

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4513165号
(P4513165)

(45) 発行日 平成22年7月28日(2010.7.28)

(24) 登録日 平成22年5月21日(2010.5.21)

(51) Int. Cl. F I
 HO4N 5/76 (2006.01) HO4N 5/76 B
 G1OL 15/00 (2006.01) G1OL 15/00 200G
 G1OL 15/00 200D

請求項の数 7 (全 8 頁)

<p>(21) 出願番号 特願2000-119854 (P2000-119854)</p> <p>(22) 出願日 平成12年4月20日 (2000.4.20)</p> <p>(65) 公開番号 特開2001-309282 (P2001-309282A)</p> <p>(43) 公開日 平成13年11月2日 (2001.11.2)</p> <p>審査請求日 平成19年1月25日 (2007.1.25)</p> <p>前置審査</p>	<p>(73) 特許権者 000002185 ソニー株式会社 東京都港区港南1丁目7番1号</p> <p>(72) 発明者 鶴飼 学 東京都品川区北品川6丁目7番35号 ソニー株式会社内</p> <p>審査官 清水 正一</p> <p>(56) 参考文献 特開平05-236506 (JP, A) 特開平11-112913 (JP, A) 特開平09-009202 (JP, A)</p>
--	--

最終頁に続く

(54) 【発明の名称】 番組記録方法及び番組記録装置及び番組記録再生装置及び番組記録再生方法

(57) 【特許請求の範囲】

【請求項1】

音声認識を行いながら番組を記録し、
 該音声認識の結果をテキストデータに変換して記録し且つ蓄積し、
 前記テキストデータから番組プログラムの内容を解析し、該解析した番組プログラムの内容を記録し且つ蓄積し、
 特定の音声信号をキーにして、蓄積されているテキストデータを検索する番組記録方法であって、

前記音声認識の結果と共に各音声認識単位の出現頻度をも記録することを特徴とする番組記録方法。

【請求項2】

音声認識を行いながら番組を記録する手段と、
 該音声認識の結果をテキストデータに変換して記録し且つ蓄積する手段と、
 前記テキストデータから番組プログラムの内容を解析し、該解析した番組プログラムの内容を記録し且つ蓄積する手段と、

特定の音声信号をキーにして、蓄積されているテキストデータを検索する手段と、を備えた番組記録装置であって、

前記音声認識の結果と共に各音声認識単位の出現頻度をも記録する手段を備えたことを特徴とする番組記録装置。

【請求項3】

音声認識を行いながら番組を記録する手段と、
該音声認識の結果をテキストデータに変換して記録し且つ蓄積する手段と、
入力したキーワードと前記蓄積されているテキストデータとを比較して一致したテキストデータに関連する番組を再生する再生手段と、を備えた番組記録再生装置であって、
前記音声認識の結果と共に各音声認識単位の出現頻度をも記録する手段を備えたことを特徴とする番組記録再生装置。

【請求項 4】

上記請求項 3 における番組記録再生装置において、
前記音声認識単位の出現頻度の高い順に表示して適宜選択し、該選択された音声認識単位のテキストデータに関連する番組を再生することを特徴とする番組記録再生装置。

10

【請求項 5】

上記請求項 3 における番組記録再生装置において、
前記テキストデータから番組プログラムの内容を解析する手段を備え、該解析したプログラムの内容を記録し且つ蓄積することを特徴とする番組記録再生装置。

【請求項 6】

上記請求項 5 における番組記録再生装置において、
前記蓄積されている番組プログラムの内容を表示して適宜選択し、該選択された番組プログラムの内容に関連する番組を再生することを特徴とする番組記録再生装置。

【請求項 7】

音声認識を行いながら番組を記録し、
該音声認識の結果をテキストデータに変換して記録し且つ蓄積し、
入力したキーワードと前記蓄積されているテキストデータとを比較して一致したテキストデータに関連する番組を再生する番組記録再生方法であって、
前記音声認識の結果と共に各音声認識単位の出現頻度をも記録することを特徴とする番組記録再生方法。

20

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、番組記録方法及び番組記録装置及び番組記録再生装置及び番組記録再生方法に関する。詳しくは、音声信号を認識しながら番組を記録する際に認識した音声信号をテキストデータに変換して記録し且つ蓄積する番組記録方法及び番組記録装置及び番組記録再生装置及び番組記録再生方法に関する。

30

【0002】

【従来の技術】

従来、放送番組を記録する時は、EPG（電子番組ガイド）等を検索して番組内容を確認し、所望の放送番組を記録している。又、記録媒体に記録している放送番組を検索する時は、記録した放送番組名一覧が表示され、その中から所望の番組を選択する方法や、記録した番組の先頭を検出して再生する頭出し機能による方法等がある。

【0003】

【発明が解決しようとする課題】

しかしながら、EPGの情報は分類方法が定型的であり、検索する時は所望のキーワードで検索することができない、又、EPGが付加されていない放送番組もある。一方、同一の記録媒体に多くの番組を記録している場合に、所望の放送番組や放送番組の見たい場面を検索したり、これらの検索を任意のキーワードで行うことができないという問題がある。

40

【0004】

従って、記録した放送番組や放送番組内の所望の場面をユーザが入力するキーワードで検索できるようにすることに解決しなければならない課題を有する。

【0005】

【課題を解決するための手段】

50

上記課題を解決するため、本発明に係る番組記録方法及び番組記録装置及び番組記録再生装置及び番組記録再生方法は、次に示す構成にすることである。

【0006】

(1) 音声認識を行いながら番組を記録し、

該音声認識の結果をテキストデータに変換して記録し且つ蓄積し、

前記テキストデータから番組プログラムの内容を解析し、該解析した番組プログラムの内容を記録し且つ蓄積し、

特定の音声信号をキーにして、蓄積されているテキストデータを検索する番組記録方法であって、

前記音声認識の結果と共に各音声認識単位の出現頻度をも記録することを特徴とする番組記録方法。

10

【0007】

(2) 音声認識を行いながら番組を記録する手段と、

該音声認識の結果をテキストデータに変換して記録し且つ蓄積する手段と、

前記テキストデータから番組プログラムの内容を解析し、該解析した番組プログラムの内容を記録し且つ蓄積する手段と、

特定の音声信号をキーにして、蓄積されているテキストデータを検索する手段と、を備えた番組記録装置であって、

前記音声認識の結果と共に各音声認識単位の出現頻度をも記録する手段を備えたことを特徴とする番組記録装置。

20

【0008】

(3) 音声認識を行いながら番組を記録する手段と、

該音声認識の結果をテキストデータに変換して記録し且つ蓄積する手段と、

入力したキーワードと前記蓄積されているテキストデータとを比較して一致したテキストデータに関連する番組を再生する再生手段と、を備えた番組記録再生装置であって、

前記音声認識の結果と共に各音声認識単位の出現頻度をも記録する手段を備えたことを特徴とする番組記録再生装置。

(4) 上記(3)における番組記録再生装置において、

前記音声認識単位の出現頻度の高い順に表示して適宜選択し、該選択された音声認識単位のテキストデータに関連する番組を再生することを特徴とする番組記録再生装置。

30

(5) 上記(3)における番組記録再生装置において、

前記テキストデータから番組プログラムの内容を解析する手段を備え、該解析したプログラムの内容を記録し且つ蓄積することを特徴とする番組記録再生装置。

(6) 上記(5)における番組記録再生装置において、

前記蓄積されている番組プログラムの内容を表示して適宜選択し、該選択された番組プログラムの内容に関連する番組を再生することを特徴とする番組記録再生装置。

(7) 音声認識を行いながら番組を記録し、

該音声認識の結果をテキストデータに変換して記録し且つ蓄積し、

入力したキーワードと前記蓄積されているテキストデータとを比較して一致したテキストデータに関連する番組を再生する番組記録再生方法であって、

前記音声認識の結果と共に各音声認識単位の出現頻度をも記録することを特徴とする番組記録再生方法。

40

【0009】

このように、番組を記録する際に、その音声信号を認識してテキストデータとして記録蓄積することにより、このテキストデータに基づいて後に番組の内容について検索したり、このテキストデータを加工してユーザに提示等する二次利用がし易くなる。

【0010】

【発明の実施の形態】

次に、本発明に係る番組記録方法及び番組記録装置及び番組記録再生装置及び番組記録再生方法の実施の形態を図面を参照して説明する。

50

【 0 0 1 1 】

本発明に係る番組記録方法及び番組記録再生装置及び番組記録再生方法を具現化する番組記録装置は、図 1 に示すように、デジタル放送番組 T S (トランスポートストリーム) を所定の packets に分別する D M U X 部 1 0 と、蓄積するデータの制御を行う制御部 2 0 と、音声認識の結果を処理する音声信号処理部 3 0 と、テキストデータを含む放送番組に関するデータを蓄積する情報蓄積部 4 0 と、情報蓄積部 4 0 に蓄積されているデータを所定の packets にする M U X 部 5 0 とから大略構成されている。

【 0 0 1 2 】

D M U X 部 1 0 は、デマルチプレクサであり、デジタル放送番組 T S (トランスポートストリーム) を画像 packets、音声 packets、カウンター値としての時間情報、その他のストリームに関する情報であるプログラム情報に分別して出力する。

10

【 0 0 1 3 】

制御部 2 0 は、プログラム情報や外部からの入力情報に従い、情報蓄積部 4 0 に対して packets の記録や消去及び出力などの指示を行う。

【 0 0 1 4 】

音声信号処理部 3 0 は、時間情報に従い音声 packets から音声信号を復元する D E C 部 3 1 と、音声信号を認識して文字データに変換する音声認識部 3 2 と、音声 packets と文字データの同期を取るための時間情報を生成するソフト部 3 3 と、文字データと時間情報を合成して出力する合成部 3 4 から構成されている。

【 0 0 1 5 】

情報蓄積部 4 0 は、制御部 2 0 の制御に従い、音声 packets や画像 packets、音声処理部 3 0 で処理された文字データ (テキストデータ)、プログラム情報等を記録媒体 4 1 に記録して蓄積する。又、制御部 2 0 から指定されたこれらの蓄積情報を記録媒体 4 1 から抽出して M U X 部 5 0 へ出力する。

20

【 0 0 1 6 】

M U X 部 5 0 は、マルチプレクサであり、制御部 2 0 からの指示に従って、情報蓄積部 4 0 から抽出された音声 packets、画像 packets、文字データ及びプログラム情報を放送番組ストリーム (T S ; トランスポートストリーム) に合成して出力する。

【 0 0 1 7 】

このような構成の番組記録装置により、デジタル放送番組を記録する際に音声認識の結果をテキストデータにして記録し、その記録したデータを検索して再生する時の動作を説明する。

30

【 0 0 1 8 】

まず、外部チューナ等から受信されたデジタル放送信号 (M P E G 2 - T S) が D M U X 部 1 0 へ送られる。 D M U X 部 1 0 はこのデジタル放送信号を解析して、画像 packets、音声 packets、システムクロックに同期する時間情報 (カウンター値)、プログラム情報に分離して出力する。

【 0 0 1 9 】

出力された音声 packets、画像 packets は情報蓄積部 4 0 へ送られ、プログラム情報は制御部 2 0 へ送られる。同時に音声 packets と時間情報は音声処理部 3 0 にも送られる。

40

【 0 0 2 0 】

制御部 2 0 は、 D M U X 部 1 0 から出力されたプログラム情報に基づき情報蓄積部 4 0 に対して送られてくる音声 packets、画像 packets を記録媒体 4 1 に記録するように指示を行う。

【 0 0 2 1 】

一方、音声処理部 3 0 は D M U X 部 1 0 からの音声 packets と時間情報により音声認識してその結果を文字データ (テキストデータ) に変換する処理を行う。この処理する過程を図 3 を併用して図 2 を参照して以下説明する。

【 0 0 2 2 】

先ず、 D M U X 部 1 0 から出力された音声 packets は、図 2 (A) に示すように、時間

50

情報 T 1 ~ T 5 をそれぞれに持った音声信号パケット A ~ E である。

【 0 0 2 3 】

D E C 部 3 1 は、このような音声信号パケット A ~ E を時間情報 T 1 ~ T 5 順に従って、ベースバンドの音声信号 A' ~ E' に復元して音声認識部 3 2 へ送る（図 2（B）参照）。音声認識部 3 2 では、音声信号 A' ~ E' を音声認識して文字データ（テキストデータ） a ~ e に変換して合成部 3 4 へ送る。音声信号から文字データに変換するまでの変換時間を要するため、認識した文字データ a ~ e の時間情報は、文字変換前に有していた時間 T 1 ~ T 5 に対して変換時間 だけ遅延することになる。又、この文字データ a ~ e は単語 w 1 ~ w n が集合したものであるため、その単語 w 1 ~ w n のそれぞれが変換時間 1 ~ n の遅延を含んだ時間 t 1 ~ t n という時間情報を持つ（図 2（C）参照）。

10

【 0 0 2 4 】

一方、シフト部 3 3 では、音声認識部 3 2 で音声認識や文字データへの変換時間の遅延を管理しカウンタ値として合成部 3 4 へ送る。合成部 3 4 は、記録媒体 4 1 に記録された音声パケットと文字データ a ~ e とを同期させる為、音声認識部 3 2 から送られてくる文字データ a ~ e の単語 w 1 ~ w n が持つ時間情報 t 1 ~ t n から、単語の変換時間 1 ~ n（遅延分）を差し引いた（t 1 - 1） ~ （t n - n）という新たな時間情報を生成する。そして、この時間情報（t 1 - 1） ~ （t n - n）を対応する各単語 w 1 ~ w n に新たな時間情報として付加する（図 2（D）参照）。

【 0 0 2 5 】

このような処理を行うことにより、図 3 に示すような音声パケットとの同期がとれた時間情報 t 1' ~ t n' をもつ文字データ a' ~ e' が生成され、この文字データ a' ~ e' を情報蓄積部 4 0 へ送る。情報蓄積部 4 0 は、制御部 2 0 に従い、D M U X 部 1 0 から送られてくる画像パケット、音声パケット、プログラム情報と、音声認識されて変換された文字データを各属性に分別して記録媒体 4 1 へ記録して蓄積する（図 4 参照）。

20

【 0 0 2 6 】

そして、このように蓄積された放送番組を検索する時に所望のキーワードを入力されると、入力されたキーワードは、制御部 2 0 で処理されて情報蓄積部 4 0 へ送られる。情報蓄積部 4 0 は、入力されたキーワードと記録媒体 4 1 に蓄積されている文字データを比較参照し、一致する文字データを検索する。そして、一致した文字データがあれば、その文字データの時間情報と関連する画像パケット、音声パケット、プログラム情報を抽出して文字データと共に M U X 部 5 0 へ出力する。M U X 部 5 0 では、制御部 2 0 の指示により情報蓄積部 4 0 から抽出された文字データ、画像パケット、音声パケットとプログラム情報を T S に合成して出力する。このキーワードは、即ち、検索するためのキーであり、特定の音声信号、例えば特定の歌手の音声信号や特定の音楽メロディーでも良く、適宜設定選択することができる。

30

【 0 0 2 7 】

このように放送番組の内容が文字データとして情報蓄積部 4 0 に蓄積されているので、所望のキーワードを入力すると、制御部 2 0 でキーワードにより蓄積された文字データを検索し、一致すれば所望の内容の場面が再生される。又、E P G を参照して文字データの比較分析を行い、番組内容概要を文字データとして表示させたり、文字データをそのまま若しくは翻訳等の変換を行って表示させることもできる。

40

【 0 0 2 8 】

尚、音声認識部 3 2 で音声認識する単位は任意の長さで良く、合成部 3 4 で付加する時間情報も単語毎でなく任意長の識別単位の文字データに付加するようにしても良い。又、この識別単位の文字データを複数まとめて時間情報を付加して時間情報のデータ量を減らせば、記憶媒体の容量節約ができる。更に自然界の音や楽器等の音声以外の音は別途区別され、これに任意に識別できる文字データを割当てて処理することも可能である。

【 0 0 2 9 】

次に、第 2 の実施の形態の番組記録装置について、第 1 の実施の形態で参照した図 1 を用いて説明する。

50

【 0 0 3 0 】

第2の実施の形態の番組記録装置は、音声認識した文字データの出現頻度をデータ化するものであり、その装置構成は第1の実施の形態と同じであるため説明は省略する。又、各部で同じ動作を行っている場合にもその説明は省略する。

【 0 0 3 1 】

図1の音声処理部30の音声認識部32は、DEC部31で復元された音声信号を音声認識して文字データに変換し、更にこの文字データの各単語A～D・・・の出現頻度a～c・・・を調べて図5のヒストグラムのようにデータ化し、この出現頻度a～c・・・を各単語A～Dに付加して、図6に示すようなHistデータを生成する。そして、このHistデータを文字データと共に情報蓄積部40へ送り、記録媒体41に記録して蓄積する。

10

【 0 0 3 2 】

そして、ユーザーによりキーワードが入力されると、そのキーワードとHistデータを参照して、記録されている放送番組からそのキーワードの出現頻度の高い放送番組名を順に一覧表示したり、更にEPGと組み合わせて記録する放送番組を検索することもできる。尚、出現頻度は単語ではなく任意長の文字データの出現頻度でもよいことは勿論である。

【 0 0 3 3 】

【 発明の効果 】

以上説明したように、番組を記録する際、番組の音声信号を音声認識させ、これを文字データとして蓄積させることにより、ユーザが任意のキーワードで記録した番組及びこの番組のある一場面を検索することができ、より効率的できめ細かい検索が可能になる。

20

【 0 0 3 4 】

更に、EPG等を利用することで番組内容の概要を知ることができる。又、EPGを持たない番組の検索も容易になり、加えて、文字データをそのまま若しくは翻訳して表示させることにより字幕として使用することもできる。

【 図面の簡単な説明 】

【 図 1 】 本発明に係る番組記録装置の主要部を示したブロック図である。

【 図 2 】 本発明に係る番組記録装置で音声パケットを音声認識して文字データに変換する時の時系列を表す説明図である。

30

【 図 3 】 本発明に係る文字データのフォーマット形式の一例を表す概念図である。

【 図 4 】 本発明に係る番組記録装置の記録媒体に記録されたデータのファイル構造を表す概念図である。

【 図 5 】 本発明に係る番組記録装置で文字データの出現頻度をヒストグラム化した状態を表す概念図である。

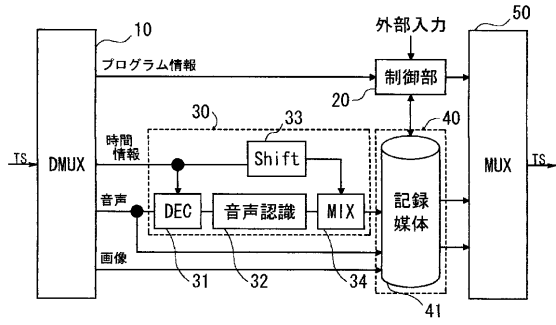
【 図 6 】 出現頻度を付加した文字データのフォーマット形式の一例を表す概念図である。

【 符号の説明 】

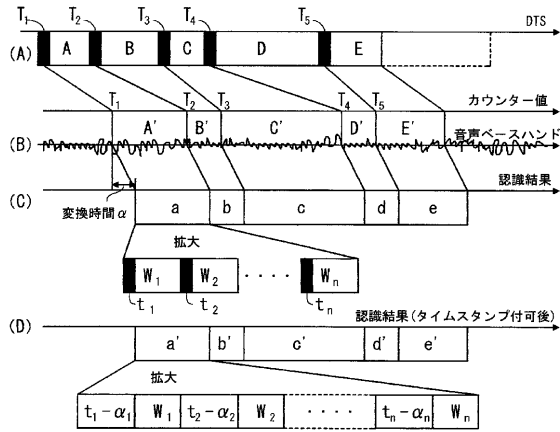
10 ; DMUX部、20 ; 制御部、30 ; 音声処理部、31 ; DEC部、32 ; 音声認識部、33 ; シフト部、34 ; 合成部、40 ; 情報蓄積部、41 ; 記録媒体、50 ; MUX部

40

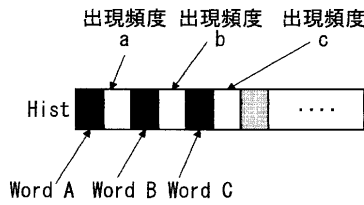
【図1】



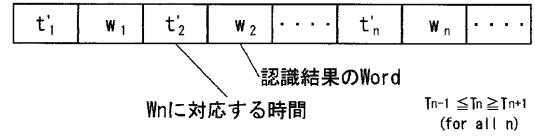
【図2】



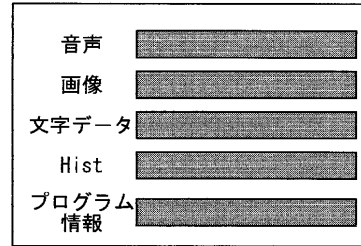
【図6】



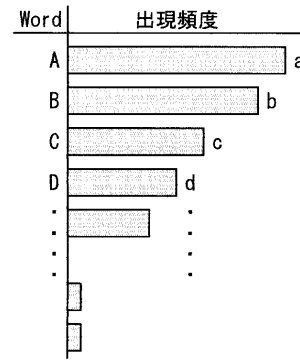
【図3】



【図4】



【図5】



フロントページの続き

(58)調査した分野(Int.Cl. , D B名)

H04N 5/76 - 5/956

G10L 15/00

G11B 20/10 - 20/16

G11B 27/00 - 27/06