

動画像からの検索を可能とするハイパーメディアシステム

1C-9

宮田功治 中村竜也

NTTデータ通信(株)開発本部

1 はじめに

ワークステーション、パソコン用のマルチメディア対応、周辺機器の充実(音声ボード、ビデオボード)、CPUの処理能力の飛躍的向上、基本ソフトウェアでのマルチメディア機能のサポート、データ形式の国際標準化の進展とLSI化による普及(JPEG、CDROM)等の状況から、本格的なマルチメディアシステム構築の基盤が整ってきている。

このような背景のもとに、筆者らはデータ検索をマルチメディアを活用したブラウジングを用いて実現する技術の検討を行なっている。ここではマルチメディアを単に格納したデータとして利用するだけでなく、それをユーザインタフェースに本格的に利用し、マルチメディアデータベースの使い勝手を向上させることを狙いとする。本論文ではその一環として動画像をデータベースのブラウジングに用いるシステムについて報告する。

2 既存システムの限界

グラフィカルなユーザインタフェースを特徴とするHyperCardのようなシステムでは、データ検索を画像に"ボタン"と呼ばれるノードを張りつけることにより対応するイメージの部分をクリックすることで関係のあるノードやデータに移ることで実現している。しかし、動画像はあくまでもノードをたどった到達点であり、その動画像による人間の空間認知能力を使用すること、すなわち動画像で提示されている映像が表すイメージあるいはその一部をソースノードとしてさらに検索を進めることができない。このためにも、動画像の表示画面をダイレクトにクリックすることで映像内の対象物を指定できるようにする必要がある。

3 解決策

現在一般的に行なわれている静止画のイメージから他のノードに移動する機能を動画像についても可能にするためには、方法として以下の二つが考えらる。

(方法1) 動画像の各フレームにボタンの張り付け。

(方法2) 映像内の対象物の形状と三次元空間内での位置を前もって用意し、動画像の表示画面のクリックされた位置と映像のカメラ位置、方向と合わせて対象物を特定する。この二つの方法の比較を表1に示す。

オーサリングの作業手順とデータの少なさにより、我々は方法2を採用した。ただし、動画像からのカメラ

の位置、方向情報を取り出し、動画像の対象物の変形と動きの予測は現行の技術で解決することが困難である。そこで、動画映像内の対象物はすべて静止している、動画のカメラ位置、方向は既知である、の二つを限定条件として付加している。

4 動画像からの検索を可能とするハイパーメディアシステム

4.1 システム構成

コンピュータでマルチメディアを利用する際の現時点での大きな問題の一つはそのデータの量の多さである。オーサラーとユーザーではマシンへの要求条件はマシン性能、価格の点で大きく異なる。そこで、オーサリングとプレゼンテーションという目的別に二つのシステムに分割している。オーサリングシステムで生成したデータをプレゼンテーションシステムで用い、動画像の表示画面をクリックすることで映像内の対象物を指定できる。

4.2 オーサリングシステム

4.2.1 アーキテクチャ

オーサリングシステムの目的は動画像から、その動画像に含まれている対象物をノードとしてプレゼンテーションシステムで使用するためのデータを生成、蓄積することである。その方法は、利用者が対象物の形状を三次元モデラーにより生成し、その三次元空間内での位置を動画像と三次元モデルより生成される画像との合成画像で決定するものである。その構成は、高性能ワークステーション、モデルの三次元画像生成のためのグラフィックエンジン、動画像の表示のためのビデオボードおよび書換え可能光ディスクからなる。

4.2.2 三次元モデルの形状、位置の決定

三次元モデル形状、位置の決定および補正、訂正の流れを図1に示す。

ワークステーションの画面に利用者が対象物として採用したい対象物が含まれるフレームを表示させる。表示されている画面をクリックすることで対象物の位置(x,y)を決定する。他の異なる角度からのフレームで同様のことをすることで対象物の動画像内の位置を特定できる。これは作業中任意に変更、補正できる。

フレームと作成中の三次元モデルの生成画像を合成することにより確認しながら行なう。モデル作成時にフレームを変化させることで、異なるカメラ位置、方向の映像での三次元モデルの生成画像とフレームの合成した画像でモデル形状の対象物との違いを認識できる。これにより、三次元モデル形状生成が容易に行なえる。

The hyper-media system with video browsing

Koji Miyata, Tatsuya Nakamura

NTT DATA COMMUNICATIONS SYSTEMS CORPORATION

モデルの形状、位置の決定後、プレゼンテーションシステムで使用する形状情報 (SuperQuardlic)、位置情報 (x、y、z) を生成し処理は終了する。

ステレオ (左右からの二つの画像) の動画像を利用すれば、モデルの表示もステレオにすることにより生成画像の三次元モデルをX、Y座標だけでなくZ座標も表現することが可能となる。これにより、さらに利用者の立体的認識を高められる。

4.2.3 フレームと三次元モデル生成画像の合成

フレームと三次元モデルの生成画像の合成の流れのイメージを図2に示す。三次元モデルのバッチ情報をグラフィックエンジンのバッファ内に格納する。動画像のフレームをディスプレイに表示し、そのフレームのカメラ位置、方向情報をグラフィックエンジンに送る。これによりグラフィックエンジンはフレームに対応する見え方、大きさについて整合性のある生成画像をグラフィックエンジンのフレームバッファに作成する。グラフィックエンジンのフレームバッファの対象物の存在する部分に対応するマスクをワークステーション側に送る。

動画像の表示はワークステーションのフレームバッファの1ビットをスイッチャーとして利用している。、動画像表示領域にグラフィックエンジンで作成されたマスクに対応する領域のスイッチをオフにすることでマスク領域を動画像のフレーム以外の映像用に確保する。

確保した領域にグラフィックエンジンで作成した生成画像を表示することでフレームと生成画像が合成された合成画像を表示している。途中、マスク情報にメッシュをかけることでフレームと生成画像を同じ領域で同時に確認できるハーフトーンの合成を実現している。

4.3 プレゼンテーションシステム

4.3.1 アーキテクチャ

プレゼンテーションシステムはオーサリングシステムによって作成された情報を利用して利用者にマルチメディア・データベースへのアクセス方法を提供するものである。その構成は、低価格化のため、小型のワークステーションと動画像の表示にワークステーションから制御可能な光ディスクとビデオボードを使用する。プレゼンテーションシステムでは、三次元モデル情報を利用して、画面上のユーザークリック位置から対象物を識別するクリックングが鍵となる。

4.3.2 クリックングアルゴリズム

プレゼンテーションの際にその利用者は動画像の提示されている画面の目的となる対象物のある地点をマウスでクリックする。このことによりスクリーンの二次元情報 clicking position(a,b) が得られる。この情報とカメラ位置方向情報をもとに対応する世界座標を特定できる。視点から画面へのベクトルを延長しオーサリングシステムで作成した対象物の形状情報と位置情報との交点を求めることで選択された目的対象物を特定できる。

5 まとめ

本報告では動画像内の対象物をノードとして使用するために必要な技術を述べた。そのうちクリックングアルゴリズムを利用したプレゼンテーションシステムと動画像と生成画像の合成画像の作成について実現している。

今後の課題はオーサリングシステムでの動画像内の対象物の三次元モデル形状と位置の決定のインタフェースの作成を行なう予定である。

参考文献

- [1] Yokoyama, S., "CAPS (Computer Aided Presentation System)" Multimedia'90, 1990.
- [2] Pentland, A., "Automatic Extraction of Part Deformable Models" International Journal of Computer Vision, 4: pp. 107-126, 1990.

	データ量	作業量	対応技術	使用できる画像
方法1	対象物×フレーム数に比例	対象物×フレーム数に比例	従来技術で対応可能	多様
方法2	対象物に比例	対象物に比例	従来技術で対応不可	対象物カメラ情報の制限

表1: 方法1と方法2の比較

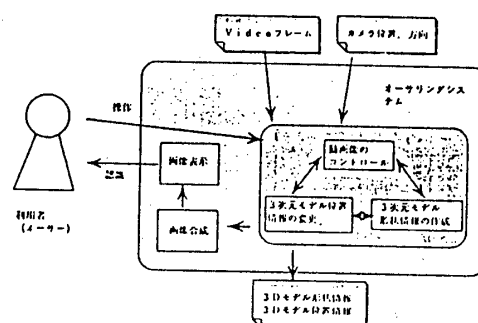


図1: 3次元モデルの作成の流れ

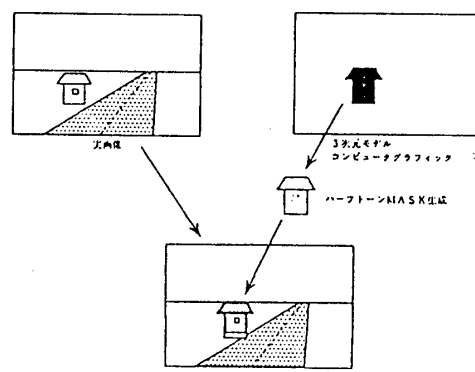


図2: ハーフトーン合成方法