

マルチメディアデータの内容検索についての一考察

宝珍輝尚

NTT情報通信処理研究所

マルチメディアデータに対する内容検索のための内容表現モデルおよび内容検索レベルを提案する。まず、各メディアデータの計算機表現についてまとめ、マルチメディアデータの対象分野全般に適用可能な内容表現モデルCORMを提案する。本モデルはマルチメディアデータの持つ構造や意味する概念をも表現可能とするものである。次に、マルチメディアデータの内容検索レベルとして、信号レベル、特徴レベル、構造レベル、論理レベル、概念レベルを提案し、各内容検索レベルでの等価性について述べる。最後に、マルチメディアデータ検索システムの実現方法について議論する。

A Study on Multimedia Data Content Retrieval

Teruhisa HOUCHIN

NTT Communications and Information Processing Laboratories

1-2356 Take, Yokosuka-Shi, Kanagawa 238-03, Japan

Content representation model and content retrieval levels for multimedia data are suggested. First, computational representations for media data are summarized. Second, the content representation model CORM is suggested. Its purpose is to describe the structural and conceptual information that multimedia data represent. Next, content retrieval levels are proposed. They are signal-, feature-, structure-, logic-, and concept-levels. Equality at each level is also described. And last, we discuss about the implementation of a multimedia data retrieval system.

1. はじめに

近年、格納媒体が多様化、低価格化⁽¹³⁾し、マルチメディアデータの検索、表示に関する研究が盛んに行われている⁽¹⁻¹²⁾。

ここで、マルチメディアデータの検索においては十分に検索者の意図が反映され、しかも、(不必要なデータは取り出さないという意味でも) 所望のマルチメディアデータのみが検索されることが望ましい。しかし、現状のマルチメディアデータの検索は、マルチメディアデータに付与された単なるキーワードに基づく条件検索にすぎず⁽¹⁴⁾、検索者の意図を十分反映できるものとは言い難い。一方、マルチメディアデータの一つである画像データについて、概略図やスケッチ画像等を用いた検索⁽¹⁻³⁾等の研究が行われているが、これらは画像であるという特性や対象分野の特徴を使用しているものが多く、マルチメディアデータ全般に適用できるものはあまりない。

本稿では、マルチメディアデータ全般に対する内容検索のための内容表現モデルCORMおよび内容検索レベルの提案を行い、各内容検索レベルでの等価性について述べる。CORMはマルチメディアデータの持つ構造や意味する概念をも表現可能とするものであり、内容検索レベルは情報検索、画像検索で提案されている検索レベルを包含したものである。

以下、2. でマルチメディアデータについてまとめ、

3. で内容表現モデルを提案し、4. で内容検索レベルと該レベルでの等価性について述べる。最後に、5. で試作について議論する。

2. マルチメディアデータ

中野はメディアを表現形態、容器、運搬形態、意味内容に分類し、マルチメディアを情報の表現形態(音声、文書、画像等)が多種混在しているという意味で使用している⁽¹⁵⁾。本稿でもこの意味でマルチメディアという語を使用する。

まず、マルチメディアデータの内容について議論可能とするため、各メディアデータの計算機上での表現についてまとめる。マルチメディアデータがデータの集まりを総称していることに着目して、各メディアデータの計算機上での表現を、最小単位、表現単位(最小単位の集まりで、何らかの意味のある単位)、全体表現(表現単位のおつまり)の3段階でとらえる。各メディアデータで対応するものをまとめて表1に示す。香り、味、触覚については、テキスト、図形、画像、音声と対応するものを示したが、これらは計算機で扱えないため以下では議論の対象外とする。

次に、各メディアデータの計算機内部表現、内部表現から抽出する特徴量の例を、各々表2と表3に示す。

表1 メディアの最小単位、表現単位と全体表現

通称	最小単位	表現単位(データ表現形式)	全体表現例(表現形式)
テキスト	数値・文字	テキスト(1次元配列)	動テキスト(時間配列)、レイアウトテキスト(空間配列)等
図形	図形素 ^{*1}	図形(混合)	動図形(時間配列)、レイアウト図(空間配列)等
音声	音素	音(混合)	音声(時間配列)、ステレオ音(空間配列)等
画像	画素	画像(2次元配列)	動画像(時間配列)、レイアウト画像(空間配列)等
香り	香り素 ^{*3}	香り(混合)	香り列(時間配列) ^{*3} 等
味	味素 ^{*3}	味(混合)	味列(時間配列) ^{*3} 等
触覚	触素 ^{*2}	触感(混合)	触感列(時間配列) ^{*3} 等

*1: 図形素は、ベクトル表現データを指す。 *2: 皮膚上の温点、冷点、痛点等に対応する感覚

*3: テキスト、図形、画像、音声との対応するものとして示したが、科学的な根拠については未調査

表2 メディアデータの内部表現の例

メディア種別	属性	データ型
テキスト	文字コード	数値
	フォント	文字列
	文字高	数値
	文字幅	数値
	文字間隔	数値
図形	図形コード	数値
	座標点	数値配列
	線の太さ	数値
	色(RGB)	数値
	線種(線種名)	文字列
音声	音素	数値
画像	画素	数値

表3 メディアデータの特徴量の例

メディア種別	特徴	データ型
テキスト	平均文字数	数値
図形	濃淡	数値
	周波数特性	数値配列
	音高	数値
音声	音色	数値
	長さ	数値
	周波数特性	数値配列
	色調	数値
	明暗	数値
画像	濃淡	数値
	大きさ	数値配列
	周波数特性	数値配列

3. 内容表現モデルCORM

従来のマルチメディアデータの内容表現モデルについて述べ、内容表現モデルCORM (Content Representation Model for multimedia data)を提案する。

3. 1 内容表現モデル

(1) 構造に着目したモデル

マルチメディアデータを文書に限定すると、文書についてはODA⁽¹⁶⁾によってその内容の表現が標準化されている。ODAでは、論理的な構造とレイアウト的な構造での文書の取り扱いを可能としている。

一方、音声や動画といった実時間性の必要なデータに対する表現では、それらのデータの同期や時間的な順序関係の表現が問題となっている⁽⁵⁻¹¹⁾。表現方法には、木構造で時間情報を管理する方法⁽⁵⁻⁷⁾、時間管理にPetri netを使用する方法^(8,9)、トリガによる方法⁽¹⁰⁾等が提案されている。

一般に、マルチメディアデータは人間の認識過程を経る前のデータであり、認識者によって認識の仕方は様々に異なる。従って、文書に対するODAのような同一マルチメディアデータに対して様々な構造に従って該データを操作する機構が必要である。また、マルチメディアデータの操作において多種のメディアデータの同期をとった出力は不可欠である。

筆者は構造的なデータ定義に基づいた視点を提案している⁽²¹⁾。この視点は、同一データに対する一以上の構造化データ定義の実現に着目したものであり、実時間性の必要なデータに対する同期、時間的順序関係の表現は考慮していない。

(2) 特徴、意味に着目したモデル

マルチメディアデータを人間の認識に近い形で検索したいという要求は常にある。この要求に応えるため、例えば、画像データに対して、画像データの形状に着目した類似検索^(2,3)、主観に基づいた検索⁽⁴⁾等の画像データの特徴量を使用した検索の研究が行われている。一方、画像の表現する意味をERモデル等を用いて表現する研究があり、記述の汎用性は高い⁽¹⁾。

筆者は、後者のアプローチを採り、テキスト、図形、画像等に対して共通的に意味内容を扱うための意味内容表現モデル⁽¹²⁾を提案している。これは、意味内容を

を概念表現と構造表現に分けてグラフ表現し、各々の対応をつけて意味内容を表現するモデルである。

3. 2 CORM

3. 2. 1 概説

(1) 従来のモデルとの関係

筆者は、マルチメディアデータの内容を表現するモデルとして構造的なデータ定義による表現と意味内容表現モデルを提案している⁽¹²⁾。しかし、マルチメディアデータ検索システムを構築するという観点からは、構造的なデータ定義による表現と意味内容表現モデルが同様な方法で定義できることが望ましい。そこで、既に提案している構造的なデータ定義による表現で時間的な関係を表現可能とし、これと、意味内容表現モデルを統合した内容表現モデルCORMを提案する。CORMでは、強化した構造的なデータ定義による表現を表層表現と呼ぶ。また、意味内容は概念表現と構造表現を用いて表現する。表層表現、構造表現、概念表現の定義は、後で述べるように、同様な機構で定義可能なものである。

(2) 概要

CORMで表現する内容の表現を内容表現と呼ぶ。内容表現は、マルチメディアデータの全体表現が表現単位によってどのように構成されるかを表す表層表現、表現単位毎に内在する構造と意味を各々表す構造表現と概念表現およびそれらの間の対応によって表される。

表層表現は、データの時間的、空間的な構造を表現可能とする。表層表現はODAのスーパーセットとなることを考えている。すなわち、1以上の構造的な視点に従って同一のマルチメディアデータを操作可能とする。また、実時間性の必要なデータの時間的な関係も表現する。時間的な関係の表現には、木構造で時間情報を管理する方法⁽⁵⁾を採用した。これらの表現の例を図1に示す。図1は、朗読された物語を表現したもので、10分という物理的な単位と章という論理的な単位で「物語」を操作可能としている。物語を構成する「内容」は、時間的な順序関係で関係付けられた構成要素として表されている。図1では示していないが、表層表現はマルチメディアデータ実体と対応付けられる。

構造表現、概念表現は、各々、表現単位毎に内在する構造と意味を表す。この表現のために、I F Oモ

ル⁽²³⁾に格支配関係⁽¹⁷⁾，同義語関係⁽²⁰⁾ならびに知識表現で使用される実体化関係(クラス-インスタンス関係)を追加した、CORMのための世界表現モデル WISDOM(World Interrelation representation model for cOrM)とWISDOMによって表される一種の意味ネットワークである世界表現を導入する。概念表現は世界表現の一部分のコピーである。ただし、世界表現におけるISA関係，実体化関係の上位の要素を下位の要素で置換することを許す。構造表現は、世界表現の一部分を構造定義に利用したデータ定義をスキーマとするデータ実体である。ここで、世界表現の一部とは、is-part-of関係，集約関係(各々，IFOモデルではcross-vertex, star-vertexに対応)を下方にたどってできる木を示す。構造表現の各点は、メディアデータの表現値や特徴値を格納可能なフレームで表現する。構造表現のスキーマとして、その木のリーフノードと表層表現においてマルチメディアデータ実体と対応つけられた点とをis-part-of関係，集約関係を使用して関係付けるものも考える。

世界表現と概念表現，構造表現の例を図2に示す。概念表現は、世界表現の「犬」のコピーである。構造表現は、世界表現の「文」以下ならびに「身体」以下を構造定義に利用したスキーマに対するデータ実体である。各々、メディアデータの記述可能な「文字列」，「曲線」という属性(スロット)を持つ。構造表現のリーフノードはデータと(実際は、表層表現のノードを介して間接的に)対応付けられる。ここでは、文字列や図形と対応付けている。また、構造表現のノードはその意味を表すために概念表現のノードと対応付けられる。これにより、構造表現上は異なるが同一の意味を持つマルチメディアデータの管理を実現する。ここでは、文字列「次郎」と右の図形が共に同一物を示し、それが「犬」であることを表している。

以上説明した表層表現，構造表現，概念表現によって構成される内容表現の概要を図3に示す。表層表現，構造表現，概念表現は一般にグラフで表現され、対応する互いのノード間が連結されて表現される。

3. 2. 2 定義

まず、WISDOMを定義する。

[定義1] (世界素) 世界素は事柄概念，事象概念

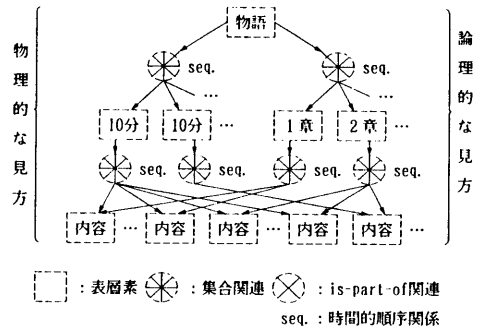


図1 表層表現の例

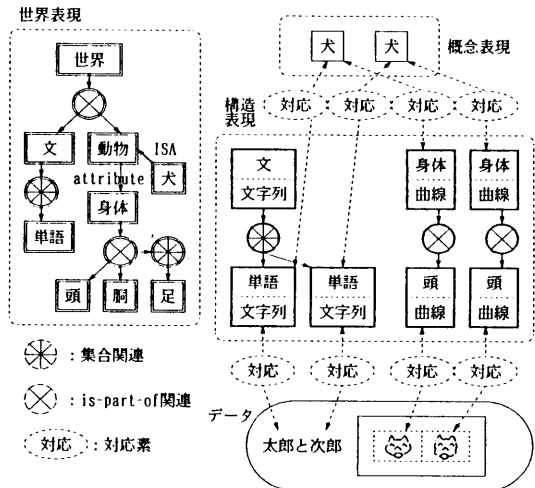


図2 世界表現と概念表現，構造表現の例

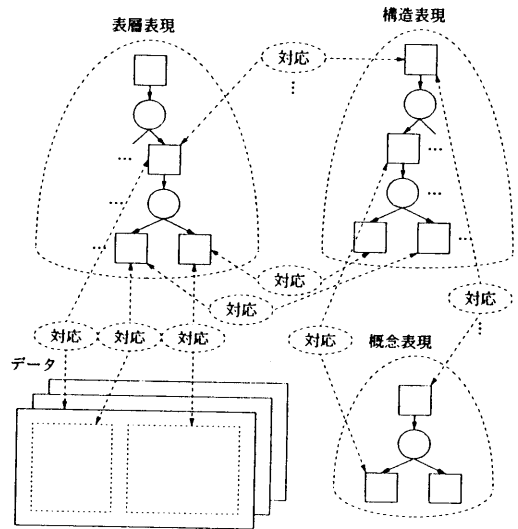


図3 内容表現の概念図

を表現する点である。事柄概念は名詞で表現される事物、事象概念は動詞で表現される動作である。

【定義2】（関係素） 関係素は、①格関係、②集合関係、③is-part-of関係、④関連を表現する点である。

【定義3】（関係枝） 関係枝は世界素間の⑤is-a関係、⑥属性関係、⑦実体化関係、⑧同義語関係を表す枝、または、世界素と関係素間を連結する枝である。

次に、準備のためにいくつかの定義を行う。

【定義4】（上位点、下位点） is-a関係、実体化関係に対する上位（下位）点とは、該関係を表す枝の終点（始点）を指す。

【定義5】（親、子） 関係素の親（子）とは、該関係素が終点（始点）となっている枝の始点（終点）を指す。

【定義6】（top-down存在制約） top-down存在制約とは、親の消去によってすべての子が消去され、親が存在しなければ子が挿入できないという制約である。

【定義7】（bottom-up存在制約） bottom-up存在制約とは、すべての子の消去によって親が消去され、子が存在しなければ親が挿入できないという制約である。

これらを使用して上記の①～⑧の関係を定義する。

【定義8】（格関係） 格関係は概念素間の格支配関係を表現する。グラベルの例⁽¹⁷⁾を表4に示す。

表4 グラベルの例

グラベル	意味	用例	グラベル	意味	用例
SUB	主体	～が	SOU	始状態	～から
OBJ	対象	～を	GOA	終状態	～へ
REC	受け手	～に	PRE	陳述	～である
PAR	相手	～と	ATT	属性	欠ける
MAN	方式	～で	TOO	手段	～で

【定義9】（is-a関係） is-a関係は専化または汎化を表現する。is-a関係は一上位点と一下位点を持つ。専化は、下位点が上位点を特殊化したものであることを表し、top-down存在制約を満たす。汎化は、上位点が下位点を一般化したものであることを表し、bottom-up存在制約を満たす。

【定義10】（is-part-of関係） is-part-of関係は集約関係または包含関係を表現する。is-part-of関係は一親と一種以上の子を持つ。集約関係も包含関係も、親が一種以上の子の集まりで構成されることを表現する。集約関係はbottom-up存在制約を満たし、包含関係はtop-down存在制約を満たす。

【定義11】（集合関係） 親が同種の子の集合（セ

ット）から構成されることを表現する。

【定義12】（属性関係） 属性関係を表す枝の終点が始点の持つ属性であることを表現する。

【定義13】（実体化関係） 下位点が上位点を実体化したものであることを表現する。

【定義14】（同義語関係） 同義語関係を表す枝の終点が始点の同義語⁽²⁰⁾であることを表現する。

【定義15】（関連） 上記以外の関係を表現する。

例えば、空間的な位置関係（①方向と②包含、接触、分離、交差）や時間的な関係（同一時刻関係、時間的前後関係）等を表現する。

次に、WISDOMを使用して表現される世界表現の定義を行う。

【定義16】（世界表現） 世界表現は、世界素および関係素を要素とする集合Vと関係枝の集合Eで表現される有向連結グラフ（V, E）である。

次に、CORMを定義する。CORMは内容表現を規定するモデルである。

【定義17】（内容表現） 内容表現は、①表層表現、②構造表現、③概念表現、④対応リンクから構成される有向連結グラフである。

【定義18】（表層表現） 表層表現は表層素、表層関係素を点とする有向グラフである。表層素は、(a)メディアデータの表現単位内の最小単位の集まりを指す点、(b)(a)を集約する格子の点を表し、フレームで表現される。表層関係素は表層素間の①is-part-of関係、②集合関係、③関連、④is-a関係を表すフレームである。

【定義19】（構造表現） 構造表現は構造素、構造関係素を点とする有向グラフである。構造素と構造関係素は、世界表現の一部を利用したスキーマに対するデータ実体である。構造素は同一メディアデータ集合を集約化する格子の点を表すフレームである。構造関係素は構造素間の①is-part-of関係、②集合関係、③関連、④is-a関係を表現する点またはフレームである。

【定義20】（概念表現） 概念表現は概念素、概念関係素を点とする有向グラフである。概念素と概念関係素は、世界表現の世界素の一部分を複写した点である。概念関係素は、概念素間の①格関係、②is-a関係、③is-part-of関係、④集合関係、⑤関連、⑥属性関係を表現する点である。

【定義2.1】(対応リンク) 対応リンクは、表層表現、構造表現、概念表現中の点集合と対応素集合で構成される2部グラフである。対応素は、概念-構造対応素、表層-構造対応素、表層-データ対応素である。

3. 2. 3 内容表現の等価性

CORMで表された内容表現は一般にグラフとなる。ここでは内容を表現する2グラフの等価性を定義する。

まず、点の等価性を定義する。点には単なる値を持つものと1以上の属性-値組を持つものがある。

【定義2.2】(点の完全等価性) 2点A, Bが単なる値をもつ場合、各々の値a, bが等しい($a = b$)ときAとBは完全等価であるといい、 $A =_c B$ と表記する。また、2点A, Bがn属性-値組を持つ場合は、両属性名の集合が等しく、同一属性に対する各々の値 a_1, b_1 が等しいとき、AとBは完全等価であるといい、 $A =_c B$ と表記する。

次に、グラフの完全等価性について定義する。

【定義2.3】(グラフの完全包含性) 2グラフA, Bにおいて、グラフとしてAがBのサブグラフであり対応する点が完全等価であるとき、BはAを完全包含するといひ、 $A \subseteq_c B$ と表記する。

【定義2.4】(グラフの完全等価性) 2グラフA, Bにおいて、AとBが枝の向きを含めて同形であり、対応する点が完全等価であるとき、AとBは完全等価であるといひ、 $A =_c B$ と表記する。

この完全包含性、完全等価性を使用して、表層表現、構造表現、概念表現、内容表現の等価性を定義できる。

3. 2. 4 表現例

CORMの内容表現モデルとしての十分性を証明するのは困難である。ここでは、CORMを様々な対象に適用した例を示す。図中の記号は図2と同様である。

(1) 物語

朗読された物語の内容を表層表現のみで表現した例を図4に示す。図4では、表層表現のみで表現し、「物語」を物理的な単位(10分)と論理的な単位(章)で表現している。「物語」を構成する「10分」および「章」は各々格納されている順序が時間的な順序(seq.)であると表現している。「10分」および「章」を構成する「内容」についても同様である。さらに、各「内容」は、同期した「左音声」と「右音声」で構

成されると表現している。「左音声」、「右音声」は各々音声データの対応する区間を指す。

(2) 写真集

生け花写真集⁽¹⁹⁾の内容を表現した例を図5に示す。表層表現では、写真集は写真の集まりで、写真は要素の集まりから構成されると表現している。実際の写真データは表層表現の点「写真」と対応付けている。構造表現の図の例では、写真は一つの生け花から成り、生け花は花、草、器から構成される等と表現されている。ここで、表層表現の写真、要素と構造表現の写真、要素の各々の対応するものが関係付けられている。さらに、概念表現では、生け花が百合、ペアグラス、扇子等から構成されていると表現され、構造表現の対応する点と関連付けられている。

(3) マニュアル

マニュアルの内容を表現した例を図6に示す。図6のマニュアルは一つの「図」と一つの「文」から構成されている。「図」は二つの「表」、一つの「矢印」と「文字列」から構成され、これを構造表現で表している。また、「文」は複数の「単語」から構成されると構造表現で表している。概念表現では、これらの意味を表現する。すなわち、「行から行を条件で選択する」を表現している。さらに、概念表現と構造表現の対応する点に対応付けられている。

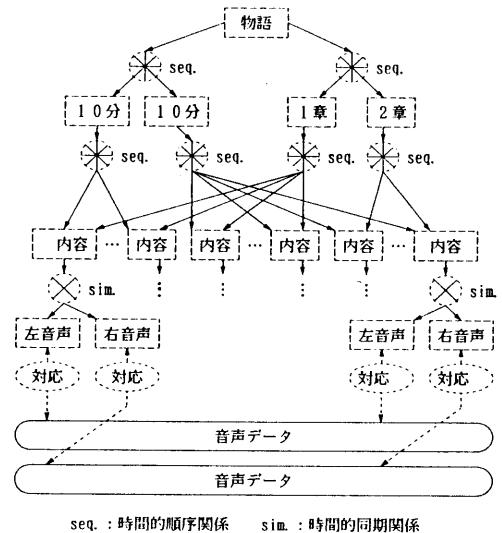


図4 物語の内容表現例

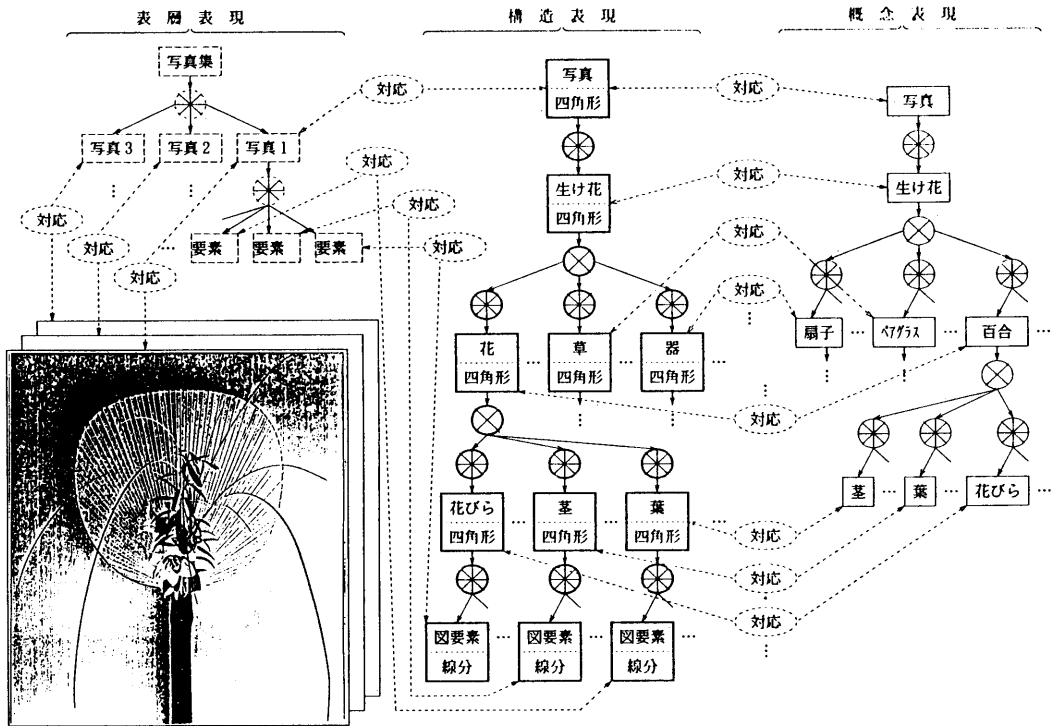


図5 写真集の内容表現例（写真は文献19中のもの）

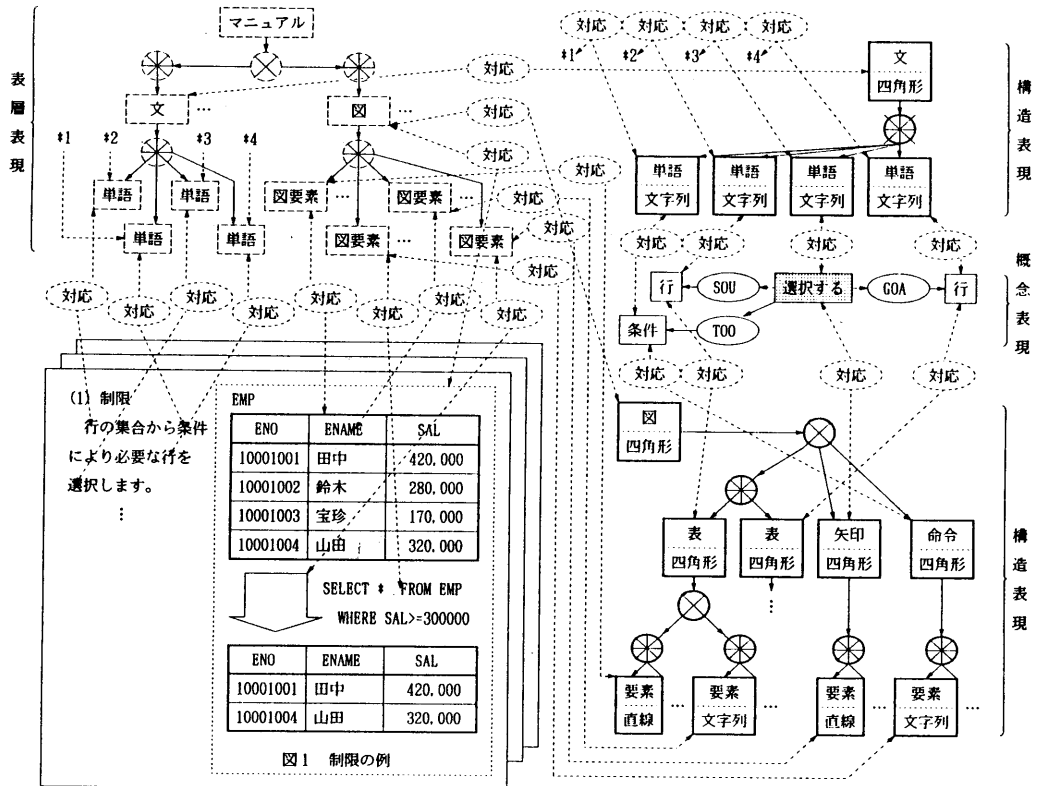


図6 マニュアルの内容表現例

4. 内容検索レベルと等価性

次に、3.の内容表現を利用したマルチメディアデータの検索方法について述べる。マルチメディアデータには様々な表現内容が内在しており、検索者の要求も多岐にわたる。従って、検索要求をレベル分けしなければ見通しの良い検索方法が得られない。そこで、まず、見通しの良い検索方法を得るために内容検索レベルを提案する。次に、ある検索レベルの検索要求の合致判定に必要な等価性について述べる。

4. 1 内容検索レベル

画像データに対する内容検索レベルとして、信号レベル、特徴レベル、意味レベルが提案されている⁽¹⁸⁾。これらは、各々、画像データを構成するデータそのものに基づいた等価性評価、画像データから抽出した特徴量に基づいた等価性評価、画像データが表現する面の構造や画の概念的な意味に基づいた等価性評価を行うことを指している⁽¹⁸⁾。

一方、従来のシソーラスを使用した情報検索では同義語を同一とみなす検索を行っており⁽²⁰⁾、これは論理レベルでの等価性の使用といえる。

本稿では、画像データや従来の数値・文字データを包含したものとしてマルチメディアデータを扱い、その内容検索を実現することを目的としている。このため、上記のいずれかの検索レベルでは不十分である。そこでこれらを包含した信号レベル、特徴レベル、構造レベル、論理レベル、概念レベルの5レベルの検索レベルを提案する。各内容検索レベルでの2データの等価性の概略を以下に示す。

- (a) 信号レベル：計算機内部表現が許容範囲のもの。
- (b) 特徴レベル：データから抽出した特徴量が許容範囲のもの。
- (c) 構造レベル：構造表現上、同一とみなされるもの。
- (d) 論理レベル：概念表現上、同義語関係を使用して同一のもの。
- (e) 概念レベル：概念表現上、同一とみなされるもの。

4. 2 信号レベルでの等価性

信号レベルでは、2データのメディア種別は同一でなければならない。信号レベルでの等価性は、2データそのものを使用して評価する。

(1) 準備

マルチメディアデータを構成するデータのマルチメディアデータ内における位置を得る関数を定義する。

[定義25] (位置関数) 表現単位F中の最小単位gに対して、位置関数pは表現形式の配列の次数kとgの配列上の位置 x_1, \dots, x_k からなる配列 (k, x_1, \dots, x_k) を返却する。また、全体表現中の表現単位Fに対しても、表現形式の配列の次数とFの配列上の位置を返却する。位置関数の返却する表現形式の配列の次数を構成数と呼ぶ。

[定義26] (位置の等価性) 位置関数が返却した位置 r_1, r_2 に対して、 r_1 と r_2 の構成数が等しい場合に限り、構成数が0のときは常に、構成数が1以上のときは対応する各々の値が等しいとき、 r_1 と r_2 は等しいといい、 $r_1=r_2$ と表記する。

(2) メディアデータの最小単位の等価性

[定義27] (最小単位の信号レベルでの等価性) 2最小単位a, bの計算機内部表現が許容範囲内であるとき、信号レベルで等価であるといい、 $a=s_{1g}b$ と表記する。

(3) メディアデータの表現単位の等価性

[定義28] (表現単位の信号レベルでの包含性) 2表現単位A, Bに対して、Aのすべての最小単位と等価な最小単位が必ずBにあり、これらを $a_1, \dots, a_n, b_1, \dots, b_n$ ($a_i \in A \wedge b_i \in B \wedge a_i=b_i$) とすると、 $n=m$ かつ $p(a_i)=p(b_i)$ であるとき、信号レベルでAはBに包含されるといい、 $A \subseteq_{s1g} B$ と表記する。

[定義29] (表現単位の信号レベルでの等価性) 2表現単位A, Bが $A \subseteq_{s1g} B$ かつ $B \subseteq_{s1g} A$ を満たすとき信号レベルでAとBは等価であるといい、 $A=s_{1g}B$ と表記する。

(4) メディアデータの全体表現の等価性

[定義30] (全体表現の信号レベルでの包含性) 2全体表現F, Gに対して、Fのすべての表現単位と等価な表現単位が必ずGにあり、これらを $f_1, \dots, f_n, g_1, \dots, g_n$ ($f_i \in F \wedge g_i \in G \wedge f_i=g_i$) とすると、 $n=m$ かつ構成数をkとするとある数列 x_1, \dots, x_k が存在して $p(f_i)=p(g_i)+(k, x_1, \dots, x_k)$ であるとき、信号レベルでFはGに包含されるといい、 $F \subseteq_{s1g} G$ と表記する。

[定義31] (全体表現の信号レベルでの等価性)

2 全体表現 F , G が $F \subseteq_{SIG} G$ かつ $G \subseteq_{SIG} F$ を満たすとき信号レベルで F と G は等価であるといい、 $F =_{SIG} G$ と表記する。

4. 3 その他のレベルでの等価性

信号レベル以外の検索レベルでは、内容表現を使用して等価性を判定する。特徴レベルでは表層表現、構造表現および対応リンクを使用し、構造レベルでは構造表現を使用する。論理レベルおよび概念レベルでは概念表現を使用する。これらの表現は一般にグラフとなる。以下ではこれらを表現グラフと呼ぶことにする。

ここで、これらのレベルでは表現グラフの等価性よりもむしろ表現グラフが部分的に合致した場合の合致度が重要である。

宇田川は、回路図の類似検索において回路図の類似度と相違度を使用して類似検索を行うことを提案している⁽²²⁾。提案された類似度、相違度の計算式は回路図に特化されたものであるが⁽²²⁾、この考え方は表現グラフの合致度の判定に利用可能である。以下に、表現グラフに対する類似度、相違度を示す。

要素（点および枝）の類似度 S_0 を 0 から F (F は任意の一定数) としたとき、相違度 D_0 を $D_0 = F - S_0$ とする。 F は利用者定義とするが、通常 1 で十分である。

表現グラフの類似度 S_G を以下に示す。

$$S_G = \sum_{i=0}^n C_{sn} \cdot S_{sn} + \sum_{i=0}^m \sum_{l=0}^k (C_{0i} \cdot S_{0i} + C_{tnl} \cdot S_{tnl}) \dots (1)$$

ここで、 n はソースノード数、 m は枝定義数、 $k(i)$ は枝定義番号 i の枝実体数であり、 C_{sn} 、 C_{0i} 、 C_{tnl} は各々、ソースノード、枝、終点の評価係数、 S_{sn} 、 S_{0i} 、 S_{tnl} は各々、ソースノード、枝、終点の類似度である。ここで、評価係数、類似度は枝定義番号 i によって異なり得る。評価係数は、利用者がどの要素を重要とするかを示すもので、 $1 \sim K$ (K は 1 以上の数) とする。

表現グラフの相違度 D_G は、式(1)の類似度の代わり相違度を使用して得られる。

5. 議論

現在、CORMに基づくマルチメディアデータ検索システムを試作中であり、以下の問題を検討している。

(1) 内容表現の入力負荷の軽減

本稿で述べた枠組みで検索システムを構築するには、

マルチメディアデータのみでなく、内容表現を入力しなければならない。完全な認識技術が確立していない現在、自動的な内容表現の生成は不可能である。しかし、なるべくデータ投入者の負荷軽減の図れる方法が望まれる。

負荷軽減方法として、テキストデータが存在する場合はテキストデータからキーワードを抽出し、該キーワードを他メディアデータの概念表現に使用することを考えている。また、構造表現に対しても半自動入力を考えている。すなわち、現状の画像認識で処理可能などころまで処理し、入力したいものと異なるときは人手で修正するという方法である。

対応リンクは、計算機で対応の候補を作成し該候補から選択させ、どの候補にもない場合、または、ある候補を修正することも可能とすることを考えている。

(2) 表現グラフの実現方法

表現グラフは文字通りグラフとなるが、グラフの計算機上での実現は簡単とはいえない。枝には複数の種類があり、種類毎に複数の枝実体が存在し得る。枝の実現方法により等価性、合致度の計算性能が左右されるので重要である。

(3) 複数の内容検索レベルでの検索時の最適化処理

複数の内容検索レベルで条件を指定された場合、どの検索レベルから評価し始めるかが問題である。すなわち、条件式評価の最適化の問題がある。これには、述語の評価順序を決定する問題と評価不要な述語の削除の問題がある。

(4) 合致度計算について

合致度計算において、定義とデータ実体が 1 : 多 (集合属性と呼ぶ) のとき、複数のデータ実体間の突合せ処理が必要となる。これは、そのまま行くと $O(n^2)$ の計算コストとなる。この計算コストを下げるには、例えばデータ実体の値によってソートしておく等の方法が考えられる。しかし、データ実体に対して単なる文字・数値のソートが常に可能とは限らず、対象分野により様々に異なる。この問題は、マルチメディアデータを扱おうとするとシステムが応用分野に依存してしまうという応用依存性の問題である。これに対しては、データ挿入手続きと突合せ手続きを利用者定義可能とし、拡張性を高めて対処することを考えている。

6. おわりに

マルチメディアデータの内容検索のための内容表現モデルCORMと内容検索レベルの提案を行った。

まず、マルチメディアデータの内容を議論するため、各メディアデータの計算機表現をまとめた。ここでは、各メディアデータの計算機表現を最小単位、表現単位、全体表現の3段階でとらえた。次に、内容表現モデルCORMについて述べた。本モデルは、マルチメディアデータの対象分野を広くカバーすることを目的としており、マルチメディアデータの持つ構造や意味する概念をも表現可能とするものである。次に、マルチメディアデータの内容検索レベルとして、信号レベル、特徴レベル、構造レベル、論理レベル、概念レベルを提案した。そして、信号レベルでのデータの等価性と他の検索レベルでの内容表現の合致度について述べた。最後に、試作についての議論を行った。

今後は、5.で議論した問題ならびに提案したCORMの評価が課題である。

謝辞 本研究を行うにあたり、熱心に議論して頂いた名古屋大学情報工学科渡辺豊英助教授ならびに上原祐介氏、竹平真則氏、都築香津生氏に深謝致します。また、日頃から議論して頂く田中豪主幹員、田中一敏主幹員をはじめとするNTT情報通信処理研究所の皆様感謝致します。

文 献

- (1) 坂内: "画像検索技術", 信学誌, Vol.71, No.9, Sept., 1988.
- (2) 加藤 他: "画像対話型商標・意匠データベースTRADEMARK", 信学論D-II, Vol. J72-D-II, No.4, pp535-544, 1989.
- (3) H. Kasahara: "Image Mnemonics - Indexing and Communication", proc. of Adv. Database sys. Sympo. '89, pp.293-300, 1989.
- (4) 栗田: "主観的類似度に適応した画像検索", 情処学論, Vol.31, No.2, 1990.
- (5) A. Poggio et al: "CCWS: A Computer-Based, Multimedia Information System", COMPUTER, 1985.
- (6) 打浪: "フィールド調査データ処理におけるマルチメディアデータベース", 情報処理, Vol.28, No.6, pp.773-783, 1987.
- (7) 森 他: "Tree構造を用いたメディアインタフェースの出力の時間制御", 1990年春季信学全大D-135, 1990.
- (8) T. D.C. Little et al: "Multimedia Object Models for Synchronization and Databases", Proc.of 6th intl. conf. on Data Eng., pp.20-27, 1990.
- (9) P. B. Berra et al: "Architecture for distributed multimedia database systems", Computer communications, Vol.13, No.4, pp.217-231, 1990.
- (10) 原田 他: "マルチメディアシナリオ記述のためのデータモデルについて", 第39回情処全大, 2J-5, 1989.
- (11) 鈴木 他: "動画・音声対応マルチメディアリアルタイムファイルシステム(3) AV編集用ユーザインタフェース", 第39回情処全大2J-8, 1989.
- (12) 宝珍: "マルチメディアデータに対する視点実現についての一考察", 信学技報DE89-7, 1989.
- (13) 石田: "マルチメディア化にいくらかかるか", bit, Vol.22, No.4, pp.373-381, 1990.
- (14) 石黒 他: "情報システムへのキーワード自動抽出機能の適用方法について", 第39回情処全大2J-8, 1989.
- (15) 中野: "マルチメディア処理の現状と将来展望", 電子工業月報, Vol.31, No.9, 1989.
- (16) 佐藤他: "オフィス文書の標準化と文書データベースの研究動向", 情報処理, Vol.28, No.6, pp.710-720, 1987.
- (17) 坂本: "格構造を中心とした用言と付属語辞書", 情処学自然言語処理研資, 38-8, 1983.
- (18) 藤村 他: "画像データの内容検索の諸問題", 第36回情処全大 2W-4, 1988.
- (19) 別冊家庭画報: "今、自然のままに花を生ける", 世界文化社, 1987.
- (20) 中井, 笹森: "情報検索システム", 日本経営出版会, 1971.
- (21) 宝珍: "データフォーマット可変な流動的データ型を用いたマルチメディアデータに対する視点の実現", 論D-I, Vol. J73-D-I, No.6, pp.580-588, 1990.
- (22) 宇田川 他: "オブジェクト指向リレーショナルデータベースADAMにおける類似検索機能", 情処学アドバンスデータベースシステム・シンポジウム, 1988.
- (23) S. Abiteboul and R. Hull: "IFO: A Formal Semantic Database Model", ACM TODS, Vol.12, No.4, pp.525-565, 1987.