Technical
Survey

技術解説

抽象的なイメージによる画像検索

正会員 小林 亜樹[†]

画像検索は、検索問合せ（クエリー）が確定すれば、パターン認識など同様の手法で実現することができる。しかし、言葉で説明することが難しい画像を、適切なクエリーへと変換するインタフェースは発展途上であり、画像検索特有の課題であると言える。本稿では、画像検索の一般的なモデルの導入と、各要素の説明を行うとともに、筆者らの提案している抽象的なイメージによる問合せを実現するインタフェースについても解説する。

キーワード：特徴量，適合性フィードバック，距離測定，選択式インタフェース，MPEG-7

1. 画像検索

1.1 画像検索とは

「^{はじめに}太初に言葉有り」で始まることで有名なヨハネ福音書だが、人類の歴史を振り返るに、少なくとも文字よりも先に絵を描き記していただろう。言葉は文字の発明によって、視覚と聴覚を通じて記録と利用がなされるようになり、情報の符号化によって情報の圧縮がおのずと行われたため、現代に至るまで極めて広範に利用されている。

しかし、同じ視覚情報である絵、すなわち画像についてはどうだろうか。記録は、絵画技術の発達を経て写真技術の発明により飛躍的な発展を遂げた。銀塩写真の時代から日本人は写真好きだと言われてきたが、デジタル化と符号化の技術進歩と相まって、日々手軽に、大量に、蓄積されている。しかし、撮ったままで見返すことも少ない写真の現状を見れば、少なくとも個人レベルでの利用は限定的にしか行われていないと言えるだろう。

情報の利用には、検索技術が不可欠であり、画像の利用が文書のそれに比べて少ないのは、とりもなおさず画像検索技術が相対的に不十分であることを示している。

1.1.1 検索の要求

一般に情報検索は利用者が何らかの情報を欲するところからはじまる。これを情報要求と呼ぶ¹⁾が、その状態を4階層に分類した提案がなされている²⁾。

Q1：要求自体を明確に意識できていない「潜在的な要求」(visceral need)

Q2：要求は意識できているが言語的説明のできない「意識的な要求」(conscious need)

Q3：要求を言語表現可能な「形式化要求」(formalized need)

Q4：要求を満たすための手段なども具体化された「妥協済み要求」(compromised need)

これは、文献検索を念頭に置いた分類であるが、画像検索において同様の分類を考えると、Q4は、検索システムが画像データベース中の画像を特定するのに用いている何らかの記述（メタデータ）によって、自己の情報要求を記述できる状態に相当する。Q3は、要求を明確に記述できるが、システムの記述とはかならずしも一致しない状態、または、要求を満たす画像自体を用意できている状態であると言える。Q2は、頭の中には欲しい画像の明確なイメージがあるが、それを外部にうまく説明できない状態、Q1は、そのイメージすら漠然とした状態を示すと言える。

1.1.2 あいまいな要求

富士山の画像を欲しい、という状況を考えてみる。さて、どんな富士山の画像を思い描いたのだろうか？ 冬の雪化粧か夏山の緑か。北斎の赤富士を、あるいはまた広重のそれかもしれない。太宰治の富嶽百景に描かれたような月見草との組合せを思い浮かべる読者もいるだろう。さらに、その構図はどうだろうか？ 富士山と月見草、その配置は？ ほかには画像を構成する要素はないのだろうか？ 画面の明るさは？ 昼間だろうか、夜だろうか、それとも朝焼けや夕焼けだろうか。北側からと南側とでも富士山は違って見える。このように、名前のついたような名画であればその名前で特定できるが、一般の写真の中から明確に1枚を指定することの困難なことがわかる。

このように、画像検索においては、文書検索系で実用化されているQ4、Q3といった状況を準備することは難しい。特にQ4のような画像を記述する（一種の）言語は、テキスト検索におけるキーワードのように、人が自然に理解できるものではない。なお、Q3の特殊な場合として、検索したい画像（とよく似た画像）が手元にある場合は、類似画像検索と呼ばれる。著作権管理などの応用も考えられ、目的とする検索画像が明確なため、研究事例も多い。

[†]東京工業大学 大学院 理工学研究科

"Image Retrieval by Abstract Images" by Aki Kobayashi (Graduate school of science and engineering, Tokyo Institute of Technology, Tokyo)

しかし、情報要求という観点から、より本質的なQ2やQ1の状況での検索を考えたい。Q2の状況とは、頭の中に完成イメージのある作品がある場合に部品となる画像を検索する状況などが当てはまり、Q1は、そのイメージ自体を明確化していくアイデアプロセッサ的な要素も含むと考えられる。具体的には、ポスター制作や放送番組の編集、身近なところでは年賀状の作成などの場面が考えられる。

画像の検索とは、蓄積された画像を利用するための手段であり、テキスト検索と比べて曖昧な状況で目的を達成しなければならない。

1.2 画像検索モデル

1.2.1 画像検索の要素

画像検索システムを構成するためにはどのような機能が必要だろうか。

まず、利用者の情報要求を読み取り、システムが理解できる形で記述する必要がある。そして、この記述に適合する画像を蓄積してあるデータベース中から探し出す。検索の言葉で言うと、システムに渡す情報要求のことはクエリー（問合せ）であり、このクエリーに適合する画像であるかどうかを判断する作業のことはマッチングと呼ばれる。

典型的なテキスト検索であれば、クエリーとしてキーワードの集合を用い、そのキーワード集合を含む文書であるかでマッチングを行う。マッチングを行う情報のことをメタデータと呼び、実用上の工夫として、メタデータによるマッチングを高速に行うため、インデックス（索引）と呼ばれるデータ構造が用いられる。

メタデータはマッチングのための情報であり、問合せ毎に異なる情報要求を反映できる情報である必要がある。しかし、文書が専らそこに書かれた言葉によって識別され、その言葉の符号化が完全に行われているのに対して、画像の場合は画像を利用する目的自体が多岐にわたるため、目的に応じてメタデータの種類から検討し、マッチング処理を考えた符号化まで設計しなければならない。

以上をまとめると図1のような課題があることがわかる。

1.2.2 検索モデル

このように画像検索では、検討しなければならない要素が多く、テキスト検索における検索モデル（マッチングモデル）、すなわち、クエリーに対してのマッチングをどのように行うのか、だけではなく、より広義の、どのようにして利用者に適切なクエリーを構成させるのか、まで含まなければならない。したがって、画像検索モデルは、

① 入力インタフェース、② クエリーの形式、③ マッチングモデル、④ メタデータの種類によって分類される。クエリーの入力方法は、単純に入力を行うのか、クエリーの修正を加えながら行うのか、に分けられ、クエリー形式は、テキスト、画像、特徴量に大別される。マッチング方式は、ベクトル空間モデルやクラスモデル、ヒストグラムを用いるものなど多数存在する。また、具体的なメタデータの

図1 画像検索の構造と課題

種類はマッチングモデルを上回る種類が用いられてきた。

1.3 検索と認識

さて、条件に適合する画像を識別するという意味では、パターン認識や画像理解という技術分野と画像検索とは同じに見える。画像内容を完全に理解したなら、識別は完全に行われるはずだから、画像検索も完全に行われるのだろうか。

パターン認識と画像検索の最も大きな違いは、あらかじめ定義されたクラスに分類すれば良いパターン認識に対して、画像検索では、利用者の情報要求を満たすクラスを良好に記述する必要がある点にある。パターン認識では、例えば文字認識のように、明確に区分された各文字を表すクラスに分類することが技術の中心となる。顔認識についても、どのような画像特徴の違いに基づいて分類を行うかをあらかじめ十分に検討しておき、その一定の基準に基づき、分類が行われればよい。

しかし画像検索では、利用者の情報要求がどのようなものであるのか予測がつかないため、きわめて多様なクラス分けに対応しなければならないだけでなく、簡単には記述できない情報要求の入力方法やあいまいな情報要求の明確化、あるいはまた、クエリーの記述からマッチング可能な形式への記述変換などに、パターン認識とは異なる課題がある（図2）。

1.4 標準化

画像検索分野の標準化の代表例として、MPEG-7とSQL/MMについて簡単に述べる。基本的にはメタデータとクエリーの基盤としての枠組みを整えてはいるものの、これらによって画像検索が完成するわけではない。

1.4.1 MPEG-7

マルチメディア検索のための標準と言われるMPEG-7は、コンテンツのメタデータの記述についての標準である。そ

図2 画像検索モデル

の中には、記述フォーマットとしてXMLを基本にすることや、効率的なバイナリー化の方式なども含まれ、実アプリケーションには有用であるが、コンテンツの検索そのものには関係しない部分と、各種コンテンツのメタデータとして、テキスト形式での内容記述などのほか、標準的な情報の種類や意味などを規格化した部分とがある。

このうち、映像コンテンツに対するメタデータとしては、色、形状、テクスチャ、動きについてそれぞれ数種類が定義されている。また、メタデータがコンテンツのどの部分を表しているのかを表現するための、時間空間双方の部分指定用の記述法も定義されている。

MPEG-7形式のメタデータが画像に付与されていれば、標準的なパーサによってその情報を読み取り、必要に応じて検索のための処理を行うことができる。ただし、注意しなければならないのは、どのようにメタデータを画像から取出すかや、利用者の情報要求からのクエリー生成やマッチング基準をどうするのか、といった点は規格外である点である。画像特徴の抽出などにリファレンスコードも用意されているが、必ずしもそれによる必要はない。

1.4.2 SQL/MM

データベースの問合せ言語の標準となっているSQLに対して、マルチメディア検索のための拡張を施したのがSQL/MMである。静止画や空間情報での問合せを記述できるようになったほか、全文検索についても拡張されている。MPEG-7とは異なり、マルチメディアコンテンツのメタデータ記述についてではなく、テーブル構造への挿入や問合せ記述にフォーカスを当てた規格である。

実際のデータ型や問合せの際の演算処理はユーザ定義として与えることで、従来のSQLの枠組みのままマルチメディアデータのメタデータを扱えるようにしている。

2. 画像内容検索

問合せと画像とのマッチングを、画像内容に基づいて行おうとするのが画像内容検索である。画像内容と呼ぶ意味は、画像に写っているものや撮影地、撮影日時などを人手によって画像に関連付けたメタデータを用いるのではなく、画像データそのものから、基本的に自動的に抽出することのできるメタデータを用いるということである。したがって、マッチングはこのメタデータの領域で行われる。

通常、画像のメタデータは、問合せによらず一定となるものを用意することで、問合せ毎に抽出する処理を軽減する。このため、実際の処理は、メタデータをあらかじめ抽出してメタデータデータベースとして格納しておき（前処理）、利用者によって作成されたクエリーとマッチングを行い、最終的に結果を出力する、という流れになる。

適合性フィードバック (relevance feedback) は、当初入力したクエリーに対する結果を利用者に提示し、その結果として利用者にクエリーの修正を直接、または間接に行わせる技法である。

図2は、メタデータとして画像から抽出した特徴量を用いた基本的な検索モデルである。画像的な情報要求を入力インタフェースと変換インタフェースを通し、画像DB中の画像を特徴量抽出フィルタによって、それぞれメタデータである特徴量での記述に変換し、別に定義したマッチング関数を適用して判断する構造であることを示している。

2.1 メタデータ抽出

画像からのメタデータ抽出は、処理が複雑で時間がかかっても検索時に高速に比較判断ができれば良い。このような画像から自動的に抽出できるメタデータは、画像特徴量とも呼ばれる。例えば、画像を構成するピクセルの色をヒストグラムとして保存しておけば、これは色に関する特徴量の一種であり、メタデータとして利用できる。

2.2 クエリー作成

例えば、色のヒストグラムをメタデータとして利用し、このヒストグラムの一致している度合をマッチングにおける類似度と決定したとしよう。この場合、画像のメタデータに対応するクエリーもまた、色のヒストグラムの情報を持つべきである。しかし、赤が12%、オレンジが7%、…などという指示を出せる人間はほとんどいない。

そこで、このような直接的な入力ではなく、「夕日」と入力されたらオレンジが20%以上、などというような人間側の入力クエリーとシステム側のクエリーとの「翻訳」が必要である。また、抽象的な画像検索を行おうというときには、このような具体的なキーワードも利用できないため、画像そのものを用いたり、印象語を利用したりされている。

2.3 マッチング

メタデータの領域でクエリーと各画像データとの類似距離を算出するのがマッチングである。マッチングの方式は

利用するメタデータによって、また、利用者から入力されるクエリーの持つ情報によって大きく異なる。人間の感覚にできるだけ近い類似距離を導く処理系が良いと言えるが、抽象的な検索では、人間の情報要求自体が不明確な状況下で、より高い満足を引き出すことが目的となるため、古典的な検索精度のような指標による評価は困難である。

2.4 結果出力

類似距離の近い順に、検索結果の画像を並べて表示するランキング表示が一般的である。この類似距離が一定の閾値以上の画像を表示しないようにすれば、簡単に2値型にもできる。もちろん、マッチング機構が2値の値しか返さない場合もありうるので、その場合はランキング表示は難しい。

1次元の類似距離ではなく、多次元のベクトル的な距離を用いる場合もある。特に適合性フィードバックにおけるクエリー修正のための結果出力では、システムのマッチング状況を利用者に伝えるための情報量が多くなるため、適切な表示を行うと全体の効率を上げることにつながる。

3. 特 徴 量

これまで多くの特徴量が研究されてきたが、MPEG-7にも採用された代表的な特徴量を中心に概説する。

まず最初は、色に関する特徴量である。基本的には、どの色のピクセルがどの程度の割合で含まれているか、という色ヒストグラムが用いられる。これに空間的な配置(位置)情報を付加したり、解像度を落としてから処理したり、頻度の少ないヒストグラムビンを保存しないなどで種類に多様性が与えられる。また、単にあるピクセルを見るのではなく、近傍のピクセルの色との組合せに特徴を見出す場合もある。

次に、テクスチャに関する特徴量である。テクスチャとは繰り返される模様を指す。特徴量としては、その細かさや方向、規則性の度合などが計算される。また、形状の特徴とも言えないものにも、エッジに関する特徴がある。エッジとは、画像の中で明るさなどが急激に変化する部分を指すが、このエッジがどの方向にどの程度含まれているか、を特徴量とする。

さらに、形状に関する特徴量がある。画像中のオブジェクトの輪郭などによって表現される。特定のキャラクタや、企業のシンボルマークなどを記述するのに向く。

動画像では、画面から検出できる動きは、あくまで動きの方向と速度だけであるが、カメラの動き(カメラワーク)やオブジェクト変化をいくつかのモデルにあてはめて記述するなどの手法がある。

MPEG-7では、これ以外に顔認識用の特別なベクトル空間による記述も定義されている。これは、人間が人の顔を識別するのに特に適応させた画像上の特徴空間と言える。このことは、これまで挙げてきたような一般的な画像特徴量

だけでは、個別の目的に充分適応させることができないということを示唆しよう。逆に言えば、MPEG-7に挙げられた以外の特徴量を研究する余地が充分にあると言える。

システム全体の性能を見る場合、特徴量空間において目的の識別・検索が行えないならば、次のマッチングにおいてどのような工夫をしようとも無駄である。

4. マッチング

マッチングは、利用者の情報要求とデータベースの画像、双方からの解釈をすり合わせる部分である。一般的なモデル化として、特徴量からなる空間をとり、各画像がその空間における点であると看做せるとき、クエリーの示す領域、あるいは点からの距離によって類似距離を求めれば良い。最終的には、特徴量空間上での利用者の情報要求に見合う領域(これを目標領域と呼ぶ)を近似するよう検索領域(システムが正解とする領域)を推定できれば良い(図3)。

利用者の入力したクエリー(外部クエリー)と、最終的なマッチングにおけるクエリー(内部クエリー)の間にはなんらかの変換が施されることが多いが、これをマッチング関数の中を含めるか、それともユーザインタフェースに含めるかはシステムとしての構成による。

4.1 距離測度

ベクトル空間的な特徴量空間が構成でき、クエリーもまたその上での1点であるならば、クエリーと各画像との距離を例えばユークリッド距離によって求めることができる。このわかりやすい例では、残念ながら、RGB色空間などで、人間の認知的距離とは大きなずれが生じ、検索精度を下げることになる。より一般的な2次形式距離、あるいはマハラノビス距離として知られる距離を用いることが多い。マンハッタン距離など非線型な距離関数を用いても構わない。

同じ特徴量空間であっても距離測度を変えることによって、システムとしての特性は変化する。クエリーの作成に含まれる誤差や、利用する特徴量の性質によって適切な距離測度を導入しなければならない。

4.2 ヒストグラム距離

ヒストグラムは情報量の圧縮と、識別性能とのバランス

図3 特徴量空間での推定

の良さから特徴量としてよく採用されるが、この各ピンを別の次元として捉えて、一般的な距離測度を適用することはできない。隣り合う（あるいは近い）ピンの間に認知的な相関があることが多く、例えば、赤とピンクの関係と赤と青の関係は大きく違うだろう。したがって、2つのヒストグラムを比較するとき、その重なり合わない部分の大きさを見る、ヒストグラムインタセクションのような単純な手法も不適当である。

色の場合、一般的な認知的相関を調べた例があり、その値を用いれば良いと思うかもしれない。しかし、一般的な関係が明らかではない特徴量では、そういった既知の値を使うわけにはいかない。

検索特有の事項として、そもそも一般的な認知的距離が有効なのかどうか、という疑問がある。確かに一般的な状況では、赤とオレンジは近い色だろう。しかし、赤い花柄の洋服が欲しいといった場合に、その人がオレンジ色の花柄をどう許容するかは、人によって、また検索によっても大きく異なると考えられる。仮に赤から見て、一般的にオレンジとピンクが同程度の距離だとしても、それが今の検索にも適用できるとは限らないのである。

4.3 クラスタリング

マッチングにおける一つの理想的なモデルとして、検索目的に適合した距離測度が推定できればよいと言える。しかし、連続的な類似距離を正確に推定するためには、非常に多くのパラメータを推定する必要がある。そのため、離散的、あるいは2値の類似距離しか与えられないが処理の簡単のためにクラスタリングを行うモデルもよく見られる。

クラスタリングモデルでは、特徴量空間を適当なクラスに区切り、各クラスに適当な距離を与えるのである。クラスターの境界は、あらかじめ与えられていても良いし、クエリーによって動的に生成されても良い。例えば、赤、オレンジ、ピンクというような色ヒストグラムしか特徴量として持っていない場合、それぞれの色に対応するクラスを用意しておけば充分である、という考え方もできる（これは必ずしも正しくはないが）。

クラスター境界を動的に定める方式として基本的なのは、検索目的に充分類似する事例と類似しない事例を集めておき、ボロノイ領域として分割する手法である。特徴量空間における認知距離に対する知識が充分でない場合、この程度のクラスタリングが限界となる。ただし、充分な数の事例を用意しなければならない点や、矛盾するような事例が近接して存在した場合にどうするか、など検索モデルに即して修正する必要がある。

5. ユーザインタフェース

画像検索がメタデータの領域、すなわち特徴量空間におけるなんらかの距離測度に基づいた類似距離によって行われることを述べてきたが、類似距離の算出にあたっては、

検索毎に異なる利用者の情報要求を反映したものである必要があることを指摘してきた。

単純な類似画像検索に見るように、目的となる画像1枚、すなわち、特徴量空間の1点だけが与えられた場合、マッチング機構としては一般的な距離の分布を用いるほかはない。その検索における許容範囲が広いのか狭いのか、あるいは、どういった違いなら広く許容するのか、といった情報もクエリーに含まれていなければならない。このような情報をクエリーに付与するのがユーザインタフェースの役割である。

5.1 クエリー作成

マッチング段階で必要となる内部クエリーはしかし、利用者に直接入力してもらうには不適切である場合が多い。例えば、45°の傾きを持ったエッジが15%、などという数値が自分の情報要求に見合ったものだとかわかるだろうか。クエリー作成は利用者の直観的な理解をもとに簡単に入力できることが求められる。

古典的なキーワード入力によるクエリー作成インタフェースでは、システムの持つ、キーワードから内部クエリーへの変換テーブルにしたがって変換を行う。固定的なテーブルに対して、形容詞などにより修正を行い、そのときの検索の目的に合わせる工夫も行われている。

より直観的な入力系として、利用者にスケッチを描かせるものや、その特殊な形として画像そのものを用意してもらう類似画像検索がある。この入力画像から、内部クエリーの要求する特徴量を抽出すれば良い。

5.2 適合性フィードバック

しかし、いずれの手法をとろうとも、基本的に特徴量空間の1点に写像されることになる。利用者の許容範囲やその偏りの情報を得るためには、複数の点の類似距離を得るが必要があり、そのための手法として適合性フィードバックと呼ばれるシステムが考案された。

適合性フィードバックでは、直前のクエリーに対する検索結果を利用者に提示し、利用者はその画像の適合度を判定する。最も簡単には、複数の検索画像を提示し、似ている画像と似ていない画像とに分類してもらう、という手法が用いられる。

複数の点の類似距離を得るだけなら、利用者に複数の画像を入力してもらうなどの手法も考えられるが、例えば、似ている絵と似ていない絵の両方を描いてもらう、などというのは入力の手間がかかりすぎる。その点、適合性フィードバックでは、システムが本来備えているマッチング機構を生かしながら、複数の類似距離を得ることができ、利用者の負担とシステムの処理の双方を軽減することができる。

こうして得られた類似距離をどのように内部クエリーに反映するかはシステムによって異なる。また、フィードバックによるクエリー修正を複数回繰り返して、より精度の高いクエリーを得ようとする手法が一般的である。

6. 抽象的な画像検索

画像検索特有の研究課題が、利用者の曖昧な情報要求をいかにすくい取るかにあることを述べた。そのためには、適合性フィードバックを基本としたクエリー修正を繰り返すインタフェースと、柔軟な類似距離測度を導入可能なマッチング機構が必要であると言える。

このような抽象的な画像検索を実現する手法として、筆者らが提案した手法³⁾⁴⁾を紹介したい。

クエリー作成を担うインタフェースは、当初から画像を用いるフィードバックインタフェースである。ただし、実際の画像を用いるのではなく、システムが把握している特徴量のみを反映した人工生成画像を提示する。その理由は、情報要求を明確化できていない状況下では、具体的な映像の提示は、その具象物に判断が引きずられる可能性があるためと、システムが把握できていない特徴量による影響を排除するためである。また、表示したい特徴量に見合った画像を生成できるため、クエリー修正のアルゴリズムが単純で効率の良いものを作成しやすいという利点もある。図4が、選択式インタフェースで表示する画像の模式図である。対象物の画面上の位置と背景色を体系的に変化させ、情報要求に近い画像を選択することで、利用者の情報要求を読取っていく。

クエリーによる目的の領域の記述は、矩形(各次元の範囲)表現や楕円体(マハラノビス距離の閾値処理)を用いた。典型的な色やエッジの特徴量を用いた、比較的低次元の特徴量空間であっても抽象的なイメージを相当程度汲み取り、画像検索を行えることを示した。

採る特徴量空間やマッチング方式についてはほかにもさまざまな手法が考えられるが、抽象的な画像検索を行う際には、意味内容によって判断を狂わされない抽象的な画像をキーとするインタフェースが合っていると思う。もちろん、

図4 選択式インタフェースの提示画像例

当初入力はスケッチ的な要素があっても良いが、現在筆者らが取組んでいるような動画像への拡張時には、扱いやすいスケッチングツールの導入が必要である。また、的確な距離測度推定のためには、何らかの形で適合性フィードバックが必要であると考えており、当初から選択式のインタフェースを用いることで、実装の簡単化やユーザインタフェースとしての扱いやすさも確保できていると考えている。

7. む す び

本稿では画像検索が、テキスト検索やパターン認識と技術基盤を共有する部分と、独自の課題が存在することを指摘し、画像検索の基本的なモデルを示した。画像検索の課題は、人の感覚にも依存する部分があり、単なる画像処理や数学モデルだけで解決できる問題ではないと考えている。それだけに、従来の理想状態へ向かうことを良しとする工学的評価にも困難が付きまとう。しかし、映像に関わる人々にとって挑戦的な課題でもあり、周辺分野の経験を持つ、新しい知恵を持って立ち向かう人の参入を期待したい。

(2005年6月6日受付)

〔文 献〕

- 1) 徳永健伸: “情報検索と言語処理”, 東京大学出版会 (1999)
- 2) R.S. Taylor: “Question-negotiation and Information Seeking in Libraries”, College & Research Libraries, 29, 3, pp.178-194
- 3) 小林亜樹, 吉田俊之, 酒井善則: “特徴量空間の領域縮小によるインタラクティブな画像検索法”, 映情学誌, 52, 7, pp.1040-1047 (1998)
- 4) 小林亜樹, 吉田俊之, 酒井善則: “距離尺度の推定によるインタラクティブな画像検索”, 映情学誌, 54, 4, pp.583-592 (2000)
- 5) 趙 剛, 小林亜樹, 酒井善則: “ラフ集合論の縮約による多グループフィードバックを用いた対話型画像検索”, 映情学誌, 59, 6, pp.884-893 (2005)
- 6) 国枝孝之, 脇田由喜, 高橋 望: “MPEG-7と映像検索”, CQ出版社 (2004)
- 7) 小川英光: “パターン認識・理解の新たな展開”, 信学誌 (1994)
- 8) 長尾真: “コンピュータのパターン認識”, 東京大学出版会 (1985)
- 9) 鳥脇純一郎: “認識工学—パターン認識とその応用—”, 映情学, コロナ社 (1993)
- 10) 坂内正夫, 大沢裕: “画像データベース”, 昭晃堂 (1987)
- 11) 亀山 渉, 花村 剛 監修: “改訂版デジタル放送教科書 下”, インプレス (2005)

小林 亜樹 1995年, 東京工業大学工学部情報工学科卒業。1997年, 同大学院理工学研究科電気・電子工学専攻修士課程修了。2000年, 同大学院博士課程修了。同年より, 同大学院集積システム専攻勤務。画像検索, ネットワーク情報検索の研究に従事。博士(工学)。