

色や形状等の表層的特徴量にもとづく画像内容検索技術

串間和彦[†] 赤間浩樹[†] 紺谷精一[†] 山室雅司[†]

デジタルカメラの普及や既存写真のデジタル化等により、大量の画像が日々蓄積されつつある。これらの効率的な管理のため、画像の内容検索技術が注目されている。

本論文では、まず、画像の内容検索技術を、逐次処理型探索モデルと蓄積型の内容検索モデルの2つに分類する。そのうちの蓄積型の内容検索モデルに対し、内容解析、照合、検索インタフェースという3つの観点から、現状の技術を分類整理する。さらに、この分類に従って、既存のシステムIllustra & VIR, QBIC, VisualSEEK, VP, ExSight を分析する。そして最後に、これらの分類と分析を通して得られた経験、および、我々が ExSight システムの実用化の過程で得られた経験をふまえ、画像の内容検索技術の今後の課題と研究の方向について考察する。

Content based Image Retrieval Techniques based on Image Features

KAZUHIKO KUSHIMA,[†] HIROKI AKAMA,[†] SEIICHI KONYA[†] and MASASHI YAMAMURO[†]

Content-based image retrieval has been addressed as one of the most promising solutions for access to rapidly growing stored images. Content-based image retrieval models are classified into two types depending on target applications; on-the-fly detection model and stored data search model. Since the latter are more strongly related to database research, this paper focuses on the stored data search model. The paper analyzes the current technologies in terms of three factors; image data analysis, similarity matching, and user interface. Then currently available or reported examples of stored model content-based image retrieval systems such as Illustra&VIR, QBIC, VisualSEEK, VP and ExSight are reviewed based on these factors. Following these reviews, the paper suggests a future research direction and issues for further study in this area.

1. はじめに

デジタルカメラの普及やスキャナの低廉化により、膨大な数の写真がデジタル化され蓄積されつつある。これらの写真を管理するために、従来は人がテキストで事前に説明を付け、キーワードにもとづいて検索を行っていた。しかし、数万枚を越える写真に人手で説明を付けるのは大変な作業である。また、人によって説明の付け方がバラバラであったり、説明自体が不十分なため、希望する写真が検索できないケースもあった。

こうした問題を解決するために、画像全体の色や写っている物の形状等、画像内容にもとづく検索技術が提案され、実用化が図られている。しかし、今後予想される数百万枚規模の画像を対象とした精度の良い検索のためには越えるべき課題が残されている。

本論文では、色や形状といった特徴量を多次元空間上で表現し、多次元空間内の近傍検索をインデックス技術の利用により高速に実行する蓄積型の画像検索モデルを対象として、従来提案されている技術を分類し、評価する。さらに、高精度かつ高速な内容検索を行う場合の問題点と解決方法について考察する。

2. 内容検索技術の概要

静止画、動画、音楽といったいわゆるマルチメディア情報に対して、従来は作品名や作者名、画像の解説等を人間がテキストで索引付けし、キーワードにもとづいて希望する画像を検索する方法がとられてきた。これに対し、図1に示すようにマルチメディア情報そのものをキーとして、それに類似した内容をもつ他の情報を検索する手法が内容検索技術である¹⁾。

例えば画像情報の場合は画像全体の色合い等にもとづいてキー画像と対象画像間の類似度を決定し、類似したものを返す。音楽の場合には、カラオケで歌いたい曲を探すために客のハミングをもとに、類似したメ

[†] NTT 情報通信研究所
NTT Information and Communication Systems Laboratories

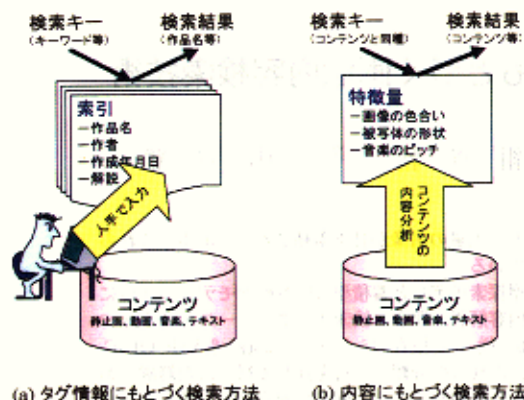


図1 内容にもとづく検索方法
Fig.1 Content Based Retrieval

ロディーの曲を検索するような場合が考えられる。これらの内容検索ではどのような特徴量を用いるかが検索精度に大きく影響する。さらに一般的にアナログ量として表現される特徴量をもとに、完全一致でなく、類似した情報を検索する必要があり、総当たりの比較計算が必要で、計算コストがかかる問題がある。

これ以降では、マルチメディア情報のうち最も多く生成されつつある画像情報を中心に論じる。内容にもとづいて画像を検索する技術として、逐次入力される入力画像の中でキーに近い部分をその都度見いだす逐次処理型の探索技術と、データベース化された大量の画像の中からキーに近い画像を高速に検索する蓄積型の検索技術がある。

2.1 逐次処理型の探索モデル

逐次処理型の探索は逐次入力される画像(イメージ)内をその都度探索し、キーとなる画像(モデル)と一致する部分を決定する。キーと一致するものを正確に見つけられる反面、入力画像中をキーの大きさや開始位置を少しづつずらしながらスキャンするため、計算コストがかかる問題があり、如何にして検索範囲を限定するかが重要となる。キー画像と入力画像の色分布をRGB尺度での画素数の積算で表し、両者の間でベクトル除算を行うことで、キーに特徴的な色使いをもつ部分を高速に決定するBack Projection法⁹⁾は代表的な手法である。さらにキー画像をもとにして最初は画像中を粗くスキャンし、類似した部分が見つかったら間隔を狭めて詳細に検査するCoarse-to-fine型¹⁰⁾の探索法や、画像中のある領域の類似度からその近傍の類似度の上限値を計算することで、近傍領域での照合を省略するアクティブ探索法¹¹⁾等が提案されている。これらは特定の前処理を必要としないため、ライブ映像や報道写真等、時時刻々と生成される画像に対しても

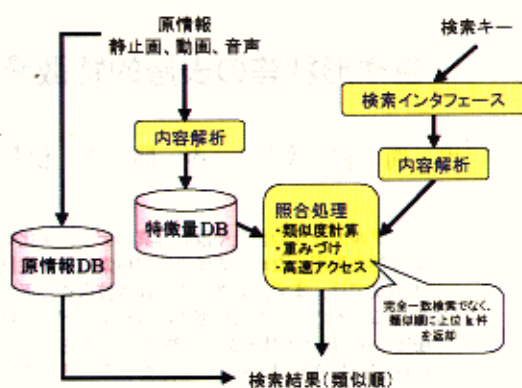


図2 内容検索のモデル(蓄積型)
Fig.2 Model of Stored Type Content Based Retrieval

即座に対応できる反面、画像を一枚毎に処理するために計算コストがかかり、数万枚を越えるような大量画像の検索には適さない問題がある。

2.2 蓄積型の内容検索モデル

これに対し、蓄積型の内容検索は事前に入力画像を解析して画像特徴を計算し、インデックスをつけたデータベースに格納することで、高速な検索を行う方式である。蓄積型の内容検索システムとしては商標・意匠パターンを対象としたTRADEMARK¹²⁾が1989年に提案された⁵⁾。その後、特定領域に依らない汎用の検索機能、かつ特徴量に対する高速なインデックス機構を備えたシステムが出現した。Illustra& VIR¹³⁾はその先駆的なシステムであり、画像全体の色合いや模様(テクスチャ)、色配置にもとづいた検索を行う。最近では、QBIC¹⁴⁾、VisualSEEk¹⁰⁾、VP¹¹⁾、ExSight¹⁵⁾~¹⁸⁾等、画像全体の特徴だけでなく、色領域の配置関係や画像内オブジェクトの形状等を用いることで、より精密な検索を指向する方式が提案されている。

本論文ではデータベース技術とより関連の深い後者の蓄積型に着目して、現状技術の課題を分析する。

3. 蓄積型内容検索のモデルと課題

蓄積型の画像検索は図2のようにモデル化できる。はじめにデータベースへのデータの入力が行われる。この段階では画像の内容を解析し、計算可能な特徴量を抽出する(内容解析フェーズ)。得られた特徴量をデータベースに格納し、同時に高速な検索のためのインデックスを作成する。次に、別の契機において、利用者が適当な検索インタフェースにより、欲しい画像に類似したキー画像を指定する。このとき、与えられたキー画像から計算された特徴量と、データベース内の特徴量間で照合が行われ、結果が返される(照合フ

ーズ) . よって、蓄積型の画像検索を行う場合のシステム側の課題は、以下の3点に要約できる。

- (1) **内容解析の課題**: 画像内容を如何に忠実に、かつ、自動的に計算可能な特徴量に変換するか。
- (2) **照合の課題**: 検索キー画像とデータベース内の画像間で特徴量の照合を如何に高速に行うか。
- (3) **検索インタフェースの課題**: 検索のキーとなる画像を如何に容易に利用者に指定させるか。

本稿では上記の3つの課題に対し、各種の画像検索システムで用いられる手法を紹介しながら、それらが上記課題に対して、どうアプローチし、どう解決しているかを論じる。

4. 内容解析

4.1 全体画検索と部分画検索

検索の目的に応じて、画像全体を対象とした全体画検索と、画像内の部分に着目した部分画検索がある。全体画検索は画像全体の色合い等にもとづいて類似画像を検索するもので、海や山の風景といった背景画やカーテン等のパターン画像の検索に適する。これに対し、部分画検索は例えば、建物や車といった特定のオブジェクトが含まれる画像を検索するのに用いられる。それぞれの場合に用いられる特徴量を図3に示す。全体画検索が画像全体を対象に特徴量を決定するのに対し、部分画検索では先ず画像内のオブジェクト(画像オブジェクト)を抽出する必要がある。

精度良く、かつ人手をかけずに画像オブジェクトを決定することを目的として種々の手法が提案されている。

本章ではまず、図3に示した各種の特徴量について、それらの概要と決定方法を述べる。続いて、部分画検索の場合に必要な画像オブジェクトの抽出方法について紹介する。

4.2 画像内容検索で用いられる特徴量

4.2.1 色情報

色は全体画検索、部分画検索のいずれに対しても有効な特徴であり、その表現にはColor Histogramが主に用いられる。RGB色尺度で表現された画像全体、もしくはその一部分(画像オブジェクト)に関して、図4に示すようにRGBの各色範囲毎に画像内の画素値を積算することで各RGB値を横軸とするヒストグラムを構成できる。実際にはRGBの256階調そのものを用いると人間にとって近い色でも計算上、別の色相として判断されるため、次元縮退が行われる。例えばR, G, Bのそれぞれについて16次元に縮退させ、計48次元程度のヒストグラムが使用される。最近では図5に示

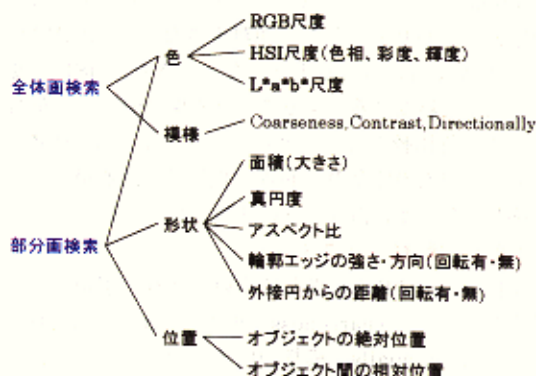
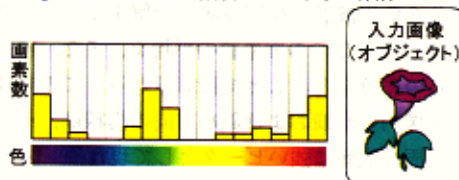


図3 内容表現に用いられる特徴量の種類
Fig.3 Features used for Content Description

Step1 画素値の出現頻度のヒストグラム作成



Step2 画像サイズでの正規化

→ n次元ベクトルとみなす

図4 色情報の抽出

Fig.4 Determining the Color Feature

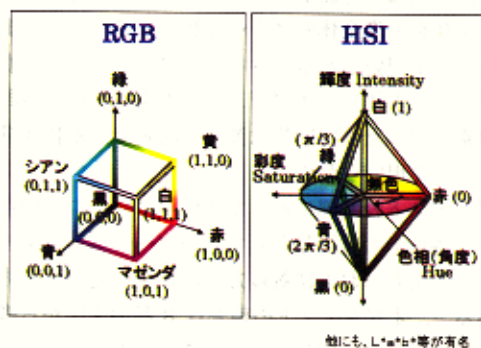


図5 色空間
Fig.5 Color Space

すように色相 (Hue), 輝度 (Saturation), 明度 (Intensity) によって色を表す HSI 尺度¹⁶⁾や, 人間の感覚に合わせて2色間の距離を調整した $L^*a^*b^*$ 尺度¹⁶⁾が多く用いられるようになってきた。HSI 尺度では色相, 彩度に着目することで画像の明るさに左右されずに色比較を行うことができる。また $L^*a^*b^*$ 尺度では空間内の任意の2色間の幾何学的距離が, 人間の感じる類似度に沿うように調整されている。

4.2.2 模様 (テクスチャ)

模様は主に全体画に関して用いられる。QBIC⁹⁾では模様を表すのに coarseness (粗さ), contrast (明瞭さ), directionally (模様の向き) の3つのパラメータを用いている。模様の取得には, 同時生起行列やフーリエ変換を用いた手法などが存在する。それらの詳細は文献 16 に譲るが, 本論文では, 近年, 注目されているウェーブレット変換^{17),18)}を用いた手法を示す。ウェーブレット変換では, 画像を平均 (低周波) と差分 (高周波) の2つに分解する過程を横方向と縦方向に交互に繰り返すことにより, 画像が持つ空間的な周波数成分を低周波層と高周波層へ分解する。ここで平均とは隣り合う2つの画素値の平均を, また, 差分とはそれらの差分をとることで計算される。この処理により, 画像内の縦横および斜め方向の模様は, そこに対応した周波数層に強いピークとして現れる。図6の例では上記分解の過程を2度 (レベル2まで) 繰り返すことで16の周波数層から16次元の特徴量を得る過程を示している。ウェーブレット変換は他の手法に比べ, 処理が容易である点と画像内の位置情報が保存される点が有利であると考えられている。

4.2.3 形状

部分画検索オブジェクトに対しては形が重要な特徴である。オブジェクトの形に関連する特徴量のうち最も単純な面積の特徴量は, オブジェクトを構成するピクセル数を画像全体のピクセル数にもとづいて正規化したものが用いられる。さらにオブジェクトの縦横比 (アスペクト比), オブジェクトの面積をその外周長の自乗で割った真円度等は簡単に決定できる特徴量であり, 広く用いられている¹⁹⁾。

オブジェクトの輪郭に沿って等間隔にいくつかの点を取り, 各点での特徴量として次の点へのベクトルの向きを上下左右, および斜め方向の8種で近似する方法がある²⁾。オブジェクトの輪郭は8種の文字から構成される有限長の文字列で表現できる。形状の類似度

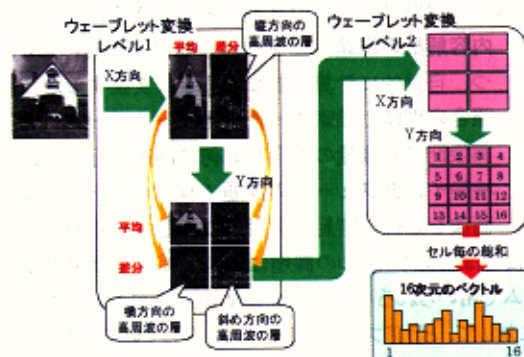


図6 ウェーブレット変換を用いたテクスチャ情報の抽出
Fig.6 Determining the Texture Feature using Wavelet Transformation

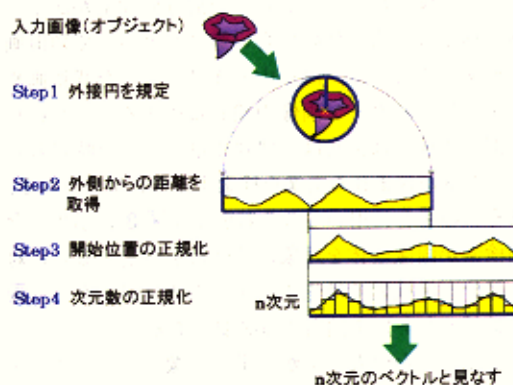


図7 外接円を使った形状情報抽出
Fig.7 Determining the Shape Feature using Outermost Circle

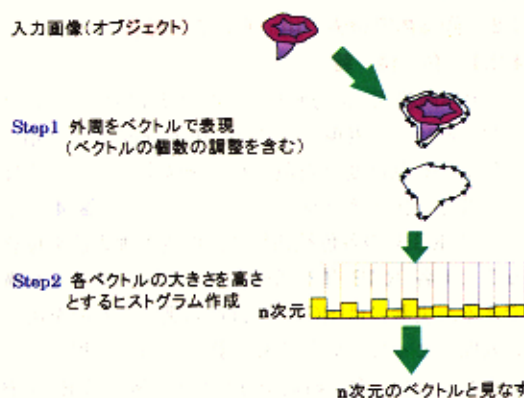


図8 外周ベクトルを使った形状情報の抽出
Fig.8 Determining the Shape Feature using Contour Vector

* HSI 色空間は RGB 色空間をマンセル表色系に合わせるように変換した色空間である。

は2つの文字列間での最大一致文字列を見出すことで計算できる。ただし、この方法は手続的な処理(文字列間の最長一致)を必要とするため、大量画像検索には適さない問題がある。

他の方法として外接円からの差分距離を用いる方法¹⁹⁾、エッジベクトルを利用する方法等がある。外接円を用いる方法では図7に示すようにオブジェクトの輪郭から外接円までの距離を計算し、それを適当な次元に区切ってヒストグラム化する。その際、形状の特徴点を適当に定め(例えば最も尖った場所)、それを原点とすることで回転に対応することもできる。エッジベクトルを利用する方法ではオブジェクトの輪郭に沿ったベクトルの方向と強さ(一般にははっきりした境界では強く、あいまいな境界では弱い)を計算する(図8)。

4.2.4 オブジェクトの位置

入力画像中に複数のオブジェクトが存在する場合、それらの位置関係は検索上重要な要素である。例えば青い服を来た人物といった検索では肌色の部分が青い部分の上にあるのを利用することで効率的な絞り込みを行うことができる。

オブジェクトの位置情報として画像中での絶対位置、および複数オブジェクト間での相対位置がある。絶対位置に関してはx, y座標を画像の大きさに応じて正規化したものを用いる。これに対し、オブジェクト間の位置関係を表現する手段としてはVisualSEEKで提案された2D-Stringがある¹⁰⁾。入力画像中にA, B, Cの三つの領域が存在し、それぞれの重心位置が左から順にA, B, Cで下から順にB, A, Cであれば、x軸方向には“A<B<C”、y軸方向には“B<A<C”という計2つの文字列で表現する。照合時にはこの文字列をもとに、検索された領域が指定された関係にあるかをチェックする。その際、2D-Stringからに基づいて2領域間の関係(領域1, 領域2, 方向)を予め全て計算しておき、領域名1にもとづくハッシングと、その中で領域名2にもとづくInverted Arrayを作ることで照合を高速化できる。さらに、2D-Stringの応用として、並びの順番以外にも包含、重なりを表現する試みがなされている。

4.3 部分画像の認識方法

4.3.1 グリッド分割法

入力画像をグリッドによって、 $N \times M$ 個の長方形のブロックに分割し、各ブロックについて特徴量を計算する。探したいオブジェクトをスケッチによりキー画像中に描画すると、キー画像自体も $M \times N$ のブロックに分割され、特徴量が計算される。照合では描画されたブロックに関して入力画像内の対応する位置のブロック、およびその近傍のブロックとの間で類似度が計算される。この方法はPartition-

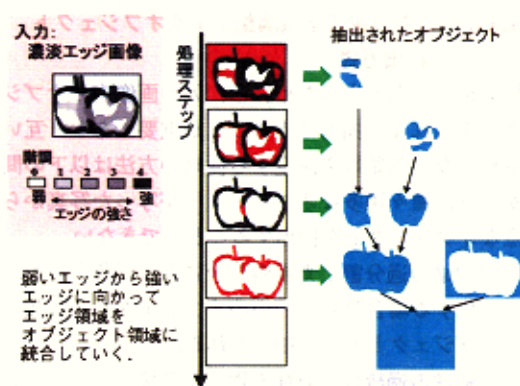


図9 しきい値の全域調整によるオブジェクト抽出
Fig.9 Object Extraction by varying Edge Magnitude Threshold

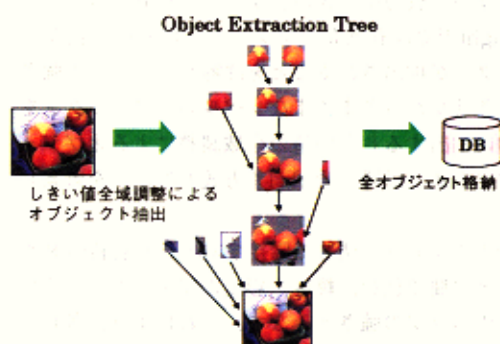


図10 オブジェクトの抽出過程の例
Fig.10 Example of Object Extraction Process

based Query-by-PaintingとしてQBICで提案された⁹⁾。

4.3.2 色情報にもとづく領域の決定

グリッド分割はオブジェクトを長方形のブロックで近似するために、近似の誤差が生じると共に大きさや位置が異なる場合に検索できない。そのため、画像内を色情報にもとづいて複数領域に分割する手法が多く用いられ、クリップアートのように色合いのはっきりした画像については有意なオブジェクトを精度良く抽出している。QBIC⁹⁾、VisualSEEK¹⁰⁾等では画像内をお互いに重複しない複数の色領域に分割し、検索対象としている。Visual-SEEKでは予め定められた色集合にもとづいてバックプロジェクションにより色領域を決定する。QBICでは画素を色と位置情報によってクラスタリングすることで色領域を決定している。しかし、一般的な写真を対象とした抽出手法は確立されておらず、例えばQBICでは人間がオブジェクトの輪郭等を指定することで切り出しを補助する手段を提供している。

4.3.3 しきい値の全域調整によるオブジェクト

抽出方式

検索精度を向上するためには、入力画像内のオブジェクトをもれなく抽出できることが重要である。互いに重複しない色領域を決定する従来の方法は以下の問題があり、種々雑多なオブジェクトが写った写真から有意なオブジェクトを抽出することができない。

- (1) **領域の過分割**: 光線の具合によって色合いが変化する場合等に、一つのオブジェクトが複数オブジェクトとして抽出される。
- (2) **領域の過統合**: 同種の物体が隣接して写っている場合に個々のオブジェクトを抽出できず、全体を一つのオブジェクトとして抽出してしまう。

上記問題を解決するためにエッジ情報を利用して色領域を分割/統合する方法が考えられる。しかし対象に依らない一般的なしきい値はなく、汎用的ではない。

ExSightTMでは有意なオブジェクト以外に冗長なオブジェクトが抽出されることを許容し、エッジの強さを変化させながらさまざまなレベルのオブジェクトを段階的に抽出するしきい値の全域調整によるオブジェクト抽出方式(以降、全域調整方式と呼ぶ)を考案した。

画像内の全画素に関して隣接する上下左右斜の8つの画素との間で色相、輝度、彩度の差を計算し、その座標でのエッジの強さとみなす。これにより、各点におけるエッジの強さを画素値とした濃淡エッジ画像を作成できる。図9に示すように画素値1以上のエッジを有効にした段階から、順次弱いエッジを削除して行くと、分割された領域が徐々に統合されて行く。2領域以上がエッジの除去により統合された場合には、統合前の各領域をオブジェクトとして抽出する。小さな領域を徐々に統合していきながら最終的に画像全体に至るまでしきい値を段階的に変化させながらオブジェクト抽出を行うことで、図10に示すようなObject Extraction Treeを構成でき、これらの全てをオブジェクトとしてデータベースに格納する。

この方法により、冗長なオブジェクトも抽出されるが、有意なオブジェクトを多く抽出できる。なお、重なりによってオブジェクトの一部が欠けている場合には、その部分が欠落したままのオブジェクトとみなされる。従って図10の一部が欠けたリングはその形状をしたオブジェクトとして扱われる。このため完全な形のリングとは形状的には類似していないと判断される。本抽出方式では意味的な取り扱いをしないため、このような誤差が発生する。

5. 照 合

5.1 照合時の距離計算

照合ではキー画像と蓄積画像の特徴量の間で類似度を距離として計算する。その際の距離計算方法としては表1に示す方法が提案されている。どの計算方法をとるかは特徴量との相性や計算コストを考慮し使い分けが行われる。

5.2 逐次的な照合方法

色情報の照合においてはヒストグラム・インターセクション法²⁾、および類似度行列(QBICではColor Similarity, VisualSEEKではQuadratic Distanceと呼んでいる)^{9),10)}が用いられる。ヒストグラム・インターセクションは特徴量をヒストグラムで表現し、類似度はそれらのヒストグラム間で重なり合った部分の面積を求めるもので、似ている部分の多さに着目するため、L1(市街地距離)、L2(ユークリッド距離)等の距離に比べオブジェクトをキー画像とした場合の背景の影響が少ない。一方、類似度行列は異なる2色間での色の隔たりを個々に定義することで2色間の距離を人間の感じる色合いの類似性に近づけることができる。しかし、この方法はキーと入力画像間でのマトリクス計算が必要でコストがかかる。このため、QBICでは最初にKL法で2~3次元に縮退を行った上でフィルタリングを行い、その後256次元の類似度行列により類似度を計算している。また、R-tree²⁰⁾の各ノードに異なる次元数を持たせ、上位ノードにはより低次元の情報を持つようにしたMulti-resolution R-tree²¹⁾等も提案されている。

5.3 多次元インデックス

照合では、キーとして与えられるオブジェクトと予め蓄積したオブジェクトとの間で特徴量を比較し、類似した順に上位k件を返却することが求められる。L1, L2等の幾何

表1 照合時の距離計算方法

Table 1 Metrics for calculating the difference between Features

	概要	利点	欠点
L1:マンハッタン距離 L2:ユークリッド距離	ベクトル空間内の2点間の幾何学的距離	多次元インデックスの利用が可能	結果が必ずしも直感に合わない
コサイン尺度	ベクトルの内積	テキストの類似検索と共用可能	
ヒストグラム・インターセクション	ヒストグラムの重なり合った面積	背景の影響が少ない照合が可能	多次元インデックス化が困難で、総当たり計算が必要
Quadratic Distance by VisualSEEK	ベクトルの各次元間に関係を2次元で定義。「赤」と「青」は「赤」と「橙」より近い等	色尺度のように隣り合ったヒストグラム間に関係がある場合に有効	
非対称距離基準 by ExSight	人間の直感に合わせて強調された特徴をもつ方が近いと判断。ヒストグラムの高い方をより近いと判定	色相の場合「半分が赤」は「全体が青」よりも、「全体が赤」により近い判定	

学的距離を用いる場合には、この性質を利用し、かつ画像特徴を多次元特徴量という形式で統一的に表現することにより、インデックスを用いた高速な照合が可能となる。上位 k 件検索は、多次元ベクトル空間内でキーに対応するベクトル点の k 近傍を求めることに対応する。最も単純な方法は、キーに対応するベクトルとデータベース中のすべてのベクトルとの距離計算を行い、上位 k 件を求めるものである(全件検索)。しかし、ベクトル間の距離計算にはコストがかかるため、全件検索では比較対象となるデータ数が増えると検索時間が長くなる問題がある。このため、全件検索を回避して k 近傍検索を行うために多次元空間インデックスが提案されてきている。

従来提案されている多次元空間インデックスは主に R-tree²⁰⁾ か k-d tree²⁰⁾ の拡張や改良版である。R-tree は図 11 に示すように空間を直方体の領域に分割し、それらを含むより上位の(つまり体積の大きな)直方体を定義していくことで階層的な木構造インデックスを作る。検索時にはルートから始めてキーが属する直方体領域を木の上位から下位に向けて選択していく。

VAM Split R-tree²⁰⁾ は検索時間性能がもっとも優れたもののひとつであると報告されている。このインデックスでは k-d tree と同様な、空間の分割を基本にした木の構築を行うことで R-tree のノード重複問題を解決して性能向上をはかっている。

5.4 C-tree による多次元インデックス

VAM Split R-tree では中間ノードに対応する MBR (Minimum Bounding Rectangle) の各次元の長さは不均一になり、近傍検索に不利になっているという問題がある(形状不均整問題)。ExSight で提案する C-tree²⁰⁾ はボトムアップで要素をクラスタリングしつつ木構造を構築するアプローチを採る。構築における各レベルでの初期クラスタを得るために VAM Split R-tree の構

築ルーチンを利用するが、先に述べた MBR の不均整形状問題に対処するために、各レベル毎にクラスタの再配置を行う。C-tree の構築法を以下に示す(図 12 参照)。

- (1) VAM Split R-tree を構築し、リーフノードの直上の中間ノード群を、リーフノードをエレメントとする第 1 レベルのクラスタとする。
- (2) 第 1 レベルのクラスタについて、Wishart 法²⁰⁾ に基づいた再配置を行う。すなわち、ある基準を超えて他のエレメントから離れているものの再配置(他のクラスタへの帰属替え)を行う。この際、ある閾値を超えてどのクラスタからも離れていると判定されたエレメントはどのクラスタに組み込まずに、保留リストに加える。
- (3) 再配置した第 1 レベルの各クラスタの代表点と保留リストにあるエレメントを対象として、(1)(2)を繰り返し、クラスタ数が特定の数以下になったら、それをルートとして終了する。

C-tree では、このように各レベルでクラスタの見直しをして中間ノードをできるだけ均整にする(球に近い形にする)。これにより検索時にチェックする木のノード数を VAM Split R-tree より 10%~15%削減でき、検索時間性能を向上させている²⁰⁾。

5.5 数値文字情報との組み合わせ検索

写真における撮影日時や撮影者といったメタ情報は画像内容からは決定できず、画像とは異なる枠組みでの検索が必要である。本章で紹介した全てのシステムではこれらテキスト情報との組み合わせ検索を実現している。ExSight ではコーディネータの配下に画像特徴にもとづく多次元ベクトル用の検索エンジンと、数値文字情報にもとづくテキスト用の検索エンジンを置き、両者の検索結果を結合することで、例えば「著者

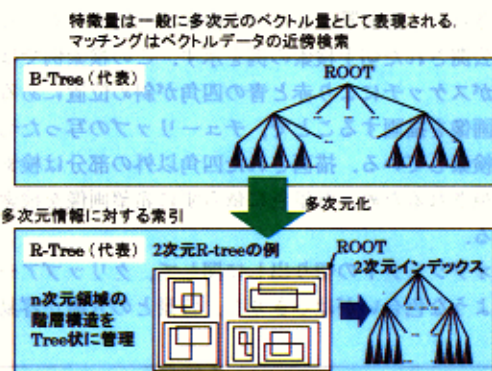


図 11 多次元インデックスによる高速な照合

Fig.11 High Performance Matching by Multi-dimensional Index

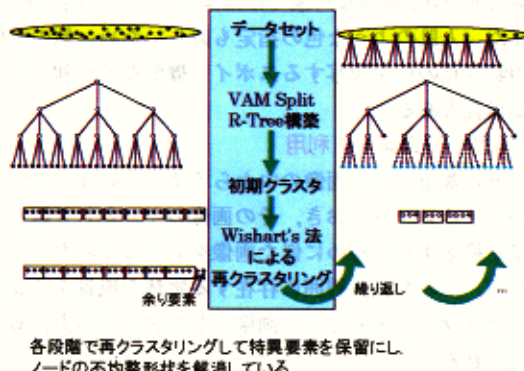


図 12 C-tree Index の構築

Fig.12 Constructing the C-tree



スケッチ	グラフィック・エディタで作図 
既存の画像を利用	① 全体画の場合 → 画像IDを指定 ② オブジェクトの場合 → 画像内のオブジェクトをクリップ 
言葉による指定	① 「サッカーシーン」 → 認識適応型データモデリング (キーワードとシーン記述とのマッピング) ② 「金髪の人物」 → 用語・特徴量変換 → 類似検索

図13 検索インタフェース
Fig.13 Retrieval Interface

がゴッホで、かつ、ひまわりのようなものが描かれた絵」といった検索を実現している。

6. 検索インタフェース

画像検索では検索キーに画像を使う場合が一般的である。その場合には検索キー画像からも同様の方法で特徴量を決定し、画像との間で照合する。しかし、検索キーと画像は必ずしも同種のデータである必要はない。例えば検索キーを言葉で与えることもできる。その場合は検索キーから画像キーへの変換を行い、得られた画像キーをもとに照合がなされる。

キー画像の入力は質問を行う利用者にとって最も重要なインタフェースであり、3つのレベルが提案されている(図13参照)。

6.1 スケッチによる入力

画像エディタを利用してスケッチにより、似た画像を作成する⁹⁾。丸や三角形等、単純な図形入力が容易であるが、複雑な対象では描画しにくい問題がある。また、人肌のような色の指定も難しいため、他の画像の特定点の色を複写するスポイト機能などが利用される場合もある。

6.2 既存画像を利用

既に蓄積された画像の中から希望する画像に近いものを予め見つけておき、その画像のキーとして似た画像を検索する。初めに似た画像を見つけるのが手間となる。このため、周囲に存在する物体や風景をデジタルカメラ等で入力しキー画像として利用する方法も考えられる。

6.3 言葉による入力

例えば「金髪の人物」といった言葉を指定すると、

辞書にもとづいてそれを画像化し、キー画像に用いる。利用者にとって最も容易な方法であるが、満足できる精度・規模の辞書の構築は容易でない。より抽象度を高めた方法として複数の被写体をシーン記述にもとづいて定義し、「サッカーシーン」という言葉から緑のフィールドと白いボールが写った写真を検索する研究²⁶⁾もなされている。なお、本論文の枠からは外れるが、画像に意味的な記述(メタ情報)を付記し、そのテキストの内容をもとに検索する内容検索システムも存在する²⁷⁾。

7. 画像内容検索システムの実例

7.1 Illustra & VIR^{*}

商用化された初めての画像内容検索システムである。Illustraはオブジェクト・リレーショナルDBMSであり、その上に拡張可能な画像検索用データブレードとしてVIRが提供され、そのアプリケーションVIVと共に市販されている。画像全体の色相や色領域の絶対的な位置、模様にもとづく検索が可能で、入力手段としては既存画像やスケッチを用いることができる。照合手法は非公開。Illustra & VIRおよびVIVによる検索例を以下に示す。図14ではハイキングをしている既存の写真をキー画像として全体的な配色が似通った他の写真を検索している。図15は同様の処理をスケッチ入力した画像をキーとして行ったものではほぼ同様の結果を得ている。このようにIllustra & VIRは全体画の類似検索を主目的に開発されており、風景やパターンといった画像に適する。ただし、部分画像の検索機能はないため、例えば図16のように画像の一部に赤くて丸い物体が写った写真等の検索はできない。

7.2 QBIC

QBICは静止画、および動画を対象として研究され、1997年にその機能の一部がDB2の拡張機能として市販された。色配置にもとづく検索が可能である。図17に公開された切手検索の例を示す。この検索例では著者がスケッチにより赤と青の四角が斜め位置にあるキー画像を描画することで、チューリップの写った切手を検索している。描画された四角以外の部分は検索で無視されるため、背景色に依らずに希望画像を検索できる。

オブジェクトの切り出しに関して、クリップアートのような色合いがはっきりし、背景との分離が容易な

* 現時点の正式名称は Informix Dynamic Server Universal Data Option (IDS/UDO) の VIR DataBlade であるが、本論文では、その初期の名称を用いる。



図 14 Illustra&VIR による検索例 (1)
Fig.14 Example of the Illustra/VIR (1)

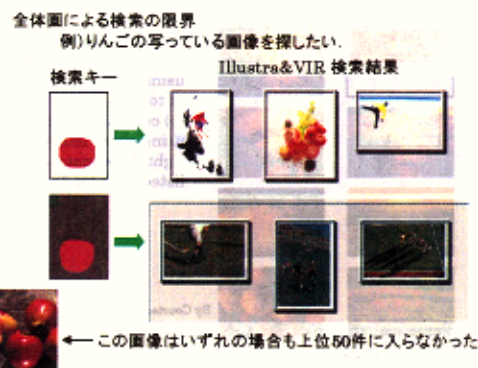


図 16 Illustra&VIR による検索例 (3)
Fig.16 Example of the Illustra/VIR (3)



図 15 Illustra&VIR による検索例 (2)
Fig.15 Example of the Illustra/VIR (2)

画像を対象としており、一般的な写真に対しては利用者の補助により切り出す手段を提供している。オブジェクトの外周を線でなぞる Outline Tool、およびオブジェクトの内と外の2点を指定する Flood-Fill Tool の2つの手段である。その後、色クラスタリングされた領域を階層的に組み合わせる方式により、写真への適用も図られている。QBICではさらに動画像に関して動画内のオブジェクトの動きやズームやパンといったカメラモーションに着目した検索ができる。

7.3 VisualSEEK

画像内の部分領域に着目した検索を行える。特に複数の領域を指定し、それらの相対的な位置関係に着目した検索ができる。図 18 では検索のキーとして画面上部にオレンジ色を背景とした黄色の領域を指定することで、夕焼けの写真を検索している。また、図 19 では上から順に、明るい青、茶、緑の領域を指定し、風景画を検索している。図 20 は各領域間の関係を利用した検索例で、領域間の包含関係、位置関係、それ



図 17 QBIC による検索例 (1995 以前の米国切手)
Fig.17 Example of QBIC -A collection of all U.S. stamps before 1995-

らの組み合わせを指定している。

VisualSEEK では領域の切り出しは自動で行われる。そのために、事前に複数個の代表色を登録し、それぞれに対して、入力画像中をバックプロジェクションにより領域分けしている。この方法はアプリケーション分野に応じて、切り出しに利用する色を予め定義できれば効果的な切り出しの手法である。VisualSEEK の応用例としては Web ロボットと組み合わせ、500 万件以上のホームページを画像により検索する WebSEEK というサービスがある。

7.4 V P

VP は東京大学浜田研究室で開発された部分画検索をターゲットとするシステムである。部分画像はデータベース入力時に自動的に切り出し、それらの特徴量を計算して格納する。切り出しでは先ず入力画像を数百から数千の細かい領域に分割する。次にそれらを領

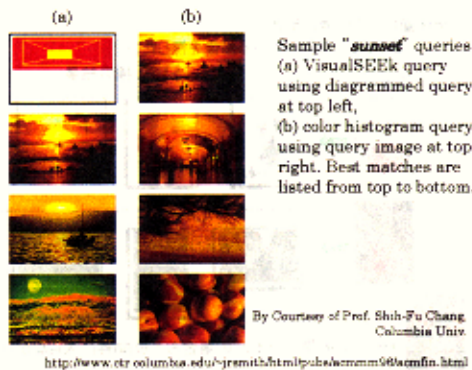


図 18 VisualSEEK による検索の例 (1)
 Fig.18 Example of VisualSEEK (1)

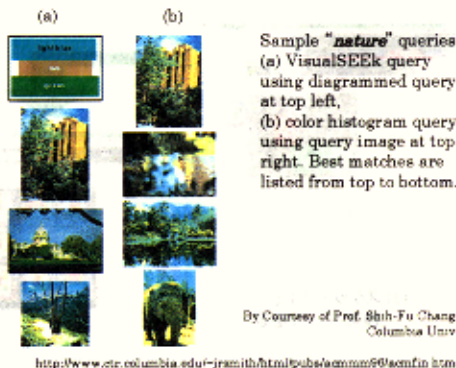


図 19 VisualSEEK による検索の例 (2)
 Fig.19 Example of VisualSEEK (2)

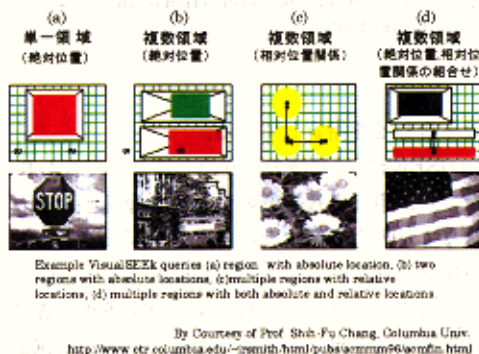


図 20 VisualSEEK による検索の例 (3)
 Fig.20 Example of VisualSEEK (3)

域の大きさや隣接する領域との間のエッジ情報をもとに併合して行き、併合の仮定で一定の基準を満たす領域をオブジェクトとして抽出する。各オブジェクトに関する画像特徴量はシグネチャ表現により、固定長のビット列として表現し、照合はプリミティブ・シグネチャによるフィルタリングと実データを用いた類似度

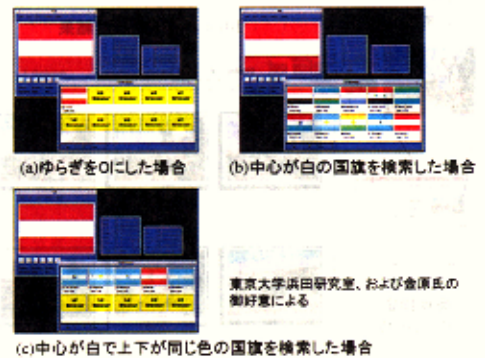


図 21 VP による検索例 (1)
 Fig.21 Example of VP (1)



図 22 VP による検索例 (2)
 Fig.22 Example of VP (2)

計算の2段階で計算している。

図 21 に VP を用いて国旗画像を検索した例を示す。図 22 は写真を対象とした検索例である。キーとして、画像内左側にあるチョコレートに乗ったクッキーを計 4 個指定している (白線)。関係はデフォルトのままとし、色、形状関係の属性は少し厳しくし、他の属性は緩くした結果、閾値 0.8 で 5 件の検索結果を得ている。

7.5 ExSight

ExSight は一般的な写真を対象として、オブジェクトにもとづく高速な内容検索を目的としたシステムである。特徴としてはノイズが含まれることを前提としたしきい値の全域調整によるオブジェクト抽出、改良型多次元インデックス C-tree による高速な照合がある。ノイズを許容した自動オブジェクト切り出しにより、色領域分割²⁸⁾による切り出しに比べて約 2 倍のオブジェクトが抽出される。このため、改良型の多次元インデックスによる高速な照合を実現している。特徴



図 23 ExSight による検索例
Fig.23 Example of ExSight

検索のキー画像	検索結果
 ナイフ	
 空  草	

図 24 ExSight による検索結果の例
Fig.24 Example of ExSight Result

量としては各オブジェクトに関して、色、形状、大きさ、位置を計算している。例えばキー画像として赤いリングを指定すると色、形状、大きさ、元画像内の出現位置をもとにして、DB内画像との照合を行い、類似したオブジェクトの写った写真を検索する。

検索例を図 23 に、検索結果の例を図 24 に示す。上の例では「ナイフ」（銀色で左に傾いた細長いオブジェクト）が写った写真を検索している。また、ExSight では人間や静物といったオブジェクトと区別することなく、空や海といった背景的な領域を抽出できる。図 23 の下の例では大きな背景領域として「青い空」と「緑の芝生」が写った写真を検索している。

8. 評価・標準化

- (1) **検索精度の評価**：実用的なシステムが複数登場するに至り、客観的な検索精度の尺度が必要となってきた。一般には、テキストの類似検索の場合と同様に再現率・適合率グラフが用いられる²⁹⁾。
- (2) **標準化**：MPEG-7 では蓄積型のマルチメディアサービス規格化のために、画像内容を表現するための特徴量の標準化が進められている。色や形状といった 4. で述べたデジタル画像情報から容易に計算可能な特徴量の形式を標準化するのが目的であり、内容の意味的な表現には踏み込まない。さらに特徴量の抽出方法や利用方法等はインプリメンタ・デファインとし、標準化の範囲外としている。現在、2002 年を目前に作業が進められており、画像内容検索分野からも多くのプロポーザルが出されている。

9. 今後の課題と展望

本論文では、画像の内容検索システムを逐次処理型の探索モデルと蓄積型の内容検索モデルに分類し、DB 技術に関連の深い後者のモデルについて、内容解析、照合、検索インタフェースの 3 つの観点から、現状技術を総括し、システムの実例の紹介を行った。

前章で示した通り、画像の内容検索技術は既に実用化の時代に入っつつある。我々の開発する ExSight も、いくつかの先進的なユーザとの共同実験を通じ、現在、実用化段階にある。そこで本章では我々の経験を踏まえた上で、画像の内容検索システムに関する今後の研究課題について、我々の観点から整理と考察を試み、本論文のまとめとする。

9.1 内容解析

内容解析については、画像処理技術の進歩に負うところが大きいですが、ここでは主にそれ以外の観点から考察する。

- (1) **新たな特徴量**：メディア系研究の進歩に追随し、常に新たな特徴量の導入も検討の必要がある。例えば、スケッチに比べ、より入力しやすい特徴量として、ExSight では、直線での検索³⁰⁾と統合を図っている。
- (2) **照合性能までを考慮した特徴量**：画像処理分野等のメディア系研究における提案の多くは精度向上を主目的としている。しかし、それらの技術は、その実用上計算オーダが高すぎる等、データベースや検索システムという観点から見た場合には、照合技術との相性が悪く、現状では適切でない手法も多々存在する。よって、大量の情報への適用を前提とした観点から、特徴量

の選別・改良を行っていく必要がある。

- (3) **複数の特徴量の扱い**: オブジェクトの色, 形状, 模様, 位置といった複数の特徴量を一度に利用する検索, さらには, 画像, 音楽, テキストといった複数のメディアを一度に利用する検索では, それらの特徴量の重みをどのように正規化し, 比較可能にすべきか, また, どのような順序で利用すべきかという問題がある. いずれも, 現在では利用者が人手で調整しており手間がかかっている.
- (4) **データベース技術を利用したアプローチ**: 例えば ExSight の場合には, オブジェクトの完全な認識を前提とする認識的アプローチではなく, 大量のノイズの存在を前提とした独自のアプローチを採用することで画像内オブジェクト検索を実現した. これに限らず, 規模を利用した DB 的アプローチが有効に機能する場面を検討していく必要がある.
- (5) **対象世界モデルを使う検索技術との統合**: 本論文では対象世界モデルを前提としない画像内容検索技術を紹介してきたが, 一方では, そのモデルを積極的に活用した画像検索技術等^{31),32)}も存在する. 対象世界のモデル化を前提とする場合, そのモデルの入力や定義の煩雑さ等のデメリットがある反面, モデルに一致した状況では極めて高精度でかつ高度な検索が可能になるという特徴を持つ. 画像検索の適用先によっては, 両手法を統合したアプローチの検討が必要であろう.

9.2 照 合

- (1) **特徴量特性に合わせた照合法の開発**: 前項でも述べたが, 特徴量の種別により, 索引の構成方法に影響がでる. 例えば, 画像でいえば色, 形状, 位置のそれぞれで索引への要求特性が異なり, 異なる索引が利用される. また, 画像, テキスト, 音楽などのメディアの違いによっても, 適切な索引は異なる. よって, 種々の特徴量に合わせ, その照合法の開発と効率的な索引構成について検討していかなければならない.
- (2) **複数メディアの扱いとアクセスパス**: テキスト検索との組合せ検索が行われることについては 5. で触れた. また, 他にも音楽・音声, モーションなどとの組合せ検索も考えられる. このように複数の特徴量やメディアが存在し, それらが個々に索引を持つ場合, それらの索引をどういう順序で, どのように使うかというアクセス

パスの最適化問題が発生する. また, システムの一部として既存の DBMS を利用する場合等も実際には多く, マルチ DB 的な視点からの問合せ最適化の検討も求められている.

- (3) **大規模化**: テキストの類似検索が WWW 検索エンジン等によってかなり大規模な情報に適用されているのに比べ, 画像検索の場合には, まだ, その対象件数は少ない. WWW にも充分適用できるだけ高速化手法の開発が必要である. なお, この場合, 特徴量抽出の高速化技術も実用上重要になってくる.

9.3 検索インタフェース

- (1) **検索キーの入力 (メディア変換)**: 画像検索の実用化に伴って, 最も重要な課題となっているのが, この検索キー入力の問題である. 6. で幾つかの手法を紹介したが, いずれも満足のいくものまでには至っていない. 多くのユーザはスケッチでの入力を望まず, テキストの入力を求めることが多い. そこで, シソーラスにも相当する大規模な用語-特徴量変換辞書の整備が求められている. また, 音声を入力キーとした画像内オブジェクトの検索などの研究も求められている.
- (2) **クラスタリング技術との統合**: 画像の特徴量を利用したクラスタリングの研究^{33)~35)}も行われている. クラスタ基準の選択だけで対象画像を絞り込める手法と, それに比べ, 入力は面倒であるものの検索キー指定の自由度の高い画像の内容検索システムとの融合が求められている.

9.4 評価・標準化

- (1) **検索精度の評価**: 再現率・適合率グラフ³⁶⁾で異なるシステムの比較評価を行うためには, 対象とする画像集合, 検索キー, および正解集合を共通にもつ必要がある. テキスト検索の世界では既に共通評価セットの確立³⁶⁾が始まっているが, 画像に関しては今後の整備が待たれる. また, 画像検索にとどまらず, マルチメディア情報検索全般に関し, そのメディア単独の場合のベンチマーク, および, 複合した検索の場合のベンチマークの構築が望まれている. これらに関し, 情報処理学会データベースシステム研究会の第 4 分科会がマルチメディア情報検索ベンチマーキング WG を設置し, 検討を開始しており, 今後の活動が期待される³⁷⁾.
- (2) **標準化**: MPEG-7 については, MPEG-4 との関係があり流動的な側面を残している. 一方, 画像内容検索技術の実システムへの適用に伴い,

類似検索に関する問合せ言語の標準化も求められている。SQL3 の範囲内で統一するのか、それを、類似検索や対話的絞り込み検索にまで拡張して統一するのかについての議論と、早期の標準化が必要である。

謝辞 本論文を書くにあたり、システム実行画面の掲載を快く承諾して戴いたQBIC開発者のIBM社 Almaden研究センター Wayne Niblack氏、VisualSEEk開発者のコロンビア大学 Shih-Fu Chang教授、VP開発者の東京大学浜田研究室および金原 史和氏（現 NEC）に深く感謝します。また、画像処理技術全般に関し、多数の有益なコメントを戴いたNTTヒューマンインタフェース研究所の倉掛 正治氏およびNTT基礎研究所の村瀬 洋氏に感謝します。

参 考 文 献

- 1) M. De Marsico *et al*: Indexing pictorial documents by their content: a survey of current techniques, *Image and Vision Computing*, Vol.15, pp.119-141 (1997).
- 2) M. J. Swain, D. H. Ballard: Color Indexing, *International Journal of Computer Vision*, Vol.7, No.1, pp.11-32 (1991).
- 3) 曾根光男, 寺田 聡, 坂内正夫: 特徴量空間とピラミッド構造とを用いた画像の高速重ね合わせ法, 電子情報通信学会論文誌 D, Vol.J71-D, No.1, pp.102-109 (1988).
- 4) 村瀬 洋, V. V. Vinod: 局所的情報を用いた高速物体検索 - アクティブ探索法 -, 電子情報通信学会技術報告, PRMU 97-17, pp.127-134 (1997).
- 5) 加藤俊一, 下垣弘行, 藤村晃明: 画像対話型商標・意匠データベース TRADEMARK, 電子情報通信学会論文誌 D-II, Vol.J72-D-II, No.4, pp.535-544 (1989).
- 6) A. Gupta: Visual Information Retrieval Technology A Virage Perspective, <http://www.virage.com>, pp.1-11 (1995).
- 7) A. Gupta, R. Jain: Visual Information Retrieval, *Communications of the ACM*, Vol.40, No.5, pp.71-79 (1997).
- 8) M. Flickner *et al*: Query by Image and Video Content: The QBIC System, *IEEE Computer*, Vol. 28, No.9, pp.23-32 (1995).
- 9) J. Ashley *et al*: Automatic and Semi-automatic Methods for Image Annotation and Retrieval in QBIC, *Proc. Storage and Retrieval for Image and Video Databases III*, Vol. 2420, SPIE (1995).
- 10) J. R. Smith, S. F. Chang: VisualSEEk: A Fully Automated Content-Based Image Query System, *Proc. ACM International Conference on Multimedia*, pp.87-93 (1996).
- 11) 金原史和, 佐藤真一, 浜田 喬: プリミティブ分解による多様な検索条件を扱うカラー画像検索, 情報処理学会論文誌, Vol.37, No.11, pp.1989-2000 (1996).
- 12) 串間和彦, 赤間浩樹, 紺谷精一, 木本晴夫, 山室雅司: オブジェクトにもとづく高速画像検索システム: ExSight, 情報処理学会論文誌, Vol.40, No.2 (1999).
- 13) 山室雅司, 串間和彦, 木本晴夫: 自動部分構造化にもとづくインデクシングによる大容量画像検索システム, 情報処理学会 第 54 回全国大会, 3Q-01 (1997).
- 14) 赤間浩樹, 三井一能, 紺谷精一, 串間和彦: 画像内オブジェクトの自動抽出を使った画像検索システム ExSight, 電子情報通信学会, データ工学ワークショップ論文集 (DEWS'97), pp.107-112 (1997).
- 15) 赤間浩樹, 紺谷精一, 三井一能, 串間和彦: 画像内オブジェクトの自動抽出を使った画像検索システム ExSight - 写真(PhotoDisc)への適用 -, 情報処理学会研究報告, 97-DBS-113, Vol.97, No.64, pp.161-166 (1997).
- 16) 高木幹雄, 下田陽久: 画像解析ハンドブック, 東京大学出版会 (1991).
- 17) 中田雄一郎, 小早川倫広, 星 守, 大森 匡: ウェーブレット変換を用いたテクスチャ解析と類似画像検索への応用, 電子情報通信学会技術報告, PRMU69-166, pp.45-52(1997).
- 18) A. Laine, J. Fan: Texture Classification by Wavelet Packet Signatures, *IEEE trans. on Pattern Analysis and Machine Intelligence*, Vol.15, No.11, pp.1186-1191 (1993).
- 19) 田邊勝義, 大谷 淳: 形状類似画像検索における類似尺度の検討, 電子情報通信学会技術報告, PRU88-68, pp.65-72 (1988).
- 20) A. Guttman: R-tree: A Dynamic Index Structure for Spatial Searching, *Proc. of the ACM SIGMOD*, pp.44-57, Boston (1984).
- 21) K. Curtis, N. Taniguchi *et al*: A Comprehensive Image Similarity Retrieval System that utilizes Multiple Feature Vectors in High Dimensional Space, 情報処理学会研究報告, 97-DBS-113-28, pp.167-172 (1997).
- 22) J. L. Bentley: Multidimensional Binary Search Trees used for Associative Searching, *Communications of the ACM*, Vol.18, No.9, pp.509-517 (1990).
- 23) D. A. White and R. Jain: Similarity Indexing: Algorithms and Performance, *Proc. SPIE IV*, Vol.2670, pp.62-75, San Jose (1994).
- 24) キャサリン カーティス, 中川純一, 谷口展郎, 山室雅司: Similarity Indexing in High Dimensional Image Space, 情報処理学会研究報告, 97-MPS-82-18, pp.99-104 (1997).
- 25) M. R. Anderberg: Cluster Analysis for Applications, Academic Press (1973).
- 26) J. Yamane, M. Sakauchi: A Construction of a New Image Database System Which Realizes Fully Automated Image Keyword Extraction, *IEICE Transactions on Information and Systems*, E76-D,10, pp.1216-1233 (1993).
- 27) A. Sheth and W. Klas (ed): Multimedia Data Management: Using Metadata to Integrate and Apply Digital Media, McGraw Hill, (1998).
- 28) 紺谷精一, 串間和彦: クラス数自動判定クラスタリングによる画像の領域分割, 情報処理学会 第 54 回全国大会, 3Q-02 (1997).
- 29) D. K. Harman: Overview of the Third Text Retrieval Conference (TREC-3), Gaithersburg, MD20899-0001, *National Institute of Standard and Technology, Special Publication* 500-255 (1999).

(1995).

- 30) 紺谷精一, 赤間浩樹, 山室雅司: 画像からの直線検出と直線をキーとした画像の検索, 情報処理学会研究報告, 98-DBS-116-1, pp.1-8 (1998).
- 31) 有澤 博: リアルワールドデータモデリングについての考察, 電子情報通信学会技術報告, DE96-4, pp.19-24 (1996).
- 32) 富井尚志, 有澤 博: マルチメディアデータベースにおける映像モデリングと操作言語, 電子情報通信学会論文誌 D-II, Vol.J79-D-II, No.4, pp.520-530 (1996).
- 33) 波多野賢治, 亀井俊之, 田中克己: 映像自己組織化構造に基づく内容記述と類似シーン検索, 情報処理学会研究報告 97-DBS-113-29, pp.173-178 (1997).
- 34) 仲川亜希, 片山幸治, 金山智一, 小西 修, 菊池時夫: 画像データベースのためのデータマイニング法の拡張, 情報処理学会研究報告, 98-DBS-115-10, pp.71-78 (1998).
- 35) W. Y. Ma, B. S. Manjunath: A Texture thesaurus for browsing large aerial photographs, *ECE Technical Report #96-10*, University of California, Santa Barbara, pp.1-32 (1996).
- 36) 木谷 強, 小川泰嗣, 石川徹也, 木本晴夫 他: 日本語情報検索システム評価用テストコレクション BMIR-J2, 情報処理学会研究報告, 98-DBS-114-3, pp.15-22 (1998).
- 37) 清木 康, 酒井哲也, 村瀬 洋, 谷口展郎, 有川正俊: マルチメディア情報検索ベンチマークの未来, 情報処理学会, アドバンス・データベース・シンポジウム'98 論文集 (ADBS'98), pp.103-108 (1998).

(平成 10 年 9 月 20 日受付)

(平成 10 年 12 月 27 日採録)

(担当編集委員 加藤 俊一)

串間 和彦 (正会員)

1957 年生, 1980 年京都大学工学部電子工学科卒業。同年, 日本電信電話公社(現 NTT)入社。知識処理用プログラミング環境の研究, 大規模クライアント・サーバシステムの実用化等を経て, 現在はマルチメディアデータベースの研究開発に従事。

赤間 浩樹 (正会員)

1988 年東海大学理学部情報数理学科卒。1990 年同大学院数学専攻修士課程修了。同年, 日本電信電話株式会社入社。以来, ネットワーク制御向け DBMS, DB 設計評価・チューニング, ニュース・オン・デマンド等の研究を経て, 現在, マルチメディア情報検索の研究に従事。人工知能学会, 日本ソフトウェア科学会, ACM 各会員。

紺谷 精一

1988 年北海道大学工学部電子工学科卒業。1990 年同大学院工学研究科電子工学専攻修士課程修了。同年日本電信電話株式会社入社。マルチメディア情報検索の研究に従事。電子情報通信学会, IEEE-CS, ACM 各会員。

山室 雅司

1985 年早稲田大学理工学部数学科卒。1987 年同大学院数学専攻修士課程修了。同年, 日本電信電話株式会社入社。1990 年コロンビア大学大学院電気工学専攻修士課程修了。以来, ネットワークオペレーション情報モデル化・ビジュアル化, データベース設計の研究に従事。現在, マルチメディア情報検索の研究に従事。1994 年電子情報通信学会学術奨励賞。電子情報通信学会, 日本ソフトウェア科学会, IEEE-CS 各会員。