

J-004

## メタデータによる映像コンテンツの再生制御

## Playback Control of Video by Metadata

加藤 雅弘† 森 靖英† 廣井 和重† 親松 昌幸† 鳥羽 美奈子† 木村 淳一†  
 Masahiro Kato Yasuhide Mori Kazusige Hiroi Masayuki Oyamatsu Minako Toba Junichi Kimura

## 1. まえがき

放送と通信の分野で、映像コンテンツを用いた多様なサービスが始まり、多彩な視聴形態を楽しむことができるようになった[1]。多様なサービスを提供するために、映像コンテンツを構成するシーンにまで立ち入ったきめ細かな検索を可能にするインデキシング技術も提案されている[2]。しかし、製造コストの制約が厳しいTVで楽しむまでに至っていない。本稿では、低コストで映像ナビゲーション機能(動画の再生と同期した関連Webページの自動表示、コーナー検索、シーン検索、ハイライト再生など)を単一のアプリとしてTVに実装する方式を提案した。映像ナビゲーション機能は、動画の分析機能と分析結果を利用した再生制御機能から構成される。本方式は動画の分析結果を記述したメタデータをTVの外部で生成することを特徴とする。この構成により、TVはメタデータの読み込みと記述されたルールに従った再生制御のみで映像ナビゲーションを行うことができ、CPUリソース0.5MHz、メモリ560KBで実行できる。

## 2. メタデータプレーヤ・アーキテクチャ

メタデータをインターフェースとしてメディア処理と再生制御を分離したアーキテクチャを本稿ではメタデータプレーヤ・アーキテクチャと呼ぶ。図2.1にメタデータプレーヤ・アーキテクチャに基づいたTV端末のシステム構成を示す。従来のシステム構成との違いはメタデータを生成するメディア処理部をTV端末の外に出した点である。そこで端末外部から提供されたメタデータを読み込み(図に「パーサ」と記述)、記述ルールに従って再生制御するモジュールを追加する。この構成を採用したTV端末は、メディア処理を行う代わりにメタデータを読み込み、記述ルールに従った再生制御を行って映像ナビゲーション機能を実行する。映像や音響の処理を伴うメディア処理[3][4]を端末外部で行うことにより端末の処理量を削減する。

この構成を採用したTV端末が新たに負担する処理量を試算すると、およそ0.5MHzとなる。これは従来のアーキテクチャにも含まれる再生制御部を除外して、メタデータ読み込み処理の処理量を試算した数値である。すなわち、XMLで記述されたメタデータの開始タグ<要素名>と終了タグ</要素名>で囲まれたデータをメタデータの最小単位としてこれに56Bを割り当てる。本研究で実験的に策定したメタデータは、このデータの最大数を10000としている。従って最大560KBのデータを読み出すことになる。560KBのデータを1回に8Bずつ読み出す場合には全てのデータを読み出すのに70000回の読み出しを行うことになる。1回の読み出しで約100HzのCPUリソースを使うので70000回の読み出しを行うと7MHz使うことになる。キャッシュによるブースト効果を考慮すると0.5MHzのCPUリソースを見積もれば十分である。

ブルーレイディスクなどのパッケージメディアと同様に、コンテ

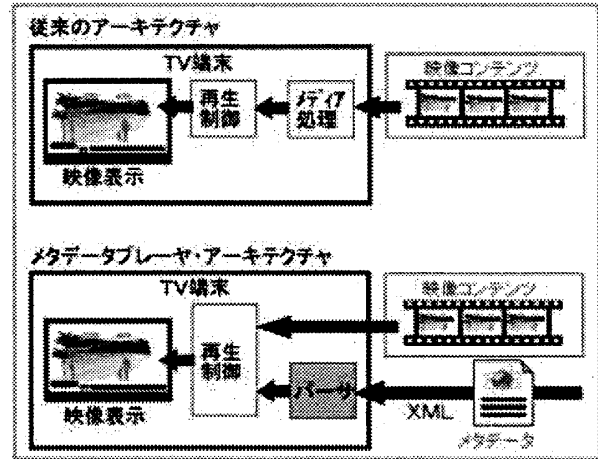


図2.1 メタデータプレーヤ・アーキテクチャ

ンツホルダがメタデータを提供することもできる。さらにポータルサイトで自動生成したメタデータを配信することもできる。このようにメタデータプレーヤ・アーキテクチャを採用すると、メタデータを端末内部で自動生成する従来方式と比べてメタデータを生成する方法の自由度を高めることができる。さらにメタデータの規格を整備すれば流通する可能性もでてくる。メタデータの流通がその品質向上に貢献すればTV端末の楽しみが広がる。

映像コンテンツの著作権を配慮して、著作権を有する映像コンテンツに付与するメタデータの利用を、ユーザが個人的に楽しむ場合に限定する。また、著作権者の許可を得た映像コンテンツに付与するメタデータは、ネットワークから配信する場合とiVDRなどの記憶媒体に保存して配布する場合を想定する。

## 3. GUIの設計

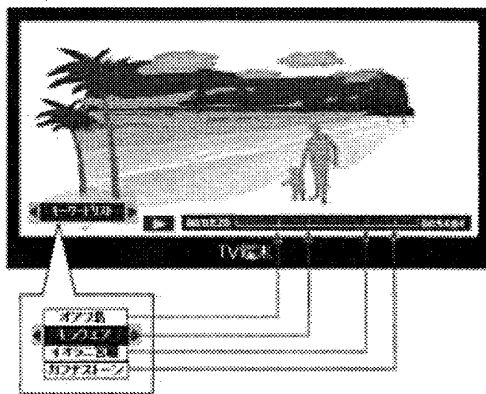
メタデータプレーヤ・アーキテクチャを採用したTV端末のGUIを設計した。メタデータプレーヤ・アーキテクチャを採用すると、安価なプロセッサで映像ナビゲーション機能を搭載だけでなく、動画の再生と同期した関連Webページの自動表示、コーナー検索、シーン検索、ハイライト再生などの異なる機能をほぼ同一のGUIで操作できる。

図3.1はシーン検索機能のGUIを表す。画面左下のメニューボックスに登録したキーワードを選択すると、これに対応付けられたチャプタに遷移する。図の例は、「オアフ島」「キラウエア火山」「イオラニ宮殿」「カフナストーン」などのキーワードをリスト表示している。これらのキーワードを「トピックス」「スポーツ」「天気」などのコーナー名に変えるとコーナー検索機能のGUIになる。また、キーワードを「5分」「10分」「20分」などのハイライト視聴の再生時間に変えるとハイライト再生機能のGUIになる。これらはメタデータを僅かに変更して利用シーンを大きく変える例でもある。

図3.2は関連Webページ表示機能のGUIを表す。

†(株)日立製作所中央研究所

番組のシーンを検索



## キーワードの例

- ・登場人物の名前
- ・アーティストの名前
- ・土地や地域の名前

図3.1 シーン検索機能

関連Webページ表示機能は動画のシーンに対応したWebページをキーワードと対応付ける。動画の再生時刻がキーワードと対応付けたチャプタ位置に到達すると、対応したWebページを表示する。このWebページは動画の再生と同期して出現するキーワードから検索する。キーワードをテロップ[4]や音声などから抽出する。一度に多数のキーワードが抽出された場合には、キーワードの出現頻度を考慮して最適なキーワードを選定する。出現頻度が低いキーワードは、珍しく馴染みが薄いキーワードである可能性が高い。このようなキーワードから検索されたWebページは有意義な情報を提供すると期待した。選定したキーワードから検索したWebページを表示する。キーワードによって特定のポータルサイトで検索できない場合があることを考慮して、ウィキペディア、Yahooニュース、Googleの順に適用して検索できた記事を表示させた。

## 4. ユーザビリティの評価

関連Webページ表示機能を中心にユーザビリティをインタビュー形式で評価した。評価で得られた主な知見を述べる。

関連Webページ表示機能は、報道番組やドキュメンタリー番組を視聴しているときに、地名、人名、時事用語などに対するちょっとした疑問を先取りして、この疑問に答えたWebページを表示するので便利である。その反面、ドラマなどのストーリーを楽しむ番組には不要であることを実感した。

また、番組本編に関連したWebページよりもCMに関連したWebページを表示した場合に、しばしば面白さを感じるようになった。CMに関連したWebページがBGM出力や動画再生を行う場合があり、CM情報を補足する効果がある。

Webページの表示間隔が7秒以下になると視聴の快適性を損なうことが分かった。

画面の表示形態については、映像ナビゲーション機能を実行するために表示したメニューやコントロールは、これらの操作目的が完了して映像視聴に没入した時点で速やかに消去しないと視聴の快適性を損なうことを実感した。

また、メタデータの記述方式については、TV-Anytimeなどの標準方式を参考にするものの、メタデータのタグの階層が増

番組の関連Webページを表示

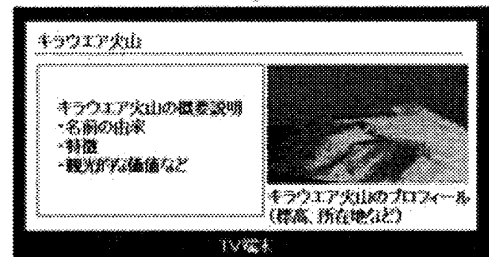
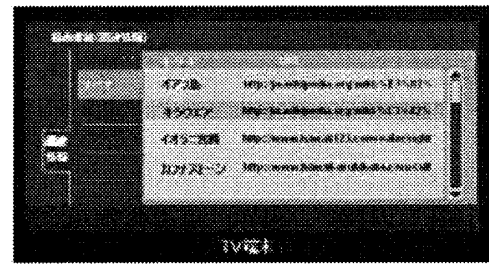


図3.2 関連Webページ表示機能

えたり、記述の冗長性が高まるとXMLパーサのメモリ使用量を増大させてTVの限られたリソースの負担となる。そこで映像ナビゲーションの機能を厳選して必要十分なメタデータを定義した。標準化された記述方式に準拠する必要が生じた場合には変換フィルタで対応する。

## 5. まとめ

(1) 動画の再生と同期した関連Webページの自動表示、コーナー検索、シーン検索、ハイライト再生などの映像ナビゲーション機能を、単一のアプリとして安価なハードウェアに実装することを目的として、メタデータをインターフェースとしてメディア処理と再生制御を分離したメタデータプレーヤ・アーキテクチャを提案した。

(2) メタデータプレーヤ・アーキテクチャをTV端末に採用した場合のメリットは、TV端末でメディア処理を行う場合と比較して要求リソースを減らせることと、機能拡張に必要なコストを低減できることである。このときに、新たに必要となるメタデータ読込処理の処理量を試算した結果、CPUリソース0.5MHz、メモリ560KBとなった。

## 参考文献

- [1] 亀山渉監修: "デジタル・コンテンツ流通教科書", インプレスR&D, 2006.
- [2] 中村大輔, 山中隆広, 片岡充照, 黒木修隆, 沼昌宏, 山本啓輔: "クローズドキャプションを用いた野球映像インデキシング", FIT2007, 一般講演論文集第3分冊, H-077, pp.183-184.
- [3] 親松昌幸, 廣井和重, 森靖英, 加藤雅弘, 木村淳一: "自動ダイジェスト再生機能のためのリプレイシーン検出手法", FIT2007, 一般講演論文集第3分冊, H-080, pp.189-190.
- [4] 平松義崇, 関本信博, 新庄広, 丸川勝美: "図形輪郭除去によるテロップ文字領域抽出", FIT2007, 一般講演論文集第3分冊, H-043, pp.103-104.