

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2016-509408
(P2016-509408A)

(43) 公表日 平成28年3月24日 (2016. 3. 24)

(51) Int.Cl.	F I	テーマコード (参考)
HO4N 21/435 (2011.01)	HO4N 21/435	5C053
HO4N 21/81 (2011.01)	HO4N 21/81	5C164
HO4N 21/431 (2011.01)	HO4N 21/431	
HO4N 5/91 (2006.01)	HO4N 5/91 N	
HO4N 5/765 (2006.01)	HO4N 5/91 L	

審査請求 有 予備審査請求 未請求 (全 43 頁)

(21) 出願番号 特願2015-552622 (P2015-552622)
 (86) (22) 出願日 平成25年8月19日 (2013. 8. 19)
 (85) 翻訳文提出日 平成27年8月11日 (2015. 8. 11)
 (86) 国際出願番号 PCT/US2013/055573
 (87) 国際公開番号 W02014/113071
 (87) 国際公開日 平成26年7月24日 (2014. 7. 24)
 (31) 優先権主張番号 61/752, 816
 (32) 優先日 平成25年1月15日 (2013. 1. 15)
 (33) 優先権主張国 米国 (US)
 (31) 優先権主張番号 61/754, 926
 (32) 優先日 平成25年1月21日 (2013. 1. 21)
 (33) 優先権主張国 米国 (US)

(71) 出願人 515166705
 ヴィキ, インク.
 V I K I, I N C.
 アメリカ合衆国 カリフォルニア州 94
 104, サンフランシスコ, ブッシュスト
 リート 225, スイート 1325
 (74) 代理人 110001302
 特許業務法人北青山インターナショナル
 (72) 発明者 ポーンブラシツァカル, ナッタナート
 アメリカ合衆国 カリフォルニア州 94
 104, サンフランシスコ, ブッシュスト
 リート 225, スイート 1325

最終頁に続く

(54) 【発明の名称】 メディアにキャプションを付けるシステム及び方法

(57) 【要約】

ストリーミングされたメディアに共同でキャプションを付ける方法であって、当該方法は、第1デバイスで音声の視覚表示をレンダリングするステップと、第1メディアセグメントに関するセグメントパラメータを第1デバイスから受信するステップと、第1デバイスとは異なる第2デバイスで音声の視覚表示をレンダリングするステップと、第1メディアセグメントに関するキャプションを第2デバイスから受信するステップと、を含む。

【選択図】 図1

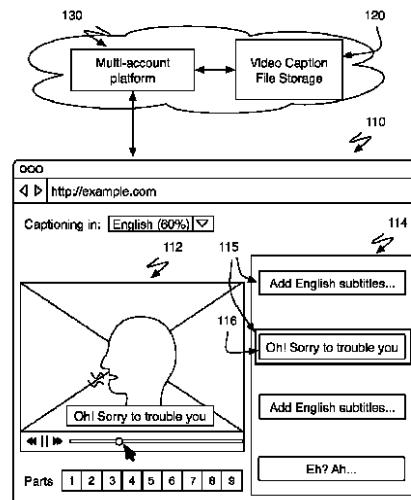


FIG. 1

【特許請求の範囲】**【請求項 1】**

映像及び音声を有するストリーミングされたメディアにキャプションを付ける方法であって、前記方法は、

遠隔メディアリポジトリからデバイスに前記メディアをストリーミングするステップと

、
第 1 デバイスで前記音声の視覚表示及び前記メディアのタイムラインをレンダリングするステップであって、前記音声の前記視覚表示は、前記タイムラインにおけるタイムスタンプ位置に対応する位置でレンダリングされる、レンダリングするステップと、

第 1 タイムスタンプ及び第 2 タイムスタンプで境界を付けられたメディアセグメントを定義するセグメントパラメータを受信するステップと、

前記第 1 タイムスタンプ及び前記第 2 タイムスタンプに関連付けられたキャプションを備えるキャプションファイルを生成するステップと、を含む方法。

10

【請求項 2】

前記遠隔メディアリポジトリから第 2 デバイスに前記メディアをストリーミングするステップと、

定義された前記メディアセグメントの前記第 1 タイムスタンプ及び前記第 2 タイムスタンプに対応したアイコンを有する前記音声の前記視覚表示をレンダリングするステップと、をさらに含む、請求項 1 に記載の方法。

20

【請求項 3】

定義された前記メディアセグメントに関連付けられた注釈を前記第 2 デバイスから受信するステップと、

前記第 1 タイムスタンプ及び前記第 2 タイムスタンプに関連付けられた前記キャプションのバージョンとして前記注釈を保存するステップを含む、前記キャプションファイルを編集するステップと、をさらに含む、請求項 2 に記載の方法。

20

【請求項 4】

注釈を受信するステップは、

定義された前記メディアセグメントの選択を受信するステップと、

選択された前記メディアセグメントに関連付けられたキャプション入力フィールドで前記注釈を受信するステップと、を含む、請求項 3 に記載の方法。

30

【請求項 5】

前記キャプション入力フィールドが、選択された前記メディアセグメント内の映像フレームに関連付けられたテキスト入力を含む、請求項 4 に記載の方法。

【請求項 6】

第 3 デバイスに前記メディアをストリーミングするステップと、

前記第 1 タイムスタンプと前記第 2 タイムスタンプとの間の映像フレーム上で、定義された前記メディアセグメントに関する前記注釈をレンダリングするステップと、をさらに含む、請求項 3 に記載の方法。

【請求項 7】

前記第 2 デバイスから境界変更を受信するステップと、変更された前記境界に対応したタイムスタンプを有する前記キャプションファイル内の前記キャプションを更新するステップと、をさらに含む、請求項 2 に記載の方法。

40

【請求項 8】

第 3 及び第 4 タイムスタンプによって境界を付けられた第 2 メディアセグメントを定義するセグメントパラメータを前記第 2 デバイスから受信するステップと、新たなキャプションに関連付けられた前記第 3 及び前記第 4 タイムスタンプを前記キャプションファイルに追加するステップと、をさらに含む、請求項 2 に記載の方法。

【請求項 9】

前記メディアセグメントを定義するセグメントパラメータを受信するステップは、前記視覚表示のセグメントの選択を受信するステップを含む、請求項 1 に記載の方法。

50

【請求項 10】

前記視覚表示のセグメントの選択を受信するステップは、
前記第1タイムスタンプで入力を受信するステップと、
前記第1タイムスタンプから始まる前記映像及び前記音声を再生するステップと、
前記入力の解除を検出して、前記入力解除された前記タイムスタンプに前記第2タイムスタンプを設定するステップであって、前記入力は、映像及び音声の再生中に維持される、設定するステップと、を含む、請求項9に記載の方法。

【請求項 11】

前記入力が、前記音声の前記視覚表示の位置のカーソル選択である、請求項10に記載の方法。

【請求項 12】

前記キャプションファイルが前記遠隔メディアリポジトリに記憶される、請求項1に記載の方法。

【請求項 13】

音声信号を含むメディアに共同でキャプションを付ける方法であって、前記方法は、
第1デバイスで前記音声信号の視覚表示をレンダリングするステップと、
第1メディアセグメントの境界に対応するセグメントパラメータを前記第1デバイスから受信するステップと、
第2デバイスで前記音声信号の前記視覚表示をレンダリングするステップであって、前記第2デバイスは前記第1デバイスとは異なり、前記視覚表示は、前記第1メディアセグメントの前記境界に対応するアイコンをさらに含む、レンダリングするステップと、
前記第1メディアセグメントに関する注釈を前記第2デバイスから受信するステップと、
を含む方法。

【請求項 14】

第2メディアセグメントの境界に対応するセグメントパラメータを前記第2デバイスから受信するステップをさらに含む、請求項13に記載の方法。

【請求項 15】

第3デバイスで前記音声信号の前記視覚表示をレンダリングするステップであって、前記第3デバイスは前記第1及び第2デバイスとは異なり、前記視覚表示が、前記第1及び第2タイムスタンプに対応するアイコンをさらに含む、レンダリングするステップと、前記第1タイムスタンプと前記第2タイムスタンプとの間で生じる映像フレーム上で前記第1メディアセグメントに関する前記注釈をレンダリングするステップと、第2メディアセグメントの境界に対応するセグメントパラメータを前記第3デバイスから受信するステップと、をさらに含む、請求項13に記載の方法。

【請求項 16】

前記第1メディアセグメントに関するセグメントパラメータ変更を前記第2デバイスから受信するステップと、前記セグメントパラメータ変更に基づいて前記第1メディアセグメント境界を更新するステップと、をさらに含む、請求項13に記載の方法。

【請求項 17】

前記セグメントパラメータが、第1タイムスタンプ及び第2タイムスタンプを示し、前記第1メディアセグメントの境界に対応するセグメントパラメータを受信するステップは、タイムラインに沿って前記メディアを進めるステップと、前記第2タイムスタンプで第2入力を受信するステップと、を含む、請求項13に記載の方法。

【請求項 18】

連続的なメディアセグメントに対応するセグメントパラメータを第1デバイスから受信するステップをさらに含む、請求項13に記載の方法。

【請求項 19】

連続的なメディアセグメントに対応するセグメントパラメータを前記第1デバイスから受信するステップは、タイムラインに沿って前記メディアを進めるステップを含み、第1入力の持続された受信中、

10

20

30

40

50

前記第 1 タイムスタンプで第 2 入力を受信するステップと、
前記第 2 タイムスタンプで第 3 入力を受信するステップと、
第 3 タイムスタンプで第 4 入力を受信するステップと、を含み、
前記第 1 メディアセグメントは、前記第 1 タイムスタンプと前記第 2 タイムスタンプとの間に定義され、第 3 メディアセグメントは、前記第 2 タイムスタンプと前記第 3 タイムスタンプとの間に定義される、請求項 18 に記載の方法。

【請求項 20】

遠隔メディアリポジトリから前記第 1 及び第 2 デバイスに前記メディアをストリーミングするステップをさらに含む、請求項 13 に記載の方法。

【請求項 21】

多言語でメディアに共同でキャプションを付ける方法であって、前記メディアは映像及び音声を含み、前記方法は、

メディアセグメントに関する第 1 言語の第 1 注釈を第 1 ユーザアカウントから受信するステップと、

前記メディアセグメントに関する前記第 1 言語のキャプションとしての前記第 1 注釈を保存するステップと、

第 2 ユーザアカウントで前記メディアセグメントに関する前記第 1 言語の前記キャプションをレンダリングするステップと、

前記メディアセグメントに関する前記第 1 言語とは異なる第 2 言語の第 2 注釈を前記第 2 ユーザアカウントから受信するステップと、

前記メディアセグメントに関する前記第 2 言語のキャプションとしての前記第 2 注釈を保存するステップと、を含む方法。

【請求項 22】

前記メディアセグメントに関する前記第 1 言語の第 1 注釈を前記第 1 ユーザアカウントから受信するステップの前に、前記メディアセグメントを定義するシステムパラメータを第 3 ユーザアカウントから受信するステップをさらに含む、請求項 21 に記載の方法。

【請求項 23】

前記メディアセグメントを定義するシステムパラメータを前記第 3 ユーザアカウントから受信するステップは、前記メディアセグメントに関する開始及び終了タイムスタンプ選択に対応するシステムパラメータを前記第 3 ユーザアカウントから受信するステップを含む、請求項 22 に記載の方法。

【請求項 24】

前記開始及び前記終了タイムスタンプ選択は、前記第 3 ユーザアカウントでレンダリングされる前記メディア音声ファイルの視覚表示で受信される、請求項 23 に記載の方法。

【請求項 25】

前記第 1、第 2 及び第 3 ユーザアカウントは、オンラインネットワークサービスの第 1、第 2 及び第 3 ユーザアカウントに関連付けられる、請求項 22 に記載の方法。

【請求項 26】

前記第 1 ユーザアカウントから前記メディアセグメントの選択を受信するステップの前に前記第 1 ユーザアカウントに前記メディアをストリーミングするステップと、前記第 2 ユーザアカウントから前記メディアセグメントの選択を受信するステップの前に前記第 2 ユーザアカウントに前記メディアをストリーミングするステップと、をさらに含む、請求項 21 に記載の方法。

【請求項 27】

前記第 1 注釈を受信するステップの前に前記第 1 ユーザアカウントから前記メディアセグメントの選択を受信するステップと、前記第 2 注釈を受信するステップの前に前記第 2 ユーザアカウントから前記メディアセグメントの選択を受信するステップと、をさらに含む、請求項 21 に記載の方法。

【請求項 28】

前記メディアセグメントの選択を受信するステップは、タイムスタンプに関連付けられ

10

20

30

40

50

たアイコンの選択を受信するステップを含み、前記タイムスタンプは、開始及び終了タイムスタンプで境界付けられるメディアセグメント期間内にある、請求項 27 に記載の方法。

【請求項 29】

第 3 ユーザアカウントで前記メディアセグメントに関する前記第 2 言語の前記キャプションをレンダリングするステップと、

前記メディアセグメントに関する前記第 2 言語の第 3 注釈を前記第 3 ユーザアカウントから受信するステップと、

前記メディアセグメントに関する前記第 2 言語の第 2 キャプションとしての前記第 3 注釈を保存するステップと、をさらに含む、請求項 21 に記載の方法。

10

【請求項 30】

前記メディアセグメントに関する前記第 2 言語の前記第 1 及び第 2 キャプションに関する分析を実行するステップと、

前記分析に基づいて前記第 1 キャプション又は前記第 2 キャプションを選択するステップと、

前記第 4 ユーザアカウントから受信した前記第 2 言語のキャプションの要求に応じて、選択された前記第 1 キャプション又は前記第 2 キャプションを第 4 ユーザアカウントでレンダリングするステップと、をさらに含む、請求項 29 に記載の方法。

【請求項 31】

前記第 2 言語の前記メディア及びキャプションを前記第 4 ユーザアカウントにストリーミングするステップをさらに含み、前記第 2 言語の選択された前記第 1 キャプション又は前記第 2 キャプションは、前記メディアセグメントの開始及び終了タイムスタンプの間で生じる映像フレームと同時にレンダリングされる、請求項 30 に記載の方法。

20

【請求項 32】

前記第 3 ユーザアカウントで前記メディアセグメントに関する前記第 2 言語の前記キャプションをレンダリングするステップの前に、前記第 3 ユーザアカウントについて前記第 2 言語を選択するステップをさらに含む、請求項 29 に記載の方法。

【請求項 33】

前記第 2 言語を選択するステップは、前記第 3 ユーザアカウントから言語選択を受信するステップを含み、前記言語選択は、前記第 2 言語に関連付けられたアイコンの選択を含む、請求項 32 に記載の方法。

30

【請求項 34】

前記第 3 ユーザアカウントで前記メディアセグメントに関する前記第 2 言語の前記キャプションをレンダリングするステップは、前記第 3 ユーザアカウントで前記メディアセグメントに関する前記第 1 言語の前記キャプションをレンダリングするステップを含む、請求項 29 に記載の方法。

【請求項 35】

映像及び音声を含むメディアに共同でキャプションを付ける方法であって、前記方法は、

メディアセグメントに関する第 1 言語のキャプションとしての第 1 注釈を第 1 デバイスから受信するステップと、

前記メディアセグメントに関する前記第 1 言語の前記キャプションとしての前記第 1 注釈を第 2 デバイスでレンダリングするステップと、

前記メディアセグメントに関する前記第 1 言語の前記キャプションとしての第 2 注釈を前記第 2 デバイスから受信するステップと、

前記メディアをストリーミングするための第 3 デバイスからの要求に応じて、前記映像、前記音声及び前記第 2 注釈を前記第 3 デバイスにストリーミングするステップと、を含む方法。

40

【請求項 36】

前記メディアセグメントに関する第 2 言語のキャプションに関する注釈を第 4 デバイス

50

から受信するステップをさらに含み、前記第 2 言語は前記第 1 言語とは異なる、請求項 3 5 に記載の方法。

【請求項 3 7】

メディアをストリーミングするための第 5 デバイスからの要求を受信するステップに応じて、前記第 1 及び第 2 言語から前記第 5 デバイスについてキャプション言語を選択するステップと、前記映像、音声及び選択された前記言語の前記キャプションを前記第 5 デバイスにストリーミングするステップと、をさらに含む、請求項 3 6 に記載の方法。

【請求項 3 8】

前記第 5 デバイスについてのキャプション言語を選択するステップは、前記第 5 デバイスに関連付けられたユーザプロファイルにアクセスするステップと、前記ユーザプロファイルに関連付けられた言語を選択するステップと、を含む、請求項 3 7 に記載の方法。

10

【請求項 3 9】

前記第 1 デバイスについてのキャプション付け言語として前記第 1 言語を選択するステップをさらに含む、請求項 3 5 に記載の方法。

【請求項 4 0】

前記メディアの前記音声は第 2 言語のものであり、前記第 1 デバイスについて前記第 1 言語を選択するステップは、前記第 1 デバイスに関連付けられたユーザプロファイルにアクセスするステップと、前記第 1 言語及び前記第 2 言語が、前記ユーザプロファイルに関連付けられた言語であることを決定するステップと、前記第 1 デバイスについての前記キャプション付け言語として前記第 2 言語を選択するステップと、を含む、請求項 3 9 に記載の方法。

20

【請求項 4 1】

前記メディアセグメントに関する第 1 言語のキャプションとしての第 1 注釈を第 1 デバイスから受信するステップの前に、前記メディアセグメントに対応するセグメントパラメータを第 4 デバイスから受信するステップをさらに含む、請求項 3 5 に記載の方法。

【請求項 4 2】

前記メディアセグメントに対応するセグメントパラメータを第 4 デバイスから受信するステップは、第 2 メディアセグメントに対応するセグメントパラメータを前記第 4 デバイスから受信するステップをさらに含み、前記第 2 メディアセグメントは前記第 1 メディアセグメントとは異なる、請求項 4 1 に記載の方法。

30

【請求項 4 3】

前記第 2 メディアセグメントに関する前記第 1 言語のキャプションに関する注釈を第 4 デバイスから受信するステップをさらに含む、請求項 4 2 に記載の方法。

【請求項 4 4】

前記第 2 メディアセグメントに関する前記第 1 言語の前記キャプションは、前記第 1 メディアセグメントに関する前記第 1 言語の前記キャプションの受信とほぼ同時に受信される、請求項 4 3 に記載の方法。

【請求項 4 5】

前記第 2 デバイスから前記第 2 注釈を受信するステップの前に、前記第 1 言語の前記映像、音声及びキャプションを前記第 2 デバイスにストリーミングするステップをさらに含む、請求項 4 5 に記載の方法。

40

【請求項 4 6】

前記第 1 言語の前記映像、音声及びキャプションを前記第 2 デバイスにストリーミングするステップとほぼ同時に、前記第 1 言語の前記映像、音声及びキャプションを第 4 デバイスにストリーミングするステップをさらに含み、前記第 4 デバイスにストリーミングされた前記第 1 言語の前記キャプションは、前記第 2 デバイスから受信された前記第 2 注釈によって動的に更新される、請求項 4 5 に記載の方法。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

50

[0001] 本発明は、概して、メディアキャプション付けの分野に関し、より具体的には、メディアキャプション付けの分野においてメディアキャプションを翻訳する新規で有用なシステム及び方法に関する。

【背景技術】

【0002】

[0002] オンライン映像ストリーミングの流行によって、人々は、世界中で創作された映像及びメディアにアクセスしている。以前では、映像は時には、1カ国又は1地域のみでの消費のために創作されていた。現在では、世界中の人々が、他国で作られたコンテンツに触れて当該コンテンツを消費している。インターネットが広範な分配を可能にするにも関わらず、メディアコンテンツに対するアクセスは言語の壁を通して依然として制限されている可能性がある。対象の視聴者によって理解される言語にメディアが翻訳されない場合、映像が視聴される可能性は低い。メディアを別の言語に翻訳するには多くの時間が必要でありまた困難なプロセスであり得る。ある翻訳が捉えることができない言語のニュアンスがたくさんある。従って、キャプション付けの分野において、メディアキャプションを翻訳する新規で有用なシステム及び方法を創出することが求められている。本発明は、そうした新規で有用なシステム及び方法を提供する。

10

【図面の簡単な説明】

【0003】

【図1】図1は、システムのメディアキャプション付けの一変形例の概略図である。

【図2】図2は、システムのメディアキャプション付けの第2変形例の概略図である。

20

【図3】図3は、システムのメディアセグメント化の変形例の概略図である。

【図4】図4は、システムのメディアセグメント化の変形例の概略図である。

【図5】図5は、システムのメディアセグメント化の変形例の概略図である。

【図6】図6は、メディアにキャプションを付ける方法のフローチャートである。

【図7】図7は、メディアにキャプションを付ける方法の一例の概略図である。

【図8】図8は、第1メディアセグメントに関するキャプションを選択するために注釈を処理する一例の概略図である。

【図9】図9は、一変形例の方法の概略図である。

【図10】図10は、複数のユーザによる複数のメディアセグメントの共同編集を含む、一変形例の方法の概略図である。

30

【図11】図11は、提示されたメディアを動的に更新することを含む、一変形例の方法の概略図である。

【図12A】図12Aは、多言語でキャプションを生成することを含む第1及び第2変形例の方法の概略図である。

【図12B】図12Bは、多言語でキャプションを生成することを含む第1及び第2変形例の方法の概略図である。

【図13】図13は、メディアをセグメント化する方法のフローチャートである。

【図14A】図14Aは、第1ユーザがメディアをセグメント化して第2ユーザがメディアにキャプションを付ける、一変形例の方法の概略図である。

【図14B】図14Bは、第1ユーザがメディアをセグメント化して第2ユーザがメディアにキャプションを付ける、一変形例の方法の概略図である。

40

【図15】図15は、一変形例の方法の概略図である。

【図16】図16は、メディアセグメントの音声及び映像とは別個の第1注釈を受信する一例である。

【図17A】図17Aは、メディアセグメントの音声及び映像とは別個の第2注釈を受信する第1例及び第2例である。

【図17B】図17Bは、メディアセグメントの音声及び映像とは別個の第2注釈を受信する第1例及び第2例である。

【発明の詳細な説明】

【0004】

50

[0018] 本発明の好適な実施形態の以下の説明は、本発明をそれら好適な実施形態に限定することを意図しておらず、当業者が本発明を行って使用することができるようにすることを意図している。

【0005】

1. 共同メディアキャプション付け及びセグメント化のためのシステム

[0019] 図1に示すように、メディアキャプションを生成するシステムは、キャプションストリームインターフェース114に制御可能に結合されたメディアプレーヤ112を含むメディアインターフェース110を含んでもよい。このシステムは、容易かつ効果的なインターフェースを形成してメディアファイルに関するキャプションを翻訳するように機能する。システムは、複数のアカウント実体の参加を通じてメディアファイルに多言語で字幕を付けるのに通常は使用される。メディアプレーヤ112でのメディアの再生は、キャプションストリームインターフェース114と協働して行われることが好ましい。システムは、1以上のセットのキャプションを生成することが好ましく、各キャプションは、異なる言語のものであることが好ましい。キャプションのセットは、(例えば図1に示すように)メディアから生成されてもよく、又は、(例えば図2に示すように)別のキャプションセットから生成(例えば翻訳)されてもよい。初期のキャプションは、メディアと組み合わせた参照として使用され、少なくとも第2セットのキャプションに関する翻訳/キャプションを収集する。システムは、複数アカウントのオンラインプラットフォーム130を通じて字幕付けを可能にすることが好ましい。オンラインプラットフォームは、ウェブサイト又は他の適切なアプリケーションを通じてアクセス可能なネットワークアクセス可能サービスであることが好ましい。複数のユーザが字幕付けに参加することができるようにすることによって、メディアに字幕付けすることの品質、効率及びスケールを向上させることができる。システムは、映像ストリーミングプラットフォームで実施されてもよい。通常の実用アプリケーションでは、映像ストリーミングプラットフォームは、多数の映像及び世界的に又は文化的に多様な視聴率を有することが可能であり、従って、多様な視聴者に対してアピールするために多数の映像の現地語化/翻訳を必要とする。従業員、ボランティア、たまたま視聴している視聴者、かなりの言語知識を有する視聴者、限られた言語知識しか有しない視聴者がすべて、メディアのキャプション付け及び/又は翻訳に参加することができる。好適な一実施形態では、キャプションストリームインターフェース114は、通常の実用プレーヤの選択可能な動作モードとして構成されてもよい。映像を観ている時、視聴者は、キャプションストリームインターフェース114を選択的に起動させることができ、かつ、メディアの全部又は一部の翻訳に寄与することができる。

【0006】

[0020] インターフェース110は、メディアの消費のために構成され、かつ、映像の少なくとも1つのセグメントをキャプチャするために構成されたインターフェースである。インターフェース110は、ウェブアプリケーション技術を通じて実装されることが好ましく、また、ウェブサイト内でレンダリングされることが可能である。インターフェース110は、代替的に、パーソナルコンピュータデバイス、モバイルコンピュータデバイス又は任意の適切なデバイス上で動作可能なネイティブアプリケーションコードを通じて実装されることが可能である。上述したように、インターフェース110は、キャプションストリームインターフェース114に制御可能に結合された少なくともメディアプレーヤ112を含むことが好ましい。

【0007】

[0021] メディアプレーヤ112は、ユーザが消費するためのメディアを再生するように機能する。メディアプレーヤ112は、例えば再生/一時停止ボタン、早送りボタン、巻戻しボタン、次チャプタボタン、前チャプタボタン、ボリューム調節、再生進行バー、再生速度調節、メディアオプションメニュー及び/又は任意の適切なメディア調節などの従来のメディアプレーヤ制御を含むことが好ましい。メディアプレーヤは、映像ファイルを再生し、オペレーティングシステムによって供給される音声チャンネルを通じて音声を出力する映像プレーヤであることが好ましい。メディアプレーヤ112は、図1に示すキャプ

10

20

30

40

50

ションストリームインターフェース 114 に隣接していることが好ましいが、メディアプレーヤ 112 及びキャプションストリームインターフェース 114 は、代替的に、任意の適切な位置レイアウト又はナビゲーションレイアウトで配列及びレンダリングされてもよい。

【0008】

[0022] キャプションストリームインターフェース 114 は、メディアセグメントにキャプションを付けるためのインターフェースを提供するように機能する。キャプションストリームインターフェースは、複数のキャプション入力 115 のスクロール可能なリスト表示であることが好ましい。キャプション入力 115 は、キャプション入力 116 を含むことが好ましく、また、(例えば異なる言語の)少なくとも1つの参考キャプションを含んでもよい。キャプション入力 115 はメディアセグメントに関連付けられることが好ましい。メディアセグメントは、メディアの特定の時間ウィンドウに関連付けられて設定キャプションコンテンツ(例えばスピーチの翻字及び/又は翻訳)を有するキャプションセットの一部であることが好ましい。キャプション入力 116 はさらに、カスタマイズされたキャプション位置決め、書式設定及び他の適切なキャプションの態様を可能にすることができる。キャプション入力 115 はさらに、キャプション採点、フラグ立て、コメント付け、編集、批評、又は、キャプション付けに関連付けられた任意の適切な動作を可能にするための他のユーザインターフェースコンポーネントを含んでもよい。キャプションストリームスクロール位置及びキャプション入力選択はメディアプレーヤ 112 に制御可能に結合されることが好ましい。キャプションストリームインターフェース 114 は、メディアの現在の再生位置のキャプション入力を示すためにスクロールすることが好ましく、また反対に、メディアの再生位置は、キャプションストリームインターフェース 114 の現在のスクロール位置又は現在選択されているキャプション入力 115 に対応するために変化してもよい。キャプションストリームインターフェース 114 は、追加的に、多くのモードで選択的に表示されてもよい。第1モードでは、キャプションストリームインターフェース 114 は実質的に非表示にされることが可能であり、また、メディアプレーヤ 112 はフレーム内に全幅で表示される。第2モードでは、フレーム/ウィンドウは、キャプションストリームインターフェース 114 及びメディアプレーヤ 112 の両方を実質的に見ることができ、また、その両方が実質的に完全に機能するように、キャプションストリームインターフェース 114 及びメディアプレーヤ 112 の間で分割される。第3モードでは、キャプションストリームインターフェース 114 は全フレーム/ウィンドウモードで表示されてもよい。第3モードでは、メディアプレーヤは、非表示にされるか、又は、実質的に小さなウィンドウ(例えば400×400ウィンドウ未満)で再生されてもよい。第3モードはさらに、例えばスペル/文法チェック、用語集、チャットウィンドウ、修正履歴、又は他のツールなどの他のキャプション編集ツールを可能にすることができる。任意の適切な数の視聴モード及び特徴構成を代替的に用いてもよい。

【0009】

[0023] システムはさらにキャプションファイル 120 を含むことが好ましい。キャプションファイルはデータオブジェクトであってもよく、又は代替的に、データモデルで特徴付けられてもよい。キャプションファイル 120 は、メディアファイルに関する少なくとも1セットのキャプションを特徴付けるように構成されることが好ましい。好ましくは少なくとも1つの参考キャプション及び/又は翻訳されたキャプションがキャプションファイル 120 に含まれる。キャプションは、複数のメディアセグメントに分割されることが好ましいが、代替的に、開始時間、終了時間、映像フレーム、キーフレーム及び/又はシーンに各々関連付けられたキャプションのリストを含んでもよい。メディアセグメントは、別のメディア(例えば映像)に対して又は別のメディア上にキャプションテキストを表示するように同期される、開始時間と終了時間との間のメディアの一部であることが好ましい。各メディアセグメントは、キャプションファイル 120 の1以上のキャプションに関連付けられることが好ましい。例えば編集履歴、投稿しているアカウント又はソース、翻訳スコア又は任意の適切なキャプション関連情報などのキャプションの他の態様は、追

10

20

30

40

50

加的に、キャプションファイル 120 内で特徴付けられてもよい。キャプションファイル 120 は、メディア再生中のキャプションを表示すること及び/又はキャプションストリームインターフェース 114 を生成することに用いられてもよい。キャプションファイル 120 は、キャプションストリームインターフェース 114 内でキャプション入力 115 の編集を通じて形成され又は更新されてもよい。キャプションファイル 120 は、各メディアセグメントのタイミング及びコンテンツを特徴付けるソフトキャプション付けファイルであることが好ましい。他の変形例では、キャプションファイル 120 は、ハードキャプションファイル(例えば、適切な映像フレームに融合されたキャプションテキストを有する映像ファイル)であってもよく、予めレンダリングされたキャプションファイル(例えば、テキストの映像フレームが映像の先頭に重ねられてもよい)であってもよく、又は、任意の適切なフォーマットであってもよい。

10

【0010】

[0024] 図3~図5に示すように、システムは、追加的に、音声レイヤセグメント化インターフェース 117 を含んでもよい。メディアプレーヤは、音声レイヤセグメント化インターフェース 117 に制御可能に結合されることが好ましい。このシステムは、メディアファイルにキャプションを付けるために容易かつ効率的なインターフェースを形成するように機能する。メディアプレーヤ 112 でのメディアの再生は、セグメント化インターフェース 117 と協働して行われることが好ましい。ユーザ入力を通じて、音声の一部は、セグメント化インターフェース 117 でのキャプション付けのために明確に記述され、メディアプレーヤ 112 でキャプションが追加されてプレビューされる。システムは、メディアでのキャプションの表示のタイミングを調整するスロープロセスに対処することが好ましく、また、システムはキャプションの入力を許容する。通常、キャプション付けは最初は母国語で実行され、また、キャプションの翻訳はその後、メディアに第2言語で字幕を付ける時に追加されてもよい。システムは、映像及び音声部分を含む映像のキャプション付けのために構成されることが好ましいが、システムは代替的に、例えば音声のみ、インタラクティブメディア他などの任意の適切な形態のメディアのために構成されてもよい。システムは、メディアのネットワーク視聴及びキャプション付けを可能にする複数アカウントのオンラインプラットフォーム 130 をさらに含んでもよい。オンラインプラットフォーム 130 は、ウェブサイト又は他の適切なアプリケーションを通じてアクセス可能なネットワークアクセス可能サービスであることが好ましい。好適な一実施形態では、システムは、通常の映像プレーヤの選択可能な動作モードとして構成されてもよい。映像を観る時、視聴者は、セグメント化インターフェース 117 を選択的に起動して、メディアの全体又は一部にキャプションを付けることができる。

20

30

【0011】

[0025] 音声レイヤセグメント化インターフェース 117 は、音声信号の視覚表示及びメディアをセグメント化するための入力インターフェースである。メディアの音声信号は、メディアプレーヤ 112 の下部のフレームに表示されたメディアタイムフレームの関数としての波形として視覚的に表示されることが好ましい。しかしながら、音声信号は、メディアプレーヤ 112 に対して任意の他の適切な位置に表示されてもよい。音声信号は代替的に、例えば対話のプレゼンスの二元表示、2以上の音声チャンネル、複数のスピーカチャンネル、又は、任意の適切な表示などの任意の適切な形態で表示されてもよい。音声信号の視覚表示によって、キャプション付けのためのメディアセグメントの容易な識別及び選択を許容する。音声レイヤセグメント化インターフェース 117 は、音声レイヤのフレームに水平にわたった音声信号表示を進展させるように構成される。音声信号表示は、メディアプレーヤでのメディアの再生に同期して進められ、その結果、セグメント化インターフェース 117 での音声レイヤ表示は、メディアの再生位置で及び再生位置の周囲での音声の反射である。メディアプレーヤ 112 での映像及び音声の再生が通常は単一の再生位置でのメディア表示のみを描写するのに対して、音声信号表示は、少なくとも時間の次元に沿って音声を提示する。音声の時間表示は、現在の映像再生位置上に中心を置くことが好ましい。

40

50

【 0 0 1 2 】

[0026] 音声レイヤセグメント化インターフェース 117 は、追加的に、ユーザセグメント化入力を受信するために構成される。通常、ユーザ入力は、音声信号の一部の選択を許容する。選択は、メディアセグメントの時間ウィンドウ生成を定義し、メディアセグメントの開始及び終了は、音声波形の一部を強調することによって割り当てられる。通常、ユーザは、最初に開始又は終了時間マーカを記して、その後、メディアセグメントの全時間を識別するために強調されたメディアセクションを拡張するために、マウス、タッチ入力、キーボード制御ポジショナ、又は、他の入力デバイスを用いる。セグメントインターフェース要素は、メディアセグメント記述の容易な調節のために構成されたコンポーネントを含んでもよい。追加的に又は代替的に、メディアのセグメント化及びキャプション付けの一部又は全部は、キーボードショートカットを通じて実行されてもよい。音声レイヤセグメント化インターフェース 117 の入力インターフェースコンポーネントは、自動的な又は半自動的なセグメント化ツールを代替的に含んでおり、また、音声レイヤのセグメントを記述するための任意の適切な代替の変形例を含んでもよい。予めセグメント化された部分は、キャプションを付けた映像の音声信号表示の進行時に適切に継続して提示される。

10

【 0 0 1 3 】

[0027] 好適な一実施形態では、システムは、メディアキャプション付けとの複数のアカウントの共同を運用するように機能する複数アカウントオンラインプラットフォーム 130 を含む。プラットフォーム 130 は、分散又はクラスター演算インフラストラクチャをホストとしていることが好ましい。プラットフォーム 130 は、ブラウザベースのウェブアプリケーションとしてインターフェース 110 をレンダリングするために必要な命令を通信することが好ましいが、代替的にネイティブアプリケーション内でインターフェース 110 をレンダリングするための命令及びデータを送達するように構成されてもよい。プラットフォーム 130 は、アカウント運用システムを含むことが好ましい。アカウントプラットフォーム 130 上のアカウントプロフィールは、言語習熟評価、キャプション入力履歴、映像視聴履歴、及び任意の適切な態様を追跡する。アカウント運用システムは、キャプションに対する変更及び編集を強制するために使用されることが可能である。

20

【 0 0 1 4 】

2. 共同メディアキャプション付け方法

[0028] 図 6、図 7 及び図 9 に示すように、メディアに共同でキャプションを付ける方法は、メディアセグメントに関するキャプションとしての第 1 注釈を第 1 ユーザから受信するステップ S 100 と、メディアセグメントに関するキャプションとしての第 2 注釈を第 2 ユーザから受信するステップ S 200 と、メディアセグメントに関するキャプションをユーザに提示するステップ S 300 と、を含む。第 1 及び第 2 注釈は、同一の言語のものであることが好ましく、メディアセグメントに関するキャプションは、メディアセグメントに関する第 1 言語のキャプションであることが好ましい。方法は、追加的に、メディアセグメントを定義するセグメントパラメータを第 3 ユーザから受信することを含んでもよい。方法は、キャプションの共同創作及び改善を可能にする。メディアキャプション付け（字幕付け）及び / 又はメディア分割の困難な仕事をクラウドソーシングすることによって、この方法は、消費者のためのキャプションを迅速に生成する方法を提供する。さらに、生成されたキャプションは、オンラインキャプション付けサービスのユーザが簡単に査読して任意の誤ったキャプションを編集することができるため、高品質なキャプションになる傾向にある。

30

40

【 0 0 1 5 】

[0029] さらに、第 1 及び第 2 注釈が音声言語とは異なる言語のものである場合、方法は、クラウドソースメディアのキャプション付けだけでなく、メディアを第 2 言語に迅速に翻訳するようにも機能する。メディアを第 2 言語に速やかに確実に翻訳することによって、この方法は、メディアプロバイダ（例えば、制作会社、メディアホストサービス等）が、第 2 言語のコンテンツを消費するユーザの他のマーケットに迅速に拡大することを可能

50

にする。

【0016】

[0030] 方法は、メディアをストリーミングするために用いられることが好ましく、キャプションは、ユーザデバイスに、映像及び音声とともにストリーミングされる。これは、新たなキャプション及び/又は編集がキャプションファイルに追加されるので、ストリーミングされたキャプションの動的更新を許容する。従って、キャプション発行者からの新たなキャプションバージョンを待つ代わりに、この方法は、メディア消費者がコンテンツを消費している際、更新されたキャプションバージョンをメディア消費者が実質的に即座に/同時に受信すること(及び理解又は文脈を深める利益を得ること)を許容する。メディアのストリーミングは、ファイル全体が送信される前にメディアが再生されることができ、絶えず受信されるマルチメディアであり、かつ、プロバイダによって配信されている間にユーザに提示されるマルチメディアであることが好ましい。メディアは、マルチキャストプロトコルを用いてストリーミングされることが好ましいが、代替的に、任意の適切なプロトコルを用いてストリーミングされてもよい。

10

【0017】

[0031] 方法は、メディアセグメントに関するキャプションとしての第2言語の第1注釈を第3ユーザから受信するステップを含み、第2言語は第1言語とは異なる。メディアに関する多言語でのクラウドソーシングキャプション生成によって、この方法は、消費者のための複数のキャプション言語オプションを速やかかつ確実に提供する。

20

【0018】

[0032] 方法は、ウェブサイト上のメディアにキャプションを付けるために実行されることが好ましいが、代替的に、ネイティブアプリケーション又は任意の他の適切なインターフェースで用いられてもよい。方法は、ストリーミングされたメディアにキャプションを付けるために用いられることが好ましく、メディアは、オンラインメディアプロバイダからキャプション付けデバイスにストリーミングされる。しかしながら、方法は、記憶されたメディアにキャプションを付けるために用いられるか、又は、任意の他の適切な形態のメディアにキャプションを付けるために用いられてもよい。メディアは、例えば映画、テレビ番組又はアニメーションなどの映像であることが好ましいが、音声、番組間のつながり、スクリーンキャスト、インタラクティブメディア及び/又は任意の適切なメディアに用いられてもよい。メディアは、録画されたメディアであることが好ましいが、代替的にライブのメディア(例えばイベントの生放送)であってもよい。方法は、メディアと同一の言語への映像のキャプション付けに用いられるか、又は代替的に、少なくとも第2言語でのメディアへの字幕付けに用いられてもよい。キャプション付けは、話された言葉に関して実行されることが好ましいが、代替的に、テキストの翻訳、音響効果のような音、音楽、補足情報の描写、及び/又は、キャプション付けの任意の適切な使用に用いられてもよい。

30

【0019】

[0033] メディアは、メディアタイムラインに沿って同期される映像ファイル及び音声ファイルを含むことが好ましい。映像ファイルは、複数の映像フレームを含むことが好ましいが、代替的に、複数のキーフレームを含んでもよく、デバイス上のメディアプレーヤはキーフレーム同士の間フレームを挿入する。メディアは、追加的に、メディアタイムラインに沿って映像及び音声ファイルに同期されるキャプションファイル(例えば、クローズドキャプション)を含んでもよい。この方法による処理の前に、メディアはキャプションを有していなくてもよく、方法がメディアに関するキャプションを生成する。代替的に、メディアは、少なくとも1セットのキャプションを有してもよく、また、複数セットのキャプションを有することがさらに好ましい(例えば、各セットのキャプションは異なる言語のものである)。メディア上でこの方法が実行された後、メディアは、1セット以上のキャプションに関連付けられることが好ましく、各セットのキャプションは、異なる言語のものであることが好ましい。複数セットのキャプションは、別個のキャプションファイルとして記憶されてもよく、包括的なキャプションファイルのサブセットとして記

40

50

憶されてもよく、又は、任意の他の適切な方法で記憶されてもよい。各セットのキャプションは、キャプションが翻字である音声に同期される映像フレーム上にキャプションが現れるように、メディアの映像及び音声に同期されることが好ましい。

【0020】

[0034] 各セットのキャプションは、複数のキャプションを含むことが好ましく、各キャプションは、キャプションファイル内の注釈（キャプション）並びに開始タイムスタンプ及び終了タイムスタンプに関連付けられることが好ましい。代替的に、各キャプションは、開始タイムスタンプ及び表示期間に関連付けられてもよく、開始映像フレーム及び終了映像フレームに関連付けられてもよく、開始音声信号及び終了音声信号に関連付けられてもよく、又は、任意の他の適切なキャプションデータ構造を含んでもよい。タイムスタンプ及び/又は期間は、キャプションが意図されるメディアセグメント/セクションを定義することが好ましい。各キャプションは、追加的に、複数のバージョンを含んでもよく、各バージョンは、異なる時間（例えば、リアルタイム、タイムスタンプではない）に受信される、又は、異なるユーザから受信されることが好ましい。各キャプションバージョンはメタデータを含むことが好ましく、メタデータは、キャプションのバージョンをサブミットしたユーザ、バージョンがサブミットされた時間、デバイス識別子、地理的位置識別子（例えば、バージョンをサブミットした時間におけるユーザ位置）を識別するユーザ識別子を含むメタデータ又は任意の他の適切なメタデータを含むことが好ましい。

10

【0021】

[0035] メディアセグメントは、開始及び終了タイムスタンプの間の映像又は音声ファイルの一部であることが好ましい。メディアは、システムのユーザによってメディアセグメントにセグメント化されることが好ましいが、代替的に、以下に説明するように、システムによってセグメントに自動的にセグメント化されてもよい。メディアセグメントは、メディアタイムラインに沿って重複しないことが好ましいが、代替的に重複してもよい。

20

【0022】

[0036] 動作時、キャプションは、開始タイムスタンプと終了タイムスタンプとの間の映像フレーム上に、又は、最初のタイムスタンプから開始する期間中の映像フレーム上に表示される（例えば重ね合わせられる）ことが好ましい。代替的に、キャプションは、映像ファイルの映像フレーム内にハードコード化されてもよい。キャプションセット内のキャプションは、重複する期間（例えば連続的である）を有しないことが好ましいが、代替的に、異なるスピーカのための異なるキャプションの場合に、重複する期間を有してもよい。代替的に、キャプションは、メディアをダビングする又はレクタリングする（lectoring）（例えば、吹き替え）のためのスクリプト又はソースとして用いられてもよい。ダビング又はレクタリングは、機械テキスト音声変換方法などを通じて自動的に生成されてもよく、又は、手動で生成されてもよく、ユーザ（第1及び第2ユーザとは異なるが、代替的に任意の適切なユーザ）は、メディアセグメント又はメディアの任意の他の適切な一部をダビング又はレクタリングするためのスクリプトとしてキャプションを用いる。

30

【0023】

[0037] メディアは、メディアレポジトリ（例えばサーバ）に記憶されることが好ましいが、代替的に、任意の他の適切な記憶手段に記憶されてもよい。メディアは、ブラウザ又はネイティブアプリケーションを支持するデバイスにストリーミングされることが好ましいが、代替的に、デバイス上に記憶されてもよい。方法は、キャプション付けネットワークサービス、メディアホストサービス、又は、任意の他の適切なメディア又はキャプション提供者によって複数アカウント共同システムで実行されることが好ましい。メディアは、オンラインキャプション付けサービスによって記憶され、オンラインキャプション付けサービスから受信デバイスにストリーミングされることが好ましく、キャプション付けサービスは、ストリーミングされたメディアに関するキャプションファイルをさらに生成することが好ましい。

40

【0024】

[0038] 代替的に、メディアは、キャプション付けサービスとは別個のオンラインメディ

50

アホストサービス（例えば YouTube）によって記憶されて当該オンラインメディアホストサービスからストリーミングされてもよく、キャプション付けサービスは、インターフェースを提供して、ストリーミングされたメディアにキャプションを付け、かつ、キャプションファイルを生成する。オンラインキャプション付けサービス及び／又はメディアホストサービスは複数のオンラインユーザアカウントを含むことが好ましく、オンラインユーザアカウントの各々は、ユーザプロフィール及びユーザ生成ポスト（例えば電子メッセージ）のコンテンツストリームに関連付けられることが好ましい。ユーザプロフィールは、ユーザ生成タイプであってもよく、又は、ユーザアカウントコンテンツストリームから、ユーザアカウントのつながりから、ユーザデバイスから、又は、ユーザアカウントに関連付けられた任意の他の適切な情報から、自動的に導出されてもよい。オンラインキャプション付けサービス及び／又はメディアホストサービスは、インターネットを通じて受信デバイスに情報（例えば映像、音声又はキャプション情報）をストリーミングすることが好ましいが、代替的に、近距離技術又は任意の他の適切なデータ伝送スキームを通じて受信デバイスにまとめて情報をストリーミング又は送信（例えばメディア再生に先立ってファイル全体を送信）してもよい。同様に、オンラインキャプション付けサービス及び／又はメディアホストサービスは、インターネットを通じてユーザデバイスからデータを受信することが好ましい。データ（例えば、キャプションに関する注釈）は区分的に受信されることが好ましく、各時間データは、メディアセグメントごとに受信されるが、代替的に、まとめて（例えばメディアの全体に関する注釈）受信されてもよい。

10

20

【0025】

[0039] 方法の一変形例では、キャプション付けサービスは、別個のメディアホストサービスからメディアをストリーミングするユーザインターフェース（例えばブラウザウィンドウ）を提供し、キャプション付けサービスは、メディアセグメント化してキャプション付けするデータを受信してキャプションファイル内に処理する。結果として得られるキャプションファイルは、次に、元のメディアとの同期化及び／又はその後のストリーミングのためにメディアホストサービスに送信され、又は、キャプション付けサービスサーバに記憶されてもよく、メディアを視聴するためのユーザリクエストは、キャプション付けサービスにユーザにキャプションをストリーミングすることを指示してもよい。

【0026】

[0040] メディアは、ユーザデバイス上のプレーヤインターフェース上で再生可能であることが好ましく、プレーヤインターフェースは、メディア視聴、リスニング及び／又は体験するためのインターフェースを提供するように機能する。プレーヤインターフェースは、上記のシステムで説明されたものと実質的に同様であることが好ましいが、プレーヤインターフェースは、メディアを消費する任意の適切なインターフェースであってもよい。プレーヤインターフェースは、メディアセグメント選択及び注釈入力を可能にすることが好ましく、及び追加的に、一時停止／再生、早送り、巻戻し、他のチャプタへの前後スキップ、ボリューム設定の変更及び／又は他のメディア再生パラメータの調節などの他の適切なメディアツールのためのインターフェースを提供してもよい。

30

【0027】

[0041] メディアはメディアプレーヤ上でさらに再生可能であることが好ましく、プレーヤインターフェースはメディアプレーヤを含んでもよい。メディアプレーヤは、一時停止／再生、早送り、巻戻し、複数の通常の再生速度での再生、メディアタイムラインに沿った前後スキップ、及び／又は、他のメディア再生パラメータの調節のためのインターフェースを提供することが好ましい。メディアプレーヤは、追加的に、クローズドキャプション選択を可能にすることが好ましく、キャプションは、クローズドキャプションアイコン選択に依存して選択的にオン／オフされてもよい。キャプションファイルが、ストリーミングされる、及び／又は、クローズドキャプションアイコン選択状態とは無関係にメディアの残りを提供されることが好ましい一方で、キャプションファイルは、クローズドキャプションアイコン選択状態に基づいて選択的にストリーミングされてもよい（例えば、クローズドキャプションアイコンがオンの時にストリーミングされ、

40

50

クローズドキャプションアイコンがオフの時にはストリーミングされない)。

【0028】

[0042] キャプション及び/又は音声の言語は、自然な人間の言語又は特定の言語体系であることが好ましい。言い換えれば、言語は、意味を伝えるための組み合わせの文法規則によって支配されたサインの形式体系であることが好ましい。しかしながら、言語は、コミュニケーションの任意の他の適切な形態であってもよい。言語は文語又は口語であってもよい。言語は、一般的用法(例えば死語ではない)の言語であることが好ましいが、代替的に、廃れた言語、祖語又は任意の他の適切な言語であってもよい。言語の例は、日本語、中国語、英語、ドイツ語、ロシア語、フランス語及びスペイン語を含む。言語は、系統樹に関係することが好ましく、系統樹内の言語は、言葉、語構造、単語発音、文字原稿、文法的構造などの共通の特性、又は、任意の他の適切な特性を共有する。例えば、英語は、ゲルマン語族の一部としてドイツ語に関係する。日本語は、漢字が多く用いられていることから中国語に関係すると考えられ得る。動作時、第1言語のキャプションは、関係する言語(例えば文語又は口語の関係による)のキャプションの生成に用いられてもよい。

10

【0029】

[0043] 各言語は、システム内の共通性ランキングに関連付けられてもよく、共通性ランキングは言語の人気又は流行を示し(例えば相関し)得る。例えば、英語は、システム内において高いランキングにあり得る一方で、チェロキー語は低いランキングにあり得る。この共通性ランキングは、ユーザアカウント(例えば、ユーザプロフィール内に記載された又はユーザアカウントがサブタイトルのための所定の言語を選択する)に関連付けられた言語に基づいて決定されてもよく、高い出現頻度を有する言語には高い共通性ランキングが付与される。各言語は、追加的に、第2言語に関する関連性ランキングを有しており、関連性ランキングは、第1及び第2言語の専門知識(例えば、初級、中級、堪能なレベルで第1及び第2言語の両方を話す、読む、書く)を有するユーザの数を示し得る。例えば、システムのさらに多くのユーザが、日本語及びロシア語の両方よりも日本語及び英語の両方の方でより堪能である場合、日本語及び英語の組み合わせに関する関連性ランキングは、日本語及びロシア語の組み合わせよりも高いことが好ましい。しかしながら、各言語又は各群の言語は、任意の他の適切なパラメータに基づいて任意の他の適切なランキングを付与されてもよい。

20

30

【0030】

[0044] メディアセグメントに関するキャプションとしての第1注釈を第1ユーザから受信するステップS100は、メディアセグメントに関するキャプションを作成又は編集するように機能する。第1注釈は、第1ユーザに関連付けられた第1デバイスから受信されることが好ましいが、代替的に/追加的に、第1デバイス又はユーザに関連付けられた第1ユーザアカウントから受信されてもよく、又は、第1ユーザに実質的に固有にマッピングする任意の他の適切な構造から受信されてもよい。第1注釈は、第1ユーザによってキャプション入力フィールドに入力されることが好ましいが、代替的に、ユーザによって選択された推薦された注釈(例えば、音声テキスト変換方法を用いて自動的に生成される、又は、スクリプトから抽出される)であってもよい。第1注釈は、テキストであることが好ましいが、代替的に、アイコン選択、リンク又は任意の他の適切な注釈であってもよい。メディアセグメントに関する注釈は、他のメディアセグメントに関する注釈とは無関係に受信されることが好ましいが、代替的に、他のメディアセグメントに関する注釈とともに受信されてもよい。

40

【0031】

[0045] 第1注釈は第1言語のものであることが好ましく、キャプション(例えば、キャプションセット)は、メディアセグメントに関する第1言語のキャプションであることが好ましい。第1言語は、音声言語であってもよく、又は、音声言語とは異なる言語であってもよい。第1注釈を受信するステップは、追加的に、第1注釈の言語を選択するステップを含んでもよい。注釈の言語を選択するステップは、第1注釈の受信に先立って、ユー

50

ザ、ユーザデバイス又はユーザアカウントから言語選択を受信するステップを含むことが好ましい。言語選択は、言語に関連付けられたアイコンの選択、ドロップダウンリストからの言語の選択、又は、任意の他の適切な言語選択であってもよい。代替的に、注釈の言語は、ユーザに関連付けられたユーザプロフィールから決定されてもよく、注釈の言語は、ユーザアカウントに関連付けられた言語のリストから選択されることが好ましい。代替的に、第1注釈の言語は、テキスト又は言語認識方法を通じて自動的に決定されてもよい。

【0032】

[0046] メディアセグメントに関するキャプションとしての第1注釈を第1ユーザから受信するステップは、第1注釈を受信するステップS110に先立って、第1デバイスに映像、音声及び任意のキャプションをストリーミングするステップを含むことが好ましい。映像、音声及びキャプションは、メディアセグメントに関する映像、音声及びキャプションに限定されてもよく、若しくは、メディアの全体又は大部分に関する映像、音声及びキャプションであってもよい。ストリーミングされたキャプションは、メディアに関する第1言語（ユーザのためのキャプション付け言語）のキャプションであることが好ましいが、代替的に、ユーザが第2言語をキャプション付け言語に翻訳する時など、キャプション付け言語とは異なる第2言語のキャプションであってもよい。

10

【0033】

[0047] 第1ユーザから第1注釈を受信するステップは、追加的に、メディアセグメントに関するキャプション入力フィールドを第1ユーザに提示するステップを含むことが好ましい。より好ましくは、キャプション入力フィールドは第1デバイスでレンダリングされる。メディアセグメントに関する注釈は、メディアセグメント内の映像フレームに関連付けられて受信されることが好ましい。より好ましくは、メディアセグメントの映像フレームは、注釈の入力/受信中に表示されることが好ましい。第1注釈は、キャプション入力フィールド（例えば、テキスト入力フィールド又は注釈入力フィールド）で受信されることが好ましく、キャプション入力フィールドは、注釈を付けられる（字幕を付けられる）メディアセグメントを表す映像クリップ、音声又は画像とともに示される又は他の方法で関連付けられる。メディアセグメント及びキャプション入力フィールドは、複数のメディアセグメントのうちの一つとして示されてもよく、又は、他のメディアセグメント及びメディアファイル全体のキャプション入力フィールドとは無関係に示されてもよい。キャプション入力フィールドは、各メディアセグメントごとに提示されることが好ましく、各メディアセグメントは、異なるキャプション入力フィールドに関連付けられる。メディアセグメントに関するキャプション入力フィールドは、メディアセグメントに関するキャプションの現在のバージョンと連動して（例えば、共通フレームの近傍で共通フレームに囲まれて）提示されることが好ましい。キャプション入力フィールドは、メディアプレーヤの上部に重ね合わされる透明レイヤであることが好ましいが、代替的に、部分的に不透明であってもよく、メディアプレーヤの映像部分上に重ね合わせられるか、映像部分のセクション上に重ね合わせられるか、又は、任意の他の適切なプロパティを有してもよい。キャプション入力フィールドの位置は、メディアプレーヤに関連して固定されることが好ましいが、代替的に調節可能であってもよく、キャプションは、キャプション入力フィールドの調節された位置に表示されることが好ましい。調節された位置は、（例えばメディアプレーヤの上部、下部及び側部に）制限されてもよく、又は、制限されなくてもよい。キャプション入力フィールドが移動可能である場合、キャプション入力フィールドのユーザ決定場所は、キャプションファイル内にも記録されることが好ましい。代替的に、キャプション入力フィールドは、メディアプレーヤの近傍に配列されてもよく、又は、任意の他の適切な関係で位置決めされてもよい。別個のキャプション入力フィールドが、各メディアセグメントごとにレンダリングされることが好ましい。キャプション入力フィールドは、キャプションストリームの一部としてレンダリングされてもよく、複数のキャプション入力フィールドがキャプションストリームのフレーム内でスクロールする。キャプションストリームのキャプション入力フィールドを進めるステップは、メディアファイルの現在の

20

30

40

50

再生位置に対応するキャプション入力フィールドが示されて、キャプションストリームインターフェースのフレーム内で任意選択的に強調されるように、キャプションストリームをスクロールさせるステップを含むことが好ましい。キャプション入力フィールドは、メディアが再生される時、定義されたメディアセグメントのタイムフレームに関してのみメディアセグメントの受信された入力を表示することが好ましい。メディアは、キャプション入力フィールドの背後で再生されてもよく、メディアセグメントのプレビューを可能にする。キャプション入力フィールドは、テキスト入力フィールドと実質的に同様に機能することが好ましい。

【0034】

[0048] 第1注釈は、代替的に、メディアセグメントとは無関係に受信されてもよくS 120、第1注釈は、メディアセグメントの映像又は音声から切り離されるキャプション入力フィールドで受信され、第1注釈の例は図16に示される。第1注釈は、対象のメディアセグメントに関連付けられるキャプションに関連付けられて受信される。キャプション入力フィールドで受信された第1注釈はその後、メディアセグメントに関するキャプションの変形例として保存され、又は、メディアセグメントに関するキャプションとして保存される。第1注釈が第1キャプションに関して受信された後、システムは、注釈に関する第2キャプションを提示することが好ましく、第2キャプションは第2メディアセグメントに関連付けられる。第2メディアセグメントは、メディアファイル内の第1メディアセグメントに隣接する(例えば後続の)メディアセグメントであってもよく、又は、異なるメディアファイルからのメディアセグメントであってもよい。第2メディアセグメントは、無作為に選択されてもよく、第1メディアセグメントとの時間的關係に基づいて選択されてもよく、又は、任意の他の適切な方法で選択されてもよい。方法の変形例は、メディアセグメントの音声又は映像を提示せずに、メディアセグメントのキャプションのみを提示することによって、字幕付け及び翻訳を促進することができる。

【0035】

[0049] キャプション及び第1注釈が関連付けられるメディアセグメントは、無作為に選択されることが好ましいが、代替的に、メディアセグメントに関する第1注釈言語の記憶されたキャプションバージョンの数に基づいて選択されてもよく(例えば、自動的に決定され又はユーザにランク付けされ、既定の閾値を上回るランクを有するキャプションバージョンを欠くメディアセグメントが選択される)、又は、任意の他の適切な方法で選択されてもよい。

【0036】

[0050] キャプション入力フィールドで受信された第1注釈は、(例えば、キャプション言語とは異なる第2言語の)キャプションの翻訳であってもよい。例えば、キャプションは英語キャプションであってもよい一方で、第1注釈は中国語のものであってもよい。代替的に、第1注釈は、キャプションの編集(例えば、キャプション言語と同一の言語のもの)であってもよく、第2言語のものであっても同一のメディアセグメントに関連付けられる二次的キャプションがまた提供されることが好ましい。例えば、キャプション及び第1注釈は中国語であってもよい一方で、二次的キャプションが英語であってもよい。第1注釈がキャプションの翻訳又はキャプションの編集であるかどうかは、言語検出又は識別アルゴリズムを用いたシステムによって自動的に検出されてもよく、又は、ユーザによって(例えば、言語ドロップダウンメニューから)選択されてもよい。キャプション言語は、共通言語(例えば、英語)であることが好ましいが、代替的に任意の適切な言語であってもよい。キャプション言語は、システムによって自動的に決定されることが好ましいが、代替的にユーザによって選択されてもよい。後者の変形例の方法では、システムは、データベースからユーザの選択したキャプション言語のキャプションを選択及び提示することが好ましい。例えば、ユーザが、翻訳されるべきキャプション言語として「スペイン語」を選択した時、システムは、スペイン語を選択して翻訳のためにユーザに提示することが好ましい。

【0037】

10

20

30

40

50

[0051] キャプションは、メディアセグメントの任意の識別子とは全く無関係に提示されることが好ましい。代替的に、リンク、サムネイル、GIF又は他のメディアセグメント識別子がキャプションとともに提示されてもよい。単一のメディアセグメントに関して編集又は翻訳されるべきキャプションは、単独で又は対象のメディアセグメントに隣接する1以上のメディアセグメント（例えば、編集又は翻訳されるべきキャプションに関連付けられたメディアセグメント）のキャプションとともに提示されてもよい。メディアセグメントに関連付けられたキャプション（例えば、メディアセグメントに関して以前に受信又は生成されたキャプション）は、キャプション入力フィールドに関連付けられて第1ユーザのデバイスでレンダリングされる。メディアセグメントに関して選択又はレンダリングされたキャプションは、（例えば、第1及び第2注釈のスコアを決定するのと同様の方法での）高いランク又はスコアを有するキャプションであることが好ましいが、代替的に、メディアセグメントに関して選択されたキャプション言語内で無作為に選択されてもよく、又は、任意の他の適切な方法で選択されてもよい。

10

【0038】

[0052] 方法は、第1ユーザから第1注釈を受信する前に第1ユーザからメディアセグメントの選択を受信するステップを含むことが好ましい。メディアセグメントの選択を受信するステップは、ユーザがメディアセグメントに関する注釈を入力することができるように、メディアセグメントに関するキャプション入力フィールドに焦点を当てることが好ましい。メディアセグメントの選択は、カーソル入力、キーストローク、タッチ又は任意の他の適切な入力として受信されてもよい。メディアセグメントの選択を受信するステップは、タイムスタンプに関連付けられたアイコンの選択を受信するステップを含んでもよく、タイムスタンプは、開始及び終了タイムスタンプによって境界付けられるメディアセグメント期間内にある。アイコンは、レンダリングされたメディアタイムライン、キャプション入力フィールドを囲むフレームに入れられたフィールド、キャプション入力フィールド、レンダリングされたキャプション、メディア音声の視覚表示のセクション、又は、任意の他の適切なアイコンであってもよい。

20

【0039】

[0053] 第1ユーザから第1注釈を受信するステップは、追加的に、第1注釈に関する植字パラメータを受信するステップを含んでもよい。植字パラメータは、（例えば映像上における）キャプション配置、キャプションフォント、キャプションスタイル（例えば、イタリック体、太字等）、キャプション色、又は、任意の他の適切な植字パラメータを含んでもよい。植字パラメータは、第1注釈に対して隔離されることが好ましいが、代替的に、所定のメディアセグメントのための第1言語のキャプションの他のバージョンまで広められてもよく、他のメディアセグメントのための第1言語のキャプションまで広められてもよく、又はそうでなければ、メディア全体を通じて広められてもよい。

30

【0040】

[0054] メディアセグメントに関するキャプションとしての第2注釈を第2ユーザから受信するステップS200は、メディアセグメントに関するキャプションの第2変形例を受信するように機能する。第2ユーザは、第1ユーザとは異なることが好ましく、第2ユーザアカウント及び第2デバイスに関連付けられることが好ましく、第2ユーザアカウント及び第2デバイスの両方が、それぞれに第1ユーザアカウント及び第1デバイスとは異なることが好ましい。第2注釈は第1言語のものであることが好ましい。第2注釈は、第1注釈と同一のメディアセグメントについて受信されることが好ましく、メディアセグメントに関する第1言語のキャプションのバージョンとして記憶されることが好ましい。代替的に、第2注釈は、メディアセグメントについての第1言語のキャプションとして記憶されてもよく、第2注釈は第1注釈を上書きする。第2注釈は、第1注釈の編集であってもよく、第1注釈は、第2注釈の受信前に第2ユーザに提示されることが好ましい。第1注釈は、メディアセグメントについての第1言語のキャプションとして第2ユーザに提示されることが好ましい。メディアセグメントについての第1言語のキャプションの現在のバージョンは、以下に説明されるように第3ユーザに対するメディアセグメントに関するキ

40

50

ャプションの提示と同様の方法で第2ユーザに提示されることが好ましいが、代替的に他の方法で提示されてもよい。第2注釈は、代替的に、第1注釈又は任意の他の適切な注釈とは別個の注釈であってもよい。第1注釈と同様に、第2注釈は、テキストであることが好ましいが、代替的に、アイコン選択、リンク又は任意の他の適切な注釈であってもよい。第1注釈を受信するステップと同様に、第2注釈を受信するステップは、追加的に、第2注釈の言語を決定するステップを含んでもよい。第2注釈の言語は、第1注釈について説明されたように決定されてもよく、又は、別の方法で決定されてもよい。植字パラメータは第2注釈について受信されてもよく、又は、第1注釈に関する植字は第2注釈に広められてもよい。代替的に、第2注釈は、第1注釈の植字パラメータにおける編集であってもよい。

10

【0041】

[0055] メディアセグメントに関するキャプションとして第2注釈を第2ユーザから受信するステップは、上述したような第1ユーザにメディアをストリーミングするのと同様の方法で、第2注釈を受信する前に第2デバイスに、映像、音声及びメディアセグメントに関するキャプションとして第1注釈をストリーミングするステップを含んでもよい。メディアセグメントに関するキャプションとしての第2注釈を第2ユーザから受信するステップは、追加的に、上述したような第1ユーザにキャプション入力フィールドを提示するのと同様の方法で、メディアセグメントに関するキャプション入力フィールドを第2ユーザに提示するステップを含んでもよい。メディアセグメントに関するキャプションとしての第2注釈を第2ユーザから受信するステップは、追加的に、上述したような第1ユーザにキャプション入力フィールドを提示するのと同様の方法で、第2ユーザから第2注釈を受信する前に、第2ユーザからメディアセグメントの選択を受信するステップを含んでもよい。

20

【0042】

[0056] メディアセグメントに関するキャプションとしての第2注釈を受信するステップは、代替的に、メディアセグメントを選択するステップと、選択されたメディアセグメント(図17に示すように)の音声及び/又は映像とは無関係のキャプション入力フィールドに関連付けて、選択されたメディアセグメントに関する第1注釈をレンダリングするステップS210と、キャプション入力フィールドにおいて第2注釈を受信するステップS220と、選択されたメディアセグメントに第2注釈を関連付けるステップと、を含んでもよい。メディアセグメントの音声及び映像とは無関係にメディアセグメントに関するキャプションとして第2注釈を受信するステップは、上述したように、メディアセグメントの音声及び映像とは無関係にメディアセグメントに関するキャプションとして第1注釈を受信するステップと実質的に同様であることが好ましいが、代替的に、実質的に異なってもよい。メディアセグメントに関するキャプションとして第2注釈を受信するステップは、追加的に、第2注釈言語を選択するステップを含んでもよく、第2注釈は、選択された言語のキャプションとして、選択されたメディアセグメントとともに保存され又は選択されたメディアセグメントに関連付けられる。メディアセグメントに関するキャプションとして第2注釈を受信するステップは、追加的に、第1注釈の言語を選択するステップを含んでもよく、選択された言語の注釈は第1注釈として提示されることが好ましい。

30

40

【0043】

[0057] メディアセグメントを選択するステップは、第2注釈言語のキャプションを欠くメディアセグメントを選択するステップを含むことが好ましいが、代替的に、既定の閾値より高いランク又はスコアを有する第2注釈言語のキャプションを欠くメディアセグメントを選択するステップ、既定の閾値より低い第2注釈言語のキャプションの数を有するメディアセグメントを選択するステップ、メディアセグメントを無作為に選択するステップ、又は、任意の他の適切な方法でメディアセグメントを選択するステップ含んでもよい。

【0044】

[0058] 選択されたメディアセグメントの音声及び/又は映像とは無関係のキャプション入力フィールドに関連付けられた選択されたメディアセグメントに関する第1注釈をレン

50

ダリングするステップS 2 1 0は、第1注釈のみをレンダリングするステップと、第1注釈とは別個のキャプション入力フィールド内で第2注釈を受信するステップと、を含むことが好ましく、第1注釈は、第2注釈とは異なる言語のものであることが好ましい。代替的に、第1注釈をレンダリングするステップは、キャプション入力フィールド外でキャプションをレンダリングするステップと、キャプション入力フィールド内で第1注釈をレンダリングするステップと、第1注釈の編集として第2注釈を受信するステップと、を含んでもよく、キャプションは第1言語のものであり、第1及び第2注釈は第2言語のものである。キャプション言語は、ユーザから受信した選択によって決定されてもよく、又は、システムによって自動的に決定されてもよい。第1注釈をレンダリングするステップは、追加的に、第1注釈を選択してレンダリングするステップを含んでもよい。メディアセグメントに関する第1注釈又はキャプションの複数のバージョン（例えば、第1注釈の様々な入力又は編集された様々なバージョン）が存在する場合、第2ユーザに提示された第1注釈のバージョンは、無作為に選択されてもよく、最も高いランクを有するバージョンであってもよく、最も低いランクを有するバージョンでもよく、又は、任意の他の適切な方法で選択された任意のバージョンであってもよい。既定の閾値よりも低いスコア又はランクを有する注釈のバージョンは、注釈又はランクのために第2ユーザに提示されるべき可能性のある注釈のキューから取り消されてもよい。しかしながら、ユーザにレンダリングされるべき第1注釈は任意の他の適切な方法で選択されてもよい。

【0045】

[0059] 第2注釈を受信するステップS 2 2 0は、キャプション入力フィールドでテキスト入力を受信するステップを含むことが好ましい。第2注釈は、新たな入力であってもよく、又は、既存の入力の編集（例えば、第1注釈の編集）であってもよい。第2注釈は、メディアセグメントに関する第2注釈言語のキャプションのバージョンとして保存されることが好ましいが、代替的に、メディアセグメントに関する第2注釈言語の単なるキャプションとして保存されてもよい。第2注釈を受信するステップは、追加的に又は代替的に、肯定的な又は否定的なランキングのような、第1注釈に関する品質指標を受信するステップを含んでもよく、品質指標はそれぞれ第1注釈のスコア又はランクを上昇又は下降させる。第1注釈がキャプション入力フィールド内でレンダリングされる方法の一例では、第1注釈のランク又はスコアは、ユーザが第1注釈を編集せずに次に選択されたメディアセグメントに関する次のキャプションに進む時に上昇させられることが好ましく、かつ、第1注釈のランク又はスコアは、ユーザが第1注釈を編集する時に下降させられることが好ましい。否定的な品質指標の受信に応じて、システムは、図17Aに示すように、キャプションに関する第2注釈（例えば、テキスト翻訳）を入力するようにユーザに促すことができる。肯定的な品質指標の受信に応じて、システムは、第2メディアセグメントに関する第2キャプションをレンダリングすることができ、第2メディアセグメントは、図17Bに示すように、無作為に選択されてもよく、又は、メディアファイル内の第1メディアセグメントとの一次的な関係に基づいて選択されてもよい。

【0046】

[0060] メディアセグメントに関するキャプションをユーザに提示するステップS 3 0 0は、ユーザにメディアセグメントに関するキャプションの現在のバージョンを表示するように機能する。ユーザは、第3ユーザであることが好ましいが、代替的に、任意の他の適切なユーザであってもよい。第3ユーザは、第1ユーザ及び第2ユーザとは異なるユーザであることが好ましく、かつ、第3ユーザアカウント及び第3デバイスに関連付けられることが好ましく、第3ユーザアカウント及び第3デバイスの両方は第1及び第2ユーザアカウント並びに第1及び第2デバイスとはそれぞれ異なることが好ましい。しかしながら、ユーザは、特に注釈の入力中又は編集中、第1ユーザ又は第2ユーザであってもよい。メディアセグメントは、第1及び第2注釈が受信されたメディアセグメントであることが好ましい。ユーザにメディアセグメントに関するキャプションを提示するステップは、メディアセグメントに関する第1言語のキャプションをレンダリングするステップを含むことが好ましく、キャプションは、ユーザデバイスでレンダリングされることが好ましいが

、そうでなければ、ユーザに提示されてもよい。キャプションは、メディアの再生中、開始タイムスタンプと終了タイムスタンプとの間の映像上にレンダリングされてもよい。代替的に、キャプションは、映像の脇のプレーインターフェース（例えば、キャプションストリーミング）の一部上でレンダリングされてもよく、キャプションは、対応のメディアセグメントの再生が終了した後に残ってもよい。キャプションは、リアルタイムのプレビューで表示されることが好ましく、かつ、注釈がユーザから受信される際に現れることが好ましい。複数アカウントの方法の実施では、キャプションは、編集、編集者及びキャプションに対する任意の他の変化を追跡するキャプション履歴によって補足されてもよい。さらに、キャプション入力コンポーネントは、キャプション、キャプション評価コンポーネント（例えば、同意/不同意、星評価）、キャプション分類コンポーネント（例えば、下手な翻訳へのフラグ立て、文法エラーへのフラグ立て等）のための考察スレッドを提供することができる。

10

【0047】

[0061] 様々なメディアセグメントに関するキャプションは、再生されたメディアと同期して進められることが好ましく、再生されたメディアは、メディアファイルを再生する際に関連のキャプションを表示するように機能する。再生されたメディアと同期してキャプションを進めるステップは、メディアの再生位置を更新するステップを含んでもよい。メディアの再生位置は、キャプションストリームに従って、メディアタイムライン上のタイムスタンプのユーザ選択に従って、メディアセグメントのユーザ選択に従って、又は、再生位置に関連付けられた任意の他の適切なメディアパラメータに従って更新されてもよい。キャプションストリームのユーザナビゲーションはメディアの再生位置を変更してもよい。例えば、ユーザがキャプションストリームを先にスクロールインし、現在の再生位置の1分先の映像セグメントに関連付けられたキャプション上をクリックすると、メディアプレーヤは1分先にメディアをスキップする。同様に、キャプションの編集又はキャプションと相互作用する時、メディアの関連付けられた部分は、キャプションに関連付けられたメディアセグメントを再生、一時停止又は代替的にループさせてもよい。メディアセグメントを自動的に再生することによって、メディアのキャプション付け又は翻訳を助けることができる。一変形例では、キャプションストリームのナビゲーションを容易にするためにキーボードのショートカットが用いられてもよい。メディアプレーヤ及びキャプションストリームは、ナビゲーションキーボード入力、マウス/カーソル入力、タッチ入力等の形態に関係なく同期されることが好ましい。キャプションストリームは、代替的に、ユーザの介在を通じてメディアと非同期化されてもよい。例えば、ユーザは、メディア再生とは独立してキャプションをブラウズするためにキャプションストリームをスクロールさせることができる。

20

30

【0048】

[0062] メディアセグメントキャプション提示は、追加的に、キャプション評価に従ってキャプションをレンダリングするステップを含んでもよい。評価は、信頼水準、キャプションの分類化（例えば、自動翻訳、原資料からのキャプション、3次翻訳）、著作者に基づいた評価、又は任意の適切な評価であってもよい。キャプション評価は、通常、少なくとも、さらなる注意が必要なキャプションと、さらなる注意が不要なキャプションと、を特徴付ける。時には不十分な出来のエンティティ（アルゴリズムツール又は有効ではないコミュニティメンバーによって）を通じて概略的にキャプションが生成されることを可能にすることによって、さらなるメディアコンテンツがさらなる言語で字幕付けされてもよい。メディアコンテンツが速やかにアクセス可能になる場合、コンテンツは、コミュニティのキャプション編集を通じて継続的に改善されてもよい。好ましくは、キャプション評価に従ってキャプションをレンダリングするステップは、キャプション評価キーに従ってテキストの書式を設定するステップを含む。テキストの色、配置、スタイル及び任意の適切な態様は、キャプション評価を示すために用いられてもよい。メディアを観る時、視聴者が通知すると、キャプションは、低いキャプション評価を有するように示され、その後、その視聴者がキャプションストリームインターフェースを可能にしてキャプションの

40

50

改善を助けることができる。キャプション評価を示すことによって、経験豊かなユーザは、さらに、編集が必要なキャプションをより優先させることができる。例示の一実施形態では、自動キャプションシステムからのキャプションは、赤フォント色でレンダリングされて機械字幕付け及び低い信頼を示し、新しいユーザから投稿されたキャプションは、グレーフォント色でレンダリングされて人のキャプション付け及び低い信頼を示し、高いキャプション付けスコアを有するアカウントユーザによって投稿又は編集されたキャプションは、白フォント色でレンダリングされて高い信頼を示し、かつ、受信された肯定的な評価を有するキャプションは、白フォントでレンダリングされて高い信頼を示してもよい。

【0049】

[0063] 方法は、追加的に、図8及び図9に示すように、メディアセグメントに関する第1言語のキャプションとしての第1注釈又は第2注釈を選択するために、第1注釈及び第2注釈の分析を実行するステップS400を含んでもよい。選択されたバージョン/注釈は、メディアセグメントに関する第1言語のキャプションの現在のバージョンとして第3ユーザに提示されることが好ましい。S400は、提示されたキャプションバージョンである注釈を自動的に選択するステップ、又は、ユーザアカウントのコミュニティからの入力に基づいて注釈を選択するステップを含んでもよい。注釈を自動的に選択するステップは、第1及び第2注釈をメディアセグメントの音声の自動機械翻訳（例えば、音声テキスト翻訳）と比較するステップと、機械翻訳に最も類似する注釈を選択するステップ、既定の閾値を超える機械翻訳との類似度を有する（例えば、50%以上類似する）注釈を選択するステップ、又は、既定の閾値を超えた機械翻訳との相違を有する（例えば、40%未満の類似度、70%を超えた相違など）注釈を除去するステップと、を含んでもよい。類似度は、語句の収束又は相違、単語の類似度（例えば、キーワード）、文法の類似度、又は任意の他の適切な言語パラメータの類似度によって決定されてもよい。提示されたキャプションバージョンであるべき注釈を自動的に選択するステップは、代替的に、各注釈を第1言語に関する既知の規則（例えば、経験的に決定される又はユーザによって設定される）と比較するステップを含んでもよく、所定の閾値内の規則を満たす注釈は、提示されたキャプションバージョンとして選択されてもよい。提示されたキャプションバージョンであるべき注釈を自動的に選択するステップは、代替的に、最も一般的に生じる注釈バージョンを選択するステップを含んでもよい。提示されたキャプションバージョンであるために注釈を自動的に選択するステップは、代替的に、複数の注釈のうちで最も一般的に生じるコンポーネントを決定するステップと、一般的に生じるコンポーネントから新しい注釈を生成するステップと、を含んでもよい。例えば、複数の注釈が所定のセットのキーワードを含み（例えば、閾値出現比率を超える）、複数の注釈が所定の文法構造を含む（再び、閾値出現比率を超える）と、システムは、文法構造を用いてキーボードセットに命令して新たな注釈を生成してもよい。代替的に、S400は、注釈を比較して、直近に受信された注釈を選択することができる。代替的に、S400は、注釈を比較して、最も高いランクのユーザアカウントによってサブミットされた注釈を選択することができる。代替的に、S400は、注釈を比較して、メディアセグメントに関する第1言語の最も高いランクの注釈を選択することができる。この変形例では、システムは、キャプションを要求している各歴代のユーザに注釈を繰返し提示すること（例えば、第1注釈を第1視聴者に、第2注釈を第2視聴者に、など）、ユーザからの注釈に関する注釈品質指標を受信すること（例えば、良い/悪い、賛成票/反対票、1~10の連続体に沿ったランク等）、注釈をランク付けする又は受信された注釈品質指標に基づいて注釈に関するランクを算出すること、かつ、メディアセグメントに関する第1言語のキャプションとして最も高いランクを有する注釈を選択することができる。代替的に、既定の閾値を超えるランク又はスコア（例えば、肯定的な入力の数）を有する注釈が選択されて提示されてもよく、既定の閾値を超えるランク/スコアを有する複数の注釈は、繰返し又は無作為に選択されて提示されてもよい。注釈は、既定の期間で繰返し提示され、各注釈に関する所定の数の提示のために繰返し提示され（例えば、各キャプションバージョンは、10分間にわたって表示される）、又は、適切な停止条件が満たされるまで繰返し提示されてもよい。注釈は、キャ

プションに関する注釈が選択された後に繰り返し提示されてもよく、注釈は、繰り返し条件が満たされた後に繰り返し提示される。繰り返し条件は、新たな注釈の受信、所定の期間の達成、既定の閾値を超えるキャプションに関するユーザリクエストの数の達成、又は、任意の他の適切な条件であってもよい。代替的に、キャプション著作者、キャプション編集、フラグ立て、評価（例えば、同意/不同意、星評価など）、キャプションコメント、及び/又は、キャプションフィードバックの任意の代替的な形態が少なくとも2つのアカウントが共同でキャプションを改善することを可能にするために用いられてもよい。共同編集は、追加的に、アカウント優先度に従って投稿されてもよい。アカウントは、承認され、及び/又は、言語熟達度、年功、優先度、承認、許可及び/又は他の管理統制を獲得してもよい。編集、コメント及びキャプション付けは、行動を起こすアカウントに従って委任されてもよい。例えば、新たなユーザは、高い優先度のアカウント（例えば、多数の高い評価の付いた翻訳をしたアカウント）によって投稿されたキャプションの編集を抑制されてもよい。新たなユーザの編集は、メッセージ/アラートとして発せられ、又は、完全に却下されてもよい。高い優先度のアカウントは、低い優先度のアカウントのいかなるキャプションを編集することを許可されてもよい。しかしながら、メディアセグメントに関する第1言語のキャプションに関する注釈は、前述のバージョンの組み合わせを用いて選択されてもよく、又は他の方法で選択されてもよい。

10

【0050】

[0064] 方法の一変形例では、所定の言語のためのメディアセグメントに関するキャプションとして提示されるべきキャプションの変形例を選択するために注釈を処理するステップは、追加的に、選択された注釈を生成したユーザアカウントに対して通知を送信するステップ、以前に選択された注釈を生成したユーザアカウントに対して通知を送信するステップ、又は、任意の他の適切なユーザに対して通知を送信するステップを含んでもよい。例えば、通知は、選択された注釈のユーザアカウントに接続されたユーザアカウント（例えば、友人、フォロワー等）に対して送信されてもよい。

20

【0051】

[0065] 方法は、追加的に、アカウントの言語熟達度を評価するステップを含んでもよく、アカウントの言語熟達度の評価は、コミュニティメンバー及び翻訳の投稿者の言語技能を測定するように機能する。アカウント名義者の言語熟達度評価は、キャプション編集の強制、翻訳スコアの測定、キャプションの表示、参照キャプションの表示及びキャプション付けされたメディアに投稿又はキャプション付けされたメディアを消費する任意の適切な態様に用いられてもよい。第1変形例では、言語熟達度は、キャプションへの投稿に基づいて少なくとも部分的に評価されてもよい。評価は、任意の適切なアルゴリズムであってもよいが、投稿されたキャプションの数、投稿されたキャプションに対して他者がなした修正の数、他のアカウントによって投稿されたキャプションに対してなされた修正の数、使用済み参照キャプションの言語、キャプションの評価、投稿されたキャプションを有するメディア視聴、及び/又は、キャプション投稿の任意の適切な態様などのパラメータを含んでもよい。第2変形例では、言語熟達度は、メディアの消費に基づいて少なくとも部分的に評価されてもよい。好ましい一変形例では、方法は、映像ストリーミングサービスによって実施され、それゆえ、方法は、言語熟達度の評価におけるメディア消費を監視するステップを含んでもよい。消費されたメディアの数/長さ、及び、ネイティブ言語及び/又は消費されたキャプション言語は、言語熟達度に組み込んでもよい。追加的に、言語熟達度の態様は、ユーザ自己査定、証明されたテストスコア、言語テスト、又は、言語熟達度を測定する任意の適切な手段に基づいて評価されてもよい。

30

40

【0052】

[0066] 方法は、追加的に、図10に示すように、様々なユーザによって複数のメディアセグメントを共同で編集するステップを含んでもよい。これは、単独のユーザに各メディアセグメントに連続的にキャプションを付けさせる場合よりもキャプションをより速く生成することを可能にする。様々なユーザによって複数のメディアセグメントを共同で編集するステップは、第2メディアセグメントに関する第1言語のキャプションを第4ユーザ

50

から受信するステップを含んでもよい。代替的に、複数のメディアセグメントを共同で編集するステップは、第2メディアセグメントに関する第2言語のキャプションを第4ユーザから受信するステップを含んでもよく、第2言語は、第1言語とは異なる言語であることが好ましい。第2メディアセグメントは、第1メディアセグメントの開始及び/又は終了タイムスタンプとは異なる開始及び/又は終了タイムスタンプを有するメディアセグメントであることが好ましいが、代替的に、同一のメディアセグメントであってもよい。第4ユーザは、第1及び第2ユーザとは異なるユーザであることが好ましいが、代替的に、任意の適切なユーザであってもよい。第1及び第2メディアセグメントに関する注釈は、実質的に同時に受信される(例えば、第2メディアセグメントに関する注釈は、第1メディアセグメントに関する注釈の受信と実質的に同時に受信される)ことが好ましいが、代替的に、異なる時間に受信されてもよい。第2メディアセグメントに関する第1言語のキャプションに関する注釈は、第1メディアセグメントに関する第1言語のキャプションを含むファイルに追加されることが好ましいが、その他の方法で記憶されてもよい。第2メディアセグメントに関する第1言語のキャプションは、追加的に、第4ユーザとは異なるユーザによって上述した方法(S200)で編集可能であることが好ましい。さらに、アカウントは、キャプションに対するメディアの特定の部分を割り当てられてもよく、当該部分は、アカウントの言語熟達度評価に基づいてもよく又は基づかなくてもよい。キャプション付けシステムのサーバは、アカウントが実質的にリアルタイムで遠隔に協力することができるように、キャプション付けインターフェースのリアルタイム同期を可能にしてもよい。

10

20

【0053】

[0067] 方法は、追加的に、図11に示すように、提示されたキャプションを動的に更新するステップを含んでもよく、動的な更新は、キャプションの直近のバージョンを提示するように機能する。キャプションは、メディア及びキャプションをストリーミングしているユーザのために動的に更新されることが好ましい。キャプションは、新たな注釈(例えば、第2注釈)の受信に応じて動的に更新されることが好ましく、新たな注釈は、それぞれのメディアセグメントに関するキャプションとして視聴しているユーザに対して提示されることが好ましい。方法の一変形例は、第2デバイスに対する映像、音声及び第1言語のキャプションのストリーミングと実質的に同時に又はわずかにずれた時間で第4デバイスに対して映像、音声及び第1言語のキャプションをストリーミングするステップを含み、かつ、第4デバイスに対してストリーミングされた第1言語のキャプションは、第2デバイスから受信された第2注釈によって動的に更新される。第2及び第4デバイスに対する映像、音声及びキャプションのわずかにずれた時間でストリーミングするステップは、メディアが第4ユーザにストリーミングされるわずかに前に第2ユーザに対してメディアをストリーミングするステップを含むことが好ましく、第2及び第4ユーザは、同時にメディアの異なる部分を視聴している。

30

【0054】

[0068] 方法は、追加的に、他の適切な協力の提供と、例えば文法チェック、スペルチェック、翻訳ツール、タスク割り当て、用語集ツール、チャット又は考察ツール、キャプションナビゲーションツール、評価ツール、フィードバックツール、又はキャプション付け作業を改善するための任意の適切なツールなどの編集ツールの提供と、を含んでもよい。タスク割り当てツールは、特定のメディアセグメント、キャプションセットの全体及びキャプションの他の部分が、異なるアカウントに割り当てられることを可能にすることが好ましい。例えば、1アカウントが、他のアカウントがそれらキャプション上への作業を重複して行わないように、キャプション付けのためにキャプションのブロックする権利を有してもよい。用語集ツールは、キャプションセットの言語を標準化するように機能することが好ましい。ある翻訳は、色々な異なる語句によって翻訳されてもよい。用語集ツールは、標準化された語句又は俗語フレーズに関する参照として役立ち得る。例えば、1文字が、「上司」、「議長」、「チーフ」又は他の適切な語句として参照されてもよいが、用語集ツールは、様々なアカウントが文字に関する標準化された用語を用いることを援助す

40

50

ることができる。用語集ツールは、追加的に、用語集ツール内の標準化された語句を変更することが他の場所の語句を更新することができるように自動更新を可能にすることができる。追加的に、ツールは、語句が用いられるべき時に意味的に検出することができ、また、用語集ツールは、語句を提案又は語句を標準化された語句に自動的に変更してもよい。キャプションナビゲーションツールは、最初/最後の欠落した翻訳をナビゲーションすること、次の欠落した翻訳に進むこと、低い信頼の翻訳のみを表示すること、特定のアカウントによって作成された翻訳を見ること、又は、任意の適切なキャプションナビゲーション、検索及び/又はフィルタリング操作を実行すること、をキャプションナビゲーションツールが可能にすることができる現在のキャプション状態に基づいて、映像がナビゲートされ、検索され、フィルタリングされることを可能にすることができる。

10

【0055】

[0069] 方法は、追加的に、第1及び/又は第2ユーザのためのキャプション付け言語(第1言語)を選択するステップを含んでもよい。方法の一変形例では、音声言語(例えば、言語タグ、ジオタグ、メタデータ、自動言語認識などから決定される)は第1言語とは異なる。第1及び/又は第2ユーザのために第1言語を選択するステップは、それぞれのユーザのユーザプロフィール内に記載された言語を選択するステップを含むことが好ましい。

【0056】

2.1 多言語におけるキャプション生成

[0070] 方法は、追加的に、多言語でキャプションを生成するステップを含んでもよい。多言語でキャプションを生成するステップは、図12Aに示すように、様々なユーザアカウントから異なる言語のキャプション(例えば、音声からの直接翻訳)を受信するステップ、第1言語のキャプションを第2言語に自動的に翻訳するステップ、図12Bに示すように、第1言語のキャプションの第2言語への手動の翻訳を容易にするステップ、又は、多言語でのキャプション生成の任意の他の適切な方法を含んでもよい。第2言語は、第1言語とは異なる言語であることが好ましい。複数のユーザが所定のメディアセグメントに関する様々な言語のキャプションを同時に作成すること及び/又は複数のメディアセグメントにわたって様々な言語のキャプションを作成することを可能にすることによって、多言語でキャプションを生成するステップが、異なる言語のキャプションセットを同時に生成することを可能にする。

20

30

【0057】

[0071] 方法の一変形例では、メディアに関する第1言語のキャプションは第2言語のキャプションに自動的に翻訳される。自動的な翻訳は、機械翻訳を含むことが好ましく、第1言語のキャプションのテキストは、機械翻訳システムに送られ、かつ、機械翻訳システムの出力は、メディアセグメント/メディアに関する第2言語のキャプションとして保存されることが好ましい。代替的に、自動キャプション翻訳は、メディアの第2ピース(piece)から上手くキャプション付けされたメディアセグメントを識別するステップと、メディアの第1ピースと第2ピースとの間でメディアセグメントをマッチングするステップと、メディアの第1ピースのメディアセグメントのマッチングのキャプションとして、メディアの第2ピースから上手くキャプション付けされたメディアセグメントのキャプションを保存するステップと、を含んでもよい。上手くキャプション付けされたメディアセグメントは、既定の閾値を超える編集又はバージョンの数(例えば、3つのキャプションバージョン)を有するメディアセグメントであってもよく、既定の閾値を超えるランク(例えば、ユーザコミュニティによってランク付けされる場合に75%を超えるスコア)を有するキャプションを有するメディアセグメントであってもよく、又は、信頼性の高いキャプション付けの指標を有する任意の他の適切なメディアセグメントであってもよい。メディアの第1及び第2ピースのメディアセグメントをマッチングするステップは、メディアセグメントの音声波形/パターンをマッチングするステップを含むことが好ましいが、メディアセグメントはその他の方法でマッチングされてもよい。メディアセグメントマッチングは、音声周波数に依存しない(例えば、語句は男声又は女声に関わらずマッチングさ

40

50

れる)ことが好ましいが、代替的に、周波数依存であってもよい(例えば、女声の同一の語句はマッチングされるが、男声の語句は女声の同一の語句とはマッチングされない)。メディアの第2ピースに関する上手くキャプション付けされたメディアセグメントに関するキャプションは、メディアの第1ピース内のメディアセグメントのマッチングに関するキャプションとして用いられることが好ましい。識別された上手くキャプション付けされたメディアセグメントは、第1メディアに関するキャプションと同一の言語のものであることが好ましい。

【0058】

[0072] 第1言語のキャプションは、メディアセグメントベースで翻訳されることが好ましいが、代替的に、キャプションファイル全体として翻訳されてもよい。自動キャプション翻訳は、第1言語に関連する第2言語(例えば、英語をドイツ語に、日本語を韓国語になど)に限定されることが好ましいが、代替的に、限定されなくてもよい。代替的に、第1言語は、英語などの「共通」言語であってもよく、共通言語から第2言語(例えば、英語からスペイン語)への翻訳に関するデータは、実質的に豊富である及び/又は信頼性が高い(例えば、既定の閾値以上)。

10

【0059】

[0073] 第1言語のキャプションを第2言語のキャプションに自動的に翻訳するステップは、追加的に、ユーザに対して第2言語のキャプションを提示するステップと、ユーザから提示されたキャプションの編集を受信するステップと、を含んでもよい。ユーザに対して第2言語のキャプションを提示するステップは、メディアに関する第2言語のキャプションを要求しているユーザのデバイスにおいて第2言語のキャプションをレンダリングするステップを含むことが好ましいが、代替的に、ユーザのための第2言語のキャプションのスク립トをレンダリングするステップ、又はそうでなければ、ユーザに対して第2言語のキャプションを提示するステップを含んでもよい。ユーザは、第2言語で読み書きができるユーザであることが好ましいが、任意の適切なユーザであってもよい。第2言語の提示されたキャプションへの編集を受信するステップは、メディアセグメントに関して第2ユーザから第2注釈を受信するステップと同様の方法で注釈を受信するステップを含むことが好ましい。しかしながら、第2言語の提示されたキャプションへの編集は、その他の方法で第3ユーザから受信されてもよい。

20

【0060】

[0074] 方法の別の変形例では、システムは、第3ユーザによって第1言語のキャプションの第2言語への翻訳を容易にする。第3ユーザは、第1及び第2ユーザとは異なるユーザであることが好ましく、かつ、第1及び第2言語に堪能である又は第1及び第2言語で読み書きができることが好ましい。第1言語のキャプションの第2言語への翻訳を容易にするステップは、メディアセグメントに関する第1言語のキャプションを第3デバイスでレンダリングするステップと、メディアセグメントに関する第2言語のキャプションとしての注釈を第3デバイスから受信するステップと、を含むことが好ましい。注釈は、第2言語に関するキャプションファイル内のメディアセグメントに関する第2言語のキャプションの第1バージョンとして保存されることが好ましい。第2言語の注釈入力のためのキャプション入力フィールドは、追加的に、第1言語のキャプションによってレンダリングされてもよい。第2言語は、自動的に認識されてもよく、又は、ユーザ選択によって指定されてもよい。

30

40

【0061】

[0075] この変形例は、追加的に、図12Bに示すように、第4ユーザに対して第1言語のキャプションを提示するステップと、メディアセグメントに関するキャプションとしての第2言語の第2注釈を受信するステップと、を含んでもよい。第4ユーザは、第3ユーザとは異なることが好ましく、かつ、第1及び第2言語に堪能又は第1及び第2言語で読み書きができることが好ましい。これは、メディアセグメントに関する第2言語のキャプションの第2変形例を受信するように機能する。第2言語の第2注釈は、メディアセグメントに関する第2言語の現在のキャプションとして保存されてもよく、又は、メディアセ

50

グメントに関する第2言語のキャプションのバージョンとして保存されてもよい。メディアセグメントに関する第2言語の第1注釈(第3ユーザから受信された)はまた、メディアセグメントに関する第1言語のキャプションが第4ユーザに提示される時に第4ユーザに提示されてもよい(例えば、第4ユーザデバイスでレンダリングされる)。この変形例では、第2注釈は、第1注釈の編集であることが好ましい。第2言語のキャプションは、上述したように第1及び第2ユーザへのキャプション提示と同様の方法で第3及び第4ユーザに提示されることが好ましいが、任意の適切な方法で提示されてもよい。

【0062】

[0076] 第2言語のキャプションを受信するステップは、追加的に、各メディアセグメントに関する第2言語の複数のキャプション変形例を処理するステップと、提示のためのメディアセグメントに関する第1言語のキャプションの変形例を選択するステップと同様に、所定のメディアセグメントに関する第2言語のキャプションとして提示するための変形例を選択するステップと、を含んでもよい。異なるメディアセグメントに関する第2言語のキャプションは、複数のユーザから実質的に同時に又は一斉に受信されてもよく、又は、連続的に(例えば、1つずつ)受信されてもよい。

10

【0063】

[0077] 第2言語のキャプションを受信するステップは、追加的に、第3又は第4ユーザのためのキャプション付け言語を選択するステップを含んでもよい。第3又は第4ユーザのためのキャプション付け言語は、第1又は第2ユーザのための第1キャプション付け言語の選択について説明された方法で選択されることが好ましい。第3又は第4ユーザのためのキャプション付け言語を選択するステップは、第1言語のキャプションを第3又は第4ユーザに対して表示するかどうかを決定するステップを含んでもよい。例えば、ユーザプロファイルが音声言語及び第2言語を含むが第1言語を含まない時、メディアセグメントに関する第1言語のキャプションはユーザに対して提示されない(例えば、ユーザデバイスでレンダリングされない)ことが好ましい。ユーザプロファイルが第1言語及び第2言語を含む時、メディアセグメントに関する第1言語のキャプションがユーザに提示されることが好ましい。

20

【0064】

[0078] 方法は、追加的に、ユーザのための提示言語を選択するステップを含んでもよく、提示言語の選択は、ユーザに対して提示される(例えば、ストリーミングされてレンダリングされる)べきキャプションセットを選択するように機能する。より好ましくは、方法は、ユーザデバイスからメディアストリーミングリクエストの受信に応じてユーザのための提示言語を選択するステップと、ユーザデバイスで選択された提示言語のキャプションをストリーミング及び/又はレンダリングするステップと、を含む。ユーザのための提示言語は手動で又は自動的に選択されてもよい。提示言語の手動選択は、ユーザデバイスから言語選択を受信するステップを含むことが好ましい。言語選択は、言語に関連付けられたアイコンの選択、ドロップダウンリストからの言語の選択、又は、任意の他の適切な言語選択であってもよい。自動言語選択は、ユーザプロファイルに関連付けられた(例えば、ユーザプロファイル内で堪能な言語として記載された)言語を選択するステップ、ユーザプロファイルに関連付けられ、かつ、既定の閾値を超えるユーザによってこれまでに選択されてきた言語を選択するステップ(例えば、ユーザが常に50%を超えるキャプション言語として英語をこれまでに選択する時、提示言語は英語に設定される)、デフォルト言語を選択するステップ、ユーザによって選択された直近の提示言語を選択するステップ、ユーザによって用いられた直近のキャプション付け言語を選択するステップ、又は、適切な既定言語を自動的に決定する任意の他の適切な方法を含んでもよい。

30

40

【0065】

2.2 メディアセグメント化

[0079] 方法は、追加的に、図13及び図15に示すように、メディアをセグメント化するステップを含んでもよく、メディアのセグメント化は、ユーザのためのメディアセグメントをキャプションに提供するように機能する。メディアは、自動的に、半自動的に又は

50

ユーザによって手動でセグメント化されてもよい。データを手動でセグメント化するユーザは、第1又は第2ユーザとは異なるユーザ（例えば、第3、第4又は第5ユーザ）であることが好ましいが、代替的に、任意の適切なユーザであってもよい。セグメントは、ユーザデバイス又はユーザアカウントから受信されることが好ましいが、代替的に、セグメント化するユーザに関連付けられた任意の適切なコンストラクト（construct）から受信されてもよい。メディアセグメント化は、メディアの同一のピースに関するメディアキャプション付けと実質的に同時に実行されてもよく、定義されたメディアセグメントは、ユーザが、識別後に（例えば、メディアの残りの部分がセグメント化されない間でも）キャプション付けをするために実質的に即座に利用可能であることが好ましい。

【0066】

[0080] 一変形例では、キャプション付けのためにメディアをセグメント化する方法は、メディアに関するタイムラインに沿って音声の視覚表示をレンダリングするステップS500と、メディアセグメントのセグメントパラメータを受信するステップS600と、選択されたメディアセグメントを提示するステップS700と、メディアセグメントの開始及び終了タイムスタンプに関連付けられたキャプションを備えるキャプションファイルを生成するステップS800と、を含んでもよい。方法は、簡単で継ぎ目のない映像のセグメント化及びキャプション付けを可能にするように機能する。方法は、追加的に、映像の共同セグメント化及びキャプション付けを可能にするように機能する。メディアタイムラインの関数として音声を視覚的に表示することによって、この方法は、キャプション付けに対して最適に可変する視覚表示器を提供する。これは、音声視覚化の特徴によって提供された視覚的キューを用いて、ユーザが音声セグメントをより簡単に識別して選択することを可能にする。方法は、当該方法の実施のためにさらに構成されたインターフェースをレンダリングするように構成されたコンピュータシステムを通じて実施されることが好ましい。

【0067】

[0081] ユーザデバイス上のメディアに関するタイムラインに沿って音声（例えば、音声レイヤ）の視覚表示をレンダリングするステップS500は、メディアタイムライン及び映像と同期して視覚音声レイヤを提供するように機能する。音声の視覚表示をレンダリングするステップはさらに、音声セグメント化を受信するための音声レイヤセグメント化インターフェースを提供するように機能する。視覚表示は、時間の関数としての画像音声レイヤとしてレンダリングされることが好ましいが、その他の方法で視覚化されてもよい。視覚表示は、少なくとも1つの時間ベースの画像次元を有する波形としてレンダリングされることが好ましい。例えば、波形は、水平時間軸上にプロットされた映像の音声振幅を表示してもよい。時間軸は、任意の適切な方向に向けられてもよく、又は、任意の適切な表示で通信されてもよい。視覚表示は、追加的に、多周波数（例えば、女声に関する1波形及び男声に関する第2の波形）、複数トラック、又は、任意の他の適切な音声特徴を表示することができる。一変形例では、各スピーカは、ユーザがメディアセグメントの境界を視覚的に識別するだけでなく、スピーカの音声ストリームに基づいてメディアを容易にセグメント化する（例えば、各スピーカごとにメディアセグメントを作成する）ことを可能にするレンダリングされた音声レイヤを有してもよい。好ましくは、音声レイヤは、ユーザが口語（例えば、語句又は文章）の境界を視覚的に検出することを可能にする。例えば、会話している2人による映像では、音声レイヤは、誰かが話している時に厳密にグループ化されたいくつかの顕著な振幅レベルを有する波形を有する。スピーカ同士の間での移行中又はスピーチの中断中、波形は、より低い又はゼロの振幅を有する。音声レイヤは、音声トラック又はメディアの結合された音声トラックであることが好ましい。代替的な実施形態では、音声レイヤは、処理された音声の表示であってもよい。音声は、スピーチ又は口語を区分けするために処理されてもよく、キャプションにとって大事ではない背景ノイズ、音楽、音響効果又は他の音声信号を除去/低減するように処理されてもよく、又は、任意の適切な方法で処理されてもよい。音声レイヤは、代替的に、左右の音声信号又は複数のスピーカの音声フィードなどの音声の複数のチャンネルを表示してもよい。

10

20

30

40

50

【 0 0 6 8 】

[0082] 音声の視覚表示は、メディアタイムラインの全体にわたって音声を視覚化することが好ましい。代替的に、メディア音声のセグメントは視覚化されてもよい（例えば、メディアの第1半部分）。音声レイヤは、映像の再生位置又はタイムスタンプ位置に同期されることが好ましい。音声レイヤは、メディアの現在の再生位置に同期された音声信号の時間ベースのプロットを表示することが好ましい。音声レイヤは、通常、現在の再生位置に先行する及び/又は後続の音声信号を表示する。より好ましくは、現在の再生位置に関連付けられた音声信号は音声レイヤの中心に配置される。再生位置における再生の進行及び変化中（例えば、早送り、再生速度の変化、先送りスキップ等）、音声レイヤは、現在の再生位置に関連して適切な音声信号をレンダリングするように進む。代替的に、現在の再生位置を示すアイコンは、実質的に不活発な音声レイヤ/音声の視覚表示に沿って進行する。一変形例では、方法は、音声レイヤのズームイン及び/又はズームアウトを可能にしてもよく、音声レイヤのズームイン及び/又はズームアウトは、音声レイヤ内に表示される音声信号の視覚可能部分及び詳細を変化させるように機能する。追加的に、同期された進行中、メディアセグメントはまた、音声レイヤ内の対応の期間でレンダリングされることが好ましい。音声レイヤはまた、映像の再生進行とともに進行（例えば、スクロール）することが好ましい。メディアセグメントは、メディア内の時間位置に対応する第1及び第2タイムスタンプ（例えば、それぞれの開始及び終了時間）によって定義されてもよい。代替的に、メディアセグメントは、開始タイムスタンプ及び期間、開始映像フレーム及び終了映像フレーム、複数の映像フレーム、音声信号、又は任意の他の適切なメディアパラメータによって定義されてもよい（例えば、境界付けされる又は識別される）。存在するメディアセグメントの画像表示は、それらが対応する音声信号の部分に調整されてレンダリングされることが好ましい。

10

20

【 0 0 6 9 】

[0083] 音声の視覚表示をレンダリングするステップは、追加的に、視覚表示を生成するためのメディアの音声信号を処理するステップを含むことが好ましい。視覚表示は、メディアの再生前に生成及びレンダリングされることが好ましいが、代替的に、メディアの再生中に生成及びレンダリングされてもよく、メディアの再生後に生成及びレンダリングされてもよく、又は、任意の他の適切な順番で生成及びレンダリングされてもよい。キャプション付けサービスがメディアをホストする方法の変形例では、キャプション付けサービスは、音声ファイルを処理し、キャプション付けサービスサーバで視覚表示を生成することが好ましく、視覚表示は、ストリーミングされたメディアとともに受信デバイスに送信される。メディアが第三者（例えば、メディアホストサービス）からストリーミングされる方法の変形例では、システムは、音声をリアルタイムで（例えば、メディアの再生中に）視覚化することができ、バッファに格納された音声を視覚化（例えば、音声がバッファに格納中又は格納後）することができ、再生された音声を視覚化（例えば、メディアの再生後に音声を視覚化）することができ、音声を記録して記録された音声を、視覚化のためにキャプション付けサービスに送信することができ、又は、任意の他の適切な方法で視覚化することができる。これらの変形例では、音声は、デバイス上のデフォルト音声出力（例えば、スピーカ）から仮想通過音声共有システムを通じて経路を切り替えられてもよく、仮想通過音声共有システムは、経路を切り替えられた音声から視覚表示を生成する。音声は、ユーザによるメディア再生前に経路を切り替えられることが好ましいが、代替的に、ユーザによるメディア再生中に経路を切り替えられてもよく、映像は、同期を維持するために対応して遅延させられることが好ましい。しかしながら、音声はその他の方法で視覚化されてもよい。

30

40

【 0 0 7 0 】

[0084] メディアセグメントのセグメントパラメータを受信するステップS600は、キャプション割り当てのためにメディアセグメントを識別して定義するように機能する。メディアセグメントの1以上のセグメントパラメータを受信するステップは、メディアセグメントの開始及び終了を定義することが好ましい。メディアセグメントの開始及び終了に

50

対応した第1及び第2セグメントパラメータが受信されることが好ましい一方で、信号セグメントパラメータ又は3以上のパラメータが代替的に受信されてもよい。セグメントパラメータは、開始タイムスタンプを示すことが好ましいが、追加的に、開始タイムスタンプ後の期間又は終了タイムスタンプを示してもよい。代替的に、セグメントパラメータは、特定の開始映像フレーム（例えば、キャプションが最初に示されるべき映像フレーム）、終了映像フレーム、メディアセグメントの間の映像フレームに関する識別子を示してもよく、又は、任意の他の適切なメディアパラメータを示してもよい。代替的に、セグメントパラメータは、タイムスタンプ、映像フレーム識別子、又は、任意の他の適切なメディアパラメータであってもよい。複数のメディアセグメント（キャプション付けセグメント又はキャプション付けのためのセグメント）が選択されてもよい。セグメントは部分的に又は全体的に重複してもよい。

10

【0071】

[0085] メディアセグメントに関するセグメントパラメータを受信するステップは、開始タイムスタンプ（メディアセグメントの開始）において入力を受信するステップと、終了タイムスタンプ（メディアセグメントの終了）を、停止イベントが検出されたタイムスタンプに設定するステップと、を含むことが好ましい。入力は、デバイスで受信されることが好ましく、かつ、キーストローク、視覚表示の第1位置上でのカーソル選択（例えば、マウス、検出されたタッチ、スタイラス）、又は、任意の他の適切な入力であってもよい。終了タイムスタンプを定義する停止イベントは、維持された入力（例えば、開始タイムスタンプを設定する入力）の解除、開始タイムスタンプとは異なる視覚表示上での終了タイムスタンプの選択又は第2位置の選択、若しくは、視覚表示上の第1位置の選択、又は、任意の他の適切なイベントであってもよい。代替的に、終了タイムスタンプは自動的に定義されてもよい。例えば、終了タイムスタンプは、開始タイムスタンプから離れた（例えば、後の）既定の期間になるように自動的に設定されてもよい。代替的に、開始タイムスタンプは、ユーザのタイムスタンプ選択の受信に応じて自動的に決定されてもよく、開始タイムスタンプは、ユーザに選択されたタイムスタンプに最も近いメディアセグメント開始特徴（例えば、振幅最小値）に対応するタイムスタンプであってもよい。終了タイムスタンプも同様に自動的に決定されてもよい。代替的に、開始タイムスタンプは、先行するメディアセグメントの終了タイムスタンプに設定されてもよい。

20

【0072】

[0086] 代替的に、予め定義されたメディアセグメントの選択が受信されてもよい。予め定義されたメディアセグメントは、別のユーザによって前もって定義された開始及び終了タイムスタンプを有するメディアセグメントであってもよい。代替的に、メディアセグメントは、少なくとも半自動的に定義されてもよい。開始及び終了タイムスタンプはユーザ入力に応じて自動的に設定されてもよい。例えば、ユーザは、音声レイヤ/視覚表示内の時点を選択することができ、処理モジュールが、音声信号を分析して、提案されたタイムスタンプ範囲を見出すことができる。さらに別の代替案では、セグメント化は、メディアの音声及び/又は映像内でセグメントを識別することによって自動的に実行されてもよい。セグメントは、（例えば、所定の音声トラック又は全体について）スピーチパターンなしで音声トラックの期間を検出すること、及び、非スピーチ分割器に基づいてセグメントをグループ化することによって識別されてもよい。自動的にセグメント化されたメディアセグメントは、ユーザ選択を通じて起動され又は有効にされてもよい。しかしながら、メディアセグメント選択はその他の方法で受信又は決定されてもよい。予め定義されたメディアセグメントの選択は、メディアセグメント境界の範囲内の時点の位置表示の選択を受信するステップを含むことが好ましいが、代替的に、メディアセグメントに関連付けられた注釈の選択を受信するステップを含む、又は、任意の他の適切なメディアセグメント識別子の選択を受信するステップを含んでもよい。

30

40

【0073】

[0087] セグメントパラメータは、音声信号の視覚表示において受信されてもよく、視覚表示セグメントが選択される。例えば、ユーザは、視覚表示のセグメントを選択して（例

50

えば、クリックをする)、対応の音声セグメントを選択することができる。代替的に、メディアセグメント選択は、視覚表示とは無関係に受信されてもよい。例えば、ユーザは、メディア再生中、予めマッピングされたキーを選択及び押し下げることができ、最初のキー選択におけるメディアタイムスタンプは開始タイムスタンプ(開始タイムスタンプ)として設定されることが好ましく、維持された入力(キー選択)が解除された(停止イベント)メディアタイムスタンプは終了タイムスタンプとして設定される。セグメントは、代替的に、任意の手動の制御インターフェースを通じて定義、選択、識別又は記述されてもよい。

【0074】

[0088] 方法の一変形例では、メディアセグメントの選択を受信するステップは、メディアの開始タイムスタンプにおける入力を受信するステップ(メディアは、再生又は一時停止されてもよい)と、入力が維持されている間に開始タイムスタンプからメディアを再生するステップと、終了タイムスタンプ(停止タイムスタンプ)における入力の解除を検出するステップと、を含む。例えば、ユーザは、時間1:12に対応する点で音声レイヤ上の「マウスダウン」を実行することができ、時間1:35に対応する点にカーソルをドラッグし、次に、「マウスアップ」を実行してもよい。マウスボタンを解放する際、メディアプレーヤは、1:12から1:35まで映像を再生することができ、その映像部分を停止させる又は代替的にループすることができる。方法の一変形例では、メディアは開始タイムスタンプから再生され、終了タイムスタンプは、第2入力(例えば、マッピングされたキーの第2入力)の受信時に設定される。

10

20

【0075】

[0089] メディアは、代替的に、セグメント選択中にメディアセグメントの前及び/又は後の追加メディアコンテンツ(メディアコンテンツが存在する場合)を再生することができる。このようなセグメントパディングは、メディアセグメントが、先行する及び後続のメディアコンテンツのあるコンテンツとともに視聴されることを可能にする。メディアコンテンツ選択は、メディアセグメントにメディア部分を追加するように又はメディアセグメントから部分を除去するように速やかに調節されてもよく、メディアセグメントは、ある変形例では、メディアパディング内に示されるメディアの部分を追加する。

【0076】

[0090] 追加的に、メディアセグメントを選択するステップは、キャプション規制を実行するステップを含んでもよい。通常、キャプション規制は、セグメント時間を案内する最小及び最大キャプション長さに関連する。最小及び最大閾値は、文字カウント、単語カウント、単語毎の読み取り速度、及び、読み取り速度の他の適切な近似値に基づいてもよい。セグメントが所定の期間にあることを阻止される一方で、代替的に、警告が起動されてキャプション期間の問題を有するユーザに通知する。例えば、選択されたメディアセグメントが所定のキャプションに対して小さ過ぎる場合、音声レイヤ内の画像表示は色を変化させてもよい。理想的には、選択されたセグメントは、メディアコンテンツの視聴者がキャプションを読み取るための適度な時間を有している。従って、最小及び最大長さ閾値は、平均キャプション長さ及び読み取り速度に従って設定されることが好ましい。読み取り速度は、個人又は集団に関する既定データから決定されてもよく、又は、経験的に決定されてもよい(例えば、ユーザの目の動きを監視することによって)。代替的に又は追加的に、メディアセグメント境界は、注釈の長さに基づいて、注釈が受信される際に動的に調節されてもよい。従って、キャプションがタイプされる際、セグメントは、メディアセグメント閾値に一致させるために自動的にサイズを拡大又は縮小することができる。代替的に、メディアセグメント期間の動的なサイズ変更は、ユーザによってオーバーライドされてもよい。

30

40

【0077】

[0091] 選択されたメディアセグメントを表示するステップS700は、選択されたメディアセグメントを視覚的に表示して、キャプションに関するタイミングを表示するように機能する。定義されたメディアセグメントは、メディアプレーヤ上でレンダリングされる

50

ことが好ましいが、代替的に、他の場所でレンダリングされてもよい。選択されたメディアセグメントは、音声信号の視覚表示上で表示される（例えば、記述される、表示される又は示される）ことが好ましいが、代替的に、タイムラインの画像表示上、映像フレームの画像表示上に表示されてもよく（例えば、音声信号のタイムライン/視覚表示に対応する複数の連続的な映像フレームがレンダリングされる）、又は、メディアタイムラインの任意の適切な表示上に表示されてもよい。メディアセグメントは、開始タイムスタンプを表示する位置にある第1境界又はアンカー（開始タイムスタンプ）、及び、終了タイムスタンプを表示する位置にある第2境界又はアンカー（終了タイムスタンプ）で表示されることが好ましい。選択されたセグメントは、それぞれのタイムスタンプに対応する位置に位置決めされた第1及び第2アイコン（例えば、時間軸、ドットなどに直行する線）によって表示されることが好ましいが、代替的に、選択されたメディアセグメントを境界付けるウィンドウによって表示されてもよく、選択されたメディアセグメント（例えば、タイムライン上、音声レイヤ上、又は映像フレーム上）を強調することによって表示されてもよく、又は、任意の他の適切な方法で表示されてもよい。メディアセグメントは、選択される際に視覚化されることが好ましい。方法の一変形例では、第1境界は、開始タイムスタンプの選択（例えば、入力最初の受信）時にレンダリングされ、第2境界は、終了タイムスタンプの選択（例えば、停止イベントの検出）時にレンダリングされる。方法の別の変形例では、第1境界及び第2境界は、開始タイムスタンプの選択時にレンダリングされ、第2境界は、開始タイムスタンプから離れた既定の期間を表示する位置でレンダリングされることが好ましい。終了タイムスタンプの連続的な選択は、第2境界を、終了タイムスタンプを表示する位置に移動させることが好ましい。メディア再生中の維持された入力の除去時に終了タイムスタンプが設定される変形例では、第2境界は、メディアの現在の再生位置と同期して移動し、メディアの現在の再生位置を表示することが好ましい。しかしながら、選択されたメディアセグメントは、任意の他の適切な方法で表示されてもよい。

【0078】

[0092] メディアセグメントの境界は、追加的に、修正され、調節され、増やされ、編集されてもよく、又は、（例えば、元のセグメント化しているユーザアカウント又は別のユーザアカウントからの）連続的に受信されたユーザ入力によるその他の方法で、更新されてもよく、方法は、追加的に、境界修正を受信するステップと、修正された境界によって表示される開始及び終了タイムスタンプを、変更された境界に関連付けられたタイムスタンプ値に設定するステップと、を含む。例えば、一旦視覚化された第1又は第2境界は、選択されてもよく（例えば、クリックされる、又は、マッピングされたキーの選択などによって選択される）、視覚表示上の新たな位置にリセットされてもよく（例えば、ドラッグされる）、リセットされた境界に関連付けられた開始又は終了タイムスタンプは、変更されて、新たな境界位置によって表示されたタイムスタンプ値を反映することが好ましい。代替的に、境界は数値インターフェースを通じて変更されてもよく、ユーザは、所定のタイムスタンプに関する新たなタイムスタンプを増加させ、減少させ又は入力することができる。メディアセグメントの開始又は終了アンカー（それぞれ第1及び第2境界）を修正している間、メディアプレーヤは、映像プレイバックを調節して、メディアセグメント内に含まれる映像部分を反映させることができる。メディアセグメントの開始又は終了タイムスタンプ（アンカー）を異なる時間部分に移動させる時、メディアプレーヤは、メディア内のその位置に関するメディアの部分をスクラブする又は提示することができる。同一のメディアセグメントに関する複数の異なる境界は、複数のユーザから受信されてもよく、通常選択された境界は、キャプションファイル内のそれぞれのタイムスタンプとして設定されることが好ましい。代替的に、境界は、モデレータによって選択され、音声ファイルのコンピュータセグメント化と最も適合するものとして選択され、又は、他の方法で選択される。

【0079】

[0093] 開始及び終了タイムスタンプに関連付けられたキャプションを備えるキャプショ

ンファイルを作成するステップS 8 0 0は、メディアファイルに関するすべてのメディアセグメントをデータオブジェクト内に特徴付けるように機能する。メディアキャプションファイルは、再生中のメディアとともにキャプションを表示するように構成されることが好ましく、また、メディアの映像及び音声に同期されることが好ましい。メディアキャプションファイルは、データファイル、データモデルのパラメータ内に記憶されたデータとして保存され、又は、任意の適切な方法で存続させられることが好ましい。一変形例では、メディアキャプションファイルは、キャプション付けがメディアファイルの再生時に利用可能であるように、メディアファイルに関連付けて記憶される。別の変形例では、メディアキャプションファイルは、ダウンロード可能又はストリーミング可能なりソースとして作成される。例えば、メディアホストサービスは、受信デバイスがメディアホストサービスからメディアを要求する時、受信デバイスにキャプションをストリーミングするように、キャプション付けサービスに要求を送信することができる。キャプションファイルは、メディアを記憶しているメディアリポジトリ上に保存されてもよく、又は、メディアを記憶するサービスとは別個のサービスに属する別個のサーバ上に保存されてもよい。キャプションファイル内のキャプションは、空（空ストリング）であってもよく、代替的に、テキストなどの注釈が投入されてもよい。キャプションファイルを作成するステップは、予め存在するキャプションファイルを更新する又は編集するステップを含むことが好ましいが、代替的に、新たなファイルを作成するステップを含んでもよい。

10

20

30

40

50

【 0 0 8 0 】

[0094] 方法は、追加的に、セグメント化しているユーザから、定義されたメディアセグメントに関する注釈を受信するステップS 9 0 0を含んでもよく、注釈を受信するステップは、メディアセグメントキャプションにコンテンツを割り当てるように機能する。セグメント化しているユーザは、第1ユーザとして扱われることが好ましいが、代替的に、異なるユーザであってもよい。代替的に、図14A及び図14Bに示すように、注釈は、セグメント化しているユーザとは異なるユーザから受信されてもよい。注釈（キャプションコンテンツ）は、テキストであることが好ましいが、代替的に、画像及び/又はナレーション/再録などのキャプション付けで用いられた他のメディアを含んでもよい。直接の入力（例えば、キーストローク又はカーソルの動き）なしのメディアセグメントのタイミングの設定後にユーザがキャプションを直接タイプすることができるように、メディアセグメントの識別/定義を開始する時又は完了するユーザ上のキャプション入力フィールドに焦点が自動的に設定されてもよい。方法のこの変形例は、メディアセグメント選択を受信するステップと、キャプションボックスを即座に作成するステップと、介在するステップなしで、メディアセグメント選択に応じてキャプションボックスに焦点を向けるステップと、を含む。

【 0 0 8 1 】

[0095] 代替的に、メディアセグメントに関する注釈は、アップロードされたスクリプトから決定され、スクリプトラインのテキストは、タイミング、性別、シーン、キーフレーム又は任意の他の適切なパラメータに基づいて関連のメディアセグメントに一致させられる。別の変形例では、注釈は、開始タイムスタンプに関連付けられた注釈テキストを含むファイルから決定され、注釈テキストは、開始タイムスタンプに基づいてキャプションファイル内の正確なキャプションにマッピングされる。例えば、注釈は、開始タイムスタンプに一致している開始スタンプを有するキャプションにマッピングされてもよく、又は、記録された注釈タイムスタンプがそれぞれの開始及び終了タイムスタンプの間で生じるキャプションにマッピングされてもよい。

【 0 0 8 2 】

[0096] 代替的に、対応の定義されたメディアセグメントを有しないメディアの部分について、ユーザは、注釈入力フィールド内にテキストを入力することができ、メディアセグメントは自動的に作成されてもよい。このような変形例では、メディアセグメントは、音声又は他の要因に基づいて、予め定義された期間をデフォルトにすることができ、又は、期間の自動決定を用いることができる。テキスト又は他の注釈コンテンツは、対応のメデ

ィアセグメントを選択することによって、又は、注釈を表示する注釈入力フィールドを選択することによって、編集されてもよい。キャプションテキストは、映像のスピーチの転写であることが好ましい。キャプションテキストは、代替的に、映像のスピーチの翻訳、補足的／主観的／情報のテキスト、署名及び他のテキストの翻訳、又は、任意の適切なコンテンツであってよい。

【 0 0 8 3 】

[0097] 追加的に、方法は、キャプションパラメータを設定するステップを含んでもよく、キャプションパラメータを設定するステップは、メディアセグメントの任意の追加の態様を設定するように機能する。キャプションパラメータは、任意の適切なユーザインターフェースを通じて設定されることが好ましい。キャプションパラメータは、スピーカ割り当て、キャプション付け分類（スピーチ、音響効果、音楽、文章翻訳など）、フォントスタイル、キャプション位置決め、キャプション推移、又は、キャプションの任意の適切な特性を含んでもよい。例えば、WYSIWYGキャプションインターフェースでは、ユーザは、キャプションを、映像セグメントの任意の適切な領域に位置決めすることができる。アカウントは、追加的に、コメントすることができ、また、メディアセグメント上の他の共同行動を実行することができる。

10

【 0 0 8 4 】

[0098] システム及び方法は、コンピュータ読み取り可能命令を記憶しているコンピュータ読み取り可能媒体を受け入れるように構成された機械として、少なくとも部分的に具現化及び／又は実装されてもよい。命令は、コンピュータ実行可能コンポーネントによって実行されることが好ましく、コンピュータ実行可能コンポーネントは、アプリケーション又はサーバホストアプリケーションのキャプションインターフェースと一体化されることが好ましい。コンピュータ読み取り可能媒体は、RAM、ROM、フラッシュメモリ、EEPROM、光学デバイス（CD又はDVD）、ハードドライブ、フロッピー（登録商標）ドライブ、又は、任意の適切なデバイスなどの任意の適切なコンピュータ読み取り可能媒体上に記憶されてもよい。コンピュータ実行可能コンポーネントは、一般的な又はアプリケーション特有のプロセッサであることが好ましいが、任意の適切なハードウェア専用又はハードウェア／ファームウェアの組み合わせデバイスが、代替的に又は付加的に、命令を実行してもよい。

20

【 0 0 8 5 】

[0099] 簡潔にするために省略したが、好適な実施形態は、様々なシステムコンポーネント及び様々な方法のプロセスのすべての組み合わせ及び配列を含む。

30

【 0 0 8 6 】

[0100] 当業者が先の詳細な説明から及び図及び特許請求の範囲から認識する際、以下の特許請求の範囲で定義された本発明の範囲から逸脱することなく、本発明の好ましい実施形態を修正及び変更することができる。

【 図 1 】

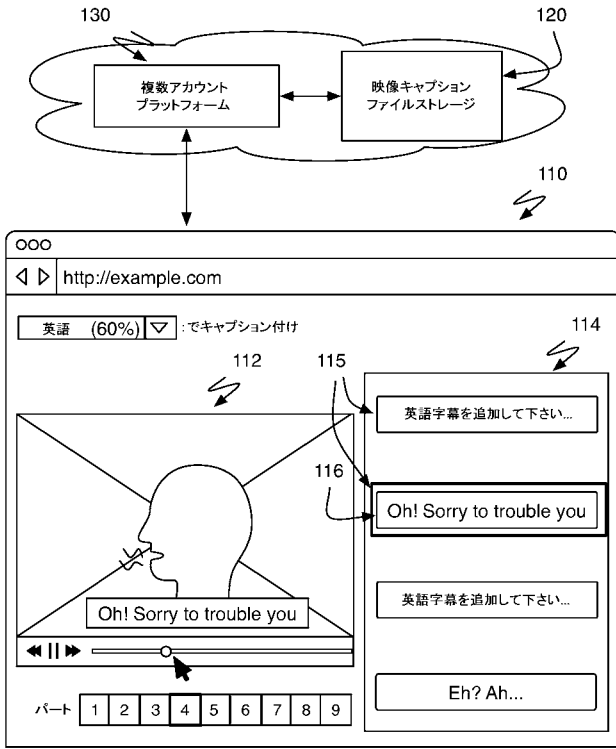


FIG. 1

【 図 2 】

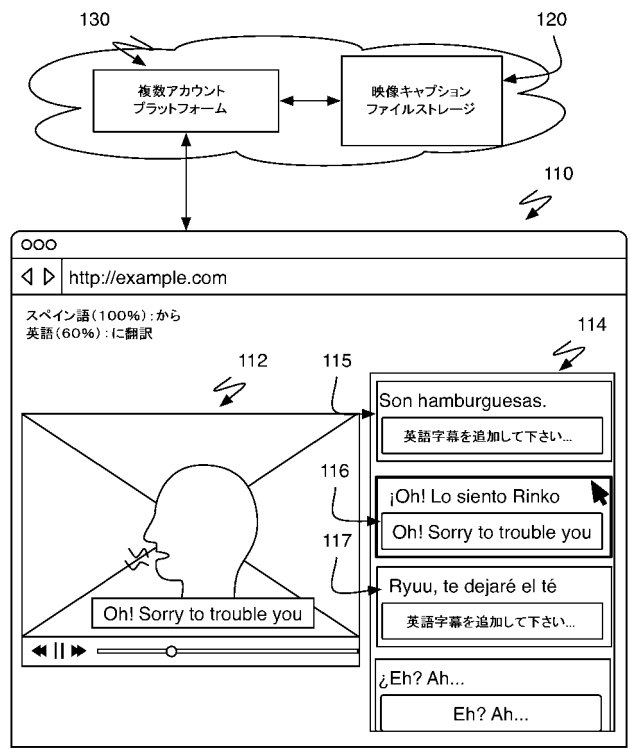


FIG. 2

【 図 3 】

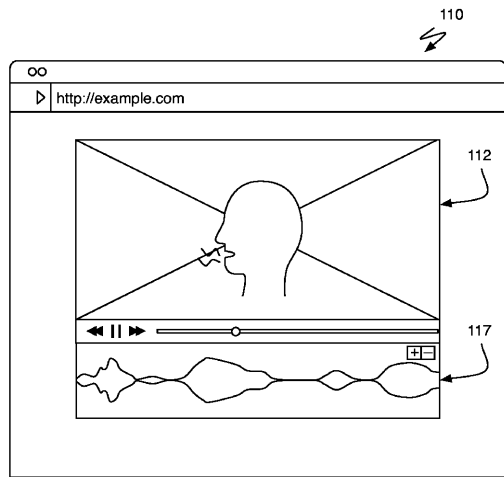


FIG. 3

【 図 4 】

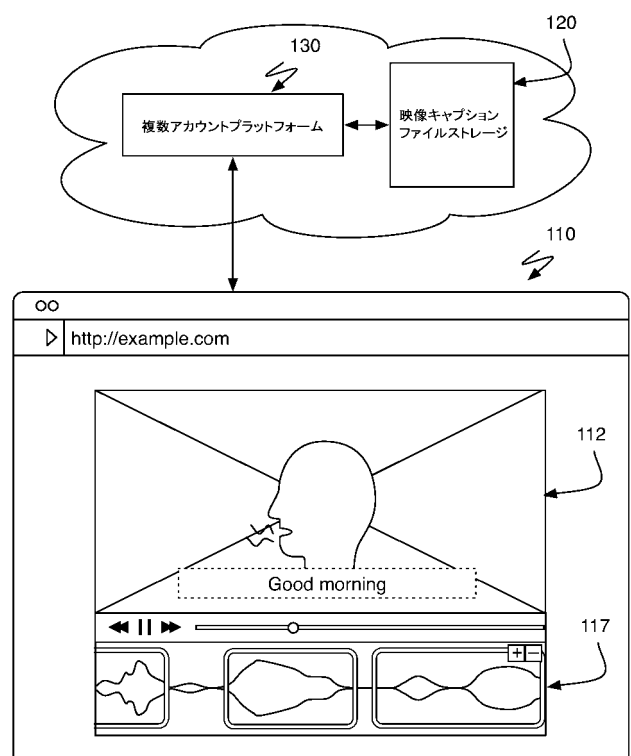


FIG. 4

【 図 5 】

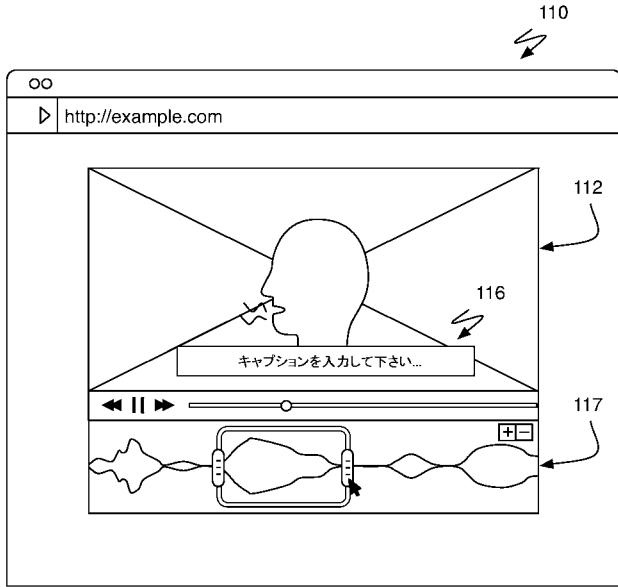


FIG. 5

【 図 6 】

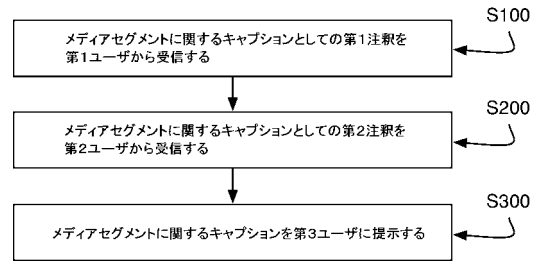


FIG. 6

【 図 7 】

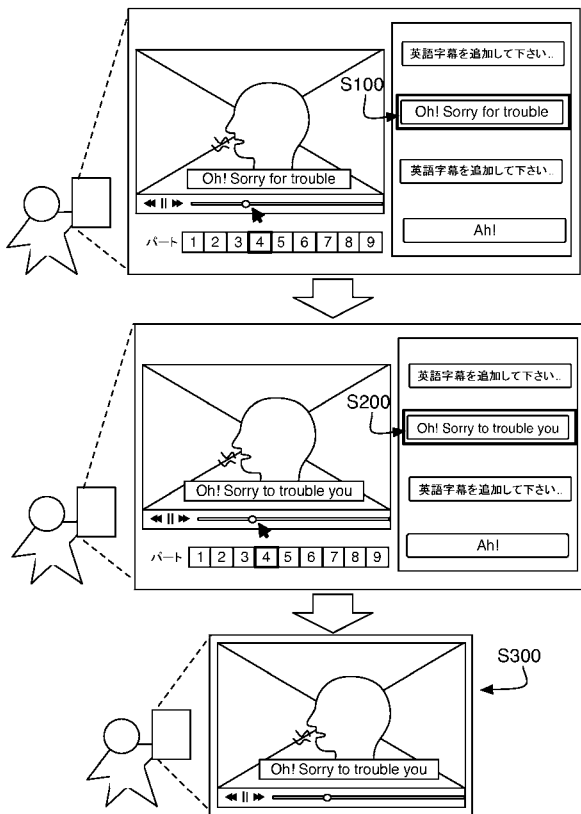


FIG. 7

【 図 8 】

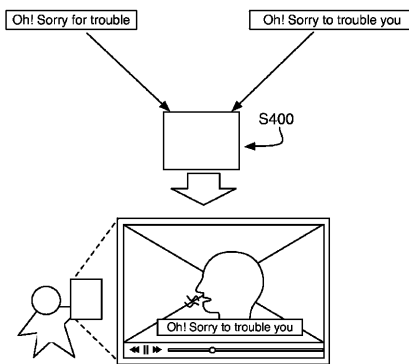


FIG. 8

【 図 9 】

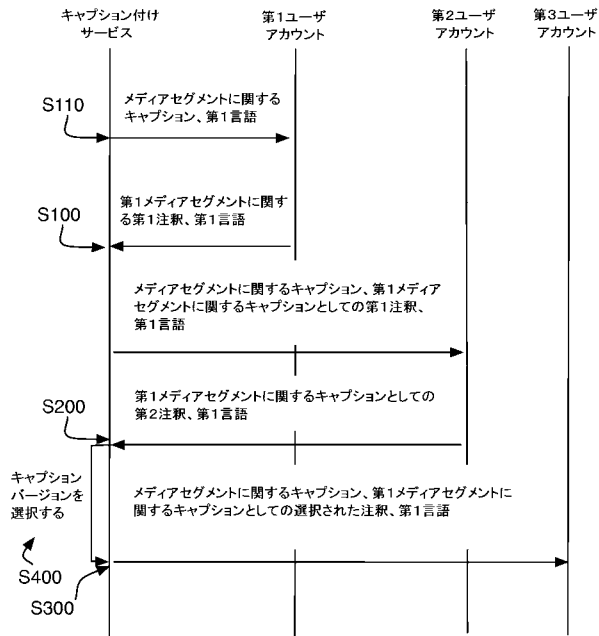


FIG. 9

【 図 1 0 】

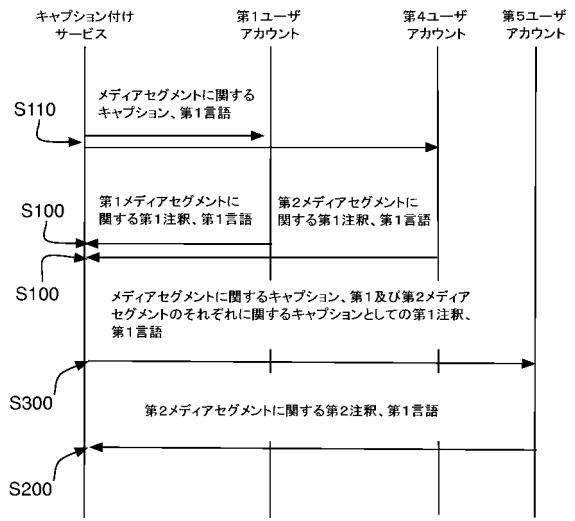


FIG. 10

【 図 1 1 】

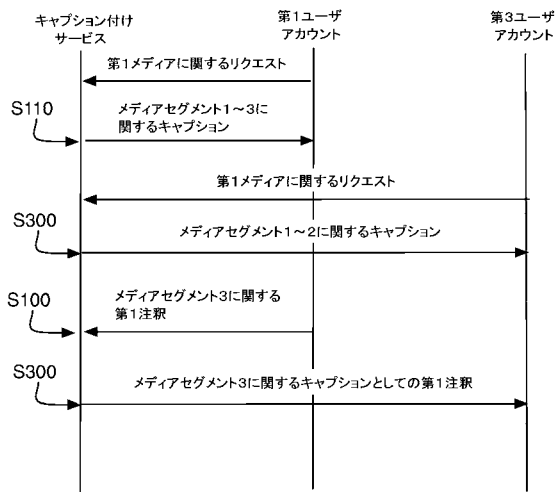


FIG. 11

【 図 1 2 】

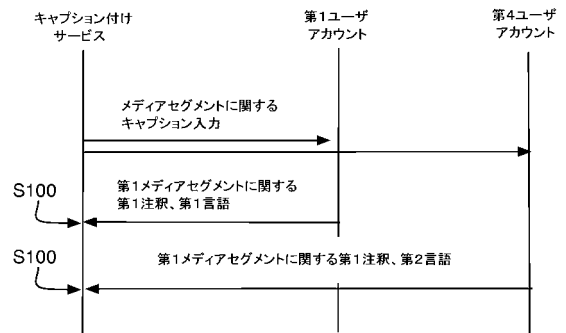


FIG. 12A

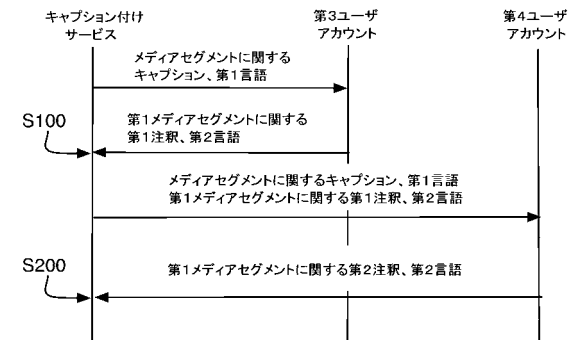


FIG. 12B

【 図 1 3 】

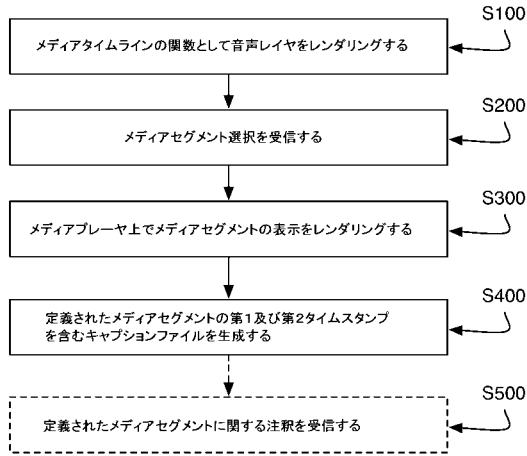


FIG. 13

【 図 1 4 】

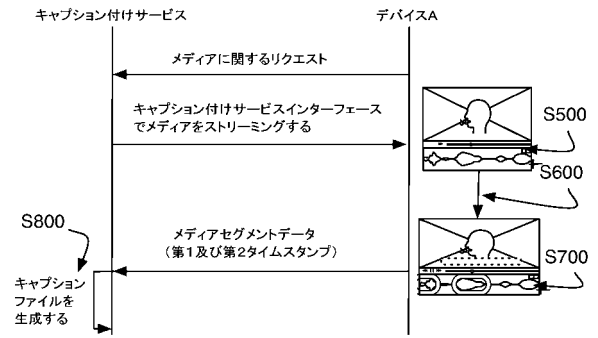


FIG. 14A

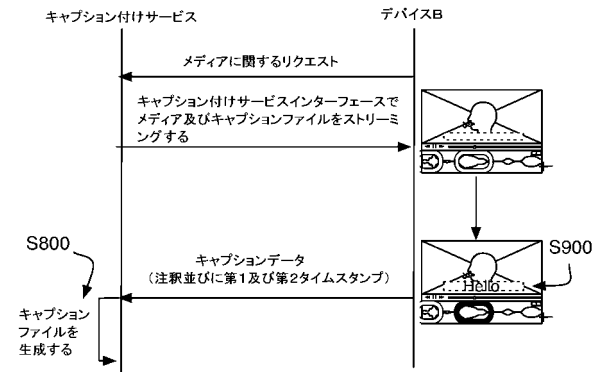


FIG. 14B

【 図 1 5 】

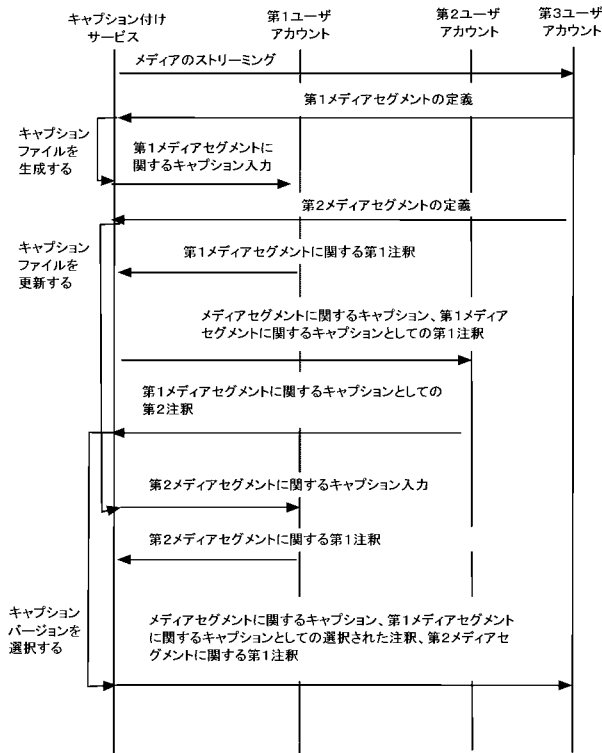


FIG. 15

【 図 1 6 】

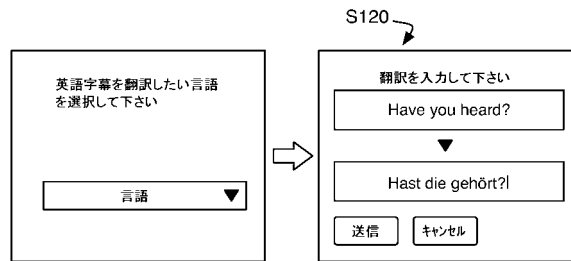
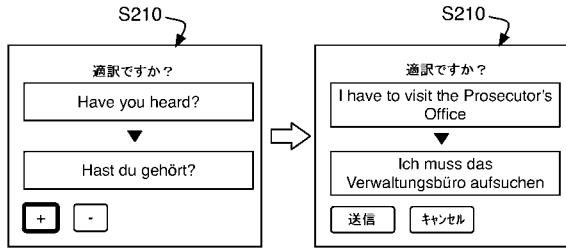
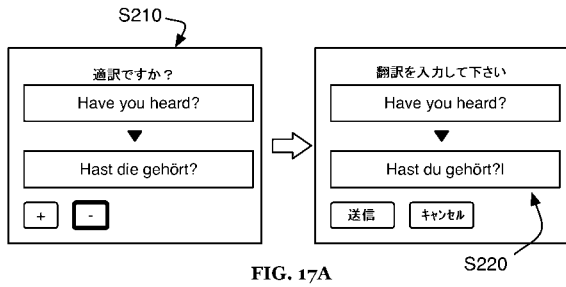


FIG. 16

【 図 17 】



【 国際調査報告 】

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US13/55573

A. CLASSIFICATION OF SUBJECT MATTER IPC(8) - G06F 3/0484 (2014.01) USPC - 386/282 According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC(8) Classification(s): G06F 3/048, 3/0484; H04N 7/173, 21/80 (2014.01) USPC Classification(s): 386/241, 278, 282 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) MicroPatent (US-G, US-A, EP-A, EP-B, WO, JP-bib, DE-C,B, DE-A, DE-T, DE-U, GB-A, FR-A); ProQuest; IEEE; Google/Google Scholar Keywords: Caption, Video, Audio, Render, Timestamp, Annotation, Input, Frames, Streaming		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2007/0136777 A1 (HASEK, C. et al.) 14 June 2007; paragraphs [0025], [0026], [0085], [0103], [0104].	1-2, 9, 12
Y		3-8, 10-11
Y	US 2003/0004968 A1 (ROMER, D. et al.) 02 January 2003; figure 5, paragraphs [0125], [0141].	3-8, 10-11
Y	US 2012/0163770 A1 (KAISER, D. et al.) 28 June 2012; paragraphs [0106], [0108], [0133], [0187].	10-11
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/>		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 06 March 2014 (06.03.2014)		Date of mailing of the international search report 07 MAY 2014
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201		Authorized officer: Shane Thomas PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US13/55573

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Group I: Claims 1-12; Group II: Claims 13-20; Group III: Claims 21-34; Group IV: Claims 35-46

****-Please see Supplemental Page-****

1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US13/55573

-Continued from Box No. III - Observations where unity of invention is lacking-

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fee must be paid.

Group I: Claims 1-12 are directed toward a method for captioning streamed media having video and audio.

Group II: Claims 13-20 are directed toward a method for collaboratively captioning media including audio.

Group III: Claims 21-34 are directed toward a method for collaboratively captioning media in multiple languages.

Group IV: Claims 35-46 are directed toward a method for collaboratively captioning media, the media including video and audio.

The inventions listed as Groups I-IV do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The special technical features of Group I include a remote media repository; a timeline of the media; a first timestamp and a second timestamp; and generating a caption file, which are not present in Groups II-IV.

The special technical features of Group II include icons corresponding to the boundaries of a first media segment, which are not present in Groups I and III-IV.

The special technical features of Group III include a first and second user account; a second language different from the first language, which are not present in Groups I-II and IV.

The special technical features of Group IV include streaming video, audio, and a second annotation to a third device in response to a request from the third device, which are not present in Groups I-III.

The common technical feature shared by Groups I-II is a method for captioning streamed media having video and audio, the method comprising: rendering a visual representation of the audio at the first device; and receiving segment parameters that correspond to a media. However, this common feature is previously disclosed by US 2009/0094659 A1 (Candelore). Candelore discloses a method for captioning streamed media having video and audio (audiovisual content stream having closed captioning; Abstract), the method comprising: rendering a visual representation of the audio at the first device (dialog of the audio portion is displayed on a display apparatus 100 (first device); Figure 1 and paragraphs [0017]-[0019]); and receiving segment parameters that correspond to a media (a database structure identifies an audiovisual program from a sequence of consecutive closed captions from the stream; Figure 6A and paragraph [0032]).

Since the common technical feature is previously disclosed by the Candelore reference, this common feature is not special and so Groups I-II lack unity.

The common technical feature shared by Groups III-IV is a method for captioning streamed media having video and audio, the method comprising: receiving a first annotation in a first language for a media segment; and receiving a second annotation. However, this common feature is previously disclosed by US 2010/0106482 A1 (Hardacker). Hardacker discloses a method for captioning streamed media having video and audio (video/audio stream includes a closed caption data format; Abstract and paragraph [0051]), the method comprising: receiving a first annotation in a first language for a media segment (a first data stream representing text in a first language is received; Abstract); and receiving a second annotation (second data stream representing text in a second language is received; Abstract).

Since the common technical feature is previously disclosed by the Hardacker reference, this common feature is not special and so Groups III-IV lack unity.

The common technical feature shared by Groups I-IV is a method for captioning streamed media having video and audio, the method comprising: rendering a first annotation at a first device; and receiving an annotation for a first media segment from a second device. However, this common feature is previously disclosed by US 2009/0094659 A1 (Candelore). Candelore discloses a method for captioning streamed media having video and audio (audiovisual content stream having closed captioning; Abstract), the method comprising: rendering a first annotation at a first device (display apparatus 100 (first device) presents additional information related to the identification of the audiovisual content; Figure 2 and paragraph [0024]); and receiving an annotation for a first media segment from a second device (audiovisual content receiver 210 receives the audiovisual content stream 112 from the audiovisual content source 104 (second device); Figure 2 and paragraph [0020]).

Since the common technical feature is previously disclosed by the Candelore reference, this common feature is not special and so Groups I-IV lack unity.

フロントページの続き

(81)指定国 AP(BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, RU, TJ, TM), EP(AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ

(特許庁注：以下のものは登録商標)

1 . Y o u T u b e

(72)発明者 リー , サンジン

アメリカ合衆国 カリフォルニア州 9 4 1 0 4 , サンフランシスコ , ブッシュストリート 2 2
5 , スイート 1 3 2 5

Fターム(参考) 5C053 FA14 GB11 GB12 HA22 HA24 HA25 HC05 JA16 LA14
5C164 MA06P UB08S UB10P UB88P UB94S