

(19) 日本国特許庁(JP)

## (12) 公開特許公報(A)

(11) 特許出願公開番号

特開2005-107529

(P2005-107529A)

(43) 公開日 平成17年4月21日(2005.4.21)

(51) Int.C1.<sup>7</sup>

F 1

テーマコード(参考)

**G10L 15/00**  
**G06K 1/12**  
**G06K 7/00**  
**G06K 19/00**  
**G06K 19/06**

G10L 3/00 551B  
G06K 1/12 B 5B072  
G06K 7/00 U 5D015  
G06T 7/00 300F 5D045  
G06K 19/00 A 5L096

審査請求 未請求 請求項の数 45 O L (全 29 頁) 最終頁に続く

(21) 出願番号	特願2004-278356 (P2004-278356)	(71) 出願人	000006747
(22) 出願日	平成16年9月24日 (2004.9.24)	株式会社リコー	
(31) 優先権主張番号	506263	東京都大田区中馬込1丁目3番6号	
(32) 優先日	平成15年9月25日 (2003.9.25)	(74) 代理人	100070150
(33) 優先権主張国	米国(US)	弁理士	伊東忠彦
(31) 優先権主張番号	506303	(72) 発明者	ジョナサン ジー ハル
(32) 優先日	平成15年9月25日 (2003.9.25)		アメリカ合衆国, カリフォルニア 94025, メンロ・パーク, サンド・ヒル・ロード 2882番, スイート 115 リ
(33) 優先権主張国	米国(US)		コー イノベーション インク内
(31) 優先権主張番号	506411	(72) 発明者	ジャメイ グラハム
(32) 優先日	平成15年9月25日 (2003.9.25)		アメリカ合衆国, カリフォルニア 94025, メンロ・パーク, サンド・ヒル・ロード 2882番, スイート 115 リ
(33) 優先権主張国	米国(US)		コー イノベーション インク内
(31) 優先権主張番号	814844		
(32) 優先日	平成16年3月30日 (2004.3.30)		
(33) 優先権主張国	米国(US)		

最終頁に続く

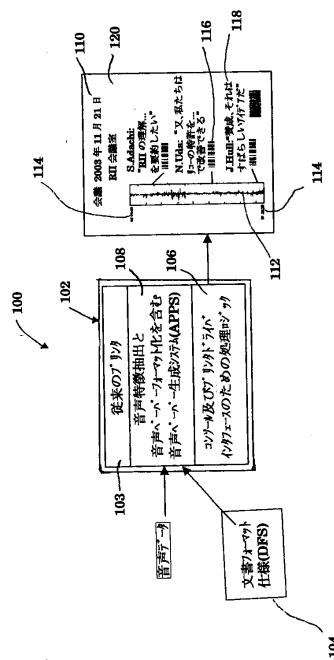
(54) 【発明の名称】 時間ベースメディアのための印刷可能表現

## (57) 【要約】

【課題】 本発明のシステムは、ユーザが時間ベースのメディアの表現を生成することを可能にする。

【解決手段】 本発明のシステムは、メディアコンテンツから特徴を抽出するための特徴抽出モジュールを含む。例えば、特徴抽出モジュールは音楽演奏におけるソロを検出することができる、又は、音楽、拍手のある発話等を検出することができる。フォーマット化モジュールはシステムにより生成されたメディア表現をフォーマットする。フォーマット化モジュールは又、表現に特徴抽出情報を適用し、表現仕様に従って表現をフォーマットする。更に、システムは、特徴抽出情報と表現仕様とに基づいて、メディア表現を生成する拡張出力装置を含むことができる。本発明の方法は、メディアコンテンツから特徴を抽出する段階と、表現フォーマットを指定するデータ構造又は仕様に基づいて及び抽出された特徴を用いて生成されるメディア表現をフォーマット化する。この方法は又、フォーマット化の結果に基づいて、メディア表現を生成する段階を含むことができる。

【選択図】 図1



**【特許請求の範囲】****【請求項 1】**

時間ベースのメディアの表現を生成するためのシステムであって：

メディアコンテンツから特徴を抽出するための特徴抽出モジュール；及び

前記システムにより生成されたメディア表現をフォーマット化するためのフォーマット化モジュールであって、前記メディア表現に特徴抽出情報を適用するために前記特徴抽出モジュールに通信可能であるように結合され、前記フォーマット化モジュールは表現仕様に従って前記メディア表現をフォーマット化する、フォーマット化モジュール；

から構成されることを特徴とするシステム。

**【請求項 2】**

請求項 1 に記載のシステムであって、前記特徴抽出モジュールは、メディアコンテンツにおける特徴を認識するためのコンテンツ認識ソフトウェアから更に構成される、ことを特徴とするシステム。

**【請求項 3】**

請求項 1 に記載のシステムであって、拡張出力装置ドライバインターフェースを制御するための処理ロジックから更に構成される、ことを特徴とするシステム。

**【請求項 4】**

請求項 1 に記載のシステムであって、拡張出力装置コンソールを制御するための処理ロジックから更に構成される、ことを特徴とするシステム。

**【請求項 5】**

請求項 1 に記載のシステムであって、前記メディア表現はデジタルフォーマットにおいて生成される、ことを特徴とするシステム。

**【請求項 6】**

請求項 1 に記載のシステムであって、前記メディア表現はペーパーフォーマットにおいて生成される、ことを特徴とするシステム。

**【請求項 7】**

請求項 1 に記載のシステムであって、前記メディア表現は、ユーザがメディアコンテンツにアクセスするようにする少なくとも 1 つのユーザ選択可能識別子を含む、ことを特徴とするシステム。

**【請求項 8】**

請求項 7 に記載のシステムであって、前記の少なくとも 1 つのユーザ選択可能識別子は、メディア表現について印刷された少なくとも 1 つのバーコードから構成される、ことを特徴とするシステム。

**【請求項 9】**

請求項 7 に記載のシステムであって、前記の少なくとも 1 つのユーザ選択可能識別子は、対応するメディアコンテンツをプレイするように選択されることができる、ことを特徴とするシステム。

**【請求項 10】**

請求項 8 に記載のシステムであって、前記の少なくとも 1 つのバーコードは、表示装置において対応するメディアコンテンツをプレイするために前記バーコードを走査することによりメディア表現において選択されることができる、ことを特徴とするシステム。

**【請求項 11】**

請求項 1 に記載のシステムであって、前記メディア表現は、タイムラインに沿ってメディアコンテンツのグラフィック表現を含む、ことを特徴とするシステム。

**【請求項 12】**

請求項 11 に記載のシステムであって、音声コンテンツのグラフィック表現は音声波形タイムラインで表示される、ことを特徴とするシステム。

**【請求項 13】**

請求項 11 に記載のシステムであって、前記タイムラインは、メディアコンテンツのユーザ選択セグメントに対応する長さに沿ってマーカーを含む、ことを特徴とするシステム

10

20

30

40

50

。

**【請求項 1 4】**

請求項 1 1 に記載のシステムであって、前記タイムラインは音声コンテンツのセグメントに対応する長さに沿ってマーカーを含み、前記セグメントはメディアコンテンツにおける特定の特徴の検索により定義される、ことを特徴とするシステム。

**【請求項 1 5】**

請求項 1 1 に記載のシステムであって、前記タイムラインはメディアコンテンツのセグメントに対応する長さに沿ってマーカーを含み、少なくとも 1 つの前記マーカーは、メディアコンテンツのセグメントを表すテキスト情報を有する、ことを特徴とするシステム。

**【請求項 1 6】**

請求項 1 1 に記載のシステムであって、前記タイムラインはメディアコンテンツのセグメントに対応する長さに沿ってマーカーを含み、少なくとも 1 つの前記マーカーは、メディアコンテンツのセグメントを表すタイムスタンプ情報を有する、ことを特徴とするシステム。

**【請求項 1 7】**

請求項 1 に記載のシステムであって、前記メディア表現は前記メディアコンテンツを表すヘッダを含む、ことを特徴とするシステム。

**【請求項 1 8】**

請求項 1 に記載のシステムであって、前記メディア表現は、データ構造に含まれるフォーマット仕様に従って生成される、ことを特徴とするシステム。

**【請求項 1 9】**

請求項 1 8 に記載のシステムであって、前記データ構造に含まれる前記フォーマット仕様は、メディア表現について印刷されたグラフィック表現のフォーマットを指定する多くのユーザ定義可能フィールドから構成される、ことを特徴とするシステム。

**【請求項 2 0】**

請求項 1 8 に記載のシステムであって、前記データ構造に含まれる前記フォーマット仕様は、メディア表現のレイアウトを指定する多くのユーザ定義可能フィールドから構成される、ことを特徴とするシステム。

**【請求項 2 1】**

請求項 1 8 に記載のシステムであって、前記データ構造に含まれる前記フォーマット仕様は、メディア表現に含まれるメディアコンテンツマーカーを指定する多くのユーザ定義可能フィールドから構成される、ことを特徴とするシステム。

**【請求項 2 2】**

請求項 2 0 に記載のシステムであって、前記データ構造に含まれる前記フォーマット仕様は、前記メディアコンテンツに適用される特徴抽出を指定する多くのユーザ定義可能フィールドから構成される、ことを特徴とするシステム。

**【請求項 2 3】**

請求項 1 に記載のシステムであって、メディア表現を生成するための拡張出力装置であって、メディア表現の生成のための命令を受けるためのフォーマット化モジュールに通信可能であるように結合されている、拡張出力装置から更に構成される、ことを特徴とするシステム。

**【請求項 2 4】**

請求項 2 3 に記載のシステムであって、前記拡張出力装置はペーパーにメディア表現を印刷するためのプリンタを含む、ことを特徴とするシステム。

**【請求項 2 5】**

時間ベースのメディアの表現を生成するための方法であって：

メディアコンテンツから特徴を抽出する段階；及び

表現仕様に従って表現をフォーマット化する段階であって、特徴抽出情報を適用する手順を含む、段階；

から構成されることを特徴とする方法。

10

20

30

40

50

**【請求項 2 6】**

請求項 2 5 に記載の方法であって、メディアコンテンツの表現を生成する段階から更に構成される、ことを特徴とする方法。

**【請求項 2 7】**

請求項 2 5 に記載の方法であって、メディアコンテンツの特徴を抽出する段階は、メディアデータにおいてキーワード検索を実行する手順から更に構成される、ことを特徴とする方法。

**【請求項 2 8】**

請求項 2 5 に記載の方法であって、メディアコンテンツの特徴を抽出する段階は、メディアデータにおける発話認識を実行する手順から更に構成される、ことを特徴とする方法。

**【請求項 2 9】**

請求項 2 5 に記載の方法であって、メディアコンテンツの特徴を抽出する段階は、メディアデータにおけるイベント検出を実行する手順から更に構成される、ことを特徴とする方法。

**【請求項 3 0】**

請求項 2 6 に記載の方法であって、表現を生成する段階は、タイムラインに沿ってメディアコンテンツのグラフィック表現を生成する手順から更に構成される、ことを特徴とする方法。

**【請求項 3 1】**

請求項 3 0 に記載の方法であって、グラフィック表現を生成する段階は、タイムラインに沿って音声コンテンツ波形を生成する手順から更に構成される、ことを特徴とする方法。

**【請求項 3 2】**

請求項 2 6 に記載の方法であって、表現を生成する段階は、ユーザがディアコンテンツにアクセスするようにするメディアコンテンツにおいて少なくとも 1 つのユーザ選択可能識別子を生成する手順から更に構成される、ことを特徴とする方法。

**【請求項 3 3】**

請求項 3 2 に記載の方法であって、前記の少なくとも 1 つのユーザ選択可能識別子を生成する手順は、メディア表現について印刷される少なくとも 1 つのバーコードを生成する手順から更に構成される、ことを特徴とする方法。

**【請求項 3 4】**

請求項 3 3 に記載の方法であって、表示装置において対応するメディアコンテンツをプレイするためにバーコードを走査することによりメディア表現において少なくとも 1 つのユーザ選択可能識別子を選択する手順から更に構成される、ことを特徴とする方法。

**【請求項 3 5】**

請求項 3 0 に記載の方法であって、タイムラインに沿ってグラフィック表現を生成する手順は、タイムラインに沿ってマーカーを生成する段であって、前記マーカーはユーザ選択メディアコンテンツに対応する、段から更に構成される、ことを特徴とする方法。

**【請求項 3 6】**

請求項 3 0 に記載の方法であって、タイムラインに沿ってグラフィック表現を生成する手順は、タイムラインに沿ってマーカーを生成する段であって、少なくとも 1 つのマーカーはメディアコンテンツから抽出された特徴に対応する、段から更に構成される、ことを特徴とする方法。

**【請求項 3 7】**

請求項 3 0 に記載の方法であって、タイムラインに沿ってグラフィック表現を生成する手順は、タイムラインに沿ってマーカーを生成する段であって、少なくとも 1 つのマーカーはメディアコンテンツを表すテキスト情報を含む、段から更に構成される、ことを特徴とする方法。

**【請求項 3 8】**

10

20

30

40

50

請求項 30 に記載の方法であって、タイムラインに沿ってグラフィック表現を生成する手順は、タイムラインに沿ってマーカーを生成する段であって、少なくとも 1 つのマーカーはメディアコンテンツを表すタイムスタンプ情報を含む、段から更に構成される、ことを特徴とする方法。

【請求項 39】

請求項 26 に記載の方法であって、表現を生成する段階は、メディアコンテンツを表すヘッダを生成する手順から更に構成される、ことを特徴とする方法。

【請求項 40】

請求項 26 に記載の方法であって、表現を生成する段階は、デジタルフォーマットで表現を生成する手順から更に構成される、ことを特徴とする方法。

10

【請求項 41】

請求項 26 に記載の方法であって、表現を生成する段階は、ペーパーフォーマットで表現を印刷する手順から更に構成される、ことを特徴とする方法。

【請求項 42】

請求項 25 に記載の方法であって、表現仕様に従って表現をフォーマット化する段階は、フォーマット仕様を有するデータ構造を用いてメディア表現のフォーマットを定義する手順から更に構成される、ことを特徴とする方法。

【請求項 43】

請求項 25 に記載の方法であって、音声コンテンツと表現仕様とを入力する段階から更に構成される、ことを特徴とする方法。

20

【請求項 44】

請求項 33 に記載の方法であって、少なくとも 1 つのバーコードを生成する手順は、識別子情報を含むバーコード画像をレンダリングするためにバーコード生成アルゴリズムを適用する段から更に構成される、ことを特徴とする方法。

【請求項 45】

請求項 33 に記載の方法であって、少なくとも 1 つのバーコードを生成する手順は、タイムスタンプ情報を含むバーコード画像をレンダリングするためにバーコード生成アルゴリズムを適用する段から更に構成される、ことを特徴とする方法。

30

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、時間ベースのメディアに対する印刷可能表現を生成するためのシステム及び方法に関する。

【背景技術】

【0002】

従来のプリンタは、種々の異なるフォーマットの文書を生成するために且つ異なるタイプのコンテンツに基づいて、現在、用いられている。しかしながら、従来のプリンタがテキストの画像及び / 又はピクチャを生成することができる一方、従来のプリンタには、マルチメディアコンテンツの表現を効果的に生成するための能力において限界がある。従来のプリンタは、紙のような固定的媒体に印刷し、それ故、従来のプリンタは時間ベースのメディアの要素を効果的に捕捉することができない。

40

【0003】

しかし、今日、一般に、時間ベースのメディアコンテンツを容易に調べるための能力に対する要求が存在している。時間ベースのメディアコンテンツにおける所望の特徴を検索するためには、コンテンツそのものを実際に検索する必要があり、所望の情報を見つけるために走り読みをする必要がある。例えば、ユーザは、特定のトピックに関するコンテンツを見つけるために又は特定の話者による議論を見つけるためにラジオのトークショーの音声記録をマニュアル操作で走り聴きする必要がある。このような従来のプリンタにおける制約のために、メディアコンテンツから興味のある特定の特徴を特定し且つ抽出するように、ユーザが長いメディアセグメントを通して検索することができる簡単な方法は、

50

現在、存在しない。更に、メディアに関する有用な情報を提供するメディアの容易に書き込み可能な表現を、ユーザが生成できる方法は存在しない。

#### 【0004】

更に、メディアコンテンツは、一般に、デジタル方式のみにおいて利用可能である。しかしながら、多くのユーザにとって、デジタルフォーマットは、どれが見るべき情報であるかの点で最適なフォーマットではない。デジタル方式のメディア情報を見ることが一部のユーザにとって適切である一方、多くのユーザは、情報が紙媒体に印刷されるとき、情報をより容易に理解することができることに気がつく。それにも拘らず、ユーザがメディアコンテンツを調べてそれにアクセスすることができる時間ベースのメディアについて紙ベースの表現を生成するための機構で現在利用可能なものは存在しない。

10

#### 【0005】

従って、必要とされるものは、紙ベースとすることができる且つマルチメディアコンテンツにおいて規定された特徴を抽出するための能力をユーザに提供することができる時間ベースのメディアの表現を生成するためのシステム及び方法である。

#### 【発明の開示】

##### 【課題を解決するための手段】

#### 【0006】

本発明は、時間ベースのメディアの表現を生成するためのシステム及び方法を用いて、先行技術の欠落点及び制約点を克服する。本発明のシステムは、メディアコンテンツから特徴を抽出するための特徴抽出モジュールを含む。例えば、特徴抽出モジュールは、音楽の演奏のソロを検出することができる、又は音楽、拍手、発話等を検出することができる。フォーマット化モジュールは、システムにより生成されるメディア表現をフォーマット化する。フォーマット化モジュールは又、メディア表現に特徴抽出情報を適用し、表現指定に従ってその表現をフォーマット化する。更に、システムは、特徴抽出情報とメディア表現指定とに基づいて、メディア表現を生成する拡張出力装置を含むことができる。メディア表現は、紙ベースのフォーマット、デジタルフォーマットまたはいずれの他の表現フォーマットの状態で生成することができる。生成されたメディア表現は、メディアコンテンツタイムラインに沿ったポイントに任意アクセスすることができるユーザ選択可能識別子を含むことができる。

20

#### 【0007】

本発明の方法は、メディアコンテンツから特徴を抽出する段階と、表現フォーマットを指定するデータ構造又は仕様に基づいて及び抽出された特徴を用いて、メディア表現をフォーマット化する段階とを含む。その方法は又、フォーマット化の結果に基づいて、メディア表現を生成する段階を含む。

30

#### 【発明を実施するための最良の形態】

#### 【0008】

時間ベースのメディアの表現を生成するためのシステム及び方法について説明する。本発明の実施形態に従って、プリンタは、特徴抽出情報を組み込むことができる時間ベースのメディアの表現を生成する。更に詳細には、プリンタは、マルチメディア情報についての視覚的表現を生成する文書を生成するために、及びマルチメディア記録におけるポイントへの任意アクセスを可能にするインデックスを生成するために、フォーマット仕様、特徴抽出及びフォーマット化アルゴリズムを組み込む。

40

#### 【0009】

本発明の目的のために、用語“メディア”、“マルチメディア”、“マルチメディアコンテンツ”、“マルチメディアデータ”又は“マルチメディア情報”は、テキスト情報、グラフィック情報、アニメーション情報、音声（オーディオ）情報、映像情報、スライド情報、ホワイトボード画像情報及び他のタイプの情報のいずれか1つ又はそれらの組み合わせをいう。例えば、テレビジョン放送の映像記録は、映像情報と音声情報とから構成されることが可能である。特定の例においては、映像記録は又、クローズドキャプションド（C C : c l o s e d c a p t i o n e d ）テキスト情報から構成され、そのC Cテキ

50

スト情報は映像情報に関する要素から構成され、多くの場合、映像記録の音声部分に含まれる発話の正確な表現である。マルチメディア情報は又、1つ又はそれ以上のオブジェクトから構成される情報を参照するために用いられ、それらオブジェクトは異なるタイプの情報を含む。例えば、マルチメディア情報に含まれるマルチメディアオブジェクトは、テキスト情報、グラフィック情報、アニメーション情報、音声（オーディオ）情報、映像情報、スライド情報、ホワイトボード画像情報及び他のタイプの情報から構成されることが可能である。

#### 【0010】

本発明の目的のために、用語“プリント”又は“プリントティング”は、あるタイプの媒体への印刷をいうとき、プリントティング、書き込み、描画、インプリンティング、エンボシング、デジタルフォーマットの生成及び他のタイプのデータ表現の生成を含むことを意図している。又、本発明の目的のために、システムにより生成される出力は、“メディア表現”、“マルチメディア文書”、“マルチメディア表現”、“文書”、“ペーパー文書”若しくは“映像ペーパー”又は“オーディオペーパー”をいう。単語文書及びペーパーが以上の用語で呼ばれる一方、本発明におけるシステムの出力は、例えば、ペーパー媒体のような物理的媒体等に限定されない。それに代えて、上記の用語は、ある有形の媒体に固定されるいずれの出力をいうことができる。一部の実施形態においては、本発明のシステムの出力は、物理的ペーパー文書においてプリントされたマルチメディアコンテンツの表現とすることができます。ペーパーフォーマットにおいて、マルチメディア文書は、ペーパーの高解像度と可搬性を利用し、マルチメディア情報の読み取り可能な表現を提供する。本発明の教示するところに従って、マルチメディア文書は又、マルチメディア情報を選択し、検索し、それにアクセスすることが可能である。他の実施形態においては、システムは、デジタルフォーマット又はある他の有形の媒体に出力することができます。更に、本発明の出力については、デジタルフォーマットにおけるマルチメディア情報を記憶するいずれの記憶ユニット（例えば、ファイル）を引き合いに出すことができる。マルチメディア情報を記憶するために、種々の異なるフォーマットを用いることが可能である。これらのフォーマットは、M P E G フォーマット（例えば、M P E G 1、M P E G 2、M P E G 4、M P E G 7等）、M P 3 フォーマット、S M I L フォーマット、H T M L + T I M E フォーマット、W M F ( W i n d o w s (登録商標) M e d i a F o r m a t ) 、R M ( R e a l M e d i a ) フォーマット、Q u i c k t i m e フォーマット、S h o c k w a v e フォーマット、種々のストリーミングメディアフォーマット、エンジニアリング共同体、専用及び慣用フォーマット並びに他のフォーマットにより発達したフォーマットを含む。

#### 【0011】

“1つの実施形態”又は“実施形態”の明確化における基準は、実施形態に関連して述べられる具体的な特徴、構成、特性が本発明の少なくとも1つの実施形態に含まれることを意味している。明細書の種々の部分における

“1つの実施形態において”との表現がある場合、必ずしも、全てが同じ実施形態を参照しない。

#### 【0012】

下の説明においては、説明目的のために、本発明を理解することを通じて提供されるように、多くの具体的な詳細例を示している。しかしながら、それらの具体的な詳細例がなくとも、本発明を実施できることは、当業者には理解できるであろう。他の例においては、本発明を分かり難くすることを回避するように、ブロック図の方式で構造及び装置を示している。例えば、本発明について、オーディオコンテンツを主に参照して説明し、プリンタにより生成される表現は、しばしば、音声ペーパーと呼ぶこととする。しかしながら、たとえ、以下の説明が音声コンテンツ及び音声ペーパーにのみ関連する特徴についてのものであっても、本発明の特徴は、いずれのタイプのメディアコンテンツに適用することができ、紙ベースのフォーマット以外のフォーマットにおけるメディア表現を対象とすることができる。

10

20

30

40

50

## 【0013】

ここで、図1を参照するに、時間ベースのメディアの表現を生成するための例示としてのシステム100を示している。この実施形態において、マルチメディア表現を生成するための拡張出力装置またはプリンタ102を示している。プリンタ102は、種々の構成要素から構成され、従来型のプリンタ103、音声ペーパー生成システム(APPS)108及びプリンタコンソールのため及びプリンタドライバインターフェースのための処理口ジック106を含んでいる。

## 【0014】

プリンタ102は、音声データのようなマルチメディアデータを受信し、このコンテンツは、システム100にアクセス可能であるマルチメディア文書に記憶されることが可能である。マルチメディアコンテンツは、システム100に直接記憶されることが可能である、又はシステム100によりアクセスされることができる外部の記憶装置又はサーバ(図示せず)に記憶される情報であることが可能である。他の実施形態においては、眼耳地メディア文書にアクセスすることに代えて、システム100は、マルチメディア情報ソースからマルチメディア情報のストリーム(例えば、ストリーミングメディア信号、ケーブル信号等)を受信することができる。システム100にマルチメディア情報を提供することができるソースの例としては、テレビジョン、テレビジョン放送受信器、ケーブル受信器、映像受信器、デジタル映像受信器、携帯デジタル端末(PDA)等がある。例えば、マルチメディア情報のソースは、マルチメディア放送信号を受信し且つシステム100にその信号を送信するように構成されるラジオとして具現化されることが可能である。この例においては、情報ソースは、システム100にライブのラジオ放送供給情報を供給するラジオ受信器/アンテナであることが可能である。情報ソースは又、システム100に記録された映像及び/又は音声ストリームを供給する、例えば、ビデオレコーダ/プレーヤ、DVDプレーヤ、CDプレーヤ等の装置であることが可能である。代替の実施形態においては、情報のソースは、システム100に捕捉されたプレゼンテーション又はミーティングの情報のストリームを供給することができる、プレゼンテーション又はミーティングレコーダであることが可能である。更に、マルチメディア情報のソースは、外部ソースからのマルチメディア情報を捕捉又は受信し、次いで、更なる処理のためにシステム100に捕捉されたマルチメディア情報を供給する(例えば、無線リンクにより)ように構成された受信器(例えば、衛星放送用アンテナ又はケーブル受信器)であることが可能である。

## 【0015】

マルチメディアコンテンツは、例えば、Real Player(登録商標)、Microsoft Windows(登録商標) Media Player等の専用の又はカスタマイズされたマルチメディアプレーヤから発せられることができる。代替の実施形態においては、システム100は、マルチメディア情報ソースにより受信されるマルチメディア情報信号を途中で捕らえるように構成されることが可能である。システム100は、マルチメディア情報ソースからマルチメディア情報を直接受信することができる、又は、それに代えて、通信ネットワーク(図示せず)により情報を受信することができる。

## 【0016】

再び、プリンタ102の構成要素を参照するに、図1においては、プリンタ102の従来型プリンタ103の構成要素を示している。プリンタ102の従来型プリンタ103の構成要素は、例えば、インクジェットプリンタ、レーザプリンタまたは他の印刷装置の一部または全てを含むことができる。従って、従来型プリンタ102は、ペーパー文書を印刷するための機能を有しており、又、ファックス装置、コピー装置及び物理的文書のための他の装置の能力を有することが可能である。印刷システムについての更なる情報は、“Networked Printing System Having Embedded Functionality for Printing Time-Based Media”と題され、2004年3月30日に、Hart等により出願された米国特許

10

20

30

40

50

出願公開第10/814,948号明細書において提供されており、この文献の援用によって発明の説明の一部を代替する。

#### 【0017】

図1においては又、本発明のこの実施形態における音声ペーパー生成システム( APPS )を示している。このシステムは音声ペーパー生成システムと呼ばれるが、それに代えて、他の実施形態における映像ペーパー生成システム又はいずれの他のタイプのマルチメディア生成システム

であってもよい。更に、 APPS108 は、そのタイトルにおいて“ペーパー”の言葉で呼んでいるが、 APPS108 は又、デジタルフォーマット及び他のタイプのフォーマットにおけるマルチメディア表現を生成するために用いられることができる。

10

#### 【0018】

APPS108 は、プリンタ102の一部であるとして図1に示されている。しかしながら、他の実施形態においては、 APPS108 は、例えば、プリンタ102に接続されることができるパーソナルコンピュータ( PC )(図示せず)に遠隔的に位置付けられる。 APPS108 は、特徴抽出能力及びフォーマット化能力を有する。音声ファイルは、入力技術及び特徴抽出技術がマルチメディアコンテンツの表現(即ち、波形の状態の音声コンテンツの表現)を生成するために適用されるとき、 APPS108 に入力される。表現または文書120は、特徴抽出の間にマルチメディアコンテンツにおいて認識される特定の特徴のためのマーカーを含むことができる。例えば、表現120は、音声タイムラインに沿って、拍手が起こる時間毎に又は音楽トラックにおけるサックス奏者のソロ演奏がある時間毎に対するマーカーを有することが可能である。適用される特徴抽出技術は、ユーザにより定義されることが可能である、又は、それに代えて、デフォルトプリンタ102設定により設定されることが可能である。 APPS108 のフォーマット化機能は、特徴抽出結果を用い、文書フォーマット仕様( DFS )に従ったフォーマット化を適用する。

20

#### 【0019】

一部の実施形態においては、ユーザは、 DFS104 に提供されるフィールドに情報を入力することにより生成される文書120に関連してフォーマット化好みを設定することができる。一部の実施形態においては、ユーザは、文書フォーマット及びレイアウト、フォントタイプ及びサイズ、各々の行に表示される情報、ヘッダに表示される情報、スケジュールカラムのサイズ及び位置、フォントの色、行間隔、行当たりの文字数、ボールディング及び拡大縮小技術、文書が印刷される言語、紙のサイズ、紙の種類等に関する好みを設定することができる。例えば、ユーザは、表示されるマルチメディアコンテンツの名前を表す、大きいボールドフォントのヘッダを含むマルチメディア文書を有するように選択することができる、ユーザは、ページ当たり表示されるマルチメディアコンテンツのグラフィク表現の配列を選択することができる。

30

#### 【0020】

DFS104 は、音声データに適用される特徴抽出と出力文書120を生成するために用いられるフォーマットガイドラインとを決定する。 DFS104 は、 PC (図示せず)におけるプリンタドライバダイアログインターフェースのような外部のアプリケーションにより供給されることができるデータ構造であり、プリンタのコンソール(図示せず)における APPS108 と対話することにより対話的に決定されることがある。 DFS104 はマルチメディアデータの表現を表す。 DFS104 は、ユーザに対して表示されるユーザインターフェースを格納するために用いられ、ユーザにフォーマット化オプションを与える。 DFS104 は、マルチメディアデータを適用することができる、ユーザに対して提供される特徴抽出オプションを決定する。 DFS104 は又、出力文書を生成するために用いられるフォーマットガイドラインを決定する。

40

#### 【0021】

DFS104 は、音声コンテンツ、音声コンテンツの構成者などのタイトルについての情報のような、音声ファイルについてのメタデータ情報を含むことができる。 DFS104

50

4は又、セグメントの開始時間及び終了時間（例えば、音声記録の開始時間及び終了時間）及びタイムラインに沿って表示されることができるマルチメディアデータノグラフィック表示（例えば、時間に対する音声信号の振幅を示す波形）のための仕様を含むことができる。D F S 1 0 4は、タイムラインに沿って表示されることが可能である各々のタイムスタンプ（即ち、バーコード、R F I D タグ、U R L又はマルチメディアデータを検索することができる位置についてのある他の表示）のためのメタデータ及びタイムスタンプマークーに対する仕様を更に含むことができる。

【0022】

図1に示す実施形態において、プリンタは、プリンタコンソール及び印刷ドライバインターフェースに対する処理ロジック106から更に構成される。処理ロジック106は、印刷ドライバダイアログインターフェース（図示せず）を介してユーザと対話する。例えば、処理ロジック106は、マルチメディアコンテンツが祭りメディア表現120の状態で表示されるフォーマット又はマルチメディアコンテンツの処理のような、特定のプリンタの動作をユーザが制御することを可能にするユーザインターフェースの表示を管理する。又、ユーザインターフェースの機能はウェブインターフェースにより提供されることができ、このウェブインターフェースによるフォーマット化イシューのようなプリンタの動作をユーザが管理することを可能にする。更に、処理ロジック106は、音声ペーパーに紙または電子フォーマットを戻すことができる。例えば、一部の実施形態においては、ユーザは、表現が印刷されるフォーマットを選択することができる。他の実施形態においては、プリンタ102は、表現のフォーマットに関してデフォルト設定素自動的に適用する。

10

20

30

【0023】

プリンタ102により生成されたマルチメディア文書120は種々のフォーマットから構成されることができる。例えば、マルチメディアフォーマット120は、図1に示す方式の音声ペーパー文書120のようなペーパー文書であることが可能である。プリンタ102により生成されるマルチメディア文書120は又、デジタルメディアに記憶されることができる。デジタルメディア書き込みハードウェアは、例えば、ネットワークインターフェースカード、D V D ライタ、セキュアデジタル（S D）ライタ、C D ライタ等を含むことができる。マルチメディアコンテンツは、例えば、フラッシュ媒体、D V D、C D等のようなデジタル媒体に記憶されることがある。

30

【0024】

マルチメディアコンテンツ120は、多くの異なるタイプのレイアウトを有することができ、種々のタイプの情報を表示することができる。図1は、音声コンテンツを表示する音声ペーパー文書120の例を提供しているが、他の実施形態においては、文書は、映像コンテンツを表示する映像ペーパー文書であることが可能である。映像ペーパー文書の生成についての更なる情報は映像ペーパーアプリケーションにおいて提供され、それらアプリケーションの各々については、そのアプリケーションの援用によって発明の説明の一部を代替する。

40

【0025】

図1の例において、音声ペーパー文書120は音声情報の音声波形112表示を示している。レイアウト及びフォーマット情報は、音声記録から抽出される音声コンテンツの長さ、メディアにおける音声波形112の配列及び他の情報を指定する。音声情報に対して、プリンタ102は、マルチメディア情報の特定なセグメントに対する音声の顕著な特徴を捕捉するセグメント（又は情報を与えるフレーム）を抽出することができる。更に、上記のように、プリンタ102は、特徴抽出能力（例えば、音声イベント検出等）を含むことが可能であり、例えば、特定の話者、音楽、笑い声又は叫び声等のような興味のあるアイテムについての音声セグメントにおいてユーザが検索することを可能にする。生成された文書120は、1つの音声波形112を表示することができる、又は2つ以上の音声波形に表示される音声コンテンツを分離することができる。図1における音声波形112は垂直方向に表示されているが、他の実施形態においては、音声波形112は他の配置に表示されることがある。

50

## 【0026】

更に、図1の音声波形112は、音声波形112に対して表示される音声コンテンツの開始及び終了にマーカーを付けるタイムスタンプマーカー114を有している。代替として、音声波形112は、全長に沿って多くのタイムスタンプマーカー114（即ち、マーカーのユーザ定義位置）を有することができる、又は、文書120はタイムスタンプマーカー114を全く有しないようにすることができる。

## 【0027】

図1の音声ペーパー120の実施形態において、文書120はヘッダ110を含むことができる。ヘッダ110は、文書120に含まれる音声コンテンツについての一般情報を提供する。例えば、ヘッダ110は、文書120に関して表示される音声コンテンツのタイプ（例えば、“会議”）、音声コンテンツの記録の日付（例えば、2003年11月21日）そして音声コンテンツが記録された場所（例えば、R I I会議室）についての情報を含むことが可能である。

10

## 【0028】

本発明の他の実施形態においては、ユーザ選択可能識別子116（例えば、バーコード又はテキストタグ）は音声波形112に対応している。図1の例においては、ユーザ選択可能識別子116は、ユーザ定義位置において音声波形112の右側に表示されているが、その代わりに、そのページのいずれかの位置に表示することができる。それらの識別子116は、インデックスマーカーとして機能し、ユーザが対応する音声コンテンツにアクセスすることを可能にしている。例えば、紙に印刷された文書120において、ユーザは、そのページのバーコード識別子116を物理的に走査することができ、このしき別氏は、音声波形112において表示されている音声コンテンツにおける音声セグメントを指示する。ユーザは、携帯電話又は携帯デジタル端末（PDA）のような、組み込まれたバーコードスキャナを有するいずれかのタイプの装置（図示せず）を用いて、ペーパー文書120における適当なバーコードを走査することによりユーザ選択可能識別子116を選択する。

20

## 【0029】

音声ファイルは、いつ文書が生成されたかを特定する技術（例えば、バーコード走査）に任意アクセスすることを可能にする装置においてプレイされることができる。例えば、バーコードを含む文書は、バーコードリーダと所定ポイントにおいて開始する音声ファイルをプレイする命令にバーコードを変換することができるソフトウェアとを有する携帯電話においてプレイされることができる。それ故、ユーザ選択可能識別子116は、ユーザがマルチメディア文書120において表示されるマルチメディアコンテンツにアクセスする又はそれを検索することを可能にするためのインターフェースとして機能する。

30

## 【0030】

一例として、図1のバーコードを走査することにより、ユーザは、音声セグメントが表示装置（例えば、テレビジョン、PCモニタ、携帯電話のスクリーン、PDA等）においてマーカーを付けられた位置からプレイを開始するようにすることができる。マルチメディア文書120は、例えば、記録された部分がプレイされている間に、PDAが連続音を立てるようにすることにより、触覚的フィードバックさえ提供することができる。他の例として、ペーパーマルチメディア文書120は又、バーコードマーカーに付加してまたはそれに代えて含む数字識別子を有することができ、ユーザは、システム100がプリンタ表示または他の装置における音声セグメントをプレイするように命令するプリンタまたは外部装置におけるキーパッド又はタッチパッド（図示せず）にそれらの数字をタイプすることができる。又、図1に示している音声ペーパー文書120がデジタルフォーマットである場合、システム100は、デジタル文書から直接プレイされる音声セグメントをユーザが選択することができるよう（即ち、プレイボタンを選択することによりマウス又は他の選択装置を用いて音声波形112における位置をチェックすることにより）、システム100を構成することが可能である。

40

## 【0031】

50

プリンタ 102 は、ユーザ選択可能識別子 116 に対応するマルチメディア情報を検索することができる。選択装置（即ち、数字識別子の状態で入力するためのバーコードスキャナ又はキーパッドを有する装置）からプリンタに通信される信号は、ユーザにより選択される音声セグメント、プレイされる音声コンテンツの場所、セグメントが選択されるようになっているマルチメディアペーパー文書、好みに及び／又はユーザにより選択された1つ又はそれ以上のマルチメディア表示装置（例えば、携帯電話）に関する情報、要求されたマルチメディア情報の検索を容易化するための他の情報等を識別することが可能である。例えば、システム 100 は、PC（図示せず）において記憶された音声ファイルにアクセスすることができ、システムはユーザの命令に関してこの音声コンテンツをプレイすることができる。

10

### 【0032】

図 1 の例は、文書 120 における音声波形 112 に沿ってマーカーを付けられた位置の隣のテキスト情報 118 を更に示している。この例において、テキスト情報 118 は、音声波形 112 に沿って、マーカーを付けられた位置に対応する会話のトランスクリプトの部分を含んでいる。このようにして、ユーザ選択可能識別子 116 を選択することにより、ユーザは、音声コンテンツがユーザ選択可能識別子 116 に対応するテキスト情報 118 の開始においてプレイを開始するようになることができる。種々の他のタイプのテキスト情報 118 は又、会話の要旨、話者の名前等のような、文書 120 における音声波形 112 タイムラインに沿って表示されることができる。

20

### 【0033】

システム 100 により生成されたマルチメディア文書 120 は、多くの異なる方法において使用されることができる。例えば、文書 120 は、ユーザに、興味をもたれている特定の音声コンテンツを検索することにより音声データを視覚的に調べる従来の方法を提供し、このような選択されたコンテンツに関するテキストとマーカーを提供し、ユーザが音声コンテンツにアクセスし且つそれをプレイするインターフェースを提供する。又、このタイプのマルチメディア文書 120 に関して、多くの変形が存在する。例えば、ユーザは、両面映像又は音声ペーパーを印刷することができる。この例においては、ユーザは、文書の両面にインクを供給することができるプリンタにおいてマルチメディア文書 120 を印刷する。オリジナルの音声又は映像ペーパーフォーマットを、文書の前面に印刷することができる。背面には、前面において表現されているデータについての2次元のバーコード表示を示すことができる。このフォーマットは、ファイリングキャビネットに格納することが可能であるスタンドアロンペーパーベース表現を提供し、それに続くマルチメディアコンテンツの検索は、オフペーパー表現への参照を必要としない。文書の前面において印刷された画像から近似のデジタルデータを抽出し、2次元バーコードにおいて符合化されるような高周波数成分でその近似のデジタルデータを補う、組み合わせの技術を用いることが可能である。

30

### 【0034】

他の例として、ユーザは、孔の開いた映像又は音声ペーパーのような有孔マルチメディア文書を作成することができる。例えば、ユーザは、紙の異なる有孔ストリップに各々印刷されたシーンにセグメント化された映像ファイルを印刷することができる。各々のストリップは、映像コンテンツからの少なくとも1つの映像フレームと、映像データのオンラインリポジトリを参照する少なくとも1つのバーコードとを含むことができる。例えば、ストリップはノートに貼り付け、掲示板に貼り出すことが可能である。有孔音声ペーパーの場合、ユーザは、話者、音声定位化、音声イベント検出等によりセグメント化された音声ファイルを印刷することができ、これらのセグメント化タイプの各々を紙の異なる有孔ストリップに印刷することができる。例えば、1つのストリップは、会議中に人が議論していたときの例を示すバーコードを含むことが可能である。各々のストリップは、音声データのオンラインリポジトリを参照する少なくとも1つのバーコードを含むことができる。しかしながら、マルチメディアデータは制限されることができるため、マルチメディアに対する完全スタンドアロン表現を提供するために2次元バーコードを用いることが可能

40

50

である。これらのストリップは切り取られ、記録のほんの小さな断片を思い出す必要がある人又は音声の記録を編集する必要のある人により容易に持ち歩くことが可能である。上記のように、ストリップは又、ノートに貼り付け、掲示板に貼り出すことが可能である。

#### 【0035】

他の例として、ユーザは、マルチメディア文書120を用いて、DVD又はCDカバーシートを作成することができる。この例においては、ユーザは、この印刷技術を用いて、DVD又はCDを印刷することができる。更に、プリンタ102は、シーンを参照するバーコード及び映像ファイルからセグメント化されたシーンからの映像フレームを示すカバーシートを自動的に生成するようにプログラムされることができる。このようなカバーシートを、例えば、プリンタ102における特定のトレイに挿入することが可能である小さい紙材に印刷することができる。又、カバーシートは標準的な紙材に印刷し、典型的なDVDホルダに適合するように、紙をどのように折り畳むべきかを示す折り畳みのマーカーを付けることができる。類似するカバーシートを、音声コンテンツのマーカーが付けられた部分を調べるバーコード及びユーザ選択コンテンツに対するマーカーを表す音声波形122タイムラインを表示する音楽CDに対して印刷することができる。マルチメディア情報の印刷可能表現を生成することについての更なる情報は、上で参照した、映像ペーパーアプリケーションにおいて提供されている。

#### 【0036】

ここで、図2を参照するに、本発明の実施形態のアーキテクチャを示している。この実施形態において、システム200は、システム200に入力される音声ファイルをしょりすることができるAPPS108を含む。APPS108はプリンタ102に位置付けられることができ、又はAPPS108はデータ処理システム(図示せず)に位置付けられることができ、それは、PC、ポータブルコンピュータ、ワークステーション、コンピュータ端末、ネットワークコンピュータ、メインフレーム、キオスク、標準的リモートコントロール、PDA、ゲーム制御器、携帯電話のような津真装置、アプリケーションサーバ又はその他のデータシステムを含むことが可能である。又、APPS108は、データ処理システムに結合されるプリンタ102に位置付けられることが可能である。

#### 【0037】

図2の例においては、APPS108は、構成要素としての特徴抽出モジュール202及びフォーマット化モジュール204から構成される。上記のように、システム200は、音声ファイルのようなマルチメディア情報にアクセスする、又はそれを受信する。そのファイルはシステム200において記憶されることができ、又はプリンタに結合されるデータ処理システム(図示せず)において記憶されることができる。図2の実施形態においては、ユーザは、ユーザがマルチメディア情報を再生する、記憶する、インデックス付ける、編集する又は操作することを可能にする種々の標準的マルチメディアプレイツールのいずれか1つを用いて、音声ファイルを聞くことができる。例として、専用の又はカスタマイズされたマルチメディアプレーヤ(例えば、RealNetworks製のRealPlayer(登録商標)、Microsoft社製のMicrosoft Windows(登録商標) Media Player、Apple社製のQuickTime(登録商標) Player、Shockwaveマルチメディアプレーヤ、その他)と、映像プレーヤ、テレビジョン、PDA等を含む

音声ファイルは、データポート206を通ってAPPS108に入ることができる。このポートは、イーサネット(登録商標)接続のような、いずれかのタイプのデータポートを含むことができ、そのデータポートに対してデータはプリンタ102に入力されることがある。更に、DFS104は接続に対してAPPS108に入力され、その接続208は、DFS104の記憶器の位置(図示せず)にAPPS108を結合している。特徴抽出モジュール202及びフォーマット化モジュール204の両方はDFS104情報を用いることができる。DFS104は、特徴抽出モジュール202によりマルチメディアコンテンツに適用される特徴抽出技術を規定し、DFS104は、フォーマット化モジュール204により用いられる文書フォーマット化情報を規定する。

## 【0038】

D F S 1 0 4 は種々の異なるタイプの情報を含む。D F S 1 0 4 は、表現が生成される音声ファイルについてのメタデータを含む。例えば、D F S 1 0 4 は、音声記録のタイトル、アーチスト、発行者等のような情報を含むことができる。D F S 1 0 4 は、記録についての開始時間及び終了時間を含むことができる。D F S 1 0 4 は又、タイムラインに沿って表示されることができる音声データのグラフィック表示のための仕様を含むことができる。例えば、そのグラフィック表示は、図1において説明したような音声波形とすることができます。その音声波形は時間に対する音声信号の振幅を示すことができ、ユーザは、必要なときに音声波形を拡大及び縮小を行うことができる。他の例は波形に対するJ P E Gである。D F S 1 0 4 は又、タイムラインに沿って表示することができる、各々のタイムスタンプ又はユーザ選択可能識別子（例えば、テキストタグまたはバーコード）に対するタイムスタンプマーカー及びメタデータのための仕様を含むことができる。

10

## 【0039】

レイアウトパラメータは又、D F S 1 0 4 において規定されることができ、そのD F S 1 0 4 において、パラメータは、生成された物理的文書1 2 0 のアピアランスを決定する。レイアウトパラメータは、例えば、文書1 2 0 の各々のページにおいて表示される、タイムラインの一部のための仕様を含むことができる。レイアウトの生成は、デフォルト挙動仕様により決定され、プリントデフォルト設定（例えば、プリント特性）において格納されることができる。これは、プリントのコンソールにおけるユーザインタフェース、ウェブページ等を用いて、対話処理又はペーパー文書1 2 0 の自立生成を含むことができる。

20

## 【0040】

特徴抽出モジュール2 0 2 は、グラフィック表現並びにD F E 1 0 4 において指定されたタイムスタンプ及びユーザ選択可能識別子1 1 6 を生成する。グラフィック表現の例として、時間に対する音声ファイルの振幅を示す曲線を含む。ユーザ選択可能識別子1 1 6 に用いられることが可能である他の特徴の例として、音楽の演奏のソロの検出、発話の認識、拍手の検出、音楽の検出等を含む。

## 【0041】

フォーマット化モジュール2 0 4 は、接続2 1 0 により特徴抽出モジュールに結合される。特徴抽出データは、文書1 2 0 のフォーマット化で用いるためにフォーマット化モジュール2 0 4 への接続2 1 0 に対して送信される。フォーマット化モジュール2 0 4 は、音声特徴及びD F S 1 0 4 を、P D F 文書のような、電子ファイルとして又は紙にレンダリングすることができる文書表現に変換する。D F S 1 0 4 は、文書フォーマット化パッケージ（例えば、M i c r o s o f t W o r d ）により典型的に生成される他の情報及び用いられるフォントについての詳細情報を含む。このレイアウト情報は、下で説明するように、D F S 1 0 4 の“レイアウト”フィールドに含まれる。

30

## 【0042】

図2のシステム2 0 0 は又、プリンタにおいて、データ信号を処理する処理器（図示せず）を含むことができる。処理器（図示せず）は、C I S C ( C o m p l e x I n s t r u c t i o n S e t C o m p u t e r ) アーキテクチャ、R I S C ( R e d u c e d I n s t r u c t i o n S e t C o m p u t e r ) アーキテクチャ又は命令集合の組み合わせを実行するアーキテクチャを含む種々のコンピューティングアーキテクチャから構成されることがある。システム2 0 0 は、1つの処理器又は複数の処理器を含むことができる。主メモリ（図示せず）は、ソフトウェア及びシステムの他の構成要素を含む処理器2 1 4 により実行されることが可能であるデータ及び/又は命令を記憶することができる。命令及び/又はデータは、ここで説明する技術のいずれか及び/又は全てを実行するためのコードから構成されることが可能である。主メモリ（図示せず）は、D R A M ( D y n a m i c R a n d o m A c c e s s M e m o r y ) デバイス、S R A M ( S t a t i c R a n d o m A c c e s s M e m o r y ) デバイス又は当該技術分野で既知のある他のメモリとすることが可能である。

40

50

## 【0043】

プリント 102 が印刷要求を受信するとき、一部の実施形態においては、その要求及び対応するマルチメディアデータは、処理器（図示せず）に転送される。処理器は入力を解釈し、適切なモジュールをアクティブにする。一部の実施形態においては、処理器は、マルチメディアコンテンツを転送するために特徴抽出モジュール 202 に結合され、それを制御する。更に、処理器は、一部の実施形態においては、文書 120 のフォーマット化を制御するためにフォーマット化モジュール 204 に結合される。APPS108 は、適切な文書ベースの表現を生成し、文書 120 生成のパラメータを修正するため及びその結果をプレビューするために印刷ドライバダイアログインターフェース（図示せず）によりユーザと対話することができる。マルチメディア変換のパラメータ及び結果は DFS104 において示される。処理器（図示せず）は又、従来型のプリンタ（図示せず）と通信し、それに印刷ジョブ情報を送信することにより文書 120 の生成を管理することができ、従来型のプリンタ（図示せず）はペーパー出力を生成する。上記のように、文書 120 は又、バーコードのようなユーザ選択可能識別子と、プリンタ 102 により記憶された又は指定されたオンラインデータベース（図示せず）において記憶されたマルチメディアデータへの他のリンクを含むことができる。10

## 【0044】

動作中、システム 200 は、マルチメディアコンテンツを印刷するための方法を提供し、図において与えられた特定の例においては、システム 200 は音声コンテンツを印刷するための方法を提供する。ここで、図 3 を参照するに、音声ペーパー生成システム 108 における処理段階を説明するフロー図を示している。下で説明するように、APPS108 は、サブルーチン処理を実行する制御プログラムに結合される。この実施形態においては、APPS108 の処理段階は、システムに音声ファイルを入力する段階 302 とシステムに DFS104 を入力する段階 302 とを含む。ユーザの命令に基づいて、APPS108 は、グラフィック表現が要求されたか否かを判定する。否定的な場合、APPS108 は、特徴抽出が要求されたかどうか、判定する方に移行する。肯定的な場合、APPS108 は、DFS104 において生成された情報と音声ファイル情報とを用いてグラフィック表現を生成するために、システム 200 の特徴抽出モジュール 202 を呼び出す 304。APPS108 は、DFS104 においてリストアップされた文書仕様フィールドの 1 つとして特徴抽出結果の表現又はシンボリックフォームを付加する 306 ことにより、DFS104 を更新する。20

## 【0045】

処理における次の段階として、APPS108 は、特徴抽出が要求されたかどうかを判定する。否定的な場合、APPS108 は、DFS104 の“レイアウト”フィールドにおいてリストアップされた DFS104 出力フォーマットにおいて指定された文書タイプを生成するためにフォーマット化モジュール 204 を呼び出す 312 方に移行する。肯定的な場合、APPS108 は、DFS104 において示された情報と音声ファイル情報とを用いて、DFS104 において要求されたマーカーを生成するために特徴抽出モジュール 202 を呼び出す 308。APPS108 は、次いで、DFS104 にマーカーデータを追加する 310。一旦、この段階が完了すると、APPS108 は、DFS104 の“レイアウト”フィールドにおいてリストアップされた DFS104 出力フォーマットにおいて指定された文書タイプを生成するためにフォーマット化モジュール 204 を呼び出す 312。30

## 【0046】

ここで、図 4 を参照するに、フォーマット化モジュール 204 の動作を説明するフローチャートを示している。下で説明するように、フォーマット化モジュール 204 は、サブルーチン処理を実行する制御プログラムに結合される。この実施形態においては、フォーマット化モジュール 204 の処理段階は、特徴抽出モジュール 202 により実施される特徴抽出の結果を入力する段階 402 を含む。DFS104 の“レイアウトページ”フィールドにおいて挙げられた各々のページに対して、フォーマット化モジュール 204 は、そ40

のページのフォーマット化が終了したかどうかを判定する。肯定的な場合、フォーマット化モジュール 204 は制御プログラムにメッセージを返信する。そのフォーマット化が終了していない場合、フォーマット化モジュール 204 は、“メタデータ配置”に関する DFS104 の“レイアウト”フィールドにおいて指定されたように、フォーマット化モジュール 204 においてメタデータをフォーマットする 404。フォーマット化モジュール 204 は、次いで、特徴抽出の結果に基づいて、DFS104 の“レイアウトタイプ”フィールドにおいて指定されたように、特徴抽出モジュール 202 により生成されたグラフィック表現をフォーマットする 406。フォーマット化モジュール 204 は、DFS104 の“マーカータイプ”フィールド、DFS104 の“マーカー頻度”フィールド及び DFS104 の“マーカー<sub>n</sub>”フィールドに従って、バーコードを生成する 408。マーカーは、次いで、バーコードを与えられるフォーマット化モジュール 204 において指定されたように、フォーマット化される 410。システムは、次いで、フォーマット化メタデータ、グラフィック表現及びマーカーが与えられたページをレンダリングする 412。一旦、この処理がページに対して終了すると、フォーマット化モジュール 204 は、次いで、“レイアウトページ”フィールドにおけるページ全てがフォーマット化されるまで、DFS104 の“レイアウトページ”フィールドにおける次のページ及び他のページ全てに対してこの処理を続ける。

#### 【0047】

ここで、図 5 を参照するに、マルチメディア文書 120 に対シルバーコードの生成について説明するフロー図を示している。下で説明するように、APPS108 は、サブルーチン処理を実行する制御プログラムに結合される。この実施形態においては、処理段階は、バーコードタイプ（例えば、2 を 5 にインターリーブする）、バーコードにおける識別子の桁の数、バーコードにおけるタイムスタンプの桁の数およびタイムスタンプ値を含む情報を入力する段階 502 を含む。システムは、次いで、フォーマット化モジュール 204 から識別子フィールドを読み取る 504、次いで、識別子を右寄せ 10 進数列に変換する 506。システムは、次いで、右寄せ識別子の長さがバーコードにおける可能な識別子の桁数より大きいかどうかを判定する。肯定的な場合、システムは制御プログラムにエラーコードに戻る。否定的な場合、システムは、タイムスタンプを右寄せ 10 進数列に変換する 508。システムは、次いで、右寄せのタイムスタンプの長さがバーコードにおいて可能なタイムスタンプの桁の数より大きいかどうかを判定する。肯定的な場合、システムはエラーコードを返信する。否定的な場合、システムは右寄せタイムスタンプを添付し 510、右寄せ識別子に対して、その右寄せタイムスタンプは左側を 0 で埋められる。システムは、次いで、指定されたタイプのバーコード画像をレンダリングし 512、識別子情報及びタイムスタンプ情報を含む。システムは、動作の終了の信号を送る制御プログラムに返信メッセージを送信する。

#### 【0048】

上記のフロー図は音声コンテンツに関連して説明したが、そのような方法は又、映像または他のメディアコンテンツに適用することができる。フロー図として示した図は、フォーマット仕様、特徴抽出及び音声ペーパー生成アルゴリズムのためのパラメータの異なる組み合わせを適用する結果の例を示している。上記のように、フォーマット仕様、特徴抽出及びパラメータは又、他のタイプのメディアコンテンツを表示する文書を生成するために用いられることができる。

#### 【0049】

図 6 は、一定間隔を置いたユーザ選択可能識別子 116 を有する音声ペーパーのグラフィック表示と、音声ペーパーを生成するための音声特徴抽出仕様 602 を有する対応する DFS104 を示している。図 6a においては、音声ペーパー文書 120 を生成するためにレイアウト及びコンテンツを指定するための DFS104 を示している。図 6a は、種々の DFS フィールドを含み、それらの DFS フィールドにおいて、レイアウト及びコンテンツに関する情報が指定されている。この例においては、文書 120（例えば、音楽のレコーディング）に含まれる音声コンテンツのタイプを挙げる“タイプ”フィールド 60

4 がある。“識別子”フィールド 608 は、バーコードまたはユーザ選択可能識別子 116 に含まれるようになっている識別化情報を挙げている。“タイトル”フィールド 610 は、音楽のレコーディング(例えば、Locomotion)のタイトルを挙げている。“アーチスト”フィールド 612 は、音声コンテンツを制作したアーチストの名前(例えば、John Coltrane)を挙げている。DFS104 は、レコーディングが含まれる音楽コレクション又はアルバム(例えば、Blue Train)を指定する“コレクション”フィールド 614 を含んでいる。DFS104 は又、誰が及びいつの日付でそのレコーディングを発行したか(例えば、Blue Note Records, 1957)を特定する、“発行者”フィールド 616 及び“発行日付”フィールド 618 を含んでいる。“開始時間”フィールド 620 及び“終了時間”フィールド 622 は、音声コンテンツが開始した時間(例えば、“00:00:00”)及び音声コンテンツが終了した時間(例えば、“00:07:14”)を挙げている。“グラフィック表現”フィールド 624 は、文書 120 に含まれる音声コンテンツのグラフィック表現のタイプ(例えば、振幅曲線)について説明している。

10

20

30

#### 【0050】

DFS104 は又、文書 120 のレイアウトしようとと共に、文書 120 に含まれるユーザ選択可能識別子 116 又はマーカーについての情報を含む。“マーカータイプ”フィールド 628 及び“マーカー頻度”フィールド 630 について示されているが、それらは、文書 120(例えば、バーコード)に含まれるマーカーのタイプ及びマーカーがグラフィック表現に沿って現れる頻度(例えば、30 秒間隔)を指定する。更に、レイアウトフィールドは文書 120 のレイアウトについての情報を与える。図 6a において、音声ペーパーのコンテンツの配置を指定する“レイアウトタイプ”フィールド 632 を示している。例えば、レイアウトタイプは、文書 120 において表示される 1 つの水平方向のタイムラインを含むことができ、又は、それに代えて、2 つの水平方向のタイムラインを含むことが可能である。“レイアウトページ”フィールド 634 は文書 120 のページ数を指定する。“レイアウトマーカー配置”フィールド 636 は、ユーザ選択可能識別子 116 またはマーカーが表示される位置(例えば、グラフィック表現の上)を指定する。更に、“レイアウトメタデータ配置”フィールド 638 は、文書 120 におけるメタデータの配置についての情報を挙げている。メタデータは、ヘッダ又は他のメタデータを含むことができる。

30

#### 【0051】

図 6a の DFS104 は、生成されたメディア表現についての情報の集合の一例を示している。他の実施形態においては、DFS104 は、例えば、ピクチャ情報、ハーパーテキスト、アーチストの経歴、アーチストの誕生日/死去日、アーチストの住所情報、表示されたメディアコンテンツをどこで購入するべきか(即ち、アルバムを購入するためのウェブサイトへのリンク)等のようなフィールドであって、それらに限定されない他の多くのフィールドを含むことができる。DFS104 の一部の他の例については又、下で説明する。これは変形の網羅的なリストではなく、多くの他のタイプの情報を組み込むことが可能である。

40

#### 【0052】

又、図 6a には音声特徴抽出仕様 602 を示している。“特徴”フィールド 606 は、音声コンテンツに適用される特徴抽出を規定する。この例において、音声特徴抽出 602 は音声振幅抽出及びグラフィック近似である。このようにして、文書 120 は音声波形 112 を示す。この例において、SVG ファイルは出力される。

#### 【0053】

図 6b において、本発明の位置実施形態に従った音声ペーパー文書 120 のグラフィック表現を示している。この文書 120 において、DFS104 における使用に従ったヘッダ情報を有するヘッダ 110 を示している。ヘッダ 110 は又、DFS104(において指定されているように位置付けられている、即ち、この場合、ページの上部に中央揃えされている)。文書 120 はタイムラインに沿って振幅曲線または音声波形 112 を表示し

50

ている。本発明の他の実施形態においては、タイムラインは1つの直線又は他の種々のグラフィック表現により表表されることができる。タイムラインは、音声レコーディングの長さに対応して、“00:00:00”から“00:07:14”まで実行される。タイムスタンプ114は音声波形112に沿って3箇所に示されており、レコーディングの開始時間と終了時間とにマーカーを付け、タイムスタンプ114と共に、セクションの中央の位置にマーカーを付けている。文書120は、ユーザの好みに従って、4つ以上のタイムスタンプ114を示すことができ、又はタイムスタンプ114を全く示さないことができる。

#### 【0054】

更に、文書120はユーザ選択可能識別子116（例えば、バーコード）を表示し、そのユーザ選択可能識別子116は、タイムラインに沿った位置においてユーザが音声コンテンツにアクセスすることができるインターフェースを提供する。一部の実施形態においては、ユーザは、各々のユーザ選択可能識別子116に対して特定の位置を指定することができる。この例において、ユーザは、文書120がタイムラインに沿って30秒毎にバーコードマーカーを有するように指定した。これらのユーザ選択可能識別子116は、各々の個々の識別子116のより容易な選択を可能にするために、1つの長い線ではなく、“階段”方式で表示されている。しかしながら、ユーザ選択可能識別子116の配列は大きく変化させることができ、DFS104において指定されることができる。上記のように、ユーザは対応する音声コンテンツをプレイするために印刷文書においていずれのユーザ選択可能識別子116を選択することができる。例えば、ユーザは、バーコードを走査することが可能であり、携帯電話又は他の表示装置においてタイムラインに沿ったいずれかの位置から開始するレコーディングをプレイするためにその位置において、バーコードスキャナを有する携帯電話を用いて、バーコードを走査することが可能である。

#### 【0055】

ここで、図7を参照するに、レコーディングにおける各々の音楽のソロのためのユーザ選択可能識別子116を有する文書120のグラフィック表示と、音声ペーパー120を生成するための音声特徴抽出仕様602を有する対応してDFS104とを示している。図7aに示すDFS104は図6aに示すそれに類似しているが、図7aの例は幾つかの変更を含んでいる。図7aのDFS104の例は、音声コンテンツに適用される特徴抽出を挙げる“特徴抽出”フィールド702を含んでいる。この場合、特徴抽出は、音声コンテンツにおける音楽のソロにマーカーを付けることを含んでおり、その特徴抽出において、出力は楽器名及びソロの開始する時間を示す。図7aの例においては、“マーカータイプ1”フィールド704及び

“マーカータイプ2”フィールド706を示しており、これらのフィールドは文書120において表示されるユーザ選択可能識別子116の2つのタイプを指定する。例えば、文書120は、タイムスタンプ114の上に示されるバーコードの上に示される楽器名を表示するマーカータイプ1を含む。この例においては、マーカータイプ2はバーコード（即ち、グラフィック表現の下の規定された位置に表示される第2バーコード）である。DFS104は又、“レイアウトマーカー1配置”フィールド710及び“レイアウトマーカー2配置”フィールド712を含む。これらのフィールドは、タイムラインの下又はタイムラインの上のように、各々のマーカーが文書120において示されるところを指定する。

#### 【0056】

図7aは又、“特徴”フィールド606における音声特徴抽出しよう602を示している。この例における音声特徴抽出しよう602は、SVGファイル出力と共に、音声振幅抽出とグラフィック近似とを含む。又、音声特徴抽出仕様602は音楽のソロの抽出を含み、各々のソロの開始時間及び各々のソロにおいて用いられる楽器を出力する。この例においては、特徴抽出の1つのタイプが音声コンテンツに適用されている。しかしながら、システムは、ある時間においていずれかの数の特徴抽出タイプを適用することができる。他の特徴抽出の例は、発話検出、話者検出、話者認識、映像／音声イベント検出、映像前

10

20

30

40

50

景 / 後景セグメント化、顔検出、顔画像適合化、顔認識、顔カタログ化、映像テキスト定位化、光学式文字認識（OCR）、言語翻訳、フレーム分類、クリップ分類、画像ステッチ、音声フォーマット変換、音声波形適合化、音声キャプション位置合わせ、映像OCR及びキャプション位置合わせ、音声定位化、ラジオ送信認識、スライダによる音声 / 映像範囲選択、話者セグメント化、プロファイル分析、カラーヒストグラム分析、クラスタリング、動き分析、距離推定、シーンセグメント化、ナンバープレート又は自動車認識並びに動き分析等を含むが、これらに限定されるものではない。以上は、多様性の網羅的なリストではなく、多くの他のタイプの抽出を本発明に組み込むことが可能である。

#### 【0057】

図7bは、図7aの音声特徴抽出仕様602及びDFS104に従った文書120のグラフィック表現を示している。文書120は、ヘッダ110と、水平方向に表示された音声波形112と、音声波形112の下部に沿った位置におけるタイムスタンプ114とを含む。この例においては又、各々のタイムスタンプ114の近くに含まれるユーザ選択可能識別子116がある。ユーザは、マーカーに位置における音声コンテンツのプレイを開始するためにこれらオンユーザ選択可能識別子を選択する（例えば、バーコードを走査する）ことができる。例えば、ユーザが時間“00:00:00”の上に示されるバーコードを走査する場合、レコーディングは初めからプレイを開始する。ソロの抽出は文書120において表示される音声コンテンツに適用されたため、音声波形112は各々の音楽のソロのためのマーカーを含む。それらのマーカーは、マーカーを付けられたソロのタイプ（例えば、サックスのソロ）と、ソロにインタフェースを提供するユーザ選択可能識別子116と、音声コンテンツにおけるソロの位置を示すタイムスタンプ114とを説明するテキスト表示714を含む。例えば、印刷文書における“サックス”ソロの下のバーコードを走査することにより、サックスのソロは表示装置において初めからプレイを開始する。

#### 【0058】

ここで、図8を参照するに、ラジオのプログラムのためのタイムラインを示す文書120のグラフィック表現と、音声ペーパー120を生成するための音声特徴抽出試用02を有する対応するDFS104とを示している。図8aに示されるDFS104は図6aに示すそれに類似しているが、図8aの例は幾つかの変更を含んでいる。図8aの例は、ラジオプログラムに対するDFS104を示しており、DFS104は、ラジオ番組に関する文書120に付加注釈を加える“注釈”フィールド802を含む。この例において、注釈は、プログラムにおけるゲストは“Bill O'Reilly”であることを示している。このようにして、ラジオトークショーが存在する時間と、ホストの名前及びコマーシャルブレークが存在するインターバルと、会話の各々の部分のためのバーコードと共に紙に印刷することができるその持続時間とのようなラジオトークショーに関する利用可能なメタ情報の要旨を示している。共演者が既知の場合、共演者の名前を含めることができる。バーコードは、システムのユーザにより別々に記録された音声データを指すことが可能であり、又は、トークショーにより提供されるウェブサイトにおける音声データを指すことが可能である。これは、文書の記録、生成、衛星放送のメディアファイル及びウェブページを後処理するソフトウェアに結合されることが可能である。トークショーが行われる間にオンラインで実行される注釈に生成処理をアクティブにリンクさせることにより更なるユーティリティが提供される。この注釈は、プログラムに電話を掛ける人の電話数のような、リスナーに利用可能でない情報へのアクセスをプロデューサが有するため、そのショーが行われている間に、このような注釈はラジオ局において実行されることがある。

#### 【0059】

図8bは、音声特徴抽出仕様602を有するDFS104に従って表示された文書120を示している。文書120は、タイトルと日付とラジオプログラムについての注釈情報とを示すヘッダ110、音声波形112、タイムスタンプ114及び“3段階”方式で表示されたユーザ選択可能識別子116を含んでいる。

10

20

30

40

50

## 【0060】

ここで、図9を参照するに、キーワードについてのマークを有するラジオプログラムに対するタイムラインを示す文書120のグラフィック表現と、音声ペーパー120を生成するための音声特徴抽出仕様602を有する対応するDFS104とを示している。図9aに示すDFS104は図8aに示すそれに類似しているが、図9aの例はマーク情報に関する変更を含んでいる。図9aの“マークータイプ”フィールド628は、キーワード、バーコード及びタイムスタンプを含むマークータイプを示している。“マークーフィールド”フィールド630は、その頻度が“ユーザ定義”されていることを示している。このようにして、この例においては、ユーザはタイムラインに沿って表示される各々のマークーを選択した。“マークー”フィールド902において、ユーザはマークー1乃至11に対する選択を行った。例えば、マークー1に対して、ユーザはバーコード、タイムスタンプ及びマークーテキストを示すテキスト（例えば、“WTC”）を含むようにマークーであって、マークーを付けられた音声コンテンツを示す、マークーを定義した。ユーザは又、タイムラインに沿って、“vert.pos.1”、“vert.pos.2”又は“vert.pos.3”のような各々のマークーの垂直方向の位置を定義した。これらの位置付けの仕様は、タイムラインの上の多くの階段状の位置から選択されたマークーが垂直方向のどこに位置付けられるかを決定する。音声特徴抽出602は又、グラフィック近似を有する音声振幅抽出である。

## 【0061】

図9bは、図9aの音声特徴週出仕様602とDFS104に従って表示された文書120を示している。文書120は、ラジオプログラムのためのヘッダ110と、音声波形112と、音声波形の下のタイムスタンプ114と、タイムラインに沿ってユーザ定義位置に表示されたマークーとを含む。マークーは、マークーを付けられたラジオプログラムにおけるコンテンツを表すためにユーザにより設定されたテキスト情報714を含む。更に、マークーは、ユーザ選択可能識別子116とタイムスタンプ114とを含む。

## 【0062】

ここで、図10を参照するに、検索語についての音声特徴抽出を有するラジオプログラムのためのタイムラインを示す文書120のグラフィック表現を示している。図10aのDFS104は、発話認識技術及びキーワード適合技術が音声コンテンツに適用された“特徴抽出”フィールド702において示されている。この例において、ユーザは、検索語として、“New York Times”又は“fair and balanced（公正で中立）”を検索した。図10aは、適合化検索語、バーコード及びタイムスタンプを含む“マークータイプ”フィールド628を示している。“マークーフィールド”630は、頻度が“ユーザ定義”であることを示している。このようにして、この例においては、ユーザはタイムラインに沿って表示される各々のマークーを選択した。“マークー”フィールド902において、ユーザはマークー1乃至11に対する選択を行った。例えば、マークー1に対して、ユーザはバーコード、タイムスタンプ及びマークーを表すテキスト（例えば、“fair and balanced”）を含むマークー並びに各々のマークーの垂直方向の位置を定義した。

## 【0063】

音声特徴抽出仕様602は又、グラフィック近時を有する音声振幅抽出を含む。音声特徴抽出仕様602は又、所定のキーワードのリストに適合する言葉と共に、発話認識を含む。このようにして、ユーザは、話者が特定の言葉を用いるラジオプログラムにおける位置と、発話の一部または全ての写しと共に、これらの位置がタイムラインに沿ってマークーを付けられた位置とを検索した。代替として、ユーザは、発話認識のみを適用することが可能であり、発話がなされた音声コンテンツにおけるいずれかのポイントを認識することが可能である。発話認識出力は騒々しい可能性があるため、認識の信頼度についての一部の表現が又、含まれることができ、それ故、ユーザは、どの単語又は文章がより正確である可能性があるかを理解することができる。例えば、文書120は、認識の信頼度を表すために色又はフォントサイズにおける変化を有することが可能である。大きい信頼度の

10

20

30

40

50

判定は赤色で 12 ポイントのフォントで表される一方、最も低い信頼度の判定は青色で 8 ポイントのフォントで表されるようになることが可能である。ユーザ選択可能識別子 116 は、最も大きい信頼度を有する判定のみ又は各々の判定に対して含まれることができる。

#### 【 0 0 6 4 】

音声コンテンツに適用されることが可能である音声特徴抽出の他の例は、話者検出と話者認識とを含む。話者検出抽出は、記録における同一の話者達のグループを認識し、同じ人が話していたときを判定することができる。このことは、限定された色のパレットを用いて注釈を付けられたセグメントによりタイムラインに沿って表されることが可能で、各々の話者に対して異なる色で、及び同じ話者に対して同じ色で示すことができる。話者認識抽出は、音声記録の間に話した実在の人物を指定する。人物のシンボリック同一性は、走査されるときにセグメントの初めから音声をプレイするバーコードと共に、タイムラインのセグメントの隣に加えられ、計算されることがある。このことは、プリントアウトを走査すること、及び誰が会議に出席したかを理解することを可能にする。代替のバージョンは名前のリストを印刷することを可能にし、それらの名前の後にバーコードを付けることを可能にする。ユーザはこれらのバーコードを走査し、その人物が話をしたときの記録のいつ部を聞くことができる。他の例は、それらの人物達についての顔画像を検索し、それらの人物の名前及びバーコードの隣にそれら顔画像を印刷する。音声データは又、2 次元バーコードにおいて具現化することができる、それ故、音声ファイルについての完全なスタンドアロン表現を提供することができる。

10

20

30

#### 【 0 0 6 5 】

図 10 b は、図 10 a の音声特徴抽出仕様 602 を有する DFS104 に従って表示された文書 120 を示している。文書 120 は、“キーワード検索語”の注釈を又示すラジオプログラムのためのヘッダ 110 と、音声波形 112 と、音声波形 112 の下のタイムスタンプ 114 と、タイムラインに沿ってユーザ定義位置に表示されたマーカーとを含む。マーカーは、マーカーを付けられたラジオプログラムにおけるコンテンツを表すためにユーザにより設定されたテキスト情報 714 を含む。この場合、滝スト情報 714 は、音声コンテンツにおいて見つけられた特定の言葉である。更に、マーカーは、ユーザ選択可能識別子 116 とタイムスタンプ 114 とを含む。ユーザ選択可能識別子を選択することにより、ユーザは、検索語が用いられた音声コンテンツを聞くことができる。

#### 【 0 0 6 6 】

図 11 において、拍手イベントについての音声特徴抽出を有するラジオプログラムのためのタイムラインを示す文書 120 のグラフィック表現を示している。図 11 a の DFS104 は、拍手検出が音声コンテンツに適用された“特徴抽出”フィールド 702 において示されている。音声特徴抽出仕様 602 は拍手検出タイムスタンプを含む。このようにして、ユーザは、拍手イベントが実施されたラジオプログラムにおける位置を検索し、そのような位置はタイムラインに沿ってマーカーが付けられる。

#### 【 0 0 6 7 】

図 11 b は、図 11 a の音声特徴抽出仕様 602 と DFS104 に従って表示された文書 120 を示している。文書 120 は、“示された注釈イベント”的注釈を又示すラジオプログラムのためのヘッダ 110 と、音声波形 112 と、音声波形 112 の下のタイムスタンプ 114 と、タイム欄に沿ってユーザ定義位置に表示されたマーカーとを含む。代替として、システムは、例えば、笑い、大声での会話、ドアを閉める音等のような音声コンテンツにおいて他のイベントが存在するときにマーカーを付けることができる。これは、例えば、議論が会議において起こったポイントで、会議の記録を即座に走査するため用いられることがある。マーカーは、ユーザ選択可能識別子 116 及びタイムスタンプ 114 を含む。ユーザ選択可能識別子を選択することにより、ユーザは、拍手が起こった音声コンテンツを聞くことができる。ユーザ選択可能識別子は、ペーパー文書において印刷された 2 次元識別子（例えば、バーコード）に盛り込まれた音声を又はオンラインで記憶された音声データを参照することができる。

40

50

## 【0068】

音声定位化技術は又、音声恩典津に適用されることが可能である。この例において、タイムライン表現は、記録がなされた部屋における場所を指す方向インジケータを含むことができる。これは、ユーザが、例えば、部屋の最南の角の人物、又はテーブルを横断してその人物に対向する人物が話しているときを判定し、タイムラインを即座に走査することを可能にする。これは、例えば、音声定位化を実行するために構成することができる複数のマイクロフォンの設置を有する一定の設定を適用することができる。又、インタビューを記録するプロにより用いられる記録器のような適切に設備される可搬型記録器と共にそれは用いられることができる。

## 【0069】

図12において、音楽イベントに対する音声特徴抽出を用いてラジオプログラムのタイムラインを示す文書120のグラフィック表現を示している。図12aのDFS104は、音楽検出が音声コンテンツに適用された“特徴抽出”フィールド702において示されている。音声特徴抽出仕様602は音楽検出タイムスタンプを含む。それ故、ユーザは、音楽イベントが存在したラジオプログラムにおける位置であって、これらの位置がタイムラインに沿ってマーカー付けされている、位置を検索した。更に、“レイアウトタイプ”フィールド632は、半分に分割された2つの垂直方向のタイムラインを含むレイアウトタイプを示している。

## 【0070】

図12bは、図12aの音声特徴抽出試用02とDFS104に従って表示された文書120を示している。文書120は、“示された音楽イベント”的注釈を又示すラジオプログラムのためのヘッダ110、音声波形112、音声波形112の近くのタイムスタンプ114及びタイムラインに沿ってユーザ定義位置に表示されたマーカーを含む。マーカーは、ユーザ選択可能識別子116とタイムスタンプ114とを含む。ユーザ選択可能識別子116を選択することにより、ユーザは、拍手が起こった音声コンテンツを聞くことができる。タイムラインは、2つの半分ずつに分割されたタイムラインであって、各々の半分は垂直方向に表示されているタイムラインと共にDFS104に示されているレイアウトタイプに従って表示されている。

## 【0071】

マルチメディアペーパーは又、音声メールメッセージの表現を生成するために用いられることができる。ユーザは、例えば、発呼者の電話番号、インターネット検索エンジン(発呼者の名前、住所及び位置を示す地図をしばしば示すことができる)における電話番号を調べた結果及びメッセージの日時と持続時間等のような、音声メールメッセージの収集についての利用可能なメタデータの要約を生成することができる。メタ情報の各々のブロックは、ネットワークの遠隔地からの音声情報を検索するバーコードの隣に印刷されることが可能であり、又は、ペーパーから直接プレイされることが可能である2次元バーコードにおいて表示されることが可能であり、それ故、いずれかのオフデバイスアクセスに対する要求を未然に防ぐことが可能である。ペーパー文書は、ペーパー文書を検索し且つそれを加えることができる付加情報(例えば、インターネット検索エンジン情報)を提供することによりユーザに有用性を提供する。又、ペーパー文書自体は、その文書に関する音声メールメッセージについての注記を書き込む能力をユーザに提供する。

## 【0072】

マルチメディアペーパーは又、日付、時間、持続時間、自動車番号、警官の名前(必要に応じて)を含む1つ又はそれ以上の公安(例えば、警察署、消防署等)無線のレコーディングについての利用可能なメタデータの要旨は、それらのレコーディングについてのオフライン表現を参照するバーコードと共にペーパーに印刷されることができる。又、音声データを直接符合化する2次元バーコードを用いることができる。このことは、ネットワーク接続を独立して使用されることができるスタンドアロン表現を提供する。メタ情報は記録された音声に適用される信号処理アルゴリズムにより計算されることができ、又は、無線で提供されるデジタルサイドチャネル情報(例えば、Motorolaデジタル無線

10

20

30

40

50

情報)から計算されることが可能である。又、メタ情報は、無線発信器コンソールにおいてデジタル的に供給されることが可能である。このシステムは、無線ログを選択的に調べる必要がある管理者を支援することが可能であり、又は、公安の手法を観測することを希望する公衆を支援することが可能である。

#### 【0073】

マルチメディアペーパーは又、航空無線の表現の生成において使用されることができる。1つ又はそれ以上の航空無線の記録についての利用可能メタデータの要旨であって、特定の伝送がなされたときの飛行の日付、時間、持続時間、便名、出発地、目的地、現在位置を含む、要旨は、音声記録のオンラインフォームを指すバーコードと共にペーパーに印刷されることが可能である。メタ情報は、適切な装置が利用可能である場合、トランスポンダが返すモードSから、直接、抽出されることがある。付加メタ情報は、リアルタイムに飛行の進行を追跡する種々のオンラインサービスから検索されることが可能である。音声記録に適用される発話認識は、オンラインデータへのリンクを計算するために用いられることが可能であるシンボリック情報を提供することが可能である。このことは、モードSへの直接的リンクに対する要求を未然に防ぎ、FAA装置にアクセスすることなく、このシステムを人が利用可能であるようにする。

#### 【0074】

本発明について特定の好適な実施形態を参照して説明したが、当業者は、種々の改良が可能であることを認識するであろう。好適な実施形態に対する変形及び改良は、同時提出の請求の範囲によってのみ規定される本発明により提供される。

10

20

20

30

40

40

50

#### 【図面の簡単な説明】

#### 【0075】

【図1】マルチメディアデータの表現を生成するためのシステムのグローバル構造図である。

【図2】図1のシステムの一実施形態の例示としてのアーキテクチャについてのブロック図である。

【図3】音声ペーパー生成システムの動作方法のフロー図である。

【図4】マルチメディアコンテンツのフォーマット化におけるフォーマット化モジュールの動作方法のフロー図である。

【図5】マルチメディア表現のためのバーコードを生成する方法のフロー図である。

【図6a】例示としての文書フォーマット仕様と音声特徴抽出の表現である。

【図6b】図6aにおいて表した仕様に基づいて生成されたマルチメディア表現のグラフィック表現である。

【図7a】例示としての音楽のソロの抽出を含む音声特徴抽出と文書フォーマット仕様の表現である。

【図7b】図7aにおいて表した仕様に基づいて生成されたマルチメディア表現のグラフィック表現である。

【図8a】例示としてのラジオプログラムのための音声特徴抽出と文書フォーマット仕様の表現である。

【図8b】図8aにおいて表した仕様に基づいて生成されたマルチメディア表現のグラフィック表現である。

【図9a】例示としてのキーワードを含む音声特徴抽出と文書フォーマット仕様の表現である。

【図9b】図9aにおいて表した仕様に基づいて生成されたマルチメディア表現のグラフィック表現である。

【図10a】例示としての話者認識及び言葉検索のための音声特徴抽出と文書フォーマット仕様の表現である。

【図10b】図10aにおいて表した仕様に基づいて生成されたマルチメディア表現のグラフィック表現である。

【図11a】例示としての拍手検出のための音声特徴抽出と文書フォーマット仕様の表現である。

【図11b】図11aにおいて表した仕様に基づいて生成されたマルチメディア表現のグラフィック表現である。

【図12a】例示としての音楽検出のための音声特徴抽出と文書フォーマット仕様の表現である。

【図12b】図12aにおいて表した仕様に基づいて生成されたマルチメディア表現のグラフィック表現である。

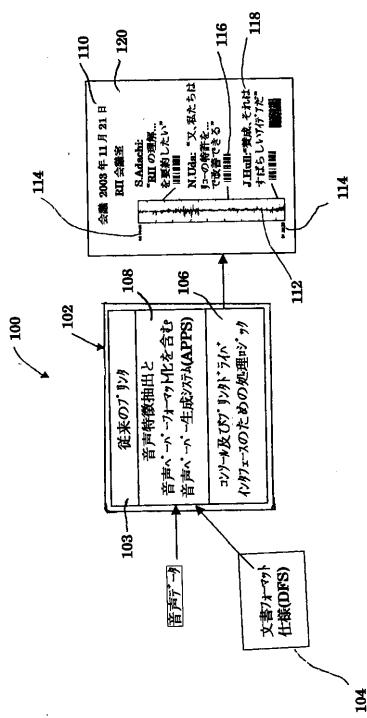
【符号の説明】

【0076】

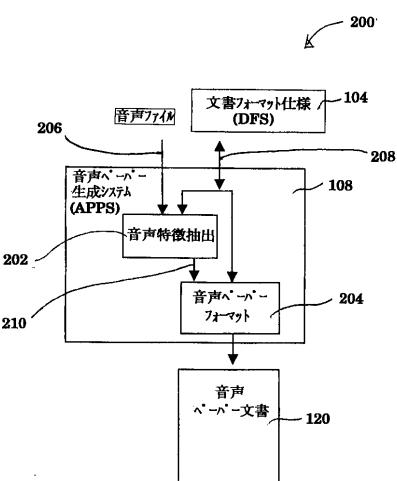
100	システム	10
102	プリンタ	
103	従来のプリンタ	
104	文書フォーマット仕様 (DFS)	
106	処理ロジック	
108	音声ペーパー生成システム (APPS)	
110	ヘッダ	
112	音声波形	
114	タイムスタンプマーク	
116	バーコード識別子	
118	テキスト情報	
120	マルチメディア文書	20
200	システム	
202	特徴抽出モジュール	
204	フォーマット化モジュール	
206	データポート	
208	接続	
210	接続	
602	音声特徴抽出仕様	
604	“タイプ”フィールド	
606	“特徴”フィールド	
608	“識別子”フィールド	30
610	“タイトル”フィールド	
612	“アーチスト”フィールド	
614	“コレクション”フィールド	
616	“発行者”フィールド	
618	“発行日付”フィールド	
620	“開始時間”フィールド	
622	“終了時間”フィールド	
624	“グラフィック表現”フィールド	
628	“マーカータイプ”フィールド	
630	“マーカー頻度”フィールド	40
632	“レイアウトタイプ”フィールド	
634	“レイアウトページ”フィールド	
636	“レイアウトマーカー配置”フィールド	
638	“レイアウトメタデータ配置”フィールド	
702	“特徴抽出”フィールド	
704	“マーカータイプ1”フィールド	
706	“マーカータイプ2”フィールド	
708	“マーカー2頻度”フィールド	
710	“レイアウトマーカー1配置”フィールド	
712	“レイアウトマーカー2配置”フィールド	50

7 1 4 テキスト表示  
 8 0 2 “注釈”フィールド  
 9 0 2 “マーカー”フィールド

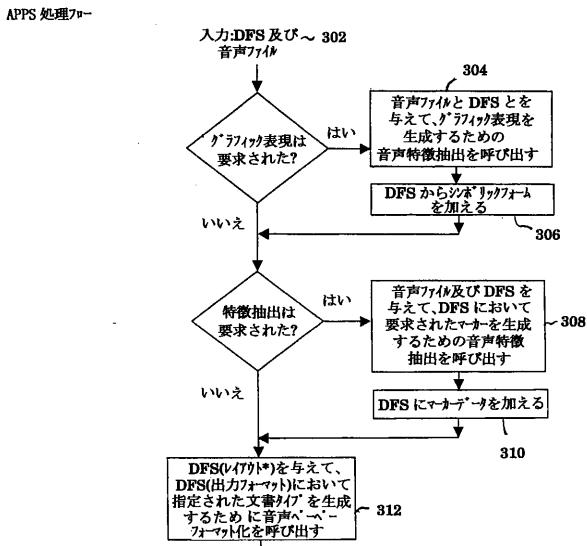
【図1】



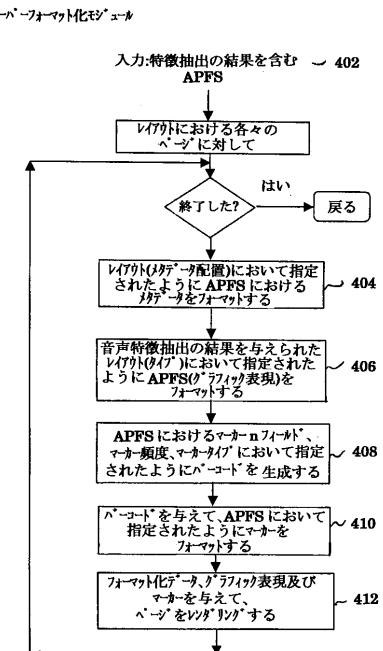
【図2】



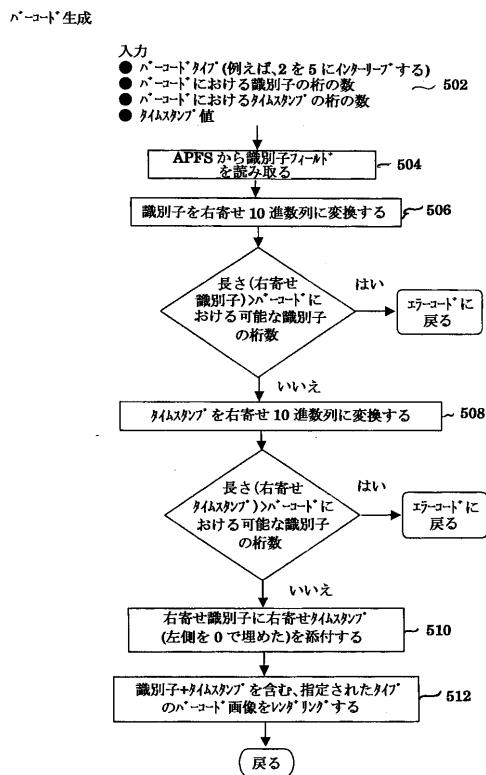
【 図 3 】



【 図 4 】



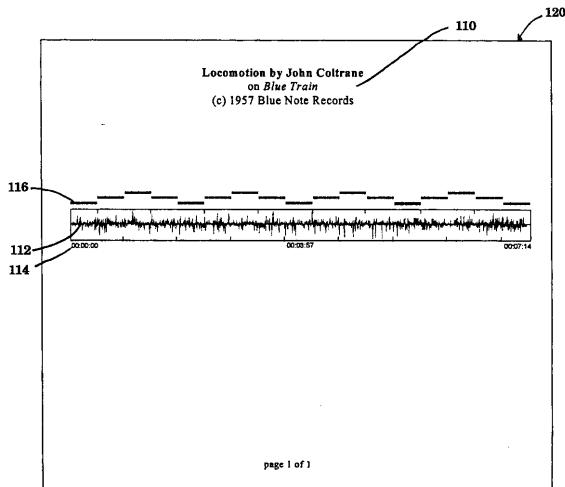
【 図 5 】



【 図 6 a 】

104	文書ワード仕様	音声特徴抽出	602
604	Type = Musical recording	音声振幅抽出及び「ワック」近似。	602
	Identifier = 1768		—606
608	Title = Locomotion		
610	Artist = John Coltrane		
612	Collection = Blue Train		
	Publisher = Blue Note Records		
614	Label = Blue Note 557		
	Begin time = 00:00:00		
614	End time = 00:07:14	624	
616	Graphic representation = Amplitude curve	628	
618	Marker type = bar code	628	
620	Marker placement = 30 sec. intervals	630	
622	Layout type = One horizontal time line	632	
	Layout pages = 1		
634	Layout marker placement = Above graphical representation	636	
	Layout meta data placement = Centered at top of page	638	

〔 図 6 b 〕

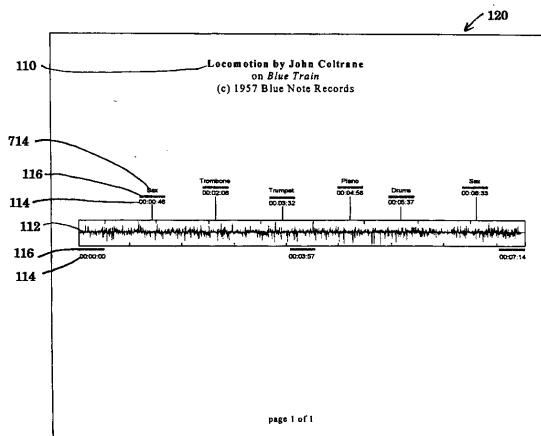


【図7a】

104 → 文書フォーマット仕様  
 Type = Musical recording  
 Identifier = 1768  
 Title = Locomotion  
 Artist = John Coltrane  
 Collection = Blue Train  
 Label = Blue Note Records  
 Publication Date = 1957  
 Begin time = 00:00:00  
 End time = 00:00:45  
 Graphical Representation = Amplitude curve  
 Feature extraction = Musical solo (output is instrument name and time when solo began)  
 Marker type 1 = instrument name above barcode above time stamp  
 Marker type 2 = bar code  
 Marker 1 frequency = 50%, 100%  
 Layout type = One horizontal time line  
 Layout pages = 1  
 Layout marker 1 placement = Above graphical representation  
 Layout marker 2 placement = Below time line  
 Layout meta data placement = Centered at top of page

602  
 音声特徴抽出  
 音声振幅抽出及びsvgファイル近似。  
 606  
 svgファイルを抽出する。又、適用する。  
 開始時間と各々の音楽の楽器を抽出する。

【図7b】

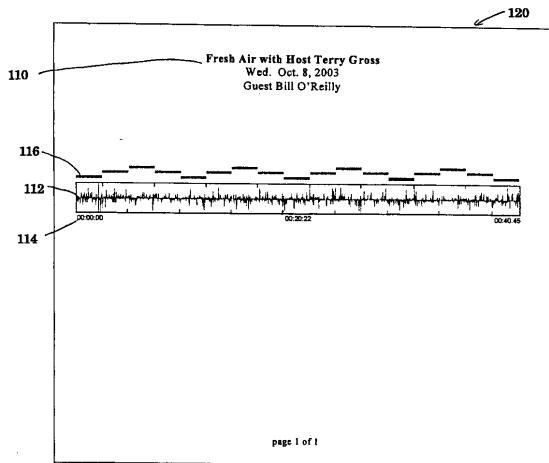


【図8a】

104 → 文書フォーマット仕様  
 Type = Radio program  
 Identifier = 1770  
 Title = Fresh Air with host Terry Gross  
 Annotation = Guest Bill O'Reilly  
 Publication Date = Oct. 8, 2003  
 Begin time = 00:00:00  
 End time = 00:40:45  
 Graphical Representation = Amplitude curve  
 Marker type = bar code  
 Marker frequency = regularly spaced  
 Layout type = One horizontal time line  
 Layout marker placement = Above graphical representation, 3-step staircase  
 Layout meta data placement = Centered at top of page

602  
 音声特徴抽出  
 音声振幅抽出及びsvgファイル近似。  
 606  
 svgファイルを抽出する。

【図8b】

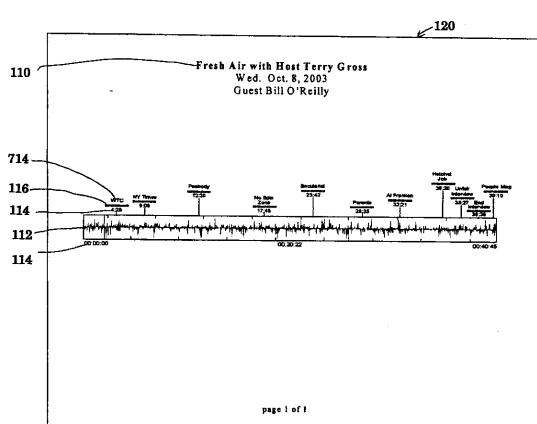


【図9a】

104 → 文書フォーマット仕様  
 Type = Radio program  
 Identifier = 1771  
 Title = Fresh Air with host Terry Gross  
 Annotation = Guest Bill O'Reilly  
 Publication Date = Oct. 8, 2003  
 Begin time = 00:00:00  
 End time = 00:40:45  
 Graphical Representation = Amplitude curve  
 Feature extraction = speech recognition and keyword match to "New York Times" and "fair and balanced"  
 Marker type = keywords, bar code, time stamp  
 Marker frequency = user-defined  
 Marker 1 = (WTC, bar code, 04:28, vert. pos. 1)  
 Marker 2 = (NY Times, bar code, 09:05, vert. pos. 2)  
 Marker 3 = (Peabody, bar code, 12:30, vert. pos. 3)  
 Layout type = One horizontal time line  
 Layout pages = 1  
 Layout marker placement = Above graphical representation, as provided  
 Layout meta data placement = Centered at top of page

602  
 音声特徴抽出  
 音声振幅抽出及びsvgファイル近似。  
 606  
 svgファイルを抽出する。

【図9b】

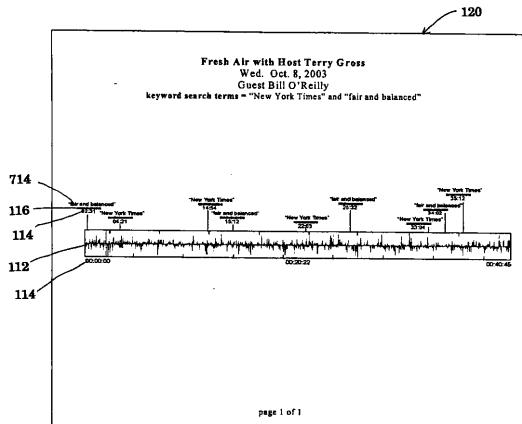


【図10a】

104 → 文書フォーマット仕様  
 Type = Radio program  
 Identifier = 1772  
 Title = Fresh Air with host Terry Gross  
 Annotation = Guest Bill O'Reilly  
 keyword search terms = "New York Times" and "fair and balanced"  
 Publication Date = Oct. 8, 2003  
 Begin time = 00:00:00  
 End time = 00:40:45  
 Graphical Representation = Amplitude curve  
 Feature extraction = speech recognition and keyword match to "New York Times" or "fair and balanced"  
 Marker type = matching search term, bar code, time stamp  
 Marker 1 = (New York Times, bar code, 02:31, vert. pos. 1)  
 Marker 2 = (Fair and balanced, bar code, 04:21, vert. pos. 1)  
 Marker 3 = (New York Times, bar code, 14:54, vert. pos. 2)  
 Marker 9 = (New York Times, bar code, 35:12, vert. pos. 3)  
 Layout type = One horizontal time line  
 Layout pages = 1  
 Layout marker placement = Above graphical representation, as provided  
 Layout meta data placement = Centered at top of page

602  
 音声特徴抽出  
 音声振幅抽出及びsvgファイル近似。  
 606  
 svgファイルを抽出する。

【図10b】

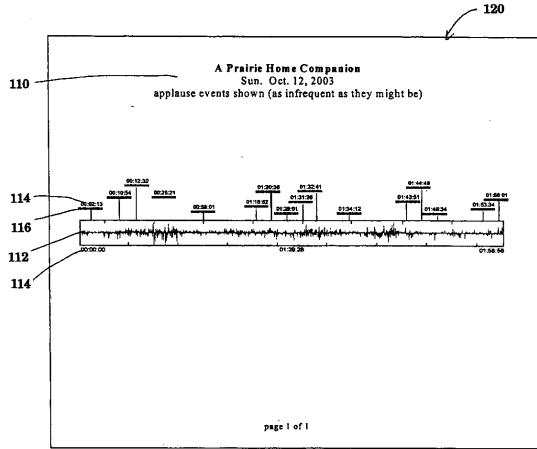


【図 1 1 a】

104 → 文書フォーマット仕様  
 Type = Radio program  
 Identifier = 1773  
 Title = A Prairie Home Companion  
 Annotation = applause events shown  
 Publication Date = Sun. Oct. 12, 2003  
 Begin time = 00:00:00  
 End time = 01:58:56  
 Graphical Representation = Amplitude curve  
 702 Feature extraction = applause detection alpha=0.27 beta=1.86  
 Marker type = time stamp, bar code  
 Marker placement = above graphical representation  
 Marker frequency = user-defined  
 Marker 1 = (00:02:13, bar code, pos. 1)  
 Marker 2 = (00:12:34, bar code, pos. 2)  
 Marker 3 = (00:12:32, bar code, pos. 3)  
 Marker 4 = (01:58:01, bar code, pos. 2)  
 Layout type = 1 horizontal time line  
 Layout pages = 1  
 Layout marker placement = Above graphical representation, as provided  
 Layout meta data placement = Centered at top of page

音声特徴抽出 ← 602  
 音声振幅抽出及び「フィック近似」  
 aif ファイルを出力する。 ← 606  
 押手検出はタイミングを出力する。

【図 1 1 b】

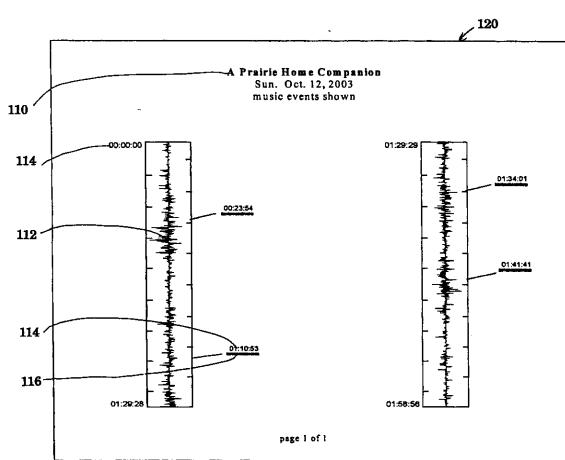


【図 1 2 a】

104 → 文書フォーマット仕様  
 Type = Radio program  
 Identifier = 1774  
 Title = A Prairie Home Companion  
 Annotation = music events shown  
 Publication Date = Sun. Oct. 12, 2003  
 Begin time = 00:00:00  
 End time = 01:58:56  
 Graphical Representation = Amplitude curve  
 702 Feature extraction = music detection delta = 12.93  
 Marker type = time stamp, bar code  
 Marker placement = right of time line  
 Marker frequency = user-defined  
 Marker 1 = (00:02:13, bar code, horiz. pos. 1)  
 Marker 2 = (01:10:53, bar code, horiz. pos. 1)  
 Marker 3 = (01:34:01, bar code, horiz. pos. 1)  
 Marker 4 = (01:41:41, bar code, horiz. pos. 1)  
 Layout type = 2 vertical time lines, split in half  
 Layout pages = 1  
 Layout marker placement = To right of graphical representation, as provided  
 Layout meta data placement = Centered at top of page

音声特徴抽出 ← 602  
 音声振幅抽出及び「フィック近似」  
 aif ファイルを出力する。 ← 606  
 音楽検出はタイミングを出力する。

【図 1 2 b】



## フロントページの続き

(51)Int.Cl.<sup>7</sup> F I テーマコード(参考)  
G 0 6 T 7/00 G 1 0 L 3/00 M  
G 1 0 L 13/00 G 0 6 K 19/00 Q

(72)発明者 ピーター イー ハート  
アメリカ合衆国, カリフォルニア 94025, メンロ・パーク, サンド・ヒル・ロード 288  
2番, スイート 115 リコー イノベーション インク内

F ターム(参考) 5B035 BB01 BB11 BC00  
5B072 BB00 CC24 DD02 JJ11 MM11  
5D015 KK02 KK04  
5D045 AA20  
5L096 FA00 HA08 JA11