

ビデオハイパーメディアにおけるオブジェクト管理法の検討

4G-2 種茂 文之 吉田 忠城 片岡 良治 井上 潮

NTT 情報通信研究所

1 はじめに

ビデオハイパーメディアとは映像中の登場物体をノードとして利用したハイパーメディアであり、利用者は興味のある登場物体を映像上で指定してその関連情報を検索できる。ビデオハイパーメディアは多種多様な情報を含む映像をインデックスとして利用できるため、マルチメディア情報の有効な検索手段と考えられる [1]。

しかし、これまでに提案されたビデオハイパーメディアでは、登場物体とその関連情報の間には固定的なリンクが張られており、物体の表示状態の時間的な変化といった映像特有の情報がナビゲーションに生かされていない [2][3]。本稿では、映像中の登場物体を複合オブジェクトとして捉え、登場物体の表示状況に応じて、動的に適切な部品情報へナビゲーションする方式について検討する。

2 ビデオハイパーメディア

2.1 定義

ビデオハイパーメディアとは、映像中の登場物体およびその関連情報をノードとしたハイパーメディアである。以降、登場物体に対応するノードをオブジェクト、関連情報に対応するノードを情報ノードと呼ぶ。

2.2 ビデオハイパーメディアの高機能化

2.2.1 複合オブジェクト管理

ビデオハイパーメディアで扱うオブジェクトの構造が単純な場合、オブジェクトと情報ノードとの間に固定的なリンクを張ることにより、利用者が指示したオブジェクトからナビゲーション先を一意に決めることができる。

しかし、扱う対象が構造をもつ複合オブジェクトの場合、その構造を考慮したオブジェクト管理が必要となる。例えば、「車」という複合オブジェクトをノード化するとき、車全体の情報だけでなく「ボディ」や「タイヤ」といった部品個別の情報の管理も必要である。すなわち、一つのオブジェクトに対して、その空間的な構造に対応した複数のノードを作成することが必要になる。

2.2.2 状況依存検索

複合オブジェクトの場合、利用者がクリック（映像上でマウスで指示）した位置だけでは一意に提示すべき情

報を決めることができない。例えば「車」が、画面の片隅に小さく写っている場面と、画面全体に大きく写っている場面を考える。利用者が両場面で同じ部品の位置をクリックした場合、前者では車全体に関心があり、後者では部品そのものに関心があると考えられるため、場面に応じて提供する情報を変えるべきである。本稿では、利用者が指示したときの状況に合わせて提示情報を変えることを状況依存検索と呼ぶ。ビデオハイパーメディアで複合オブジェクトを扱う場合、この状況依存検索の技術が不可欠である。

検索に利用できる状況としては、オブジェクトの表示状況、ユーザの検索履歴など様々なものが考えられるが、本検討では、オブジェクトの表示状況の一つである画面中でのオブジェクトの大きさを利用することにする。

3 提案法

3.1 CG技術の応用

ビデオの多くはビデオカメラを用いて撮影されたものであり、ビデオ中のオブジェクトは、本来3次元の構造をもつオブジェクトを中心射影（透視射影）で2次元に射影したものである [4]。オブジェクトを3次元モデル化できれば、ビデオ中のオブジェクトの動きをCG（コンピュータグラフィック）でシミュレーションすることが可能である。そこで本稿では、CG技術を応用して複合オブジェクトに対応したノードを実現する。つまり、(1) オブジェクトの3次元モデルを作成し、(2) CGシミュレーションを作成し、(3) ビデオとシミュレーションを連動させる。(3)で利用者がビデオの画面上で行なったクリックイベントがCGのシミュレーション画面に伝わるようにすれば、CGのピッキング機能でクリック位置にある3次元モデルの部品を選択できる。

このようなノード作成方法には、二つの大きなメリットがある。一つは、オブジェクトの位置決めをすれば、その全ての構成部品に対する位置決めができ、オブジェクトが複雑になってもノード作成の負担が変わらない。もう一つは、後述する状況依存検索で利用するオブジェクトや部品の表示の大きさが、3次元モデルでの部品の大きさおよび位置の情報から計算で求められることである。

この手法では、ビデオとCGの位置合わせが問題となるが、その作業はビデオとCG映像を見ながらインタラクティブに行うことができる。なお、モデルの基準点をいくつか決め、それがフレーム画面中のどの座標にくるか入力する方法をとれば、より正確な位置合わせが可能である。

3.2 ルール記述による状況依存検索

前節で述べた手法により、利用者がクリックした複合オブジェクトの部品と、その表示状況（部品の大きさ）が求まる。ここでは、これらの情報を用いて提示する情報ノードを決定する状況依存検索アルゴリズムを提案する。

状況依存検索で選択される情報ノードの候補は、クリックされた複合オブジェクトの部品および、部品階層中でその上位の部品すべてである。状況依存アルゴリズムは、それらの部品の中から、表示面積がビデオ表示画面中で占める割合が、一定の値（最適割合）に最も近いものを選ぶとする。

アルゴリズムを実現するために、まず部品の大きさの評価関数を決める。部品 p の表示面積を $Area(p)$ とし、 p の評価関数 $E(p)$ を

$$E(p) = e^{-(Area(p) - S \times WindowSize)^2}$$

で与える。ここで、 $WindowSize$ はビデオ表示画面全体の大きさであり、 S は最適割合 ($0 < S \leq 1$) である。評価関数 E は、表示がビデオ画面中に占める割合が S に近いほど大きくなる。

部品 p の上位部品を $Parent(p)$ とすると、状況依存検索のアルゴリズムは以下のようになる。

```

program condition-sensitive(p);
var result, upper : part;
    max : real;
begin
    max := 0; upper := p;
    while ( upper ≠ nil);
    begin
        if E(upper) > max then
            begin
                max := E(upper);
                result := upper;
            end
            upper := Parent(upper);
        end
    return result;
end.

```

4 プロトタイプ

3章で提案した方法の実現性をプロトタイプを作成して検証した。システム構成を図1に示す。ビデオ表示システムは、ビデオの再生および利用者のクリックイベント処理を行なう。CG処理システムはクリックされた部品を識別し、3.2節に示した状況依存検索アルゴリズムに従って、利用者に対して情報ノードの提示を行なう。

プラットフォームはSGIのIndigo2を使った。本プロトタイプでは、1台の車が走っているビデオを素材にした。車を複合オブジェクトとし、車の各部品に対して情報ノードを付加した。車は約20点の部品から構成される。プロトタイプにおけるユーザインタフェースを図2に示す。

プロトタイプ作成を通して、提案手法が既存のプラットフォーム上で容易に実現可能であることが確認できた。

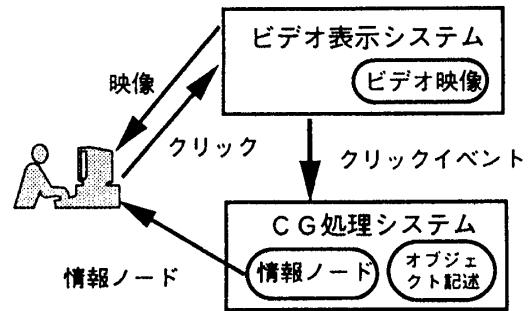


図1: システムの構成

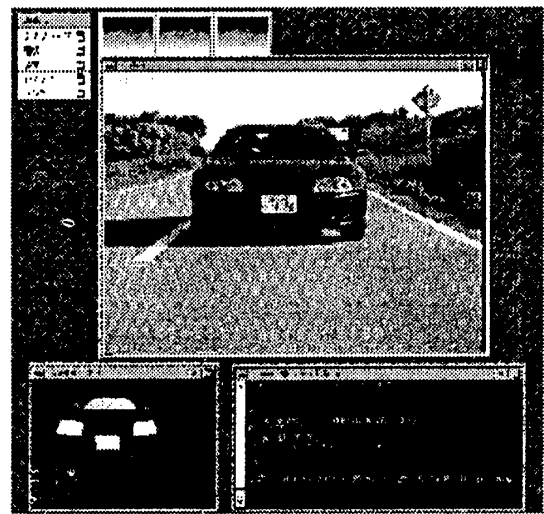


図2: プロトタイプのインタフェース

5 おわりに

本稿では、ビデオハイパーメディアの高機能化を、複合オブジェクトの管理と状況依存検索により実現する方法について提案した。また、ビデオハイパーメディアのプロトタイプを作成し、提案法の実現性を確認した。今後、我々の状況依存検索が利用者の意図に対して妥当であるか評価する予定である。また、表示の大きさ以外に状況依存検索で利用可能な状況を整理、検討し、それらの有用性を検証する予定である。

参考文献

- [1] 平田 恭二, 川崎 成人, 高野 元, 原 良憲, “ネットワーク環境下における動画ハイパーメディアの実装方式”, ADBS'94, pp.165-173(1994).
- [2] 高野 元, 的場 ひろし, 原 良憲, “ビデオ・ハイパーメディアのナビゲーション方式”, 第8回ヒューマンインタフェース・シンポジウム, pp.607-612(1992).
- [3] 佐藤 哲司, 片岡 良治, 井上 潮, “ビデオリアリティ: 映像を用いた情報検索手法の高度化”, 情報処理学会データベースシステム研究会 99-38(1994).
- [4] 日本写真測量学会編, “写真による三次元測定 - 応用写真測量 -”, 共立出版, pp.171-178(1983).