

(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2006-238147

(P2006-238147A)

(43) 公開日 平成18年9月7日(2006.9.7)

(51) Int. Cl.	F I	テーマコード (参考)
HO4N 7/173 (2006.01)	HO4N 7/173 630	5C052
G11B 20/10 (2006.01)	G11B 20/10 E	5C053
G11B 27/00 (2006.01)	G11B 20/10 321Z	5C164
HO4N 5/85 (2006.01)	G11B 27/00 D	5D044
HO4N 5/93 (2006.01)	HO4N 5/85 Z	5D110
審査請求 有 請求項の数 21 O L		(全 25 頁) 最終頁に続く

(21) 出願番号 特願2005-50987 (P2005-50987)

(22) 出願日 平成17年2月25日 (2005.2.25)

(71) 出願人 000003078

株式会社東芝

東京都港区芝浦一丁目1番1号

(74) 代理人 100058479

弁理士 鈴江 武彦

(74) 代理人 100091351

弁理士 河野 哲

(74) 代理人 100088683

弁理士 中村 誠

(74) 代理人 100108855

弁理士 蔵田 昌俊

(74) 代理人 100075672

弁理士 峰 隆司

(74) 代理人 100109830

弁理士 福原 淑弘

最終頁に続く

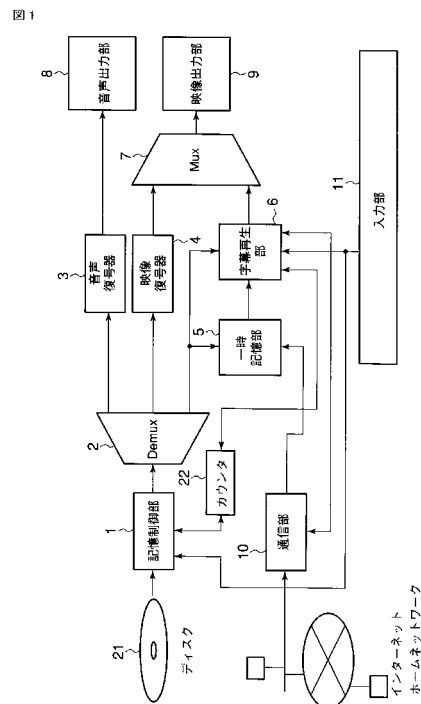
(54) 【発明の名称】 コンテンツ再生装置、字幕再生方法及びプログラム

(57) 【要約】

【課題】記録媒体上のコンテンツデータを再生すると同時に、現在再生されているコンテンツデータに対応する所望の言語の字幕を容易に再生・表示することができ、ランダムアクセスに適したコンテンツ再生装置を提供する。

【解決手段】コンテンツデータを記憶する記録媒体からコンテンツデータを再生する場合、再生されるコンテンツデータの時間区間毎に、記録媒体から、あるいはネットワークを介して、当該時間区間内に再生すべき字幕テキストを含む構造化文書記述言語で記述された分割字幕XMLデータを取得し、この取得された分割字幕XMLデータに基づき、コンテンツデータの時間区間内の各字幕テキストを再生する。

【選択図】 図1



**【特許請求の範囲】****【請求項 1】**

少なくとも、映像データ及び音声データを含むコンテンツデータと、当該コンテンツデータの再生時間を区分した各時間区間に再生すべき各字幕テキストを含む構造化文書記述言語で記述された分割字幕データを時間区間毎に記憶する記録媒体から読み出された前記コンテンツデータを再生する再生手段と、

前記再生手段で再生される前記コンテンツデータの時間区間毎に、前記記録媒体から読み出された当該時間区間に対応する分割字幕データに基づいて、当該時間区間内の各字幕テキストを再生する字幕再生手段と、

を具備したことを特徴とするコンテンツ再生装置。

10

**【請求項 2】**

前記記録媒体は、前記コンテンツデータの各時間区間と、各時間区間に対応する前記分割字幕データを示した字幕再生制御情報をさらに記憶し、

前記字幕再生手段は、前記記録媒体から読み出された前記字幕再生制御情報を基に、前記記録媒体から読み出された分割字幕データのうち、前記再生手段で再生される前記コンテンツデータの時間区間に対応する分割字幕データを用いて、当該時間区間内の各字幕テキストを再生することを特徴とする請求項 1 記載のコンテンツ再生装置。

**【請求項 3】**

前記分割字幕データに含まれる各字幕テキストは、当該字幕テキストの再生開始時を、その直前の字幕テキストの表示終了時から当該再生開始時までの時間で示す属性情報を含み、

20

前記分割字幕データは、字幕の表示領域毎に、当該分割字幕データに対応する時間区間内で最初の字幕テキストの前記再生開始時を補正するための時間補正情報を含み、

前記字幕再生手段は、字幕の各表示領域において、前記再生手段で再生される前記コンテンツデータの時間区間内で最初の字幕テキストを再生する際には、当該最初の字幕テキストの前記属性情報と前記時間補正情報とを用いて補正された再生開始時に、当該最初の字幕テキストの再生を開始し、2 番目以降の字幕テキストを再生する際には、その直前の字幕テキストの表示終了時から当該 2 番目以降の字幕テキストの前記属性情報が示す時間が経過したときに当該 2 番目以降の字幕テキストの再生を開始することを特徴とする請求項 1 記載のコンテンツ再生装置。

30

**【請求項 4】**

前記補正された再生開始時は、前記再生手段で再生される前記コンテンツデータの時間区間の開始時から前記最初の字幕テキストの再生開始時までの時間を示すことを特徴とする請求項 3 記載のコンテンツ再生装置。

**【請求項 5】**

前記記録媒体は、字幕の各表示領域及び表示形式を示すスタイル情報を含み前記分割字幕データを再生するために必要な共通字幕データをさらに記憶し、

前記字幕再生手段は、前記再生手段で再生される前記コンテンツデータの時間区間毎に、前記記録媒体から読み出された当該時間区間に対応する分割字幕データと前記記録媒体から読み出された前記共通字幕データとから、再生可能な分割字幕データを生成し、この再生可能な分割字幕データを用いて、字幕の各表示領域に当該時間区間内の各字幕テキストを再生することを特徴とする請求項 1 記載のコンテンツ再生装置。

40

**【請求項 6】**

少なくとも、映像データ及び音声データを含むコンテンツデータを記憶する記録媒体から読み出された前記コンテンツデータを再生する再生手段と、

前記再生手段で再生される前記コンテンツデータの時間区間毎に、ネットワークを介して当該時間区間内に再生すべき字幕テキストを含む構造化文書記述言語で記述された分割字幕データを取得する取得手段と、

前記取得手段で取得された分割字幕データに基づいて、前記再生手段で再生される前記コンテンツデータの時間区間内の各字幕テキストを再生する字幕再生手段と、

50

を具備したことを特徴とするコンテンツ再生装置。

【請求項 7】

前記記録媒体は、前記コンテンツデータの各時間区間と、各時間区間に対応する前記分割字幕データを示した字幕再生制御情報をさらに記憶し、

前記取得手段は、前記記録媒体から読み出された前記字幕再生制御情報を基に、前記再生手段で再生される前記コンテンツデータの時間区間に対応する分割字幕データを取得することを特徴とする請求項 6 記載のコンテンツ再生装置。

【請求項 8】

前記分割字幕データに含まれる各字幕テキストは、当該字幕テキストの再生開始時を、その直前の字幕テキストの表示終了時から当該再生開始時までの時間で示す属性情報を含み、

前記分割字幕データは、字幕の表示領域毎に、当該分割字幕データ対応する時間区間内で最初の字幕テキストの前記再生開始時を補正するための時間補正情報を含み、

前記字幕再生手段は、字幕の各表示領域において、前記再生手段で再生される前記コンテンツデータの時間区間内で最初の字幕テキストを再生する際には、当該最初の字幕テキストの前記属性情報と前記時間補正情報とを用いて補正された再生開始時に、当該最初の字幕テキストの再生を開始し、2 番目以降の字幕テキストを再生する際には、その直前の字幕テキストの表示終了時から当該 2 番目以降の字幕テキストの前記属性情報が示す時間が経過したときに当該 2 番目以降の字幕テキストの再生を開始することを特徴とする請求項 6 記載のコンテンツ再生装置。

【請求項 9】

前記補正された再生開始時は、前記再生手段で再生される前記コンテンツデータの時間区間の開始時から前記最初の字幕テキストの再生開始時までの時間を示すことを特徴とする請求項 8 記載のコンテンツ再生装置。

【請求項 10】

前記記録媒体は、字幕の各表示領域及び表示形式を示すスタイル情報を含み前記分割字幕データを再生するために必要な共通字幕データをさらに記憶し、

前記字幕再生手段は、前記再生手段で再生される前記コンテンツデータの時間区間毎に、前記取得手段で取得された当該時間区間に対応する分割字幕データと前記記録媒体から読み出された前記共通字幕データとから、再生可能な分割字幕データを生成し、この再生可能な分割字幕データを用いて、字幕の各表示領域に当該時間区間内の各字幕テキストを再生することを特徴とする請求項 6 記載のコンテンツ再生装置。

【請求項 11】

前記記録媒体に記録される前記字幕再生制御情報は、各時間区間に対応する前記分割字幕データを前記ネットワークを介して取得するための前記コンテンツデータの再生開始時からの時刻をさらに含み、

前記取得手段は、各分割字幕データを取得するための前記時刻に、当該分割字幕データを前記ネットワークを介して取得することを特徴とする請求項 7 記載のコンテンツ再生装置。

【請求項 12】

前記取得手段は、前記ネットワークを介して、字幕の各表示領域及び表示形式を示すスタイル情報を含み前記分割字幕データを再生するために必要な共通字幕データをさらに取得し、

前記字幕再生手段は、前記再生手段で再生される前記コンテンツデータの時間区間毎に、前記取得手段で取得された当該時間区間に対応する分割字幕データ及び前記共通字幕データとから、再生可能な分割字幕データを生成し、この再生可能な分割字幕データを用いて、字幕の各表示領域に当該時間区間内の各字幕データを再生することを特徴とする請求項 6 記載のコンテンツ再生装置。

【請求項 13】

前記分割字幕データは、字幕の各表示領域及び表示形式を示すスタイル情報を含み、

前記字幕再生手段は、前記再生手段で再生される前記コンテンツデータの時間区間に  
対応する分割字幕データの前記スタイル情報に基づいて、各表示領域に当該時間区  
間内の各字幕テキストを再生することを特徴とする請求項 1 または 6 記載の  
コンテンツ再生装置。

【請求項 1 4】

各時間区間に  
対応する分割字幕データが前記コンテンツデータに多重化されて前記記録  
媒体に記憶されていることを特徴とする請求項 1 記載のコンテンツ再生装置。

【請求項 1 5】

映像データ及び音声データを含むコンテンツデータと、

当該コンテンツデータの際せ時間を区分した各時間区間に再生すべき各字幕  
テキストを含む構造化文書記述言語で記述された、各時間区間毎の分割字幕  
データと、

前記コンテンツデータの各時間区間と各時間区間に  
対応する前記分割字幕データを示した字幕再生制御情報と、

字幕の各表示領域及び表示形式を示すスタイル情報を含み、前記複数の  
分割字幕データを再生するために必要な共通字幕データと、

を記憶し、

各分割字幕データに含まれる各字幕テキストは、当該字幕テキストの再生  
開始時を、その直前の字幕テキストの表示終了時から当該再生開始時  
までの時間で示す属性情報を含み、各分割字幕データは、字幕  
テキストを表示するための表示領域毎に、当該分割字幕データ  
対応する時間区間で最初の字幕テキストの前記再生開始時を補正する  
ための時間補正情報を含むことを特徴とする記録媒体から読み出され  
た前記コンテンツデータを再生する再生手段と、

前記記録媒体から読み出された前記字幕再生制御情報を  
基に、前記再生手段で再生される前記コンテンツデータの時間区間  
毎に、前記記録媒体から読み出された当該時間区間に  
対応する分割字幕データ及び前記共通字幕データに基づき、字幕の  
各表示領域に当該時間区間内の各字幕テキストを再生する字幕再生  
手段と、

を具備したことを特徴とするコンテンツ再生装置。

【請求項 1 6】

少なくとも、映像データ及び音声データを含むコンテンツデータを記憶する  
記録媒体からコンテンツデータを再生する再生手段を備えたコンテンツ再生  
装置における字幕再生方法において、

前記再生手段で再生される前記コンテンツデータの時間区間毎に、前記  
記録媒体から、あるいはネットワークを介して、当該時間区間に再生  
すべき字幕テキストを含む構造化文書記述言語で記述された分割  
字幕データを取得する第 1 のステップと、

前記第 1 のステップで取得された分割字幕データに基づき、前記再生  
手段で再生される前記コンテンツデータの時間区間内の各字幕  
テキストを再生する第 2 のステップと、

を有することを特徴とする字幕再生方法。

【請求項 1 7】

前記記録媒体は、前記コンテンツデータの各時間区間と、各時間区間に  
対応する前記分割字幕データを示した字幕再生制御情報をさらに記憶し、

前記第 1 のステップは、前記記録媒体から読み出された前記字幕再生  
制御情報を基に、前記再生手段で再生される前記コンテンツデータの  
時間区間に  
対応する分割字幕データを、前記記録媒体から、あるいはネットワーク  
を介して取得することを特徴とする請求項 1 6 記載の字幕再生方法。

【請求項 1 8】

前記分割字幕データに含まれる各字幕テキストは、当該字幕テキストの  
再生開始時を、その直前の字幕テキストの表示終了時から当該再生  
開始時までの時間で示す属性情報を含み、

前記分割字幕データは、字幕テキストを表示するための表示領域毎に、  
当該分割字幕データ対応する時間区間で最初の字幕テキストの前記  
再生開始時を補正するための時間補正情報を含み、

10

20

30

40

50

前記第2のステップは、字幕テキストを表示するための各表示領域において、前記再生手段で再生される前記コンテンツデータの時間区間で最初の字幕テキストを再生する際には、当該最初の字幕テキストの前記属性情報と前記時間補正情報とを用いて補正された再生開始時に、当該最初の字幕テキストの再生を開始し、2番目以降の字幕テキストを再生する際には、その直前の字幕テキストの表示終了時から当該2番目以降の字幕テキストの前記属性情報が示す時間が経過したときに当該2番目以降の字幕テキストの再生を開始することを特徴とする請求項16記載の字幕再生方法。

【請求項19】

前記記録媒体は、字幕の各表示領域及び表示形式を示すスタイル情報を含み前記分割字幕データを再生するために必要な共通字幕データをさらに記憶し、

10

前記第2のステップは、前記再生手段で再生される前記コンテンツデータの時間区間毎に、前記第1のステップで取得した当該時間区間に対応する分割字幕データと前記記録媒体から読み出された前記共通字幕データとから、再生可能な分割字幕データを生成し、この再生可能な分割字幕データを用いて、字幕の各表示領域に当該時間区間内の各字幕テキストを再生することを特徴とする請求項16記載の字幕再生方法。

【請求項20】

前記分割字幕データは、字幕の各表示領域及び表示形式を示すスタイル情報を含み、

前記第2のステップは、前記再生手段で再生される前記コンテンツデータの時間区間に対応する分割字幕データの前記スタイル情報に基づいて、各表示領域に当該時間区間内の各字幕テキストを再生することを特徴とする請求項16記載のコンテンツ再生方法。

20

【請求項21】

少なくとも、映像データ及び音声データを含むコンテンツデータを記憶する記録媒体からコンテンツデータを再生する再生手段を備えたコンピュータに、

前記再生手段で再生される前記コンテンツデータの時間区間毎に、前記記録媒体から、あるいはネットワークを介して、当該時間区間に再生すべき字幕テキストを含む構造化文書記述言語で記述された分割字幕データを取得する第1のステップと、

前記第1のステップで取得された分割字幕データに基づき、前記再生手段で再生される前記コンテンツデータの時間区間内の各字幕テキストを再生する第2のステップと、

を実行させるためのプログラム。

【発明の詳細な説明】

30

【技術分野】

【0001】

本発明は、記録媒体上のコンテンツデータに対応する字幕データの再生方法及びそれを用いたコンテンツ再生装置に関する。

【背景技術】

【0002】

現行のDVDのビデオ規格ではビットマップ情報の字幕をサブピクチャとして光ディスク上に記憶できる。次世代DVDビデオ規格では、これをさらに多言語化できるようにすること、あるいは、ネットワークからダウンロードして追加できるようにすることが求められている。字幕をビットマップ情報とするとサイズの問題があるため、テキスト形式で保持することが求められている。テキスト形式での保持方法には独自に規定する方法もあるが、標準的な方法でXML規格を用いた方法が期待される。

40

【0003】

一方、W3C Timed Text Working groupでは、マルチメディアの再生順序を記述するSMILを応用して、字幕情報をXML文書化するマークアップ言語であるTimed Textを策定中である(例えば、非特許文献1参照)。この限定されたサブセットが2004年11月にワーキングドラフト(DFXP)として公開されている(例えば、非特許文献2参照)。

【0004】

Timed TextはHTMLのサブセットの構文で字幕の情報とレイアウトを記述し、段落

50

ごとに表示開始時刻情報と表示継続時間を記述したものである。

【非特許文献1】 [online]、インターネット <<http://www.w3.org/AudioVideo/TT/>>

【非特許文献2】 Timed Text (TT) Authoring Format 1.0 ? Distribution Format Exchange Profile (DFXP)、 [online]、2004年11月、インターネット <<http://www.w3.org/TR/ttaf1-dfxp/>>

【発明の開示】

【発明が解決しようとする課題】

【0005】

光ディスクはヘッドを移動させるシーク時間が遅い。映像データや音声データなどのコンテンツデータにマルチプレックスして随時読み込むようにする方が、ランダムアクセスも効率よく行える。特に多言語にわたる複数の字幕ファイルを効率よく扱うには再生開始時に全ての字幕データを読み込むのではなく、再生区間に応じて分割された字幕データを扱うほうが良い。

10

【0006】

字幕情報が、Timed TextのようなXML文書である場合においても、ネットワークからダウンロードする際には、再生区間に対応してXMLデータを分割してダウンロードするほうが効率が良い。このためXML文書を複数のファイルに分割し、断片だけから再生することが求められる。

【0007】

しかし、XML文書は構文上の関係で単純には分割できないこと、さらに、Timed Text形式では、各字幕パラグラフの表示開始時刻が、その直前の字幕パラグラフの表示終了時からの時間で表されているため、分割した断片だけからは表示時刻が再現できなかった。

20

【0008】

このように、従来は、字幕データがXMLなどで記述されていたり、ネットワークからダウンロードする場合には、分割された字幕データを逐次読み込んで、記録媒体上のコンテンツデータの再生に同期した字幕の再生が容易に行えないという問題点があった。

【0009】

そこで、本発明は、字幕データを分割記憶し、記録媒体上のコンテンツデータの再生に同期した字幕の再生が容易に行え、ランダムアクセスに適した字幕再生方法及びそれを用いたコンテンツ再生装置を提供することを目的とする。

30

【課題を解決するための手段】

【0010】

映像データ及び音声データを含むコンテンツデータを記憶する記録媒体からコンテンツデータを再生する場合、再生されるコンテンツデータの時間区間毎に、上記記録媒体から、あるいはネットワークを介して、当該時間区間内に再生すべき字幕テキストを含む構造化文書記述言語で記述された分割字幕データを取得し、取得された分割字幕データに基づき、再生されるコンテンツデータの時間区間内の各字幕テキストを再生する。

【発明の効果】

【0011】

記録媒体上のコンテンツデータの再生に同期した字幕の再生が容易に行える。

40

【発明を実施するための最良の形態】

【0012】

以下、本発明の実施形態について図面を参照して説明する。

【0013】

ここでは、映像や音声データを含むコンテンツデータと字幕データなどが記録された光ディスクなどの記録媒体を再生するコンテンツ再生装置を例にとり説明する。

【0014】

本実施形態に係るコンテンツ再生装置は、記録媒体上のコンテンツデータを再生しながら、記録媒体上に記録された字幕データやネットワークからダウンロードされる字幕データを表示するようになっている。

50

## 【0015】

ここでは、各記録媒体（以下、ディスクと呼ぶ）に記録されているコンテンツデータに対しては、言語毎の字幕データが用意され、各言語の字幕データは、XML（Extensible Markup Language）やHTML（hyper text markup language）などの構造化文書記述言語で記述されている。ここでは、例えば、DFXPに従って、Timed Textで記述されたXML（Extensible Markup Language）文書であるとする。

## 【0016】

まず、Timed Textで記述された字幕データのXML文書について説明する。

## 【0017】

図2は、字幕データのXML文書の一例を示したものである。図3は、ディスクに記録された映像データと、図2のXML文書の字幕データとを図1のコンテンツ再生装置に接続された所定のディスプレイ装置で表示する場合の表示画面上の表示例を示したものである。図3において、映像上の2箇所の領域301、302に字幕を重ねて表示した例である。通常、領域301のように、ディスプレイ装置の下辺中央付近に音声を葺き替えた字幕を表示するが、さらに、映像中に表示されている他言語の文字での看板、文章、状況説明を領域302のように右端に別に表示することがある。

## 【0018】

図2のXML文書は、画面（映像）上の複数の位置に字幕を表示する場合の例である。Timed TextではHTML文書と似た構文を持ち、大きく分けて<head>要素605と<body>要素606を含む。

## 【0019】

<head>要素605には、CSS（Cascade Style Sheet）を応用したスタイル属性のルール定義を含む。この例は2つのスタイルルール定義601、602を含み、ID属性として示されている「region1」に対応する図3の領域301に表示される要素603、ID属性として示されている「region2」に対応する図3の領域302に表示される要素604に対して、スタイルをそれぞれ定義している。<head>要素には、字幕の表示領域（ここでは、領域301、302）毎に、字幕を画面上に表示するためのレイアウトに関するスタイル情報として、文字のフォントの種類、大きさ、色などの字幕の表示形式に関する情報のほかに、字幕の表示領域の大きさ及び位置、テキストの配置などの字幕の表示領域に関する情報が含まれている。

## 【0020】

図2において、スタイルルール定義601は<div>要素603が、図3の301の位置に表示されることを、スタイルルール定義602は<div>要素604が、図3の302の位置に表示されることを示している。

## 【0021】

<body>要素606は表示される字幕データとその表示タイミングが記述される。<body>要素のtimeContainer属性の値が「par」である場合には、<body>要素の子要素である<div>要素の時間が並行して進むことを示している。つまり、図2に示す例では、<div>要素603と<div>要素604の2つの字幕が並行して時間管理されて表示される。

## 【0022】

<div>要素603、604には、timeContainer属性の値として「seq」が指定されている。これは、これらの子要素である各<p>要素が順に表示されることを示している。<p>要素は、一表示単位の字幕テキストを含み、begin属性とdur属性が指定されている。begin属性は、当該要素の表示（再生）がいつ開始されるのかを示す表示開始時刻（再生開始時刻）を示し、dur属性は、当該要素が何秒間表示されるかといった表示時間を示している。ただし、各要素のbegin属性で指定する時刻は、当該要素の直前の要素の表示終了時から当該要素の表示開始時（再生開始時）までの時間、すなわち、当該要素の直前の要素の表示終了時からカウントされる相対時刻であり、ディスク上の映像や音声（コンテンツデータ）の再生の開始時からカウントされる

10

20

30

40

50

再生時刻とは異なるものである。

【0023】

図4は図2のXML文書で記述された字幕データを説明するための図である。横軸は、ディスク上の映像や音声(コンテンツデータ)の再生時刻を示し、縦軸には、図3の2つの領域301、302に対応する「region1」と「region2」のそれぞれで表示される字幕をならべて示している。図4において、各字幕テキストを矩形で囲み、矩形の位置と大きさで、当該字幕のテキスト表示(再生)開始時、表示終了時及び表示時間を示している。各字幕テキストの矩形の下にある「17-20」というような数字は、この字幕テキストが再生時刻17秒から再生時刻20秒の間に割り当てられている(表示されている)ことを示している。

10

【0024】

「region1」において、最初に表示される<p>要素は、begin="10s"、dur="6s"となっているため、再生時刻10秒から6秒間、「おはよう!」という字幕(テキスト)が割り当てられて、この間に表示される。次の<p>要素は、begin="1s"、dur="3s"であるから、直前の「おはよう!」という字幕の表示終了時の再生時刻である16秒から1秒経過した時点から3秒間、つまり、再生時刻17秒から再生時刻20秒までの間に「やあ...」という字幕が割り当てられる(表示される)。

【0025】

同様に、「region2」では、再生時刻2秒から再生時刻14秒までの間、「ワシントン州」という字幕が割り当てられ(表示され)、再生時刻38秒から再生時刻48秒までの間「号外 事故発生」という字幕が割り当てられる。

20

【0026】

(構成)

図1は、映像や音声データを含むコンテンツデータと字幕データなどが記録された光ディスクなどの記録媒体を再生するコンテンツ再生装置の構成例を示したもので、記録媒体上のコンテンツデータを再生しながら、記録媒体上に記録された字幕データやネットワークからダウンロードされる字幕データを表示するようになっている。

【0027】

図1において、コンテンツ再生装置は、ディスクが挿入されると、記憶制御部1は、当該ディスク上から各種情報(VOBU(後述)、再生制御情報など)を読み取る。分離部(Demux)2は、記憶制御部1でディスクから読み出されたVOBU(後述)から音声データ、映像データ、分割字幕XMLデータを分離し、音声データを音声復号器3へ出力し、映像データを映像復号器4へ出力し、分割字幕XMLデータを一時記憶部5へ出力する。また、記憶制御部1でディスクから読み出された字幕再生制御情報や共通字幕XMLデータなどの字幕の再生に必要な情報などは、字幕再生部6へ出力する。

30

【0028】

音声復号器3は、入力された音声データを復号し、音声出力部8へ出力する。

【0029】

映像復号器4は、入力された映像データを復号し、合成部(Mux)7へ出力する。

40

【0030】

カウンタ22は、再生時刻をカウントするもので、ディスクの再生が開始されると(記憶制御部1がディスクからコンテンツデータなどの読み取りを開始すると)、再生時刻のカウントを開始する。

【0031】

字幕再生部6は、字幕再生制御情報、共通字幕XMLデータ、一時記憶部5に記憶されている分割字幕XMLデータを用いて、字幕テキストを表示するための表示データ(字幕表示データ)を生成し、合成部(Mux)7へ出力する。なお、後述するように、ネットワーク上のサーバから、字幕再生制御情報、共通字幕XMLデータ、分割字幕XMLデータなどを取得する場合には、通信部10から当該サーバへアクセスして、所望のデータを

50



取得するようになっている。

【0032】

合成部7では、字幕表示データを映像復号器4で復号された映像データの上に合成し、映像出力部9に出力する。

【0033】

入力部11は、記憶制御部1や、字幕再生部6などに対し、ディスク21の再生開始や終了などの指示入力などを行うためのものである。

【0034】

本実施形態では、ディスクに記録されたコンテンツデータに対応する(複数の言語のうちある1つの言語の)字幕データ全体を含む1つのXML文書を複数に分割して記憶する方法と、その分割された部分文書だけから、コンテンツデータが再生される間の指定された時間区間内では同じ振る舞いをする分割字幕XMLデータを作成し、字幕を再現する方法について説明する。

10

【0035】

(データ構成とディスク上の記憶方法)

図5は、ディスク上に分割字幕XMLデータを保存するひとつの方法を示す。例えば、DVD-Video規格では映像データと音声データ、ビットマップ字幕であるサブピクチャー、ナビゲーション情報などが多重化して記憶されている。すなわち各メディアデータを時間で分割し、PACK情報として包み、それを0.4から1秒単位のVideo Object Unit (VOBU)に多重化して記憶している。このVOBUの列でVideo Objectが構成される。

20

【0036】

本実施形態でも同様の構成をとるが、VOBUのサブピクチャーを入れるSP\_PACKET 202に、サブピクチャーの代わりに、分割字幕XMLデータ201を記憶する。なお、SP\_PACKETではなく新たな別のPACKとして記憶しても構わない。

【0037】

分割字幕XMLデータにヘッダ情報(packet header)206を付加してPACK(パック)データを構成する。ヘッダ情報は、PACKデータの長さや、PACKデータ内に含まれる情報の種別を示す種別情報などを含む。分割字幕XMLデータ201は、所定の時間区間に表示する各字幕テキストとその表示タイミングと時間の補正情報を含む。図6に分割XMLデータの記述例を示す。

30

【0038】

分割字幕XMLデータ201は映像の再生の進行に応じて、逐次ディスクから字幕情報一時記憶部5に供給される。また、共通字幕XMLデータ205と、分割字幕XMLデータの再生制御情報を保持する字幕再生制御情報203とがディスク上の別の場所に記憶されている。共通字幕XMLデータ205と字幕再生制御情報203は、ディスクの再生開始時に読み込まれ、再生中に分割字幕XMLデータ201と組み合わせて字幕再生に利用される。共通字幕XMLデータ205には、字幕データの元のXML文書の<head>部分の情報やスタイルルール定義(スタイルシート)などを含む。図7に共通字幕XMLデータの一例を、図8に字幕再生制御情報の一例を示す。

40

【0039】

字幕再生制御情報203は、ディスク上のコンテンツデータの再生を制御する再生制御情報204にビデオの再生順序などと一緒に記憶される。なお、字幕再生制御情報203は再生制御情報204とは異なるファイル内に記憶されていてもよい。

【0040】

共通字幕XMLデータは、それ自体別個のファイルとして、ディスク上のコンテンツデータの再生開始時に読み込んでも良いし、各分割字幕XMLデータ201と一緒に当該分割字幕XMLデータ201と同一のPACKデータ内に記憶させるようにしてもよい。また、共通字幕XMLデータのうち、スタイルルール定義をスタイルシートとして、それ自体別個のファイルに記憶し、<head>要素内部ではスタイルシートへの参照だけとし

50

て、各分割字幕XMLデータ201と一緒に当該分割字幕XMLデータ201と同一のパックデータ内に記憶してもよい。

#### 【0041】

(字幕データの時間軸での分割方法、分割字幕XMLデータ及び字幕再生制御情報)

ここで、ディスクに記録されたコンテンツデータに対応する字幕データ全体を含む、図2に示したような1つのXML文書を、時間区間毎に分割する方法を図4及び図9を参照して説明する。ここでは、図5の時間軸を、3つの時間区間に分割する場合を例にとり説明する。

#### 【0042】

時間軸での分割はシーンの切れ目などをもとに行うことで、字幕表示していないところで区切る。すなわち、図9では、再生時刻0秒から再生時刻22秒までの時間区間401と、再生時刻22秒から再生時刻50秒までの時間区間402と、再生時刻50秒から再生時刻70秒まで時間区間403に分割した場合を示している。

#### 【0043】

図8は、図9に示したように時間軸を分割する場合の字幕再生制御情報を示したものである。図8に示す字幕再生制御情報では、再生時刻0秒から再生時刻22秒までの時間区間を「Subtitle1.xml」というファイル名の分割字幕XMLデータを再生し、再生時刻22秒から再生時刻50秒までの時間区間を「Subtitle2.xml」というファイル名の分割字幕XMLデータを再生し、再生時刻50秒から再生時刻70秒までの時間区間を「Subtitle3.xml」というファイル名の分割字幕XMLデータを再生することを示している。例えば、再生時刻37秒から再生するには、「Subtitle2.xml」の15秒目の位置から実行すればよい。

#### 【0044】

図8に示す字幕再生制御情報では、各時間区間に再生する分割字幕XMLデータのファイル名をURI (Uniform resource identifier) で参照しているが、予めディレクトリにPACKデータ内の分割字幕XMLデータへの参照を保持しても良いし、PACKデータをディレクトリの名前空間にマップするマップ情報を別に保持することでファイル名を参照できるようにしても良い。あるいは、PACKデータの識別番号などを用いて字幕再生制御情報から直接PACKデータを参照しても良い。

#### 【0045】

図6は、図9の時間区間402に対応する分割字幕XMLデータを示したもので、この分割字幕XMLデータは、元のXML文書の<body>要素のうち、当該時間区間の<p>要素のみを含む各表示領域301(「region1」)、302(「region2」)に対応する<div>要素と、(表示領域毎の)時間補正情報が含まれている。

#### 【0046】

時間補正情報は、図6の例では、図3の表示領域301、302に対応する字幕テキストの<p>要素を含む<div>要素1501、1502の「begin = "-2s"」などの元のXML文書には含まれていないbegin属性として含まれている。

#### 【0047】

図9に示したように、時間軸を複数の時間区間に分割して、元の字幕データのXML文書から各時間区間に対応する部分文書を取り出しただけでは、再生実行することはできない。例えば、「元気ないな」の字幕を含む<p>要素は、分割前はbegin = "5s"と前の字幕の表示終了時から5秒後に開始することになっていたが、分割したときは、再生時刻22秒から再生時刻50秒までの時間区間に対応する「Subtitle2.xml」というファイル名の分割字幕XMLデータの開始位置から3秒後に開始する必要がある。この時間(例えば、ここでは-2sec)を補正時間として追加して記憶する必要がある。

#### 【0048】

例えば、「region1」に着目すると、「Subtitle1.xml」という分割字幕XMLデータに対応する時間区間401は、「やあ...」という字幕の表示が終了し

た後2秒後に終了する。「Subtitle2.xml」という分割字幕XMLデータに対応する時間区間402では、最初の「元気ないな」という字幕は「やあ...」という字幕の表示が終了した後5秒後に表示されるから、「Subtitle2.xml」の分割字幕XMLデータの開始時刻を2秒前にずらせば、あるいは、「元気ないな」という字幕を時間区間402の開始時から $5 - 2 = 3$ 秒後に再生開始すれば、「元気ないな」という字幕は、「やあ...」という字幕の表示が終了した後5秒後に表示されることになる。

【0049】

このような補正時間（各表示領域において時間区間内で最初の字幕の再生開始時刻を2秒前にずらすことを意味する「-2秒」）を、「Subtitle2.xml」の「region1」に対応する<div>要素のbegin属性の値として、「begin = “ - 2 s ”」と表す。なお、「region1」の「元気ないな」という字幕の後の字幕である「これを見てくれ」は前の字幕の終了時からの相対時間で表示開始時刻を示しているため、補正する必要はない。

10

【0050】

時間区間402の「region1」に対応する<div>要素の補正時間は、時間区間402の直前の時間区間401の「region1」に対応する<div>要素の最後の子要素である<p>要素（字幕「やあ...」）の表示終了時を示す再生時刻「20秒」から、時間区間402の開始時に対応する再生時刻「22秒」を減算すること「-2秒」と求めることができる。

【0051】

より一般的には、 $n$ （ $n$ は2以上の整数）番目の時間区間のある<div>要素の補正時間は、 $n$ 番目の時間区間の直前の（ $n - 1$ ）番目の時間区間の当該<div>要素の最後の子要素である<p>要素の表示終了時を示す再生時刻から、 $n$ 番目の時間区間の開始時に対応する再生時刻を減算することで求めることができる。

20

【0052】

「region2」についても同様で、「Subtitle1.xml」という分割字幕XMLデータに対応する時間区間401は、「ワシントン州」という字幕の表示が終了した後8秒後に終了する。「Subtitle2.xml」という分割字幕XMLデータに対応する時間区間402の最初に再生される「号外 事故発生」という字幕は「ワシントン州」という字幕の表示が終了した後24秒後に表示されるから、「Subtitle2.xml」の分割字幕XMLデータの開始時刻を、8秒前にずらせば、あるいは、時間区間402の最初の「号外 事故発生」という字幕を、時間区間402の開始時から $24 - 8 = 16$ 秒後に再生開始すれば、「号外 事故発生」という字幕は、「ワシントン州」という字幕の表示が終了した後24秒後に表示されることになる。このような補正時間（時間区間内の最初の字幕の再生開始時刻を8秒前にずらすことを意味する「-8秒」）を、「Subtitle2.xml」の「region2」に対応する<div>要素のbegin属性の値として、「begin = “ - 8 s ”」と表す。

30

【0053】

各分割字幕XMLデータに、上記のように、表示領域毎に、当該分割字幕XMLデータに対応する時間区間の開始時刻あるいは当該時間区間内で最初の字幕の再生開始時を補正するための補正時間を含む時間補正情報を付加することで、元のXML文書の<body>要素内を任意の時間区間で分割しても、コンテンツデータの再生時刻に同期して、各部分文書内の各字幕テキストを表示することができる。

40

【0054】

従って、コンテンツデータの最初から最後まで字幕データ全体のXML文書を作成し、それをファイルサイズに応じて複数の部分文書に分割することも容易となる。

【0055】

（分割字幕XMLデータから時間区間単位の再生実行可能な分割字幕XMLデータの生成）

図6に示したような分割字幕XMLデータそのものは、分割前の元のXML文書にある

50

ような < head > 要素が含まれていない。すなわち、分割字幕XMLデータのみでは、当該時間区間内の字幕を再生実行することはできないのである。そこで、各時間区間に分割字幕XMLデータ内の各字幕テキストを再生する際には、図10～図12に示すように、元のXML文書にある < head > 要素（共通字幕XMLデータ）を含む再生実行可能な（再生可能な）分割字幕XMLデータとする必要がある。

【0056】

図10～図12は、時間区間単位の再生実行可能な分割字幕XMLデータの例を示したもので、それぞれ、図9の時間区間401、時間区間402、時間区間403の各時間区間内の字幕だけを再生するXMLデータである。

【0057】

図10は再生時刻0秒から、図11は再生時刻22秒から、図12は再生時刻50秒から再生することが可能なTimed Text形式のファイル形式で記述されており、それぞれの時間区間内の字幕を記述してある。

【0058】

再生実行可能な分割字幕XMLデータは、前述したように、分割前の元のXML文書の < body > 要素内の各 < div > 要素から、当該時間区間以外の < p > 要素を除いたものに、残りの、すなわち、当該時間区間内の < p > 要素のみを含む < div > 要素に、（表示領域毎の）時間補正情報を付加したものである。

【0059】

図10、図11、図12の例では、図3の表示領域301、302に対応する字幕テキストの < p > 要素を含む < div > 要素801、802、901、902、1001、1002の「begin = “ - 4 ”」などの元の字幕XMLデータには含まれていないbegin属性の値として、補正時間を示している。

【0060】

図10は、図9の再生時刻0秒から再生時刻22秒までの時間区間401に対応する再生実行可能な分割字幕XMLデータであり、再生時刻0秒から開始するものなので、表示領域毎に時間区間401の開始時刻あるいは時間区間401内の最初の < p > 要素の再生開始時刻の補正は不要であるため「begin = “ 0 s ”」あるいは時間補正情報そのものが含まれていない。

【0061】

図11は、図9の再生時刻22秒から再生時刻50秒までの時間区間402に対応する再生実行可能な分割字幕XMLデータである。「region1」に対応する < div > 要素901は、時間補正情報として「begin = “ - 2 s ”」を含み、「region1」において、時間区間402の開始時刻あるいは時間区間402内の最初の < p > 要素の再生開始時刻を2秒前に補正している。

【0062】

同様に、「region2」に対応する < div > 要素902は、時間補正情報として「begin = “ - 8 s ”」を含み、「region2」において、時間区間402の開始時刻あるいは時間区間402内の最初の < p > 要素の再生開始時刻を8秒前に補正している。

【0063】

図12は、図9の再生時刻50秒からの時間区間403に対応する再生実行可能な分割字幕XMLデータである。「region1」に対応する < div > 要素1001は、時間補正情報として「begin = “ - 6 s ”」を含み、「region1」において、時間区間403の開始時刻あるいは時間区間403内の最初の < p > 要素の再生開始時刻を6秒前に補正している。

【0064】

同様に、「region2」に対応する < div > 要素1002は、時間補正情報として「begin = “ - 2 s ”」を含み、「region2」において、時間区間403の開始時刻あるいは時間区間403内の最初の < p > 要素の再生開始時刻を2秒前に補正し

10

20

30

40

50

ている。

【0065】

図10～図12に示したような、再生実行可能な分割字幕XMLデータは、図1の字幕再生部6で、字幕再生制御情報で示した各時間区間の開始時間に、Timed Textファイルとして再生される。このとき、表示領域毎に、各時間区間の開始時間あるいは各時間区間で最初に表示される字幕テキストの再生開始時を表示領域毎の補正時間で補正することで、当該時間区間内の字幕が再生可能になる。

【0066】

(共通字幕XMLデータ及び分割字幕XMLデータを生成するための手順)

図10～図12に示したような再生実行可能な分割字幕XMLデータを、そのまま分割字幕XMLデータ201として用いてもよいが、ヘッダ情報(<head>要素)が重複している。これを共通字幕XMLデータ205として、図5に示すように、分割字幕XMLデータ201とは別個に記憶して分割しておく方がよい。ここでは、字幕データ全体を含む図2に示すようなXML文書から、図6に示すような<head>要素のみの共通字幕XMLデータと、<head要素>を含まない図8に示すような分割字幕XMLデータを生成するための手順について、図17に示すフローチャートを参照して説明する。

【0067】

まず、図13に示すように、字幕データ全体を含むXML文書1101から<head>要素1102と<body>要素1103を抜き出して(ステップS1)、<head>要素1102を共通字幕XMLデータとする(ステップS2)。

【0068】

次に、抜き出された<body>要素1103の各<div>要素から、時間区間ごとに、当該時間区間内に表示される<p>要素を抽出し(ステップS3)、各時間区間内の各<div>要素について、前述のように補正時間を算出する(ステップS4)。そして、図14に示すように、各時間区間について、各<div>要素の補正時間を示す時間補正情報(begin属性)と、当該時間区間内に表示される<p>要素とを含む分割字幕XMLデータを生成する(ステップS5)。

【0069】

図14では、<body>要素1103を、図9に示したように、3つの時間区間401～403のそれぞれに対応する3つの分割字幕XMLデータ1202～1204を生成する場合を示している。

【0070】

各分割字幕XMLデータ1202～1204には、<body>要素1103中の図3の各表示領域301、302にそれぞれ対応する<div>要素1104、1105から取り出された、同じ時間区間内に表示される字幕テキストを含む<p>要素と、当該時間区間内の各<div>要素について(表示領域毎に)算出された補正時間を示す時間補正情報(begin属性)1202a、1202b、1203a、1203b、1204a、1204bとが含まれている。

【0071】

ここで、<body>要素1103から、図9に示したように、3つの時間区間のそれぞれに対応する3つの分割字幕XMLデータ1202～1204を生成する手順についてより詳細に説明する。まず、<body>要素1103中の各<p>要素に含まれる表示開始時刻、表示時間を図4に示したように、ディスク上の映像や音声(コンテンツデータ)の再生時刻に沿って展開する。その上で、<body>要素1103から、図9に示したように、各時間区間401～403について、表示領域(「region1」、「region2」)毎に、当該時間区間内の各<p>要素をコピーする。さらに、表示領域毎に、補正時間を算出して、時間補正情報として追加することで、分割字幕XMLデータを生成する。この分割字幕XMLデータがSP\_PACKET202に記憶される。

【0072】

(分割字幕XMLデータから、再生実行可能な分割字幕XMLデータへ変換する方法)

10

20

30

40

50

前述したように、図14の分割字幕XMLデータ1202～1204だけでは、各時間区間について字幕を再生することはできない。そこで、次に、図14の分割字幕XMLデータ1202～1204を再生実行可能な分割字幕XMLデータへ変換する方法（分割字幕XMLデータから再生実行可能な分割字幕XMLデータを生成する方法）につて、図15を参照して参照して説明する。

【0073】

ここでは、図14の分割字幕XMLデータ1203を再生実行可能な分割字幕XMLデータに変換する場合について説明する。

【0074】

分割字幕XMLデータ1203の前に、共通字幕XMLデータ1102を連結させる。連結させてできたXMLフラグメントを開始タグ<tt>と終了タグ</tt>で囲うことで、図9の時間区間402に対応した再生実行可能な分割字幕XMLデータが得られる。

【0075】

（ディスク上の字幕の再生手順）

図5に示したように、ディスク上に、字幕再生制御情報203と共通字幕XMLデータ205と、映像データや音声データと多重化された分割字幕XMLデータ201が記憶されている場合の字幕再生部6の字幕の再生手順について、図18に示すフローチャートを参照して説明する。

【0076】

ディスクがコンテンツ再生装置に挿入されて、ディスクの再生が開始されると（ステップS11）、カウンタ22により、再生時刻のカウントが開始される。また、記憶制御部1は、ディスクの再生開始時、もしくは、タイトルの再生開始時に、再生制御情報とともに字幕再生制御情報と、共通字幕XMLデータをディスクから読み込み、分離部（Demux）2を経て、字幕再生部6に渡される（ステップS12）。

【0077】

光ディスク上のVOBデータは再生中、記憶制御装置1で読み込まれ、Demux2が映像データと音声データと分割字幕XMLデータを分離する。分割字幕XMLデータは一時記憶装置5に記憶される。

【0078】

字幕再生部5は、字幕再生制御情報中の各時間区間と、カウンタ22の値（ディスクの再生時刻）とを比較し（ステップS13）、現在の再生時刻に対応する、あるいは現在の再生時刻に取得すべき分割字幕XMLデータを取得する（ステップS14）。このとき、新たな分割字幕XMLデータが必要な場合には、一時記憶部5に当該分割字幕XMLデータがあるかどうかチェックする。もし、なければ、処理をせずに、分割字幕XMLデータが届くまで待つ。

【0079】

現在の再生時刻に対応する、あるいは現在の再生時刻に取得すべき分割字幕XMLデータが一時記憶部5で記憶されている場合には、それを取得し、当該分割字幕XMLデータと共通字幕データから再生実行可能な分割字幕XMLデータ生成する（ステップS15）。

【0080】

字幕再生部6は、字幕を表示するための各表示領域において、再生実行可能な分割字幕XMLデータに対応する時間区間内で最初の字幕テキスト（<p>要素）を再生する際には、当該最初の字幕テキストの再生開始時刻を、時間補正情報を用いて補正する（ステップS16）。すなわち、当該時間区間の開始時から当該最初の字幕テキストの再生開始時までの時間を算出する。例えば、図9に示すように、時間区間402に対応する分割字幕XMLデータ内の各字幕テキストを再生する場合、「Region1」において、時間区間402で最初の字幕テキスト「元気ないな」を再生するとき、当該字幕テキストの<p>要素では、その属性情報として、begin="5s"が含まれており、また、当該分

割字幕XMLデータの「Region1」の時間補正情報として、「-2s」が含まれているから、時間区間402で最初の字幕テキスト「元気ないな」の再生開始時として、時間区間402の開始時から $5 - 2 = 3$ 秒後と算出する。また、「Region2」において、時間区間402で最初の字幕テキスト「号外 事故発生」を再生するとき、当該字幕テキストの<p>要素では、その属性情報として、begin="24s"が含まれており、また、当該分割字幕データの「Region2」の時間補正情報として、「-8s」が含まれているから、時間区間402で最初の字幕テキスト「号外 事故発生」の再生開始時として、時間区間402の開始時から $24 - 8 = 16$ 秒後と算出する。

**【0081】**

字幕再生部6は、再生実行可能な分割字幕XMLデータを用いて、当該再生実行可能な分割字幕XMLデータに対応する時間区間内の各字幕テキストを再生する(ステップS17)。まず、字幕テキストを表示するための各表示領域において、当該時間区間内で最初の字幕テキストを再生する際には、当該時間区間の開始時から上記ステップS16で求めた時刻に当該最初の字幕テキストの再生を開始し、当該最初の字幕テキストを表示するための字幕表示データを生成し、合成部(Mux)7へ出力する。また、2番目以降の字幕テキストを再生する際には、その直前の字幕テキストの表示終了時から当該2番目以降の字幕テキストの<p>要素のbegin属性として含まれている時間が経過したときに当該当該2番目以降の字幕テキストの再生を開始し、当該2番目以降の字幕テキストを表示するための字幕表示データを生成し、合成部(Mux)7へ出力する。

**【0082】**

各字幕テキストを表示するための字幕表示データは、コンテンツ(映像)の表示される表示画面上で、当該再生実行可能な分割字幕データに含まれている<head>要素により指定されている字幕の表示領域内(表示位置)に、当該<head>要素により指定されている表示形式(文字のフォントの種類、大きさ、色など)で、当該字幕テキストを表示するためのものである。

**【0083】**

MUX7では、入力された字幕表示データを、Demux2で分離され、映像復号器4で復号された映像データの上に合成し、映像出力部9へ出力する。映像出力部9から出力された字幕表示データの合成された映像データは、表示装置に送られて表示される。

**【0084】**

字幕再生部6での再生実行可能な分割字幕XMLデータの実行処理は、字幕再生制御情報に記述された当該分割字幕XMLデータの時間区間から、全ての<p>要素の再生表示が終了するまで行い、当該時間区間内の最後の<p>要素の表示時間が終了した場合には、次の新たな分割字幕XMLデータを一時記憶部5に届くのを待つ。

**【0085】**

上記ステップS13~ステップS17の処理は、ディスクの再生の終了指示が入力されるまで繰り返される(ステップS18)。

**【0086】**

(分割字幕XMLデータをネットワークからダウンロードして取得する場合の字幕の再生手順)

以上の説明は、字幕データがディスク上に記憶されている場合であったが、次に、字幕データがネットワーク上のサーバにあり、そこから逐次供給される場合について、図18を参照して説明する。この場合は、字幕再生制御情報と共通字幕XMLデータと複数の分割字幕XMLデータの全て、あるいは、そのうちの一部分がサーバ上に存在し、必要に応じて、ディスク再生中にこれらのダウンロードをサーバへ要求して取得するための処理を行う以外は、前述のディスク上に全ての字幕データがあるときと同様に、字幕の再生を行う。ディスク上に、字幕再生制御情報と共通字幕XMLデータと複数の分割字幕XMLデータの全てが記憶されている場合と異なる点は、図18のステップS12と、ステップS14の処理動作である。

**【0087】**

10

20

30

40

50

字幕再生制御情報、共通字幕XMLデータ及び複数の分割字幕XMLデータの全てがネットワーク上のサーバに保持されていてもよいし、あるいは、これらのうちの一部がネットワーク上のサーバに保持され、残りがディスク上に記憶されていてもよい。

【0088】

また、接続先のサーバを示すURL (Uniform Resource Locator) などのアドレス情報は、字幕再生制御情報、共通字幕XMLデータ (<head>要素)、各分割字幕XMLデータ、各分割字幕XMLデータに<head>要素が含まれている場合には、この<head>要素内などのうちの少なくとも1つに含まれている。あるいは、ディスク上の予め定められた特定領域内に予め記憶されていてもよい。さらに、この特定領域内に記憶されているアドレス情報に対応するサーバから、字幕再生制御情報、共通字幕XMLデータ及び複数の分割字幕XMLデータの全て、あるいは、そのうちの一部を取得できるサーバのアドレス情報が通知されてもよい。コンテンツ再生装置は、例えばディスクから読み出されたあるいはネットワークから取得したアドレス情報を用いて、サーバへ接続し、所望の情報を取得する。

10

【0089】

コンテンツ再生装置の字幕再生部6は、ディスクの再生が開始されると(ステップS11)、字幕再生制御情報及び共通字幕XMLデータがディスク上の記憶されている場合には、前述同様、ディスクから読み出された字幕再生制御情報及び共通字幕XMLデータを取得する(ステップS12)。一方、字幕再生制御情報及び共通字幕XMLデータが、ディスクにではなく、ネットワーク上のサーバに存在する場合には、通信部10を用いて、当該サーバから字幕再生制御情報及び共通字幕XMLデータを取得する(ステップS12)。

20

【0090】

字幕再生部5は、字幕再生制御情報中の各時間区間と、カウンタ22の値(ディスクの再生時刻)とを比較し(ステップS13)、現在の再生時刻に対応する、あるいは現在の再生時刻に取得すべき分割字幕XMLデータを取得する(ステップS14)。このとき、現在の再生時刻に対応する、あるいは現在の再生時刻に取得すべき分割字幕XMLデータが、サーバ上に存在する場合には、通信部10を用いて、当該サーバから分割字幕XMLデータを取得する(ステップS14)。

【0091】

すなわち、ステップS14では、通信部10は、当該サーバから分割字幕XMLデータを受け取ると、それを一時記憶部5に記憶し、字幕再生部6に通知する。字幕再生部6は、現在の再生時刻に対応する、あるいは現在の再生時刻に取得すべき分割字幕XMLデータを、一時記憶部5から読み出すと、当該分割字幕XMLデータと共通字幕データから再生実行可能な分割字幕XMLデータ生成する(ステップS15)。

30

【0092】

そして、字幕再生部6は、字幕テキストを表示するための各表示領域において、再生実行可能な分割字幕XMLデータに対応する時間区間内で最初の字幕テキストの再生開始時刻を時間補正情報を用いて補正した後(ステップS16)、再生実行可能な分割字幕XMLデータを用いて、当該再生実行可能な分割字幕XMLデータに対応する時間区間内の各字幕テキストを前述同様にして再生する(ステップS17)。すなわち、各字幕テキストの再生開始時に当該字幕テキストの再生を開始し、各字幕テキストを表示するための字幕表示データを生成し、合成部(Mux)7へ出力する。

40

【0093】

このように、図18のステップS12とステップS14以外の処理動作は、ディスク上に、字幕再生制御情報と共通字幕XMLデータと複数の分割字幕XMLデータの全てが記憶されている場合と同様である。

【0094】

なお、ある時間区間に対応する分割字幕XMLデータをサーバへ要求して、当該サーバからダウンロードしてもらう場合、当該時間区間に入る前の所定の再生時刻に、前もって

50



サーバへ当該時間区間に対応する分割字幕XMLデータをサーバへ要求する方が望ましい。この場合、各時間区間に対応する分割字幕XMLデータをサーバへ要求する、あるいは、サーバから取得する時刻（プリフェッチ時刻）を字幕再生制御情報や再生制御情報などに記述しておく。

【0095】

図16は、各時間区間（時間区間の開始時刻と終了時刻）と当該時間区間に対応する分割字幕XMLデータのファイルを示す字幕再生制御情報に、さらに、各時間区間について、当該時間区間に対応する分割字幕XMLデータをサーバに要求する、あるいは、サーバから取得する時刻であるプリフェッチ時刻が含まれている場合を示している。図16では、各時間区間に対応する<Clip>要素のprefetch属性により、プリフェッチ時刻を示している。

10

【0096】

図9の時間区間402に対応する図16の<Clip>要素1601では、再生時刻が「10秒」のときから、当該時間区間に対応する分割字幕XMLデータをサーバへ要求すること、あるいはサーバから取得することを示している。すなわち、この分割字幕XMLデータは、再生時刻22秒から再生時刻50秒までの時間区間に対応するので、<Clip>要素1601は、この時間区間の開始時から10秒前の再生時刻10秒から50秒の間に、サーバへ当該分割字幕XMLデータの要求がなされておらず、当該分割字幕XMLデータが一時記憶部5に記憶されていない場合に、当該分割字幕XMLデータのダウンロードをサーバへ要求すること、あるいは、サーバから取得することを示している。従って、図18のステップS14において、字幕再生部6は、再生時刻10秒のときに、当該分割字幕XMLデータが一時記憶部5に記憶されていない場合には、当該分割字幕XMLデータのダウンロードをサーバへ要求して、当該分割字幕XMLデータを取得する。

20

【0097】

同様に、図9の時間区間403に対応する<Clip>要素1602では、再生時刻40秒のときから、当該時間区間に対応する分割字幕XMLデータをサーバへ要求することを示している。すなわち、この分割字幕XMLデータは、再生時刻50秒から再生時刻70秒までの時間区間に対応するので、<Clip>要素1602は、この時間区間の開始時から10秒前の再生時刻40秒から70秒の間に、サーバへ当該分割字幕XMLデータの要求がなされておらず、当該分割字幕XMLデータが一時記憶部5に記憶されていない場合に、当該分割字幕XMLデータのダウンロードをサーバへ要求すること、あるいはサーバから取得することを示している。

30

【0098】

従って、図18のステップS14において、字幕再生部6は、再生時刻40秒のときに、当該分割字幕XMLデータが一時記憶部5に記憶されていない場合には、当該分割字幕XMLデータのダウンロードをサーバへ要求して、当該分割字幕XMLデータを取得する。

【0099】

このように、各時間区間の開始時よりも所定時間前の時刻に、当該時間区間に対応する分割字幕XMLデータをサーバへ要求して取得する（プリフェッチする）ことにより、各時間区間に、当該時間区間の字幕が、通信の混み具合により大きく変動する通信時間などの遅延時間に影響されずに表示されることが期待できる。

40

【0100】

以上説明したように、上記実施形態によれば、映像データ及び音声データを含むコンテンツデータを記憶する記録媒体からコンテンツデータを（音声復号器3、映像復号器4、などを含む再生手段で）再生する場合、再生されるコンテンツデータの時間区間毎に、記録媒体から、あるいはネットワークを介して、当該時間区間内に再生すべき字幕テキストを含む分割字幕XMLデータを取得し、この取得された分割字幕XMLデータを用いて、コンテンツデータの時間区間内の各字幕テキストを再生する。

【0101】

50

また、記録媒体は、コンテンツデータの各時間区間と、各時間区間に対応する分割字幕XMLデータを示した字幕再生制御情報をさらに記憶し、字幕再生部6は、記録媒体から読み出された上記字幕再生制御情報を基に、再生されるコンテンツデータの時間区間に対応する分割字幕XMLデータを、記録媒体から、あるいはネットワークを介して取得する。

#### 【0102】

記録媒体上のコンテンツデータの再生を開始する前に、入力部11から言語の種類（日本語、英語など）が入力され、記録媒体から、あるいはネットワークを介して、分割字幕XMLデータを取得する場合、図18のステップS14（及びステップS12）では、この入力された言語の分割字幕XMLデータ（及び字幕再生制御情報）を取得する。

10

#### 【0103】

また、字幕再生部6は、コンテンツデータの時間区間内の各字幕テキストを再生する際、記録媒体から、あるいはネットワークを介して取得された、分割字幕XMLデータを再生するために必要な共通字幕XMLデータと、当該時間区間に対応する分割字幕XMLデータとから、再生可能な分割字幕XMLデータを生成し、この再生可能な分割字幕XMLデータを用いて、当該時間区間内の各字幕テキストを再生する。

#### 【0104】

記録媒体から、あるいはネットワークを介して取得される分割字幕XMLデータは、例えば、図6に示したように、コンテンツデータの再生時間を区分けした時間区間毎に、当該時間区間内に再生すべき各字幕テキスト（<p>要素）を含むが、さらに、図10～図12に示したように、字幕の表示領域や表示形式に関する情報を含む<head>要素が含まれている場合もある。前者の場合の字幕の再生手順は図18を参照して既に述べたように、ステップS15において、共通字幕XMLデータを用いて再生実行可能な分割字幕XMLデータを生成する必要があるが、後者の場合には、分割字幕XMLデータが、そのまま再生実行可能な分割字幕XMLデータであるので、ステップS15はスキップされる。それ以外の字幕の再生手順は、前者の場合と同様である。

20

#### 【0105】

分割字幕XMLデータに含まれる各字幕テキストは、当該字幕テキストの再生開始時を、その直前の字幕テキストの表示終了時から当該再生開始時までの時間で示す属性情報を含み、分割字幕XMLデータは、字幕を表示するための表示領域毎に、当該分割字幕XMLデータ対応する時間区間内で最初の字幕テキストの再生開始時を補正するための時間補正情報を含み、字幕テキストを表示するための各表示領域において、再生されるコンテンツデータの時間区間内で最初の字幕テキストを再生する際には、字幕再生部6は、当該最初の字幕テキストの属性情報と時間補正情報とを用いて補正された再生開始時に、当該最初の字幕テキストの再生を開始し、2番目以降の字幕テキストを再生する際には、その直前の字幕テキストの表示終了時から当該2番目以降の字幕テキストの属性情報が示す時間が経過したときに当該2番目以降の字幕テキストの再生を開始する。

30

#### 【0106】

このような構成により、光ディスクなどの記録媒体上のコンテンツデータに対し、（Ti med TextのようなXMLで記述された）多くの言語の字幕データが用意されている場合に、記録媒体上のコンテンツデータを再生すると同時に、現在再生されているコンテンツデータに対応する所望の言語の字幕を容易に再生・表示することができる。また、現時点のディスク上の再生位置から別の再生位置へジャンプしてコンテンツデータの再生を行うような、ランダムアクセスにも適用することができる。

40

#### 【0107】

また、字幕データをネットワークからダウンロードする場合にも、記録媒体上のコンテンツデータの再生に同期した字幕の再生が容易に行える。

#### 【0108】

本発明の実施の形態に記載した本発明の手法は、コンピュータに実行させることのできるプログラムとして、磁気ディスク（フレキシブルディスク、ハードディスクなど）、光

50

ディスク（CD-ROM、DVDなど）、半導体メモリなどの記録媒体に格納して頒布することもできる。

【0109】

例えば、一時記憶部5などの記憶手段や通信部10などを備えたコンピュータが、映像・音声データ及びプログラムを含むコンテンツデータを記憶する記録媒体に記憶されたコンテンツデータを再生する再生手段（音声復号器3や映像復号器4など）、字幕再生部6の機能を実現させるためのプログラムを実行することにより、図1のコンテンツ再生装置が実現可能である。

【0110】

また、上記実施形態では、字幕データがXMLで記述されている場合について説明したが、この場合に限らず、他の構造化文書記述言語で記述されていても上記同様である。 10

【0111】

なお、本発明は上記実施形態そのままに限定されるものではなく、実施段階ではその要旨を逸脱しない範囲で構成要素を変形して具体化できる。また、上記実施形態に開示されている複数の構成要素の適宜な組み合わせにより、種々の発明を形成できる。例えば、実施形態に示される全構成要素から幾つかの構成要素を削除してもよい。さらに、異なる実施形態にわたる構成要素を適宜組み合わせてもよい。

【図面の簡単な説明】

【0112】

【図1】コンテンツ再生装置の構成例を示した図。 20

【図2】字幕データ全体を含むXML文書の一例を示した図。

【図3】コンテンツ再生装置に設けられたディスプレイ装置の表示画面上での字幕の表示例を示した図。

【図4】図2のXMLで記述された字幕データを説明するための図。

【図5】ディスク上のコンテンツデータと字幕データの記憶例を示した図。

【図6】分割字幕XMLデータの一例を示した図。

【図7】共通字幕XMLデータの一例を示した図。

【図8】字幕再生制御情報の一例を示した図。

【図9】コンテンツデータの各時間区間と、各時間区間内に再生される字幕テキストを再生するための当該時間区間に対応する分割字幕XMLデータについて説明するための図。 30

【図10】再生可能な分割字幕XMLデータの一例を示した図。

【図11】再生可能な分割字幕XMLデータの他の例を示した図。

【図12】再生可能な分割字幕XMLデータのさらに他の例を示した図。

【図13】字幕データ全体を含むXML文書と、当該XML文書から取り出された<head>要素のみの共通字幕XMLデータと、<head要素>を含まない部分を示した図。

【図14】字幕データ全体を含むXML文書から取り出された<head要素>を含まない部分から、分割字幕XMLデータを生成するための手順を説明するための図。

【図15】分割字幕XMLデータから再生可能な分割字幕XMLデータを生成する手順を説明するための図。 40

【図16】分割字幕XMLデータ毎のプリフェッチ時刻を含む字幕再生制御情報の一例を示した図。

【図17】字幕データ全体を含むXML文書から、<head>要素のみの共通字幕XMLデータと、<head要素>を含まない分割字幕XMLデータを生成するための手順を説明するためのフローチャート。

【図18】字幕再生部の字幕の再生処理動作を説明するためのフローチャート。

【符号の説明】

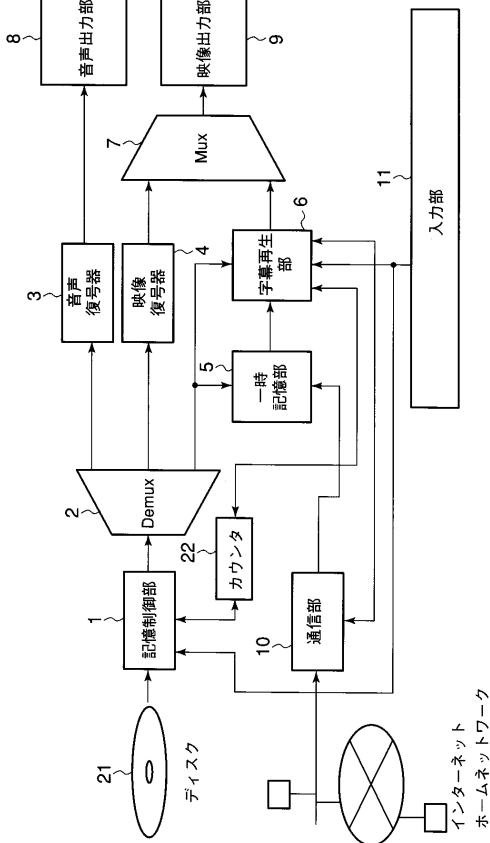
【0113】

1...記憶制御部、2...分離部(Demux)、3...音声復号器、4...映像復号器、5...一時記憶部、6...字幕再生部、7...合成部(Mux)、8...音声出力部、9...映像出力部 50

、 1 0 ... 通信部、 1 1 ... 入力部、 2 1 ... ディスク、 2 2 ... カウンタ。

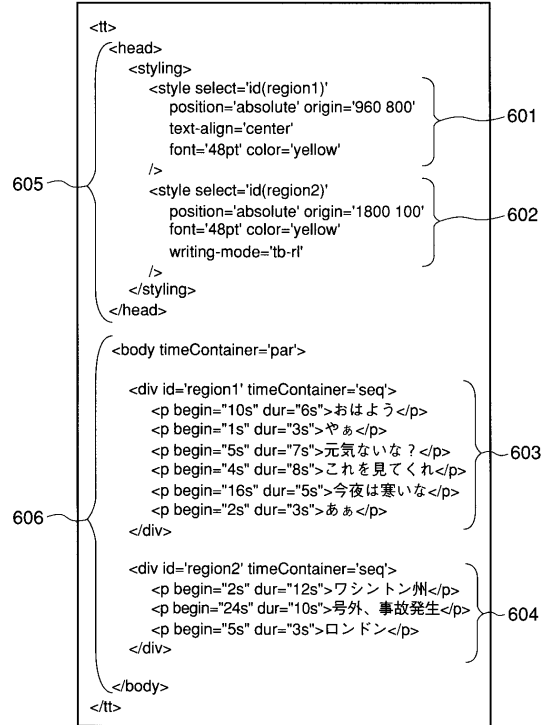
【 図 1 】

図 1



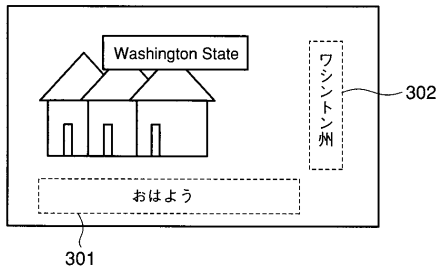
【 図 2 】

図 2



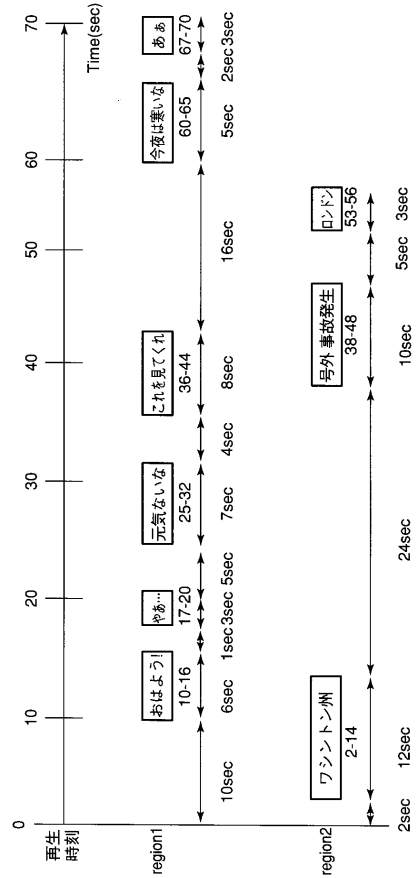
【 図 3 】

図 3



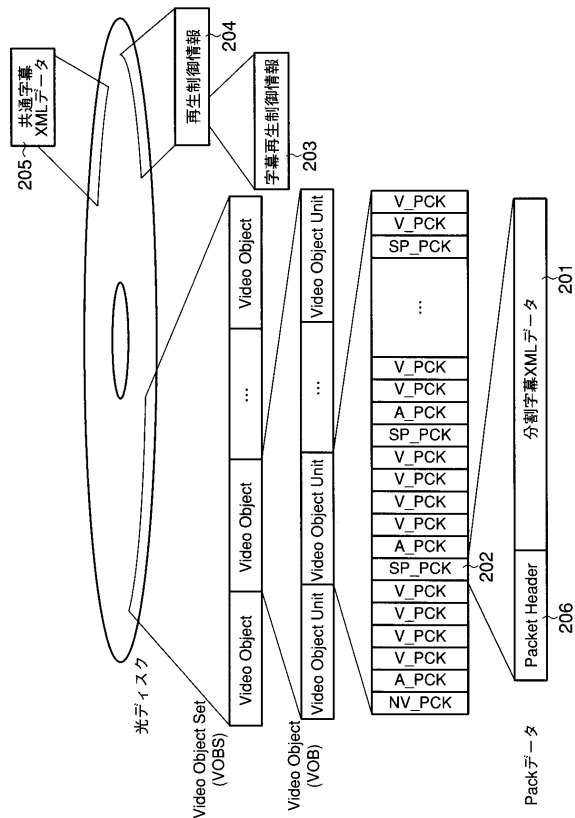
【 図 4 】

図 4



【 図 5 】

図 5



【 図 6 】

図 6

```

分割字幕XMLデータ
Subtitle2.xml
<-body timeContainer='par'>
  <div id='region1' timeContainer='seq' begin='2s'>
    <p begin='5s' dur='7s'>元気ないな?</p>
    <p begin='4s' dur='8s'>これを見てくれ</p>
  </div>
  <div id='region2' timeContainer='seq' begin='8s'>
    <p begin='24s' dur='10s'>号外、事故発生</p>
  </div>
</body>
  
```

Labels 1501 and 1502 point to the first and second regions of the XML data respectively.

【 図 7 】

図 7

```

共通字幕XMLデータ
<head>
  <styling>
    <style select='id(region1)'
      position='absolute' origin='960 800'
      text-align='center'
      font='48pt' color='yellow'
    />
    <style select='id(region2)'
      position='absolute' origin='1800 100'
      font='48pt' color='yellow'
      writing-mode='tb-rl'
    />
  </styling>
</head>
  
```

【 図 8 】

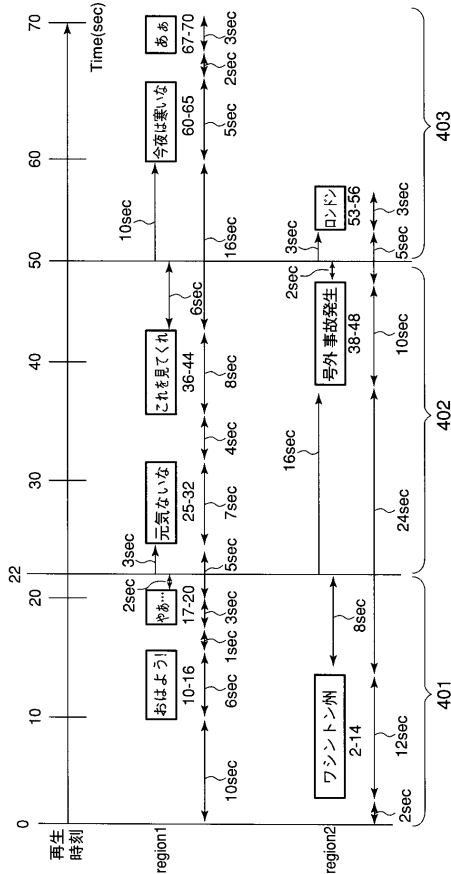
図 8

```

字幕再生制御情報
<Subtitle>
  <Clip from='0s' to='22s' src='Subtitle1.xml' />
  <Clip from='22s' to='50s' src='Subtitle2.xml' />
  <Clip from='50s' to='70s' src='Subtitle3.xml' />
</Subtitle>
  
```

【 図 9 】

図 9



【 図 10 】

図 10

時間区間401に対応する再生実行可能な分割字幕XMLデータ

```

<tt>
  <head>
    <styling>
      <style select='id(region1)'
        position='absolute' origin='960 800'
        text-align='center'
        font='48pt' color='yellow'
      />
      <style select='id(region2)'
        position='absolute' origin='1800 100'
        font='48pt' color='yellow'
        writing-mode='tb-rl'
      />
    </styling>
  </head>

  <body timeContainer='par'>

    <div id='region1' timeContainer='seq'>
      <p begin="10s" dur="6s">おはよう</p>
      <p begin="1s" dur="3s">やあ</p>
    </div>
    <div id='region2' timeContainer='seq'>
      <p begin="2s" dur="12s">ワシントン州</p>
    </div>
  </body>
</tt>

```

801  
802

【 図 11 】

図 11

時間区間402に対応する再生実行可能な分割字幕XMLデータ

```

<tt>
  <head>
    <styling>
      <style select='id(region1)'
        position='absolute' origin='960 800'
        text-align='center'
        font='48pt' color='yellow'
      />
      <style select='id(region2)'
        position='absolute' origin='1800 100'
        font='48pt' color='yellow'
        writing-mode='tb-rl'
      />
    </styling>
  </head>

  <body timeContainer='par'>

    <div id='region1' timeContainer='seq' begin='-2s'>
      <p begin="5s" dur="7s">元気ないな？</p>
      <p begin="4s" dur="8s">これを覚えてくれ</p>
    </div>
    <div id='region2' timeContainer='seq' begin='-8s'>
      <p begin="24s" dur="10s">号外、事故発生</p>
    </div>
  </body>
</tt>

```

901  
902

【 図 12 】

図 12

時間区間403に対応する再生実行可能な分割字幕XMLデータ

```

<tt>
  <head>
    <styling>
      <style select='id(region1)'
        position='absolute' origin='960 800'
        text-align='center'
        font='48pt' color='yellow'
      />
      <style select='id(region2)'
        position='absolute' origin='1800 100'
        font='48pt' color='yellow'
        writing-mode='tb-rl'
      />
    </styling>
  </head>

  <body timeContainer='par'>

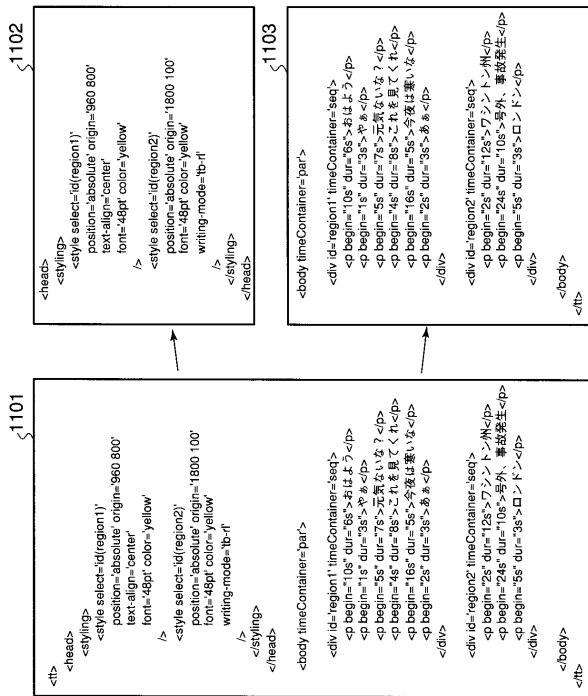
    <div id='region1' timeContainer='seq' begin='-6s'>
      <p begin="16s" dur="5s">今夜は寒いな</p>
      <p begin="2s" dur="3s">ああ</p>
    </div>
    <div id='region2' timeContainer='seq' begin='-2s'>
      <p begin="5s" dur="3s">ロンドン</p>
    </div>
  </body>
</tt>

```

1001  
1002

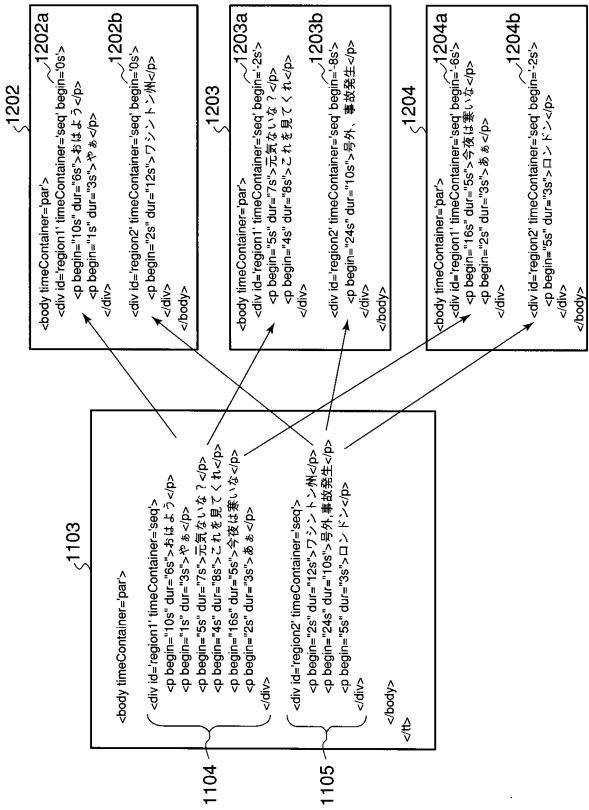
【 図 1 3 】

図 13



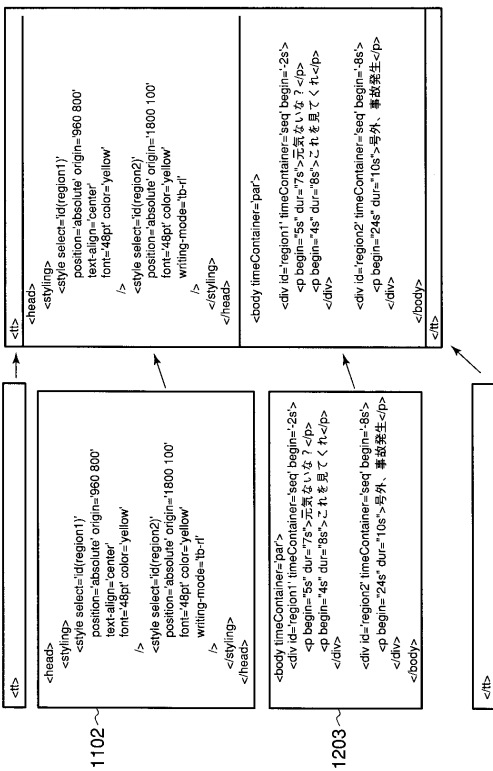
【 図 1 4 】

図 14



【 図 1 5 】

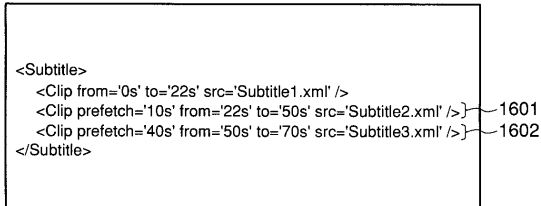
図 15



【 図 1 6 】

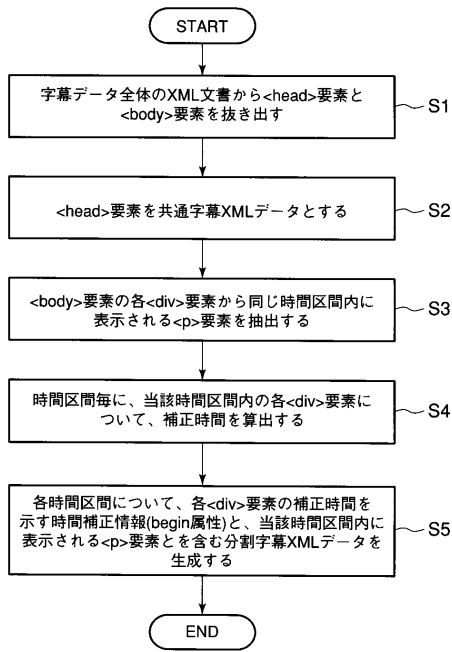
図 16

字幕再生制御情報



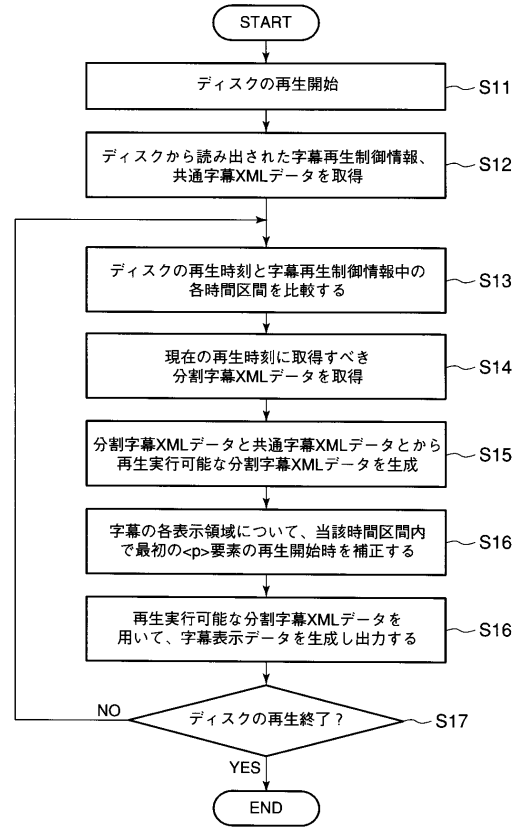
【 図 17 】

図 17



【 図 18 】

図 18





## フロントページの続き

(51) Int.Cl.	F I	テーマコード(参考)
	H 0 4 N 5/93	Z
	H 0 4 N 5/93	E

(74)代理人 100084618

弁理士 村松 貞男

(74)代理人 100092196

弁理士 橋本 良郎

(72)発明者 外山 春彦

神奈川県川崎市幸区小向東芝町1番地 株式会社東芝研究開発センター内

(72)発明者 上林 達

神奈川県川崎市幸区小向東芝町1番地 株式会社東芝研究開発センター内

(72)発明者 三村 英紀

神奈川県横浜市磯子区新杉田町8番地 株式会社東芝横浜事業所内

Fターム(参考) 5C052 AA04 AB03 AB04 DD10

5C053 FA24 GB06 GB12 JA16 LA14

5C164 MA06S MC07S SA21S UA03S UA22S UA53S UD01P UD12P

5D044 AB09 BC02 CC04 DE18 EF05 FG09 FG18 HL11

5D110 AA14 AA26 AA28 BB01 DA17 DB02