

ハイパーメディアシステムのための映像シーン記述言語

4 J-5

佐藤 隆, 山根 淳, 龔 怡虹, 坂内正夫

東京大学生産技術研究所

1 はじめに

近年、画像や映像の自動認識技術を応用した、高度なハイパーメディアシステムへの期待が高まっている [1]。我々も、映像シーン記述言語 (VSDL) を用いた動画像検索システムを提案し、色とその動きを用いたシーンの一般的な記述方法を検討している [2],[3]。我々の提案したシステムを応用し、次のような高度なハイパーメディアを構築することを考える。

- 蓄積情報、放送情報から能動的に情報を見つけ出し選択することができる
- ユーザとのやりとりを通じてシステムが発展する
- 画像や映像内容にふみこんで、素材情報を獲得できる

このハイパーメディアシステムでの情報の流れを、図1に示す。ユーザは蓄積情報や放送情報から、検索したい情報を想定し VSDL によってそのシーンを記述する。システムは入力されている動画像からシーンを切りだし、ユーザの検索意図とマッチングをとり、マッチングに成功したシーンを出力する。出力結果は、認識結果に対応してシーン/フレーム/物体と重層的になっており、それぞれのレベルで利用できる。また、出力結果を改良し、次の検索意図に活用することで、検索意図をより自分の意図にあつたものに行うことができる。

このようなシステムを構築するためには、ユーザの意図を反映しつつ、情報を解析し、蓄積する統合的な手段が必要となる。そこで、VSDL をこれまでのスクリプト言語から発展させ、以下に述べる通り、認識・推論機能やユーザインタフェース機能を持たせるように拡張している。

2 ユーザの検索意図を記述する機能

VSDL では (1) 色の定義 (2) 色セグメントの定義 (3) 検索シーンクラスの定義の3つの階層で検索意図を表現する。このとき、色の色相・明度・彩度パラメータや、セグメントの大きさ・位置を指定するパラメータなどを、数値によって指定するには、少なからず無理があった。

そこで、色、色セグメントとシーンクラス、それぞれのオブジェクトについて、表裏一体のユーザインタフェースを設け、視覚的に検索意図を構築することを可能にする。例えば、色セグメントをマウスによって移動し配置できるようにしている。

3 認識のプロセスを記述する機能

VSDL を解釈しシーンを検出するシステムは、シーンクラス (SC)、メジャメントクラス (MC)、サブエリアメジャメントクラス (SAMC) の3つの部分から構成さ

Video Scene Description Language for Hypermedia System

Takashi SATOU, Jun YAMANE, Yihong GONG, Masao SAKAUCHI

Institute of Industrial Science, University of Tokyo

れている。これまでは、これらは VSDL のいわば「外部」にあつて、VSDL は単なるスクリプト言語にすぎなかった。しかしながら、より複雑な検索を可能とするためには、VSDL 自体に認識・推論を可能とするプログラム記述能力をもたせ、システムの定義までも VSDL によって記述できることが望ましい。

例えば、パラメータの値を、実行時に計算で求める必要がある場合には、値を宣言的に記述するだけのスクリプト言語では手も足も出なかった。しかし、新しい VSDL ではそれが可能となる。

また、これまでは色セグメントの形状について記述することができなかったが、高度な認識のための推論機能をもった新しい VSDL では、形状も考慮に入れたシーン検出が可能となる。

この他、シーンクラスにローカルな処理を記述しておいて、システムの挙動を部分的に変更することが可能である。

4 認識結果の利用を容易にする機能

動画像検索システムでは、最終結果として動画像のシーンを出力するだけであつた。しかしながら、再利用可能な素材情報として欲しいのは、シーン全体よりはむしろ、その中に登場する人物、動物や物体であることが少なくない。逆に言うと、これらシーン構成要素は、シーンを検索する際のキー情報となっているものであつて、検索結果に重層的に含めることが可能である。

シーン全体や、その構成要素を視覚的に表示し、ダイレクトマニピュレーションを可能にすれば、映像中の物体をハイパーテキストに切り貼りして柔軟に保存管理したり、新たな検索意図を作成するときの部品に活用したりできるであろう。

5 認識・検出結果を蓄積するための内部表現としての機能

シーンやその構成要素を、再利用可能な素材情報にするためには、何らかの方法でその内容を保存する必要がある。このため、これらのオブジェクトを記述する VSDL に永続性をもたせ、蓄積のための内部表現として働くようにする必要がある。

オブジェクトを保存しておくことは、素材情報を蓄積する以外にも有用である。例えば、蓄積情報から再びシーンを検索する場合に、映像をすべて解析しなおすのではなく、以前の解析結果を利用して効率よく処理を行うことができる。

6 実装

以上の考察をまとめると、オブジェクト指向言語を用いて、色、色セグメント、あるいはシーンクラスなどをオブジェクトとして定義し、そのオブジェクトにユーザインタフェースと永続性を持たせる構成で VSDL を実装することが考えられる。

ここでは、オブジェクト指向の拡張を行なった Prolog 上位互換言語 GOLS (Graphic Oriented Language System) を開発し、それを用いて VSDL を記述している。

GOLS のオブジェクトは、画面上のウィンドウやアイコンと対応づけ、ユーザのダイレクトマニピュレーションを可能とする。

GOLS は次のライブラリをもつ。また、ユーザインタフェースライブラリの使用例を図2に示す。

- ユーザインタフェースライブラリ
ウィンドウやアイコンを表示し、ユーザ入力を受け付ける。プッシュボタン、ラジオボタン、1行エディタ、スライダー、動画アイコンなど。
- ハイパーメディア構築ライブラリ
ハイパーテキストのノードとなるウィンドウ。ウィンドウをデザインするインタフェースビルダ、ドロウツール、お絵かきツール、色パレット。
- 画像処理ライブラリ
フィルタリングやラベリング、拡大・縮小などの画像処理モジュール。
- データベース操作ライブラリ
OODBMS とのインタフェース、BD 木による図形空間管理など。

新しい VSDL は以上に述べた GOLS の機能を使い、ライブラリとして実装される。

7 おわりに

画像・映像の自動認識技術を応用した高度なハイパーメディアシステムに必要な、映像シーン記述言語 (VSDL) の機能拡張について考察し、その実現のためのプラットフォーム GOSL を開発した。今後は、新しい VSDL によるシーン記述と実験を行なう予定である。

参考文献

- [1] 君山, 清末, 大庭: “動画像中物体に対する自動記述生成の検討”, 情処第44回全大 5B-8 (1992)
- [2] 山根, 佐藤, 巖, 坂内: “映像シーン記述言語を用いた動画検索システムの提案” 情処44回全国大会 (1992)
- [3] 佐藤, 山根, 巖, 坂内: “シーン記述言語を用いた動画検索システムとその評価”, 第129回画像電子学会研究会, (1992)

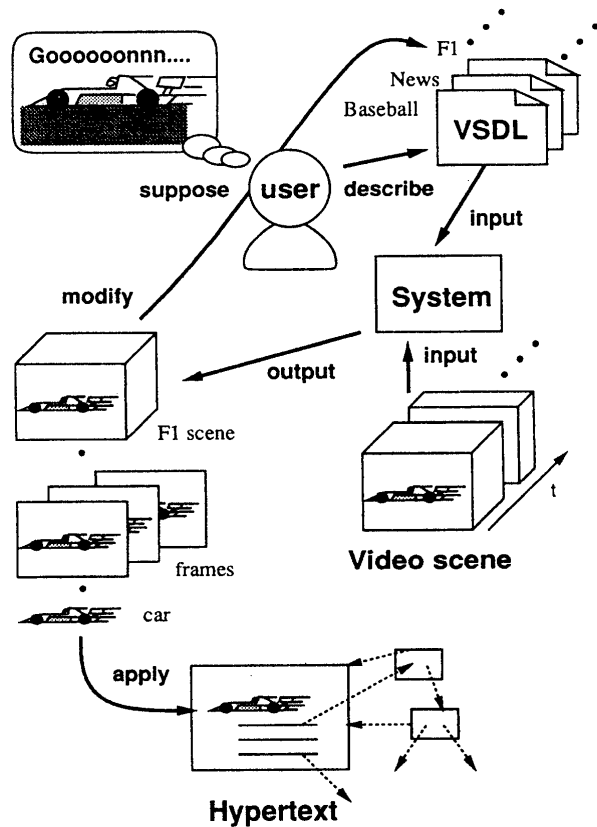


図1: ハイパーメディアシステム

```

@ my_win : c_window { % ウィンドウの定義
    shape(0,0,300,200). % 大きさ (300x200)

    .
    .
    .

@ my_button : c_button { % ボタンの定義
    shape(50, 20, 50, 20). % 位置と大きさ
    label(hello). % ラベル
    state(0). % 初期状態
    action :- % ボタン押下の反応
        write('Hello, World!'), nl.
    }.

@ my_image : c_image { % 画像アイコン
    shape(120, 20, 128, 128). % 位置と大きさ
    image('mori3.img'). % 画像ファイル名
    }.

}.

?- my_win::draw, my_win::x_map. % ウィンドウの表示
?- x_run. % イベントループ起動
    
```

図2: GOLS の記述例