

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6416698号
(P6416698)

(45) 発行日 平成30年10月31日 (2018. 10. 31)

(24) 登録日 平成30年10月12日 (2018. 10. 12)

(51) Int. Cl.	F I
HO 4 N 21/236 (2011. 01)	HO 4 N 21/236
HO 4 N 21/434 (2011. 01)	HO 4 N 21/434
HO 4 H 20/28 (2008. 01)	HO 4 H 20/28
HO 4 H 60/07 (2008. 01)	HO 4 H 60/07

請求項の数 4 (全 15 頁)

(21) 出願番号	特願2015-108638 (P2015-108638)	(73) 特許権者	000004352
(22) 出願日	平成27年5月28日 (2015. 5. 28)		日本放送協会
(65) 公開番号	特開2016-6958 (P2016-6958A)		東京都渋谷区神南2丁目2番1号
(43) 公開日	平成28年1月14日 (2016. 1. 14)	(74) 代理人	100064908
審査請求日	平成30年1月26日 (2018. 1. 26)		弁理士 志賀 正武
(31) 優先権主張番号	特願2014-113439 (P2014-113439)	(74) 代理人	100108578
(32) 優先日	平成26年5月30日 (2014. 5. 30)		弁理士 高橋 詔男
(33) 優先権主張国	日本国 (JP)	(72) 発明者	馬場 秋継
特許権者において、実施許諾の用意がある。			東京都世田谷区砧一丁目10番11号 日 本放送協会放送技術研究所内
早期審査対象出願		(72) 発明者	砂崎 俊二
			東京都世田谷区砧一丁目10番11号 日 本放送協会放送技術研究所内
		最終頁に続く	

(54) 【発明の名称】 送信装置および受信装置、ならびにプログラム

(57) 【特許請求の範囲】

【請求項 1】

提示タイミングが指定されたタイムドテキストデータを提示タイミンググループごとに M M T 伝送方式における一つのメディアプロセッシングユニットに格納するとともに、複数の前記メディアプロセッシングユニットから成るテキストアセットを生成するテキスト設定部と、

前記テキスト設定部が生成した前記テキストアセットを少なくとも含むコンテンツデータを多重化して配信データを生成する多重化部と、

前記多重化部によって多重化して得られた配信データを送出する送信部と、
を具備する送信装置であって、

前記提示タイミンググループは、複数の前記提示タイミングを含み得るものであり、

前記テキスト設定部は、前記提示タイミンググループに対応する前記タイムドテキストデータと当該タイムドテキストデータから参照されるリソースファイルとの各々を、同一の前記メディアプロセッシングユニット内の、別のメディアフラグメントユニットにそれぞれ格納することによって前記テキストアセットを生成する、

ことを特徴とする送信装置。

【請求項 2】

外部から配信データを含む信号を受信する受信部と、

前記受信部が受信した前記配信データから、M M T 伝送方式における一連のメディアプロセッシングユニットとして成るテキストアセットを分離する分離部と、

前記分離部によって分離された前記テキストアセットから前記メディアプロセッシングユニットを順次抽出し、前記メディアプロセッシングユニットに含まれるタイムドテキストデータを抽出するテキスト処理部と、

前記テキスト処理部によって抽出された前記タイムドテキストデータから提示タイミングの情報を取得するとともに、前記タイムドテキストデータに含まれるテキストを前記提示タイミングの情報にしたがうタイミングで提示するよう制御する提示制御部と、

を具備する受信装置であって、

前記テキストアセットは、提示タイミンググループごとのメディアプロセッシングユニットから成り、

前記提示タイミンググループは、複数の前記提示タイミングを含み得るものであり、

前記メディアプロセッシングユニットは、前記提示タイミンググループに対応する前記タイムドテキストデータと、当該タイムドテキストデータから参照されるリソースファイルとの各々を、当該メディアプロセッシングユニット内の、別のメディアフラグメントユニットにそれぞれ格納して成るものであり、

前記テキスト処理部は、前記メディアプロセッシングユニット内の前記タイムドテキストデータおよび前記リソースファイルを抽出し、

前記提示制御部は、前記テキスト処理部によって抽出された前記タイムドテキストデータと前記リソースファイルに基づいて提示用のレイアウトを生成し、前記提示タイミングの情報にしたがうタイミングで提示する、

ことを特徴とする受信装置。

【請求項 3】

提示タイミングが指定されたタイムドテキストデータを提示タイミンググループごとに M M T 伝送方式における一つのメディアプロセッシングユニットに格納するとともに、複数の前記メディアプロセッシングユニットから成るテキストアセットを生成するテキスト設定手段、

前記テキスト設定手段が生成した前記テキストアセットを少なくとも含むコンテンツデータを多重化して配信データを生成する多重化手段、

としてコンピューターを機能させるプログラムであって、

前記提示タイミンググループは、複数の前記提示タイミングを含み得るものであり、

前記テキスト設定手段は、前記提示タイミンググループに対応する前記タイムドテキストデータと当該タイムドテキストデータから参照されるリソースファイルとの各々を、同一の前記メディアプロセッシングユニット内の、別のメディアフラグメントユニットにそれぞれ格納することによって前記テキストアセットを生成する、

プログラム。

【請求項 4】

受信手段が受信した配信データから、M M T 伝送方式における一連のメディアプロセッシングユニットとして成るテキストアセットを分離する分離手段、

前記分離手段によって分離された前記テキストアセットから前記メディアプロセッシングユニットを順次抽出し、前記メディアプロセッシングユニットに含まれるタイムドテキストデータを抽出するテキスト処理手段、

前記テキスト処理手段によって抽出された前記タイムドテキストデータから提示タイミングの情報を取得するとともに、前記タイムドテキストデータに含まれるテキストを前記提示タイミングの情報にしたがうタイミングで提示するよう制御する提示制御手段、

としてコンピューターを機能させるプログラムであって、

前記テキストアセットは、提示タイミンググループごとのメディアプロセッシングユニットから成り、

前記提示タイミンググループは、複数の前記提示タイミングを含み得るものであり、

前記メディアプロセッシングユニットは、前記提示タイミンググループに対応する前記タイムドテキストデータと、当該タイムドテキストデータから参照されるリソースファイルとの各々を、当該メディアプロセッシングユニット内の、別のメディアフラグメントユ

10

20

30

40

50

ニットにそれぞれ格納して成るものであり、

前記テキスト処理手段は、前記メディアプロセッシングユニット内の前記タイムドテキストデータおよび前記リソースファイルを抽出し、

前記提示制御手段は、前記テキスト処理手段によって抽出された前記タイムドテキストデータと前記リソースファイルに基づいて提示用のレイアウトを生成し、前記提示タイミングの情報にしたがうタイミングで提示する、

プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

10

本発明は、送信装置および受信装置、ならびにそれらのプログラムに関する。特に、テレビ放送の、送信装置、受信装置、プログラムに関する。

【背景技術】

【0002】

テレビ放送における字幕テキスト等のデータは、テキストそのものと、そのテキストを視聴者側で提示するタイミングの情報とを含むものである。このような提示タイミングが指定されたテキストデータを、タイムドテキスト(timed text)、またはタイムドテキストデータと呼ぶ。このような提示タイミングに関する指定を伴うテキストの表現方法の一つとして、TTML(Timed Text Markup Language、タイムドテキストマークアップ言語)がある。TTMLを用いることにより、テキストの表示に関して、表示タイミングや、表示レイアウト属性(表示位置など)を指定して記述することができる。デジタルテレビ放送における、字幕テキストやその他のスーパーインポーズなどの情報を伝送するためにこのTTMLを利用することも検討されている。なお、字幕テキストは、例えば、放送コンテンツにおける音声発話の内容を、音声発話と同じ言語やあるいは異なる言語で表示するために用いられる。

20

【0003】

非特許文献1には、TTMLについての規定が記載されている。その第8章にはスタイリング(文字のフォントを含む)について記載されている。また第9章にはレイアウトについて記載されている。また第10章にはタイミングについて記載されている。

【先行技術文献】

30

【非特許文献】

【0004】

【非特許文献1】「Timed Text Markup Language 1 (TTML1) (Second Edition) W3C Recommendation 24 September 2013」, [online], 2013年9月24日, World Wide Web Consortium(W3C), [平成26年(西暦2014年)5月25日検索], インターネット<URL: <http://www.w3.org/TR/ttaf1-dfxp/>>

【発明の概要】

【発明が解決しようとする課題】

【0005】

非特許文献1は、TTMLデータそのものに関する規約を定めるものであるため、データ形式について記載されているが、そのデータを伝送する手順等については記載されていない。従来技術では、映像コンテンツの字幕等のタイムドテキストをTTMLで記述して伝送する場合には、コンテンツ全体についてのタイムドテキスト(例えば、1時間の映像コンテンツであれば、1時間分の映像に対応するタイムドテキスト)を予め伝送していた。

40

【0006】

しかしながら、そのような従来技術による方法では、受信装置側で蓄積しなければならないタイムドテキストのデータ量が膨大になり、必要となるメモリ等の資源が多くなってしまいう問題がある。また、コンテンツ全体についてのタイムドテキストを予め伝送するために、例えば、番組開始前に十分な時間の余裕をもってタイムドテキストの伝送を

50

行わなければならないという問題がある。また、放送の場合には、番組の途中から受信装置の電源がオンされる場合もあり、そのような場合にもタイムドテキストを遅滞なく伝送することは技術的な困難を伴う。

【 0 0 0 7 】

本発明は、上記の課題認識に基づいて行なわれたものであり、放送等において、適切なタイミングでタイムドテキストを伝送することのできる、送信装置および受信装置、ならびにそれらのプログラムを提供するものである。

【 0 0 0 8 】

また、本発明のさらなる課題は、上記のタイムドテキストの伝送において、画像、音声、ウェブフォントなどを組み合わせた字幕データあるいは文字スーパーのデータを確実に表示させるための、送信装置および受信装置、ならびにそれらのプログラムを提供することである。

【課題を解決するための手段】

【 0 0 0 9 】

[1] 上記の課題を解決するため、本発明の一態様による送信装置は、提示タイミングが指定されたタイムドテキストデータを複数の提示タイミンググループに分割するとともに、分割後の前記タイムドテキストデータを M M T 伝送方式における一つのメディアプロセッシングユニットに格納し、前記提示タイミンググループごとの複数のメディアプロセッシングユニットから成るテキストアセットを生成するテキスト設定部と、前記テキスト設定部が生成した前記テキストアセットを少なくとも含むコンテンツデータを多重化して配信データを生成する多重化部と、前記多重化部によって多重化して得られた配信データを送出する送信部と、を具備することを特徴とする。

【 0 0 1 0 】

[2] また、本発明の一態様は、上記の送信装置において、前記テキスト設定部は、分割後の前記タイムドテキストデータと、当該分割後の前記タイムドテキストデータから参照されるリソースファイルとの各々を、同一のメディアプロセッシングユニット内の、別のメディアフラグメントユニットにそれぞれ格納することによって前記テキストアセットを生成する、ことを特徴とする。

【 0 0 1 1 】

[3] また、本発明の一態様は、上記の送信装置において、前記テキスト設定部は、あるメディアプロセッシングユニット内におけるメディアフラグメントユニットの総数を表すデータを、当該メディアプロセッシングユニット内における所定のメディアフラグメントユニット内に格納して、前記テキストアセットを生成する、ことを特徴とする。

【 0 0 1 2 】

[4] また、本発明の一態様は、上記の送信装置において、前記テキスト設定部は、あるメディアプロセッシングユニット内における各々のメディアフラグメントユニットについて、メディアフラグメントユニットが格納するファイルのデータタイプおよびデータサイズを表すデータを、当該メディアプロセッシングユニット内における所定のメディアフラグメントユニット内に格納して、前記テキストアセットを生成する、ことを特徴とする。

【 0 0 1 3 】

[5] そして、本発明の一態様による受信装置は、外部から配信データを含む信号を受信する受信部と、前記受信部が受信した前記配信データから、M M T 伝送方式における一連のメディアプロセッシングユニットとして成るテキストアセットを分離する分離部と、

前記分離部によって分離された前記テキストアセットから前記メディアプロセッシングユニットを順次抽出し、前記メディアプロセッシングユニットに含まれるタイムドテキストデータを抽出するテキスト処理部と、前記テキスト処理部によって抽出された前記タイムドテキストデータから提示タイミングの情報を取得するとともに、前記タイムドテキストデータに含まれるテキストを前記提示タイミングの情報にしたがうタイミングで提示するよう制御する提示制御部と、を具備することを特徴とする。

【 0 0 1 4 】

〔 6 〕また、本発明の一態様は、上記の受信装置において、前記テキストアセットは、複数の提示タイミンググループに分割された結果として得られた前記提示タイミンググループごとのメディアプロセッシングユニットから成り、前記メディアプロセッシングユニットは、分割後の前記タイムドテキストデータと、当該分割後の前記タイムドテキストデータから参照されるリソースファイルとの各々を、当該メディアプロセッシングユニット内の、別のメディアフラグメントユニットにそれぞれ格納して成るものであり、前記テキスト処理部は、前記メディアプロセッシングユニット内の前記タイムドテキストデータおよび前記リソースファイルを抽出し、前記提示制御部は、前記テキスト処理部によって抽出された前記タイムドテキストと前記リソースファイルに基づいて提示用のレイアウトを生成し、前記提示タイミングの情報にしたがうタイミングで提示する、ことを特徴とする。

10

【 0 0 1 5 】

〔 7 〕また、本発明の一態様は、上記の受信装置において、前記テキスト処理部が抽出した前記リソースファイルを、当該リソースファイルが属するメディアプロセッシングユニットの処理の間、一時的に蓄積するとともに、当該メディアプロセッシングユニットの処理の終了後には一時的に蓄積した前記リソースファイルを消去するリソース一時蓄積部、をさらに具備することを特徴とする。

【 0 0 1 6 】

〔 8 〕また、本発明の一態様は、上記の受信装置において、前記テキスト処理部が前記リソースファイルのうちの少なくとも一つの抽出に失敗したときには、前記提示制御部は、前記タイムドテキストと、抽出に失敗しなかった前記リソースファイルとに基づいて、提示用のレイアウトを生成する、ことを特徴とする。

20

【 0 0 1 7 】

〔 9 〕また、本発明の一態様は、上記の受信装置において、前記テキストアセット内のメディアプロセッシングユニットにおける所定のメディアフラグメントユニットは、当該メディアプロセッシングユニット内におけるメディアフラグメントユニットの総数を表すデータを格納しており、前記テキスト処理部は、前記メディアプロセッシングユニットを処理する際に、当該メディアプロセッシングユニット内におけるメディアフラグメントユニットの総数と、当該メディアプロセッシングユニット内における受信済みのメディアフラグメントユニットの数とに基づいて、当該メディアプロセッシングユニットにおける全てのメディアフラグメントユニットの受信を完了したか否かを判定し、前記提示制御部は、前記テキスト処理部による判定の結果に基づき、当該メディアプロセッシングユニット内の全てのメディアフラグメントユニットの受信が完了している場合に提示用のレイアウトを生成する、ことを特徴とする。

30

【 0 0 1 8 】

〔 1 0 〕また、本発明の一態様は、上記の受信装置において、前記テキストアセット内のメディアプロセッシングユニットにおける所定のメディアフラグメントユニットは、当該メディアプロセッシングユニットにおける各々のメディアフラグメントユニットについて、メディアフラグメントユニットが格納するファイルのデータタイプおよびデータサイズを表すデータを格納しているものであり、前記テキスト処理部は、前記メディアプロセッシングユニットにおける前記所定のメディアフラグメントユニットを取得した際に、当該メディアプロセッシングユニットにおける各々のメディアフラグメントユニットについてメディアフラグメントユニットが格納するファイルのデータタイプおよびデータサイズを表すデータを参照することによって、当該メディアプロセッシングユニットに含まれる全てのメディアフラグメントユニットに含まれるファイルを処理するために必要な計算資源を確保する処理を行う、ことを特徴とする。

40

【 0 0 1 9 】

〔 1 1 〕また、本発明の一態様は、提示タイミングが指定されたタイムドテキストデータを複数の提示タイミンググループに分割するとともに、分割後の前記タイムドテキスト

50

データをMMT伝送方式における一つのメディアプロセッシングユニットに格納し、前記提示タイミンググループごとの複数のメディアプロセッシングユニットから成るテキストアセットを生成するテキスト設定手段、前記テキスト設定手段が生成した前記テキストアセットを少なくとも含むコンテンツデータを多重化して配信データを生成する多重化手段、としてコンピューターを機能させるプログラムである。

【0020】

[12] また、本発明の一態様は、受信手段が受信した配信データから、MMT伝送方式における一連のメディアプロセッシングユニットとして成るテキストアセットを分離する分離手段、前記分離手段によって分離された前記テキストアセットから前記メディアプロセッシングユニットを順次抽出し、前記メディアプロセッシングユニットに含まれるタイムドテキストデータを抽出するテキスト処理手段、前記テキスト処理手段によって抽出された前記タイムドテキストデータから提示タイミングの情報を取得するとともに、前記タイムドテキストデータに含まれるテキストを前記提示タイミングの情報にしたがうタイミングで提示するよう制御する提示制御手段、としてコンピューターを機能させるプログラムである。

【発明の効果】

【0021】

本発明によれば、映像等のコンテンツに付随するタイムドテキストを、提示タイミングのグループごとに分割して、適切なタイミングで伝送することができる。つまり、コンテンツ全体に対応するタイムドテキストをまとめて予め伝送する必要がない。

また、タイムドテキストから参照されるリソースファイル等を確実に伝送し、提示することができる。

【図面の簡単な説明】

【0022】

【図1】本発明の実施形態による送信装置および受信装置の各々の概略機能構成を示すブロック図である。

【図2】同実施形態による送信装置から受信装置へ伝送される、字幕等のタイムドテキストを伝送するためのデータ構成（アセットの構成）を示す概略図である。

【図3】同実施形態による送信装置から受信装置へ伝送される、字幕伝送アセット内の各MFUにおける、字幕サンプルヘッダーのデータ構成を示す概略図である。

【図4】同実施形態による字幕サンプルヘッダーにおける、データタイプの値とその意味を示す概略図である。

【発明を実施するための形態】

【0023】

次に、図面を参照しながら、本発明の実施形態について説明する。

図1は、本実施形態における送信装置および受信装置の各々の概略機能構成を示すブロック図である。なお、テレビ放送システムにおいて、実際には受信装置2は多数存在するが、本図では便宜上1台の受信装置のみを示している。

図示するように、送信装置1は、送信部11と、多重化部12と、映像・音声設定部14と、テキスト設定部15とを含んで構成される。また、受信装置2は、受信部21と、分離部22と、映像・音声処理部24と、テキスト処理部25と、リソース一時蓄積部26と、テキスト提示制御部27と、合成表示・出力部28とを含んで構成される。

なお、送信装置1および受信装置2は、テレビ放送の信号をそれぞれ送受信するものである。テレビ放送の信号の送受信のためには、テキストアセットだけではなく、映像や音声や制御情報などの処理のための機能を各装置は有するが、ここでは、本発明に直接関係する機能のみについて説明し、その他の機能については説明を省略する。

【0024】

まず送信装置1内の各部の機能から説明する。

送信部11は、多重化部12によって多重化して得られた配信データを変調して放送信号として送出するものである。送信部11は、電波（地上波や放送衛星波）あるいはテレ

10

20

30

40

50

ビ用ケーブルにより、放送信号を送信する。

多重化部 1 2 は、テキスト設定部 1 5 が生成したテキストアセットを少なくとも含むコンテンツデータを多重化して配信データを生成する。多重化部 1 2 は、MMT (MPEG Media Transport) 伝送方式による多重化を行う。具体的には、多重化部 1 2 は、テキストアセットとともに、映像・音声設定部 1 4 によって生成される映像アセットや音声アセットを多重化し、配信データを生成する。ここで、テキストアセットとは、字幕アセットや文字スーパーアセットである。

映像・音声設定部 1 4 は、供給される映像信号や音声信号を取得し、符号化し、それぞれ、映像アセットや音声アセットのコンテンツデータとして生成する。映像や音声はストリーム型のデータであり、これらを総称してストリーム型コンテンツデータと呼ぶ。

10

【0025】

テキスト設定部 1 5 は、提示タイミングが指定されたタイムドテキストデータを複数の提示タイミンググループに分割するとともに、分割後のタイムドテキストデータを MMT 伝送方式における一つのメディアプロセッシングユニットに格納し、提示タイミンググループごとの複数のメディアプロセッシングユニットから成るテキストアセットを生成する。

【0026】

ここで、テキストアセットとは、字幕アセットあるいは文字スーパーアセットである。タイムドテキストデータは、受信機側の提示のタイミング (年月日、時分秒など) の情報を含んでいる。提示タイミンググループとは、1 つまたはそれ以上の提示タイミングの集合である。提示タイミンググループが複数の提示タイミングを含むとき、それらは、隣接するタイミングの一群である。

20

なお、テキストアセットの内部のデータ構成については後で図面を参照しながら詳述する。

【0027】

テキスト設定部 1 5 は、また、分割後のタイムドテキストデータと、当該分割後のタイムドテキストデータから参照されるリソースファイルとの各々を、同一のメディアプロセッシングユニット内の、別のメディアフラグメントユニットにそれぞれ格納することによってテキストアセットを生成する。

【0028】

テキスト設定部 1 5 は、また、あるメディアプロセッシングユニット内におけるメディアフラグメントユニットの総数を表すデータを、当該メディアプロセッシングユニット内における所定のメディアフラグメントユニット内に格納して、テキストアセットを生成する。具体的には、テキスト設定部 1 5 は、メディアプロセッシングユニット内の先頭のメディアフラグメントユニットにおけるヘッダー部分に、その総数を表すデータを格納する。

30

【0029】

テキスト設定部 1 5 は、また、あるメディアプロセッシングユニット内における各々のメディアフラグメントユニットについて、メディアフラグメントユニットが格納するファイルのデータタイプおよびデータサイズを表すデータを、当該メディアプロセッシングユニット内における所定のメディアフラグメントユニット内に格納して、テキストアセットを生成する。具体的には、テキスト設定部 1 5 は、メディアプロセッシングユニット内の先頭のメディアフラグメントユニットにおけるヘッダー部分に、後続する各メディアフラグメントユニットが格納するファイルのデータタイプとデータサイズ (バイト数) の情報を格納する。この情報は、テキストアセットを受信する受信装置側では、すべてのメディアフラグメントユニットのデータの処理のためのヒント情報となり得る。つまり、受信装置側ではヒント情報に基づいて、リソースファイルを処理するための計算資源 (受信機側における、メモリやコンピューター処理の資源) を準備することができる。

40

なお、テキスト設定部 1 5 は、ヒント情報の格納を省略するようにしても良い。

【0030】

50

送信装置 1 から送出された放送信号は、電波（地上波や放送衛星波）により、あるいはケーブルにより、多数の受信装置 2 に伝送される。

【0031】

次に、受信装置 2 内の各部の機能について説明する。

受信部 2 1 は、外部から放送信号を受信する。この放送信号には、送信装置 1 によって送信された配信データが含まれている。受信部 2 1 は、受信した放送信号を復調する。

分離部 2 2 は、受信部 2 1 が受信した配信データから、MMT 伝送方式における一連のメディアプロセッシングユニットとして成るテキストアセットを分離する。また、分離部 2 2 は、配信データから、映像アセットや音声アセットなどを分離する。

映像・音声処理部 2 4 は、分離部 2 2 によって分離された映像アセットや音声アセットをそれぞれ復号し、映像信号および音声信号を生成する。映像・音声処理部 2 4 は、同時に、映像および音声の提示タイミングに関する情報を、それぞれ、映像アセットおよび音声アセットから抽出する。

【0032】

テキスト処理部 2 5 は、分離部 2 2 によって分離されたテキストアセットからメディアプロセッシングユニットを順次抽出し、メディアプロセッシングユニットに含まれるタイムドテキストデータを抽出する。

テキスト提示制御部 2 7 は、テキスト処理部 2 5 によって抽出されたタイムドテキストデータから提示タイミングの情報を取得するとともに、そのタイムドテキストデータに含まれるテキストを指定された提示タイミングで提示するように制御する。なお、テキスト提示制御部 2 7 は、テキストを、指示されたレイアウトに従って配置する。テキストアセットは、後述するように、画像のリソースファイルや、音声のリソースファイルや、フォントファイルを含んでいる場合があり、テキスト提示制御部 2 7 は、これらの各リソースをも統合してレイアウトする。

なお、受信機側の表示画面において、字幕プレーンと文字スーパープレーンとは別のプレーンである。したがって、テキスト提示制御部 2 7 は、字幕アセットとして取得したテキストについては字幕プレーンにおいて提示するようにし、文字スーパーアセットとして取得したテキストについては文字スーパープレーンにおいて提示するようにする。つまり、テキスト提示制御部 2 7 は、これら両プレーンを互いに独立に制御する。

合成表示・出力部 2 8 は、テキスト提示制御部 2 7 によって出力される字幕プレーンおよび文字スーパープレーンと、映像・音声処理部 2 4 から出力される映像プレーンとを合成して画面を表示するとともに、映像・音声処理部 2 4 から渡される音声を出力する。

【0033】

なお、テキストアセットは、複数の提示タイミンググループに分割された結果として得られた提示タイミンググループごとのメディアプロセッシングユニットから成る。また、メディアプロセッシングユニットは、分割後のタイムドテキストデータと、当該分割後のタイムドテキストデータから参照されるリソースファイルとの各々を、当該メディアプロセッシングユニット内の、別のメディアフラグメントユニットにそれぞれ格納して成るものである。

テキスト処理部 2 5 は、メディアプロセッシングユニット内のタイムドテキストデータおよび単数または複数のリソースファイルを抽出する。なお、テキスト処理部 2 5 は、抽出したリソースファイルをリソース一時蓄積部 2 6 に一時的に記憶させる。

リソース一時蓄積部 2 6 は、抽出されたリソースファイルを一時的に記憶する。つまり、リソース一時蓄積部 2 6 は、いわゆるキャッシュメモリの役割を果たす。本実施形態では、リソースファイルが参照されるのは 1 つのメディアプロセッシングユニットの範囲内においてのみである。したがって、あるメディアプロセッシングユニットの処理が完了して、次の新たなメディアプロセッシングユニットの受信および処理に移る際には、リソース一時蓄積部 2 6 の内容は一旦消去される。つまり、あるメディアプロセッシングユニットから次のメディアプロセッシングユニットの処理に移る段階で、リソース一時蓄積部 2 6 の内容を毎回クリアすることができる。即ち、本実施形態が複数のメディアプロセッシ

10

20

30

40

50

ングユニット（各々は提示タイミングのグループに対応）に分割してタイムドテキストを伝送する方式をとっているため、リソースファイルを蓄積するメモリの容量が最小限で済む。

つまり、リソース一時蓄積部 26 は、テキスト処理部 25 が抽出したリソースファイルを、当該リソースファイルが属するメディアプロッシングユニットの処理の間、一時的に蓄積するとともに、当該メディアプロッシングユニットの処理の終了後には一時的に蓄積したリソースファイルを消去するものである。

そして、テキスト提示制御部 27 は、テキスト処理部 25 によって抽出されたタイムドテキストとリソースファイルに基づいて提示用のレイアウトを生成し、提示タイミングの情報にしたがうタイミングで提示するように制御する。

10

【0034】

テキスト処理部 25 がリソースファイルのうちの少なくとも一つの抽出に失敗したときには、テキスト提示制御部 27 は、タイムドテキストと、抽出に失敗しなかったリソースファイルとに基づいて、提示用のレイアウトを生成する。

つまり、伝送におけるメディアフラグメントユニットの単位と、論理的なファイル（タイムドテキストデータのファイルおよび画像や音声やフォントなどのリソースファイル）の単位とを一致させる。これにより、仮に一部のメディアフラグメントユニットの受信に失敗した場合にも、他のメディアフラグメントユニット（即ち他の論理的なファイル）への影響を及ぼすことなく、字幕や文字スーパーの提示を行うことができるようになる。

【0035】

20

また、テキストアセット内のメディアプロセッシングユニットにおける所定のメディアフラグメントユニットは、当該メディアプロセッシングユニット内におけるメディアフラグメントユニットの総数を表すデータを格納している。

テキスト処理部 25 は、メディアプロセッシングユニットを処理する際に、当該メディアプロセッシングユニット内におけるメディアフラグメントユニットの総数と、当該メディアプロセッシングユニット内における受信済みのメディアフラグメントユニットの数とに基づいて、当該メディアプロセッシングユニットにおける全てのメディアフラグメントユニットの受信を完了したか否かを判定する。なお、メディアプロセッシングユニット内におけるメディアフラグメントユニットの連番を付与して、メディアフラグメントユニット内に格納することにより、テキスト処理部 25 は、受信したメディアフラグメントユニットが、当該メディアプロセッシングユニットにおける何番目のものであるかを知ることができる。

30

テキスト提示制御部 27 は、テキスト処理部 25 による判定の結果に基づき、当該メディアプロセッシングユニット内の全てのメディアフラグメントユニットの受信が完了している場合に提示用のレイアウトの生成を開始することができる。

【0036】

また、テキストアセット内のメディアプロセッシングユニットにおける所定のメディアフラグメントユニットは、当該メディアプロセッシングユニットにおける各々のメディアフラグメントユニットについて、メディアフラグメントユニットが格納するファイルのデータタイプおよびデータサイズを表すデータを格納しているものである。

40

テキスト処理部 25 は、メディアプロセッシングユニットにおける所定のメディアフラグメントユニットを取得した際に、当該メディアプロセッシングユニットにおける各々のメディアフラグメントユニットについてメディアフラグメントユニットが格納するファイルのデータタイプおよびデータサイズを表すデータを参照することによって、当該メディアプロセッシングユニットに含まれる全てのメディアフラグメントユニットに含まれるファイルを処理するために必要な計算資源（メモリやCPU）を確保する処理を行う。

具体的には、メディアプロセッシングユニット内の先頭のメディアフラグメントユニットのヘッダー部分に後続する各メディアフラグメントユニットのデータタイプおよびデータ長の情報を格納することにより、テキスト処理部は、そのメディアフラグメントユニットを受けた段階で、予め、計算資源の確保を行っておくことができる。

50

【 0 0 3 7 】

次に、伝送されるデータについて説明する。

図 2 は、字幕等のタイムドテキストを送送するためのデータ構成（アセットの構成）を示す概略図である。本実施形態では、送信装置 1 は、映像や音声やテキストなどからなるテレビの放送信号を M M T（MPEG Multimedia Transport）方式で多重化して送信する。同図では、字幕アセットの例を示しているが、文字スーパー（文字によるスーパーインポーズ）アセットの場合も同様である。映像アセットや、音声アセットや、字幕アセットや、文字スーパーアセットなどが、ひとつのパッケージを構成する。あるパッケージ内における各アセットの所在は、M P テーブル（M M T パッケージテーブル；不図示）において指し示される。

10

【 0 0 3 8 】

図示するように、字幕アセットは、複数の M P U（Media Processing Unit，メディアプロセッシングユニット，メディア処理ユニット）の時系列として構成される。各 M P U は、複数の M F U（Media Fragment Unit，メディアフラグメントユニット，メディア断片ユニット）に分割して伝送することができる。ここで、M F U への分割のしかたは一般的には任意であるが、本実施形態では、T T M L 文書や、T T M L 文書から参照される画像ファイル、音声ファイル、フォントファイルなど、これらの論理的なファイルの単位を、M F U の単位と一致させる。また、各 M P U 内における先頭の M F U は、T T M L 文書ファイルとする。つまり、1 つの M P U は、1 つの T T M L 文書ファイル（1 つ目の M F U）と、その T T M L 文書から参照するリソース（画像ファイル、音声ファイル、動画フ

20

ァイル、フォントファイル等であり、各々のファイルが M F U に相当）から成る。但し、T T M L 文書から参照されるこれらのリソース数は、0 個以上（0 個であってもよい）である。なお、T T M L 文書の中には、T T M L の規約に基づく提示タイミングの情報が含まれる。1 つの T T M L 文書内に、隣接する複数の提示タイミングが含まれていても良い。なお、1 つの T T M L 文書内に複数の提示タイミングが含まれている場合にも、M P U は、それらのタイミングのうちの代表的タイミングの情報、例えば複数の提示タイミングのうち最初の提示タイミングを M U P タイムスタンプとして有することができる。上記のように、字幕伝送アセットに含まれる各 M P U は、1 つまたは複数の提示タイミングをもつ T T M L 文書による字幕データとその関連データをカプセル化したものである。

30

【 0 0 3 9 】

また、本実施形態では、各 M F U は、字幕サンプルヘッダーを有している。各 M P U における 1 つ目の M F U（これは、T T M L 文書の M F U である）の字幕サンプルヘッダーには、当該 M F U に関する情報だけではなく、当該 M P U 内の他の M F U に関する情報をも格納することができる。字幕サンプルヘッダーについては、図 3 を用いて説明する。

【 0 0 4 0 】

図 3 は、送信装置 1 が送信する字幕伝送アセット内の各 M F U における、字幕サンプルヘッダーのデータ構成を示す概略図である。文字スーパーにおける各 M F U のヘッダーも、同様のデータ構成を有する。同図では、データ構造と、そのデータ構造に含まれる各データ項目のビット数およびデータ表記を示している。同図では、データ構造は、ブロック構造を有する形式言語で表現されている。なお、データ表記における「u i m s b f」は

40

、「unsigned integer, most significant bit first」、即ち、符号なし整数であって最上位ビットが先頭であることを表す。また「b s l b f」は、「bit string, left bit first」、即ち、ビット列であって左ビットが先頭であることを表す。

【 0 0 4 1 】

字幕サンプルヘッダー内の各フィールドについて以下に説明する。

（1）subtitle_tag（字幕識別タグ）：字幕や文字スーパーのストリームを識別するためのラベルである。

（2）subtitle_sequence_number（字幕シーケンス番号）：当該字幕アセット中で送る字幕データ全体のシーケンス番号を示す。同一 M P U 内において共通の番号が格納される。また、連続する M P U 間で 1 ずつインクリメントされる。

50

(3) subsample_number (字幕データ番号) : 当該 M P U での、字幕データ(subsample)の番号を昇順で表す。当該 M P U 中の最初の M F U で送られる字幕データの番号は「 0 」である。同一 M P U 内において連続する M F U 間で 1 ずつインクリメントされる。

(4) last_subsample_number (最終字幕データ番号) : 当該 M P U で送る、字幕データの最後の subsample_number を示す。

(5) data_type (データタイプ) : 本 M F U で送る字幕データ(subsample)のデータタイプを示す。データタイプの種類とその値については、後述する。

【 0 0 4 2 】

(6) length_extension_flag (データ長拡張フラグ) : 下に現れるデータ長のフィールドのビット数を拡張するか否かを示すフラグである。ビット数を拡張して 3 2 ビットとする (0 から、4 , 2 9 4 , 9 6 7 , 2 9 5 まで表現可能) 場合にはその値は「 1」、ビット数を拡張せずに 1 6 ビットとする (0 から、6 5 , 5 3 5 まで表現可能) 場合にはその値は「 0 」とする。

10

(7) subsamples_info_flag (字幕データヒント情報フラグ) : 当該 M P U で伝送する、全ての字幕データ(subsample)に関するヒント情報を記述するか否かを表すフラグ情報である。ヒント情報が含まれる場合には場合の値は「 1」、含まれない場合の値は「 0 」とする。

【 0 0 4 3 】

(8) data_size (データサイズ) : 当該 M F U で送る字幕データのサイズ (バイト数) を示す。

20

(9) subsample_i_content_type (i 番目字幕データタイプ) : i 番目の字幕データ (M F U) のデータタイプを示す。データ型の種類とその値については、後述する。

(1 0) subsample_i_data_size (i 番目字幕データサイズ) : i 番目の字幕データ (M F U) のサイズを示す。

(1 1) data_byte (データバイト) : 当該 M F U の字幕データのデータバイトである。

【 0 0 4 4 】

なお、当該 M F U の字幕データ番号が「 0 」 (即ち、先頭の M F U) であって、且つ (4) の最終字幕データ番号が 0 より大きく (即ち、T T M L 文書から参照されるリソースファイルの M F U が 1 個以上存在し)、且つ (7) の字幕データヒント情報フラグが「 1 」である場合に、そしてその場合にのみ、(9) の i 番目字幕データタイプと (1 0) の i 番目字幕データサイズのペアのループは存在する。

30

また、(6) のデータ長拡張フラグは、(8) のデータサイズと (1 0) の i 番目字幕データサイズの、両方のビット数に対して効く。

【 0 0 4 5 】

図 4 は、上述した字幕サンプルヘッダーにおける、データタイプ (図 3 におけるデータ構造の第 6 行) および i 番目字幕データタイプ (図 3 におけるデータ構造の第 1 9 行) の値とその意味を示す概略図である。なお、データタイプはビット列表現である。

データタイプの値が「 0 0 0 0 」であるとき、その M F U は、A R I B - T T M L の文書ファイルを格納している。A R I B - T T M L は、T T M L の規格の一つである。

40

データタイプの値が「 0 0 0 1 」であるとき、その M F U は、P N G (Portable Network Graphics) 形式の画像ファイルを格納している。

データタイプの値が「 0 0 1 0 」であるとき、その M F U は、S V G (Scalable Vector Graphics) 形式の画像ファイルを格納している。

データタイプの値が「 0 0 1 1 」であるとき、その M F U は、P C M (A I F F - C) 形式の音声ファイルを格納している。P C M は、パルスコード変調を表す。A I F F は、Audio Interchange File Format の略である。

データタイプの値が「 0 1 0 0 」であるとき、その M F U は、M P 3 (MPEG-1 Audio Layer-3) 形式の音声ファイルを格納している。

データタイプの値が「 0 1 0 1 」であるとき、その M F U は、M P E G - 4 A A C (

50

Advanced Audio Coding, 先進的音響符号化)形式の音声ファイルを格納している。なお、MPEGは、Moving Picture Experts Groupの略である。

データタイプの値が「0110」であるとき、そのMFUは、SVG(Scalable Vector Graphics)形式のフォントファイルを格納している。SVGは、ベクターフォントの一種である。

データタイプの値が「0111」であるとき、そのMFUは、WOFF(Web Open Font Format)形式のフォントファイルを格納している。

データタイプの値の「1000」から「1111」までの値は、将来の使用のためにリザーブされている。

【0046】

なお、上述した実施形態における送信装置、受信装置の機能の少なくとも一部をコンピュータで実現するようにしても良い。その場合、これらの機能を実現するためのプログラムをコンピュータ読み取り可能な記録媒体に記録して、この記録媒体に記録されたプログラムをコンピュータシステムに読み込ませ、実行することによって実現しても良い。なお、ここでいう「コンピュータシステム」とは、OSや周辺機器等のハードウェアを含むものとする。また、「コンピュータ読み取り可能な記録媒体」とは、フレキシブルディスク、光磁気ディスク、ROM、CD-ROM等の可搬媒体、コンピュータシステムに内蔵されるハードディスク等の記憶装置のことをいう。さらに「コンピュータ読み取り可能な記録媒体」とは、インターネット等のネットワークや電話回線等の通信回線を介してプログラムを送信する場合の通信線のように、短時間の間、動的にプログラムを保持するもの、その場合のサーバーやクライアントとなるコンピュータシステム内部の揮発性メモリのように、一定時間プログラムを保持しているものも含んでも良い。また上記プログラムは、前述した機能の一部を実現するためのものであっても良く、さらに前述した機能をコンピュータシステムにすでに記録されているプログラムとの組み合わせで実現できるものであっても良い。

【0047】

以上説明した本実施形態の構成によれば、MMTによる伝送において、画像や音声、ウェブフォントなどを組み合わせた字幕あるいは文字スーパーを確実に表示させることができるようになる。

【0048】

また、本実施形態の構成によれば、メディアを意識してMFUを構成し、つまり、論理的なファイル(TTMLファイルやリソースファイル)の単位とMFU単位とを一致させて伝送するようにしていることで、パケットロスのような伝送品質低下が発生したときの誤りの伝搬を抑えることが可能である。例えば、画像ファイルの部分のMFUがロスしても、最低限、字幕テキストの文字は表示可能であるなどといった、耐障害性が得られる。

【0049】

また、本実施形態の構成によれば、subsample_number(MPU内における当該MFU(受信中のMFU)の通し番号)とlast_subsample_number(MPU内におけるMFUの総数の情報)を、各MFUのヘッダー内に入れるようにしたことで、当該MFUがMPU中の最後のMFUかどうかを判別でき、これにより全てのデータが揃ったタイミングで確実に字幕の提示ができる。なお、仮にこのヒント情報がないと、MPUのデータが揃ったかわからず、タイムリーな提示ができない。

【0050】

また、本実施形態の構成によれば、最初のMFUにつける字幕サンプルヘッダー中に、後続のMFUに入る字幕情報に関する情報(データタイプおよびデータサイズ)を入れることで、受信機が予めメモリを確保するなどの準備ができるようになる。

【0051】

以上、実施形態を説明したが、本発明はさらに次のような変形例でも実施することが可能である。

例えば、本発明をテレビ放送以外に適用しても良い。テレビ放送に限らず、MMT伝送

10

20

30

40

50

方式を用いてストリーム型コンテンツおよびタイムドテキストを配信するサービス（例えば、インターネットを利用したコンテンツ配信サービス）等において、上述の実施形態と同様の構成を適用することができる。

【 0 0 5 2 】

以上、この発明の実施形態について図面を参照して詳述してきたが、具体的な構成はこの実施形態に限られるものではなく、この発明の要旨を逸脱しない範囲の設計等も含まれる。

【産業上の利用可能性】

【 0 0 5 3 】

本発明は、例えば、テレビ放送のためのシステム（送信装置や受信装置）のために利用可能である。また、本発明は、テレビ放送に限らず、タイムドテキストを伴ったストリーム型のコンテンツを配信するためのシステムのために利用可能である。

10

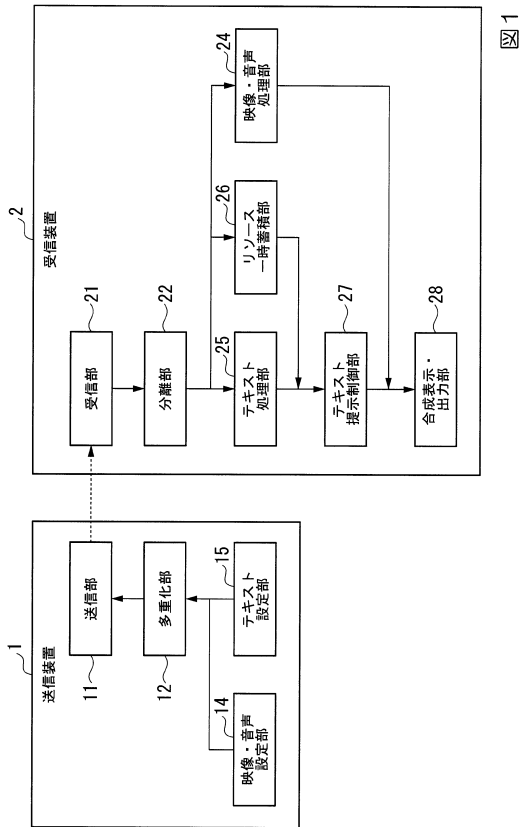
【符号の説明】

【 0 0 5 4 】

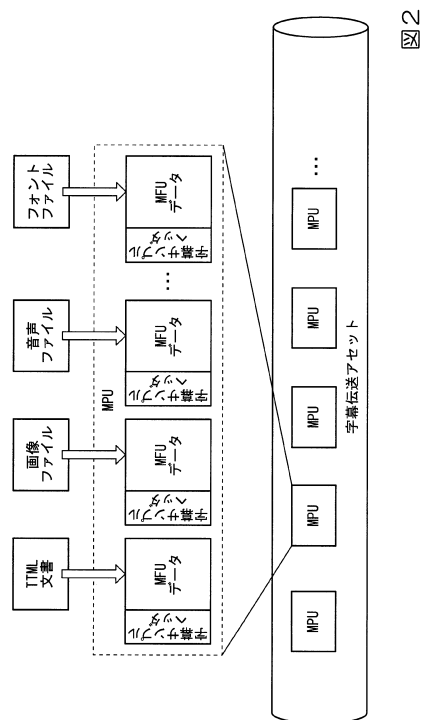
- 1 送信装置
- 2 受信装置
- 1 1 送信部
- 1 2 多重化部
- 1 4 映像・音声設定部
- 1 5 テキスト設定部
- 2 1 受信部
- 2 2 分離部
- 2 4 映像・音声処理部
- 2 5 テキスト処理部
- 2 6 リソース一時蓄積部
- 2 7 テキスト提示制御部
- 2 8 合成表示・出力部

20

【図 1】



【図 2】



【図 3】

	データ構造	ビット数	データ表記
1	Subtitle_sample_header() {		
2	subtitle_tag	8	uimsbf
3	subtitle_sequence_number	8	uimsbf
4	subsample_number	8	uimsbf
5	last_subsample_number	8	uimsbf
6	data_type	4	bslbf
7	length_extension_flag	1	bslbf
8	subsamples_info_flag	1	bslbf
9	Reserved	3	bslbf
10	if (length_extension_flag==1) {		
11	data_size	32	uimsbf
12	} else {		
13	data_size	16	uimsbf
14	}		
15	if (subsample_number==0&		
16	last_subsample_number>0&		
17	resources_info_flag==1) {		
18	for (i=1; i<last_subsample_number+1; i++) {		
19	subsample_i_content_type	4	bslbf
20	reserved	4	bslbf
21	if (length_extension_flag==1) {		
22	subsample_i_data_size	32	uimsbf
23	} else {		
24	subsample_i_data_size	16	uimsbf
25	}		
26	}		
27	}		
28	for (j=0; j<N; j++) {		
29	data_byte	8	bslbf
30	}		
31	}		

図 3

【図 4】

data_typeの値	意味
0000	ARIB-TTML 文書ファイル
0001	PNG形式の画像ファイル
0010	SVG形式の画像ファイル
0011	PCM (AIFF-C) 形式の音声ファイル
0100	MP3形式の音声ファイル
0101	MPEG-4 AAC形式の音声ファイル
0110	SVG形式のフォントファイル
0111	WOFF形式のフォントファイル
1000-1111	将来使用のためのリザーブ

図 4

フロントページの続き

(72)発明者 所 洋一

東京都渋谷区神南二丁目2番1号 日本放送協会放送センター内

審査官 後藤 嘉宏

(56)参考文献 特開2015-216524(JP,A)

特開2015-220517(JP,A)

米国特許出願公開第2012/0272281(US,A1)

特開2006-238147(JP,A)

青木 秀一, スーパーハイビジョンの放送に向けたメディアトランスポート技術MMT, 情報処

理学会 研究報告 オーディオビジュアル複合情報処理(AVM)2013-AVM-083, [online], 日本, 情報処理学会, 2013年11月28日, No.30

デジタル放送におけるデータ放送符号化方式と伝送方式 ARIB STD-B24, 日本, 一般社団法人 電波産業会, 2012年 9月25日, 5.6版(第一分冊), p.170-180

(58)調査した分野(Int.Cl., DB名)

H04N 21/00-21/858

H04H 20/28

H04H 60/07