

## 考古学データを対象とした画像検索手法の研究

5R-2

北幸典 岡安光彦 石川佳治 植村俊亮

奈良先端科学技術大学院大学 情報科学研究科

### 1. はじめに

現在、考古学の分野では、遺跡発掘情報の電子化が盛んに行なわれようとしている。遺跡発掘に際しては、様々な種類のデータが大量に生じる。結果として、電子化され、蓄積される考古学情報は莫大なものとなる。そこで、それら電子化され、大量に蓄積された考古学情報を効果的に利用していくために、それをサポートしてくれるような情報システムが期待されている<sup>6)</sup>。

考古学分野の報告資料には、出土品の図や写真などの情報が数多く含まれるため、その電子化に際しては画像や図の管理・検索機能が求められる。本研究では、考古学の報告書に埋め込まれた画像データを対象とし、類似検索を行なう考古学画像データの検索システムを提案する。

### 2. 背景

近年、マルチメディア情報システムが大きく注目され、画像データの管理の重要性が高まってきている。本研究が対象とする考古学分野においても、今後膨大な量の画像データが活用されると考えられ、大量の画像データの管理・検索機能が求められると考えられる。

#### 2.1 画像検索のアプローチ

これまでの画像検索システムは、画像に付加した文字列で表現される属性（ファイル名、タイトル名、キーワードなど）だけに基づいた検索アプローチをとることが多かった。しかしながら、画像のような直観的なデータを文字だけで表現しようとするのは困難である。

一方、画像の客観的な物理的特徴に基づくアプローチが存在する。この画像の内容表現には様々なものがあり、テクスチャ、色、形状、位置などが存在し、これらの組合せによって画像の内容が表現される<sup>4)</sup>。

実際には、対象分野の画像データや要求に合わせて、文字情報と画像の特徴量を併用して検索の効率や精度を高めることが求められる。

#### 2.2 画像検索の効率化手法

画像の検索処理には主に以下の二つが考えられる<sup>3)</sup>。

1. 問合せ画像と、データベースに格納している全ての画像とを個々に比較する。
2. 検索を効率的にする索引付け手法を用いる。

多くのシステムで1.の逐次的に比較を行なっていく手法が用いられているが、画像データベースに格納されている画像の数が増え、データベースのサイズが少

しでも大きくなると、すぐに手がつけられないほどの遅さになってしまう。

本研究では2.の索引を用いる手法を選択し、画像の特徴量に基づく検索に空間インデックス (spatial index) を利用することを考え、効率的な画像の内容検索が行えるようにすることを目標としている。

有力な空間インデックスとしては、(a)R-木、(b)quad tree、(c)グリッドファイル (grid file) などがあり、それらを改善した手法も数多く提案されている。これらの手法のいくつかは、次元数に対して指数関数的に爆発してしまう、という欠点をもつ。画像データベースでは一般に高次元のベクトル情報を扱う必要があり、次元に対するロバスト性は重要である。本研究では、もっとも次元に対してロバストであるといわれるR-木に類する手法の利用を想定する。

### 3. 考古学データを対象とした画像検索システム

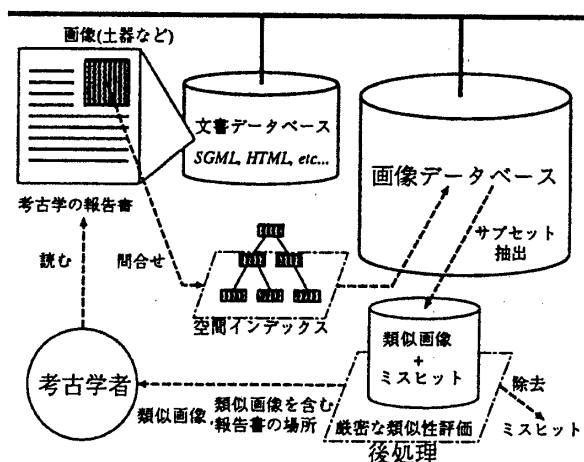


図1 考古学画像検索システム

一般に考古学の報告書類には図表や写真が豊富に含まれており、考古学者には、「この画像に似た画像はないだろうか?」、「この土器と似たような土器に関する報告書はないだろうか?」といった問合せを行いたいという要求がある。

そこで、電子化された考古学の報告書に埋め込まれた画像(例:土器など)を対象とした画像検索システム(図1)を提案する。画像の内容表現に従って、各画像を特徴ベクトルへと変換する。すなわち、画像は特徴空間中の点として表現される<sup>2)</sup>。この特徴空間を空間インデックスに対応させることによって、画像データベースは空間インデックスを利用して効率的に類似検索を

A Retrieval Technique for Archaeological Image Data  
Yukinori KITA, Mituhiko OKAYASU, Yoshiharu  
ISHIKAWA and Shunsuke UEMURA  
Graduate School of Information Science,  
Nara Institute of Science and Technology (NAIST)

行うことができ、データベース自身に格納している類似画像や、類似画像を含むような報告書の場所を返す。

このシステムで考慮すべき問題点には以下のものがある。

1. 文字で表現される属性だけに基づいて問合せを行なうことは困難(不十分)である。
2. 大量の画像データが存在しうる。
3. 考古学者による検索要求は多様である。

1. に関しては、例えば報告書に埋め込まれた画像そのものを問合せとして、類似画像の検索を行なうことで、より自然で直観的な問合せが可能となる。ただし、文字だけに基づく検索と比較して、画像の特微量に基づく検索は計算機にとって重たい処理である。したがって、いかに効率的に画像の特微量に基づく検索を行なうか、という別の問題が起こる。また、2. の問題では、莫大な数の画像データに対して効率的に画像の特微量に基づく検索を行なえる手法が求められている。そこで、これらの二つの問題を同時に解決するために、次のような2段階のアプローチをとる。

1. フィルタリング 空間インデックスを用いて、大雑把な内容に基づいて類似検索を行ない、類似画像の候補(ミスヒットを含む)としてのデータベースのサブセットを抽出する。
2. 後処理 より厳密に類似性を評価し、類似画像の候補からミスヒットを取り除く。

この手法を適用することにより、画像データを効率的に検索できるだけでなく、データベースのサイズの成長に対しても有効な検索システムの構築が可能となる<sup>1)2)</sup>。

3. の問題は、例えば土器全体の形状で検索したい場合もあれば、ある特定の部分の形状(例:土器の底面部)に注目したい場合も考えられる、ということである。そのため様々な特微量を要求に応じてうまく組み合わせる必要がある。

#### 4. 考古学情報システムへの統合

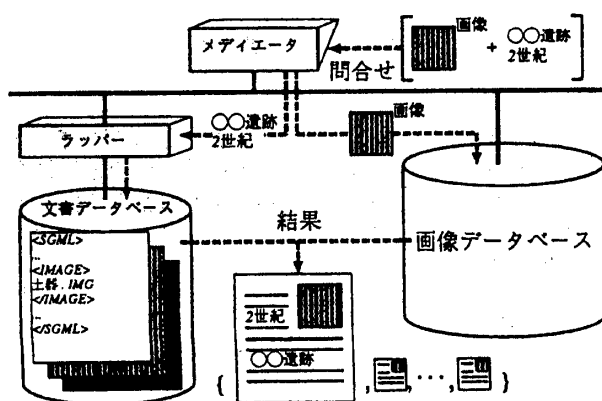


図2 考古学情報システムにおける問合せ例

本研究室では考古学情報を対象とした情報システムの研究を行なっている<sup>3)</sup>。考古学の分野では、情報の電子化が各地の調査団体ごとに独立して行なわれているため、異なる情報の画一化をサポートする統合化された情報システムが必要とされている。

5) では、構造化文書(SGML文書)として各サイトに格納された報告書のデータをメディアエータ(mediator)により統合するアーキテクチャを提案している。特に、利用者からメディアエータに与えられた問合せを構造化文書データベースへの問合せに変換するラッパー(wrapper)機能について論じている。

本研究で提案する画像検索システムは、最終的にはこの考古学情報システムに統合され、考古学者に対する高度な研究環境を提供する。構造化文書として電子化された報告書からは、5) のラッパー機能により画像部分が抽出され、画像検索システムに送られ、索引情報が生成される。これにより、例えば、「○○遺跡で発掘された、2世紀の土器で、この土器(画像)に似たようなものに関する報告書」のような、SGML文書による報告書の内容に関する問合せと、報告書に埋め込まれた画像に関する問合せを融合した問合せ(図2)を行なうことが可能になる。

#### 謝辞

本研究を進めるにあたり有意義な御指導・御討論を頂いた植村研究室の皆様へ感謝いたします。

#### 参考文献

- 1) Rakesh Agrawal, Christos Faloutsos, and Arun Swami: Efficient Similarity Search in Sequence Databases. In FODO Conference, Evanston, Illinois, October 1993.
- 2) H.V. Jagadish: A Retrieval Technique for Similar Shapes. *Proc. ACM SIGMOD Conf.*, pages 208-217, May 1991.
- 3) Youping Niu, M. Tamer Özsu, Xiaobo Li: A Study of Image Indexing Techniques for Multimedia Database Systems. Technical Report TR 95-19, University of Alberta, July 1995.
- 4) C. Faloutsos, R. Barber, M. Flickner, J. Hafner, W. Niblack, D. Petkovic, and W. Equitz: Efficient and Effective Querying by Image Content. *Journal of Intelligent Information Systems*, pages 231-262, 1994.
- 5) 古館文裕, 岡安光彦, 石川佳治, 植村俊亮: 構造化文書データベースに対するラッピング手法の提案. 情報処理学会データベースシステム研究会研究報告, July 1996.
- 6) 岡安光彦, 石川佳治, 植村俊亮: データベースを基盤とする考古学ネットワークシステム. 情報処理学会データベースシステム研究会研究報告, July 1996.