

## 音楽 CD・WWW 連携システム ～WebSync/CD の開発～

大泉 俊雄 的場 ひろし 前野 和俊  
NEC C&C メディア研究所

本論文では、音楽 CD と WWW 上のホームページを連携させたサービスを提供するための基盤システムである WebSync/CD について述べる。WebSync/CD は、音楽 CD の再生に合わせて、予め作成したシナリオに従った一連のホームページの連動表示や、音声・動画ファイルの同期再生が可能であり、音楽 CD を軸とした新しいマルチメディアコンテンツが作成可能である。WebSync/CD には、CD-EXTRA 等の特殊な CD フォーマットを採らない通常の音楽 CD に対してもコンテンツ提供できる、WWW 上のコンテンツを随時更新できる、と言ったメリットがある。さらに、音声パターンマッチング技術を開発し、ラジオ放送、オーディオテープ等からアナログ音声信号を PC に入力する形態においても、楽曲の特定、並びに、再生位置の検出を可能とした。

### A Development of WebSync/CD: A Method for Synchronization of Music and WWW

Toshio OIZUMI, Hiroshi MATOBA, Kazutoshi MAENO  
C&C Media Research Laboratories, NEC Corporation  
4-1-1 Miyazaki, Miyamae-ku, Kawasaki, Kanagawa, 216-8555, Japan  
E-mail: {ohizumi,matoba,maeno}@ccm.cl.nec.co.jp

This paper describes a development of "WebSync/CD", which provides synchronization of Music and WWW. By using the WebSync/CD system, creators can produce a new type of contents. According to play back the music, the system show a series of related homepages and play sound or video files downloaded from WWW. The advantage of the WebSync/CD system is that it can apply to CD-DAS which don't adopt a special structure such as CD-EXTRA and it can update the contents on WWW at any time. Even if users use analog equipments such as a radio or an audio tape, WebSync/CD can be used by inputting the sound signal from analog equipments to a PC. In that case, the system can identify the music and its position currently play back by sound pattern matching technology.

#### 1.はじめに

本論文では、音楽 CD と WWW(ホームページなど)の連携による音楽 CD のマルチメディア化を実現した PC ベースのシステムである"WebSync/CD"について述べる。WebSync/CD は、CD-EXTRA フォーマットを利用した従来の音楽 CD のマルチメディア化とは異なり、音楽 CD に含まれる楽曲の再生と、WWW 上のホームページの表示や、音声・動画の再生を関連付けることにより、マルチメディアコンテンツの制作を可能にするシステムである。さらに、音声パターンマッチング技術<sup>(1)</sup>を用いて、ラジオ放送やオーディオテープ等のアナログ音声信号に対してもコンテンツ提供を可能にしている。

以下、従来手法と比較した WebSync/CD の特徴について 2.で、楽曲管理・コンテンツ管理方式、コンテンツ

連動方式を中心とする WebSync/CD の動作方式について 3.で、音声パターンマッチング技術とその性能について 4.で、WebSync/CD コンテンツのオーサリング環境について 5.で、それぞれ述べた後、6.で、WebSync/CD の実用化状況と今後の課題についてまとめる。

#### 2.WebSync/CD の概要と特徴

##### 2.1.既存の類似システムとその問題点

既存の音楽 CD のマルチメディア化手法として、CD-EXTRA フォーマットの利用が挙げられる。CD-EXTRA フォーマットとは、CD 媒体にサウンドトラックとデータトラックを共存させるフォーマットであり、このデータトラックには、macromedia DIRECTOR<sup>(2)</sup>

などで作成したアプリケーションソフトやコンテンツデータを格納しておくことが可能である。この手法は、CD-ROM 制作と同様の考えで、十分な期間と費用をかけてマルチメディアコンテンツを制作する場合には有効だが、現在の音楽 CD 制作体系の中では、以下のような問題点がある。

- (1) プレス用原盤を作成する前に、データトラックに格納すべき内容（アプリケーションソフトやコンテンツデータ）を完成させなければならない（コンテンツ制作日程は、楽曲のレコーディング日程に左右されることになる）。また、レコード会社以外の第三者によるコンテンツの提供は困難である。
- (2) 一旦、プレスされるとデータトラックの内容の書き換えは不可能なので、コンテンツの内容の更新は困難である(DIRECTORにあるShockedCDの仕組みを使えば、一部のコンテンツデータをWWW上から取り込むことが可能であるが、アプリケーションソフトが規定するコンテンツ全体のシナリオを変えることは出来ない)。
- (3) データトラックに格納されるアプリケーションソフトでは、PC内蔵CD-ROM装置で再生する音楽CDの楽曲に対してのみ、コンテンツを対応させている(外部オーディオ機器で音楽CDが再生される場合は、この手法によるマルチメディア化の対象外である)。

## 2.2. WebSync/CD によるコンテンツサービス形態

WebSync/CD のサービス形態を図 1 に示す。コンテンツ提供者は、サービス対象となる音楽 CD の楽曲に合わせたコンテンツ(ホームページデータや音声・動画データなど)を作成して、WWW サーバ上に公開する。ユーザが、PC 内蔵 CD-ROM 装置に WebSync/CD コンテンツが公開されている音楽 CD をセットすると、その音楽 CD に対応したコンテンツが WWW ブラウザに表示される。そして、音楽 CD を再生すると、コンテンツ提供者の作成したシナリオに従って自動的に、ホームページが更新され、音声・動画データが再生される。

また、ラジオ放送やオーディオテープから流れる楽曲に合わせて、ホームページの連動表示や音声・動画の同期再生を実現する場合は、オーディオ機器と PC の間をオーディオケーブルで接続する。WebSync/CD は音声パターンマッチング技術により、PC に入力されたアナログ音声信号から楽曲の再生位置を特定するため、ユーザは、PC 内蔵 CD-ROM 装置で音楽 CD を再生する場合と同様のコンテンツサービスが受けられる。

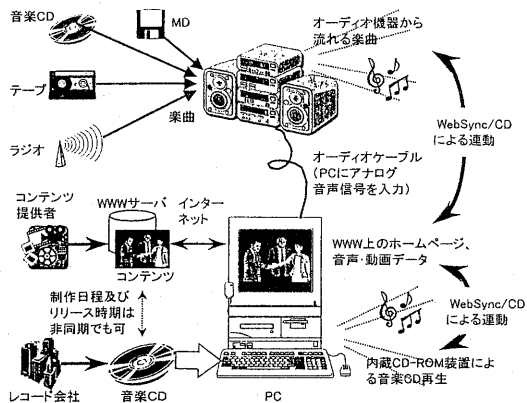


図 1. WebSync/CD サービス形態

図 2 は、Netscape Navigator、Internet Explorer 上で動作する Plug-in ソフトウェアとして実装した WebSync/CD ビューワによるコンテンツ画面例である。



音楽CDの再生に連動するホームページ(各フレーム毎に、独立して連動が可能)

図 2. WebSync/CD コンテンツ画面例

## 2.3. WebSync/CD の特徴とメリット

WebSync/CD による音楽 CD のマルチメディア化には、以下のような特徴がある。

- (1) 連動に必要なファイルを音楽 CD の媒体中に含めず、WWW 上から取得するファイルをもとにメディア・コンテンツ連携を実現している。これにより、
  - ・ 楽曲のレコーディング日程に左右されずに、コンテンツ制作が可能、レコード会社以外の第三者であっても、コンテンツを提供可能、既に発売済みの音楽 CD に対しても、コンテンツが提供可能
  - ・ 異なる音楽 CD に収録された同じ楽曲を同一の楽曲として認識可能、同じ楽曲に対して複数のコンテンツを提供可能

といった、メリットがある(2.1の(1)の問題点を解決)。

- (2) 音楽 CD(楽曲)の再生に合わせて、シナリオに従った一連のホームページの連動表示や、音声・動画の同期再生が可能である。これにより、
- ・サーバ上のシナリオを変更することで、いつでもホームページが連動表示するタイミングや、音声・動画データを同期再生するタイミングを変更可能
  - ・サーバ上のホームページや、音声・動画データを変更することで、いつでも最新の情報を提供可能といった、メリットがある(2.1の(2)の問題点を解決)。
- (3) 音声パターンマッチング技術により、楽曲がアナログ音声信号により PC に入力される場合でも、その楽曲を特定し再生位置を検出できる。これにより、
- ・ PC 内蔵 CD-ROM 装置でなく、外部のオーディオ機器で音楽 CD を再生した場合でも、ホームページとの連動が可能
  - ・ ラジオ放送で流れる楽曲や、オーディオテープや MD などに録音した楽曲を再生する場合でも、ホームページとの連動が可能
- といったメリットがある(2.1の(3)の問題点を解決)。

### 3.WebSync/CD の動作方式

#### 3.1.WebSync/CD の動作の流れ

図 3は、PC 内蔵 CD-ROM で音楽 CD を再生して、ホームページを連動表示させる場合の WebSync/CD の動作の流れを示している。WebSync/CD ビューフは、CD 制御、楽曲管理、コンテンツ管理、シナリオ解釈の4つのモジュールからなる。WebSync/CD ビューフ自身は、Plug-in ソフトウェアの他、ActiveX コントロール、JavaApplet、独立したプロセスを持つアプリケーションとしても実装できる。

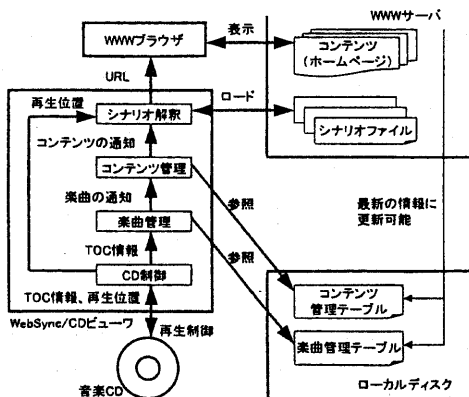


図 3. WebSync/CD の動作の流れ

ユーザが CD をセットすると、CD 制御モジュールが音楽 CD の TOC (Table of Contents) 情報を読み取り、楽曲管理モジュールに通知する。その後、楽曲管理モジュールは、楽曲管理テーブルを参照して楽曲を特定し、コンテンツ管理モジュールは、コンテンツ管理テーブルを参照してコンテンツを決定する(楽曲管理・コンテンツ管理方式)。シナリオ解釈モジュールは、コンテンツ毎に準備されたシナリオをロードして、音楽 CD の再生に合わせたホームページの URL を WWW ブラウザに通知する(シナリオによるコンテンツ連動方式)。

以下、この2つの方式に関して、詳しく説明する。

#### 3.2. 楽曲管理・コンテンツ管理方式

WebSync/CD では、楽曲管理テーブル、コンテンツ管理テーブルの2つの表データファイルをローカルディスクに保持している。各テーブルは、コンテンツ参照時に WWW サーバ上から最新情報が追加される。

楽曲管理テーブルは、TOC 情報により識別された個々の音楽 CD の各トラックごとに、楽曲を対応させた表である。楽曲の特定を TOC 情報から直接行うのではなく、このような対応表を介在させることにより、異なる音楽 CD に収録された同じ楽曲を同一のものと認識できるようにする(図 4.参照)。

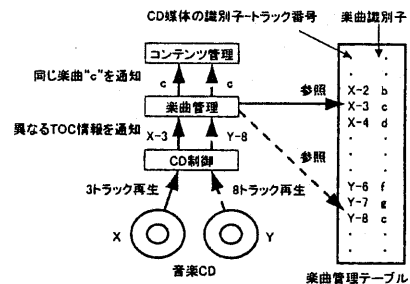


図 4. 楽曲の特定

コンテンツ管理テーブルは、各コンテンツごとに、楽曲を対応させた表であり、1つの楽曲に、複数のコンテンツが対応可能である。楽曲管理モジュールから通知を受けた楽曲に、複数のコンテンツが対応していた場合、コンテンツ管理モジュールは、ユーザにコンテンツ選択を要求し、選択されたコンテンツをシナリオ解釈モジュールに通知する。コンテンツ管理テーブルに、コンテンツの公開日時、最終アクセス日時などを付加情報として記録し、時間的に最後に公開されたコンテンツ、ユーザが最後にアクセスしたコンテンツなどの条件で、コンテンツ管理モジュールが自動的にコンテンツを選択することも可能である(図 5.参照)。

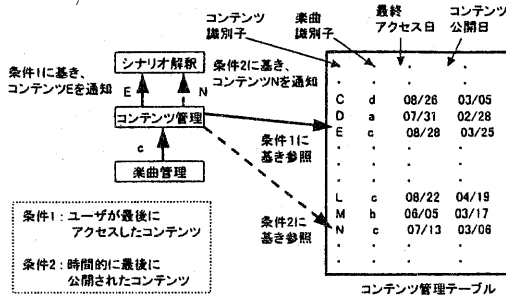


図 5. コンテンツの選択

### 3.3. シナリオによるコンテンツ連動方式

WebSync/CD では、楽曲の再生位置 (タイミング) と表示すべき URL (ホームページ) の対応表であるシナリオファイルを、コンテンツごとに WWW サーバ上に用意する (フレーム構造を利用したコンテンツの場合は、タイミングに対応する URL は、各フレーム毎に個別に記述できる)。

シナリオ解釈モジュールは、コンテンツ管理モジュールから通知を受けたコンテンツの存在する WWW サーバにアクセスし、そのコンテンツのシナリオファイルをロードする。ロード後、シナリオ解釈モジュールは、シナリオファイルの記述に従って、CD 制御モジュールから通知される再生位置に応じた URL を WWW ブラウザに通知する。WWW ブラウザが、通知に基づいてホームページを連動表示することにより、音楽 CD とホームページの連携コンテンツが実現する (図 6. 参照)。

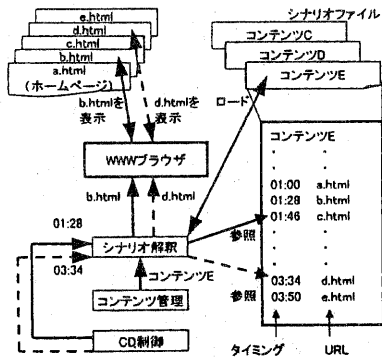


図 6. 音楽 CD とホームページの連動表示

シナリオファイルには、タイミングと URL の対応表の他に、タイミングと、音声・動画データのファイル名、再生位置、音量などの対応表を保持することができる。シナリオ解釈モジュールは、URL を WWW ブラウザに通知するのと同様に、音声・動画データのファイル名とその再生位置、音量などを音声・動画データ再生外部モジュール (WebSync/CD ビューワの公開する接続 IF を

利用した外部モジュールで、Plug-in として実装することでホームページ上に表示可能) に通知する。音声・動画データ再生外部モジュールが、通知に基づいて音声・動画データを同期再生することにより、音楽 CD と音声・動画データの連携コンテンツが実現する (図 7. 参照)。

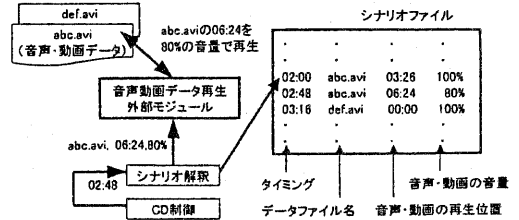


図 7. 音楽 CD と音声・動画データの同期再生

シナリオファイルや、コンテンツデータ (ホームページや音声・動画データなど) は、WWW サーバに存在するので、コンテンツ提供者がいつでも更新可能である。シナリオ解釈モジュールは、シナリオファイルのロードと共にコンテンツデータのキャッシングを開始し、連動表示や同期再生のサーバアクセスによる遅延を最小限に押さえている。

## 4. 音声パターンマッチング処理方式

### 4.1. 音声パターンデータ

WebSync/CD は、以下に述べる手法により、楽曲全体に渡って生成した音声パターンデータを記録した音声パターンファイルを、楽曲毎にローカルディスクに保持している (最新の音声パターンファイルは、WWW サーバから随時入手できる)。

図 8 は、音声パターンデータを作成する過程を示したものである。PC に取り込まれるアナログ音声信号を 1 秒毎に A/D 変換し、PCM データを生成する。PCM データからは、サンプリング周波数 66kHz の FFT (Frequent Fourier Transform) によってスペクトラムを生成する。スペクトラムを 3 バンド (低域: 43Hz ~ 430Hz、中域: 430Hz ~ 860Hz、高域: 860Hz ~ 5.5MHz) に分割して、サンプリングデータを各バンドごとに 2 値に量子化したデータが音声パターンデータである。

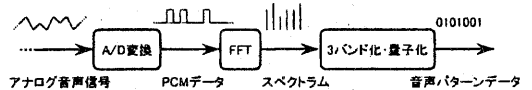


図 8. 音声パターンデータの生成

生成される音声パターンデータは、1 時間あたり容量 90K バイトとなり、PCM データと比べると非常に小さなデータ量となるので、データロードの負荷を軽減し、マッチング処理計算の高速化をはかれる。

#### 4.2. 音声パターンマッチング処理方式

図 9 は、1 秒毎に処理される音声パターンデータを処理し、候補となる音声パターンファイル中の該当位置を検出する様子を表した図である。処理は、検索フェーズと確認フェーズに分けられる。

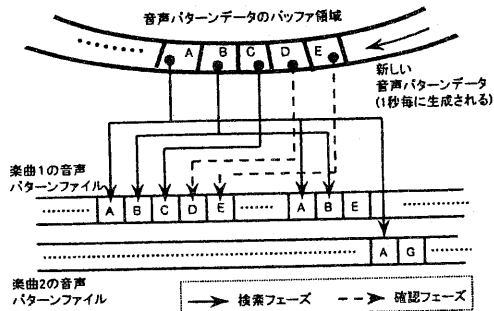


図 9. 音声パターンマッチング処理

アナログ音声信号の入力があると、4-1.で説明した手法により、音声パターンデータが1秒毎に生成され、最初、検索フェーズで処理される。検索フェーズにおいては、各音声パターンファイルに含まれる各音声パターンデータとの照合を行い、合致する位置すべてを候補位置として抽出する。次の1秒の処理においては、この候補位置の次の1秒の音声パターンデータに対してのみ照合を行い、基本的にこの過程を繰り返すことで最終的に候補位置を1つに絞り込む(1つに絞り込まれた候補位置を確定位置と呼ぶ)。

確定位置が存在する場合、次の1秒の処理は、確認フェーズで行う。確認フェーズでは、確定位置の次の1秒の音声パターンデータのみと照合を行い、合致しているかどうか判定する。合致していれば、その位置を確定位置として、次の1秒の処理を引き続き確認フェーズで行い、合致していなければ、次の1秒の処理は再度、検索フェーズから始める。

上記の2つのフェーズの処理を行うモジュールが、音声パターンマッチング処理モジュールであり、確定位置を楽曲の再生位置として上位モジュールに通知する。WebSync/CD ビューワ内においては、音声パターンマッチング処理モジュールは、CD 制御モジュールと同様の位置づけであり、楽曲の再生位置特定後の動作は、3.で説明した PC 内蔵 CD-ROM 装置で音楽 CD を再生した場合と同様である(図 10.参照)。

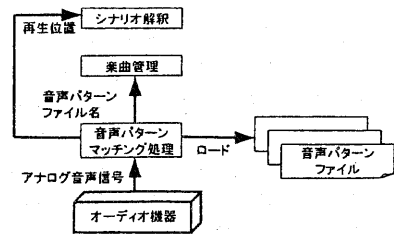


図 10. 音声パターンマッチング処理モジュール

#### 4.3. 音声パターンマッチングの性能

音声パターンマッチング処理方式の性能を確認するために、PC(Pentium200MHz の Windows95 マシン)とオーディオ機器をオーディオケーブルで接続した環境で、適当に選択した 30 曲の楽曲(1曲あたり平均 4 分 52 秒)の音声パターンファイルを作成した(結果として、144 分相当の音声パターンデータが検索対象となる)環境で、以下の実験を行った。

- (1) 上記の環境下で、候補となる音声パターンデータの容量を変えて、選択された 30 曲の楽曲のみが録音されたオーディオテープ中の任意の 50 箇所を再生して再生位置が特定されるまでの所要時間(検索フェーズでの処理時間)の平均を計測した結果が図 11.である。音声パターンデータの容量が大きくなるほど、所要時間が増大する。

音声パターンデータ容量 (楽曲数)	検索フェーズでの 平均処理時間
46 分 35 秒相当 (10 曲)	2.4 秒
99 分 12 秒相当 (20 曲)	3.5 秒
144 分相当 (30 曲)	5.0 秒

図 11. 音声パターンデータ容量と処理時間の関係

- (2) 同じ環境下で、候補となる音声パターンデータの容量を変えて、選択された 30 曲が全く含まれない楽曲が録音されたオーディオテープを 120 分再生したときに、再生位置が特定された(誤答した)回数を、測定した結果が図 12.である。音声パターンデータの容量が大きいくほど、誤答回数が増える。

音声パターンデータ容量(楽曲数)	誤答回数
46 分 35 秒相当 (10 曲)	35
99 分 12 秒相当 (20 曲)	64
144 分相当 (30 曲)	89

図 12. 音声パターンデータ容量と誤答回数の関係

- (3) WebSync/CD では、検索フェーズで確定位置が決定しても、即、再生位置として上位モジュールに通知せず、確認フェーズで数秒、確定位置の継続が

確認されてから、再生位置として上位モジュールに通知することによって、誤答を減らす方策をとっている(なお、仮に誤答が起こっても、確認フェーズで誤答を検出できるため、誤答状態が長時間続くことはない)。図 13.は、候補となる音声パターンデータ容量が 144 分(30 曲)相当分だった場合に、確認フェーズでの待ち時間を変えて、誤答回数を測定した結果である。待ち時間を増やすと、誤答回数は減る。

確認フェーズでの待ち時間	誤答回数
1 秒	26
2 秒	9
3 秒	1

図 13. 待ち時間と誤答回数の関係

以上の結果から、現在の WebSync/CD では、

- ・ 1 曲あたり平均 5 分程度の楽曲 30 曲を対象
- ・ 2 時間程度の楽曲再生中に誤答が平均 10 回以下を実用上の基準と仮定すると、検索フェーズでの所用時間平均 5 秒と、確認フェーズでの待ち時間 2 秒の合計、平均 7 秒で、楽曲の再生位置を特定できると言える。

## 5. WebSync/CD におけるコンテンツ制作

WebSync/CD におけるコンテンツとは、一連のホームページの集合である。コンテンツを作成するには、既存のホームページを利用する場合と、新しくホームページを作成する場合とが考えられるが、新しく作成する場合でも、簡易なホームページ作成ツールは多く世に存在するので、ホームページ作成自体は困難な作業ではなくなってきた。

WebSync/CD 固有の作業として、大部分の工数を占めるのは、楽曲の再生位置(タイミング)とホームページ(URL)の対応表であるシナリオファイルの作成である。そこで、シナリオファイルの作成支援に重点を置いたオーサリングツールを開発した(図 14.参照)。

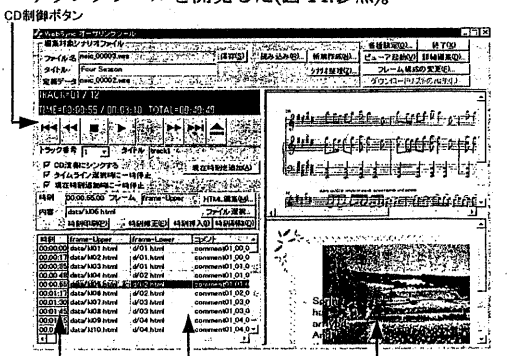


図 14. WebSync/CD オーサリングツール利用例

このオーサリングツールには、

- ・ 実際に楽曲を再生しながら、ホームページを表示するタイミングを設定できる
- ・ 表示するホームページを楽曲の再生に合わせてプレビューできる
- ・ グリッド表示することで、タイミングの一部分をカット&ペーストしたり、一括してタイミングをずらしたりできる

といった特徴があり、簡単にシナリオファイルが作成できる。

## 6. まとめ

本論文では、WWW との連携による音楽 CD のマルチメディア化を実現した PC ベースのシステムである「WebSync/CD」の動作方式について説明し、CD-EXTRA フォーマットを利用した従来の音楽 CD のマルチメディア化手法に比べて、メリットがあることを示した。また、音声パターンマッチング技術による楽曲の再生位置の特定を行うことで、PC 内蔵 CD-ROM 装置で再生される音楽 CD のみでなく、ラジオやテープから出力される楽曲に対しても、サービス可能であることを示した。

WebSync/CD は、1998 年 8 月から実用化されており、「01MESSENGER」(サザンオールスターズ、Victor Entertainment 発売)、「SURVIVE」(B'z、RoomsRecords 発売)をはじめとする、7 タイトルの音楽 CD に対して、8 つのコンテンツが公開されている<sup>6)</sup>。

また、WebSync/CD に用いられている音声パターンマッチング技術は、TV 番組や、VCR に記録された映像と WWW(ホームページ)との連携サービス<sup>4)</sup>としても展開されており、Discovery Channel で 1998 年 1 月に放送された番組「Titanic～タイタニック号悲劇の検証～」に対してコンテンツ提供サービスが行われた。

今後、音楽 CD とストリーミングデータとの連携、異なるマシン間での音楽 CD の再生制御の連携など、機能拡張実装を進め、さらに充実したコンテンツ制作が可能なる環境を整備していく予定である。また、音声パターンマッチング技術においては、マッチング精度向上と、検索速度向上のための改良をさらに進めていく。

### 参考文献

- 1 H. Matoba, T. Oizumi, K. Maeno: "WebSync: A System for Synchronization of Analog Sequential Media and WWW", IEEE PROMSMmNet'97
- 2 <http://www.macromedia.co.jp>
- 3 <http://www.neic.co.jp/websync>
- 4 大泉, 的場, 前野 "TV 番組・WWW 連携システム WebSync の開発", 情報処理学会第 55 回全国大会