

12 画像検索

12.1 画像検索

コンピュータの性能向上やメディア技術の進展に伴って、地図や、リモートセンシング、医療、バイオメトリックス、電子出版、デザイン、ゲーム、電子商取引などさまざまな分野で画像が利用されるようになってきました。このような画像データを多量に蓄積し利用するためのデータベースが画像データベースになります。画像データベース中のデータは、画像データとそれを効率的に管理するためのメタデータから構成されます。画像データは、画像ファイルを直接データベースに格納して管理することができます。また、画像ファイルの保存先のアドレスを用いて管理する方法もあります。例えば、画像ファイルの保存先としてWebサーバを利用することができます。この場合、画像データベースでは、その画像ファイルのURLを用いて画像の管理を行います。このような画像を効率的に管理するために、メタデータが利用されます。メタデータは、画像データの効率的な管理や検索などに利用される付加的な情報です。例えば、写真の場合、タイトルや撮影者、撮影場所、撮影日時、画像の内容を表すキーワードなどがメタデータとして利用されます。また、画像サイズ、画像形式、色情報といった画像に特有の情報や、著作権情報などもメタデータとして利用されます。メタデータを利用することで、膨大な情報の中から、利用者が必要な情報を検索することができるようになります。通常は、利用者が思いついた自由なキーワードを入力し、それに合致する画像を検索結果として提示します。

一方、画像データには、文字データに比べて、はるかに多くの情報が含まれており、必ずしもすべてを言葉として表現できるわけではありません。このため、画像の内容を解析することで検索を実現しようという試みもあります。以下、このような画像検索の例についてお話しします。

画像検索は、利用者の検索意図に基づき、大きく2つに分類することができます。1つは、検索したい画像のイメージを、明確に表現できる場合です。利用者が「こんな画像」という検索要求を、明示的に与えることができる場合とも言えます。これに対し、検索要求を、明示的に与えることができない場合もあります。これは、利用者が検索内容のイメージを表現することが難しい場合や、画像を連想的に検索する場合などに対応します。これらの検索は、必要な機能や役割が異なっており、それぞれ、さまざまな手法が提案されています。

まず、検索したい画像のイメージを、明確に表現できる場合ですが、このような検索の場合、利用者の検索要求がはっきりしていることから、検索システム側では、検索する画像内容を、比較的容易に推定することができます。このため、このような検索では、どのように有効な検索キーを選び出し、利用するかが重要になります。ここでは、検索キーの違いにより、(1)画像の特徴量に基づく検索、(2)画像内の物体やその配置に基づく検索、(3)画像と他の関連情報を利用した検索の3つに分類して見てみましょう。

画像の特徴量に基づく検索とは、画像から、色、大きさ、形状、模様の繰り返しのパターンであるテクスチャ、などの特徴量を抽出し、それらの類似度により、検索を実現するものです。色に関する特徴量としては、例えば、明るさのヒストグラムなどを用いることができます。また、色相のヒストグラムを作成して、画像特徴として利用することもできます。画像の部分ごとの色情報も、特徴量として利用できます。

大きさや形状に関する特徴量としては、例えば、面積、周囲長、縦横比などがあります。また、外接する長方形のうち、どれぐらいが埋められているかを表す充填率も特徴量として利用できます。円に近いか細いかを表す円形度といった特徴量を利用することもできます。円形度は、 $(4 \times \text{円周率} \times \text{面積}) / (\text{周囲長} \times \text{周囲長})$ で割ったものとして定義されます。円の場合は、円形度は、円の半径によらず、最大値である1となります。面積が0である直線等は、円形度は最小値である0になります。模様の繰り返しのパターンであるテクスチャが、どのようなパターンであるかも、画像の特徴量として利用できます。例えば、画像中に小さな格子模様が何度も繰り返し現れるのであれば、空間的な周波数が高い画像と言えます。逆に、格子模様が大きく、繰り返し回数が少なければ、空間周波数は低くなります。また、どの方向に繰り返しパターンが現れるかも、特徴量として利用できます。このようなテクスチャ情報の解析には、フーリエ変換や、高次局所自己相関特徴などが利用されます。

ここで、画像の特徴量に基づく検索の例として、色相のヒストグラムによる画像検索について、もう少し詳しく見てみましょう。まず、画像の色を表現する赤(R)、緑(G)、青(B)の値を色相、明度、彩度に変換します。図12-1は、色相について、横軸の分割数が12のヒストグラムを作成した例を示しています。横軸は、赤、黄、緑、シアン、青、マゼンタといった色に対応し、縦軸がそれぞれの色に対応する度数を表しています。

色相のヒストグラムの類似度の尺度としては、しばしば、カイ二乗検定に類似した式が使われます。この値を計算し、この値が小さいほど、画像は類似しているとすることができます。

図12-2は、CG画像を対象にして検索を行った例を示しています。この例では、左上の画像と類似する順に、検索結果の画像が並べて表示されています。上位に同じCGの画像が表示されており、おおむね正しく検索できていることが分かります。

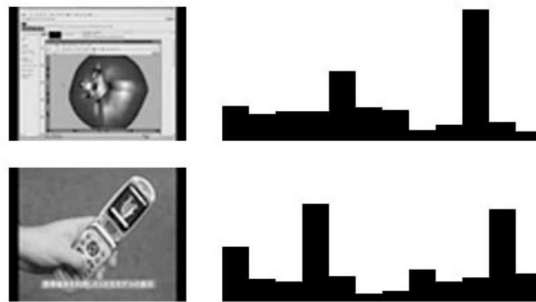


図12-1 色相のヒストグラム

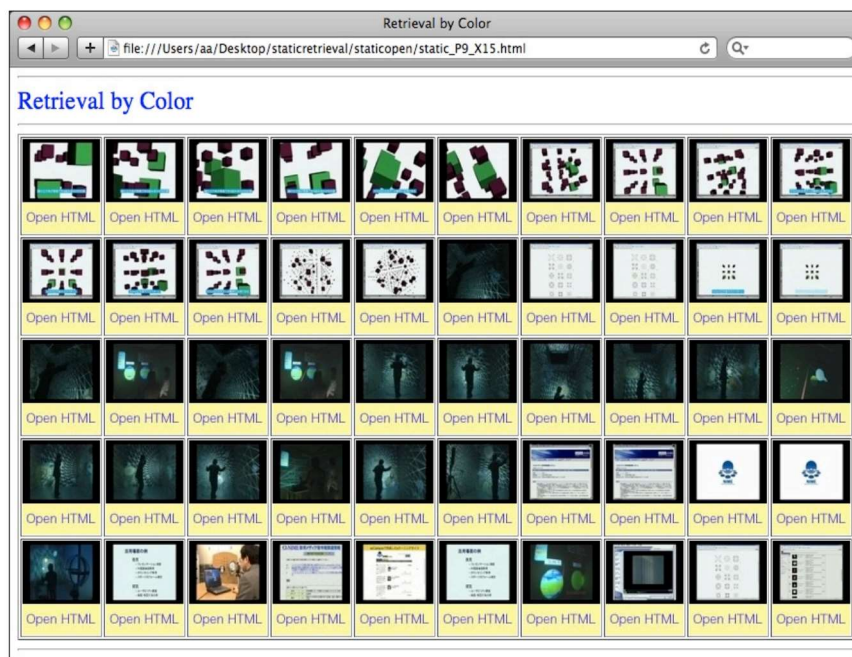


図12-2 色相のヒストグラムによる検索

画像内の物体や配置に基づく検索とは、画像内の物体や、それらの位置関係等の指定に基づく内容検索を実現するものです。例えば、人工衛星で撮影した画像を対象に、道路や川、空港等の主要な目標物を自動抽出し、それらの位置関係を指定することにより画像検索を行った例などがあります。このような画像内の物体やその配置に基づく検索の問題点として、自動認識により対象物体の抽出を行うことが、必ずしも容易ではないことが挙げられます。このため、自動認識が完全にうまくいかない場合でも、不完全な認識レベルをも利用して画像検索を行うシステムが提案されています。例えば、サッカーの画像を対象とした場合、認識の最初の段階では、画像中から緑色、白、黒といった領域を抽出します。これは、画像の色情報を見て、その色の領域がどのぐらいの数、連続しているかを調べることで抽出することができます。そして、緑色の領域のうち、大きな領域があれば、それに対して「大きな緑色」というキーワードを付与します。更に直線で囲まれていれば「フィールド」というようにキーワードを付与していきます。また、白い領域のうち、白い縦長の領域があれば、それに対して「ゴールポスト」のキーワードを付与します。また、同じぐらいの大きさの白い領域と黒い領域がいくつか集まっていれば、その領域に対して「ボール」のキーワードを付与します。そして「フィールド」や「ゴールポスト」や「ボール」が同時に存在すれば、「サッカー」というキーワードを付与します。検索の際は、不十分かもしれない認識水準を前提にして、いわば「それなりに」検索に利用します。例えば、「サッカー」というキーワードで検索を行ったものの、「サッカー」というキーワードが抽出されていなかった場合、システムは「サッカー」を構成する、より低位の概念である「ゴールポスト」や「フィールド」といった状態をも使って検索を実行します。これにより、完全なキーワードが無い場合は、精度は落ちますが、不完全なキーワードも用いてそれなりに検索を行うわけです。

画像と他の関連情報を利用した検索とは、一般には困難である画像の認識、キーワード抽出を補うものとして、他の関連情報を利用するものです。これらの検索の場合、画像に関連する主に書誌情報などのテキスト情報を利用します。例えば、画像を管理するリレーショナルデータベースからのキーワードと、画像特徴（主に色情報）とを用いて画像検索を実現するシステムがあります。色情報としては、色のヒストグラムを利用しています。多量の画像を扱う場合、質問のたびに画像処理を行うことは、現実的ではありません。このため、このシステムでは、データの投入時に色数を20色に限定し、その色ヒストグラムを属性として利用しています。ユーザは、特別な質問用の言語を用いて「1994年1月1日以降に撮影されて、場所がサンフランシスコ湾の近辺、色がオレンジ色を含む画像」といった検索条件を指定して検索を行います。色に関しては、「SomeOrange」とすることで、オレンジ色が画像の一部に含まれること、「MostlyOrange」とすることにより、オレンジ色が画像全体に含まれることを表すことができます。また、概念を、リレーショナルデータベース中の記述や画像特徴を用いてあらかじめ登録しておくことにより、概念による検索も実現できるようになっています。例えば、「日の入り」という概念を、(a)「日の入り」というキーワードが付与されている、あるいは、(b) 赤、黄、紫の色領域が大きな画像というように、あらかじめ登録しておくことにより、概念による検索も実現できるようになっています。

一方、検索要求を、明示的に与えることができない場合もあります。これは、利用者が検索内容のイメージを表現することが難しい場合や、あるいは、画像を連想的に検索する場合などに対応します。例えば、雑誌をばらばらとめくるように、検索しながら検索イメージを明確化していきます。このような検索の場合、検索結果の関係性の表示方法も重要となります。

ここでは、色が類似する画像を3次元空間に配置することを考えてみましょう。例えば、色特徴として、色相のヒストグラムを用いる場合、画像中のRGBを色相、明度、彩度に変換し、ヒストグラムを作成します。もし、横軸の分割数が12のヒストグラムを作成した場合、それぞれの度数を並べた12次元のベクトルをそれぞれの画像の特徴と考えます。そして、主成分分析を用いて、12次元の色相のベクトルを、3次元ベクトルへと変換します。これにより、画像の色特徴に応じて、3次元空間中に画像を配置することができるようになります。

画像を3次元ベクトルで表される位置に表示するためには、例えば、VRMLやOpenGL、WebGLといった3次元空間を記述するための技術を利用することができます。図12-3は、VRMLを利用して3次元空間の表示を行った例です。類似した画像が近くに集まって表示されていることが分かります。ビューアの機能により、ユーザは、この3次元空間内を水平移動、回転、拡大、縮小などをしながら、自分の必要な画像を探し出すことができます。

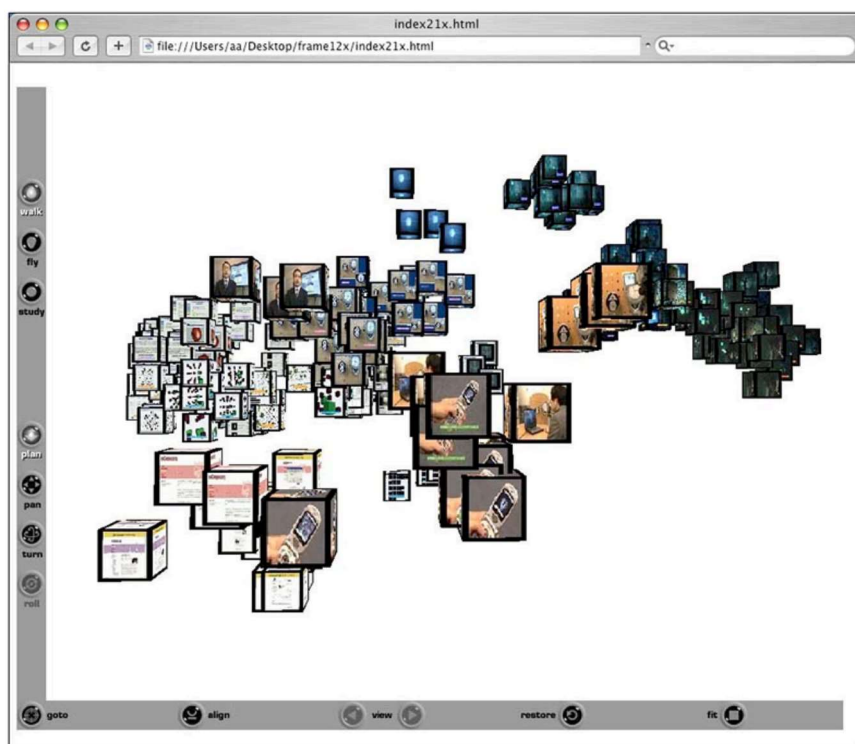


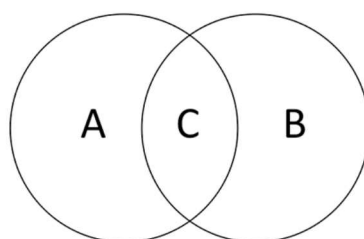
図12-3 画像の3次元配置

12.2 検索性能の評価

ここで検索性能の評価について考えてみましょう。検索性能の指標として代表的なものとしては、適合率と再現率があります。図12-4は、適合率と再現率の関係を示したものです。図のAで表される円は、検索システムによって検索されたデータとします。また、Bで表される円は、データベース中の検索されるべき正解データとします。Cは、AとBの共通部分で、検索されたデータのうちの正解データになります。

適合率 (precision rate) は、(検索されたうちの正解データ数/検索されたデータ数) で求められます。図で言えば、Aに含まれるデータのうちの、Cに含まれる割合を表します。これは、検索結果がどれだけ正確かを表す指標になります。適合率は、単に精度と呼ばれる場合もあります。

再現率 (recall rate) は、(検索されたうちの正解データ数/検索されるべき正解データ数) で求められます。図で言えば、Bに含まれるデータのうちの、Cに含まれる割合を表します。これは、検索されるべきデータのうち、どれだけ検索できたかを表す指標になります。



- A:検索されたデータ
- B:検索されるべき正解データ
- C:検索されたうちの正解データ

図12-4 検索性能の評価

検索結果を提示する場合、例えば、より確からしい、少数の検索結果のみを提示することができます。この場合、一般に、適合率は上がるものの、再現率は下がってしまいます。一方、より信頼度の低い結果も含めて多くの検索結果を示す場合、一般に、再現率は上がるものの、適合率は下がってしまいます。このように、適合率と再現率はトレードオフの関係があるため、検索性能の評価の際は、適合率と再現率の一方だけを評価するわけにはいきません。このような適合率と再現率を同時に評価するための指標として、F値があります。F値は、適合率と再現率の調和平均で定義されます。F値が大きいほど、適合率と再現率の値が同時に大きいと言え、検索性能が高いと言えます。