

# AVON : サイバブックに “ 編み込んだ ” インタラクティブ映像本

外村 佳伸、 小澤 英昭、 谷口 行信、 南 憲一

NTTヒューマンインタフェース研究所

横須賀市光の丘 1-1

e-mail: tonomura@aether.hil.ntt.co.jp

概要：本型インタフェース（ブックメタファ）を用いた情報環境に映像情報を ‘ 編纂 ’ することにより、映像にインタラクティブな利用環境を提供する映像の電子本「AVON」( Abstracted Video On Network )を提案する。AVONの主要な特徴は、本型の利用環境プラットフォームであるサイバブックにおける身近で使いやすい情報環境を映像に対して利用できること、映像解析による特徴情報の構造化により自動的に作成できることである。従来の映像利用インタフェースに比べると、映像が内在する情報を利用している点で、表現およびアクセス粒度の階層性を始め、効率的、効果的なインタラクティブ環境を提供できる。本報告では、サイバブックさらに AVON におけるインタラクティブ性の考え方、AVON 作成処理体系、実際の応用例について述べる。

## AVON: Interactive Video Book by ‘Compiling’ Video in CyberBook

Yoshinobu Tonomura, Hideaki Ozawa, Yukinobu Taniguchi, and Kenichi Minami

NTT Human Interface Laboratories

1-1 Hikarino-oka, Yokosuka, Kanagawa, Japan

e-mail: tonomura@aether.hil.ntt.co.jp

Abstract: This paper proposes an electronic video book called “AVON”, Abstracted Video on Network, made by ‘compiling’ video into the CyberBook. Main features of AVON are that familiar and easy-to-use interactive interface on the CyberBook can be utilized and that the book can be created automatically based on video feature analysis. The AVON realizes an efficient and effective interactive video interface such as browsing with hierarchical granularity. The concept of interactivity in CyberBook and AVON, processing architecture, and examples of AVON are presented.

### 1 . はじめに

映像など時間軸を持つ連続メディアを、その特徴を十分に生かし、高いインタラクティブ性を持って利用できる道具はまだない。映像はテキストに比べると扱える自由度が低く、使い勝ってもよくない。

そこで本論文では、映像に高いインタラクティブ性を与えることのできるインタフェース候補としてブックメタファに着目し、著者らが開発したマルチメディア情報プラットフォームとしての電子本「サイバブック」上に、映像作品を自動的に編纂した映像本「AVON」を提案する。本論文では、連続メディアにインタラクティブ性を持たせるための構造化情報を基本とした処理体系、電子本のインタラクティブ性、映像本の特徴と作成処理、応用例等について述べる。

### 2 . メディアの構造化とインタラクティブ性

#### 2 . 1 映像ハンドリング

レーザーディスク等の再生制御から始まったコンピュータによる映像ハンドリング(1)においては、映像がデジタル化され、映像圧縮符号化の標準化、OSレベルでの映像ファイルの扱いなどが進展し、ある程度の量のデータ蓄積が可能になった最近、ようやく本格的な利用が想定されるようになってきた。

初期のデジタル映像ハンドリングでは、映像の効率的な蓄積、伝送、表示、再生等の基本機能を、アプリケーションレベルで直接利用することが中心であった。これに対し最近では、ブラウズ、編集、検索、インデクシングなど、基本処理環境としてのハンドリング技術に関しさまざまな試みが行われている(2)(3)(4)(5)(6)(7)。しかしながら、まだ想定利用環境が特殊で、身近な利用側に立ったとき、具体的にどのように役立つ

かについては課題が多い。そこで我々は、もっと身近な観点で映像を利用するために必要な道具、環境としてのインタフェースについて検討し、その一つの解としてブックメタファ(8)の利用を提案する。

## 2.2 連続メディアの構造化

メディアを上手に扱うには、そのメディアが含む情報に関する手がかりが必要である。構造の少ない平板な情報を、扱いが便利になるよう手がかりを付加し特化していくのが構造化である(2)。一般文書の場合には文字がコンピュータ可読になっていることや、今までの情報処理がテキスト中心で発展してきたことから、検索、フィルタリング、意味解析、生成など実用的な面でもテキストについては研究が進んでいる。これに対し特に音声や映像など、いわゆる時間軸を持つ連続メディアは、コンピュータ上の扱いの歴史が短いこともあり、各種利用に有効な手がかりの発見、構造化による利用はまだ不十分である。

構造化処理の基本は、カット点検出等、映像ハンドリングの基本を定義するセグメンテーション(9)(10)(11)、さまざまな特徴情報を抽出し、インデクスとして記述、利用する技術である(12)。どのような情報をインデクスとして記述できれば、どう映像が扱いやすくなるか、またそのインデクスはどう生成するかが、使える映像ハンドリングのための大きな課題である。映像制作の立場では、映像に付加可能な情報として、映像の意味レベル、表現の意図レベルをハンドリングする研究もあるが(13)、今回の我々の報告では、身近で汎用的な利用を想定し、主に特徴レベルを検討する。(図1)

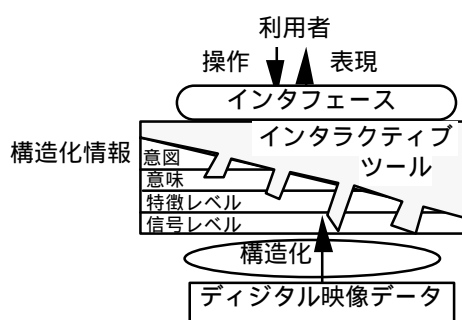


図1 構造化映像ハンドリング

## 2.3 インタラクティブ性付与とハンドリング粒度

映像利用にインタラクティブ性を付与し、さらに高度化するためには、利用方法に照らした情報の構造化、および構造化情報をうまくインタラクティブに利用するためのツールが必要となる。きめ細かなインタラクションを可能にするためには、きめ細かな構造化が必

要となる。主なインタラクティブ性には、例えばアクセス、表現制御、加工・編集に関するものがある。

情報の構造化とインタラクティブ性に関して最も基本的な点は、扱える情報の単位である。理想的には目的により、マイクロ～マクロ、固定的～可变的、連続～不連続と自在に扱えるのが望ましい。テキストで言えば文字、単語、段落、節、章などの単位が操作対象となる。これが音声や映像の場合、その連続性、多義性のために、扱う情報の単位が一意に決まりにくい。我々の処理系では、映像等の比較的汎用的な特徴情報から、利用形に応じた情報の単位を定義する。操作上で扱う情報の粗さ(あるいは細かさ)をハンドリングの粒度(グラニュラリティー)(14)と呼ぶ。我々は、当面数段階の情報単位の粒度で映像を自由に扱えることを目標としている。

## 3. インタラクティブ電子本「サイバーブック」

サイバーブックは、本型のインタフェース(ブックメタファ)を用いた電子本であり、わかりやすく簡単に使えるインタラクティブな情報利用環境として、特にインターネット上のマルチメディア情報利用に威力を発揮する。

### 3.1 ブックメタファインタフェース

構造化したマルチメディア情報を利用する基本環境として、歴史的に培われてきた本型の使い勝手と、電子手段ならではの特徴を生かすよう設計したインタフェースである。ブックメタファを利用する電子本の研究例は、多くはないが以前から行われてきた(15)(16)。しかしWWW等の機能面重視のインタフェースに押された感があった。最近になって、インターネットを経由して電子的な本を3次元的なグラフィクスでリアルに表現する例(17)、本を電子的に出版する立場から、表現上の豊富な機能を有するソフトツールも市販されている(18)。これらに対し、我々の目的はむしろ、マルチメディア情報利用の観点から、作成・流通・利用(特にブラウザとして)の流れを考慮した機能、シンプルな使い易さの実現に主眼を置いている(19)。

### 3.2 サイバーブックのインタラクティブ性

サイバーブックのマルチメディア情報利用に適した重要な特徴、機能には、まず直接アクセス操作に関し、

- ・順次アクセス：ページ制御による順次アクセス
  - ・ランダムアクセス：任意のページを開ける
  - ・リンクアクセス：特定情報に直接ジャンプする
- の3つのアクセスモードを基本とし、さらにこれを支える機能として、

- (1)ページ閲覧制御機能：ページ送り / 戻しなど
- (2)ページ厚み表現：厚み感から勘でページを開く
- (3)情報タグ付加機能：付箋付加と付箋指示ジャンプなど、紙の本にならった機能を有する。また、
- (4)目次機能：生成とジャンプ
- (5)索引機能：生成とジャンプ

などリンク生成によって即該当ページへジャンプできる機能は電子ならではである。また、人間の空間的な認知処理に関わる情報表現機能として、

- (6)情報全体量呈示：厚みによる情報量の全体感提供
- (7)現在位置呈示：全体の中での利用者視点等を備えている。更にコンピュータ利用の特性として、
- (8)映像・音声等マルチメディア利用機能
- (9)ハイパーリンク機能：任意点間のリンク設定を備える。

### 3.3 サイバーブックの生成

サイバーブックを作成するには、専用のオーサリングツールを用いて、WYSIWYG で作成する方法と、HTML 等から自動的に変換・作成する方法がある。

通常インターネット上で発信されている HTML ベースの情報を、専用の記述言語の BMDL (BookMetaphor Description Language) (19) に自動変換してサイバーブックで見られることもできる。スクロールバー操作で情報の位置が変わり、複雑なリンクで画面の切り替えが頻繁に生じる汎用 Web ブラウザと比べると、サイバーブックでは情報が視野内でページ単位で表現されるため、情報の位置が直感的に分かり、記憶として後のアクセスに役立つ点、情報の全体量、現在位置が厚みでわかる点、等々が絡んで、操作・内容把握上効率のよい環境であることが評価実験の結果でもわかっている(20)。

### 3.4 ブックウェア

ブックメタファは単に情報の本格的表現によるのみ有用なのではなく、さらに情報利用上の周辺環境と相まって便利になる。我々はこうした周辺も含めてブックウェア (BookWare) と呼んでいる(21)。ブックウェアにおいて、本以外の重要な要素には本棚がある。我々の設計している本棚 (InterShelf) (図2) は、インターネット情報をパーソナルな環境として目に見える形で可視化する役割と同時に、利用者が利用しやすいように、好きな場所に情報を本として置いておく位置記憶支援の実現、必要な情報を持つ本の自動配置の他、本棚をブラウズするために手がかりとなる情報特性の表現 (色、大きさ、柄、見出し・・・) などを重視して設計している。



図2 ブックウェアにおける本棚

## 4. AVON (映本)

AVON (Abstracted Video On Network) は、上記サイバーブックを基本環境として、ビデオ (映像作品) から自動的に作るビデオ抜粋本である。

### 4.1 映像を“編み込む”

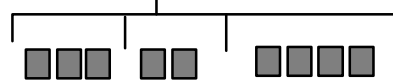
AVON 作成では、元のビデオを解析し、抽出された代表画像および特徴情報をインデクスとして記述し、それらを使ってあるルールでサイバーブックのページにレイアウトする (図3)。作成処理は、以下に示す MediaFarm と呼ぶ処理体系の中で行う。

#### ●ビデオシーケンス



特徴抽出、セグメンテーション、特徴レベル構造化

#### ●インデクス



#### ●AVON

ルールに基づく編さん (選択、加工、レイアウト)

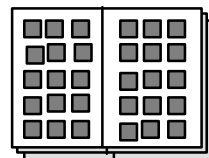


図3 AVONの作成概要

### 4.2 MediaFarm：付加価値型メディア処理体系

処理環境基本処理体系 MeidaFarm は、特に特徴情報レベルの情報処理を基本に、元の情報にさまざまな

付加価値をつけることを実現する環境である(図4)。

MediaFarmの主要構成要素は、マルチメディアデータを解析して特徴レベルのインデクス記述を生成するMediaParser、マルチメディアデータおよび付加情報を格納するMediaDB、MediaDBの情報をを使って利用上意味のある形にするMediaProducerおよびその結果を格納するMediaProductDB、そして最終的に利用するクライアント環境としてのMediaFront(情報ビューア等)である。

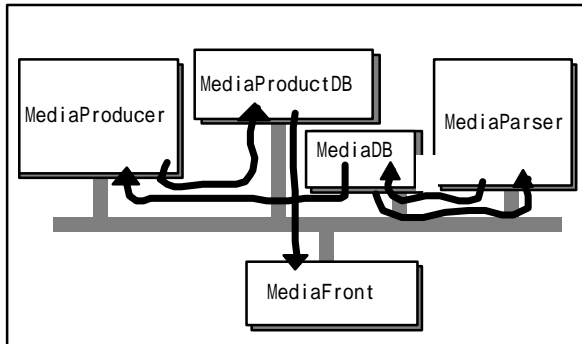


図4 AVON作成環境MediaFarmの処理体系

(1) MediaParser: マルチメディアインデクシング

この処理ノードにおける主要処理は、データからの特徴情報の自動抽出である。

特徴情報自動抽出処理: カット点検出、カメラ操作情報抽出、キャプション区間検出、音声区間検出、音楽区間検出を中心に、それらのもととなる特徴情報抽出とともに進行。カット点検出に関しては、我々を含め今までさまざまな試みが行われているが(10)(2)(9)

(22)、我々自身の手法を総合して、瞬時カットはもちろん、フラッシュ光に強く、ディゾルブ等の漸次変化もとらえる方式(23)を用いている。カメラ操作情報抽出に関しては、被写画像のマクロな動きを調べ、まずカメラ操作が安定な区間を検出し、その区間についてはカメラ操作情報の種類や程度をインデクスとする(23)。キャプション区間検出は、一定時間以上静止状態にある、ある条件内の領域をとらえることを行っている。音声および音楽区間検出については、まず音声スペクトラム上で安定したピークを持つ区間を検出することで音楽の存在区間を検出し、次に音楽スペクトラムを除いた中から音声特有のハーモニック構造を検出することで音声区間を検出する(24)。

インデクス: 抽出した特徴情報を、後の処理が容易な形でインデクスとして記述しておく。特に昨今のインターネットでの利用を考慮してHTML形式での記述での出力を可能としている。インデクス内容として

は、抽出イベント(例えばカット点)とそれに対応する区間情報、付加的な属性などを、複雑な構造を持たさないで、フラットな形で出力する。

(2) MediaProducer: マルチメディア情報の“編纂”

対象とするマルチメディアデータ自身(映像等)、およびMediaParserによって抽出されたインデクスを利用して、目的に合うよう構造的に“編纂(Compile)”する。対象データに対し、論理的な関係と表現上のレイアウト関係を中心にアレンジする。この処理に必要なルール(条件)、例えば対象情報の絞り込み、順序決め、レイアウトのルールを目的に応じてあらかじめ作成しておき、処理時に選択する。特徴情報のみを基本とした比較的単純なルールによる処理から、特別な記述上のタグを定義、挿入して、それを元に複雑な処理を行うルールなどの拡張も考えることができる。処理結果はMediaProductDBに格納するが、このとき最終利用系に即した記述形式で格納する。例えば、サイバースタックでは専用記述言語BMDLで記述している。

4.3 AVONのインタラクティブ映像インタフェース

AVONはサイバースタック環境上で表現されているため、サイバースタックで実現している基本インタフェース機能が利用可能である。例えば、ページを進めたり、戻したり、ぱらぱらめくりなどは当然可能である。これらの機能は本として見ると自然な機能であるが、映像に関する機能として見ると、従来の映像利用インタフェースにはない非常に重要な機能が実現されていることがわかる(図5)。

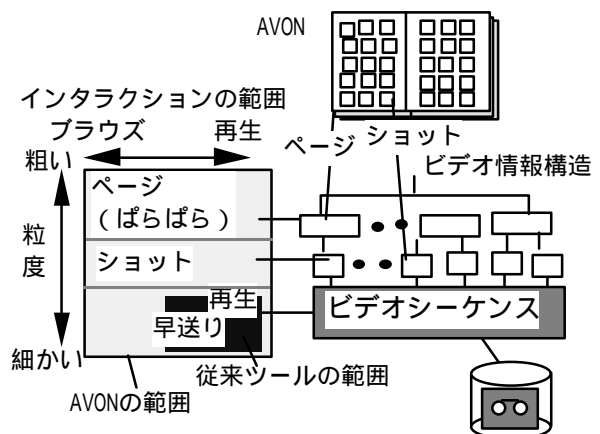


図5 AVONのインタラクティブインタフェース

(1) 階層的ブラウジング・アクセスインタフェース

映像内容を把握しようとする際に課題となるのは、時間的に(しかも長時間)内容の変化する映像をいかに要約して見せるかである。これを幅広い種類の映像

に対して汎用的に実現するために特徴レベルで実現するのが我々の目的である。まず何らかの観点で映像をセグメント化して各セグメントを代表画像で表すことが有効である。このため従来カット点で映像を抽出して並べることが行われてきたが(2)、長時間映像に関しては画像の数も多くなり、より少ない数で代表するべく工夫が試みられてきた(25)。これに対し、AVONでは代表画像をすくなくすることにこだわらず、むしろ数の多さを積極的に利用する。

AVONでは、ある数の代表画像の集まりで構成されるページ単位の切れ目が入ることで、より粗い粒度(グラニュアリティ)でブラウズし、アクセスすることができる。一方で、各ページ内の代表画像で表される対応映像は、そのアイコンをクリックするだけで再生される。従って、本棚から映像をアクセスする手段として考えると、その粒度は、本単位/ページ/代表画像/映像の4段階で実現される。例えば、ショットの数で各段階の粒度を見ると、10冊ほどAVONが並んでいるとし、1冊約300ショットの例で考えると、3000/300/18/1ショットの階層的ブラウジング・アクセスインタフェースが実現されている。これが、最初に述べたインタラクティブ性の付与に関する粒度の制御の実現例のひとつである。

#### (2) 映像情報位置の固定化と利用者マーク

ページ単位で付けることのできる利用者タグ(付箋)や、ページ内の任意の位置に貼れるメモ書きは、映像に対する利用者のマーキングとなる。また、これらの情報は、クライアント側の情報として保存できるため、元の情報が特に変わることなく多くの利用者間共有できる。AVONをビデオライブラリー等に利用することを想定した場合、よるこうした機能は利用者にとって大変便利なものになるが、従来の映像利用環境では操作対象の処理単位が定義できず実現し得なかった。

AVONでは、時間的に展開される映像を空間的に広げると同時に、展開の範囲をページに区切ってさらに重ねることで、扱う情報の空間的位置関係を固定していることになる。このように連続メディアを扱うには、何らかの不連続性(この場合さらに段階的不連続性)を持ち込み、位置関係を固定することが、利用者側での認知的な情報の定位を実現するためのインタフェースとして重要である。

#### (3) 映像の多義的インタフェース

映像の見せるものには、本来多義性があり、人によって、場合によって持つ意味内容が変化する。そこで

AVONでは、どのような側面で映像を把握するかによって、さまざまな編纂の仕方ができると考えている。例えば、次に示すような異なるタイプの本を編纂することが可能である。

#### 4.3 AVON ア・ラ・カルト

以下で示すさまざまなタイプの表現による本を、特徴情報のフィルタリングや表現の前処理、レイアウト処理のルールの違いで創り出す。

##### (1) 単純ショットシーケンス本

カット点毎に抽出される代表画像を時間系列に並べたことをレイアウトの基本とする。これは最も単純なタイプで、場合によっては代表画像が必ずしも代表にならないケースもあるが、全体的に映像内容がわかればよい目的に対しては、何より処理が簡単で、レイアウト上の充填率も高い。(図6)

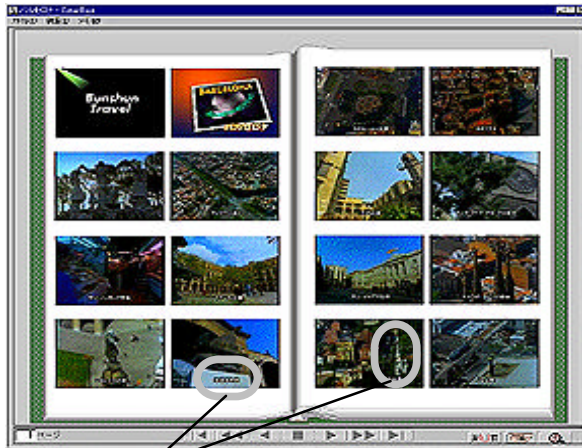


図6 AVON:単純ショットシーケンス

##### (2) キャプションシーケンス本

検出されるキャプションの存在区間毎に抽出される代表画像を時系列で並べたもの。解説的なビデオや字幕入りの映画等では内容を具体的に表す抜粋となる。対象とする映像のキャプションの付けられ方によっては非常に効率的な要約となる。(図7)





キャプション  
 図7 AVON : キャプションシーケンス本

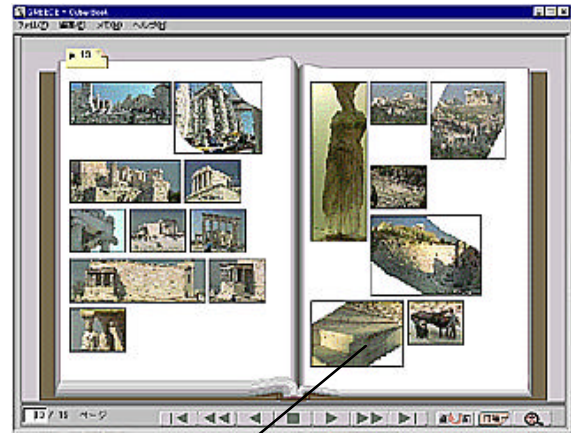
(3) パノラマ表現型ショットシーケンス本

カメラ操作が安定で、操作内容がパン、チルト、ズームあるいはそれらの組み合わせであると判断される区間に対しては、パノラマ画像を作成して代表画像とするものである。それ以外の区間には単純ショットシーケンス本と同じであり、各ショットの先頭などの代表画像(キーフレーム)を選び、時系列でパノラマ画像による代表画像とともにレイアウトする。(図8)

レイアウト上では、各代表画像が矩形に限らないため、時系列性を保存したまま、しかもページ上の表現空間が有効に利用されるためのレイアウトアルゴリズムを開発した。この表現をパノラマ画像を含んだ抜粋という意味でPnoramaExcerptsと呼んでいる。(23) 落ち着いたカメラ操作の多い旅行・紀行もののビデオなどには、パノラマ表現の合う被写対象の景色がよく現れるため、大変効果が高い。

(4) 音楽本 / 会話本

音解析結果により抽出される音声区間あるいは音楽区間毎に代表画像を抽出し、時間系列に並べるものである。対話ばかり、あるいは音楽ばかりを集めた映像の本ができる。対象とする映像内容が音楽ビデオなどの利用インタフェースには効果が高い。



ショットのパノラマ表現  
 図8 AVON : パノラマ表現型ショットシーケンス (PanoramaExcerpts)

以上紹介したものは、インデックスの基本形を比較的素直に用いるものであるが、この他にも、映像内容、利用目的をふまえていろいろな処理ルールを考えることで、ユニークで役立つAVONを創ることができる。

5. 考察

(1) 評価

AVONの最も大きな特徴は、今までの映像利用にないインタラクティブ性を自然なインタフェースとして実現していることと、こうした映像の抜粋による電子本が自動的に作成できることである。これが実用上どれだけの利用効果を生むかは、現在実際のニーズに基づく検討を行っており、その評価を待たねばならないが、少なくとも従来なかったインタフェース機能を映像に付加したことは事実である。

利用上の評価に間接的につながるであろう、表現上での評価例として、パノラマ表現型ショットシーケンス本における一結果を挙げておく。紀行もののビデオ1本(15分)に含まれる総ショット数(180)のうち、22.8%に当たる41ショットに対してパノラマ画像を代表画像として自動レイアウトがされた。これにより、単純シーケンス本に比べて、ページ数にして\*%の増加となっている。2割以上のショットに対して、映像内容把握上より有効なパノラマ表現がなされたと解釈できる。またページ数の増加については、この程度であれば本におけるばらばらめくりの効率から考えると大きな影響はないと考えている。

自動作成につきものの、100%理想的には行かない部分をどうカバーできるかの問題がある。現在のところ、映像の本的な利用に対する正解を定義すること

が難しいため評価は難しいが、個々の特徴処理の精度の高さから考えると、全体としては概要の把握とアクセスに対しては大きな問題はないと考えている。

## (2) 課題と発展性

今回紹介した例では、代表画像のレイアウトによる表現を行っている。実際には、映像中の時間情報や抽出特徴情報も含め、併せて表現すると効果的な工夫がまだまだ可能である。また、代表画像の表現寸法についても、何らかの特徴情報を用いて可変にすることも考えられる。ページ単位での区切り方にも音楽の切れ目で行うなどの特徴情報の利用の仕方も可能である。

通常の本にある索引的な機能を、映像に対してはテキストにおける索引とは違った観点で設けることも課題の一つである。例えば色その他の特徴情報ベースで引くことのできる索引も可能である。映像本中における検索機能についても同様に、テキスト本とは異なる方法による機能を盛り込むことも考えている。

できた AVON を原型として映像に関してもっと人が手をかけて情報を付加すれば、さらに豊富なマルチメディア電子本になるとの利用者の意見もある。これに対しては、MediaParser の段階、あるいは MediaProducer の段階でのオーサリング環境の利用により実現する事ができる。ただし、大きな手間暇をかけず、ある程度の目的は十分達成できる範囲で自動的に作成可能な AVON 設計が、多くの情報が今後流通、利用される社会を考えると重要であるとも考えている。いずれにしても、目的やコストに見合ったトレードオフが必要である。

応用上の観点からは、利用するインタフェースがブックメタファ故、特にマニュアルを見なくてもすぐに利用できるメリットを持ち、電子図書館、ビデオライブラリーへの応用が有望である。

ハイパーリンク的な機能の実現は、特徴レベルでの自動作成からだけでは難しく、人手ないしはオブジェクトレベルの認識に基づく処理が必要となる。これをは一般的な映像に関して実現するのは難しいが、コンテンツ範囲を特定したものであれば内容によっては現実味を帯びる。

MeidaFarm と呼ぶシステム構築体系は、処理ノード間でデータインタフェースを設けて行うため、またさまざまな利用系にカスタマイズが容易であるとともに、同機能の分散環境での実現も容易である。

## 6. さいごに

本論文では、映像解析結果を利用し、ブックメタフ

ァ型インタフェースであるサイバーブック上で映像を表現し、新しいインタラクティブ性を実現した映像本 (AVON) の考え方、作成処理について述べた。映像の内在する構造を十分利用することができなかった従来のインタフェースに比べ、アクセスする粒度の階層性など、特徴レベルの情報ハンドリングでも、効果的なインタフェースを実現することができた。

映像を本にしたことで、今後、暗に本以上の利用機能を付加することが発想的に難しくなることも考えられるが、我々としては、映像のさまざまな使われ方への発想を豊かにして、ブックメタファだけにこだわらず新しいインタフェースも提案していくつもりである。

(本文中図面中の画像のもととなるビデオは文芸春秋社の提供によるものです)

## 謝辞

研究を進めるに当たり、日頃ご指導頂く NTT ヒューマンインタフェース研究所映像処理研究部徳永部長をはじめ討論頂いた同研究部各位に感謝致します。

## 文献

- (1)Hodges E.M., Sasnett M. R., "Multimedia Computing," Addison-Wesley 52029, 1993.
- (2)Tonomura Y., Akutsu A., Taniguchi Y., Suzuki G., "Structured Video Computing," IEEE Multimedia, Vol. 1, No. 3, pp.34-43, Fall 1994.
- (3)Ueda H., Miyatake T. and Yoshizawa S., "IMPACT: An Interactive Natural-Motion-Picture Dedicated Multimedia Authoring System," in Proc. of CHI'91, pp.343-350, 1991.
- (4)Meng J., Chang S., "CVPS-A Compressed Video Editing And Parsing System," Proc. of ACM Multimedia96, pp.43-53, 1996.
- (5)Arman F., Depommier R., Hsu A., Chiu M-Y., "Content-based Browsing of Video Sequences," Proceedings of ACM Multimedia'94, pp.97-103, 1994.
- (6)Chua T., Lim S., Pung., H., "Content-based Retrieval of Segmented Images," Proc. of ACM Multimedia94, pp.211-218, 1994.
- (7)Davis, M., "Media Streams: An Iconic Visual Language for Video Annotation," Proc. of IEEE Symposium on Visual Languages, CS Press, pp.196-202, 1993.
- (8)米村,小川, "HI 設計ガイドラインデータベース: ブックメタファの可視化とその効果", 情処論, Vol.35, No.10, pp.2159-2169, 1994.
- (9)Zhang H., Kankanhalli A., Smoliar S., "Automatic partitioning of full-motion video," ACM Multimedia Systems, 1, pp.10-28, 1993.
- (10)Otsuji K., Tonomura Y., "Projection-detecting filter for video cut detection," ACM Multimedia Systems, 1, pp.205-210, 1994.
- (11) Lienhart R., "Automatic Text Recognition for Video Indexing", Proc. of ACM Multimedia 96, pp.11-20.
- (12)Tonomura Y., Abe S., "Content Oriented Visual Interface Using Video Icons for Visual Database Systems," Journal of Visual Languages and Computing, Academic Press, 1, 183-198, 1990.

- (13)柴田, “映像の内容記述モデルとその映像構造化への応用”, 信学論 D-II, Vol.J78-D-II, No.5, pp.754-764, 1995.
- (14)Davenport G., Aguiere, S. and Pincever, N., "Cinematic Primitives for Multimedia," IEEE Computer Graphics and Applications, 11, 4, pp.67-74, 1991.
- (15)Weyer S. A., "Searching for Information in a Dynamic Book," *Xerox Technical Report SCG-82-1*, 1982.
- (16)Egan D. E. et al., "Formative Design Evaluation of SuperBook," ACM Trans. on Information Systems, Vol. 7, No. 1, pp.30-57, 1989.
- (17)Card s. K., Robertson G. G., York W., "The WebBook and the Web Forager: An Information Workspace for the World-Wide Web," Proc. of CHI96, pp.111-117, 1996.
- (18)<http://www.voyager.co.jp/>
- (19)小澤,吉宗,浜田,小川, “ブックメタファ: マルチメディア情報の閲覧における「本」インタフェース”, 画像電子学会誌, 第25巻, 第5号, pp.454-463, 1996.
- (20)Ueno K., Suzuki K., Ozawa H., "Book based view and scroll based view: Which is suited for HTML viewers?,"Proc. of HCI International'97, Vol.2, pp.795-798, 1997.
- (21)鈴木,小澤,外村, “BookWare: 本と本棚メタファによるマルチメディア情報ハンドリング” ヒューマンインタフェース, 72-12, pp.67-72, 1997.
- (22)Nagasaka A., Tanaka Y., “Automatic video indexing and full-video search for object appearances,” in Proc. of 2nd Working Conference on Visual Database Systems, pp.119-133, October, 1991.
- (23)Taniguchi Y., Akutsu A., Tonomura Y., "PanoramaExcerpts: Extracting and Packing Panoramas for Video Browsing," in Proc. of ACM Multimedia97, pp.427-436, 1997.
- [24]Minami K., Akutsu A., Hamada H., Tonomura Y., “Enhanced Video Handling based on Audio Analysis,” Proc. of IEEE Multimedia Computing and Systems '97, pp. 219-226, 1997.
- (25)青木,堀 “映像構造を利用した代表フレーム表示方法”, インタラクシオン 97 論文集, pp.9-16, 1997.