

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B1)

(11) 特許番号

特許第5860575号  
(P5860575)

(45) 発行日 平成28年2月16日(2016.2.16)

(24) 登録日 平成27年12月25日(2015.12.25)

(51) Int.Cl. F I  
H O 4 N 5/91 (2006.01) H O 4 N 5/91 C

請求項の数 6 (全 15 頁)

(21) 出願番号	特願2015-87381 (P2015-87381)	(73) 特許権者	512290388 株式会社シフトワン 東京都千代田区神田神保町3-27-8
(22) 出願日	平成27年4月22日(2015.4.22)	(74) 代理人	100158883 弁理士 甲斐 哲平
(62) 分割の表示	特願2014-223700 (P2014-223700) の分割	(72) 発明者	黒岩 司 東京都千代田区神田神保町3-7-28 株式会社シフトワン内
原出願日	平成26年10月31日(2014.10.31)	(72) 発明者	渡邊 賢悟 東京都武蔵野市境1-22-9レジテンス 武蔵境301 渡辺電気株式会社内
審査請求日	平成27年7月9日(2015.7.9)	審査官	松元 伸次
(31) 優先権主張番号	特願2014-194496 (P2014-194496)		
(32) 優先日	平成26年9月24日(2014.9.24)		
(33) 優先権主張国	日本国(JP)		
早期審査対象出願			

最終頁に続く

(54) 【発明の名称】 音声録音プログラム、音声録音端末装置、及び音声録音システム

(57) 【特許請求の範囲】

【請求項1】

コマ送りされる画像の連続表示を含む「映像」を確認しながら、音声を録音する機能をコンピュータに実行させるプログラムであって、

前記映像及びタイムラインを有する「コンテンツ」を読み出すコンテンツ読出し処理を、前記コンピュータに実行させる機能を備えるとともに、

前記タイムラインには、2以上の声役が設定されるとともに、該声役ごとに前記映像中に割り当てられる録音時間帯が設定され、

2以上の前記声役のうち、録音対象とする主声役を選択する主声役選択処理と、

前記映像を表示する映像表示処理と、

2以上の前記声役のうち前記主声役を除くそれぞれの「従声役」に対して記憶された2以上の録音ユーザによる音声のうち、所望の録音ユーザの音声を選択する従音声選択処理と、

前記主声役の録音時間帯内に録音された音声を、当該録音時間帯と関連付けて記憶する音声記憶処理と、を前記コンピュータに実行させる機能を、備えたことを特徴とする音声録音プログラム。

【請求項2】

前記映像とともに音声を出力する動画再生処理を、前記コンピュータに実行させる機能をさらに備え、

前記動画再生処理は、前記主声役に対応する前記録音時間帯では前記音声記憶処理で録

10

20

音された音声を出力し、前記従声役に対応する前記録音時間帯では前記従音声選択処理で選択された種類の音声を出力する、ことを特徴とする請求項 1 記載の音声録音プログラム。

【請求項 3】

録音する際、前記従声役に対して設定された録音時間帯で、前記従音声選択処理で選択された種類の音声を出力する従音声出力処理を、前記コンピュータに実行させる機能をさらに備え、

前記従音声出力処理によって前記従音声が出力されることで、主声役の録音タイミングを計りやすくした、ことを特徴とする請求項 1 又は請求項 2 記載の音声録音プログラム。

【請求項 4】

前記タイムラインに基づいて、前記主声役の録音時間帯になると録音速度計を表示する録音支援処理を、前記コンピュータに実行させる機能を、さらに備え、

前記録音速度計は、前記主声役の録音時間帯における進行状況及びノ又は残り時間を動的に示すことで、音声の録音を支援する、ことを特徴とする請求項 1 又は請求項 2 記載の音声録音プログラム。

【請求項 5】

コマ送りされる画像の連続表示を含む「映像」を確認しながら、音声録音する端末装置であって、

前記映像及びタイムラインを、有する「コンテンツ」を読み出すコンテンツ読出し手段を、備えるとともに、

前記タイムラインには、2 以上の声役が設定されるとともに、該声役ごとに前記映像中に割り当てられる録音時間帯が設定され、

2 以上の前記声役のうち、録音対象とする主声役を選択する主声役選択手段と、

2 以上の前記声役のうち前記主声役を除くそれぞれの「従声役」に対して記憶された 2 以上の録音ユーザによる音声のうち、所望の録音ユーザの音声を選択する従音声選択手段と、

前記映像を表示する映像表示手段と、

音声を録音する音声録音手段と、を備えたことを特徴とする音声録音端末装置。

【請求項 6】

コマ送りされる画像の連続表示を含む「映像」及びタイムラインを、有する「コンテンツ」を入力するコンテンツ入力手段と、

入力された前記コンテンツを記憶するコンテンツ記憶手段と、

前記コンテンツが有する前記映像を確認しながら、音声を録音する音声録音端末装置と、

前記音声録音端末装置で録音された音声を、記憶する音声記憶手段と、を備え、

前記タイムラインには、2 以上の声役が設定されるとともに、該声役ごとに前記映像中に割り当てられる録音時間帯が設定され、

前記音声録音端末装置は、

前記コンテンツ記憶手段から前記コンテンツを読み出すコンテンツ読出し手段と、

2 以上の前記声役のうち、録音対象とする主声役を選択する主声役選択手段と、

2 以上の前記声役のうち前記主声役を除くそれぞれの「従声役」に対して記憶された 2 以上の録音ユーザによる音声のうち、所望の録音ユーザの音声を選択する従音声選択手段と、

前記映像を表示する映像表示手段と、

音声を録音する音声録音手段と、を具備し、

前記音声記憶手段は、録音時間帯内に録音された音声を、当該録音時間帯と関連付けて記憶する、ことを特徴とする音声録音システム。

【発明の詳細な説明】

【技術分野】

【0001】

10

20

30

40

50

本願発明は、映像に合わせて音声を録音する技術に関するものであり、より具体的には、設定された複数の声役のうち所望の声役を選んで音声録音することのできる音声録音プログラム、音声録音端末装置、及び音声録音システムに関するものである。

【背景技術】

【0002】

従来、動画といえば映画やテレビが主流であったが、近年では情報技術の飛躍的進歩に伴い、コンピュータを用いて表示する動画も多く利用されている。コンピュータ上で動作する「動画ファイル」はコンテナとも呼ばれ、一般的に「映像データ」と「音声データ」で構成される。そして、動画再生用のソフトウェアを使って動画ファイルを再生するわけである。動画ファイルを再生する端末機器としては、パーソナルコンピュータ（パソコン）をはじめ、タブレット型端末やスマートフォンなど多種多様なものが利用されている。

10

【0003】

映像は、静止画とは異なり人や物の動きを連続的に表現するものであり、実際の動作等を写し取った実写や、アニメーションが代表例として挙げられる。この映像は、多数の静止画を高速で切り替えることで動きを表現するのが主流であるが、ここでは、複数の静止画を断続的に切り替える、いわゆるコマ送りによる表現も映像に含めることとする。したがって、本願の出願人は特許文献1に示す「動く漫画」を発明しているが、この「動く漫画」もここでいう映像に含まれる。

【0004】

ところで、動画には映像とともに音声が含まれることは既に説明したとおりであるが、映像中の登場人物（人に限定されないため、以下ここでは「キャラクター」という。）に関する音声（例えば台詞など）は、映像中の者とは異なる者が担当することも少なくない。外国映画を日本語に吹き替えるケースや、アニメーションのキャラクターに対して台詞を入れるケース（いわゆるアテレコ）などはよく知られている。また、演劇やアニメーション、動く漫画の場合、通常は複数のキャラクターが登場することから、キャラクターの数だけ音声担当者（例えば声優）が用意される。なお、ここでは便宜上、台詞などキャラクターに関する音声の担当を「声役」ということとする。つまり、AというキャラクターとBというキャラクターが登場する場合、声役Aと声役Bが必要になるわけである。

20

【0005】

映画やアニメーションでは、専門の声優が声役を担当し、これまで視聴者（ユーザ）が声役として参加する余地はなかった。一方、コンピュータを用いて表示する動画の場合、音声データを映像データと分離して構成することができることから、ユーザが声役として参加することも十分考えられる。音声データのうち特定の声役に対して、ユーザの音声に書き換えることができれば、ユーザ自身の音声で反映された動画が再生できるわけである。

30

【0006】

特定の声役にユーザ自身の声音を反映させるということは、つまり実際の声優とユーザが共演した動画を作成するということである。昨今では声優を志望する者が急増しており、このように実際の声優との共演を実現できる仕組みは広く要望されることが予想される。特許文献2でも、ユーザ自身がナレーションし、その音声を録音した紙芝居などを作成する方法について提案している。

40

【先行技術文献】

【特許文献】

【0007】

【特許文献1】特許第5327823号公報

【特許文献2】特開2005-267065号公報

【発明の概要】

【発明が解決しようとする課題】

【0008】

特許文献2では、紙芝居などのナレーションとしてユーザ自身の音声を反映することが

50

できる。しかしながら、通常の動画には複数のキャラクターが登場し、すなわち多くの声役が用意されるため、ユーザの音声を反映させる声役をいずれか選択しなければならないが、特許文献2では所望の声役を選んで音声を反映させることができない。また、ユーザ自身の音声と実際の声優が共演した動画を作成したいと思う場合、一つの声役に対して様々な声優が選択できれば好適であるが、このような技術は、特許文献2を含め未だ提案されることがなかった。

【0009】

さらに、音声を録音する際、映像を確認しながら行う方が、作品に適した音声を記録することができるが、例えば話すべき台詞と、その台詞に許される時間（いわゆる尺と呼ばれるもの）との兼ね合いは、映像だけで判断することは難しい。

10

【0010】

本願発明の課題は、従来技術が抱える問題を解決することであり、すなわち、動画中に表れる複数のキャラクターの中から、所望のキャラクターを選んでユーザ自身の音声を反映させるとともに、他の声役を複数の音声から選択でき、しかも映像を確認しながら音声を録音する際、その速度を案内することのできる音声録音プログラム、音声録音端末装置、及び音声録音システムを提供することである。

【課題を解決するための手段】

【0011】

本願発明は、動画中に登場する複数のキャラクターから所望のものを選んで音声録音し、しかも映像を再生しながら録音するという点に着目して開発されたものであり、従来にはない発想に基づいてなされた発明である。

20

【0012】

本願発明の「音声録音プログラム」は、映像（コマ送りされる画像の連続表示を含む）を確認しながら音声を録音するプログラムであって、コンテンツ読出し処理と、主声役選択処理、映像表示処理、従音声選択処理、音声記憶処理をコンピュータに実行させる機能を備えたものである。このうちコンテンツ読出し処理は、映像及びタイムラインを有する「コンテンツ」を読み出す処理である。このタイムラインには、2以上の声役が設定されるとともに、声役ごとに映像中に割り当てられる録音時間帯が設定されている。主声役選択処理は、2以上の声役のうち録音対象とする主声役を選択する処理であり、映像表示処理は、映像を表示する処理である。また、従音声選択処理は、それぞれの「従声役（2以上のうち主声役を除く声役）」に対して記憶された2以上の録音ユーザによる音声のうち所望の録音ユーザの音声を選択する処理であり、音声記憶処理は、主声役の録音時間帯内に録音された音声をその録音時間帯と関連付けて記憶する処理である。

30

【0013】

本願発明の「音声録音プログラム」は、動画再生処理を、コンピュータに実行させる機能をさらに備えたものとすることもできる。この動画再生処理は、映像とともに音声を出力するもので、主声役に対応する録音時間帯では音声記憶処理で録音された音声を出力し、従声役に対応する録音時間帯では従音声選択処理で選択された種類の音声を出力する。

【0014】

本願発明の「音声録音プログラム」は、従音声出力処理を、コンピュータに実行させる機能をさらに備えたものとすることもできる。この従音声出力処理は、録音する際、従声役に対して設定された録音時間帯で、従音声選択処理で選択された種類の音声を出力するもので、従音声出力処理によって従音声出力されることで、主声役の録音タイミングが計りやすくなる。

40

【0015】

本願発明の「音声録音プログラム」は、録音支援処理を、コンピュータに実行させる機能をさらに備えたものとすることもできる。この録音支援処理は、タイムラインに基づいて主声役の録音時間帯になると録音速度計を表示するものである。録音速度計は、主声役の録音時間帯における進行状況や残り時間を動的に示すもので、これを表示することで音声の録音を支援することができる。

50

## 【0016】

本願発明の「音声録音端末装置」は、映像（コマ送りされる画像の連続表示を含む）を確認しながら音声を録音する端末装置であって、コンテンツ読出し手段と、主声役選択手段、従音声選択手段、映像表示手段、音声録音手段を備えたものである。このうちコンテンツ読出し手段は、映像及びタイムラインを有する「コンテンツ」を読み出す手段である。このタイムラインには、2以上の声役が設定されるとともに、声役ごとに映像中に割り当てられる録音時間帯が設定されている。主声役選択手段は、2以上の声役のうち録音対象とする主声役を選択する手段であり、従音声選択処理は、それぞれの従声役に対して記憶された2以上の録音ユーザによる音声のうち所望の録音ユーザの音声を選択する処理である。映像表示手段は、映像を表示する手段であり、音声録音手段は、音声を録音する手段である。

10

## 【0017】

本願発明の「音声録音システム」は、映像（コマ送りされる画像の連続表示を含む）を確認しながら音声を録音するシステムであって、コンテンツ入力手段と、コンテンツ記憶手段、音声録音端末装置、音声記憶手段を備えたものである。コンテンツ入力手段は、コンテンツを入力するものであり、コンテンツ記憶手段は、入力されたコンテンツを記憶するものであり、音声記憶手段は、音声録音端末装置で録音された音声を記憶するものである。コンテンツには、映像及びタイムラインが含まれ、このタイムラインには、2以上の声役が設定されるとともに、声役ごとに映像中に割り当てられる録音時間帯が設定されている。また、音声録音端末装置は、コンテンツが有する映像を確認しながら音声を録音するものであって、コンテンツ読出し手段と、主声役選択手段、従音声選択手段、映像表示手段、音声録音手段を具備している。このうちコンテンツ読出し手段は、コンテンツ記憶手段からコンテンツを読み出すものであり、主声役選択手段は、2以上の声役のうち録音対象とする主声役を選択するものであり、従音声選択処理は、それぞれの従声役に対して記憶された2以上の録音ユーザによる音声のうち所望の録音ユーザの音声を選択する処理である。映像表示手段は、映像を表示するものであり、音声録音手段は、音声を録音するものである。なお、音声記憶手段は、録音時間帯内に録音された音声をその録音時間帯と関連付けて記憶する。

20

## 【発明の効果】

## 【0018】

本願発明の音声録音プログラム、音声録音端末装置、及び音声録音システムには、次のような効果がある。

30

(1) 同じ映像であっても、ユーザ自身の音声を録音することで、オリジナルの動画を作成することができる。

(2) さらに、台詞をアレンジすることによって、より独創的な動画を作成することができる。

(3) 映像を確認しながら音声を録音することができるので、その作品（コンテンツ）に応じた音声（例えば、感情を込めた音声）を録音することができる。

(4) 音声録音の際、従声役（主声役以外の声役）の音声を再生することで、さらにその作品に応じた臨場感をもって音声録音することができる。

40

(5) 録音速度計が、主声役の録音時間帯における進行状況や残り時間を動的に示すので、あらかじめ定められた録音時間帯に適切に音声録音することができる。

(6) ユーザ自身が録音した音声と、あらかじめ入力された声優の音声を再生することによって、声優との共演を実現することができる。

(7) 無線又は有線による通信手段を通じて、複数の音声録音端末装置の接続が可能な音声録音システムとすると、同一の映像に対して、見知らぬ者どうしが非同期で（時間と場所を選ばず）共演することができる。

## 【図面の簡単な説明】

## 【0019】

【図1】タイムラインを説明するためのモデル図。

50

【図2】本願発明の音声録音プログラムのうち主に音声録音に関する処理の流れの例を示すフロー図。

【図3】本願発明の音声録音プログラムのうち主に動画再生に関する処理の流れの例を示すフロー図。

【図4】録音速度計の1例を示すモデル図。

【図5】本願発明の音声録音システムのうち主に音声録音に必要な構成を示すブロック図。

【図6】本願発明の音声録音システムのうち主に動画再生に必要な構成を示すブロック図。

【図7】本願発明の音声録音システムを使用した1例を示すブロック図。

10

【発明を実施するための形態】

【0020】

本願発明の音声録音プログラム、音声録音端末装置、及び音声録音システムの例を図に基づいて説明する。

【0021】

1. 定義

本願発明の実施形態の例を説明するにあたって、はじめにここで用いる用語の定義を示しておく。

【0022】

(動画とコンテンツ)

20

動画は、映像と音声を表示したものである。なおここで映像とは、既述したとおり、多数の静止画を高速で切り替えることで動きを表現するもののほか、複数の静止画を断続的に切り替える、いわゆるコマ送りによる表現も含まれる。一方、コンテンツは、映像データとタイムラインデータを含むもので、アニメーションや物語のタイトル(作品)ごとに用意される。タイムラインは映像中の音声時間を設定するものであり、このタイムラインに従って音声を録音し、あるいは出力する。つまり、コンテンツを動画再生手段で再生したものが、「動画」として表示されるわけである。

【0023】

(タイムラインと声役)

タイムラインについて、図1を参照しながらさらに詳しく説明する。図1は、タイムラインを説明するためのモデル図である。タイムラインは、この図に示すように、映像の開始から終了までの間、どのタイミングで、しかもどの程度の時間幅(いわゆる尺)で、音声を表現するかを設定するものであり、いわば音声時間を決める設計図の役割を果たすものである。

30

【0024】

通常、一つの動画作品(タイトル)には、複数のキャラクターが登場し、当然ながらキャラクターごとに台詞やその出力タイミング等は異なる。なお、ここでは便宜上、キャラクターの音声担当を「声役」としている。例えば、図1に示すタイトルでは6種類のキャラクターが登場し、それぞれのキャラクターに対して、声役A~声役Fが割り当てられている。このように複数の声役が用意されている場合、当然ながらタイムラインは、それぞれの声役に対して、音声を入出力する時間帯(以下、「録音時間帯」という。)を設定することとなる。図1では、例えば、声役Aに対して3つの録音時間帯が設定され、声役Dでは2つの録音時間帯が設定されている。

40

【0025】

タイムラインは、登場するキャラクターに係る音声に限らず、効果音(SE: Sound Effect)や背景音楽(BGM: Back Ground Music)の録音時間帯である「音響時間帯」も必要に応じて設定する。図1では、効果音として7つの音響時間帯、背景音楽として3つの音響時間帯が設定されている。もちろん、効果音や背景音楽も、声役と同様に2種類以上(例えば、効果音1、効果音2など)用意することもできる。

50

## 【 0 0 2 6 】

## 2 . 音声録音プログラム

次に、本願発明の音声録音プログラムについて、図を参照しながら説明する。図2は、本願発明の音声録音プログラムのうち主に音声録音に関する処理の流れの例を示すフロー図であり、図3は、主に動画再生に関する処理の流れの例を示すフロー図である。なお、それぞれのフロー図は、中央の列に実施する処理を示し、左列にはその処理に必要な入力情報を、右列にはその処理から発生する出力情報を示している。また、ここで示す処理は、具体的にはコンピュータによって実行される。

## 【 0 0 2 7 】

まず、ユーザが所望のタイトル（動画作品）に係るコンテンツを選択し、さらに所望の10  
声役（キャラクター）に対して音声録音する処理について、図2に基づいて説明する。  
はじめに、映像とタイムラインが含まれたコンテンツを、例えばコンテンツ記憶手段（コ  
ンテンツサーバ）から読み出す（Step 101）。通常、コンテンツサーバには複数の  
タイトルに係るコンテンツが記憶されており、したがってユーザはそのうち所望のタイ  
トルに係るコンテンツを選んで読み出す。

## 【 0 0 2 8 】

読み出されたコンテンツに含まれるタイムラインには、2以上の声役が設定されている  
ため、ユーザは担当したい声役、すなわち自身の音声録音したい声役を選択する（St  
ep 102）。なお、便宜上ここでは、複数の声役のうち音声録音対象として選択された  
20  
声役を「主声役」という。主声役が選択されると、主声役を除く他の声役は「従声役」と  
して設定される。

## 【 0 0 2 9 】

従声役に対して、既に記憶された音声がある場合、その声を読み出すこともできる。さ  
らに、一つの声役に対して2以上の音声録音が記憶されていれば、ユーザが所望する音声を選  
択することもできる（Step 103）。もちろん、複数の従声役がある場合は、それぞ  
れの声役に対して所望の音声を選択する。一方、ユーザが従声役に対して音声を選択し  
ない場合は、既定の（デフォルトの）音声録音が当該従声役の音声として設定される。こ  
こで、従声役の音声として選択（設定）されたものが「従音声」である。なお、従音  
声を  
30  
選択する場合、その従声役として記憶された複数の音声のうち一部（あるいはすべて）の音  
声を試聴したうえで、所望の音声を選択することもできる。この場合、記憶された音  
声を  
読み出す処理と、その音声を再生する処理が必要となる。

## 【 0 0 3 0 】

あるいは、従音声を選択する場合、音声録音した者（以下、「録音ユーザ」という。  
）に係る属性情報（以下、「ユーザ情報」という。）を表示することもできる。例えば、  
ユーザ情報を参照すれば、支持する声優に係る音声を従音声として選択することができ  
るわけである。この場合、ユーザ情報を記憶するユーザ情報記憶手段と、記憶された音  
声に係るユーザ情報を読み出す処理、そのユーザ情報を例えば一覧形式で表示する処理が必要  
となる。

## 【 0 0 3 1 】

ここまでの処理が終わると、いよいよ音声録音が開始する（Step 104）。例え  
ば、録音開始のトリガーとなる操作を行うと、映像が映像表示手段に映し出される（St  
ep 105）。そして、主声役の録音時間帯になるとタイムラインにしたがって録音支援  
40  
処理が開始される（Step 106）。この録音支援処理は、具体的には録音速度計を表  
示するものであり、また録音速度計は、主声役の録音時間帯における進行状況を動的に示  
すものである。録音速度計は、主声役の録音時間帯における進行状況に代えて（あるいは  
加えて）残り時間を動的に示すものとすることもできる。

## 【 0 0 3 2 】

図4は、録音速度計の1例を示すモデル図である。なお、この図では「マリノ」という  
キャラクターが主声役として選択されている。この図に示すように、主声役に対して設定  
された録音時間帯になると、録音速度計の動的表示が開始される。この図の録音表示計は  
50

、左右に長手方向を有する長方形枠の中を、移動針が左から右に向かって移動するものである。また、この録音表示計は、既に経過した時間帯を濃い色で、残りの時間帯を薄い色で示しており、その境界が移動針として表示されている。この図の録音速度計は、長方形枠の幅（移動針が移動する範囲）を一定長としているため、録音時間帯が比較的長い場合は移動針の移動速度を緩速とし、録音時間帯が比較的短い場合は移動針の移動速度を急速としている。なお、録音速度計は、図4に示す形式に限らず、主声役の音声録音を開始するタイミングを示すとともに、録音時間帯における進行状況又は残り時間を動的に示すことができ、例えば時計表示や円グラフ表示、あるいは数字のみの百分率表示など、種々の形式を採用することができる。

#### 【0033】

10

さらに録音速度計は、その録音時間帯で話すべき台詞を表示することもできる。もちろん、台詞どおり話すのではなく、アドリブとして独創的な台詞を話すこともできるし、声ではなく楽器その他の音を発することもできる。ここで話された（発せられた）音声は、録音手段によって録音され（Step107）、主声役に係る音声（以下、「主音声」という。）として記憶される（Step108）。このとき、主音声は当該録音時間帯と関連づけて（紐づけて）記憶される。例えば、図1で声役Aが主声役とすると、録音時間帯A1で録音された主音声は、録音時間帯A1に紐づく主音声として記憶され、録音時間帯A2で録音された主音声は、録音時間帯A2に紐づく主音声として記憶される。

#### 【0034】

主音声を録音するに当たっては、図4にも示すように、タイムラインに従って従声役の録音速度計を表示することもできる。加えて、従声役の録音速度計に当該録音時間帯の台詞を表示することもできるし（Step109）、Step103において選択（設定）された従音声を出力することもできる（Step110）。もちろん、タイムラインに従って効果音や背景音楽を出力することもできる。従声役の録音速度計を表示し、従音声や効果音等を出力することで、主声役を録音するタイミングが計りやすくなるとともに、臨場感が向上してより適した音声を録音することができる。ただしこの場合、主声役で表示される録音速度計と、従声役で表示される録音速度計を識別可能に表示すると良い。主声役と従声役の録音速度計を同様にすると、主声役を録音するタイミングが却って計りにくくなるからである。例えば、主声役の録音表示計の濃淡表示（経過時間と残時間）は赤色とし、従声役の録音表示計の濃淡表示はグレーとすると良い。なお、異なる従声役の録音表示計は、それぞれ識別可能にしてもよいし、統一して識別不可とすることもできる。

20

30

#### 【0035】

さらに、主音声を録音するに当たっては、タイムラインに従って主音声を出力しながら録音することもできる。例えば、外国語の台詞の場合、ヒアリングしながらその発音に合わせて録音するわけである。この場合も従音声と同様、主声役に対して2種類以上の音声記憶されていれば、記憶された音声のうち所望の種類を選択して出力させることができる。その際、記憶された複数の音声のうち一部（あるいはすべて）の音声を試聴したうえで、所望の主音声を選択することもできるし、表示されたユーザ情報を参照しながら所望の主音声を選択することもできる。

#### 【0036】

40

ところで図4の映像は、いわゆるコマ送りによる映像を表示しており、この画像に関する複数の録音速度計の長方形枠が、一覧形式で表示されている。この場合、画像が切り替わる（つまりコマ送りされる）タイミングで長方形枠の一覧を総入れ替えすることもできるし、画像の切り替えにかかわらず、時間の経過とともに長方形枠の一覧を上方にスクロール移動させることもできる。もちろん、画像を高速切り替えする映像の場合も同様である。

#### 【0037】

映像が終了する（図1のEND）まで、つまりすべての録音時間帯（例えば、図1のA1～A3）に対して繰り返し音声を録音する処理が終わると、一連の録音処理が終了する（Step111）。

50

## 【 0 0 3 8 】

つぎに図 3 に基づいて、ユーザが所望のタイトルに係るコンテンツを選択し、そのコンテンツを再生する処理について説明する。はじめに、映像とタイムラインが含まれたコンテンツを、例えばコンテンツ記憶手段（コンテンツサーバ）から読み出す（Step 201）。通常、コンテンツサーバには複数のタイトルに係るコンテンツが記憶されており、したがってユーザはそのうち所望のタイトルに係るコンテンツを選んで読み出す。なお、ここで読み出すコンテンツに係る録音時間帯には、既に音声録音されている。したがって、図 2 のフロー図で説明したコンテンツサーバとここで説明するコンテンツサーバは異なるものとして用意してもよいし、1のコンテンツサーバを用意して、これに記憶されるコンテンツが録音済みであるか、あるいは未録音のものか識別できるように記憶させることもできる。

10

## 【 0 0 3 9 】

所望のコンテンツを読み出すと、声役に対して出力する音声を選択する（Step 202）。なお、動画を再生する場合は、音声を録音する場合と異なり、主声役と従声役に分けて設定する必要はない。複数ある声役それぞれに対して、2以上記憶された音声のうち所望の音声を選択する。もちろん、一つの声役に対して1の音声のみが記憶されている場合はそのまま指定する。一方、ユーザが声役に対して音声を選択しない場合は、既定の（デフォルトの）音声が当該声役の音声として設定される。ここで、声役の音声として選択（設定）されたものがそれぞれの声役の「再生音声」である。なお、音声を選択する場合、その声役として記憶された複数の音声のうち一部（あるいはすべて）の音声を試聴したうえで、所望の音声を選択することもできるし、表示されたユーザ情報を参照しながら所望の音声を選択することもできる。

20

## 【 0 0 4 0 】

ここまでの処理が終わると、いよいよ動画再生が開始する（Step 203）。例えば、再生開始のトリガーとなる操作を行うと、映像が映像表示手段に映し出される（Step 204）。そして、Step 201で読み出したコンテンツのタイムラインに規定された録音時間帯に応じて、Step 202で選択され読み出された再生音声出力される（Step 205）。

## 【 0 0 4 1 】

動画再生中に、それぞれの声役に対応する録音速度計をタイムラインに従って表示し、さらに対応する台詞を表示させることもできる（Step 206）。映像が終了する（図 1 の END）まで、つまりすべての声役のすべての録音時間帯に対して繰り返し音声出力され、一連の動画再生処理が終了する（Step 207）。

30

## 【 0 0 4 2 】

## 3. 音声録音端末装置、及び音声録音システム

ここでは、本願発明の音声録音端末装置、及び音声録音システムについて、図を参照しながら説明する。なお、音声録音プログラムで説明した内容と重複する説明は避け、音声録音端末装置、及び音声録音システムに特有の内容のみ説明することとする。すなわち、ここに記載されていない内容は、音声録音プログラムで説明したものと同様である。

## 【 0 0 4 3 】

## （音声録音端末装置）

はじめに、図 5 を参照しながら、音声録音システムを構成する音声録音端末装置 100 について説明する。図 5 は、本願発明の音声録音システムのうち主に音声録音に必要な構成を示すブロック図である。音声録音端末装置 100 は、音声録音プログラムの処理（全部または一部）を実行するものであり、専用のもので製造することもできるが、汎用的なコンピュータ装置を利用することもできる。このコンピュータ装置は、パーソナルコンピュータ（PC）や、iPad（登録商標）といったタブレット型端末やスマートフォン、あるいは PDA（Personal Data Assistance）などによって構成することができる。コンピュータ装置は、CPU等のプロセッサ、ROMやRAMといったメモリを具備しており、さらにマウスやキーボード等の入力手段やディスプレイ

40

50

(映像表示手段103)を含むものもある。なお、一般的なPCであればマウスやキーボード等のデバイスから入力するが、タブレット型端末やスマートフォンではタッチパネルを用いた操作(タップ、ピンチイン/アウト、スライド等)で入力することが多い。

#### 【0044】

音声録音端末装置100は、図5に示すように、コンテンツ読出し手段101と、主声役選択手段102、映像表示手段103、録音支援手段104、音声録音手段105を具備している。コンテンツ読出し手段101は、コンテンツ記憶手段(コンテンツサーバ200)からコンテンツを読み出すものである。通常、コンテンツサーバには複数のタイトルに係るコンテンツが記憶されているので、コンテンツ読出し手段101は複数の中から所望のタイトルに係るコンテンツを選んで読み出すことができる。

10

#### 【0045】

読み出されたコンテンツに含まれるタイムラインには、2以上の声役が設定されている。そこで、主声役選択手段102が、この2以上の声役から主声役を選択する。また、映像表示手段103は、例えばディスプレイなどのように映像を表示し、録音支援手段104は、図4に示すような録音速度計を表示させる。そして、音声録音手段105を用いて音声を録音することができる。なお図5では、主音声を記憶する音声記憶手段(音声サーバ300)を音声録音端末装置100の外部に設けているが、これに限らず音声録音端末装置100が音声サーバ300を具備することもできる。音声サーバ300は、主音声と、当該主音声の録音時間帯を、それぞれ関連づけて(紐づけて)記憶する。

#### 【0046】

20

また、コンテンツサーバ200からコンテンツを読み出すために、無線又は有線による通信手段を通じてコンテンツを受信する「受信手段106」を、音声録音端末装置100に具備させることもできる。同様に、録音時間帯と紐づいた主音声を、無線又は有線による通信手段を通じて、音声サーバ300(この場合は音声録音端末装置100の外部に設けている)に送信する「送信手段107」を、音声録音端末装置100に具備させることもできる。

#### 【0047】

さらに、録音ユーザに係る属性情報である「ユーザ情報」を入力する「ユーザ情報入力手段108」を、音声録音端末装置100に具備させることもできる。この場合、入力されたユーザ情報は、送信手段107によってユーザ情報記憶手段(ユーザ情報サーバ400)に記憶される。このユーザ情報サーバ400は、主音声として記憶された音声と、当該音声を録音した録音ユーザに係るユーザ情報を、それぞれ関連づけて(紐づけて)記憶する。

30

#### 【0048】

##### (音声録音システム)

音声録音システムは、図5に示すように、音声録音端末装置100と、コンテンツ入力手段201、コンテンツサーバ200、音声サーバ300で構成される。ここでコンテンツ入力手段201は、コンテンツサーバ200にコンテンツを入力するためのものである。なお、既述のとおりコンテンツサーバ200は音声録音端末装置100に具備させることもできるし、同様に、コンテンツ入力手段201も音声録音端末装置100に具備させることができる。さらに、ユーザ情報サーバ400を含めて、音声録音システムを構成することもできる。

40

#### 【0049】

図6は、本願発明の音声録音システムのうち主に動画再生に必要な構成を示すブロック図である。この図では、図5に示す音声録音システムの構成に加えて、動画再生装置500が設けられている。なおこの図では、動画再生装置500が音声録音端末装置100から独立して(別体として)設けられているが、これに限らず、動画再生装置500を音声録音端末装置100に含めて(一体として)設けることもできる。

#### 【0050】

動画再生装置500は、図6に示すように、再生音声選択手段501と、再生音声読出

50

し手段502、動画再生手段503を具備している。再生音声選択手段501は、コンテンツ読み出し手段101で読み出した所望のコンテンツで設定される複数の声役に対して、それぞれの声役について出力する音声を選択する。ここで選択された音声が、再生音声として設定される。再生音声読み出し手段502は、再生音声選択手段501で選択(設定)した再生音声を読み出し、動画再生手段503は、コンテンツに含まれる映像を、映像表示手段103に表示して再生する。さらに動画再生装置500は、各々の声役で選択した再生音声を一つの組み合わせ(以下、「音声グループ」という。)として登録する音声グループ登録手段を具備したものとすることもできる。この場合、再生音声選択手段501は、個々の声役ごとに再生音声を選択することもできるし、既登録のものがあれば音声グループを選択することもできる。配役どうしの相乗効果によってより優れた動画作品となることもあり、音声グループを登録しておけば容易に同じ声役の組み合わせで動画を閲覧することができて好適となる。

10

#### 【0051】

再生音声選択手段501で出力音声を選択する際、音声サーバ300に記憶された複数の音声のうち一部(あるいはすべて)の音声を試聴したうえで、所望の再生音声を選択することもできるし、表示されたユーザ情報を参照しながら所望の再生音声を選択することもできる。この場合、再生音声選択手段501が、音声サーバ300から音声を読み出して出力し、あるいはユーザ情報サーバ400からユーザ情報を読み出して表示する。

#### 【0052】

また、無線又は有線による通信手段を通じて、音声サーバ300から音声を受信する、あるいはユーザ情報サーバ400からユーザ情報を受信する「受信手段505」を、動画再生装置500に具備させることもできる。さらに、動画再生後に、選択した再生音声を録音した録音ユーザに対して、評価する手段として「評価手段504」を設けることもできる。ここで評価した結果(例えば、テキストデータ)は、「送信手段506」を用い、無線又は有線による通信手段を通じて、ユーザ情報サーバ400に送信することもできる。なお、受信手段505は、図5に示す受信手段106と兼用することもできるし、送信手段506は、図5に示す送信手段107と兼用することもできる。

20

#### 【0053】

(音声録音システムの使用例)

図7は、本願発明の音声録音システムを使用した1例を示すブロック図である。この図では、クライアントがコマーシャル(CM)用に映像を作成し、その映像の声役に適した音声役(ユーザ)を探し出すケースを示している。クライアントは、映像とともにタイムラインを作成し、映像とタイムラインからなるコンテンツをコンテンツサーバ200に登録する。このとき、映像中、採用したい声役(つまり主音声)を明確にしておくのが望ましい。

30

#### 【0054】

一方、録音ユーザは、音声録音端末装置100を用いて、クライアントが目的としている声役を主声役として、自身の音声を録音する。そして録音された音声は、音声サーバ300に記憶される。なお、ユーザ情報はあらかじめユーザ情報サーバ400に記憶されている。

40

#### 【0055】

何人かの録音ユーザによって音声サーバ300に音声記憶されると、今度はクライアントが、動画再生装置500を用いて動画再生する。このとき、音声試聴やユーザ情報を参考に再生音声を選択したうえで動画再生を行う。そして、何人かの録音ユーザに係る再生音声を確認し、その中から採用したい録音ユーザを選択する。ここでクライアントが採用したい録音ユーザを選択すると、その採用した結果を「ユーザ選択情報」として当該録音ユーザに通知する。具体的には、図7に示す「選択情報送信手段601」によってユーザ選択情報を送信し、「選択情報受信手段602」によって録音ユーザが選択情報を受信する。なお、選択情報を送受信するとともに、ユーザ情報に含まれる銀行口座等の情報に基づいて、出演フィーとしてクライアントが当該登録ユーザの銀行口座に入金する仕組み

50

とすることもできる。

【産業上の利用可能性】

【0056】

本願発明の音声録音プログラム、音声録音端末装置、及び音声録音システムは、実写の物語（ドラマ）や、アニメーション、あるいは動く漫画に利用できるほか、外国語の会話練習、テレビコマーシャルの声優選び、あるいは映像付きの音楽演奏にも利用することができる。すなわち、今後、様々な産業界で応用が期待できる発明である。

【符号の説明】

【0057】

100	音声録音端末装置	10
101	コンテンツ読出し手段	
102	主声役選択手段	
103	映像表示手段	
104	録音支援手段	
105	音声録音手段	
106	受信手段	
107	送信手段	
108	ユーザ情報入力手段	
200	コンテンツサーバ	
300	音声サーバ	20
400	ユーザ情報サーバ	
500	動画再生装置	
501	再生音声選択手段	
502	再生音声読出し手段	
503	動画再生手段	
504	評価手段	
505	受信手段	
506	送信手段	
601	選択情報送信手段	
602	選択情報受信手段	30

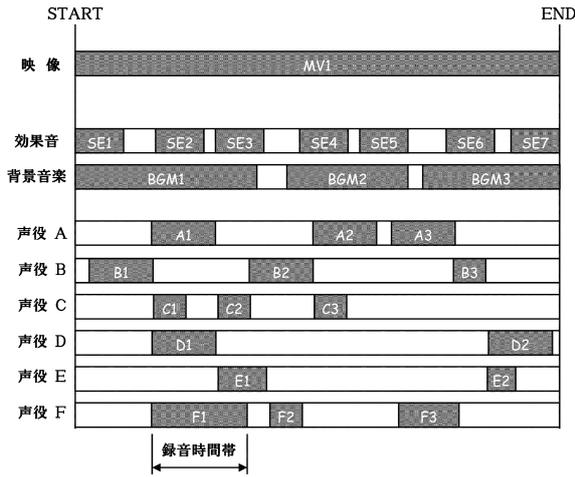
【要約】

【課題】本願発明の課題は、従来技術が抱える問題を解決することであり、すなわち、動画中に表れる複数のキャラクターの中から、所望のキャラクターを選んでユーザ自身の音声を反映させるとともに、他の声役を複数の音声から選択でき、しかも動画を確認しながら音声を録音する際、その速度を案内することのできる音声録音プログラム、音声録音端末装置、及び音声録音システムを提供することである。

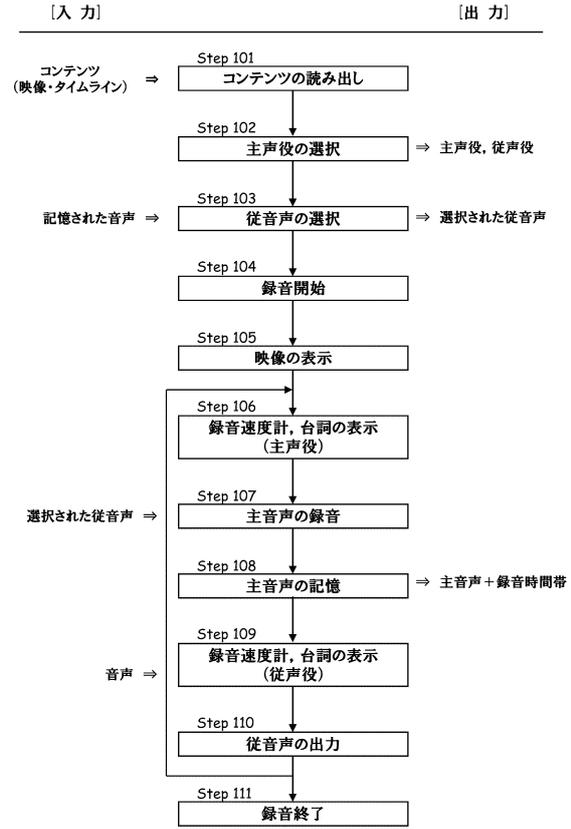
【解決手段】本願発明の「音声録音プログラム」は、動画（コマ送りされる画像の連続表示を含む）を確認しながら音声を録音するプログラムであって、コンテンツ読出し処理と、主声役選択処理、動画表示処理、録音支援処理、音声記憶処理をコンピュータに実行させる機能を備えたものである。このうちコンテンツ読出し処理は、動画及びタイムラインを有する「コンテンツ」を読み出す処理である。

【選択図】図4

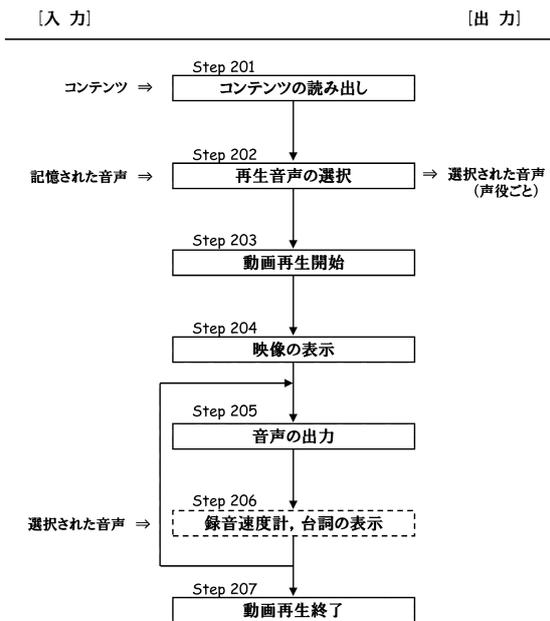
【図1】



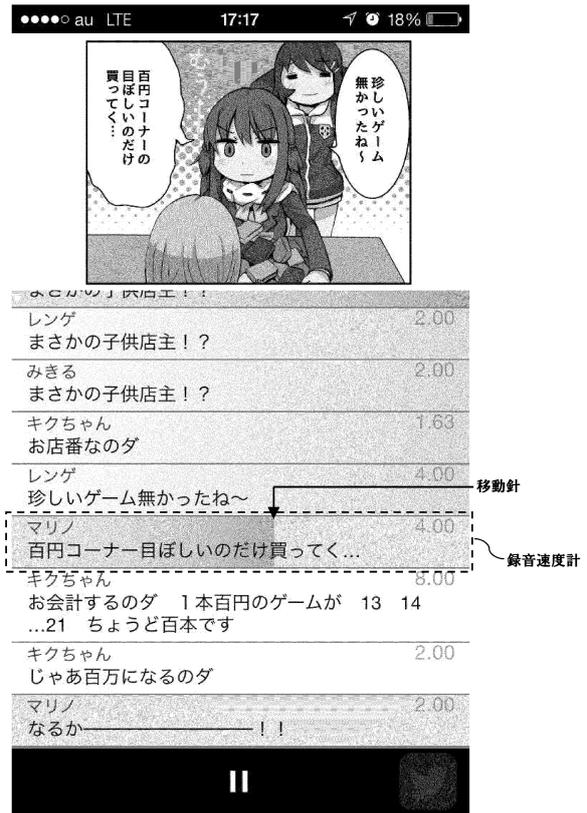
【図2】



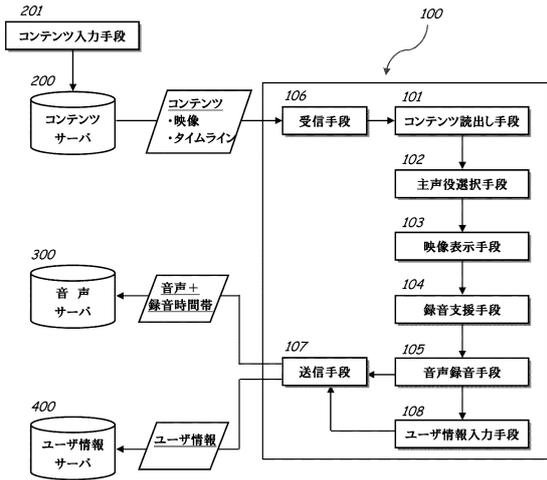
【図3】



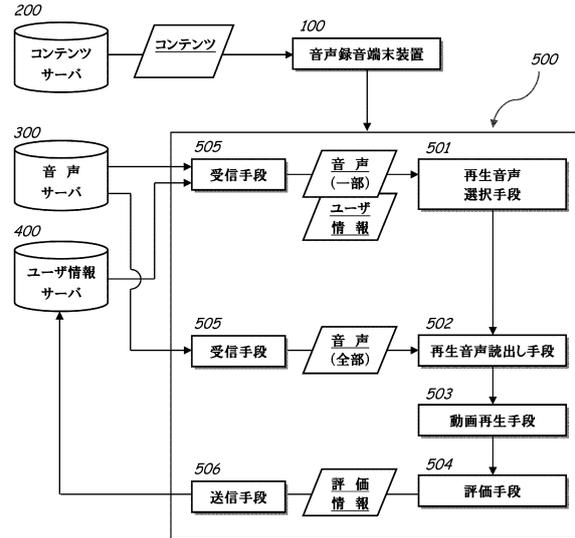
【図4】



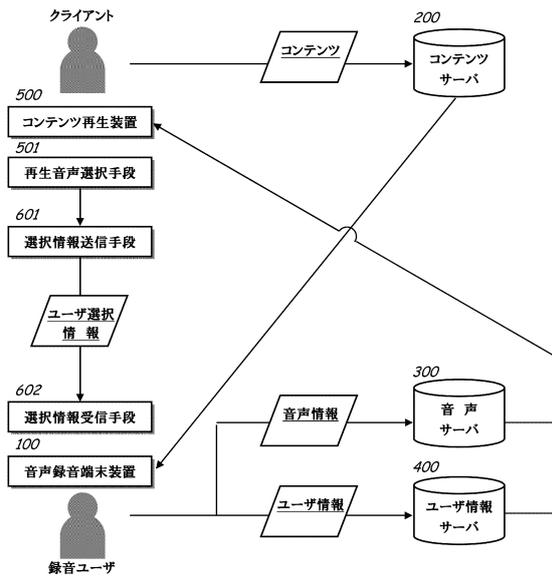
【図5】



【図6】



【図7】



---

フロントページの続き

(56)参考文献 特開2006-346284(JP,A)  
特開2007-328849(JP,A)  
特開2002-346211(JP,A)  
特開2009-122667(JP,A)  
特開2003-122554(JP,A)

(58)調査した分野(Int.Cl., DB名)

A63F9/24  
13/00 - 13/98  
H04N5/765  
5/91  
5/915  
5/92  
5/922  
5/928 - 5/93  
5/937 - 5/94  
5/95 - 5/956