

データベースを中心とした番組制作支援システム

住吉英樹 有安香子 望月祐一 佐野雅則 井上誠喜
{sumiyosi,ariyasu,motiduki,sanom,sinoue}@strl.nhk.or.jp
NHK放送技術研究所 マルチメディアサービス

デジタル放送で必要となるマルチメディアコンテンツの制作支援と同時に、番組映像の2次利用を目的とした映像データベースを効率的に構築するためにデータベースを中心に配置した新しい番組制作システムを提案、試作した。本報告では、このシステムで用いるデータ構造の概念と、入力となる素材映像への効率的なインデックス情報付加の仕組み、蓄積されたデータの応用例について報告する。

A Database-equipped Support System for TV Program Production

Hideki Sumiyoshi Kyoko Ariyasu Yuichi Mochizuki Masanori Sano Seiki Inoue
NHK Science and Technical Research Laboratories

We have constructed a trial system for supporting production of multimedia contents required for digital broadcasts. The system is based around a video database whose purpose is to store already broadcast video and text information in such a way so that they can be easily accessed and reused in multimedia contents. This report explains the concept behind the structure of the stored data (i.e., text/video), describes how efficient indexing information is added to the videos inputted into the database, and presents an example application of the stored data.

1. はじめに

2000年末に予定されている衛星デジタル放送、また、その後続く地上デジタル放送など、本格的なデジタル放送時代を迎えようとしている。デジタル化によるチャンネル数の増加で、番組コンテンツ需要の増大が予想されている。デジタル化は映像や音声信号の伝送方式を変えるだけでなく、図1のような新しいマルチメディアコンテンツの配信を可能にする。また、インターネットの普及によって、図2のようなWebでの番組紹介や補足情報の提供など、新しいコンテンツの制作を支援するシステムへの期待が高まっている。

一方、現在の放送局では、映像をハードディスクに記録して編集などに利用するノンリニア編集機や特殊効果装置などに代表されるデジタル化、コンピュータ化された高機能な制作機器が導入、利用されている。しかし、これらの機器は映像・音声の処理に主眼が置かれており、制作過程で利用されている台本などの文字情報を操作することは考慮されていない。このため、制作過程における文字情報の共有や再利用が困難な状況にある。

本報告では、番組情報を蓄積するデータベースを中心に置き、制作時に使用される情報を制作の過程で整理、蓄積し、情報の共有と活用を目的と



図1 新しいデジタル放送コンテンツの例

した番組制作情報データベースを構築する新しい番組制作システムとそこで用いるデータ構造、映像データベース構築に不可欠なインデックス情報の効率的な入力方法、蓄積された情報の効果的な応用例について報告する。

2. 番組制作で利用される情報

番組制作時に利用される情報は、(1)映像音声情報、(2)ナレーションやスーパーインポーズなどの文字情報、(3)それらを互いに結び付ける番組構造情報の3つに分類できる。最終的にこれらの情報は、映像と音声として1つにまとめられて放送されており、直接、表に現れることのない文字情報と構造情報は、制作作業中には極めて重要な要素として取り扱われているにもかかわらず、放送後に廃棄されることが多い。また、映像を別番組で利用する場合には、個々の映像に対する著作権情報が非常に重要なものとなるが、数が多い上に映像ごとに利用条件が異なるなど、管理が難しい情報である。

ここでNHKでの一般的な番組制作作業の工程と情報の流れを図3に示し、以下に説明する。

- ・企画：番組で何を伝えるかを定める
(番組提案書などの作成)
- ・調査、取材：番組テーマに関する情報を収集
(資料映像、文献、書籍などの情報収集)
- ・構成：情報の選択、話題の展開、提示順序検討
- ・撮影：構成をもとに撮影
- ・編集：映像に合わせ、構成を変更し映像を繋ぐ
- ・台本作成：ナレーション原稿作成
- ・スタジオ収録：編集済み映像とナレーション、スタジオ映像などを組み合わせる
- ・関連資料の作成：広報資料、出演者、著作権管理
- ・放送

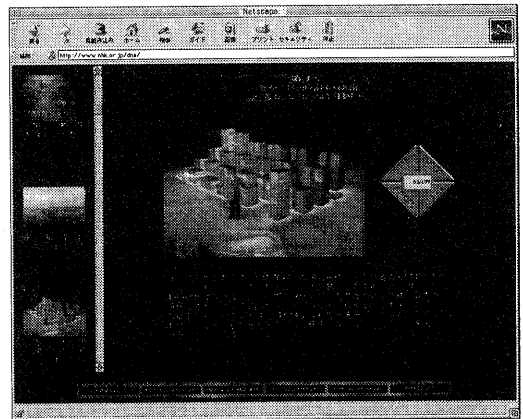


図2 Webでの番組情報の提供例
(NHK 驚異の小宇宙 人体III 遺伝子・DNA)
<http://www.nhk.or.jp/dna/>

・番組データベース登録

ここで図3の上部は、映像・音声の操作を中心とした作業領域で、現在の制作機器で支援される作業である。下部は、番組で使用される文字を中心とした情報の種類と作業の流れを示している。下部に示した作業では、何をどのように伝えるかという番組意図や、その意図を映像音声で構成するために必要となる情報が、台本や構成表、取材メモやノートといった形で扱われている。特に複数の人間が協調して作業を行う番組制作では、番組の目的や撮影すべき映像などの情報を作業間で共通の認識にするため、台本や構成表といった資料が不可欠である。また、番組完成後には、視聴者対応を行う部門へ出演者の連絡先などを配付したり、番組データベースの登録のためにも管理部門へ番組概要を記した資料を提出する必要がある。これらの情報の間には、共通する項目が数多くあるにもかかわらず各作業段階で別個に作成、利用されているのが現状である。

これら番組制作の過程で用いられる情報をデータベース化して共有し、再利用することが制作作業の効率化につながるものと考えられる。また、調査資料などの中には、映像の出所(著作権者)や番組のテーマに関するより詳しい情報や関連情報などが多く含まれている。現在十分に蓄積、管理されていないこれらの情報を映像・音声と統合してデータベース化することが、番組制作作業を効率化し、新しいコンテンツの制作を支援にするものだと考える。また、印刷物やWeb、DVDなどのマ

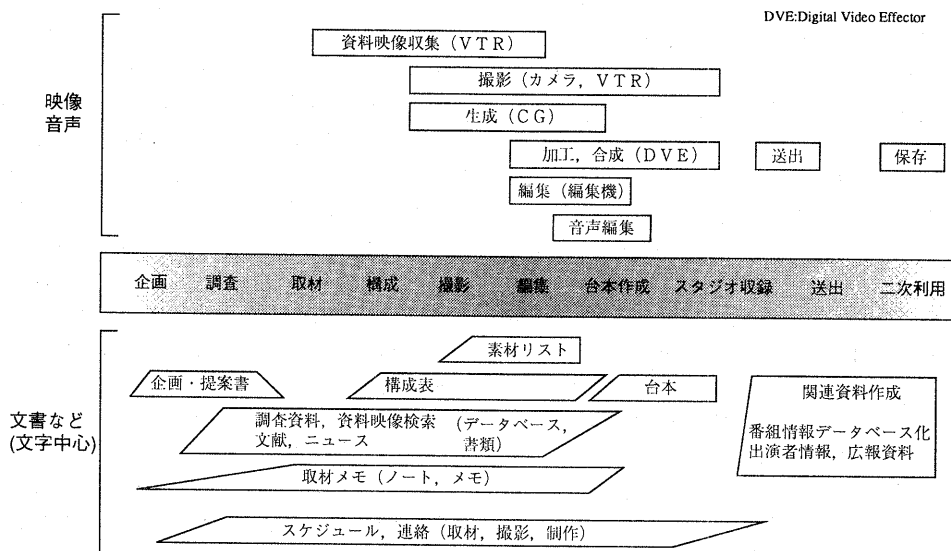


図3 番組制作の作業工程と情報の流れ

ルチメディア応用を進めていくためには、これらの放送には使用しなかったより詳しい情報が映像と関連づけ蓄積されるデータベースを構築することが重要となる。

3. データベースを中心とした制作システム

前述のような番組制作情報をデータベース化する際に、放送局が持つメリットとして「何を撮影したのか」、「このシーンには、この意味を持たせる」など各情報の内容や意味を考慮しながら映像コンテンツを作り出している点があげられる。この制作段階の知識を電子化し、整理、関連づけて蓄えることで、正確な情報の蓄積とデータベース構築作業の効率化が可能となる。

このような映像制作者側の特長を活かした情報管理システムの研究には、上田[1]らの映像編集システムに文字情報の管理機能を統合し、映像コンテンツを作る過程で著作権情報などの文字情報を付加しながら統合的に管理する手法などがある。

筆者らは上田らの映像を表わすアイコンに対して情報付与するという方法ではなく、放送現場で使用されている番組の設計図に相当する構成表を情報管理の枠組みとした番組情報データベースを中心に置き、このデータベースに映像を含む番組情報を入力しながら番組制作作業を行う手法を採

用した。このデータベースを中心とした情報の操作概念を用いる DTPP (Desk-Top Program Production) と呼ぶ番組制作システムを提案している[2]。

3.1 DTPP の情報管理構造

番組で使用する情報をデータベースとして効率的に蓄積するためには、番組の情報構造を的確に表現する構造定義が必要である。一方、番組で何を伝えたいのか明確にするためには、ストーリー展開を構造的に考えながら番組を作ることが重要であると言われている。

我々は番組の設計図として制作者がよく利用する図4に示すような「構成表」に注目した。構成表には文章の段落に相当するような項目名とおおまかな説明内容が記述されており、一目で話題の流れが把握できる。図4の構成表からも明らかなように番組には階層的な構造がある。このような番組の意味的な情報のまとまりを示す構造を基に番組制作作業、特にポストプロダクションと呼ばれる映像編集を中心とした作業とそこで用いられる情報の性質を考慮して、階層的な情報管理構造を定義し、企画段階から編集、二次利用の間で使用される情報を蓄積するデータベースの枠組みとした。具体的には、映像の最小単位(切れ目が無い)である「カット」と、「カット」を複数組み合わせ

て制作者の伝えたい意図を表現する映像のまとまり「シーン」を基本要素として、編集時のデータ操作性を考慮して図5(c)に示すような6階層のデータ構造を定義した。構造要素は、上位から「プログラム」、「シーングループ」、「シーン」、「カットグループ」、「カット」、「情報(Info)」である。この構造に番組で使用される映像・音声、文字(アナウンス、スーパー)を素材要素として関連づけることで番組データベースを構成する。素材要素としては、映像、アナウンスコメントなど9種類を定義している。映像などの素材情報は、それぞれに付与された検索情報や時間情報、著作権情報などを保持するため、構造要素の最下層「情報(Info)」を介して結合する構造とした。

3.2 DTPPシステム

DTPPは、図3の企画立案段階から編集、二次利用までの制作作業と情報の管理を支援するために番組情報を蓄積管理、操作するための枠組みを設定し、ネットワークを利用したコンピュータ環境で、作業間、作業者間での情報の共有と活用を目指している。

前述のように放送局は、番組の作り手であり番組の意味的構造とそこで使用する映像の内容を知っている。一方で、近年のパーソナルコンピュータやワードプロセッサの普及により、台本や構成表など作業中に使用される多くの文書は、制作者により電子データとして作成されている。

そこで番組情報を的確に表現できる階層的な情報の枠組みをデータベースのモデルとしてシステムの中心に置き、作業を効率的に行うことができるユーザーインターフェイスを前面に配することで、作業者がデータベースを意識することなく、番組情報をデータベースの要素として入力することができるDTPPシステムを試作した。直接放送されない取材情報なども統合化し、番組構造と関連付けて蓄積し、データベース化することができる。

DTPPシステムのユーザーインターフェイスの一つは、文字情報と番組構造を中心に操作するもので、図5(a)のように従来の「構成表」を作成する感覚で各情報が入力、編集操作できる。下位の映像や文字を複数含んだまま移動、削除などの操作ができるので、番組の話題展開の検討を中心とした編集が可能となっている。もう一つのユーザーインターフェイスは、図5(b)の簡易なノンリニア映像編集機能を持つもので、フレーム単位での映像切替え点の指定やアナウンスコメントの読みはじめの時間を指定するなど、映像編集操作を中心に時間軸上での各情報の関係づけ作業を行うことができる。

これらのユーザーインターフェイスを使用して入力、作成された番組データベースからは、番組制作に必要となる構成表や台本など各種の書類や編集データなど必要となるフォーマットに合わせて情報を取りだせるので、何度も同じような情報を入力する労力を減らすことができる。

4. 映像素材データベース

素材要素の一つである映像をDTPPシステムで用いるためには、撮影VTRから映像をコンピュータに取り込み、撮影カット毎に分割し、データベース化しておく必要がある。試作システムによる番組制作実験では、この作業に関する時間と労力の大きさが指摘された。また、番組制作にお

月の地震で地球の起源を探る
文部省宇宙科学研究所・水谷仁さん

項目	映像	内容	時間
1. イントロ			
●オープニングVTR	<ul style="list-style-type: none"> 名物研究室タイトル ベネトレイター2 I アポロ計画での月面調査 月に近づいていくルナA衛星 月に近づくベネトレイター 何十本ものテスト用のベネトレイターの中に立つ水谷さん 	アポロ計画で人類が初めて月に降り立ちってから27年。今、人工衛星で月を調査するという日本で初めての試みが進められている。ベネトレイターと呼ばれる観測衛星を打ち込み、月の地震を測ることで、その誕生の謎に迫るルナA計画。その研究チームの中心となっているのが、宇宙科学研究所の水谷仁さんである。	
2. ルナA計画とは?			
●水谷さん紹介	<ul style="list-style-type: none"> 宇宙科学研究所ロング ロビーに入っていくリポーター 陳列されている人工衛星 コントロールルームに入るリポーター 水谷さんとあいきつ 水谷さんUP 	<p>神奈川県横浜市の宇宙科学研究所を訪ねました。(この宇宙科学研究所は、独自のロケットを開発して様々な宇宙の謎を解くための科学衛星を打ち上げてきた世界的にもユニークな研究所です。ロビーには、現在開発中の科学衛星などが陳列されています。)</p> <p>(こんにちは、～)</p> <p>宇宙科学研究所の水谷仁さん。異年、太陽系の成り立ちを研究してきた水谷さんは、10年前から、月を調査するルナA計画に関わっています。</p>	
●ルナA衛星紹介	<ul style="list-style-type: none"> つなぎを巻くリポーター 椅子をかぶる エアを浴びる ルナA衛星に近づいていく ルナAUP ノズル 	<p>早速、現在開発中のルナA衛星を見せてもらうことにしました。精密な機器をほごりから守るために、衛星はクリーンルームに置かれています。後で修理することができないため、わずかな故障でも命取りになるのです。(初めてルナA衛星を見た感想。大きさなど。)</p> <p>これが、月探査に使われるルナA衛星です。高さ2メートル、重さは520キロ。3本のベネトレイター</p>	

図4 構成表(例)

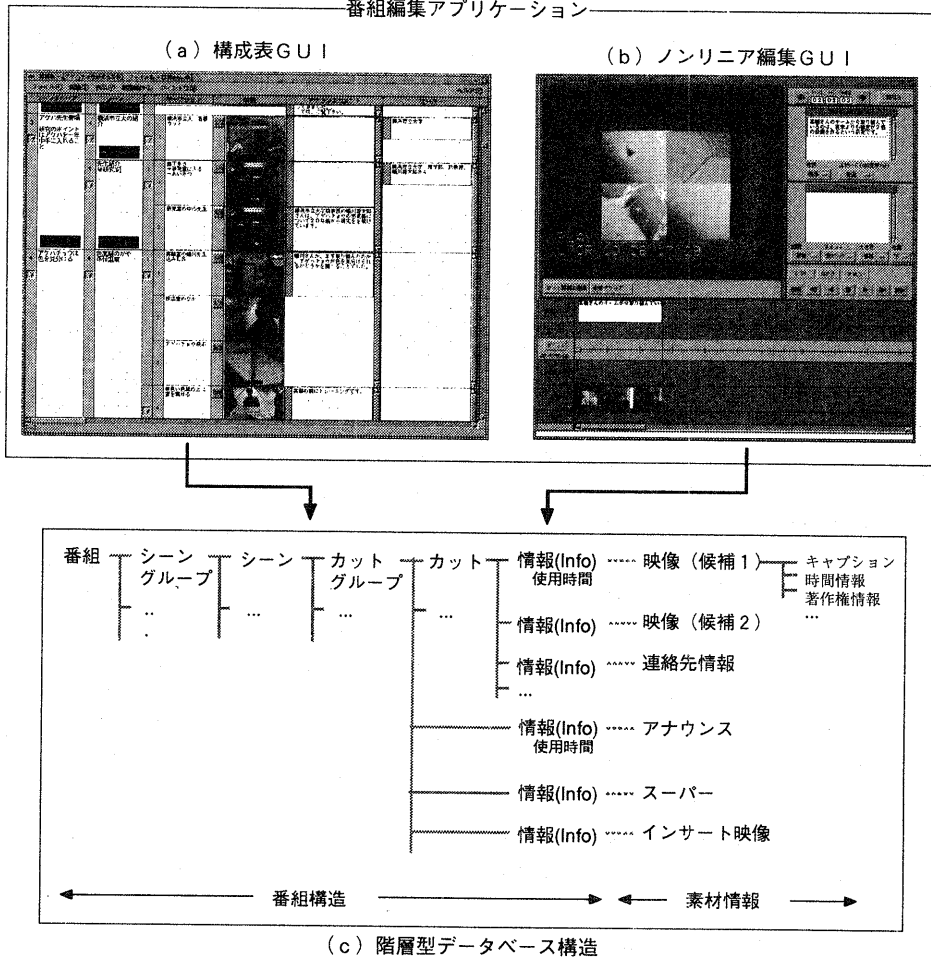


図5 データベースを中心とする制作システム

いては、一般的に放送される時間の十倍から百倍以上の時間に相当する何百、何千カットもの映像が撮影される。編集作業では、この大量の映像の中から番組制作意図をうまく表現する映像を選び出さなくてはならない。現在、この作業は編集マンと呼ばれる専門の作業者が人手で行っている。編集マンは、撮影されたビデオテープを事前に試写し、どのテープのどの部分に何が収録されているのか、テープ上の時間と映像内容を素材リストとしてノートに書き起こして管理している。通常、編集マンは撮影に同行しないため被写体が何かを正確に知るには、撮影に同行した制作者に確認する必要がある。

各カットを代表する静止画がサムネイル画像として表示されるノンリニア編集環境やDTPPで

あっても、カット数が数百から数千にもなれば、希望する映像を探しだすのは現実的ではない。効果的に素材管理をするためには、検索用のインデックスを各カットに付けた映像データベースが求められている。しかし、この映像素材データベースは、基本的に番組制作期間中しか利用されないもので、データベース作成(=インデックス付け)に大きなコストをかけるわけにはいかない。短時間かつ簡単に、できれば自動的にデータベース化する仕組みが必要とされている。

4.1 映像素材へのインデキシング

映像素材データベースに求められるインデックス情報は、主に次のようなものである。

(1)いつ撮影されたのか

- (2)どこで撮影されたものか
- (3)被写体が何(誰)であるか
- (4)被写体は何をしているのか
- (5)どのようなカメラアングル,カメラワークか
- (6)番組のどこで使用するための映像か

これらのインデックスの中で,(3)の情報の生成には,認識技術を使用した,いわゆる映像インデキシング技術が利用できる.人物[3]や建物[4]に対しての映像認識技術,文字認識技術を使用した自動インデキシング技術が研究されている.また,静止画像に含まれる文字部分をOCR技術によって文字情報を得る翻訳カメラ[5]も提案されており,映像内の看板や地名の表示から撮影場所や対象の情報を得てインデックス化することが期待できる.しかし,さまざまなアングルや照明環境で撮影された実際の番組素材に適用するには,まだ多くの課題が残されている.また,与えられた映像のみから対象物が「何」であるか認識するには至っていない.

コンテンツ制作サイドのメリットを活かしたインデックス付加技術としては,取材用のVTR一体型カメラで撮影中のカメラマンが使用できると判断した映像にマークを付けられるものがある.撮影対象や撮影目的を知っている制作者やカメラマンが撮影時に情報を付加すれば,後の制作作業が効率的になり素材の2次利用などを行う際の検索情報や,著作権,肖像権管理などに結びつく極めて重要な情報になる.撮影時点での作業量は増えるものの,制作作業全体を考慮した場合,正確な情報が最も低コストで付加できる方法だと考える.

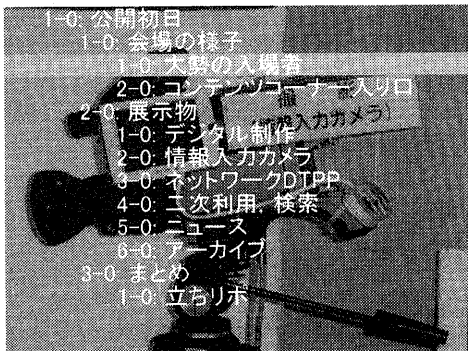


図6 情報入力カメラのビューファインダーイメージ (構成選択状態)

有川[6]らは撮影時のカメラの状態を各種のセンサーを用いて時空間記述情報として生成し,得られた時空間記述情報を実時間3次元コンピュータグラフィックスとして可視化することで,大量のビデオデータの管理,検索を可能にするハイパーメディアシステムの研究を行っている.

我々は,より番組制作へ特化した情報付加装置として,前述のDTPPシステムを用いて企画・構成段階で入力された番組構成情報を撮影時に利用する次のような撮影システムを考案,試作した.撮影時に(1),(2),(5),(6)の情報を付加して映像素材データベースを構築し,効率的な制作作業を実現する実験を行っている[7].

(a)構成情報による撮影映像の管理

撮影現場で内容が変更されることも多いが,おおまかな番組構成と撮影内容(予定)は,番組の撮影前に決まっている.この番組構成と撮影内容を映像の管理情報として映像と同時に記録し,(6)に対応する情報を得ようとするものである.

映画撮影などでシーン番号やカット番号が書かれた「カチンコ」と呼ばれる板を撮影し,映像の管理に利用しているが,その電子版と言えるものである.

試作情報入力カメラでは,付属のコンピュータ

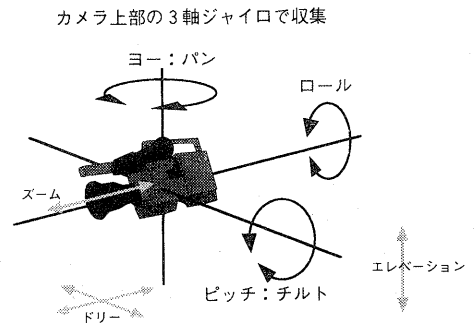


図7 カメラワーク用語とカメラの動き

表1 素材に付加されたカメラワーク用語

	番組1	番組2
総クリップ数	209	163
右パン	8	14
左パン	14	9
チルトアップ	7	6
チルトダウン	8	6
ズームイン	44	35
ズームアウト	26	24

に番組構成情報を読み込ませ、図6のようにカメラ映像に構成情報を重畳して表示し、簡単なボタン操作でこれから撮影する映像がどの構成に対応するか選択して撮影を行う。

実験では、図5(c)の中のシーングループ、シーン、カットという番組内の主要な3階層に撮影映像を対応させて記録し、映像を構成順に並べ替える簡単な編集の自動化を行っている。

(b)機械的に得られる情報を用いたインデキシング
制作者は映像を表現する「パン」や「ズーム」といった図7のようなカメラワーク用語を良く用いて映像を表現している。これらのカメラワーク用語は、映像の特徴としての動きを表すと同時に、そのカメラワークが行われている映像区間を指定する意味もある。

実験的に作成された番組の映像素材データベースには、表1に示すように多くのカメラワーク用語がインデックスとして付加されていた。これらのカメラワーク用語単独では検索情報にはできないが、サムネール画像などからでは、区別しにくい映像を絞り込むのに有効な情報になると考える。また、通常カットを最小単位とする映像をカメラワークを利用して、さらに小さな単位にセグメントに分割し、映像編集作業にも役立てることが可能となる。

試作したカメラでは、3軸ジャイロを上部に付加し、得られたヨー、ピッチ、ロールのデータを映像のフレーム番号と合わせて記録して、(5)に対応するパン、チルトなどのインデックス情報を生成している。ヨー軸に対し一定時間以上、一定方向に変化していればパン、ピッチ方向の動きを同様にチルトと判断して、映像素材データベースのインデックス情報として出力される。さらに、撮影時刻と屋外であればGPSで測定された緯度経度データを記録し、地図情報データベースから地名を生成して(1),(2)の、「いつ」「どこで」に相当する情報をインデックスとして付加する実験を行っている。

5. 番組情報のXML表現による汎用化

蓄積された番組情報を多用途に利用することを目的として、XML(eXtensible Markup Language)を用いて構造文章化し、汎用的に扱えるようにした。XMLには、以下のような特長がある。

- ・階層構造の表現に適する
- ・任意のタグが指定できる
- ・W3Cで規格化された言語で汎用性があり、処理ツールが豊富

これらの特長を活かし、DTPPで用いている階層的なデータをXMLを用いて表現し、様々な用途に利用できる形式とした。同時に検索や2次利用時には不要となる編集用の情報を取り除き、図8に示すようなタグ付けを行った。XMLを用いて番組制作情報を表現することで、システムへの依存性を減らし、他用途への効果的な活用を目指している。[8]

このXMLデータの活用例として、検索システムを試作した。検索システムでは、番組制作情報内

```
<?xml version="1.0" encoding="shift_jis"?>
<!-- DTPP2XML Script Version 1.0.0-->
<番組ファイル>
  <DTPP情報>
    <DTPPバージョン>2.2</DTPPバージョン>
  </DTPP情報>
  <番組情報>
    <番組付属情報>
      <番組名>サイエンスアイ</番組名>
    </番組付属情報>
    <番組リソース/>
    <シーングループ>
      <シーングループの時間 単位="フレーム">12589</シーングループの時間>
      <シーングループの説明>体へ応用される組織再生</シーングループの説明>
    </シーングループ>
    <シーンの時間 単位="フレーム">750</シーンの時間>
    <シーンの説明>東京歯科大学¥n山田教授紹介</シーンの説明>
    <カットグループ>
      <カットグループの時間 単位="フレーム">245</カットグループの時間>
    </カットグループ>
    <カットの時間 単位="フレーム">245</カットの時間>
    <カットの説明>東京歯科大学ロング</カットの説明>
    <動画 使用="Y">
      <動画ファイル名>takyo3.mov</動画ファイル名>
      <著作権管理情報>
        <著作権情報>NHK内</著作権情報>
      </著作権管理情報>
      <ロール番号>0194</ロール番号>
      <代表フレーム番号>0</代表フレーム番号>
      <代表静止画>VID0194001_0</代表静止画>
      <開始時間 単位="フレーム">119</開始時間>
      <終了時間 単位="フレーム">363</終了時間>
    </動画>
  </番組情報>
</番組ファイル>
..... 略 .....
```

図8 XMLデータの例

の「情報の関連性」を利用した検索が可能である。DTPPシステムで蓄積された番組制作情報には、番組に関する個別の情報と、それらの関係を示す情報とが含まれている。例えば、映像データと共に、その映像の説明（キャプション）や映像に付けるアナウンスコメントなどの情報が関連付けられている。この関連づけられた情報に対して、図9のような検索条件設定画面を用いて検索を行うことができる。

「カットの説明」に

検索文字列 富士山 を含む
代表画像 とその 著作権情報

従来の映像データベースのように、映像に付加された単純なキーワードだけではなく、番組内で使用された情報の関連を用いた条件を入力し、必要な情報を検索し図10のように表示することができる。

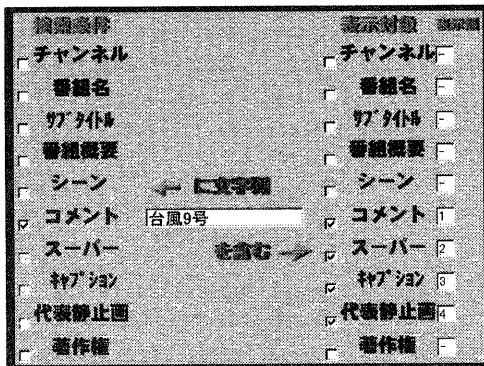


図9 検索条件入力画面

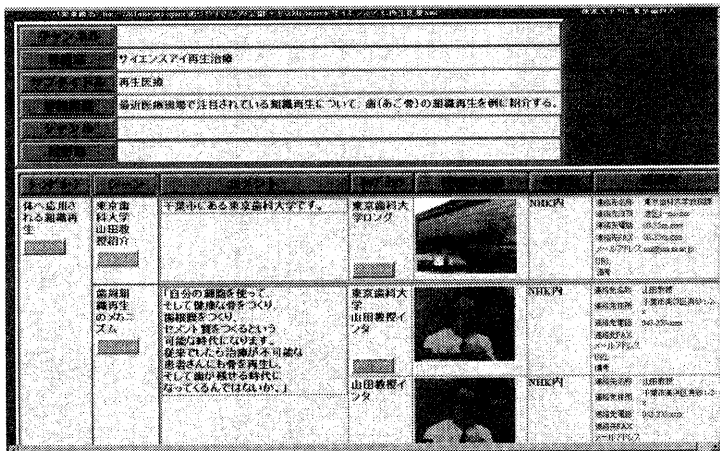


図10 検索結果表示画面例

6. まとめ

放送番組の制作過程における情報管理と活用を目的として、データベースを中心に置く新しい番組制作システムで用いているデータ構造と、要素となる映像素材を効率的に管理し編集に結び付ける撮影システムについて報告した。また、蓄積したデータの高機能な検索応用について述べた。

今後は、より高機能な素材管理データベース構築のためのインデキシング技術の検討を進めるとともに、蓄積された情報の効果的な利用方法について検討を行っていく。

参考文献

- [1]上田：コンピュータを駆使した最新の放送番組制作技術，情報処理，Vol. 40 No.11 (1999)
- [2]住吉，有安，望月，佐野，井上：階層化番組制作手法における番組情報構造，電子情報通信学会総合大会，SMD-2，(1998)
- [3]Shin' ichi Satoh, Yuichi Nakamura, Takeo Kanade, Name-it: Naming and Detecting Faces in News Videos, IEEE Multi Media, pp.22-35, Jan-March 1999.
- [4]谷田部，川崎，坂内：実世界映像の自動構造化およびこれを用いた3次元都市の自動生成，信学技報，IE99-15, PRM99-39, MVE99-35 (1999)
- [5]渡辺，岡田，金，武田：翻訳カメラ，第3回知能情報メディアシンポジウム，pp.271-278 (1997)
- [6]有川：カメラの時空間記述情報を利用したビデオデータの空間ブラウジング，情報処理学会研究報告，98-DBS, 116-56 (1998)
- [7]住吉，有安，井上：デジタル放送時代のコンテンツ制作技術，画像電子学会第172回研究会，99-02-4, pp17-20 (1999)

[8]有安，住吉，望月，佐野，井上：番組制作情報を用いた高機能検索システム，電子情報通信学会総合大会，SMD-1，(1998)