

映像アーカイビング

児島 治彦

NTTサイバーソリューション研究所

博物館、自治体、放送局など、大量のコンテンツを保有している団体・組織がデジタルアーカイブを構築する動きが盛んである。今後はデジタルアーカイブのネットワーク化など、コンテンツ流通プラットフォームにおける「知の集積」として、デジタルアーカイブが果たす役割は高まっていく。本稿では、デジタル映像アーカイブ構築に不可欠な映像インデクシング技術と映像表現技術（ブラウジング&インタフェース）の最新プロダクトを紹介する。

【 デジタルアーカイブをめぐる 】

大量のコンテンツを保有している組織・団体が、電子博物館、電子美術館、電子図書館などのように、コンテンツをデジタルアーカイブ化する動きが盛んである。デジタルアーカイブとは、有形、無形の文化資産をデジタル形式で記録し、その情報をデータベースにまとめたものである。デジタル情報であることから、保存性に優れ、検索や再加工がしやすいなどのメリットがある。文化資産保存の観点から国際的に関心が高まっており、日本でも1996年4月に文化庁、通産省などが中心となって「デジタルアーカイブ推進協議会」が設立されて以来、京都デジタルアーカイブのように自治体の単位で全国的に構築する機運が高まっている。

一方、大量の映像を保有している放送事業者や映像制作事業者も、映像アーカイブ構築を進めつつある。BSデジタル、地上波デジタルと、今後のテレビ放送多チャンネル化時代において想定されるのは、著しいコンテンツ不足である。放送事業者にとってのアーカイブは、①局内の再放送・再構成素材としての利用、②番組公開ライブラリとしての利用、③2次利用としての活用、が挙げられる。

個々のデジタルアーカイブに閉じたコンテンツの活用だけでなく、複数のデジタルアーカイブをまたがった検索・再利用、あるいは映像コンテンツの2次利用としての活用を進めるうえで、デジタルアーカイブのネットワーク化が重要になってくる。

上記で示したコンテンツの例では、編集・加工されて2次著作物化されるケースは少ないと考えられるが、権利保護が確立している流通形態が存在すれば、素材映像

など2次著作物化されることを前提としたコンテンツのデジタルアーカイブの構築が加速化されるものと考ええる。このような流通形態をコンテンツリサイクルマートと呼ぶ¹⁾。図-1に示すとおり、コンテンツリサイクルマートでは、デジタルコンテンツはまず生成・蓄積（知の集積）され、種々のルートを通じて流通・販売（知の流通）された後、編集・加工（知の再生産）され、2次著作物として利用される。本稿で紹介する技術はこのうちの「知の集積」を実現するためのものである。

話をデジタルアーカイブ構築に戻すと、映像アーカイブ構築には単なるデジタル化だけでなく、後での活用をふまえたインデクス付与が重要である。インデクス付与により、蓄積された映像に対するブラウジングやインタフェースの面で新しい可能性が開ける。

本稿では、デジタル映像アーカイブ構築に不可欠な映像インデクシング技術と映像表現技術（ブラウジング&インタフェース）について最新プロダクトを紹介する。

【 映像インデクシング 】

映像は、ビデオテープや映画フィルムを想起していたらだければお分かりのとおり、「巻物」として捉えることが

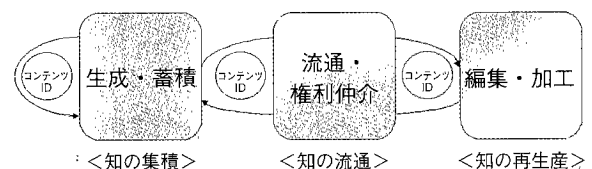
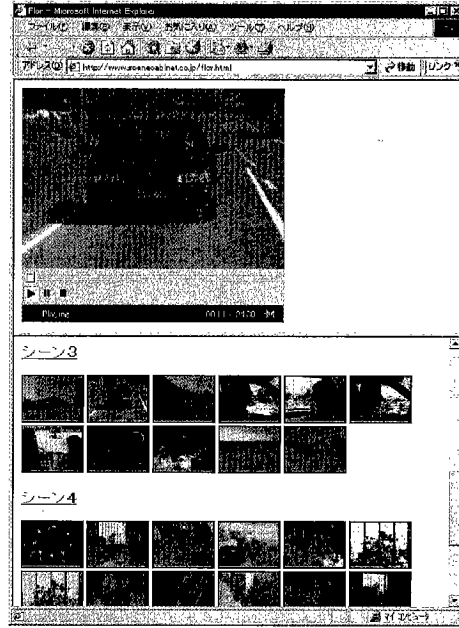


図-1 コンテンツリサイクルマート（文献1）より



(a) インデクス一覧例



(b) Webブラウザ表示例

図-2 SceneCabinet Pro II

でき、①全体が把握しにくい、②どこに何があるか分からない、③見たい場面がすぐには見えない、といった問題がある。

これに対して書籍の場合は、章節で構成されている、目次・索引がある、という具合に構造化されていることから、一覧・概要把握・検索が容易である。映像も同じように構造化されていれば、一覧・概要把握・検索が容易になる。映像を書籍のごとくハンドリングするためには、構造化を支援する映像インデクシング技術が必要である。映像インデクシング技術を用いたプロダクトは国内外を問わずいくつも紹介されているが、ここでは代表例としてSceneCabinet Pro II(図-2)を紹介する。

SceneCabinet Pro IIでは、映像を取り込みながら、あるいはあらかじめ蓄積された映像から、カット点²⁾、テロップ³⁾、カメラワーク⁵⁾、音声区間、音楽区間⁴⁾を自動検出してインデクスを生成することができる。カット点検出ではディゾルブやワイプなど、漸次シーン変化を検出可能であり²⁾、音楽と音声重なった場合でも各々を独立に検出可能である⁴⁾。生成されたインデクスに対するコメント付与、修正、一覧、印刷はもちろんのこと、生成されたインデクスからWebページを自動作成し、イントラネット上で映像を閲覧可能にする機能も有している。

【 パノラマ展開型ブラウザ 】

映像から代表画像を自動的に抜粋して映像一覧を生

成するブラウザは、先述のSceneCabinet Pro IIを含めて、種々開発されてきている。しかしながら、代表画像としてシーンの先頭画像を用いることが必ずしもよいとは限らない。たとえば、パン・チルトのようなカメラワークを含むシーンでは、図-3に示すようにシーンから合成したパノラマ画像を代表画像として一覧表示することにより、映像の一覧性が増す。この図-3の機能を実現しているのが、PanoramaExcerpts⁵⁾である。

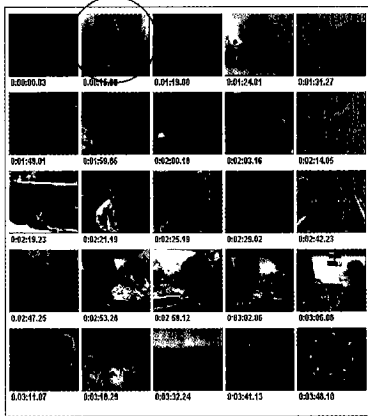
PanoramaExcerptsでは、まず映像解析によりカット点検出、カメラパラメータ推定、安定したカメラワークの検出を行う。次に検出されたカメラワーク区間についてパノラマ画像を自動合成し、これを代表画像とする。さらにパノラマ画像を含むことによって、形状、大きさが変動した代表画像を適切に配置する自動レイアウトの処理を行う。PanoramaExcerptsでは、カメラワーク区間の自動検出・パノラマ画像の自動生成・代表画像の自動レイアウトを、PC上で実時間以上の処理速度、かつロバストに行っている。

【 テロップ認識を用いた映像検索 】

音声・字幕に含まれる言語情報はインデクスとして重要であり、言語を検索キーとする検索には必須のものである。

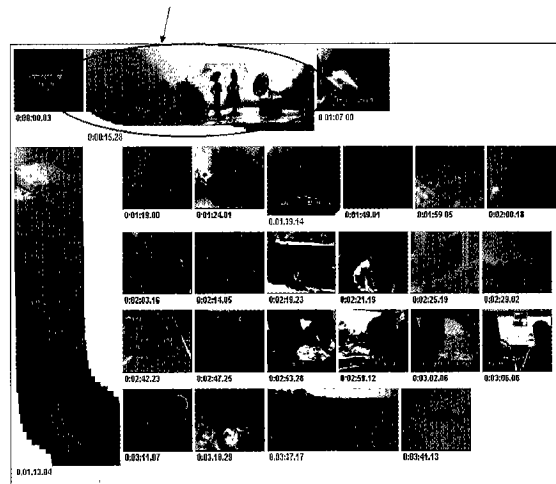
自動付与できるインデクスとして、カット点、テロップ、カメラワーク、音声区間、音楽区間があることを紹介した。音声区間が自動検出できれば、音声認識もでき

カメラワーク（パン）が存在するため、シーン先頭画像にはシーンの一部しか写っていない。



(a) シーン一覧型

カメラワーク（パン）を自動検出し、パノラマ画像化することで、一覧性が増す効果がある。



(b) パノラマ画像展開型 (PanoramaExcerpts)

図-3 映像一覧表示

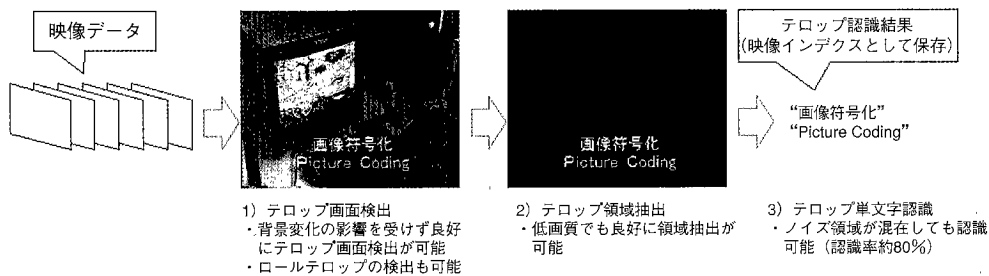


図-4 テロップ認識フロー

るのでは、という意見がよく聞こえてくるが、背景雑音、語彙など、種々の課題があって、一部の例外を除き、実用化には多くの壁がある。

ここでは、映像中のテロップ文字検出に加えて、テロップ文字を自動認識し、認識結果をインデックスとして保存し、これを用いて映像を検索・閲覧するテロップ認識システムを紹介する。テロップ文字には映像の内容を直観的に説明する役割があり、映像検索のキーとして有効と考えられる。

テロップ認識のフローを図-4に示す。まず映像の中からテロップ画面を検出した後、テロップ領域を抽出し、テロップ領域から文字切り出しを行ったうえで、テロップ単文字認識を行う。テロップ文字の色に対する制限はなく、縦書き、横書き、ロールテロップにも対応している。認識結果は映像のインデックスとして保存され、これを検索キーとして所望の映像の自動再生が可能となる。

大量映像ハンドリング

放送事業者・映像制作会社など、大量の映像を保有している組織が映像アーカイブを構築する際、以下のニーズがあると考えられる。

- ・大量の映像を高速に自動構造化したい。
- ・手入力情報付加における操作者の負担を減らしたい。
- ・大量の映像アーカイブを一覧検索したい。

ところで、大量のデータからモデルを発見する手法としてデータマイニングの研究がある。また、新聞や書籍などのデータを大量に集積したコーパスを整備し、統計分析に基づく定型的パターン抽出や、分類、認識、検索を行う研究がある。映像処理の分野でも、大量の映像データを付加情報とともに蓄積したコーパスを構築すれ

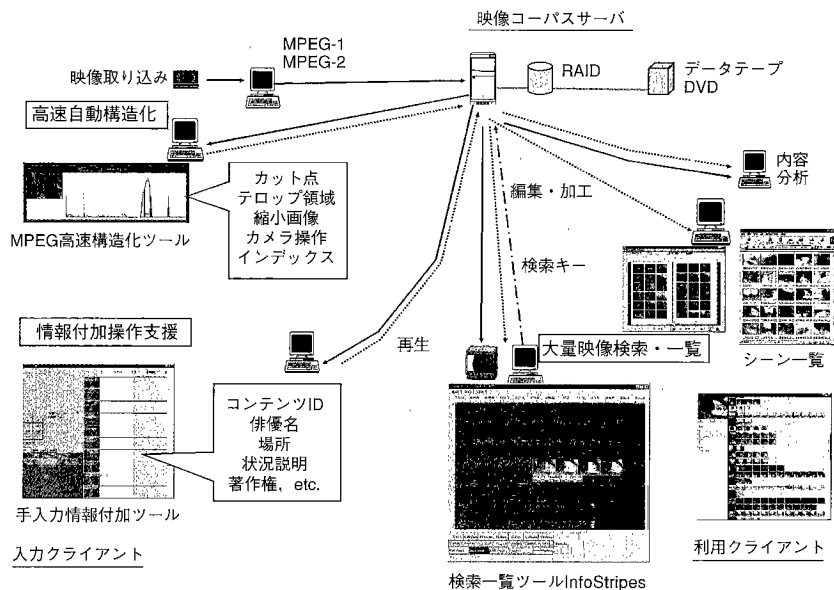


図-5 映像コーパスのシステム構成

ば、実用的な分類、検索、可視化技術の開発や、アルゴリズムの比較・評価に有効であるとの期待から、映像コーパスの研究が行われた⁶⁾。この研究の過程で開発したツールが、実は上記のニーズに対する回答になっているので、以下に紹介する。

図-5に映像コーパスの構成を示す。

(1) MPEG高速構造化ツール

MPEG-2符号化映像を直接解析し、カット点、テロップ、縮小画像などのインデックスを自動付与する。

(2) 手入力情報ツール

俳優名、状況説明、著作権など、自動で付与できない情報は、オペレータによる手操作入力が必要になる。ここでは、表とプルダウンメニュー形式で簡単に種々のインデックス情報を付加できるツールを開発している。

(3) 大量映像検索・一覧ツール

大量データを対象とした場合、マクロな観点から大雑把にみて特徴的な部分が見つかったとき、それがどのような意味を持つかはミクロなデータを見て理解できるものである。そこで、数百時間規模の映像を対象として、インデックスによる検索とその結果を可視化するツールInfoStripesを開発している。InfoStripesは実映像の数フレームを1画素に対応付け、そのフレームのインデックス情報を色に置き換えて、多数の映像をストライプ上に表示する。またカーソル位置のインデックス情報を小ウィンドウに表示し、さらにクリックすれば、対応する映像の再生ができる。InfoStripesはメニューを用いた絞り込み検索も可能で、検索結果を特定の色で表示できる。

直観的映像インタフェース

最後に映像コンテンツの活用という観点から、インタラクティブな映像操作インタフェースについて紹介する。

インタラクティブな映像操作としては、フレーム画像中のアンカー領域(ホットスポット)を用いるものや、時間軸対応のスライダを用いるものがある。前者は被写体との空間的対応に優れているが、早送りや巻き戻しといった時間軸上の操作には不向きである。一方、後者は連続的な時間操作ができるものの、空間的対応がとりにくい。通常スライダは左から右へと水平配置されているので、たとえば、スライダのつまみを右へ動かしているのに、被写体は左に動くといった違和感を抱くことがある。

映像の時空間構造に直観的な操作性をもたらす新しい映像インタフェースとしてCyberCoasterを紹介する⁷⁾。CyberCoasterは映像操作のためのアンカーとして埋め込まれた、カメラワークや物体移動などの軌跡を近似した折れ線スライダである。CyberCoasterを用いると、あたかも映像中の物体を直接つかんで動かしているかのような、映像の可変速再生が可能である。

図-6はCyberCoasterをゴルフのスイング撮影ビデオに適用した例である。図中の折れ線が、映像に埋め込まれた折れ線スライダ; CyberCoasterである。マウスなどのポインティングデバイスを用いてCyberCoasterを操作すると、CyberCoasterの出力値に対応するフレーム画

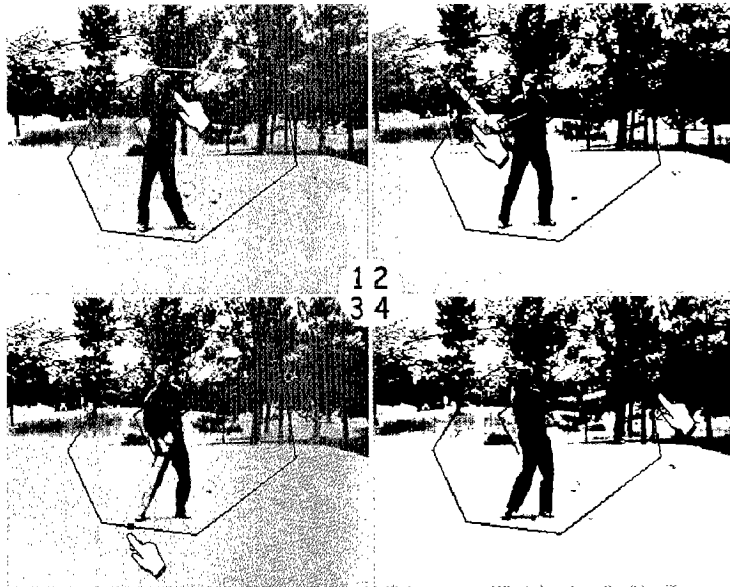


図-6 CyberCoasterをゴルフスイング映像に展開した例

像を表示する。この場合、クラブのヘッドをつかんで動かしているかのように再生できる。CyberCoasterは手動で設定することも、また被写体追跡技術を用いて自動で設定することも可能である。

今後について

映像アーカイビング技術の最新状況を紹介してきたが、最後に今後の課題について考察する。

放送事業者では、映像アーカイブにおける活用の利便性と構築時の入力コスト削減の観点から、映像制作時点で映像制作者が付属情報をキー入力するシステム作りを行っているところもあると聞く。一方、制作現場からは、付属情報を細かくキー入力している暇はない、という声も聞こえてくる。

本稿で示したテロップ認識システムは、これに対する1つの解決策である。しかしながら、映像解析による映像自動インデクシング技術は100%の精度でできるわけではない。また、自動で付与できるインデクスの種類も限られている。したがって現段階においては、映像アーカイブ構築時には手入力によるインデクス付与や、自動インデクシング結果に対する確認修正など、オペレータの介在が必須である。膨大な映像データのアーカイブを構築するうえで、この運用コストは大きな課題である。2次利用の促進による投資回収という言葉ではあまりにリスクが大きい。2次利用の可能性が高いものから順に詳細にインデクシングしていくというのが現実解であろう。2次利用の促進という意味では、冒頭でも述べた

とおり、著作権などの権利保護、決済・課金系も含めたコンテンツリサイクルマートの確立が望まれる。

研究開発面での今後の展望としては、自動で抽出・認識できるインデクスの種類を増やす、個々のインデクス抽出精度を高める、といった従来の技術開発の延長と、話題抽出、内容解析など、意味レベルまで踏み込んだ映像構造化技術の開発がなされるであろう。後者においては、ニュース、スポーツ中継など対象を限定すれば、ドメインに依存した知識を用いて意味的なインデクスを抽出することは比較的容易である。一方、映像の表現方法についてはアイデア勝負といった感もあるが、さまざまな可能性があり、今後も世の中を驚嘆させる成果が続々登場することが期待できる。

参考文献

- 1) 山中, 中村, 小川, 高嶋, 曾根原: 著作権保護技術の動向—コンテンツリサイクルマート創出の基盤技術—, 情報処理, Vol.41, No.4, pp.382-387 (Apr. 2000).
- 2) 谷口, 外村, 浜田: 映像ショット切換え検出法とその映像アクセスインタフェースへの応用, 信学論, Vol.J79-D-II, No.4, pp.538-549 (1996).
- 3) 佐藤, 新倉, 谷口, 阿久津, 外村, 浜田: MPEG符号化映像からの高速テロップ領域検出法, 信学論, Vol.J81-D-II, No.8, pp.1847-1855 (1998).
- 4) Minami, K., Akutsu, A., Hamada, H. and Tonomura, Y.: Video Handling with Music and Speech Detection: IEEE MultiMedia, Vol.5, No.3, pp.17-25 (1998).
- 5) Taniguchi, Y., Akutsu, A. and Tonomura, Y.: PanoramaExcerpts: Extracting and Packing Panoramas for Video Browsing, ACM Multimedia 97, pp.427-436 (1997).
- 6) 佐藤, 児島, 阿久津, 外村: 映像コーパスの構築と分析, 信学論, Vol.J82-D-II, No.10, pp.1552-1560 (1999).
- 7) Satou, T., Kojima, H., Akutsu, A. and Tonomura, Y.: CyberCoaster: Polygonal Line Shaped Slider Interface to Spatio-Temporal Media, ACM Multimedia 99 (1999).

(平成11年11月12日受付)

