

映像データベースにおける情報の入力と管理

草場 匡宏 高橋 淳一 洪 政国

日本アイ・ビー・エム株式会社 東京基礎研究所

映像は博物館が管理する資料の中でもデータの質・量で従来扱われてきた文字や静止画などのデータと比較し膨大な情報を持ち、それらを管理するためのデータベースの構築・管理は緊急の課題となっている。我々は映像をシーンの集合としてとらえ、従来人手にのみ頼られてきたシーン情報の入力と管理を支援するためのユーザインタフェースを研究している。現在国立民族博物館との共同研究において映像データベースのプロトタイプを構築し、ユーザーによる試用によってその問題点を明らかにしようとしている。本稿では我々の映像データベースの構築に対する基本的なアプローチと問題点、その解決法と共に映像データベースの在り方について考察する。

A method of a video scene information entry and its management

Masahiro KUSABA Junichi TAKAHASHI Jung-Kook HONG

IBM Research, Tokyo Research Laboratory

This paper introduces a method of archiving motion video and a prototype system called M-CIDB (Motion Color Image Data Base). Because of enormous volume of information and data, there are some serious problems in motion video archiving including (1) scene indexing and database creation, and (2) searching scenes of interest. To solve these problems, the following techniques are proposed which help people input the scene information of motion video and manage it interactively: (1) a data structure of describing scene information, (2) an automatic scene change detection, and (3) a scene information editor. A prototype system has been developed by covering multimedia data in a joint research program with National Museum of Ethnology.

1 はじめに

近年ハードウェアの低価格・高機能化により、これまでの文字・数値を中心としたデータベースから静止画像まで扱えるデータベースが普及しつつある。一方映像・音声のデータベースについてはその膨大な情報量のために、ビデオをはじめとするアナログの映像機器の普及に対してコンピュータにおける利用は大幅に遅れている。人文科学とりわけ博物館では、本質的に単一のメディアでは取り扱えない文字・画像・音声・映像などを組み合わせたマルチメディアを対象としており、これらのデータは現在も着実に蓄積されるはや人手では管理できない状態になっている。放送業界から家庭に至るまでビデオ機器の普及することによって映像情報の重要性が社会的に認識されているにも関わらず、それを管理するための有効な手段は確立していない。映像の編集・再利用の立場から幾つかのアプローチが試みられているが[1, 2, 3]、映像情報を満足する形で管理するには至っていない。映像は、静止画および音声の集合から成り立っているがそれらの単体では意味をなさず、ある連続した時間を単位とするシーンの集合であると考えられる[4]。我々はそのような観点から、(1) 動画像からシーンを自動的に検出し、(2) 得られたシーン情報に対するインデックスの付与とデータベースの作成を対話的に行い、(3) ユーザーが目的とするシーンを検索するためのアプローチを提案し[5]、国立民族博物館(民博)の所有する映像資料を対象にプロトタイプを構築している。以下本システムを M-CIDB(Motion Color Image Database)と呼ぶことにする。

本稿ではまず映像データベースを構築する上での問題点について考察し、次にその問題解決を追求する過程で生まれたシーン情報を入力するためのインタフェースおよびシーン情報を管理するためのデータ構造と方法を示す。続いて M-CIDB の紹介と民博の保有する映像資料を対象とした予備試験の結果から明らかにされた現プロトタイプの問題点を挙げ、その解決方法について考察する。最後に以上の研究成果を踏まえ博物館における映像データベースの在り方について述べる。

2 映像データベースにおける問題点

近年映像データのソースの主体がフィルム(映画)からビデオに移行することによって、放送業界はもとより家庭内においても映像ソースの蓄積が加速的に進んでいる。これらの総数は既に本数にして100万本以上と推定されており、その時間的な総量は19万時間を越えると考えられている[5]。しかしフィルムやビデオテープなどの映像の管理は、もっぱらタイトル、収録時間、内容の概要といったラベル付けによって管理され収録庫に納められた。またラベルの管理についても文字ベースのデータベースによって行われ、実際にその内容物に関し

てユーザーが任意の部分を利用するためには人間の記憶力に頼るか膨大な労力をかけて逐一探し出さなければならない。一方 TV コマーシャルや番組の制作をはじめ、プロモショナルビデオの制作等映像を扱う様々な分野で過去に貯えられてきた映像の再利用が盛んに行われるようになり、必要な映像をより短時間に効率良く探し出す手段が求められてきている。また博物館など人文科学の分野でも、従来の文字や写真などの資料に加えて映像資料の価値が重視されるにつれて研究を進める上で必要な映像を探し出すための時間が無視できなくなってきており、映像を直接管理する映像データベースの重要性が増している。

映像すなわち動画像を扱うデータベースやオーサリングシステムに関してこれまでに幾つか研究が行われている[6, 7, 8, 9, 10]。その中で映像データベースに対して必要とされる要件として次のようなものが揚げられている。

- シーンとその属性を記述するためのデータ構造
- シーンやフレームをスムーズに検索する方法
- 直感的なシーンの検索を可能にするユーザーインタフェース
- 動画像と音声等他のデータの同期法
- 膨大なデータを扱うのに適するインデックスデータの作成
- 従来のシステムとの互換性を持つシステム設計

しかしこれらの課題に対する解答を与える研究はまだ十分に行われていない。そこで我々は以上の要件を満たしながら、特に映像データベースの構築時の問題点、すなわち

- シーン情報の入力(インデクシング)
- 対象シーンの検索

に注目して研究を行った。これらの問題の解決はシーン情報の入力を容易にすると共に、必要とするシーンを自在に効率よく探し出す上で不可欠な課題である。

3 シーン情報の入力

3.1 シーン情報構造化

映像データベースを構築するためにまず土台となるデータ構造を定義する。動画像は、2次元の静止画像からなるフレーム画像の連続集合として捉えられる[6]。各フレームには先頭から連番あるいは時間単位の位置を示すフレーム番号が付けられている。シーンは動画像全体の一部であり、連続したフレーム番号によって全体での位置を特定できる。そこで個々のシーンに対してフレー

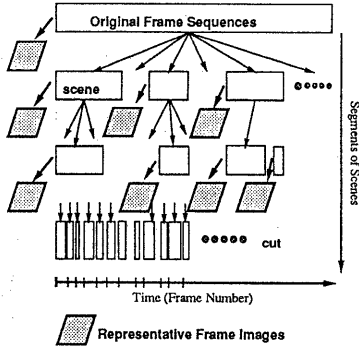


図 1: シーンの論理構造

ム情報を持たせると共にシーンを代表するフレーム画像とシーン属性を定義する。またシーンに対して階層構造を持たせることにより、シーンを単位とした動画の再構成や効率的な管理を行う。またフレーム情報とは別の時系列(フレーム)を基準とした内容に関する属性をキーワードとして記述する。これらはそれぞれ(1)フレーム情報、(2)シーン階層情報、(3)時系列キーワード情報をシーン情報として管理する。

3.1.1 フレーム情報

シーンは連続したフレームで構成されるから、その先頭と終了を示すフレーム番号の対によって特定される。個々のシーンに対してそれを代表する1つ又は複数のフレームを代表フレーム(群)として定義する。代表フレームはシーンを検索する手掛かりとして利用する。

3.1.2 シーン階層情報

シーンの論理構造は、図1に示されるように動画像はより小さい単位のシーンによって分割され、さらにそれぞれのシーンはより小さいシーンから成り立ち、最終的には物理的にこれ以上分割できないカットシーンにいたるまで構造的に捉えられる。そこでシーンの論理構造として(図2-a)のようなシーン階層を定義する。各ノードはシーンのフレーム情報を持ち一方親子関係を示す階層情報は枝に対して与える。シーンの属性を特定したり親子関係のリンクを走査することで対象とするシーンを検索できる。また階層の各枝には3.1.1で示した代表フレーム(群)が付けられており、階層間を移動して概覧するのに役立つ。フレーム情報がシーンの時間的な位置や順序だけを示すのに対して、シーンの階層情報は、シーン同士の間を独立した情報をもつため[8]、入力においてシーン情報の一貫性を保持するには特別な方法が必要である。これについては次章4.1で述べる。

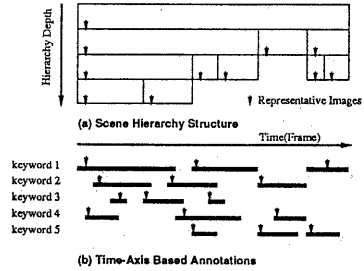


図 2: シーンの階層構造と時系列データ

Scene Attribute and Representative Frame Table

SID	SF	EF	SA	RF	Pr
R	s1	e1	---	f1	f1
A	s2	e2	---	f1	f2
B	s3	e3	---	f2	f3
C	s4	e4	---	f3	f4
A1	s5	e5	---	f4	f1
A2	s6	e6	---	f4	f2
A12	sn	en	---	f4	f4

* SID: Scene ID, SF: Start Frame, EF: End Frame, SA: Scene Attribute
RF: Representative Frame, Pr: Pointer

Hierarchy Table		Keyword Table		
PID	CID	Keyword	SF	EF
R	A	K1	s1	e1
R	B	K1	s2	e4
R	C	K1	s6	e6
A	A1	K2	s1	e3
---	---	-	-	-
A1	A12	---	---	---

* PID: Parent Scene ID, CID: Child Scene ID

図 3: 時系列キーワード情報テーブル

3.1.3 時系列キーワード情報

前節までは物理的なシーン単位を基準とした情報を扱ってきたが、シーンの意味的な内容を表わす属性は独立して存在する。例えば動画像に現れるある特定の物体、背景、事象に注目する時、これらがシーンの限られたフレームにのみ出現するだけでなく、異なるアングルで撮影される場合のように物理的に異なったシーンにまたがり登場する可能性がある(図2-b)。

図3はキーワードによってそのような事象や対象を代表させて、時系列(フレーム)上に配置した例を示している。このようにフレーム情報とは独立した時系列情報を各キーワードに与えることで特定の内容を持ったシーンを選びだすことが出来る。

3.1.4 シーン情報入力のインタフェース

シーン情報の入力によってユーザーが任意のシーンを探す労力は減少するが、入力するための労力は増大する。これに対しコンピュータが動画像の内容を抽出して自動的にシーン情報を入力することは困難であり、内容の意味を理解する一般的な方法は見つかっていない。シーン

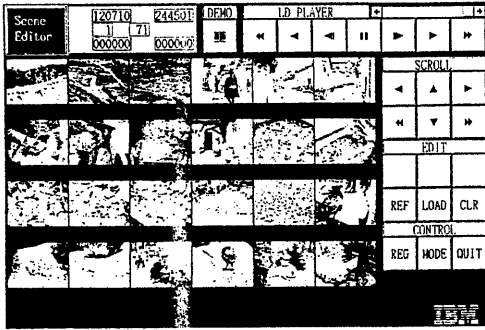


図 4: シーンディテクタの表示画面

に対するインデクシングには必ず人間による操作が介在するため [4]、対話的にシーン情報の入力を支援する方法が必要である。そこで映像データベース構築の基礎となるシーン情報の入力するためのインタフェースについて M-CIDB の例を示しながら説明する。

M-CIDB では、現在シーン情報入力に関して大きくわけて (1) シーンディテクタと (2) シーンエディタの二つのインタフェースを持っている。シーンディテクタによって各カットシーンの先頭を自動的に検出する。本プロトタイプでは検出するシーン変化はカット変化に限っている。また先頭フレームの静止画像はシーンの代表画像として VCA/A(ビデオ変換アダプター, IBM Co.) によりデジタル化され PC に取り込まれる。このとき同時にシーンのフレーム情報が記録される。進行の状況はリアルタイムに PC に表示される。検出終了後代表フレームの一覧をみながら、シーンエディタによりシーンの修正、キーワードの付与等のシーン情報入力を行う。

3.1.5 シーンディテクタ

シーン変化には物理的なものと意味的なものがあるが、後者が内容の認識による変化であるのに対して前者は撮影時のシーン変化(カット、ズーム、パニング等)と編集時のシーン変化(カット、デゾルブ、ワイプ等)の組み合わせに起因する変化である。シーンディテクタは、デジタル化されたフレーム画像の変化を輝度等の幾つの特徴量を追跡することによって物理的なシーン変化を検出し [11]、シーンの先頭フレーム番号を調べる。検出されたフレームの直前は 1 つ前のシーンの終了フレームに相当する。また先頭フレームは代表フレームのデフォルトとしてコンピューターに取り込まれる。途中経過は PC スクリーン上に表示される (図 4)。シーン変化が物理的な場合にはフレーム情報の入力を大幅に省力化できる。

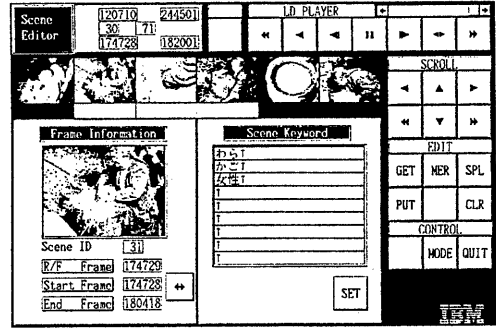


図 5: シーンエディタの表示画面

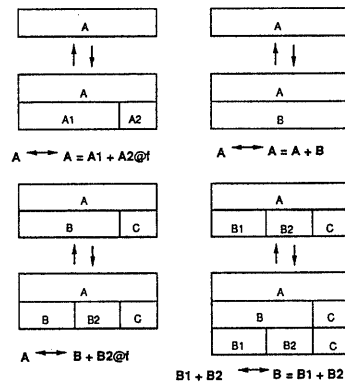


図 6: シーン階層の構造化

3.1.6 シーンエディター

シーンディテクタの出力結果には検出ミスや誤検知が含まれる可能性がある。またデフォルトで使用される代表画像もユーザーの要求と一致しない場合もある。したがってシーンディテクタによって得られたフレーム情報は最終的なものではなくユーザーによる確認と修正が必要である。シーンエディターは、視覚的ユーザーインタフェースによってユーザーが対話的にこれらの操作を行う環境を提供する。ディテクタが検出したシーンを修正するためにはシーンの合成(マージ)、分割(スプリット)を行うことができる。またシーンに対しキーワードの入力やシーンの階層構造を構築しシーン情報を入力する。シーンは代表画像をシンボルとするオブジェクトとして表示され(図 5)、ユーザーはすべての操作をオブジェクトに対する直接操作として行う。

	Frame Number					
Keyword 1	F	T	F	T	F	F
Keyword 2	F	F	F	T	T	F
Keyword 3	F	F	T	T	F	F
Keyword 4	F	T	F	F	F	T
Keyword 5	T	F	F	T	F	T
(K1 and K2) or K4	F	T	F	T	F	T

図 7: 真理表

4 シーン情報の管理

4.1 階層情報の管理

4.1.1 シーン階層の構造化

シーンは図 6 に示される、8 つの上下方向の合成と分割によって構造化される。図 6 の左上の例では、シーン A が分割操作によって A1, A2 の二つの子シーンを持つ階層が作られる。A2 はフレーム f を開始点とする。逆に A1, A2 に対して合成操作を行うことにより A1, A2 が削除されもとのシーン A に戻る。

4.1.2 階層構造の巡回

ある階層のノードから他の階層ノードへ移動するために、現在のノードに対し親・子・兄弟・子孫・前・後にあたるノードへ移動する 6 つの操作を組み合わせて階層間を任意に巡回する。

4.1.3 代表フレームの変更

階層構造の変更に伴って代表フレームは次のように自動的に変更される。

- (1) 分割によって新たな階層のシーンが作られる場合には、フレーム情報を比較することによってもとのシーンの代表フレームを含む方のシーンに代表フレーム番号を複写する。他方のシーンに対しては代表フレームを定義しない。
- (2) シーンの代表フレームが変更された場合には (1) の規則に従い、そのシーンの先祖・子孫のシーンに対して代表フレームを再定義する。

4.2 時系列キーワードの管理

キーワード情報は、キーワードと開始・終了フレーム番号の対からなるキーワードテーブルによって管理する(図 3)。シーンの検索には問い合わせの各キーワードについてフレーム番号にしたがってキーワードの含まれるフレームを真、含まれないフレームを偽として縦にキーワードを横にフレーム番号をとった真理表を作る(図 7)。

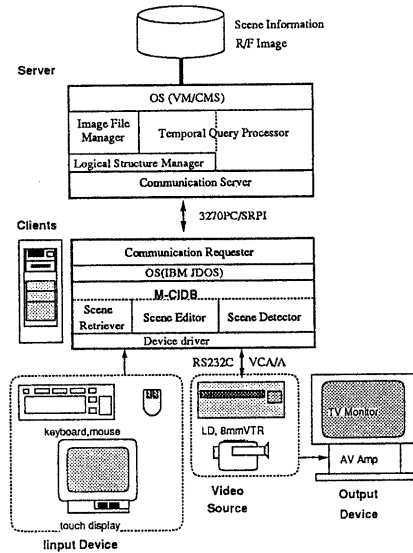


図 8: M-CIDB システム構成図

与えられたキーワードに対応した真理表に対して、論理積 (AND)、論理和 (OR)、否定 (NOT) を組み合わせた論理演算を行うことによって、指定キーワードの含まれる時系列(フレーム)を得られる。この検索は次の手順によって行う。

- (1) 指定キーワードを持つレコードをキーワードテーブルから探す。
- (2) 開始フレームによってレコードを昇順に並び変えたテーブルを作成する。
- (3) 1レコードづつテーブルを読みだし、時系列にそって真理表を展開する。
- (4) 展開されたフレーム区間に対し検索条件の真偽値を評価する。
- (5) (3) にもどって操作を繰り返す。

5 システム構成とプロトタイプ

M-CIDB は PC 側で (1) シーンディテクタによってビデオ機器の制御と画像およびフレーム情報の入力を行い、(2) シーンエディタによってシーン情報の入力を行う。大型計算機(ホスト)側では (3) 静止画像とフレーム・キーワード情報からなるシーン情報を基に画像・テキストデータの一括管理を行う環境を提供する(図 8)。

我々はプロトタイプの構築にあたり次の点に留意した。

- 対象とする動画は NTSC 規格のビデオ映像とし毎秒 30 フレームの精度で管理を行う

デジタルメモリーを備えたLDプレーヤーによってタイムコードを用いたフレーム情報の管理を行った。動画データの取り扱いは、NTSCのビデオ画像をデジタル化して管理する方法とアナログのまま管理する方法がある。前者に対して(1)MPEGやRTV/DVI(Intel.Co)技術を用いた完全なデジタル化の方式、一方後者に対し(2)M-Motion(IBM Co.)等のビデオ入力ボードによってコンピュータディスプレイ上でビデオ画像の合成表示する方式と(3)TVモニター上でビデオ画像を表示し他の情報をコンピュータで扱うハイブリッド方式が考えられる。M-CIDBではプロトタイプ構築の容易さ、操作性、またPC側のコストを抑えるために(3)のハイブリッド方式を採用した。

- 画像データはPC上にいったん蓄積しシーン情報の入力後ホストに一括転送し管理する

ハイブリッド方式では代表画像のみ画像データとして管理するため、(1)のデジタル方式に比べデータ容量は少ない。したがってシーンディテクタによって取り込まれた画像もPCのハードディスクに蓄積できるため、シーン情報の入力はPC上で行いその後一括してホストへ転送し管理する。

- スクリプト言語によるウインドウシステムを使用しユーザーインターフェースの統一とシステムの拡張を容易にする

ユーザーが出来る限り少ない学習で操作するために画面表示を統一しなければならない。またプロトタイプによるシステムの構築で改良が短時間で簡単に行えるためにはスクリプト言語(ハイパーカードのハイパートークなど)によるウインドウシステムが望ましい。本プロトタイプではハイパーリトリーバー(IBM Co.)を使用した。

- 操作によってパネルを変え必要最少限の情報を表示するとともに使用できる機能を視覚化する

狭いスクリーンに大量の情報を一度に表示することはかえってユーザーを惑わせ操作性を低下させる。操作に応じて必要最少限の情報を示すとともにその時点で使用できる機能をユーザーに明示する[12]。

- タッチパネルを前提として操作ボタンの大きさと位置や操作性を考慮する

常時システムを使用するデータの管理者に対して、一時的に使用する一般ユーザーは入力手段としてマウスよりもタッチパネルの方が直感的に使用できる。操作ボタンが小さすぎるとスクリーン上での指示が困難になるし、大きすぎると他の表示領域が小さくなる。そこでディスプレイのサイズにあわせて操作ボタンの大きさ位置を決めタッチパネルでの操作性に留意した。

6 プロトタイプの子備試験と考察

本研究は主として人文科学系のエンドユーザーを対象とするために、文献[13]で明らかにされた問題を解決する立場から進めている。ユーザーインターフェースの設計にあたり文献[14]で試された繰り返し設計法を用いてシステムの改善を図るための予備試験を行った。使用したデータは民博より提供されたビデオテープ(来館者用のビデオ番組)4タイトルをCLVフォーマットのLDの形式で一枚にまとめたものである。本データは52分・274カットシーンからなり802レコードのキーワードテーブルからなっている。カットシーンの長さは平均11秒でシーンディテクタによる検出率は95%程度であった。

6.1 シーン情報入力時間の測定

まず手作業でシーン情報の入力を行い作業時間を見積もり、次にプロトタイプを用いて測定し手作業に比べて時間短縮がどのくらい行われたか測定した。シーンの検出や画像の取り込みはCPUの能力に依存するためおおよその目安である。

手作業でのシーン情報の入力は次の通り行った。シーンの概観と位置の確認のため(1)全体を普通再生し(2)シーンの開始位置まで1/2倍速再生する。次にシーン情報の入力のため(3)カットシーンの先頭・終末で3秒のコマ送り再生し(4)代表画像の取り込みと(5)フレーム情報の記録を行い、最後に作業の確認のため(6)全体を倍速再生した。合計時間は15時間51分(=A)要した。

M-CIDBを用いた場合は次の通りである。シーンディテクタによって(1)シーン自動検出とフレーム情報・代表画像の入力を行い、次にシーンディテクタによって作業の修正を(2)シーンの選択・スクロール、(3)各シーン1/2倍速再生、(4)誤ったシーンの修正と代表画像の変更の順に行い最後に(3)各シーンを再生して作業を確認した。合計時間は11時間28分(=B)でうち人手によるものは2時間58分(=C)であった。

AとBを比較することでおよそ全体で1.4倍の時間短縮が図られていることがわかるが、実際には人間が直接作業した時間を比較すべきである。そこでAとCを比較すると4.9倍時間の短縮を実現したことになる。現実には連続して作業することは出来ないで細かい休憩時間を含めると全体で2-3倍、手作業で6倍以上の効率の向上が期待される。

6.2 シーン検出率

3.2.1で述べたようにM-CIDBのシーンディテクタは、現在カット変化を対象にしているが輝度の変化を追跡するために、(1)動きの非常に激しいシーンに対しては実際のカット変化よりも誤検知してしたり、(2)全体の輝度が小さい(背景が暗い)シーンにたいしては検出できない場合がある。しかしこれらのシーン変化は意味的に

捉えなければ分けられない場合もあり必ずしもシーンエディタの検出ミスとは言えずデータによって調整が必要である。

6.3 シーン情報入力操作の評価

民博の協力を得て、M-CIDBの予備試験を行った結果について報告する。試験したのはシーンディテクタとシーンエディターである。映像資料や標本資料の作成に携わるスタッフ4名を被験者とした。このうち映像資料係は専門家であり映像資料の分類・管理に習熟しているが、標本資料係は一般ユーザーとして映像資料を利用する。

シーンディテクタについては、限られたデータでの試験ではあったが民博において実際の業務として行われている「カット割り作業に有効である」との意見を得た。一方シーンエディタに関しては、映像資料係と標本資料係のスタッフで意見が分れた。

映像資料の立場からは、

- キーワードを含め内容記述の全自動化が必要である
- カットとシーンでは内容記述の対象が異なるため使用するキーワードも異なる
- 代表フレームだけでは動きの情報がわからない
- 複数の代表をフレームを用いたりフレームを間引いた動画再生は動きの内容を知る上では効果が不明である
- 作業画面をよりシンプルにして余計な表示をしない

標本資料の立場からは、

- 代表フレームはカット内容を思い出す上で十分効果がある
- キーワードの入力では他のシーンに付けられたキーワードを検索して複写する操作が必要である
- キーワードのマクロ・エイリアスによって冗長な入力操作を省く
- 作業モードを細かく設定したい
- キーワード情報を画像によって確認できるほうが良い

など特にキーワードの入力と操作ボタンの表示について問題点が指摘された。現在のプロトタイプではまだ十分に実現できない機能があるとはいえ、キーワードなどの文字情報をいかに少ない操作で入力するかという点に現場からの強い要望が感じられた。またインターフェースに関しては、専門家(映像資料係)と一般のユーザー

(標本資料係)の間で意見が違いがあった。作業の効率化を追求する専門家にとっては1つのパネルに1つの情報を表示・編集する方が望ましいが、あまり操作に慣れていない一般ユーザーにとっては必要な情報を表示して現在行っている作業状況が常に確認できる環境が求められていることがわかった。

6.4 これからの課題とその解決

以上のような予備試験の中で明らかになった課題をもとに次のプロトタイプ構築に向けて現在次のような解決案を考えている。

シーンディテクタに関してはアルゴリズムの改良により(1)検出率の向上と同時に、(2)検出可能なシーン変化の種類をズーミング・パニングなどに対応させて増やす必要がある。現在は使用されるデータやハードウェアの性能によって検出率が変化するため、今後もより多くの種類のデータに対して実験を行いながら改善してゆく。

シーンエディタに関しては(1)キーワードの入力と表示方法の改良、(2)表示する情報や操作機能をユーザーに応じてカスタマイズできるようにする。(1)についてはマクロやエイリアスを使用したり、キーワードをアイコンで表示することによってキーボード入力を減らして操作性を向上させる。またシーン間での複写など冗長な入力操作の繰り返しを省く。(2)に対しては当面映像資料の専門家を対象として、必要な機能を限定して単一機能の操作性の向上を目指す。最終的にはユーザーに対応して視覚的なインターフェースをカスタマイズできる柔軟なシステムを目指す。

またシーンの検索についても多くの問題が残っているが、上に挙げた解決も含めて今後の課題とし機会を改めて紹介したい。

7 むすび

本稿の中で我々は、データを入力する立場すなわち情報管理者の観点に立ち映像データベースについて論じてきたが、今後のシステムの発展を踏まえて最後に博物館にとつての映像データベースの在り方について考える。文献[15]で論じられているように、博物館の機能は(1)資料の収蔵施設、(2) 学術研究の場、(3) 資料の公開の場の3つの機能を持つと考えられる。(1)の機能が正しく働くことによってはじめて資料そのものが価値を持つようになり、研究に有効に活用されその結果得られる成果が資料と共に一般に公開され博物館が社会的に機能する。したがってこのサイクルを有効に発揮させるにはまず(1)の機能が満たされなくてはならない。コンピュータによる資料の情報化によって、例えば民博においては多くの標本資料が静止画像化され博物館の機能は飛躍的に高まったが[16]、映像資料に対してはその

膨大な情報量により未だに(1)の機能さえも十分に發揮出来ていない。そこで何よりも映像資料を収蔵管理するためのデータベースが必要となる。そのためには次のような段階を経過することが期待される。

第一段階として映像資料にあったデータの入力と整理を行う。ここでは映像資料の管理者が最も効率的に入力を行えるシステムの整備が必要である。ユーザーインタフェースはなるべく単純にして必要最小限の機能を持たせる。

第二段階は入力作業と平行して、研究者が自分の必要とする資料を自由に検索し結果を映像以外に様々な形で利用できる環境が望ましい。高度な検索を行えると同時に様々な情報を視覚的インタフェースによって容易に引き出せるシステムを構築する。また各研究者が自分専用のデータベースを作るための機能の拡張性も必要となる。

第三段階は研究者や学芸員が、個人用データベースからデータを抜き出し教材として、あるいは一般者への展示公開を目的とするハイパーメディアを作るためのシステムである。それには、マルチメディアを前提としたシステム間のデータの互換性や親しみやすいユーザーインタフェースなど新たな問題を解決しなければならない。

以上の三つの段階を経て、博物館における映像データベースの基礎が完成するであろう。まだ研究は始まったばかりであるが映像データは日々刻々と蓄積されている。今後は映像データを中心としたマルチメディア化を目標にプロトタイプを構築する中で課題を一つ一つ解決しながら一日も早く実際の博物館に成果を活かしてゆきたい。

8 謝辞

本稿は、国立民族博物館(民博)と日本IBM(株)との共同研究「博物館におけるマルチメディアの有効利用」の一環として実施している研究成果の一部であり、使用されたデータはすべて民博に所属する。ご指導いただいた民博の杉田繁治教授、データの準備等の協力をいただいた民博の鈴木明・中川隆・宇治谷恵諸氏および、シーン検出のアルゴリズムを担当された当研究所・井岡幹博氏に対し感謝いたします。

参考文献

- [1] Narashima, A. Desai and Christodoulkis, S.(eds.), "Multimedia Information System," IEEE Computer, vol.24, no.10, 1991.
- [2] Fox, E. A. (ed.), "Interactive Technologies," Communications of the ACM, vol.32, no.7, 1989.
- [3] Fox, E. A. (ed.), "Digital Multimedia Systems," Communications of the ACM, vol.34, no.4, 1991.
- [4] Kawagoe, K., "Continuous Media Data Management," ACM SIGMOD, Records, vol.20, no.3, pp.74-75, 1991.
- [5] Hong, J-K., Takahashi, J. and Kusaba, M., "A Motion Archiving Technology and its Application in an Ethnology Museum," 3rd Int'l Conf. Database and Exp. Sys. Applications (DEXA92), Sep.2-4, 1992, Valencia, Spain.
- [6] 大喜多, 有澤, "意味データベースに基づく映像情報のモデリング," 電子情報通信学会, データベースシンポジウム, pp.7-16, Feb., 1990.
- [7] 外村, 阿部, "動画像データベースハンドリングに関する検討," 信学技報, IE89-33, pp.49-56, 1989.
- [8] 上田, "インタラクティブな動画編集方式の提案," 信学技報, IE90-56, pp.39-46, 1990.
- [9] 池田, 中村, 内山, "動画を含むハイパーメディア管理システムの構築技法とアプリケーション," 信学技報, DE89-32, pp.25-32, 1989.
- [10] 大友, 渥美, 田中, "ビデオデータベースの概念モデリングとビジュアルインターフェイスについて," 信学技報, DE89-50, pp.17-24, 1990.
- [11] Ioka, M., "Method of detecting scene changes in moving picture," IBM Technical Disclosure Bul., Vol.34, No.10A, March, 1992.
- [12] 笠原, 岸本, "画像データベースと視考支援環境," 情処研報, 89-CH-3, 1989.
- [13] 杉田, "人文科学とコンピュータ," 情処研報, 89-CH-1, 1989.
- [14] 橋原, 洪, 杉田, "繰り返し設計法によるユーザ・インタフェースの向上性," 情処研報, 89-CH-2, 1989.
- [15] 宇治谷, "博物館資料の整理と活用におけるコンピュータの役割," 情処研報, 90-CH-4, 1990.
- [16] Hong, J-K. and Sugita, S., "A color image database for an ethnology museum," in Computers in the Humanities and the Social Sciences, pp.53-60, SAUR, München, 1991.