研究に従事。正会員。

ふじい まひと
藤井 真人 1981年、名古屋大学工学部電気工学科卒業。1983年、同大学院修士過程修了。同年NHK入局。1987年より、同放送技術研究所に勤務。1998～2001年、ATR人間情報通信研究所へ出向。主に視覚のバイオサイバネティクス、画像認識、立体画像情報処理の研究に従事。現在、同ヒューマンサイエンス研究部副部長。正会員。

いとう たかゆき
伊藤 崇之 1977年、京都大学工学部電気工学科卒業。1979年、同大学院修士過程修了。同年NHK入局。1981年より、同放送科学基礎研究所、および放送技術研究所に勤務。主に視覚のバイオサイバネティクス、画像認識、立体画像情報処理の研究に従事。現在、同研究企画部担当部長。工博。正会員。