

## 映像情報の動的構造化とストリーミング配信に関する検討

An Approach to Dynamic Structuring  
and Streaming Distribution of Audiovisual Contexts

加藤 寛<sup>†</sup> 青島 大悟<sup>†</sup> 藤城 卓己<sup>†</sup> 坂根 裕<sup>‡</sup> 竹林 洋一<sup>‡</sup>  
 Hiroshi Kato Daigo Aoshima Takumi Fujishiro Yutaka Sakane Yoichi Takebayashi

## 1. はじめに

筆者らは、一般的な学問の学習だけでなく格闘技や楽器演奏など、スポーツや趣味の「学び」を複数の映像や音声、センサデータ、知識を融合した「マルチモーダル知識コンテンツ」によって活性化させるための研究を推進してきた [1]。効果的な学習環境をデザインするには、取り組みを妨げず学習意欲を向上させるように、意味のあるコンテンツの作成と適切なタイミングや手段で提示できる仕掛けが重要となる。これまで閲覧用 PC や DVD を利用しコンテンツの閲覧環境を実現した (図 1) が、カスタマイズされた PC や再編集が難しい DVD メディアではシステム構成が取り組みに制限を与えることがあった。

本稿では、より広くコンテンツを利用できる環境を実現するため、インターネットによるオンデマンド方式のコンテンツ編集・配信・閲覧環境を実現するプラットフォームを提案する。ストリーミングサーバによるインターネット配信は、ある決められた映像や音声 (ファイルやライブカメラ、マイクなど) をネット帯域に応じてクライアントに届けるシステムである。本研究ではサーバに蓄積したコンテンツの編集、内容や帯域に応じたストリームデータの生成、クライアントでの提示方式など広く全体を見据えたデザインが重要であり、新しいコンテンツ配信環境を実現するものである。

## 2. マルチモーダル知識コンテンツをストリーミング配信するための要件

マルチモーダル知識コンテンツの有効性を損なわずにストリーミング配信するためには、単に映像やテキスト、センサ情報を配信するのではなく、ユーザの要求に合わせて構成した、意味のあるコンテンツを提供する必要がある。そのためには、マルチモーダル知識コンテンツをストリーミング配信するための要件について検討を行い、以下の機能が必要であることがわかった。

## コンテンツの動的な構築

従来の VOD 配信方式では、予め静的なコンテンツを用意しておくのが普通であるが、マルチモーダル知識コンテンツでは、ユーザの目的や状況等によって、構成や提示方法が変化するため、静的なコンテンツを扱う場合、マルチモーダル知識コンテンツの有効性を損なわずに提供することが難しい。単純に、蓄積してある複数のモダリティの情報を全て提示するだけでは、情報量が多すぎてユーザの負担になってしまうためである。そこで、ユーザの要求に合わせて、複数のモダリティ情報を適切に取捨選択してコンテンツを構成する必要がある。つまり、



図 1: マルチモーダル知識コンテンツ

編集者やユーザが知識等の付加情報を加えたときに、即座に配信できるようにするためには、オンデマンドで動的にコンテンツを構成する機能が必要である。

## ストリーミング形式への変換

映像をストリーミング配信する枠組みはすでに世の中に普及しているが、センサ情報やテキストを映像に同期させてストリーミング配信できる枠組みは、まだ実現されていない。複数のモダリティの異なるデータを統合して、1本のストリーミングデータを生成する機能が必要である。

## コンテンツの提示

ユーザが受信したストリーミングデータを閲覧する際に、ユーザの目的やコンテンツの種類によって、モダリティの数や種類も変化するため、閲覧用アプリケーションはさまざまな提示方法が必要になる。そこで、閲覧用アプリケーションに柔軟性を持たせるために、ストリーミングデータを映像、センサデータ、テキスト等に分離してアプリケーションに渡す機能が必要になる。

## 3. マルチモーダル知識コンテンツ配信環境の試作と検討

前章で説明した機能を持つ配信プラットフォームの構成の設計について検討し、コンテンツの編集機能を実装し、利用可能なマルチモーダル知識コンテンツを構築した。

## 3.1 システム構成

図 2 に示すように、マルチモーダル知識コンテンツ配信のためのプラットフォームを、コンテンツ配信サーバで動作するサーバプログラムと、ユーザが使用する閲覧用のアプリケーションに組み込んで動作するクライアントモジュールに機能を分割した。サーバプログラムは、

<sup>†</sup>静岡大学 大学院 情報学研究科  
<sup>‡</sup>静岡大学 情報学部

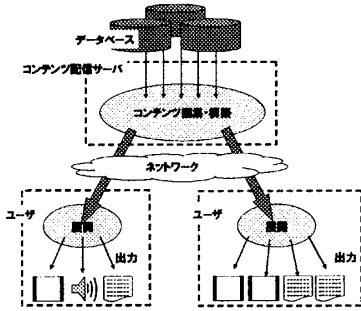


図2: システム構成

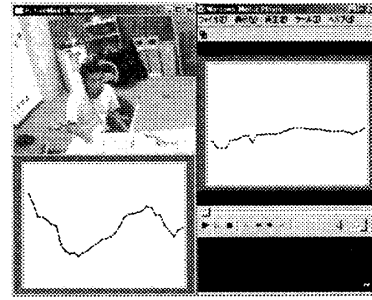


図3: しぐさ研究支援コンテンツ

コンテンツの編集・構築と配信を行い、クライアントモジュールは、閲覧用のアプリケーションに組み込んで、サーバプログラムと通信し、受信したコンテンツをもとの映像やテキスト、センサデータ等に分離し、レンダラや出力のためのデバイスに渡す役割である。映像のレンダリング方法や、センサデータの出力方法は、閲覧用アプリケーションの実装に依存する。そのため、実装は、どのように表示、出力するかという点に注力でき、様々な表現方法を生み出すことが容易であるというメリットがある。

### 3.2 コンテンツ編集機能の実装

今回は、複数の異なるフォーマットのデータを、統合して一つのコンテンツとして編集・構築することのできる機能の実装を行った。複数の映像やテキストなどをコンテンツとして構築するために、複数の映像、音声に加え、テキストなどを1本のストリーミングデータとして構築することが可能であるASFフォーマット[6]を用い、任意の映像と音声、テキストを読み込み、それぞれの再生時間に合わせて順番にサンプルをインターリーブし、1本のストリーミングデータを構築できる。

これにより、センサデータをグラフにプロットしたものと、映像を合わせて見ることができるコンテンツを作成した(図3)。ヘッドセットに装着したカメラと地軸センサ、加速度センサから撮影した視点の映像と、首の傾きのセンサデータからうなずきのタイミングを抽出し、会話の成り立ちとの関係性を検証する[5]際に利用できる。試作したコンテンツを利用することで、映像とセンサデータの変化などを同時に把握することができ、研究のための観測ツールとして有効である知見が得られた。

### 3.3 コンテンツ編集のための構成情報

関連や意味のつながりを考慮してコンテンツを編集するためには、関連のあるデータ、再生時間などをマルチモーダル知識コンテンツの構成情報として保持する必要がある。筆者らの研究グループでは、センサ情報の収集と処理、蓄積を目的としたセンサデータベース[4]を開発している。センサデータベースでは、各種センサデータに対し、利用しやすい形に加工したり、生データに対し解析を施し意味付けといった処理をかけることが可能である。このセンサデータベースをコンテンツデータベースに組み込んで利用することにより、マルチモーダル知識コンテンツ構成情報の作成と参照が容易となると考え

られる。

## 4. まとめ

インターネットによるマルチモーダル知識コンテンツ配信は、これまで難しかった教室や道場、屋外でのコンテンツ閲覧環境を実現することが可能になり、学習環境デザインに与える影響は大きい。モーションデータ(センサデータ)を映像に変換し取り組み映像と合成するような「新しい視点」でのコンテンツ提示は、閲覧者に新たな課題を気づかせ学習意欲を向上させる、奥深い学習環境を実現できるという見通しを得た。

今後は、さらにシステムの実装を進め、実用に則した閲覧用アプリケーションとコンテンツを作成し、社会にとって有益なマルチモーダル知識コンテンツの利用環境の実現を目指す。

## 参考文献

- [1] Y.Takebayashi, T.Sugiyama, and Y.Sakane: Multi-modal Knowledge Creation in Ubiquitous Learning, UEM 2003 at ATR (invited), pp.49-52 (2003.9).
- [2] 高島政実, 白木顕介, 黒木孝志, 坂根裕, 竹林洋一: 空手の型の極意を学ぶための音楽フィードバックの利用, 情報処理学会インタラクティブ2004 インタラクティブ発表, pp.233-234 (2005.3).
- [3] 川口紗季, 坂根信一, 坂根裕, 竹林洋一: 親子のコミュニケーションを促進する幼児教育の一検討, 情報処理学会第67回全国大会, 5A-2 (2005.3).
- [4] 青島大悟, 大関毅, 坂根裕, 竹林洋一: マルチモーダル知識構築のためのセンサデータベース, HIS2004.
- [5] 山本剛, 坂根裕, 竹林洋一: マルチモーダルヘッドセットを用いたうなずき検出と会話の重要箇所把握, 情報処理学会ヒューマンインタフェース研究会, pp.13-19 (2003.9).
- [6] Microsoft:AdvancedSystemsFormat(ASF)Specification (<http://www.microsoft.com/asf>)