

マルチメディアシステムにおける情報アクセス

Information Access Methods in Multi-Media Systems

坂 内 正 夫*

Masao SAKAUCHI

マルチメディアシステムは、メディアコンピューティングの分野で特に大きな注目を集めている。映像を中心とするマルチメディアシステムで重要な問題は、いかに情報の中身を認識してモデルをつくりデータベース化し、魅力的に使用していくかということである。そこでこの解説では、そのための基本技術を概観し、ついでそのなかでキー技術となる情報アクセス方法、検索方法について新しい視点の技術を紹介している。

1. ま え が き

1960年代中頃に、「ハイパーメディア」という用語を最初に用いた米国の Ted Nelson は、長年高度な情報保存/取出しシステム「Xanadu」(誰もが自由にこのシステム上で出版でき、それに誰もがアクセスでき、いわば蓄積した膨大な情報ワールドに「窓」を設定できるシステム)を構想し、一部をリリースしはじめている。Xanaduでは、ユーザにより作成された文書、著作物だけではなく、映像や映像物といった情報が蓄積され、情報ワールドを形成している。

この Xanadu をひきあいに出すまでもなく、われわれにとって多くの情報をもつ文書や図面、写真や映像物、リアルタイムの映像等の画像情報、さらには音声情報を含めたマルチメディア情報をコンピュータを通して自由に、かつ便利にアクセスすることができるようになれば、われわれにとってのコンピュータの世界、メディアの世界は飛躍的に拡大することになる^{1),2)}。この目的のために、多量の画像やマルチメディア情報を蓄積管理しておき、ユーザが高度に活用できるための情報アクセス手段や演算操作手段を提供するシステムを、マルチメディアシステムと呼んでいる^{5),12)}。本稿では、画像情報を中心とするマルチメディアシステムについて、特に情報アクセス=(情報/メディア世界への「フレンドリーな窓」の設定)という観点に重点をおいて、今後の必要性が高まると考えられる視点を展望してみたい。

2. マルチメディアシステムと情報アクセス手段

2.1 マルチメディアシステムの要素

図1は1.で述べたマルチメディアシステムを概念的に示したものである。ユーザは、形成されたメディア情

報の空間に自由にアクセスし、そのユーザワークスペース上で自由な情報利用を行う。図2はこのようなマルチメディアシステムを実現するための構成要素を示している。

対象とする「マルチメディア情報素材」は静止画、動画などの一般画像から、図形データ、文書、図面、地図等までを含む形状や相互関係などの幾何学的情報が本質である各種の情報、およびそれらの結合情報である。それらは、生活や文化・ビジネスにとって有意・有効な情報空間形成のベースとなる。また、映像への人間の欲求を反映して今後は、より実用的、より高品位、より高い具像レベルや現実感のあるデータへの指向が強まると考えられ、カラー動画、ハイビジョン映像、立体映像なども重要な対象となっていく。

「データベース形成」は、対象メディアからの必要情報の抽出、メディア変換、情報結合やキーワード抽象などの情報空間へのアクセス路設定のための準備を行う機能である^{5),6)}。これを効率的に実行していくためには、外界世界のシーンを理解する画像・映像の認識理解技術が必要である。マルチメディアシステム(データベース)のような応用指向性の強いシステムでは、将来的にもハードウェア、ソフトウェアに比してデータの占めるウェイトが高まり、特に、このデータベース獲得コストが支配的になるケースが多くなる(マッピングデータベースやハイパーメディアでは、すでに多大なコストがこのデータベース形成に投入されている)⁸⁾。

画像符号化・蓄積・データ構造化⁹⁾、モデル化などの「データベース管理機能」は、画像メディアのもつデータ量の多さや幾何学性に対処するためのものである。このうち特に、画像を一塊として扱うときに重要なデータ圧縮符号化については、ディジタルコサイン変換をベースにする方式が、画像通信、医用DB等で標準的になり

*東京大学生産技術研究所 第3部

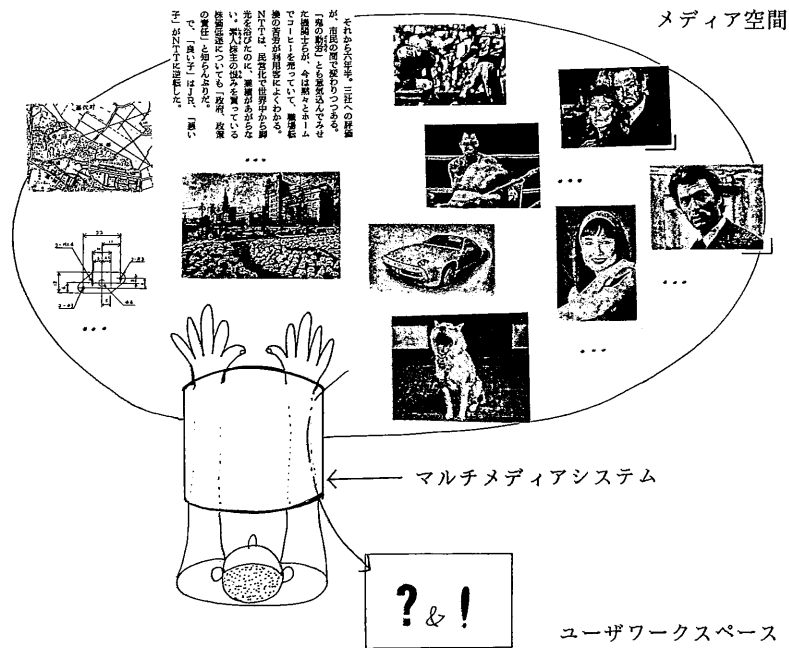


図1 マルチメディアシステムのイメージ

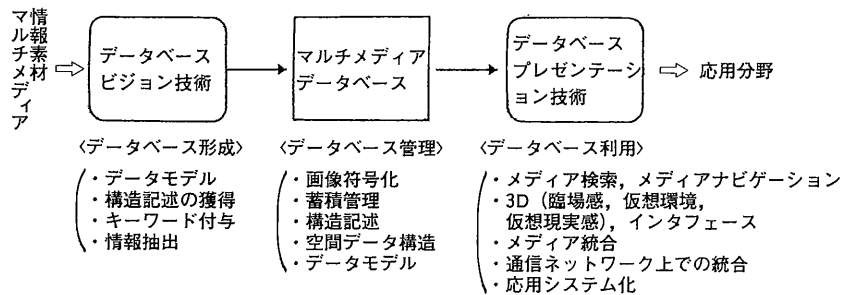


図2 マルチメディアシステムの技術要素

つつあり、また光ディスク、CDなど画像ファイル装置も一般化してきていて、研究としては一定の成熟をみたといえよう。

メディア結合、メディア検索インタフェース、通信との融合などの「データベース利用」機能はマルチメディア情報を応用し、外部に生かす演算・操作全般に対応するものである。特に、メディア空間から所望の情報を取り出すために、いかにフレンドリーに情報アクセスが達成されるかが、当面重要な課題となる。この際、ユーザの多様な発想やヒューマンインタフェースの充実、マルチメディア情報本来のもつ魅力とおもしろさの提供を十分考慮しておく必要がある⁵⁾。

また「応用システム化」は、個別の応用の下に、以上の要素を組み合わせていくわけであるが、経済的・人的・社会的・文化的側面などの各種の技術以外の要素をも考慮して、適応的・動的にシステムの最適化、立ち上

げを行っていく技術である^{2), 13)}。マルチメディア情報は、応用によってその活用のされ方が異なり、また応用による歴史もあるため、個別的な視点をもつことが不可欠である。また、マルチメディアデータベースでは著作権の扱いが、1つのポイントになる。これは、「おもしろい」映像情報がいわばプロの作成によるものが多いこと、改ざん・変形をともなって初めてマルチメディア情報本来の付加価値発掘ができること等が背景となっている。

以上のシステム要素は、いずれもシステム形成のために不可欠であるが、メディア空間から自由に情報が取り出せることを重視すれば、特に、情報アクセス路の設定が重要である。以下、これに重点をおいてさらに詳しくみてみよう。

2.2 フレンドリーな情報アクセス手段

画像/マルチメディア情報が蓄積された「メディア空

間」から、どのように情報アクセスができればよいのだろうか。図2に模式的に示すように、「マルチメディア情報空間」ユーザの間を、「アクセスツール」が結合しフレンドリーなインタフェースを実現する。マルチメディア情報空間は、文書やグラフィック情報、静止画や動画、さらには放送メディア等のライブ情報などから形成されている。

ユーザは情報収集、デザインの発想の獲得（デザイン支援）、アミューズメント、教養獲得、計画評価など多様な目的でこの空間にアクセスしてくる。

たとえば、「恐竜時代の冒険体験をしたい」、「軽やかな夏の挨拶状の参考になる資料をさがしたい」、「重厚感のある建築デザインのための参考資料をさがしたい」、あるいは「どのTVチャンネルでもいいから、天気予報になるか、大相撲の小綿と曙戦になったらアクセスしたい」等々である。ユーザは独自の判断や好み、感性を生かして情報選択ができることを望むだろうし、また、情報選択の結果が次の情報選択行動に変化や影響を与えることも当然有りえる。

「アクセスツール」は、このようなユーザの多様で個性的、動的な情報選択を、情報空間内のメディア情報に対して実行できる手段を提供する必要がある。典型的には、図3にあるように、ユーザに対し、まずデータiを、次にj, k, l...という様にデータ系列を提示、提供する。これらは、ユーザの目的により、特定の目標情報に向かって検索を行っていく経路を形成したり、また、このデータ系列自身が1つの情報体験を与えたり、また何かのHow to教育の効果を与えたりすることになる。このような経路をユーザにできる限り便利に選択させることをいかに実現するかが、アクセスツールの中心課題である。このアクセス経路はその形態からみると、次のようなものに分けて考えることができる¹³⁾。

(1) ダイレクトアクセス：通常のデータ検索で、検索条件をユーザから得て、それを満たすデータ、またはデータ群に直接的にリンクを設定する形態。

(2) ハイパーリンクアクセス：図2に示された形態のように、ユーザを現在参照しているデータ（群）から次々に別のデータへと導いていく形のリンクを設定する形態。この形態は、リンクのはり方が最大の問題である。現在いわゆる「ハイパーメディア」ソフトでは主として、作成者が固定的にリンク（選択的ではあるが）を設定したものである。しかし、この固定リンクをより柔軟なリンクに変更したり¹²⁾、さらには(1)のような自由な条件設定やそれまでの経路履歴を反映した動的なリンクをはる¹⁷⁾ことが今後必要となってくる。

2.3 情報アクセス手段実現の課題

では、上記の(1)や(2)のアクセス形態を魅力ある形で実現する上での課題は何であろうか？

たとえば、次のものが考えられる。

(a) 検索あるいはリンク設定の条件の与え方と対応する検索方法。（これについては13)等を参照されたい。）

(b) おもしろく、フレンドリーなインタフェース環境。

(c) (a)(b)を自動化された環境で実現すること。

特に(c)はシステム実現の上で今後重要性を増大してくる。たとえば(a)で画像検索のためのキーワードを抽出・利用する段階で現在のように人手に頼る形では非行率的でコストがかさみ、マルチメディアシステムを一般的に普及させるのに大きな障害となるからである。また、マルチメディアシステムの有力な対象として考えられるリアルタイム性の高いメディア情報に対してはあらかじめ人手によるキーワードやリンク設定を行うことには限界がある。そこで以下はこの(c)にポイントをしぼり、筆者が行っている2つの「自動化された環境でのマルチメディア検索」のアプローチを御紹介する形で、メディア情報へのアクセス路設定の実例を述べよう。

3. 映像シーンへの自動アクセス

蓄積されていたり、放送型で提供されている動画像群に対して、ユーザの所望する検索条件やシーン状況を指定して自由にアクセスでき、しかもそれが人手を介さずに実現できれば、2.2で述べたメディア情報のフレンドリーなアクセスの一面を達成することができ、魅力的なマルチメディアシステム提供の基礎になる^{3), 4)}。この考の下に、筆者らが開発中の映像シーンへの自動アクセス方式¹⁶⁾を簡単に紹介しよう。（詳細は、本小特集の速報「シーン記述言語を用いたマルチメディア検索システム」を参照されたい。）

情報アクセス路設定のためには、まずユーザの意志の記述が不可欠である。このシステムではシーンセグメントの色の種類や、配置と位置関係および色の動き情報などをベースにするビデオ映像のシーン記述言語（Video Scene Description Language）を用意している。VSDLは(1)カラー定義層、(2)セグメント定義層、(3)シーン定義層の3層構造をもち、VSDL シンタックスによって相互に関連づけられている。ユーザはまず検索あるいは分類

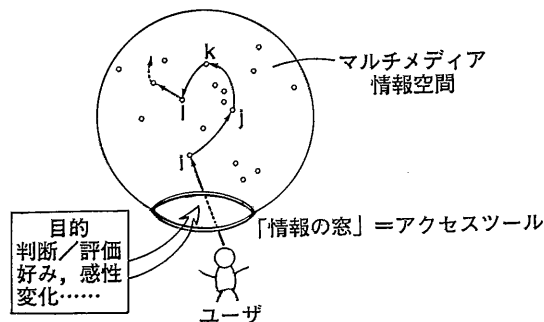


図3 マルチメディア情報空間へのアクセス

したいシーンを想定し、シーン記述言語を用いて想定したシーンを記述する。一度に複数のシーンを記述してもよい。

システムはユーザが定義したシーン記述を解釈し、蓄積または、入力されるビデオ映像とこれらのシーン記述との間で自動的にマッチングをとる。マッチングが成功すれば、ユーザ定義のシーンが入力映像から検出されたとみなす。マッチング処理は同一場面を単位とし、場面の変わり目（シーンチェンジ）が検出されたらマッチング処理を最初から繰り返す。

システムは拡張性を考慮し、オブジェクト指向的な構造を取り、Measurement Class (MC)、Sub Area Measurement Class (SAMC) と Scene Class (SC) の3種類のクラスから構成されている。MC はデータ部には画像から抽出されたすべての色情報を格納するためのデータベース、メソッド部には画像から色情報を抽出するための“Detector”群と、他のクラスと情報交換のためのメソッド群が含まれている。SC は検出および分類したいビデオシーンに対応しており、データ部は対応するビデオシーンの属性を記述するための属性テーブル、メソッド部は MC から送られてくる情報を属性テーブルマッチングをとるための“Comparator”群等から構成されている。色の分布と配置を効率よく記述するために、本手法の現バージョンでは画面を 3×2 の6面に分割した6つの SAMC によりそれぞれの領域ごとに調査を行っている。動き情報は色単位で抽出される。システムはファイル“Scene Def. sdl”に格納してあるユーザのシーン定義を解析し、シーン定義に対応する SC 群を生成する。その後、MC は6つの SAMC の助けを仰ぎながら上述の色情報や動き情報を対象映像から順番に抽出し、それぞれの SC にブロードキャストする。各 SC は MC から送られてきた結果と自分の属性テーブルに記述されている対応属性と比較して、照合を実現している。

この方式により、さまざまなビデオの映像シーン19例を対象に記述分類の評価実験を行っている。その結果、「カーレースの会場全景」や「バレーボールのネットをはさんだ攻防シーン」の記述困難な2例を除いて、的確に分類することができた。また、自動処理の時間もリアルタイム化が可能な水準であった。

ここで御紹介した例はより一般的な映像を扱う場合、まだ十分とはいえないが、映像メディアへの新しいアクセス路設定方式の一環として発展させられると考えている。

4. 「それなり論理」によるキーワード自動抽出を行う 画像データベース

メディア情報への自在なアクセス路設定の他の試みと

して、画像の内容理解を完全自動の形でいき、抽出された画像の内容をキーワードとして持たせるタイプの自動化を考えることができる。これが実現できれば画像情報を単にファイルにボンボンいれていくだけで利用者は自分の所望の条件をシステムに示して画像の内容に依存した自在なアクセスを行うことができる。

しかし、ここに大きな問題がある。一般の画像に対して、現在の画像理解技術は内容理解を十分満足のいく水準まで実行することができないことである。そこで、その突破方法の1つとして筆者らによる「認識適応型画像モデル形成と利用」、いいかえれば「それなり論理」による画像データベース形成方式を紹介しよう^{14), 15)}。

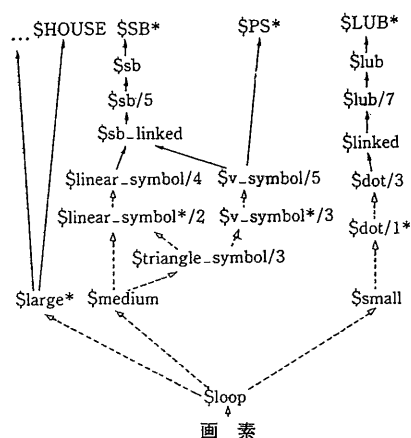
この方式では複雑な画像映像メディアの内容理解は、当分の間進歩はするがまだ十分でないことを前提としたデータベース形成法をとっている。認識・理解は、図4の状態遷移図形認識手法^{9), 13)}のように、対象メディアの認識レベルの上昇が陽な形で表わされるように実行する。(図4はボトムアップルールとトップダウンルールを併用した認識状態の変化の条件を Prolog により記述したもので、それ自身汎用性のある認識方式の例である)。データベースとして形成されるモデルは、この認識システムで実現されたものであれば、それが中間的、不完全な認識レベルであっても許容する。

すなわち、図4でいえば、一般には最上位の状態が望ましい認識水準であるが、その途中の中間状態で認識がとどまっても、画像メディアのそのオブジェクト部分については、中間状態の記述を行い、データモデルとして利用するわけである。

こうして設定された画像データモデルを利用するシステムは、この中間的かつ不十分かもしれない認識水準を前提にして、各種の利用演算を記述する、つまりいわば「それなりに活用する」ことになる。

この枠組により、各種のスポーツシーンを対象として静止画像の検索システムを構成し、約70枚の種々の画像を対象に実験を行った。このシステムでは、各画像キーワードの抽出は、図4の形の画像理解により完全に自動的に行われる。対象理解の不十分さを利用目的が達成されればよいという「それなり論理」の考え方でカバーするために、この状態遷移型の画像理解シナリオを検索操作システム側でも共有する形としている。ユーザは、たとえば「サッカー」等の高位の概念で与えられた内容検索コマンドを入力する。しかし一般にはこの水準のキーワードは抽出されていない。この場合、システムは自動認識が実現できているより低位の概念をも使って、換言すれば、画像理解シナリオを認識水準が上がっていく一種のシソーランスであると考えて、知的に実行することになる。

図5に、ユーザが「サッカーのシーン」というコマン



—— トップダウンプロセス
 ---- ボトムアッププロセス
 図4 状態遷移図型認識モデル



図5 「それなり論理」による画像検索例

ドを入れて検索をかけた出力例を示している。画像から自動抽出されて得られたキーワードには「サッカー」という高水準なものがないため、システムは認識シーランスを低レベルへと下がっていった照合をとっている。その結果、所望の3枚のサッカーシーンとごみとしての1枚のラグビーシーンが得られている。こうして「それなり論理」により、いわば70-80点の性能ではあるが完全自動の検索システムが実現されたわけである。画像理解部が今後充実すれば、より高い性能と段階的に発展していける「成長性」のあるシステムでもある。

5. む す び

以上、マルチメディアシステムについて情報アクセス

路の方法を中心に述べた。これ以外に既述のインタフェースや汎用システム化（プラットフォーム）等の問題も重要であり、統合化されたシステムが新たな情報利用の世界を開いていくものと期待している。

(1992年9月1日受理)

参 考 文 献

- 1) Lippman, A. and Bender, W.: "News and Movies in the 50 Megabit Living Room", Globecom, IEEE, Tokyo, Japan (Nov. 1987)
- 2) Mackay, E. W. and Davenport, G.: "Virtual Video Editing in Interactive Multimedia Applications" communications of the ACM, Vol. 32, No. 7, pp. 802-810, July, 1989
- 3) 上田博唯: "イクタラクティブな動画編集方式の提案", 信学技報, IE90-6 (1990)
- 4) 長坂晃明, 田中譲: "カラービデオ映像における自動索引付け法と物体探索法", 情報学論, Vol. 33, No. 4, pp. 543-549 (1992)
- 5) 坂内正夫: "これからの画像・図形情報データベース化", 生産研究, 42, 4, pp. 227-234 (1990)
- 6) 坂内正夫, 大沢裕: "画像データベース", 昭晃堂 (1987)
- 7) 坂内正夫: "地図データベース", 情報処理, 27, 10 (1986)
- 8) M. Sakauchi: "Two Interface in Image Database Systems", Proc. of MIV-89, pp. 22-27 (1988)
- 9) S. Satoh, M. Sakauchi, et al.: "Drawing Image Understanding Framework using State Transition Models", Proc. 10th ICPR, pp. 491-495 (1990)
- 10) 坂内正夫, 大沢裕: "画像データベースにおけるデータ表現・管理方式", 信学論, J68-D, 4 (1985)
- 11) 坂内正夫: "画像検索技術", 信学誌, 71, 9 (1988)
- 12) 増永: "マルチメディアデータベース総論", 情報処理, 28, 6 (1987)
- 13) 坂内正夫, 佐藤真一: "画像データベースにおけるモデル形成", 信学誌, J74-DI, (新しいデータベース特集号招待論文) (1991)
- 14) M. Sakauchi and J. Yamane: "Realization of fully automated keyword extraction in Image Database Systems", SPIE's 1992 International Symposium on Optical Applied Engineering (1992)
- 15) 坂内正夫: "マルチメディアデータベース", テレビジョン誌, 45, 8, pp. 931-937 (1991)
- 16) Y. Gong and M. Sakauchi: "A Method for color Moving Image Classification Using the Color and Motion Features", Proc. of ICARCV '92 (1992)
- 17) 山根, 坂内: "Implicit Link によるハイパーメディアのフレンドリ化", 情報学第44回全大, IC-10 (1992)