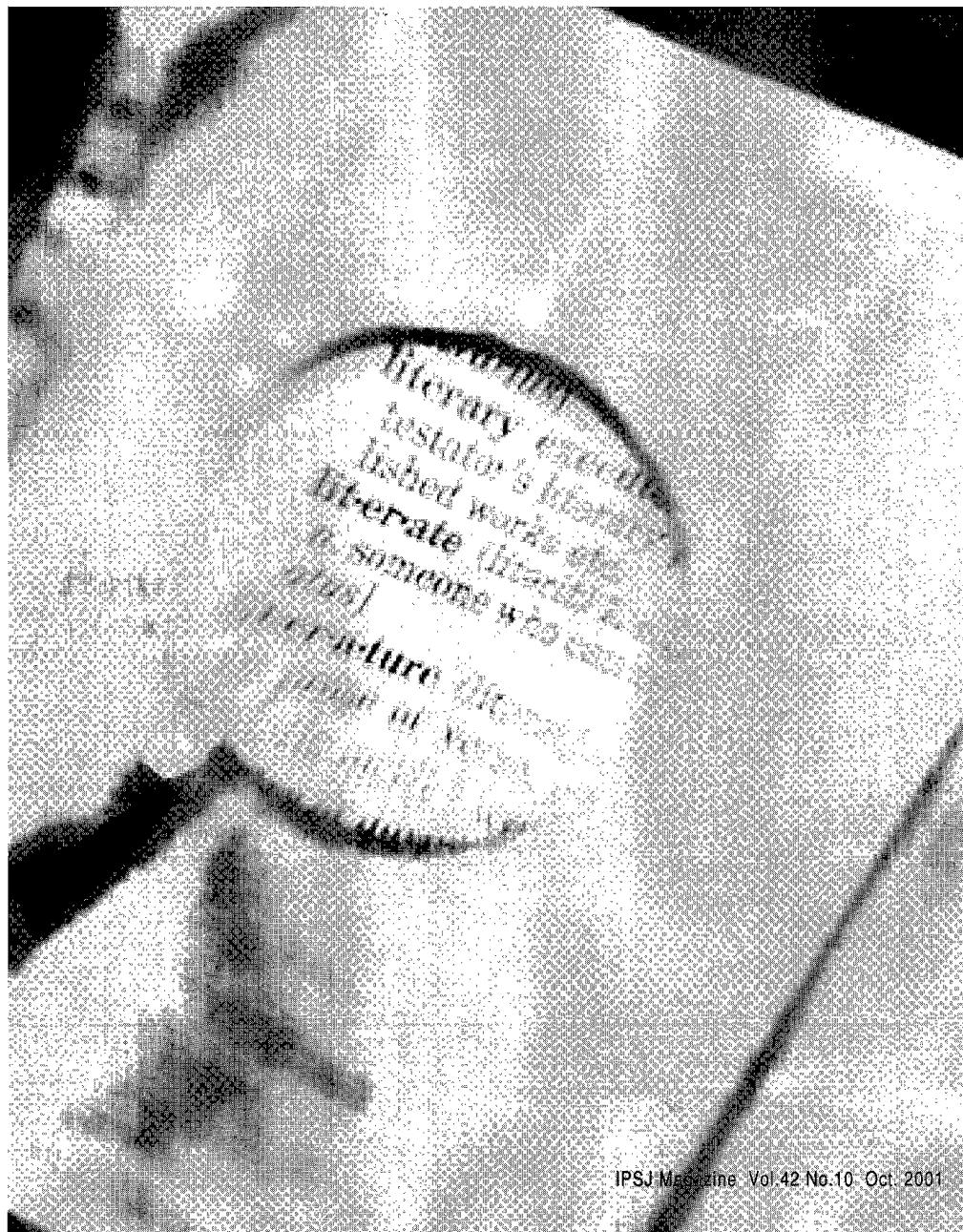


## 編集にあたって

有次 正義 群馬大学工学部情報工学科 aritsugi@cs.gunma-u.ac.jp



いうまでもないが、計算機技術の発達はめざましい。データ工学の分野では、数値や文字などの簡単なデータのみを扱う時代はとうに過ぎ、現在のアプリケーションでは非常に複雑で大量のデータを高速に操作することが常に求められている。本特集では、マルチメディアデータを扱う上での検索技術を分かりやすく解説することを目指す。

データが大量に集まると、その中から所望のものを高速に検索したいという要求が生じる。この要求を満足するためには索引は必要不可欠である。テキスト、図形、画像、音声、動画といった、いわゆるマルチメディアデータに対しても同様である。しかし、マルチメディアデータに対しては、キーを1次元の直線上の点とすることが困難なことが多い。また、範囲検索が多次元にわたるという問題があり、従来の数値・文字に対して用いられてきた索引技術をそのまま適用するというわけにはいかない。

本特集では、大量のマルチメディアデータから所望のものを高速に検索するための、最新の索引技術を解説・紹介する。ここでは、画像等に対する類似検索のための索引技術、地理情報システムで扱われる空間データに対する索引技術、時間的に変化するデータに対する索引技術、ならびに文書データを対象とした索引技術について述べる。各著者には、技術的詳細にあまり立ち入ることなく、直感的な理解を得やすいような記述をしていただいた。

ここで紹介される新しいインデックス技術は、従来技術に基づき、それらを拡張・応用したものである。そのため、その詳細を理解するためにはB木(B-trees)や多次元空間などに関する基礎知識が必要となる。基礎的な索引技術に関しては、文献3)などに詳しい。R木(R-trees)などについては、比較的新しい教科書として文献1)や文献2)などがあるので、詳細を知りたいときにはそちらを参照していただきたい。また逆に記事の中では、あまりその詳細に立ち入らずにトピックの紹介と関連する文献の参照をするだけにとどめている部分も少なくないが、これは、できるだけ幅広い解説にするためとご理解いただきたい。

以下、特集の構成について簡単に紹介する。

第1編は、マルチメディアデータのための索引技術の設計開発について述べている。ここでは、マルチメディアデータを、特微量を用いた表現をした場合を例にとり、その性質や利用形態を考えた時に索引に要求される事項をまとめている。また、多次元データに関する最新

の索引技術のエッセンスを簡単に紹介している。複雑で大量なマルチメディアデータに対する研究課題と最近のトレンドが簡単に分かるようになっている。

第2編は、画像や映像の類似検索のための索引技術を述べている。画像や映像に対する類似検索では、特微量ベクトルに基づく手法が広く用いられている。その基本的な概念と、実際に特微量データがどのように扱われているかをやさしく解説している。最近のトレンドや、何がどこまでできるようになり、どのような技術的課題が残されているかを述べている。第1編と重複する内容もあるが、特に画像などの類似検索という面からより詳細に述べられている。

第3編は、空間データ、すなわち地理情報の効率的管理と索引技術について述べている。地理情報システムでは、大量の空間データを効率的に管理し、それに対する空間位置、距離、接続・位置関係に基づく検索を高速に処理することが必要である。ここでは範囲検索処理を高速化する空間データ索引技術について解説している。本構造に基づく種々のデータ構造を紹介し、各々の特徴が理解できるようになっている。

第4編は、時間的に変化するデータに対する索引技術を述べている。実世界では時間的に変化するデータは数多くあり、その変化の情報もまた検索の対象となる。ここでは、データベースにおける時間のデータを扱う意味や目的に始まり、そのデータを効率的に扱うための検索技術をいくつか取り上げ紹介している。

そして第5編では、文書データを対象とした索引技術について述べている。文書データの扱いは、WWW、メール、電子出版などを考えるとこれからますます重要な技術である。ここでは、キーワードを対象とした検索技術、文字列を対象とした検索技術、および構造化文書を対象とした検索技術について最新のトピックを紹介する。

#### 参考文献

- 1) Bertino, E., Ooi, B. C., Sacks-Davis, R., Tan, K.-L., Zobel, J., Shidlovsky, B. and Catania, B.: *Indexing Techniques for Advanced Database Systems*, Kluwer Academic Publishers (1997).
- 2) Garcia-Molina, H., Ullman, J. D. and Widom, J.: *Database System Implementation*, Prentice Hall (2000).
- 3) 西尾章治郎、横田一正、北川博之、石川佳治、有川正俊、井田昌之: 情報の表現(岩波講座マルチメディア情報学3), 岩波書店 (2000).

(平成13年8月27日)