

動画像データの内容記述と その構造化処理に関する検討

麻植 周 前田 哲郎 福田 慶郎 上原 邦昭

神戸大学 工学部 情報知能工学科
meg@jedi.seg.kobe-u.ac.jp

内容梗概

動画像データを利用するためには、動画像が持つ時間、動作、ストーリーなどの内容を記述でき、かつ適度に制限された形式のデータモデルが必要である。本研究では、内容記述として概念依存表現を用いたシステム HierOGlyph と属性値リンク法を用いたシステム VRS の実験を通して、動画像内容記述形式の検討を行なう。また、HierOGlyph において、NA と呼ばれるネットワーク構造化手法を用いて、グラフ表現された動画像内容記述データを構造化し、動画像のサブストーリーに当たる部分構造を発見する手法を提案する。最後に、これらの過程で生じた問題を挙げ、各問題点を検討する。

Content-Based Video Description Model and Its Structurization of Video Data

Meguru Oe Tetsuro Maeda Yoshio Fukuda Kuniaki Uehara

Department of Computer and Systems Engineering,
Faculty of Engineering, Kobe University
meg@jedi.seg.kobe-u.ac.jp

Abstract

In this paper, we propose two different approaches for content-based video description model. Content-based video description model consists of representations of time, action and story. One approach is that video description is assumed to be given just by structured representation (*i.e.* conceptual dependency representation). According to this idea, we implemented the system HierOGlyph which can organize video descriptions into hierarchical structure by NA (Network Algorithm). The other approach is that video description is assumed to be given by collections of attribute-value pairs. VRS is the alternative which can store and retrieve video description in the framework of database management system Postgres95. The retrieval of video description given by structured representation may become computationally intractable, although HierOGlyph can perform intelligent content-based retrieval compared with VRS.

1 はじめに

動画像のようなマルチメディアデータを利用するために、マルチメディアデータベースの研究が盛んに行なわれている [1]. マルチメディアデータベースでは、定型的な内容記述形式が定められていないために、ユーザは各々の主観と動画像の内容記述形式に基づいて、問合せを作成しなければならない。しかしながら、すべてのユーザの主観にあわせた内容記述を、予め用意しておくことは不可能である。そのため、動画像データを記述するには適度な制約が必要となる。逆に、動画像データの記述を必要以上に制限すると、問合せを限定してしまい、ユーザに対して問合せの作成コストをかけることになる。このような問題に適応する動画像内容記述形式にはいくつか考えられるが、ユーザにとっての利便性とデータベース構築コストなどを考慮した上で、本研究では、まず Schank の提案した自然言語の意味表現形式の一つである概念依存 (Conceptual Dependency, 以下 CD と略す) 表現を使用して [2], 動画像データを構造化し、その動画像を検索するシステム HierOGlyph [3] を開発した。

また、HierOGlyph は計算コストが高いという問題点があるため、検索の計算コストが低く、CD 表現と等価性を持つような属性値表現による動作記述及び場面検索が可能な、属性値リンク法を用いた動画像検索システム VRS [4] を並行して構築した。

両者の実験を通して、動作や意味、ストーリーを表現することのできる動画像内容記述とその構造化について検討する。さらに、その構造化処理及び検索処理で生じたいくつかの問題点について検討する。

2 HierOGlyph における動画像内容記述とその構造化

動画像の内容記述を行なう場合、現在の技術では人間が画像を見てその意味をキーワードや自然言語という形で付加するしかない。この付加された画像の意味が検索のインデックスに対応し、動画像データの内容検索において重要な役割を果たす。しかしながら、時系列を伴う動画像はデータ間に時間関係や因果関係が存在している。これらの関係を記述するためには、グラフ表現のような要素間の関係を記述することができる表現力の高い形式が適して

いる。

HierOGlyph では、まず動画像の内容を Schank の提案した自然言語の意味表現形式の一つである CD 表現 (概念依存表現) を用いて記述する。CD 表現は、次節で述べる様に、非常に高い記述の制約を持っており、この形式を利用することで記述を限定し、検索や構造化処理を比較的容易に行なう事ができると考えられる。

2.1 CD 表現 (概念依存表現)

CD 表現は、動的な概念を ACTION, ACTOR, OBJECT などの要素で表現しており、特に文章中の動詞に対応する ACTION を、所有権の移動を表す PTRANS (give, take, ...) や位置の移動を表す PTRANS (go, come, ...) などの 11 種類のプリミティブと呼ばれる構成要素に分類している。また、静的な概念は対象物と状態により表現されている。なお、CD 表現では概念間の関係として 9 種類の因果タイプが定義されているが、本研究では、after, before といった Allen による時間区間に関する 13 種の関係を追加している [5]. CD 表現による記述例を図 1 に示す。

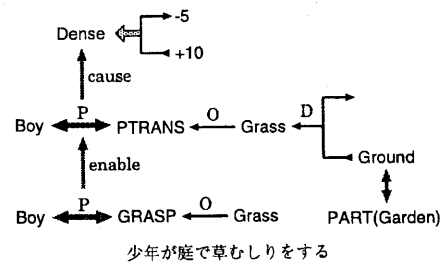


図 1: Schank の CD 表現による記述例

CD 表現を用いる方法の最大の長所は、プリミティブのみによって動作を記述できることである。これによって、データ構築や検索の際に必要な知識を、ユーザ側でも計算機側でも限定することができる。CD 表現は、一般的には図 1 のようなネットワーク表現をとっているが、HierOGlyph では、枝分かれした辺を扱うことができないため、また節点数を減らして、計算コストを下げるために、行為や状態などの概念を節点で、概念間の関係を辺で表現

し、図2のようなラベル付き有向グラフへ変換している。

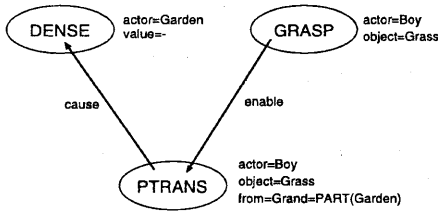


図 2: 本研究における CD 表現による記述例

2.2 HierOGlyph による構造化と検索

動画画は一般にストーリーを持っており、そのストーリーはいくつかの部分的なサブストーリーから構成され、そのサブストーリーもさらに分割することができる。このように、動画画は一つの木構造で表現することができる [1]。類似したストーリーを持つ動画画を複数集めれば、そこには部分グラフの階層構造が存在し、その類似した部分グラフは、ストーリー中の各サブストーリーやさらに分割された場面に当ると予想される。

HierOGlyph は、複数の動画画を CD 表現を用いて内容記述したグラフ集合を、事例ベース推論の分野で用いられる NA と呼ばれる階層構造を表すネットワークの構造化手法を用いて、グラフ集合の構造化ならびに検索を行なうことができる。

NA によって、ネットワークへ構造化されたグラフ集合における検索は次のように行なわれる。まず、入力されたグラフの各節点は、ネットワークの入力ノードから各節点ノードへ渡され、マッチする節点ノードのインスタンスとして格納される。さらに、各節点ノードはその子ノードへインスタンスを渡していく。そして、二つの親ノードに呼び出された子ノードは、インスタンス間の辺がそのノードの制約にマッチした場合には、親ノードからインスタンスを受け取り自分のインスタンスとして格納し、さらに自分の子ノードを呼び出す。このプロセスをインスタンスがグラフノードに到達するか、インスタンスがそれ以上子ノードに渡されなくなるまで繰り返す。このように、グラフ集合を構造化することで、共通部分に対するマッチングを一回で済ませ、効率的な事例検索を実現できる [6][7]。

しかしながら、このネットワークには、仮にグラフ集合と共通の部分グラフを入力されたグラフが持っていたとしても、ノードとして含まれない部分グラフにはマッチしないという問題点がある。したがって、ネットワークの構造によっては検索できない部分構造が存在する。また、ネットワークの構造は検索処理の計算量に大きな影響を与えており、最適なネットワークへの構造化が必要とされる。

NA によってネットワークへ構造化されたグラフ集合での検索例を図3に示す。図3の例では、まず Graph 1 と Graph 2 が入力された結果、Graph 1 と Graph 2 の部分グラフである internodes が生成される。これに対して Graph 3, Graph 4 の検索を行なう場合、Graph 3 では、完全に match する internodes が存在するので、match する internodes に対応する Graph Nodes の Graph 1 と Graph 2 に match する。これに対して、Graph 4 の場合、Graph 2 の部分グラフであるにも関わらず、該当する internodes が存在しないために、検索を行なうことができない。

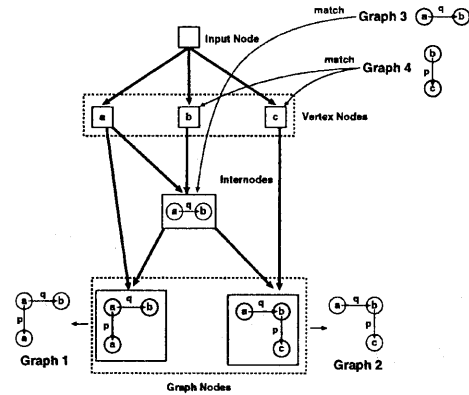


図 3: NA によるネットワーク化

効率的な検索を行なうためには、複数のグラフ間またはグラフ内部に共通する構造を捉えてグラフ集合を単純化する必要がある。そこで、HierOGlyph では、MDL (Minimum Description Length) 基準に従い、グラフ集合を最も単純に記述する部分構造の組合せを探索するアプローチをとっている。MDL 原理は、データから未知の情報源の確率モデルを推定するための選択原理の一つとして、情報理論の枠

組で提案されたものである [8]。MDL 基準における最良のモデルとは、データを説明するモデル自身の記述長とそのモデルを用いてデータを記述した際の記述長の和が最小となるようなモデルである。以上のような考え方に基づいて、部分グラフの組合せである階層構造を用いてグラフ集合を表現した際の記述長でその構造を MDL 基準により評価し、最も記述長を短くする部分グラフの組合せを最適な構造として選択するようにしている [9]。

2.3 構造化による階層的場面の構築

CD 表現にしたがって、動画像の内容をグラフ表現で記述し、ネットワークによる構造化を行なった。構造化の際に発見した部分グラフを用いて、元の動画像を記述した例を図 4 に示す。図 4 の枠で囲んでいる部分が発見された部分グラフを示している。図 4 から分かるように、一つの事例はいくつかの部分グラフにより表現され、その部分グラフもさらに小さな部分グラフにより表現されている。

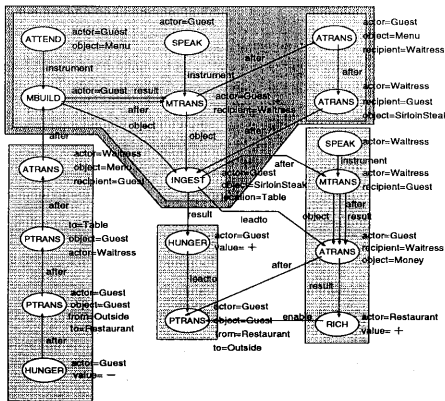


図 4: 結果の一例

2.2 節で述べたように、ストーリーはサブストーリーやさらに分割された場面の階層構造を持っておるが、このサブストーリーや場面が構造化によって得られた部分グラフに相当すると考えられる。また、部分グラフで表される場面やサブストーリーは、因果関係や時間関係を示す辺で結ばれているため、それぞれの関係が特定されているので、構造化によって得られた部分グラフの階層構造は、ストーリーを構成す

る場面やサブストーリーを表す概念の階層構造となっている。図 2 で示したように、HierOGlyph においては、節点であるプリミティブに actor や object などの属性を付加したグラフになっているので、場面やサブストーリーを表す概念は、それぞれの属性を取り除いた節点とその節点間を結ぶ辺のラベルによって作られる関係で表される。

現在の HierOGlyph には導入されていないが、このストーリーの概念構造を利用すれば、ストーリーを考慮した検索を行なうことができるようになると思われる。ストーリーを考慮した検索を行なうためには、動画像に対して以下のような 2 つの階層の内容記述を考える必要がある。

第 1 階層 CD 表現を用いた動画像全体の動作レベルの内容記述

第 2 階層 場面やサブストーリーの概念を節点とするネットワークグラフ

第 1 階層は、従来の HierOGlyph と同様に、CD 表現によって動画像全体を記述して得られるものである。さらに、第 1 階層の内容記述を構造化することで得られる階層構造において、場面やサブストーリーを表す部分グラフすべてに、その概念を適切に表現するような新たなラベルを付加し、その記述を部分グラフ自体のラベル、すなわち概念のラベルとする。このとき、概念を 1 つの節点とみなすと、それぞれの概念は因果関係や時間関係を示す辺で結ばれているので、第 2 階層を得ることができる。このような内容記述を行なった場合の第 2 階層の記述例を図 5 に示す。

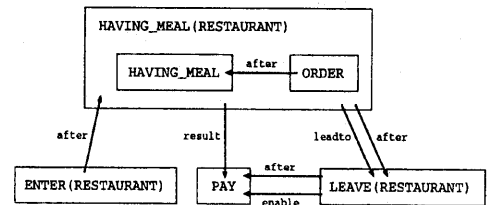


図 5: 第 2 階層の記述例

このような記述を行なうと、一連の動作の検索などの従来の検索に加えて、場面やサブストーリーの検索やストーリーを考慮した検索が可能になる。た

たとえば、図5で表されるストーリーを含む動画像の場合、HAVING_MEAL(RESTAURANT)と属性値 Hamburger の和集合を検索すれば、レストランでハンバーガーを食べるシーンの検索ができ、またPAYを含まない検索を行えば、レストランでの飲食で支払を忘れたストーリーを検索できるようになる。

2.4 グラフ検索および構造化のコスト

HierOGlyphにおいて、実際にCD表現で内容記述を付加した場合、図4の様なまとまりを1場面とすると、7分程度の動画像に対して、1場面平均9節点で約80場面、節点数総計約700個程度になった。このような動画像に対して、HierOGlyphを用いて構造化及びグラフ検索を行なったが、2つのグラフの部分的なマッチングはNP完全問題として知られており、実際に複数の動画像を記述した数百個のグラフ集合に対して構造化及び検索することは困難であった。また上記の例は、類似したストーリーのグラフ集合を用いた結果であり、さまざまな動画像内容記述を行なったグラフ集合に適用すると、上記の例ほどははっきりした場面構造を抽出することが困難であることが分かった。これは、主にネットワーク構造化の際に、計算コストを軽減する為に、中間層において、存在可能なすべての部分グラフを用いるのではなく、ランダムに選択した部分グラフにより、マッチングを行なっているために、場面を適切に抽出できなくなること、異なった内容の動画像を対象にすると、共通する場面に相当する部分グラフが非常に小さくなってしまふことが原因であると考えられる。

3 VRS における動画像内容記述

CD表現による動画像内容表現は表現力が高い反面、グラフマッチングを行わなければならないため、計算コストが高いという欠点があることが分かった。これに対して、直観的なキーワードをインデックスとして動画像の内容を記述しておけば、検索においては属性値マッチングとなり、計算コストを抑えることができる。VRSでは、属性値表現による属性値リンク法を用いて、上記のようなキーワードをインデックスとした動画像内容記述を実現したシステムである。

3.1 属性値リンク法

属性値リンク法では、動画像で表現されている内容は、location, actor, action, objectの4つの属性の属性値とそれらのリンクによって表現することができる。このように表現された動画像の例を図6に示す。図6の例は、「ある女性が縁側でしゃべりながら牛乳が載っているお盆を運んでいる」シーンを記述したものである。ここで、フレームとは各属性値が動画像内で存在している時間を、SGIムービーとよばれるシリコングラフィックス社固有の動画像記録方式に依存するフレーム番号によって表現したものである。リンクは、各属性値がリンクしている属性値のID番号を示している。これらのリンクは、それぞれの属性値間で相互に参照され、ストーリー上の各場面に相当するシーンを構成する。

ID	属性	属性値	フレーム	リンク
229	location	veranda	(1548 1582)	(230 231 232 233 234 235)
230	actor	woman	(1548 1582)	(231 232 233 234 235 229)
231	action	carry	(1548 1582)	(230 233 234 229)
232	action	speak	(1558 1582)	(230 229)
233	object	milk	(1548 1582)	(230 231 233 229)
234	object	tray	(1548 1582)	(230 231 232 229)

ある女性が縁側でしゃべりながら牛乳が載っているお盆を運んでいる

図6: 属性値リンク法による記述例

VRSにおいて、ある属性値に対して検索要求があった場合、その特定の属性値の存在時間のみでは、時間が短過ぎて意味をなさない場合が多い、たとえばspeakをキーにして検索した時に、speakという属性値が対応する部分の動画像だけを表示したのでは、ユーザには口を動かしているシーンにしか見えない。そこで、リンクを参照して、その属性値の属するシーンまたはlocation全体を表示すれば、ユーザは前後関係から判断して、「話す」部分の動画像を得ることができる。

このように、VRSでは、actorやaction, objectの関係を等価に扱って記述することで、個々の属性や属性値間の関係を1対1で記述する必要がなくなり、データ構築のコストや検索の計算量を軽減することができる。また、リンクを参照することで、断片的な動画像を表示するのではなく、ユーザにとって意味のある検索を可能にすることができる。

3.2 属性値とプリミティブの等価性

属性値リンク法では、図6に示すように action 属性の記述を通常の動詞を用いて行なったが、システムに表1のような action 属性に対する CD 表現の動作プリミティブへの変換辞書を組み込めば、記述形式が異なるだけで等価の表現として考えることができる。実際には、1つの action 属性が2つ以上のプリミティブに対応する場合の属性値間のリンクの扱いなどが未解決であるため、システムに実装してはいないが、今後の研究において、特に HierOGlyph と VRS の統合の際には考慮する必要がある。

属性値	プリミティブ
carry	PTRANS + PROPEL
speak	SPEAK

表 1: プリミティブへの変換辞書の例

CD 表現を用いる場合、表現力が高く、記述が限定されており、データ構築が容易である反面、検索や構造化の計算コストが高い。属性値リンク法を用いる場合、記述の制限がほとんどなく、表現が発散しやすいが、検索などの計算コストは比較的低い。そのため、より効率的なシステムを組む事を考慮するならば、双方の形式をうまく生かし、表現と計算コストとのトレードオフを取る必要がある。

3.3 VRS による検索

VRS は、図7に示すような各モジュール Query-Manager, Query-FrontEnd, MoviePlayer, DBMS および Video Server からなるシステムである。VRS ユーザは、問合せをキーワードの形で、Query-Manager に入力する。Query-FrontEnd は、Query-Manager から受けとったキーワードを SQL に変換し、DBMS (Postgres95 [10]) においてユーザの要求を満たすまでリンクをたどりながら、データの集合を検索する。その結果、DBMS はそのデータである属性値集合を返し、Query-FrontEnd は、映像要求を MoviePlayer へ送る。MoviePlayer は、ワークステーション SiliconGraphics Indy 上で動画像表示用のライブラリである IRIS Digital Media library を使用して、フレーム指定により Video Server から

ら得た動画像を、ユーザに対して出力する。

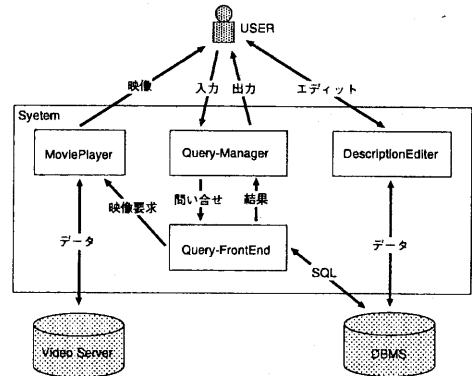


図 7: システム構成

実際の動画像に、属性値リンク法を用いて内容記述を付加した結果、7分程度の動画像に対して、シーンを表す location 属性の数が約 75 個、1シーン平均属性値数 13 個、全体で千個程度の属性値が必要であった。データ数で比較すると、CD 表現を用いた場合と大きな差はないが、1つ1つのデータの大きさが、グラフ表現と属性値表現では大きく異なることから、データ量を大幅に圧縮できたと言える。また、検索に要する時間も、HierOGlyph では実時間では困難であるのに対して、VRS を用いた場合、属性値表現であるために、計算コストが低く、複数の動画像に対しても、数秒以内に行なうことができた。しかしながら、VRS では、HierOGlyph で可能だった場面間の関係を記述することはできない。これは、場面間の関係を記述するリンクを設定していないためである。

4 動画像内容記述とその構造化における問題点

4.1 動作の記述

たとえば、「歩く」という動作を表現する際に、VRS では単に walk を用いるが、HierOGlyph の場合、「行為者が空間的に移動する」とも「足を交互に動かす事によって、行為者を空間的に移動させる」とも表現できる。このように、すべての動作の表現は、その表現の詳細さのレベルを持っている。

ある動画像中の動作の記述をする際には、この詳細さのレベルを揃える必要がある。しかし、明確な詳細さの尺度が存在していないために、CDによる記述においても、属性値による記述においても、具体的な解決方法は存在しない。

また、内容記述が人間の手によるために、記述を付加した人間個人の主観に依存するという問題も生じた。例えば、VRSにおいて、ある人物が移動している動画像を記述している場合、その動作が歩いているか走っているかは、動画像を見た人間の主観に依存してしまう。記述を付加した人間が「歩く」と記述した動作に対して、ユーザが「走る」という問合せを行なう場合、このマッチングには、「歩く」と「走る」がほぼ同じ概念を含んでいるという知識を用意しておく必要が生じるが、すべての動作の概念関係をすべて知識として用意することは不可能である。前述の例は、VRSを用いた時の例であるが、HierOGlyphを用いた場合にも同様の問題が生じる。

さらに、ある人物や物体が全く動作をしない場合の記述をどのように行なうかという問題も存在する。VRSでは、この問題に対して「動作していない」ことを表す記述を導入して、問題を回避している。実際には、動作していない人物や物体を検索することはほとんどないが、たとえば動画像全体にある人物がどの程度存在しているかといった問合せがあった場合、このような「動作していない」ことを表す記述が必要となる。

4.2 時間の記述

VRSにおいては、すべての属性値の存在時間をSGIムービーの形式に依存するフレーム番号によって記述することは既に述べたが、動画像の時間軸に沿ったそれぞれの属性値の時間関係、たとえば前後関係などは、陽に記述しなくとも、J. F. Allenなどによって研究されている時間論理を用いて、必要な場合ごとにフレーム番号から推論可能であることが分かっている[5]。しかしながら、動画像中の意味的時間関係やストーリー上の時間関係は、動画像の時間軸と一致しない場合が多い。たとえば、動画像中に過去の回想や空想のシーンがいくつも挿入されている場合などには、表面上の時間軸とストーリー上の時間軸は一致しない。

この問題を解決する方法の1つとして、それぞれ

の属性値に対して、個別に意味的時間関係を記述する方法が考えられるが、その記述はすべての属性値について行なわなければならないが、現実的な方法ではない。また、登場人物の空想などの時間関係が存在しないシーンの扱いも問題となる。

別の方法として、動画像に沿った時間軸とは別に、意味的な時間軸を導入する方法が考えられるが、その場合、それぞれの属性値がどの程度の時間幅を持つか、あるいはそれぞれの属性値間の時間的距離はどの程度かといった基準がなく、動作の記述の場合と同様に、人間の主観の問題が関わってくる。

これに対して、HierOGlyphでは、行為や状態のみを各節点に対応付けたCD表現を用いているので、行為者や物体などの存在時間を記述せず、各節点間を13種の関係で結ぶことで時間関係を明確にしている。しかし、HierOGlyphにおける各時間関係でも、過去や想像などのシーンでは、人間の主観の問題から離れることはできない。

4.3 その他の問題

4.3.1 因果関係とストーリー

因果関係の問題とは、ある行動がそれより後のシーンの行動の原因や動機となる場合やある行動によってその後の行動が存在しなくなるような場合のシーンや行動の関係をすべて記述することが困難であるという問題である。たとえば、ある人物の行動がその後のすべてのシーンのすべての人物の行動に影響する場合などがそれに当たる。

この問題に対して、ストーリーの持つ階層構造に着目し、動画像中の行動の記述と別に用意したストーリーの階層構造との対応関係を用いて、因果関係を記述する方法が提案されているが[11]、この方法においてもストーリーに現れない因果関係を記述することはできない。たとえば会話のシーンで、Aが発話しているシーンについて記述する場合、その前後にBやCと同時に存在しているシーンがあったとしても、Aの発話がBに向けられたものか、Cに向けられたものか、あるいは両方に向けられたものかを動画像から特定することはできない。この問題は、記述の問題を離れ、認識の問題としてとらえても特定することは不可能であると考えられる。

また、画面効果として用いられるアップの扱いの問題も存在する。アップによって、人物や物体の判

別が不可能な場合、たとえば、ある人物の手のアップが挿入された場合、その前後のシーンから判断しなければならない。因果関係に関する問題すべてに当てはまるが、個別のシーンからは判断不能な属性やノードのラベルを記述すべきかどうかは、記述方法の問題よりも高次の問題である。

4.3.2 独自性

因果関係とは独立して、各属性値やノードのラベルの独自性に関する問題が存在する。たとえば、ある路地のシーンの場合、その路地が実際には大阪のある路地であったとしても、その直前のシーンに東京タワーが存在するシーンがあれば、東京のどこかの路地であると判断できてしまうことが独自性の問題である。この問題は、ある俳優が同じ動画内で別の役柄を演じているような場合に特に顕著になる。つまり、現在用いている方法では、実際の属性と異なっても、画像自体の属性はその記述に完全に依存してしまうことが、その問題の本質といえる。

本実験では、ユーザが一般的な背景知識として、人物関係や世界観を有していると仮定したため、独自性を考慮した記述を行っていないが、実際には、あるシーンの人物 A と別のシーンの人物 A が全くの同一人物であるという知識または記述がなければ、独自性を考慮した場合、二人の A は別の人物としてとらえなければならない。

5 おわりに

本研究では、グラフ表現と属性値表現による動画内容記述について検討した。HierOGlyph では、グラフ表現を用いるため、計算コストが高いという問題点がある反面、柔軟な表現が可能であり、場面間の関係を記述することが可能であることが特徴と言える。これに対して、VRS では属性値表現を用いることで、計算コストを下げ、実時間内での検索を可能としたが、属性が制限されているために、柔軟な表現が困難であり、また場面間の関係も記述できない。このように、両システムとも問題点を有しており、今後の研究では、計算コストと表現力の双方の問題点を解決するシステムを目指す。

また、本研究で扱わなかったが、せりふなどの音声も、動作やストーリーなどと並んで、動画像に必要

な表現である。音声についても、その内容を記述する適切な形式はなく、新たに研究が必要である。

参考文献

- [1] 美濃導彦: 知的映像メディア検索技術の動向, 人工知能学会誌, Vol. 11, No. 1, pp. 3-9 (1996).
- [2] Schank, R. C. and Riesbeck, C. K., 石崎俊監訳: 自然言語理解入門 — LISP で書いた 5 つの知的プログラム —, 総研出版, pp. 12-47 (1986).
- [3] 福田慶郎: グラフ表現によって記述された事例集合からの MDL 基準による階層構造の発見, 情報処理学会研究会報告, 96-AI-103, Vol. 96, No. 4 (1996).
- [4] 前田哲郎: 動画像情報の動作に着目した内容記述と検索方式, 平成 7 年度神戸大学大学院自然科学研究科修士論文 (1996).
- [5] Allen, J. F. and Ferguson, G.: Actions and Events in Interval Temporal Logic, Technical Report, 521, University of Rochester (1994).
- [6] Messmer, B. T. and Bunke, H.: A Network Based Approach to Exact and Inexact Graph Matching. Technical Report, IAM-93-021, University of Berne (1993).
- [7] Bunke, H. and Messmer, B. T.: Similarity Measures for Structured Representations, *Lecture Notes in Artificial Intelligence*, Vol. 837, pp. 106-118 (1994).
- [8] 山西健司: MDL 入門: 計算論的学習理論の立場から, 人工知能学会誌, Vol. 7, No. 3, pp. 435-442 (1992).
- [9] Cook, D. J. and Holder, L. B.: Substructure Discovery Using Minimum Description Length and Background Knowledge, *Journal of Artificial Intelligence Research*, Vol. 1, pp. 231-255 (1994).
- [10] Yun, A. and Chen, J.: The POSTGRES95 User Manual Version 1.0, the Regents of the University of California (1995).
- [11] 柴田正啓: 映像の内容記述モデルとその映像構造化への応用, 電子情報通信学会論文誌 D-II, Vol. J78-D-II, No. 5, pp. 754-764 (1995).