

テレビ型 Web 視聴システムの開発

The Development of The Viewing-and-listening System of Television Type Web

灘本 明代¹⁾
Akiyo NADAMOTO

1) 神戸大学大学院自然科学研究科情報メディア科学専攻(〒657-8501 神戸市灘区六甲台町
1-1 E-mail: nadamoto@db.cs.kobe-u.ac.jp)

ABSTRACT. In this paper, I develop the system of the passive Web-browsing by the TV-program metaphor. This facilitates users to browse Web pages like watching TV programs. During the conversion process, which we refer to as programmization, voice and image-data presentations must be synchronized carefully so that they effectively convey the original information. Moreover, the overall presentation needs to follow a particular presentation structure, such as news and variety shows, in order to improve the familiarity of the program. Our approach solves these issues by developing a novel method of converting HTML documents into programs, and designing a new mark-up language, Scripting XML (S-XML), which is suitable for describing programs.

1. 背景

近年のインターネットの目覚ましい発展により Web 上には種々の情報が氾濫し、膨大な情報量となっている。現在使用されている Web の閲覧環境は、利用者に「読む」「スクロールする」「クリックする」等の能動的な操作を要求する Read and Click 型インターフェイスである。この Read and Click 型ユーザインターフェイスでは、多数クリックして複数の Web ページを取得し、それら Web ページ内の情報を読まなければならない。その為、利用者の負担が大きく、Web 情報を容易に閲覧できるとは言い難い。

一方、現在のテレビは「見る」「聞く」の受動的な操作で番組を容易に視聴することができるが、得られる情報はあくまで限られた製作者側の人々から発信された情報である。

そこで、一般の人々が発信する膨大な Web 情報をテレビを視聴するように容易に視聴できると望ましいと考え、チャンネルを選択するように Web ページを選択し、テレビ番組のような番組風コンテンツを作成し Watch and Listen 型インターフェイスであるテレビ型 Web 視聴システムの開発を行った。

テレビ型 Web 視聴システムとは、これまでテキストや画像で表示されていたコンテンツを、音声合成・読み上げやキャラクタアニメーションを用いて表現した番組を作成して視聴することである。

以下、受動的視聴のための番組を作成することを番組化と呼ぶ。

図 1 に番組化によるイメージ図を示す。

図 2 に Web の受動的視聴の例を示す。図 2 の左半分は現在の Web ブラウザによって表示されたコンテンツであり、図 2 の右半分はそのコンテンツを受動的視聴した画面を示している。

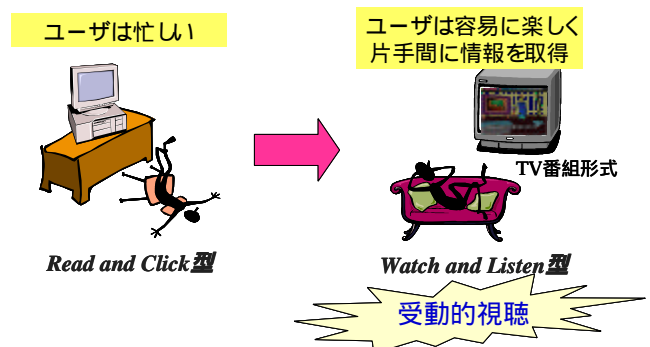


図 1 番組化のイメージ図



図 2 Web の番組化例

2. 目的

本開発では、膨大な Web 情報を容易に閲覧する事を目的とし、チャンネルを選択するように Web ページを選択し、テレビ番組のような番組風コンテンツを作成し Watch and Listen 型インターフェイスで受動的視聴の可能となる方式の開発を行った。

Web の番組化による利点を下記に示す .

- 容易に Web の閲覧が受動的に出来る .
- 楽しみながら Web の閲覧が出来る .
- コンピュータを技術的に利用できない人も Web 情報を取得出来る .

これまで Web の受動的な情報取得には PointCast^[1] のような PUSH 型のものがある . これに対し , 番組化は Web 情報の音声読み上げを行うため , 聞き流したり , 別のことをしながら片手間に容易に Web の情報を取得できるだけでなく , キャラクターアニメーションを使用しているため , 番組として楽しみながら視聴する事が可能である . また , これまで「クリック」することすら出来ない人も , Web の情報の取得が可能となる .

本事業では , 上記背景および目的に基づき下記 3 機能の開発を行った .

1. 番組表作成機能開発
2. 自動変換機能開発
3. 番組メタファ機能開発

図 3 に開発イメージ図を , 図 4 に開発機能関係図を示す .

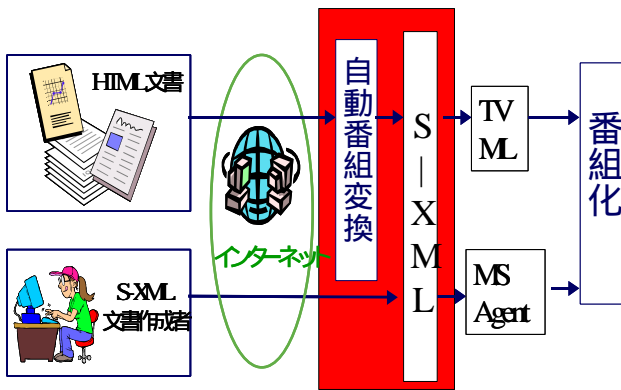


図 3 . 開発イメージ図

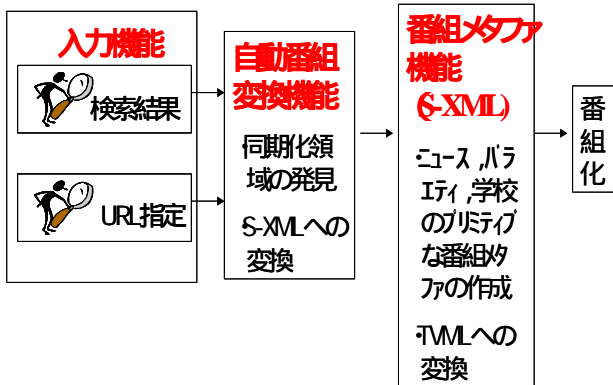


図 4 開発機能関係図

3 . 番組表作成機能

本機能は , 番組化を行いたい Web を指定するユーザーインターフェイス部である . URL を直接指定することにより , 指定された Web ページの番組化を行う機能を開発した . ユーザーインターフェイスを図 5 に示す .

URL 指定部は直接 URL を指定する . 指定された URL の Web ページを画像を含めてダウンロードする .

プリミティブ指定部ではニュース / パラエティ / 教室 , エージェントのあらかじめ用意したプリミティブな演出を選択する . 選択された演出は S-XML に渡され , S-XML にて演出を行う .

保存先指定部は番組化変換を行ったファイルの保存先指定をおこなう . 実際には画像 , タグ解析結果ファイル , S-XML ファイルもこの指定されたディレクトリに保存される .

言語指定部は指定した URL の Web ページが英語のページの場合選択する .

タグ解析結果表示部は自動変換時に解析したタグの木構造を表示する . 図 6 に表示したウィンドウを示す .

番組自動変換部は指定した URL の Web ページを番組に自動変換する .

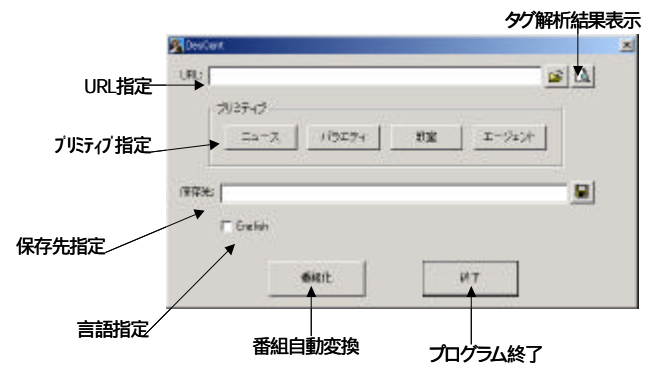


図 5 ユーザーインターフェイス

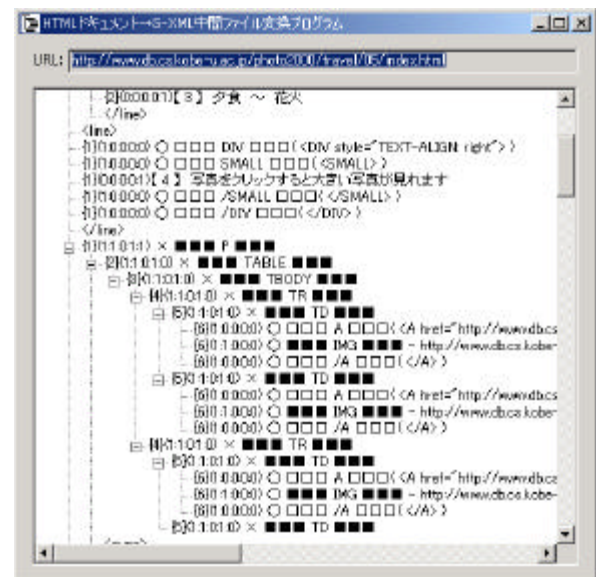


図 6 タグ解析結果表示ウィンドウ

4 . 自動変換機能

ユーザが番組表より指定した Web ページからこれまで , 我々が開発し , 本事業においてバージョンアップを行った , スクリプト作成マークアップ言語である Scripting-XML (以下 , S-XML という)^{[2][3]} に変換する機能である .

自動変換機能においては時間軸上の同期の技術的課

問題を解決した。時間軸上の同期とは、通常 Web ブラウザは時間軸を気にする必要のない 2 次元空間での表示であるのに対し、番組化することにより、Web 上の画像と文章の同期や文章と文章の同期などの時間軸を考慮する必要が生じる。

現在の Web 情報はほとんどが HTML で作成されている為、これらの時間軸上での同期は考慮されていない。しかしながら、通常画像とそれに関連する文章は HTML 文書中の近いところにあると考え、本開発では、これまで研究してきた HTML のタグ解析による同期化領域の発見方法^{[2][3]}を用いて解決した。

図 6 に自動変換の手順を示す。

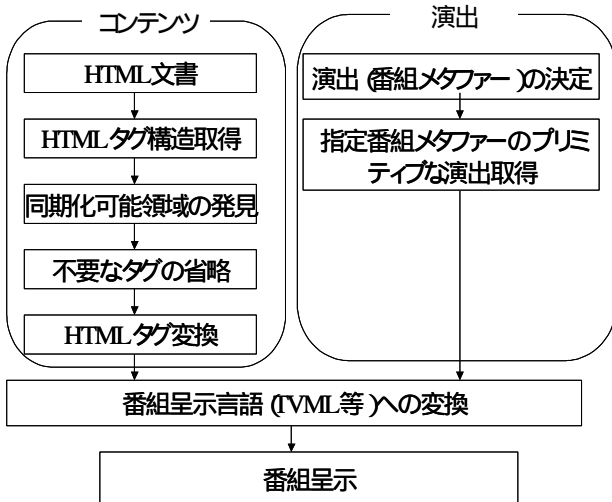


図 6 自動変換の手順

自動変換の機能の特徴を下記に示す。

- HTML 文書のタグ解析を行う。
- HTML 文書の論理構造の縦型探索のシリアライゼーションを行う。
- 画像と文章、文章と文章の時間軸同期を行う。
- リンク先ページの情報を要約してリンク元ページの番組にマージする。
- 番組メタファによるプリミティブな演出を用意する。

(1) HTML タグ解析

現在の Web ページはテキストを主にして情報を表示する為、HTML タグはテキスト表示のためのタグ構成に

なっている。これに対し、番組化では、テキスト表示をキャラクタ間の対話形式の音声に変え、時間軸の同期を取り、そして演出を加えて番組として Web 情報を呈示する。HTML 内の文章は台詞として変換するが、番組化を行うために下記の 4 種類のタグ解析が必要となる。

- 演出効果となる HTML タグ
- 表示不要となる HTML タグ
- 画像の HTML タグ
- 台詞となる HTML タグ

演出効果となる HTML タグ

HTML タグには、テキスト表示のため、や<I>、<BIG>に代表されるフォントに関するタグがある。

HTML 文書内では、これらの開始タグには、各種パラメータの指定が含まれている。

現在のテキスト表示ベースの Web ページ上ではこれらのタグを用いて文章や単語を強調している。強調した文章や単語は Web ページ内のキーワードである可能性が高い。

したがって、番組化においても強調して利用者に呈示しなければならない。

そこで、これらのタグはパラメータごとに、音声や音量を設定する。各タグの設定パラメータが変わることで、音声を変えたり、音量を大きくして強調してユーザに知らせるようになる。

表示不要となる HTML タグ

<P>や<CENTER>等の Web ページ上のテキスト整形に用いられるタグは音声にて情報を呈示する番組化には必要ない。これらのタグは HTML から番組化への変換時に不要となるため削除する。

画像の HTML タグ

タグは画像ファイルをリンクしている。この画像と同期化可能な文章領域を発見するために画像タグを取得する。

台詞となる HTML タグ

上記で示したタグ群以外のタグは台詞となるタグ領域として取得する。

表 1 にタグ変換の一覧を示す

表 1 HTML タグ変換

タグ	HTML における特徴	番組化への操作
演出効果となる HTML タグ群		
<H >	見出し文字サイズを指定するものであり、パラグラフのキーワードとして利用可能。	見出し語は各ページの特徴を表わすものとして、タイトルと同程度に重要であると考えられる為、音量を大きくする。
	SIZE や COLOR の指定ができるが、これは HTML ソースの作成者が意図して行っていることで、文書全体における重要度は高い。	SIZE や COLOR の指定が行われている部分は、「」(括弧)をつけた言葉に置き換え、番組上でキャラクタの台詞が表示された時に、タグで囲まれていた部分を強調するため、音声や音量を変えて演出を行う。
<I><BIG>等	これらは全て語句を強調する為に用いるタグ。	これらをユーザが意図する強調レベルに応じて抜き出し、音量を変えて番組に変化を与えていく。

<CITE>	語句の引用を行うタグ。	ユーザがこのタグを利用する事を明示した場合においては引用文を番組に導入する。
 	これらは直前の文書の説明を箇条書きにしたものが多い。	直前部分の文章の内容に応じて表現するメタファーを変えることが可能。また、箇条書きということから、内容的に対等なものの列挙であることが多いので、これに適したメタファーの適用が考えられる。
不要となるHTML タグ群		
<U><SUP><TT>等	アンダーライン等	番組呈示には不要となるタグ群である
<P><HR><CENTER> <PRE>等	これらは Web ページ内のテキスト整形を行う。	番組呈示にはテキスト整形は必要ないため不要となるタグ群である。
<SCRIPT><FORM><FRAMESET><ADDRESS>等	これらは Web ページ内各種設定を行う。	番組呈示には不要となるタグ群である。

(2) 同期化可能領域の発見

これまで研究してきた同期化領域の発見方法^{[2][3]}を用いて、本システムでは画像と同期する文章を発見した。

(1) で述べたように、HTML のタグ構造を解析し、そのタグ構造で表される論理構造の縦型探索を行い番組化するが、時間軸上に情報を呈示する番組化において、時間軸的同期が非常に重要な問題となってくる。画像とそれに関係する文章の同期が取れていなければ、その Web 情報は誤って伝わることになる。

同期が必要である画像と文章は、HTML 文書において互いに近い距離にあると考えられる。そして、HTML タグは階層構造をもっており、文書や画像をある程度グループ分けすることが可能である。これらのことに注目し HTML タグ構造を解析すると、画像と文章を含む極小なタグ領域を発見することにより、画像と文書の同期化可能領域を発見できる。

画像 i に対する、極小タグ領域 $R(i)$ を下記に示す条件から $S(i)$ をすべて満たすものと定義した。 $R(i)$ は任意のタグ $\langle X \rangle$ と $\langle /X \rangle$ で囲まれた領域である。すなわち

極小タグ領域 $R(i)$ は任意のタグ $\langle X \rangle$ と $\langle /X \rangle$ で囲まれた領域である。

$$R(i) = \langle X \rangle \dots \langle /X \rangle$$

$R(i)$ は画像 i を含む。

$$R(i) \quad i$$

$R(i)$ は少なくとも一つの文章 s を含む。

$$R(i) \quad s$$

$R(i)$ のいかなる部分タグ領域 $\langle Y \rangle, \dots \langle /Y \rangle$ も $S(i)$ を満たさない

画像 i に対する極小タグ領域 $R(i)$ を用いて、画像 i と同期すべき文章を含む同期化可能領域 $S(i)$ を次のように定義する。

[a] $S(i)$ は $R(i)$ に含まれる。

$$S(i) \quad R(i)$$

[b] $S(i)$ は i と少なくとも一つの文章 s を含む。

$$S(i) \quad i \quad (s) S(i) \quad s$$

[c] $S(i)$ は (a), (b) を共に満たす極小な部分領域である。

ここで (c) で述べた「部分領域」とは任意の連続部分領域を意味しており、先に定義した極小タグ領域とは異なる。

り、先頭と末尾が $\langle X \rangle, \dots \langle /X \rangle$ のように対応するタグになっている必要はない。

5. 番組メタファ機能

番組化において、キャラクタの台詞や動き、カメラや照明の配置、スタジオセット等の演出が必要である。現在の Web にはそのような演出は考慮されていないため、本開発ではニュース、バラエティ、学校、エージェントの番組メタファを用いて、あらかじめプリミティブな演出を S-XML を用いて解決した。

また、自動番組変換機能では、既存の Web ページを自動番組変換により番組化を行うことについて述べた。しかしながら、現在の HTML では同期化や演出的要素を明示的に記述することができないため、自動的な番組変換を行ってもすべての Web ページに対してそれが十分であるとは限らない。そこで、Web 作成者が番組化を意識した上で Web ページを作成することが望ましいと考え、本章では、すでに我々が提唱してきた番組化の台本作成のための、スクリプト作成マークアップ言語である Scripting-XML (S-XML)^{[2][3]} の改良を行った。

S-XML は画像や動画と台詞の同期化や演出的要素の設定などをふまえた、番組化を行うことを目的とした言語であり、以下にその利点を記す。

- Web 作成者が番組化を意識した上で文書を作成することが可能である。番組化において重要なポイントである同期と演出を設定するタグセットを用意しているため、Web 作成時に番組化の設定を行うことが可能である。
- TVML や SMIL, FLASH 等の種々の同期化コンテンツ再生環境に対応可能である。画像と文章との同期タグセットを用意しているため、S-XML のタグを種々の同期化コンテンツ再生環境のタグに変換するだけでコンテンツを再生することが可能である。現在の S-XML は通常の Web ブラウザと TVML により表示することができるが、将来的には、SMIL や FLASH 等の同期化コンテンツを用いての表示も可能とする予定である。
- XML をベースにしているため、種々の文書とのリン

クが可能である。S-XML は XML の拡張タグ集合であるため、既存の XML 文書とのリンクが容易である。このため S-XML は Web 上にある文書のみならず XML で書かれた、もしくは変換されたプレゼンテーション用資料やデータベースのデータ等とリンク構造を用いて関連付けを行うことで番組化することが可能である。さらに、逆にそれらの XML 文書を番組を構成する素材として管理し、別の番組に再利用することも可能である。

- 将来的にデジタル放送を意識した様々な番組化が可能なアプリケーションに適応させることができる。すでに標準化されているデジタル放送用の拡張制限を入れた BML や BXML、またダイジェスト映像を対象として研究されている PPML との融合が可能である。

S-XML は、演出を容易に設定すること、XML で書かれた他の文書を容易に番組化できること、および番組化を目的とした文書を他のアプリケーションでも使用できることの3つの理由から、番組内容を記述するための要素をコンテンツ識別系タグとスタイル系タグの2つの構成に分別して定義する。

(1) コンテンツ識別系タグ

番組化コンテンツ識別系タグとは、番組のタイトル、XML 文書内の文章や画像、動画などの番組で使用するコンテンツの指定、選択されたコンテンツ同士の同期対応付け、強調部分の指定、要約に必ず使用する部分の指定を行うためのタグである。コンテンツ識別系タグの一覧とその特徴を表2に示す。

表2 コンテンツ識別系タグの特徴

タグ名	特徴
title	タイトルを囲み、キーワードを属性として持つ。キーワードは検索に利用する。
word	ある特定の単語に対応させて表示する画像を指定する。その指定したファイルを同時に提示させたい単語を checkword タグで囲む。
line	台詞としてキャラクターにしゃべらせたい部分を囲む。
img	ファイル名、同期を取るための番号、複数画像の提示方法、配置位置(中央、左方、右方、上方、下方)を指定する。複数画像の提示方法は以下の中から選択する。
movie	ファイル名、同期を取るための番号、動画の開始、終了、再生時間を指定する。
insert	挿入したいファイルを指定する。指定したファイルが複数ある場合は、指定した順に挿入する。
hide	XML 文書としては使用したいが番組では表示したくない部分を囲む。
hlink	リンクのアンカー文字をタグで囲み、番組ではタイトル、Web ページではリンク先を表すアンカー文字として扱う。

(2) スタイル系タグ

スタイル系タグとは、XML 文書に番組化のための演出を付加するタグ群である。スタイル系タグはテンプレート系タグとプリミティブ系タグがあるが、本開発において

はテンプレート系タグの改良を行った。

システムは、あらかじめ Character タグで番組内で登場するキャラクターのモデル、声の初期設定を行う。そして、news、variety等のテンプレート系タグで、それぞれのテンプレートに当てはめた演出を付加したい部分を囲む。さらに、各テンプレート系タグ固有の子要素タグを利用すれば、ストーリー展開やキャラクターの台詞の割り振りを指定することができ、文書作成者の意図を番組に反映させることも可能である。番組形式によって演出や進行方法は異なると考えられるため、テンプレート系タグで指定した範囲のコンテンツに適用する、スタジオセットやカメラワーク、キャラクターの台詞の割り振り等の演出の自動生成は、各テンプレートに応じたアルゴリズムに従って行われる。例えば、variety タグの子要素として、起承転結タグがある。このタグはストーリー展開を指定するタグであり、このタグに囲まれたコンテンツは、はじめは落ち着いていたキャラクターがそわそわし始めて、騒ぎだし、最後には力づくでおとなしくさせられてしまうような演出が付加される。以下に各テンプレート系タグの基本コンセプトとそれぞれに固有の子要素タグの説明を加える。

- news メインキャラクターとサブキャラクターに分け、話す量に違いを与える。堅い雰囲気。カメラは基本的に話し手のキャラクターをズームアップ。
 - topic: 主題用の演出部分。
 - detail: 主題の詳細用の演出部分。
 - main: 予め指定しておいた main キャラクターにしゃべらせたい部分。
 - sub: 予め指定しておいた sub キャラクターにしゃべらせたい部分。
- school 先生と生徒がいるある教室での出来事。先生と生徒の間で質疑応答を繰り返す。
 - teacher: 先生にしゃべらせたい部分
 - student: 生徒にしゃべらせたい部分
- variety 相づちを入れ、落ち着かないキャラクターが、常に動き回ったりそわそわしたりする。
 - intro: 起。まだ落ち着いている。
 - development: 承。少しそわそわし始める。
 - turnstory: 転。さらに動きが激しくなる。
 - conclusion: 結。強引におとなしくさせる。

6. 考察

本開発において、音声合成、CG アニメ等を用いて Web 情報をテレビ番組のように受動的に容易に視聴できる「テレビ型 Web 視聴システム」の開発を行うことを目的とし、下記の機能の開発を行った。

- 番組表作成機能開発
- 自動変換機能開発
- 番組メタファ機能開発

番組表作成機能開発では、URL 指定による Web の取得システムと検索結果受動的視聴呈示による Web の取得システムを開発した。

本機能の開発により、Web の番組化を作成するにあたり、URL を直接指定して番組化する方法と番組表のような検索結果を用いて番組化する方法の2つの方法を可能とした。

自動変換機能開発では、Web の HTML 文書のタグ解析を

行うことにより，画像と文章の同期化可能領域の発見を可能とし，HTML 文章を自動で番組記述マークアップ言語である S-XML に変換することを可能とした．

番組メタファ機能開発においては，プリミティブな番組を S-XML にて作成し，番組記述言語（NHK 技術研究所開発）ある TVML^{[4][5]} に変換し，番組化することを可能とした．

しかしながら，本開発において，開発した同期化可能領域の発見はあらゆる HTML 文書に対応しているものではなく，今後更なる同期化領域の発見のアルゴリズム開発が必要である．

また，Web を番組化するにおいて，平叙文である HTML 文書を会話文に変換し，より通常の番組に近づける課題が残り，今後更なる開発を行ってゆきたいと思う．

7．参考文献

- [1] PointCast ホームページ：
<http://www.pointcast.com/>
- [2] 灘本 明代，服部 多栄子，近藤 宏行，沢中 郁夫，田中 克己：Web コンテンツの受動的視聴のための自動変換とスクリプト作成マークアップ言語，情報処理学会論文誌，データベース，Vol.42No.SIG1 (TOD8)，p.p. 103-116，2001 ．
- [3] 服部 多栄子，沢中 郁夫，灘本 明代，田中 克己： Web の受動的視聴のための同期化可能領域の発見と番組化用マークアップ言語 S-XML ，情報処理学会研究報告，Vol.2000，No.44 00-DBS-121-2，pp.9-16 2000 年 5 月 ．
- [4] NHK 放送技術研究所: TVML ホームページ
<http://www.strl.nhk.or.jp/TVML/indexj.html>
- [5] 林 正樹: 番組記述言語 TVML を使った情報の番組化，情報処理学会 DBS 研究会技術報告，Vol.2000，No.10，2000-DBS-120-13，pp.91-98，2000 年 1 月 ．