

マルチモーダルナレッジ技術を利用した 科学館ガイドシステム

Science Museum Guidance System Using Multimodal Knowledge Technology

福井 美佳 鈴木 優 藤井 寛子

FUKUI Mika

SUZUKI Masaru

FUJII Hiroko

東芝は、来館者が展示物に近づくと Bluetooth™(注1)を介して展示物の説明映像が手元の端末に配信される科学館ガイドシステムを開発し、2002年9月から東芝科学館で一般公開を開始した。端末に向かって音声で質問すると、更に詳しい情報を検索できる。このシステムは、映像と文字情報を構造化したコンテンツを簡単に作成し、オフィス、家庭、モバイル環境など様々な場所でオンデマンドに利用する、マルチモーダルナレッジ技術の実証実験のために開発した。ほかに、社員研修などの教育関連、店舗における商品説明、観光地の説明など、様々な応用に展開可能である。

Toshiba has developed a multimodal knowledge and information on demand system as an application of ubiquitous computing. This system provides nuggets of multimodal knowledge composed of video/audio streaming data and text data to users via wireless network. It delivers whatever knowledge is appropriate to the user's situation in the ubiquitous computing environment to a terminal in the user's hand. The user can input questions using voice and can retrieve detailed information.

For practical verification of this system, we have developed a science museum guidance system and have been conducting an open experiment since September 2002 at the Toshiba Science Institute in Kawasaki City, Japan.

1 まえがき

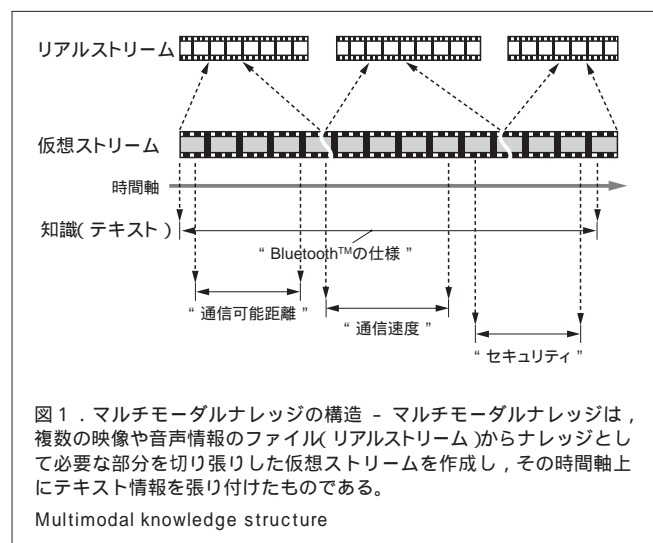
Bluetooth™などの無線通信技術や携帯端末機器の発達により、時間や場所にとらわれずコンピュータ資源にアクセスできるユビキタスコンピューティングの世界が実現しつつある。今後は、ユビキタス環境を生かした実サービスを指向するシステム構築と実証実験が重要になる。

東芝は、ユビキタス環境での有用なアプリケーションとして、映像・音声とテキスト情報を構造化し共有する、マルチモーダルナレッジ技術を利用した科学館ガイドシステムを開発した。このシステムは、ユーザーが場所や状況に応じて適切なナレッジに簡単にアクセスできるよう Bluetooth™技術、自然言語検索、音声認識、MPEG-7(Moving Picture Experts Group-phase 7)など、多様な技術を結集して実現されている。ユーザーが移動中でも、目の前の物に関する詳しい情報が、映像を含むコンテンツとしてユーザーの手元の端末に配信され、更に必要な情報があれば、音声で質問を入力し関連情報を検索することができる。このシステムは、2002年9月から、川崎市にある東芝科学館で実験運用を行っている¹⁾⁻⁽³⁾。

ここでは、マルチモーダルナレッジ技術と科学館ガイドシステムについて解説し、実験運用によって得られた知見について述べる。

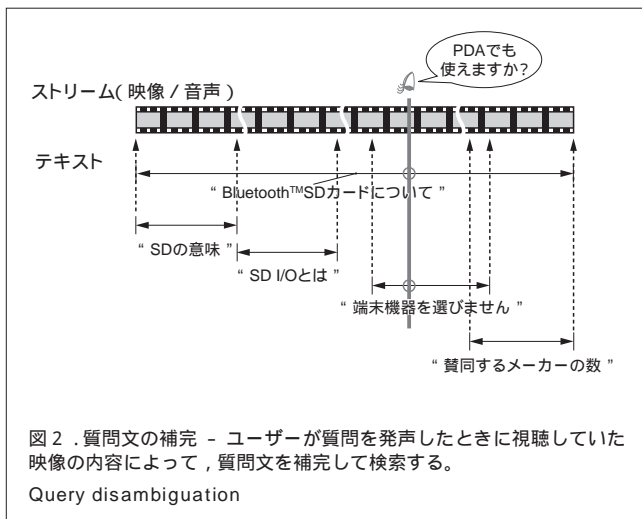
2 マルチモーダルナレッジ

マルチモーダルナレッジは、複数の映像や音声情報のファイル(リアルストリーム)からナレッジとして必要な部分を切り張りした仮想ストリームを作成し、その時間軸上にテキスト情報を張り付けたものである。リアルストリームと仮想ストリームの関連付けや仮想ストリームへのテキストの張付けは、マルチメディアコンテンツ内容記述の国際基準である MPEG-7に基づいて記述している。マルチモーダルナレッジの構造を図1に示す。



(注1) Bluetoothは、Bluetooth SIG, Inc.の商標。

現在、オフィスなどで普及しているナレッジマネジメントシステムの多くは、伝達する知識の形式として文書をベースにしている⁽⁴⁾。映像や音声も利用することにより、文書のようなテキスト中心の形式では伝わりにくい、機器の具体的な操作方法や細かいニュアンスが伝えやすくなる。また、映像とテキストが関連付けられているので、従来の自然言語検索技術を適用して簡単に目的の映像を検索できるようになる。更に、ユーザーの履歴や状況を考慮した検索も可能になる。例えば、図2のストリームを視聴中に、ユーザーが「PDAでも使えますか?」と質問した場合、ユーザーが現在見ている映像に関連付けられたテキストの「Bluetooth™ SDカードについて」と「端末機器を選びません」によって、ユーザーの質問を補完することにより、質問文のあいまい性を解消して検索を行うことができる。



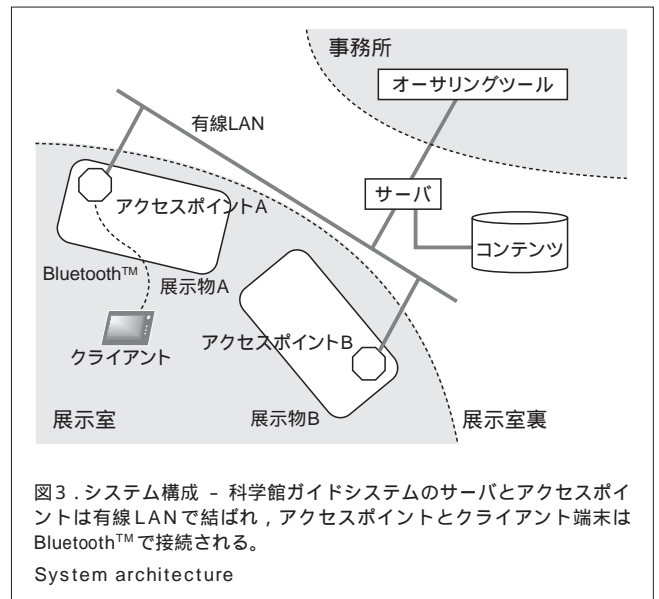
3 科学館ガイドシステムの概要

3.1 システム構成

科学館ガイドシステムは、サーバ、アクセスポイント、クライアント端末により構成される。サーバとアクセスポイントは有線LANで結ばれ、アクセスポイントとクライアント端末はBluetooth™で接続される。システム構成を図3に示す。

サーバは、マルチモーダルナレッジの登録、配信を行うコンテンツサーバと、コンテンツのインデキシングと検索要求の処理を行う知識エンジンにより構成される。将来的には屋内外のいたるところにアクセスポイントが設置されることを想定しているが、現在は展示物の近くにノートパソコン(PC)を1台設置しアクセスポイントを実現している。クライアント端末は、小型・軽量でBluetooth™を内蔵するタブレット型PCのWIA(Wireless Internet Appliance)を利用して(図4)。

来館者がクライアント端末を持って展示物に近づくと、クラ



クライアント端末は近隣のアクセスポイントを検索し、自動的に接続する。将来、展示ごとにアクセスポイントが設置された場合、サーバはアクセスポイントの近くの展示に関するコンテンツを優先的に配信する。来館者は、配信されたコンテンツに対して、タブレットをタッチするか、タイピン型マイクを介して音声で操作を行い、更に詳しい情報を入手する。クライアントの画面例を図5に、東芝科学館での実験のようすを図6に示す。

3.2 マルチモーダルナレッジのオーサリングツール

ユビキタス環境でナレッジを共有するためには、ユーザーの様々な状況に対して適切な情報を提供できるよう、豊富なコンテンツが必要になる。すなわち、マルチモーダルナレッジのオーサリング支援が重要な課題となる。

従来のビデオ編集ツールの用途は、エンターテインメント用あるいは映像編集のプロ向きであり、マルチモーダルナレッジのオーサリング作業には不必要な機能が多く、使いづらい。

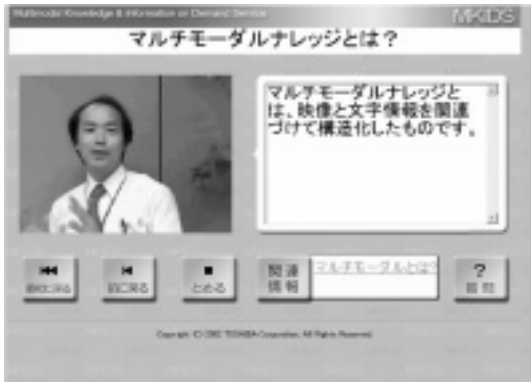


図5. クライアント画面 - クライアント端末の画面には、展示内容の説明映像とともに、映像のタイトル、説明文、関連情報へのリンクが表示される。

Appearance of client screen



図7. オーサリングツール - 映像や音声データから必要なシーンを切り出し、タイトル、説明文、関連情報へのリンクを付加して、マルチモーダルナレッジを作成する。

Authoring tool



図6. 東芝科学館での実験 - 東芝科学館で実際に来館者に操作してもらいながら実験を行っている。

Experiment at Toshiba Science Institute

そこで、最低限必要と思われる以下の機能に絞ったオーサリングツールを新たに開発した。

- (1) 映像、音声及びテキストの3トラックに限定
- (2) 各トラックの分割、削除、結合操作が可能
- (3) テキストトラックの1分割単位ごとに、タイトル、説明文、キーワード、関連情報へのリンクを格納

オーサリングツールの画面イメージを図7に示す。

4 実験運用

4.1 コンテンツの作成、評価

展示案内者による展示物の概要説明(5分程度)1本、展示機器の実演(1分程度)1本、及び来館者からよく聞かれる質問への技術者の回答(1分程度)10本を撮影し、これらの素

材をもとに、展示案内者と展示企画者の3名に、このオーサリングツールを使ってオーサリング作業を行ってもらった。

オーサリングツールの操作性に関しては、初めて使う場合でも特に問題なく使えることを確認した。また、オーサリング作業のうち、素材の中から回答として価値のある映像部分を切り出す作業に手間が掛かることがわかった。

ユーザーを自然に関連情報に導いていくために、どの素材をどの素材の関連情報にするかを決めるのは難しい。展示案内者からは一度に複数の関連情報を関連付けたいという要望がでたが、それにより画面の可読性も低下するため、最適な文字数、画面デザイン、表示時間などを慎重に検討する必要がある。

4.2 科学館展示担当者へのインタビュー

上記の展示案内者と展示企画者の3名に、以下の項目についてインタビューを行った。

質問1: 今後も利用したいか? 主な利用目的は?

回答: 利用したい。展示の無人化が主目的ではなく、展示案内者が同行説明できない来館者への補足説明として利用し、来館者全体の満足度を上げたい。

質問2: コンテンツを作成したいか?

回答: 作成したくないが2名、作成したいが1名。作成したくない理由は作業がめんどろ、作成したい理由はおもしろそう。

質問3: 機能の改善や拡張の要望は?

回答: 映像の配信時間を短く、一度に利用できる端末を増やしたい、端末を小型化したい、自由に音声で質問をしたい、など。

4.3 アンケート調査

当社のCS(顧客満足)評価センターにて、成人63名を対象にこのシステムのデモンストレーションを行い、システムについての感想とどんな用途に適しているかのアンケート調査を

行った。

- (1) 感想 好意的な感想が25件と多かった。中でも、音声インタフェースが便利そうだという意見は多かった。否定的な意見としては、端末が大きい、表示が遅い、映像や音声がきれいではない、などがあった。
- (2) どんな用途に適しているか 展示案内や交通案内がいちばん多かったが、家庭での生活や趣味に関する情報交換に使いたいという希望も意外に多かった。以下に内訳を示す(複数回答)。
 - ・展示案内, 交通案内, 道案内など 42件
 - ・生活・趣味一般(料理, 本の紹介など) 35件
 - ・買い物(通信販売や店舗内で検索) 18件
 - ・医療・福祉 18件
 - ・教育 13件
 - ・製品の不具合対応, 操作説明 9件

4.4 課題

4.4.1 クライアント端末 このシステムは、ユビキタス環境でモバイル端末を介して、マルチモーダルナレッジというリッチなコンテンツを利用しようという野心的なシステムである。一般に、端末の重さと画面の大きさ、表示する映像のリッチさと画面更新スピードはトレードオフとなる。また、タッチパネル、音声コマンド、キーボード、テンキーなど、どの入力インタフェースをメインと考えるかによって、最適な端末の大きさも違ってくる。

このシステムでは、クライアント端末でタッチパネル、音声認識、映像の再生などをすべて可能にするためWIAを採用したが、図5のようなコンテンツであれば、回答者の顔映像を表示しなくてもいいのではないかという指摘もあった。逆に、高齢者用に更に大きくきれいな映像を表示すべきだという指摘もあった。来館者が持ち歩きやすく、使いやすいインタフェースを実現するために、クライアント端末としてどの程度のスペックを実現すべきかを、再度検討する必要がある。

持ち歩き用にはもっと小さい端末を用意し、映像を見せるときは展示物のそばに常設したキオスク端末を利用するという、2種類の端末の併用案も検討していきたい。

4.4.2 新しいナレッジが増えない 導入時にはナレッジを作成するが、その後、実験運用を行いながら、新たなナレッジを増やしていくには障壁があるようである。素材を撮るのがめんどろという点と、オーサリングするのがめんどろという二つの問題がある。

前者に対しては、来館者の質問に該当する回答が登録されていない場合に、展示説明員や専門知識を持つ技術者が

その場で回答し、その回答映像を素材としてインタラクティブにナレッジを追加していく枠組みを検討している。後者に対しては、撮った映像素材から回答シーンを切り出す作業を自動化することにより省力化を図っていく。

5 あとがき

マルチモーダルナレッジ技術を利用した科学館ガイドシステムを開発し、実験運用を行った。今後は、ユビキタスコンピューティング環境の到来に備え、マルチモーダルコンテンツの量産化のためのオーサリング支援技術の開発に注力する。これにより、誰にでも新しい価値の創出が可能となる豊かな社会の実現を目指す。

また、東芝科学館での実験運用の結果を踏まえて、実用化する分野を拡大していく。実用化が望まれる分野は、ユビキタス環境に限らず、家庭、ショッピング、医療・福祉、社内教育など多岐にわたっている。

文献

- (1) 宮澤隆幸,ほか. Bluetooth機器に映像・ナレッジをオンデマンドで配信するシステム MKIDS(Multimodal Knowledge and Information on Demand Service). インタラクシオン2002 予稿集 . 2002 , p.51 - 52 .
- (2) 藤井寛子,ほか. 博物館案内システムにおけるマルチモーダルナレッジオーサリング . 情報学シンポジウム2003 予稿集 . 2003 , p.49 - 52 .
- (3) 鈴木 優,ほか. マルチモーダルナレッジ技術の展示案内システムへの適用 . 人工知能学会誌 . 18 , 2 , 2003 , p.177 - 182 .
- (4) 中山康子,ほか. 知識情報共有システム(KIDS)の開発と実践 - 組織におけるノウハウ共有の促進 . 人工知能学会誌 . 16 , 1 , 2001 , p.64 - 68 .



福井 美佳 FUKUI Mika

研究開発センター 知識メディアラボラトリー 研究主務。
ヒューマンインタフェース技術の研究・開発に従事。情報処理学会会員。
Knowledge Media Lab.



鈴木 優 SUZUKI Masaru

研究開発センター 知識メディアラボラトリー 研究主務。
情報検索技術の研究・開発に従事。情報処理学会会員。
Knowledge Media Lab.



藤井 寛子 FUJII Hiroko

研究開発センター 知識メディアラボラトリー。分散コンピューティング技術の研究・開発に従事。日本ソフトウェア科学会会員。
Knowledge Media Lab.