

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2015-176592
(P2015-176592A)

(43) 公開日 平成27年10月5日(2015.10.5)

(51) Int.Cl.	F I	テーマコード (参考)
G06T 13/40 (2011.01)	G06T 13/40	5B050
G06F 3/16 (2006.01)	G06F 3/16 330K	
G10L 13/00 (2006.01)	G10L 13/00 100S	
	G10L 13/00 100V	

審査請求 未請求 請求項の数 12 O L (全 23 頁)

(21) 出願番号 特願2014-55203 (P2014-55203)
(22) 出願日 平成26年3月18日 (2014. 3. 18)

(71) 出願人 000208891
KDDI株式会社
東京都新宿区西新宿二丁目3番2号
(74) 代理人 100114258
弁理士 福地 武雄
(74) 代理人 100125391
弁理士 白川 洋一
(72) 発明者 高山 伸也
埼玉県ふじみ野市大原二丁目1番15号
株式会社KDDI研究所内
(72) 発明者 池田 和史
埼玉県ふじみ野市大原二丁目1番15号
株式会社KDDI研究所内

最終頁に続く

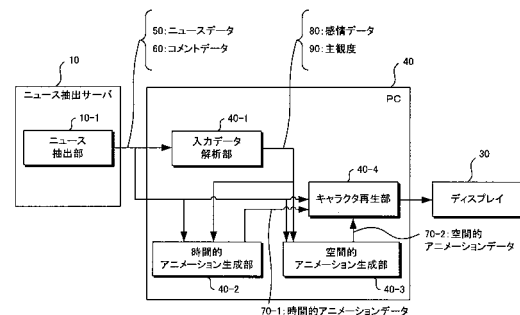
(54) 【発明の名称】 アニメーション生成装置、アニメーション生成方法およびプログラム

(57) 【要約】

【課題】 話題性の高いニュースおよびニュースに対するコメントをWEBサイトから抽出すると共に、適切なアニメーションを持つキャラクタが、ニュースだけでなくニュースに対するコメントまでも提示する。

【解決手段】 本発明のアニメーション生成装置は、任意の情報に基づいて、キャラクタのアニメーションを生成するアニメーション生成装置であって、解析された任意の複数の入力データに含まれる感情または主観度の動的特徴量に基づいて、時間的アニメーションを生成する時間的アニメーション生成部40-2と、解析された前記複数の入力データに含まれる感情または主観度の静的特徴量に基づいて、空間的アニメーションを生成する空間的アニメーション生成部40-3と、を備える。

【選択図】 図2



【特許請求の範囲】**【請求項 1】**

任意の情報に基づいて、キャラクタのアニメーションを生成するアニメーション生成装置であって、

解析された任意の複数の入力データに含まれる感情または主観度の動的特徴量に基づいて、時間的アニメーションを生成する時間的アニメーション生成部と、

解析された前記複数の入力データに含まれる感情または主観度の静的特徴量に基づいて、空間的アニメーションを生成する空間的アニメーション生成部と、を備えることを特徴とするアニメーション生成装置。

【請求項 2】

前記入力データを W E B (World Wide Web) サイトから抽出する情報抽出部を更に備えることを特徴とする請求項 1 に記載のアニメーション生成装置。

【請求項 3】

前記複数の入力データにおける感情または主観度を解析する入力データ解析部を更に備えることを特徴とする請求項 1 または請求項 2 に記載のアニメーション生成装置。

【請求項 4】

前記入力データを提示すると共に、前記生成されたキャラクタのアニメーションを再生するキャラクタ再生部を更に備えることを特徴とする請求項 1 から請求項 3 のいずれかに記載のアニメーション生成装置。

【請求項 5】

ナレータが読み上げた前記入力データの音声を取得する音声収録部を更に備えることを特徴とする請求項 1 から請求項 4 のいずれかに記載のアニメーション生成装置。

【請求項 6】

前記入力データに対応する音声データを合成する音声合成部を更に備えることを特徴とする請求項 1 から請求項 5 のいずれかに記載のアニメーション生成装置。

【請求項 7】

前記生成されたキャラクタのアニメーションと共に、前記取得された音声または前記合成された音声を再生する音声再生部を更に備えることを特徴とする請求項 1 から請求項 6 のいずれかに記載のアニメーション生成装置。

【請求項 8】

前記入力データの音声の有音区間を検出し、前記キャラクタの台詞の開始時刻および前記台詞の終了時刻を判定する台詞時刻判定部を更に備えることを特徴とする請求項 1 から請求項 7 のいずれかに記載のアニメーション生成装置。

【請求項 9】

前記台詞の開始時刻および前記台詞の終了時刻に応じて、前記アニメーションデータを変換する時間長調整部を更に備えることを特徴とする請求項 8 に記載のアニメーション生成装置。

【請求項 10】

前記アニメーションは、前記キャラクタが全身動作または表情変化する際の、任意の時刻における前記キャラクタの任意のボーン (bone) またはポリゴン (polygon) の空間座標であることを特徴とする請求項 1 から請求項 9 のいずれかに記載のアニメーション生成装置。

【請求項 11】

任意の情報に基づいて、キャラクタのアニメーションを生成するアニメーション生成方法であって、

解析された任意の複数の入力データに含まれる感情または主観度の動的特徴量に基づいて、時間的アニメーションを生成するステップと、

解析された前記複数の入力データに含まれる感情または主観度の静的特徴量に基づいて、空間的アニメーションを生成するステップと、を少なくとも含むことを特徴とするアニメーション生成方法。

10

20

30

40

50

【請求項 1 2】

任意の情報に基づいて、キャラクタのアニメーションを生成するアニメーション生成装置のプログラムであって、

解析された任意の複数の入力データに含まれる感情または主観度の動的特徴量に基づいて、時間的アニメーションを生成する処理と、

解析された前記複数の入力データに含まれる感情または主観度の静的特徴量に基づいて、空間的アニメーションを生成する処理と、の一連の処理をコンピュータに実行させることを特徴とするプログラム。

【発明の詳細な説明】**【技術分野】**

10

【0001】

本発明は、話題性の高いニュースおよびニュースに対するコメントをWEB (World Wide Web) サイトから抽出すると共に、適切なアニメーションを持つキャラクタがニュースおよびコメントを提示する技術に関する。

【背景技術】**【0002】**

従来から、キャラクタを利用してニュース情報を提供する技術が知られている。例えば、特許文献1には、リアルタイムに伝えられるニュース情報を音声に変換し、音声に対応する口形態と顔表情をキャラクタに適用して、キャラクタニュースを生成するシステムが開示されている。

20

【0003】

また、特許文献2には、ニュースに関連するSNS (Social Networking Service) のコメントを複数の指標に基づいて評価し、ニュースに合わせて各指標の重みを設定することで最適なSNSのコメント提示順序を決定する技術が開示されている。

【先行技術文献】**【特許文献】****【0004】**

【特許文献1】特許第4489121号明細書

【特許文献2】特許第4134975号明細書

【発明の概要】

30

【発明が解決しようとする課題】**【0005】**

しかしながら、特許文献1に開示されている技術では、予め定められたキャラクタがニュースを読み上げるアニメーションは生成できるが、SNS等で表現されるニュースに対するコメント内容やそのコメントに含まれる感情表現を組み合わせた適切なアニメーションは生成できない。

【0006】

一方、特許文献2に開示されている技術では、話題性の高いニュースまたはコメントを抽出することは可能だが、それらに適合したアニメーションを持つキャラクタによって、ニュースまたはコメントを提示することは困難である。

40

【0007】

本発明は、このような事情に鑑みてなされたものであり、話題性の高いニュースおよびニュースに対するコメントをWEBサイトから抽出すると共に、適切なアニメーションを持つキャラクタが、ニュースだけでなくニュースに対するコメントまでも提示するアニメーション生成装置、アニメーション生成方法およびプログラムを提供することを目的とする。また、本発明では、ニュースに対するコメントを予め集約することで、集約されたコメントを踏まえたアニメーションを持つキャラクタによるニュース本文の提示が可能となる。さらに、本発明は、表情変化だけでなくキャラクタの全身動作を加えることで、テレビ局が制作する本当のニュース番組に近いコンテンツを提供することができる。

【課題を解決するための手段】

50

【 0 0 0 8 】

(1) 上記の目的を達成するために、本発明は、以下のような手段を講じた。すなわち、本発明のアニメーション生成装置は、任意の情報に基づいて、キャラクタのアニメーションを生成するアニメーション生成装置であって、解析された任意の複数の入力データに含まれる感情または主観度の動的特徴量に基づいて、時間的アニメーションを生成する時間的アニメーション生成部と、解析された前記複数の入力データに含まれる感情または主観度の静的特徴量に基づいて、空間的アニメーションを生成する空間的アニメーション生成部と、を備えることを特徴とする。

【 0 0 0 9 】

このように、解析された任意の複数の入力データに含まれる感情または主観度の動的特徴量に基づいて、時間的アニメーションを生成し、また、解析された前記複数の入力データに含まれる感情または主観度の静的特徴量に基づいて、空間的アニメーションを生成するので、入力された情報に応じて、適切なアニメーションを有するキャラクタを作成することが可能となる。

10

【 0 0 1 0 】

(2) また、本発明のアニメーション生成装置は、前記入力データを W E B (World Wide Web) サイトから抽出する情報抽出部を更に備えることを特徴とする。

【 0 0 1 1 】

このように、入力データを W E B サイトから抽出するので、 S N S 等で表現されるニュースに対するコメント内容に応じたキャラクタのアニメーションを生成することが可能となる。

20

【 0 0 1 2 】

(3) また、本発明のアニメーション生成装置は、前記複数の入力データにおける感情または主観度を解析する入力データ解析部を更に備えることを特徴とする。

【 0 0 1 3 】

このように、複数の入力データにおける感情または主観度を解析するので、 S N S 等で表現されるニュースに対するコメントに含まれる感情表現に応じたキャラクタのアニメーションを生成することが可能となる。

【 0 0 1 4 】

(4) また、本発明のアニメーション生成装置は、前記入力データを提示すると共に、前記生成されたキャラクタのアニメーションを再生するキャラクタ再生部を更に備えることを特徴とする。

30

【 0 0 1 5 】

このように、入力データを提示すると共に、生成されたキャラクタのアニメーションを再生するので、入力された情報に応じて、適切なアニメーションを有するキャラクタを表示することが可能となる。

【 0 0 1 6 】

(5) また、本発明のアニメーション生成装置は、ナレータが読み上げた前記入力データの音声を取得する音声収録部を更に備えることを特徴とする。

【 0 0 1 7 】

このように、ナレータが読み上げた前記入力データの音声を取得するので、人間の声に基づいた音声の再生を行なうことが可能となる。

40

【 0 0 1 8 】

(6) また、本発明のアニメーション生成装置は、前記入力データに対応する音声データを合成する音声合成部を更に備えることを特徴とする。

【 0 0 1 9 】

このように、入力データに対応する音声データを合成するので、人工的に音声を作り上げることが可能となる。また、ナレータによる読み上げが不要となるため、製作コストを削減することが可能となる。

【 0 0 2 0 】

50

(7) また、本発明のアニメーション生成装置は、前記生成されたキャラクタのアニメーションと共に、前記取得された音声または前記合成された音声を再生する音声再生部を更に備えることを特徴とする。

【0021】

このように、生成されたキャラクタのアニメーションと共に、取得された音声または合成された音声を再生するので、キャラクタのアニメーションと共に、音声を出力することができる。これにより、例えば、ニュースキャスターのキャラクタにニュースを読ませるサービスを提供することが可能となる。

【0022】

(8) また、本発明のアニメーション生成装置は、前記入力データの音声の有音区間を検出し、前記キャラクタの台詞の開始時刻および前記台詞の終了時刻を判定する台詞時刻判定部を更に備えることを特徴とする。

【0023】

このように、入力データの音声の有音区間を検出し、前記キャラクタの台詞の開始時刻および前記台詞の終了時刻を判定するので、音声を入力した場合であってもキャラクタが読み上げる台詞(テキスト)データを生成することが可能となる。

【0024】

(9) また、本発明のアニメーション生成装置は、前記台詞の開始時刻および前記台詞の終了時刻に応じて、前記アニメーションデータを変換する時間長調整部を更に備えることを特徴とする。

【0025】

このように、台詞の開始時刻および前記台詞の終了時刻に応じて、アニメーションデータを変換するので、キャラクタの台詞が動的に生成される場合であっても、画像信号と音声信号の同期ズレを解消して、キャラクタのアニメーションを生成することができる。

【0026】

(10) また、本発明のアニメーション生成装置において、前記アニメーションは、前記キャラクタが全身動作または表情変化する際の、任意の時刻における前記キャラクタの任意のボーン(bone)またはポリゴン(polygon)の空間座標であることを特徴とする。

【0027】

このように、アニメーションは、キャラクタのボーン(bone)またはポリゴン(polygon)の任意の時刻における空間座標を示すデータであるので、台詞に応じた複雑なアニメーションを持つキャラクタを生成することができる。

【0028】

(11) また、本発明のアニメーション生成方法は、任意の情報に基づいて、キャラクタのアニメーションを生成するアニメーション生成方法であって、解析された任意の複数の入力データに含まれる感情または主観度の動的特徴量に基づいて、時間的アニメーションを生成するステップと、解析された前記複数の入力データに含まれる感情または主観度の静的特徴量に基づいて、空間的アニメーションを生成するステップと、を少なくとも含むことを特徴とする。

【0029】

このように、解析された任意の複数の入力データに含まれる感情または主観度の動的特徴量に基づいて、時間的アニメーションを生成し、また、解析された前記複数の入力データに含まれる感情または主観度の静的特徴量に基づいて、空間的アニメーションを生成するので、入力された情報に応じて、適切なアニメーションを有するキャラクタを作成することが可能となる。

【0030】

(12) また、本発明のプログラムは、任意の情報に基づいて、キャラクタのアニメーションを生成するアニメーション生成装置のプログラムであって、解析された任意の複数の入力データに含まれる感情または主観度の動的特徴量に基づいて、時間的アニメーションを生成する処理と、解析された前記複数の入力データに含まれる感情または主観度の静

10

20

30

40

50

的特徴量に基づいて、空間的アニメーションを生成する処理と、の一連の処理をコンピュータに実行させることを特徴とする。

【0031】

このように、解析された任意の複数の入力データに含まれる感情または主観度の動的特徴量に基づいて、時間的アニメーションを生成し、また、解析された前記複数の入力データに含まれる感情または主観度の静的特徴量に基づいて、空間的アニメーションを生成するので、入力された情報に応じて、適切なアニメーションを有するキャラクタを作成することが可能となる。

【発明の効果】

【0032】

本発明によれば、入力された情報に応じて、適切なアニメーションを有するキャラクタを作成することが可能となる。

【図面の簡単な説明】

【0033】

【図1】第1の実施形態に係るキャラクタ情報提示装置の概略構成を示す図である。

【図2】第1の実施形態に係るキャラクタ情報提示装置の機能を示すブロック図である。

【図3】第1の実施形態に係るキャラクタ情報提示装置の動作を示すフローチャートである。

【図4】第2の実施形態に係るアニメーション生成装置の概略構成を示す図である。

【図5】第2の実施形態に係るアニメーション生成装置の機能を示すブロック図である。

【図6】第2の実施形態に係るアニメーション生成装置の動作を示すフローチャートである。

【図7】テキストデータ、開始時刻および終了時刻から構成される台詞データを示す図である。

【図8】アニメーションデータを示す図である。

【図9】感情データを示す図である。

【図10】第2の実施形態に係るデータフォーマットを示す図である。

【発明を実施するための形態】

【0034】

[第1の実施形態]

本発明の実施形態に係るキャラクタ情報提示装置は、任意の情報をWEBサイトから抽出して入力し、複数の入力データにおける感情または主観度を解析し、解析された複数の入力データに含まれる感情または主観度の動的特徴量に基づいて時間的アニメーションを生成し、解析された前記複数の入力データに含まれる感情または主観度の静的特徴量に基づいて空間的アニメーションを生成し、入力データを提示すると共に、生成されたキャラクタのアニメーションを再生する。

【0035】

この構成により、入力データがSNSのコメント群のような碎けたテキストデータであっても、適切なアニメーションを持つキャラクタにより提示することができる。また、複数の入力データを包括的に解析すれば、コメント群を反映したニュース本文の提示のように、集約されたデータを踏まえたアニメーションを持つキャラクタによる情報提示が可能となる。さらに、キャラクタのアニメーションとして、キャラクタが全身動作または表情変化する際の、任意の時刻におけるキャラクタの任意のボーンまたはポリゴンの空間座標を制御すれば、テレビ局が制作する本当のニュース番組に近いコンテンツを提供することができる。

【0036】

なお、本実施形態では、WEB上のニュースおよびコメント群を用いるが、本発明の技術的思想は、これに限定されるわけではなく、SNS上コミュニティ内の発言、メッセージングによるチャットでの発言、各種紹介サイトにおける口コミ、街角インタビューの結果、交通機関の運行情報・状態情報、気象情報、占い・運勢を用いても十分である。

10

20

30

40

50

【 0 0 3 7 】

図 1 は、本実施形態に係るキャラクタ情報提示装置の概略構成を示す図である。キャラクタ情報提示装置 1 は、ニュース抽出サーバ 1 0 と、スピーカ 2 0、ディスプレイ 3 0 と、P C (Personal Computer) 4 0 とから構成されている。まず、ニュース抽出サーバ 1 0 は、話題性の高いニュースおよびコメント群を W E B サイトから抽出し、ニュースデータ 5 0 およびコメントデータ 6 0 として P C 4 0 に入力する。なお、図 1 では、ニュース抽出サーバ 1 0 が、P C 4 0 に接続されているが、本発明の技術的思想は、これに限定されるわけではなく、P C に対してオフラインによりニュースデータ 5 0 またはコメントデータ 6 0 を入力しても十分である。

【 0 0 3 8 】

一方、P C 4 0 は、ケーブル 4 0 a を介してキャラクタ情報提示装置としてのスピーカ 2 0 およびディスプレイ 3 0 に接続されている。P C 4 0 は、入力されたニュースデータ 5 0 およびコメントデータ 6 0 における感情データ 8 0 または主観度 9 0 を解析する。また、P C 4 0 は、解析された感情データ 8 0 または主観度 9 0 に含まれる感情または主観度の動的特徴量に基づいて時間的アニメーションデータ 7 0 - 1 を生成する。さらに、P C 4 0 は、解析された感情データ 8 0 または主観度 9 0 に含まれる感情または主観度の静的特徴量に基づいて空間的アニメーションデータ 7 0 - 2 を生成する。そして、P C 4 0 は、入力されたニュースデータ 5 0 およびコメントデータ 6 0 および生成されたキャラクタの時間的アニメーションデータ 7 0 - 1 および空間的アニメーションデータ 7 0 - 2 における画像信号を随時ディスプレイ 3 0 に発信する。

【 0 0 3 9 】

その際、P C 4 0 から発信された画像はディスプレイ 3 0 に A 1 として映写される。なお、本発明の実施形態に係るキャラクタ情報提示装置では、入力されたニュースデータ 5 0 およびコメントデータ 6 0 の音声を予め収録または合成し、キャラクタの時間的アニメーションデータ 7 0 - 1 および空間的アニメーションデータ 7 0 - 2 の映写と同時に、収録音声信号または合成音声信号を随時スピーカ 2 0 に発信するが、音声信号を発信しなくても良い。例えば、キャラクタが表示され、いわゆる“口パク”のように、音声はミュートされ、字幕が表示されている状態である。

【 0 0 4 0 】

図 2 は、第 1 の実施形態に係るキャラクタ情報提示装置 1 の機能を示すブロック図である。ニュース抽出サーバ 1 0 のニュース抽出部 1 0 - 1 は、話題性の高いニュースおよびコメント群を W E B サイトから抽出し、ニュースデータ 5 0 およびコメントデータ 6 0 として P C 4 0 に入力する。

【 0 0 4 1 】

P C 4 0 の入力データ解析部 4 0 - 1 は、ニュース抽出サーバ 1 0 より入力されたニュースデータ 5 0 およびコメントデータ 6 0 における感情データ 8 0 または主観度 9 0 を解析する。P C 4 0 の時間的アニメーション生成部 4 0 - 2 は、解析された感情データ 8 0 または主観度 9 0 に含まれる感情または主観度の動的特徴量に基づいて時間的アニメーションデータ 7 0 - 1 を生成する。P C 4 0 の空間的アニメーション生成部 4 0 - 3 は、解析された感情データ 8 0 または主観度 9 0 に含まれる感情または主観度の静的特徴量に基づいて空間的アニメーションデータ 7 0 - 2 を生成する。P C 4 0 のキャラクタ再生部 4 0 - 4 は、入力されたニュースデータ 5 0 およびコメントデータ 6 0 および生成されたキャラクタの時間的アニメーションデータ 7 0 - 1 および空間的アニメーションデータ 7 0 - 2 における画像信号を再生する。

【 0 0 4 2 】

図 3 は、第 1 の実施形態に係るキャラクタ情報提示装置 1 の動作を示すフローチャートである。初めに、ニュース抽出サーバ 1 0 において、W E B 上にあるニュースの収集を行なう。(ステップ S 1)。R S S 等を利用し、対象とするニュースの U R L、タイトル、記事本文、ジャンル、配信日時等の情報を収集する。

【 0 0 4 3 】

次に、ニュース抽出サーバ10において、ステップS1で収集したニュースのURLやタイトル、またはTFIDF等を用いて抽出したニュースを特徴づける単語をSNS上から検索することで、ニュースに関連するコメント群を収集する(ステップS2)。ここで、ニュースに関連するコメント群は、ニュースのURLやタイトルを、SNSが提供する検索API等を用いて検索することで、収集できる。また、ニュース記事のタイトルや本文から当該ニュースを特徴づける単語を、TFIDF等を用いて取り出し、検索に利用しても良い。このとき、TFは各ニュース記事のタイトルまたは記事本文中の各単語の出現頻度、DFには、全記事における各単語の出現頻度などを設定する。

【0044】

次に、ニュース抽出サーバ10において、ステップS2で収集したコメント群を解析し、話題性の高いニュースデータ50およびコメントデータ60を抽出する(ステップS3)。ここで、ニュース抽出サーバ10は、主観指標抽出機能と、主観度算出機能を有する。また、主観指標抽出機能は、顔文字抽出、文体抽出、盛り上がり抽出を実行する。

10

【0045】

顔文字抽出では、SNSユーザのコメント中に顔文字が含まれるコメント群を抽出する。顔文字の抽出方法は、事前に登録した辞書を用いて、コメント中に感情表現が含まれるか否かを判別すれば良い。ここで、顔文字の個数をsとする。

【0046】

[顔文字を含む主観的なSNSコメント群の例]

“えっ(;´・`・)誰得?”

20

“欲しいかも(´・`・。)”

“(;´・`・)・・・うわぁ、誰得・・・”

文体抽出では、SNSコメント群を文体に基づいて複数に分類する。文体を判別する方法は、文末に着目し、語尾の活用形(命令形、仮定形、未然形など)を利用することができる。形態素解析器を用いて文章を解析し、文末の形容詞、形容動詞、動詞の活用形を抽出する。活用形の種類数Nに対し、各活用形の出現の有無を $w_1 \dots w_n$ で表現する。例えば、 w_1 :命令形、 w_2 :仮定形、...、とし、文体が命令形の場合、 $w_1 = 1$ 、 $w_2 \dots w_n = 0$ とする。

【0047】

[文体の例]

30

“わぁ、誰得な仕様はやめろよ 命令形”

“欲しいけど、もうちょっと安かったらなぁ 仮定形”

“こんな誰得携帯は買わない。 未然形”

盛り上がり抽出では、SNSコメント群から盛り上がりを表す表現を抽出する。盛り上がりを表す表現は、文字の連続性に注目することで抽出する。同一の文字が3文字以上出現することを検出する、または同一形態素が連続して出現することを検出する。各コメントにおいて、連続する文字数または形態素数の最大値を当該コメントの盛り上がり度をeとする。

【0048】

[盛り上がり表現の例]

40

“この機能、誰得wwwwwwww 8文字のため、 $e = 8$ ”

“うおおおおお!欲しいいい! 5文字のため、 $e = 5$ ”

“おいおいおい誰得だよ 6文字(「おい」3回)のため、 $e = 6$ ”

一方、主観度算出機能では、各コメントから抽出された主観指標に基づいて、各コメントの主観度Pを算出する。主観度PはSNSコメントから抽出した(丸1)顔文字含有度合いs、(丸2)文体の特徴 $w_1 + w_2 + \dots + w_n$ 、(丸3)盛り上がり度合いeを用いて、例えば、式(1)のような線型結合の計算式で算出できる。ここで、Pの値が閾値以上であるコメント群を話題性の高いコメントデータ60として抽出する。また、このとき、当該ニュースをニュースデータ50として抽出する。

【0049】

50

【数 1】

$$P = s + w_1 + w_2 + \dots + w_n + e \quad \dots (1)$$

また、主観度算出機能では、評価指標重み付け部を含んでもよい。この場合、WEBニュースの特徴、具体的には(A)ジャンル、(B)配信元、(C)配信日時に基づいて、主観指標に重み付けをすることで、主観コメント群を選別する。WEBニュースのジャンル、配信元、配信日時などは、RSSを用いてWEBニュースを収集する際に同時に取得することが可能である。重みづけを考慮した主観度Pの算出式の例を、式(2)に示す。ここで、重み係数 α_A 、 α_B 、 α_C はそれぞれ、コメント中の顔文字表現に対する、ジャンルによる重み、配信元による重み、配信日時による重み、となる。同様に、重み係数 γ_A は文体、盛り上がり表現、に対するそれぞれの重みとなる。

10

【0050】

【数 2】

$$P = (\alpha_A + \alpha_B + \alpha_C) * s \\ + (\beta_{A1} + \beta_{B1} + \beta_{C1}) * w_1 + (\beta_{A2} + \beta_{B2} + \beta_{C2}) * w_2 + \dots + (\beta_{An} + \beta_{Bn} + \beta_{Cn}) * w_n \\ + (\gamma_A + \gamma_B + \gamma_C) * e$$

… (2)

20

例えば、ジャンルがスポーツの場合、より共感するコメントを優先的に表示するため、盛り上がり表現の重み α_A を大きく設定すると有効である。一方、ジャンルが政治・経済の場合、冷静なコメントが求められることから、 α_A の値は小さくなる。また、週刊誌が配信するニュースは、読者の感情を引き出すような見出しや文章が多いことから、感情的なコメントを優先的に表示する α_B の重みを大きくすることで、より共感できる主観コメントが得られる。さらに、休日に配信されるニュースは、平日に配信されるニュースと比べて穏やかな内容のニュースが多く、読者もそのようなコメントを望む可能性が高いことから、命令形の文体を排除するよう α_C の重みを設定する、といったコメント群の選別方法が有効である。

【0051】

30

次に、ステップS3で抽出されたニュースデータ50およびコメントデータ60をPC40に入力し、入力されたニュースデータ50およびコメントデータ60における感情データ80または主観度90を解析する(ステップS4)。ここでは、入力されるニュースデータ50およびコメントデータ60が単語列の場合を説明するが、本発明の技術的思想は、単語に限定されるわけではなく、句であっても文であっても良い。本実施形態では、感情分類とそれぞれの強度から構成される感情語データベースを使用する。なお、感情語データベースは、日本語辞書などの大量の単語データベースに存在する全ての単語に対して、人間の表情形成に用いられる感情分類「喜び」「悲しみ」「怒り」「嫌悪」「恐怖」「罪」「恥」「興味」「驚き」の9つがどれくらいの割合で存在するか規定し、それぞれの感情強度を0~1の範囲で、0.1刻みに10段階で指定して、予め形成されている。

40

【0052】

また、入力されるニュースデータ50またはコメントデータ60が句または文である場合は、単語の場合と同様に、句または文に全体における感情分類とそれぞれの強度から構成される感情句データベースまたは感情文データベースを使用すれば良い。ここで、「喜び」の強度を S_1 、「悲しみ」の強度を S_2 、「怒り」の強度を S_3 、「嫌悪」の強度を S_4 、「恐怖」の強度を S_5 、「罪」の強度を S_6 、「恥」の強度を S_7 、「興味」の強度を S_8 、「驚き」の強度を S_9 と表す。ただし、式(3)を満たす。

【0053】

【数 3】

$$S_i(i=1, \dots, 9) = \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\} \quad \dots (3)$$

そして、入力されるニュースデータ 50 およびコメントデータ 60 と一致または類似する単語を感情語データベースの中から検索し、その単語における感情分類と強度を抽出し、感情データ 80 を生成する。すなわち、ニュースデータ 50 における感情分類は、式 (4) で表わされる。

【0054】

【数 4】

$$(S_1(N), S_2(N), S_3(N), S_4(N), S_5(N), S_6(N), S_7(N), S_8(N), S_9(N)) \quad \dots (4)$$

10

コメントデータ 60 における感情分類は、式 (5) で表わされる。

【0055】

【数 5】

$$(S_1(C_k), S_2(C_k), S_3(C_k), S_4(C_k), S_5(C_k), S_6(C_k), S_7(C_k), S_8(C_k), S_9(C_k)) \quad \dots (5)$$

20

このように、9次元のベクトルで示される。ここで k は、入力されるコメントデータ 60 の ID を示している。また、入力される台詞のニュースデータ 50 およびコメントデータ 60 における全体強度 W は、式 (6) で表わされる。

【0056】

【数 6】

$$W = \sum_{i=1}^9 S_i(N) \quad W = \sum_{i=1}^9 S_i(C_k) \quad \dots (6)$$

30

例えば、コメントデータ 60 として、「泣き笑い」を入力すると、式 (7) が感情データ 80 として生成される。

【0057】

【数 7】

$$(0.4, 0.6, 0, 0, 0, 0, 0.2, 0, 0), \quad W = 1.2 \quad \dots (7)$$

一方、主観度 90 は、式 (2) を用いて、ニュースデータ 50 の主観度 P(N) およびコメントデータ 60 の主観度 P(C_k) を算出する。

40

【0058】

次に、PC40 において、ステップ S4 で解析された感情データ 80 または主観度 90 に含まれる感情または主観度の動的特徴量に基づいて時間的アニメーションデータ 70 - 1 を生成する (ステップ S5)。本実施形態では、まず、キャラクタの全身動作または表情変化の際のアニメーションデータと、日本語辞書などの代表的な単語データベースに存在する各単語をパラメータとして算出した各アニメーションデータにおける類似度パラメータ、のペアで構成されるアニメーションデータベースを用意する。なお、ここで言うアニメーションデータは、キャラクタの全てのボーンまたはポリゴンが任意の時刻において何れの空間位置座標に存在するかが記述されている。

【0059】

50

ニュースデータ50およびコメントデータ60における類似度パラメータを算出し、アニメーションデータベースの中にある全てのアニメーションデータにおける類似度パラメータとのコサイン類似度を求め、最も値の大きいアニメーションデータをキャラクタのアニメーションデータ70として選定する。ここで、選定されたアニメーションデータ70にキャラクタにおける任意のボーンの空間座標が記録されている場合について述べる。本実施形態では、予め収録された全ての全身動作または表情変化における無感情のアニメーションデータおよび各感情分類のアニメーションデータを学習させ、無感情のアニメーションデータから各感情分類のアニメーションデータへの変換を予め定義し、選定されたアニメーションデータ70を、ステップS4で解析された感情データ80の分散 V_i (式(8))または主観度90の分散 V_i (式(9))が閾値を超えた感情分類*i*に対し、動的特徴量 Δ_i (式(10)、式(11))に応じて合成変換することで、キャラクタの全身動作または表情変化における時間的アニメーションデータ70-1を生成する。

【数8】

【数8】

$$V_i = \frac{1}{k+1} \sum_{N, C_k} \left(S_i(N, C_k) - \frac{\sum_{N, C_k} S_i(N, C_k)}{k+1} \right)^2 \quad \dots (8)$$

10

20

【数9】

$$V_i = \frac{1}{k+1} \sum_{N, C_k} \left(P_i(N, C_k) - \frac{\sum_{N, C_k} P_i(N, C_k)}{k+1} \right)^2 \quad \dots (9)$$

【数10】

$$\Delta_i = S_i(C_k) - S_i(N) \quad \text{または} \quad S_i(C_{k+1}) - S_i(C_k) \quad \dots (10)$$

30

【数11】

$$\Delta_i = P_i(C_k) - P_i(N) \quad \text{または} \quad P_i(C_{k+1}) - P_i(C_k) \quad \dots (11)$$

本明細書では、一例として、主成分分析を用いてキャラクタの全身動作または表情変化に感情を付与するが、本発明の技術的思想は、主成分分析に限定されるわけではなく、非線形形状空間写像や機械学習等、別の方法を用いて全身動作または表情変化に感情を付与しても良い。まず、無感情および人間の表情形成に用いられる感情分類「喜び」「悲しみ」「怒り」「嫌悪」「恐怖」「罪」「恥」「興味」「驚き」の9つの感情の全身動作または表情変化におけるアニメーションデータを収録等により予め複数用意し、学習用データベースに登録しておく。登録された全ての無感情のアニメーションデータおよび各感情分類のアニメーションデータを学習し、線形回帰手法により、無感情のアニメーションデータから各感情分類のアニメーションデータに変換するためのパラメータを算出する。すなわち、予め用意された全身動作または表情変化を m ($m = 1, 2, \dots$) とすると、全身動作または表情変化 m における無感情のアニメーションデータの第 j 主成分座標の微分値 k^j

40

50

(m) を用いて、式 (12) により線形回帰演算を行なって、変換パラメータ a_i^j 、 b_i^j を算出する。

【0061】

【数12】

$$q_i^j(m) = a_i^j k^j(m) + b_i^j \quad \dots (12)$$

ただし、 $q_i^j(m)$ は全身動作または表情変化 m における各感情のアニメーションデータの第 j 主成分座標の微分値を示しており、 $i = 1$ は感情分類が「喜び」の場合を、 $i = 2$ は感情分類が「悲しみ」の場合を、 $i = 3$ は感情分類が「怒り」の場合を、 $i = 4$ は感情分類が「嫌悪」の場合を、 $i = 5$ は感情分類が「恐怖」の場合を、 $i = 6$ は感情分類が「罪」の場合を、 $i = 7$ は感情分類が「恥」の場合を、 $i = 8$ は感情分類が「興味」の場合を、 $i = 9$ は感情分類が「驚き」の場合をそれぞれ表す。

10

【0062】

次に、算出された a_i^j に対応する変換パラメータ a_i^j 、 b_i^j を用いて、選定されたアニメーションデータを合成変換し、キャラクタの全身動作または表情変化における時間的アニメーションデータ 70 - 1 を生成する。すなわち、時間的アニメーションデータ 70 - 1 の第 j 主成分座標の微分値 $p^j(h)$ は、式 (13) となる。

【数13】

20

$$p^j(h) = \sum_{i=1}^9 \frac{\Delta_i}{W} (a_i^j k^j(h) + b_i^j) \quad \dots (13)$$

これにより、例えば、続々と入力されるニュースデータ 50 が段階的に暗くなる場合は、徐々に沈静化したアニメーションを時間的アニメーションデータ 70 - 1 として生成することが可能となる。また、主観度 90 を用いることで、例えば、ニュースデータ 50 のジャンルがスポーツの場合、引き分けの試合の結果を伝えるニュースデータ 50 および「まずまずの結果」というコメントデータ 60 に対して、主観度 90 が大きい場合は、アニメーションデータベースの中からポジティブなアニメーションを象徴する「親指を上げる」アニメーションを時間的アニメーションデータ 70 - 1 として生成することができる。一方、主観度 90 が小さい場合は、ネガティブなアニメーションを象徴する「肩をすくめる」アニメーションを時間的アニメーションデータ 70 - 1 として生成することができる。

30

【0063】

次に、PC 40 において、ステップ S 4 で解析された感情データ 80 または主観度 90 に含まれる感情または主観度の動的特徴量に基づいて時間的アニメーションデータ 70 - 2 を生成する (ステップ S 6)。ステップ S 5 で選定されたアニメーションデータ 70 を、ステップ S 4 で解析された感情データ 80 の分散 V_i が閾値を下回った感情分類 i に対し、静的特徴量 E_i (式 (14)、式 (15)) に応じて、背景色の変更、キャラクタの立ち位置の変更等の空間的アニメーションデータ 70 - 2 を生成する。

40

【数14】

$$E_i = \frac{\sum_{N, C_k} S_i(N, C_k)}{k+1} \quad \dots (14)$$

【数 15】

$$E_i = \frac{\sum_{N, C_k} P_i(N, C_k)}{k+1} \quad \dots (15)$$

例えば、 E_5 の値が大きく算出された場合、「恐怖」が継続的であるので、背景色を少し暗くし、入力されるニュースデータ 50 またはコメントデータ 60 が明るい内容であっても、暗い印象を与えることが可能となる。また、 E_7 の値が大きく算出された場合、「恥」が継続的であるので、キャラクタの立ち位置を少し遠ざけ、入力されるニュースデータ 50 またはコメントデータ 60 の内容に応じず恥ずかしい印象を与える可能となる。

10

【0064】

次に、PC 40 において、ステップ S 5 およびステップ S 6 で生成されたキャラクタの時間的アニメーションデータ 70 - 1 および空間的アニメーションデータ 70 - 2 における画像信号を生成する（ステップ S 7）。まず、ニュースデータ 50 を提示するキャラクタの描画を開始して画像信号の生成を開始する。次に、ステップ S 5 およびステップ S 6 で生成されたニュースデータ 50 における時間的アニメーションデータ 70 - 1 および空間的アニメーションデータ 70 - 2 を反映したキャラクタを描画し、画像信号を生成する。その後、ニュースデータ 50 における画像信号の生成が終了すると、コメントデータ 60 を提示するキャラクタの描画を開始して画像信号の生成を開始する。

20

【0065】

次に、ステップ S 5 およびステップ S 6 で生成されたコメントデータ 60 における時間的アニメーションデータ 70 - 1 および空間的アニメーションデータ 70 - 2 を反映したキャラクタを描画し、画像信号を生成する。全てのコメントデータ 60 に対して同様の処理を繰り返し、コメントデータ 60 における画像信号を生成する。最後のコメントデータ 60 に対しての処理が終了した後、別のニュースを提示する場合は、ステップ S 1 に戻り、次のニュースデータ 50 およびコメントデータ 60 に対しての処理を開始する。なお、本発明の技術的思想は、ニュースデータ 50 を提示するキャラクタとコメントデータ 60 を提示するキャラクタは、それぞれ同じであっても異なってもどちらでも良い。次に、未処理データがあるかどうかを判断し（ステップ S 7 - 2）、ある場合は、ステップ S 1 に遷移し、無い場合は、ステップ S 8 に遷移する。

30

【0066】

最後に、ステップ S 4 で入力されたニュースデータ 50 およびコメントデータ 60 の画像信号と共に、ステップ S 7 で生成されたキャラクタの画像信号をディスプレイ 10 に映写する（ステップ S 8）。

【0067】

このように、本実施形態によれば、話題性の高いニュースおよびコメント群を WEB サイトから抽出し、ニュースデータ 50 およびコメントデータ 60 として PC 40 に入力し、入力されたニュースデータ 50 およびコメントデータ 60 における感情データ 80 または主観度 90 を解析し、解析された感情データ 80 または主観度 90 に含まれる感情または主観度の動的特徴量に基づいて時間的アニメーションデータ 70 - 1 を生成し、解析された感情データ 80 または主観度 90 に含まれる感情または主観度の静的特徴量に基づいて空間的アニメーションデータ 70 - 2 を生成し、入力されたニュースデータ 50 およびコメントデータ 60 および生成されたキャラクタの時間的アニメーションデータ 70 - 1 および空間的アニメーションデータ 70 - 2 における画像信号を再生するので、入力データが SNS のコメント群のような碎けたテキストデータであっても、適切なアニメーションを持つキャラクタにより提示することができる。また、複数の入力データを包括的に解析すれば、コメント群を反映したニュース本文の提示のように、集約されたデータを踏ま

40

50

えたアニメーションを持つキャラクタによる情報提示が可能となる。さらに、キャラクタのアニメーションとして、キャラクタが全身動作または表情変化する際の、任意の時刻におけるキャラクタの任意のボーンまたはポリゴンの空間座標を制御すれば、テレビ局が制作する本当のニュース番組に近いコンテンツを提供することができる。

【 0 0 6 8 】

[第 2 の実施形態]

第 2 の実施形態に係るアニメーション生成装置は、台詞である音声信号を入力し、入力された音声信号を時系列で記録した音声データに基づいて台詞のテキストデータおよび台詞の開始時刻および終了時刻を時系列で生成する。この生成された台詞のテキストデータおよび台詞の開始時刻および終了時刻に基づいてキャラクタのアニメーションデータを時系列で生成する。アニメーションデータには感情データを伴っていても良く、生成された感情データに基づいてアニメーションデータに感情を付与し、台詞の開始時刻および終了時刻に基づいてアニメーションデータの時間長を調整する。この生成された台詞の開始時刻および終了時刻に応じて、音声データにおける音声信号および生成されたアニメーションデータにおける画像信号を生成する。

10

【 0 0 6 9 】

これにより、キャラクタの台詞が動的に生成される場合であっても、画像信号と音声信号の同期ズレを解消して、キャラクタを生成することができる。また、アニメーションデータとしてキャラクタの任意のボーンまたはポリゴンの空間座標を制御した画像信号を生成するので、台詞に応じた複雑なアニメーションを持つキャラクタを生成することができる。

20

【 0 0 7 0 】

図 4 は、第 2 の実施形態に係るアニメーション生成装置の概略構成を示す図である。このアニメーション生成装置は、マイクロフォン 2 1 0 と、スピーカ 2 2 0、ディスプレイ 2 3 0 と、P C (Personal Computer) 2 4 0 とから構成されている。そして、マイクロフォン 2 1 0 により P C 2 4 0 に音声データ 2 5 0 が入力される。なお、図 4 では、マイクロフォン 2 1 0 が、P C 2 4 0 に接続されているが、本発明の技術的思想は、これに限定されるわけではなく、任意手段により P C に対して音声データが入力されれば十分である。

30

【 0 0 7 1 】

P C 2 4 0 は、ケーブル 2 4 0 a を介してアニメーション生成装置としてのスピーカ 2 2 0 およびディスプレイ 2 3 0 に接続されている。P C 2 4 0 では、台詞として入力された音声信号を時系列で記録した音声データ 2 5 0 に基づいて台詞のテキストデータまたは台詞の開始時刻および終了時刻を時系列で生成する。また、生成された台詞のテキストデータおよび台詞の開始時刻および終了時刻に基づいて、キャラクタのアニメーションデータおよび感情データを時系列で生成する。さらに、生成された感情データに基づいてアニメーションデータに感情を付与すると共に生成された台詞の開始時刻および終了時刻に基づいてアニメーションデータの時間長を調整する。

【 0 0 7 2 】

そして、P C 2 4 0 は、生成された台詞の開始時刻および終了時刻に応じて、音声データにおける音声信号を随時スピーカ 2 2 0 に発信する。一方、P C 2 4 0 は、生成されたアニメーションデータにおける画像信号を随時ディスプレイ 2 3 0 に発信する。その際、P C 2 4 0 において生成された画像はディスプレイ 2 3 0 に A 1 として映写される。なお、本発明の実施形態に係るアニメーション生成装置では、音声データにおける音声信号またはアニメーションデータにおける画像信号をスピーカ 2 2 0 またはディスプレイ 2 3 0 に発信しなくても構わない。

40

【 0 0 7 3 】

図 5 は、第 2 の実施形態に係るアニメーション生成装置の機能を示すブロック図である。P C 2 4 0 の台詞テキスト生成部 2 4 0 - 1 は、マイクロフォン 2 1 0 により台詞として入力された音声信号を時系列で記録した音声データ 2 5 0 を認識して、台詞のテキスト

50

データ260-1を時系列で生成する。また、PC240の台詞時刻判定部240-2は、音声データ250の有音区間を検出して、台詞の開始時刻および終了時刻260-2を時系列で判定する。また、PC240のアニメーション生成部240-3は、台詞のテキストデータ260-1の内容を解析して、台詞の開始時刻および終了時刻260-2毎に、キャラクタのアニメーションデータ270を時系列で生成する。

【0074】

また、PC240の感情生成部240-4は、台詞のテキストデータ260-1の内容を解析して、キャラクタの感情データ280を時系列で生成する。また、PC240の感情付与部240-5は、感情データ280に応じて、アニメーションデータ270を変換する。また、PC240の時間長調整部240-6は、台詞の開始時刻および終了時刻260-2に基づいて、アニメーションデータ270を変換する。また、PC240のキャラクタ再生部240-7は、台詞の開始時刻および終了時刻260-2に応じて、音声データ250における音声信号および生成されたアニメーションデータ270における画像信号を再生する。

10

【0075】

図6は、第2の実施形態に係るアニメーション生成装置の動作を示すフローチャートである。初めに、PC240に、マイクロフォン210により音声信号を入力し、入力された音声信号を時系列で記録して音声データ250を形成する(ステップS21)。次に、PC240において、音声データ250を認識して、台詞のテキストデータを時系列で生成する(ステップS22)。

20

【0076】

次に、PC240において、音声データ250の有音区間を検出して、台詞の開始時刻および終了時刻260-2を時系列で判定する(ステップS23)。音声データ250において、音圧レベルが閾値より大きい状況が一定時間以上継続した場合、該当区間を有音区間とし、区間の先頭を台詞の開始時刻、区間の終端を台詞の終了時刻と判定する。また、ステップS22で生成された台詞のテキストデータを、判定された台詞の開始時刻から終了時刻までを1つの単位として、台詞のテキストデータ260-1に分割する。

【0077】

次に、PC240において、ステップS23で分割された台詞のテキストデータ260-1の内容を解析して、ステップS23で判定された台詞の開始時刻および終了時刻260-2毎に、「目を閉じる」「お辞儀する」「手を挙げる」等のキャラクタのアニメーションデータ270を時系列で生成する(ステップS24)。本実施形態では、キャラクタの全身動作または表情変化の際のアニメーションデータと、日本語辞書などの代表的な単語データベースに存在する各単語をパラメータとして算出した各アニメーションデータにおける類似度パラメータ、のペアで構成されるアニメーションデータベースを用意する。

30

【0078】

なお、ここで言うアニメーションデータは、キャラクタの全てのボーンまたはポリゴンが任意の時刻において何れの空間位置座標に存在するかが記述されている。台詞のテキストデータ260-1における類似度パラメータを算出し、アニメーションデータベースの中にある全てのアニメーションデータにおける類似度パラメータとのコサイン類似度を求め、最も値の大きいアニメーションデータをキャラクタのアニメーションデータ270として時系列で生成する。

40

【0079】

次に、PC240において、ステップS23で分割された台詞のテキストデータ260-1の内容を解析して、感情分類および強度から構成されるキャラクタの感情データ280を時系列で生成する(ステップS25)。ここでは、入力される台詞のテキストデータ260-1が単語列の場合を説明するが、本発明の技術的思想は、単語に限定されるわけではなく、句であっても文であっても良い。本実施形態では、感情分類とそれぞれの強度から構成される感情語データベースを使用する。

【0080】

50

なお、感情語データベースは、日本語辞書などの大量の単語データベースに存在する全ての単語に対して、人間の表情形成に用いられる感情分類「喜び」「悲しみ」「怒り」「嫌悪」「恐怖」「罪」「恥」「興味」「驚き」の9つがどれくらいの割合で存在するか規定し、それぞれの感情強度を0～1の範囲で、0.1刻みに10段階で指定して、予め形成されている。また、入力される台詞のテキストデータ260-1が句または文である場合は、単語の場合と同様に、句または文に全体における感情分類とそれぞれの強度から構成される感情句データベースまたは感情文データベースを使用すれば良い。ここで、「喜び」の強度を S_1 、「悲しみ」の強度を S_2 、「怒り」の強度を S_3 、「嫌悪」の強度を S_4 、「恐怖」の強度を S_5 、「罪」の強度を S_6 、「恥」の強度を S_7 、「興味」の強度を S_8 、「驚き」の強度を S_9 と表す。ただし、次式を満たすものとする。

【0081】
【数16】

$$S_i (i = 1, \dots, 9) = \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$$

…(16)

そして、入力される台詞のテキストデータ260-1と一致または類似する単語を感情語データベースの中から検索し、その単語における感情分類と強度を抽出し、感情データ280を生成する。すなわち、台詞のテキストデータ260-1における感情分類は、次式のように、9次元のベクトルで示される。

【0082】
【数17】

$$(S_1, S_2, S_3, S_4, S_5, S_6, S_7, S_8, S_9) \quad \dots(17)$$

また、入力される台詞のテキストデータ260-1における全体強度 W は、次式で表わされる。

【0083】
【数18】

$$W = \sum_{i=1}^9 S_i \quad \dots(18)$$

例えば、台詞のテキストデータ260-1として、「泣き笑い」を入力すると、次式が感情データ280として生成される。

【0084】
【数19】

$$(0.4, 0.6, 0, 0, 0, 0, 0.2, 0, 0) \quad W = 1.2 \quad \dots(19)$$

10

20

30

40

50

次に、PC230において、感情データ280に応じてアニメーションデータ270を変換し、キャラクタの全身動作または表情変化に感情を付与する（ステップS26）。まず、アニメーションデータ270にキャラクタにおける任意のボーンの空間座標が時系列で記録されている場合について述べる。本実施形態では、予め収録された全ての全身動作または表情変化における無感情のアニメーションデータおよび各感情分類のアニメーションデータを学習させ、無感情のアニメーションデータから各感情分類のアニメーションデータへの変換を予め定義し、ステップS24で生成されたアニメーションデータ270を、ステップS25で生成された感情データ280に応じて合成変換することで、キャラクタの全身動作または表情変化に感情を付与する。

【0085】

10

本明細書では、一例として、主成分分析を用いてキャラクタの全身動作または表情変化に感情を付与するが、本発明の技術的思想は、主成分分析に限定されるわけではなく、非線形形状空間写像や機械学習等、別の方法を用いて全身動作または表情変化に感情を付与しても良い。まず、無感情および人間の表情形成に用いられる感情分類「喜び」「悲しみ」「怒り」「嫌悪」「恐怖」「罪」「恥」「興味」「驚き」の9つの感情の全身動作または表情変化におけるアニメーションデータを収録等により予め複数用意し、学習用データベースに登録しておく。登録された全ての無感情のアニメーションデータおよび各感情分類のアニメーションデータを学習し、線形回帰手法により、無感情のアニメーションデータから各感情分類のアニメーションデータに変換するためのパラメータを算出する。

【0086】

20

すなわち、予め用意された全身動作または表情変化を m ($m = 1, 2, \dots$) とすると、全身動作または表情変化 m における無感情のアニメーションデータの第 j 主成分座標の微分値 $k^j(m)$ を用いて、次式により線形回帰演算を行なって、変換パラメータ a^j_i 、 b^j_i を算出する。

【0087】

【数20】

$$q_i^j(m) = a_i^j k^j(m) + b_i^j \quad \dots(20)$$

30

ただし、 $q^j_i(m)$ は全身動作または表情変化 m における各感情のアニメーションデータの第 j 主成分座標の微分値を示しており、 $i = 1$ は感情分類が「喜び」の場合を、 $i = 2$ は感情分類が「悲しみ」の場合を、 $i = 3$ は感情分類が「怒り」の場合を、 $i = 4$ は感情分類が「嫌悪」の場合を、 $i = 5$ は感情分類が「恐怖」の場合を、 $i = 6$ は感情分類が「罪」の場合を、 $i = 7$ は感情分類が「恥」の場合を、 $i = 8$ は感情分類が「興味」の場合を、 $i = 9$ は感情分類が「驚き」の場合をそれぞれ表す。次に、ステップS25で生成された感情データ280に対応する変換パラメータ a^j_i 、 b^j_i を用いて、ステップS24で生成されたアニメーションデータ270を合成変換し、キャラクタの全身動作または表情変化に感情を付与する。すなわち、感情が付与されたアニメーションデータ270の第 j 主成分座標の微分値 $p^j(m)$ は、次式のようになる。

40

【0088】

【数21】

$$p^j(m) = \sum_{i=1}^9 \frac{S_i}{W} (a_i^j k^j(m) + b_i^j) \quad \dots(21)$$

時系列で記録されたアニメーションデータ270の全てに対して同様の操作を実行し、

50

アニメーションデータ270を再生成する。

【0089】

次に、アニメーションデータ270にキャラクタにおける任意のポリゴンの空間座標が時系列で記録されている場合について述べる。本実施形態では、ステップS24で生成されたアニメーションデータ270を、ステップS25で生成された感情データ280に応じて合成変換することで、キャラクタの全身動作または表情変化に感情を付与する。まず、人間の表情形成に用いられる感情分類「喜び」「悲しみ」「怒り」「嫌悪」「恐怖」「罪」「恥」「興味」「驚き」の9つの感情の全身動作または表情変化におけるアニメーションデータに対して、無感情のアニメーションデータからの変化量を予め定義しておく。ここで、アニメーションデータの中のあるポリゴンPの座標に対して、「喜び」における無感情との差分を (X_1, Y_1, Z_1) 、「悲しみ」における無感情との差分を (X_2, Y_2, Z_2) 、「怒り」における無感情との差分を (X_3, Y_3, Z_3) 、「嫌悪」における無感情との差分を (X_4, Y_4, Z_4) 、「恐怖」における無感情との差分を (X_5, Y_5, Z_5) 、「罪」における無感情との差分を (X_6, Y_6, Z_6) 、「恥」における無感情との差分を (X_7, Y_7, Z_7) 、「驚き」における無感情との差分を (X_8, Y_8, Z_8) 、「驚き」における無感情との差分を (X_9, Y_9, Z_9) 、と表す。

10

【0090】

そして、ステップS25で生成された感情データ280に応じて、ステップS24で生成されたアニメーションデータ270を合成変換し、キャラクタの全身動作または表情変化に感情を付与する。すなわち、アニメーションデータ270におけるPの座標を (X_0, Y_0, Z_0) とすると、感情が付与されたアニメーションデータ270におけるPの座標 (X_E, Y_E, Z_E) は、次式で表わされる。

20

【0091】

【数22】

$$(X_E, Y_E, Z_E) \\ = (X_0 + \sum_{i=1}^9 \frac{S_i}{W} X_i, Y_0 + \sum_{i=1}^9 \frac{S_i}{W} Y_i, Z_0 + \sum_{i=1}^9 \frac{S_i}{W} Z_i)$$

30

…(22)

時系列で記録されたアニメーションデータ270の全てに対して同様の操作を実行し、アニメーションデータ270を再生成する。

【0092】

次に、PC240において、ステップS23で判定された台詞の開始時刻および終了時刻260-2に基づいて、ステップS24またはステップS26で生成されたアニメーションデータ270を変換し、アニメーションデータ270の時間長を調整する(ステップS27)。本実施形態では、キャラクタの一連の全身動作または表情変化におけるアニメーションデータ内のボーンまたはポリゴンの空間座標を「開始部」「主要部」「収束部」等の動きの内容に応じて予め分割して、それらの境界となる時刻をキーフレームと指定し、キーフレーム間の遷移が規定された「アニメーショングラフ」を用いて、アニメーションデータ270を変換する。

40

【0093】

本明細書では、一例として、アニメーショングラフを用いてアニメーションデータ270の時間長を調整するが、本発明の技術的思想は、アニメーショングラフに限定されるわけではなく、機械学習やHidden Markov Model等の確率モデルも用いてアニメーションデータ270の時間長を調整しても良い。まず、ステップS24で用意されたアニメーションデータベースにおける全てのアニメーションデータに対して、予め「開始部」「主要部」「収束部」を定めてキーフレームを指定し、上述したアニメーショングラフを用意する

50

。

【0094】

次に、ステップS23で判定された台詞の開始時刻および終了時刻よりステップS23で分割された台詞のテキストデータ260-1の時間長を求める。そして、ステップS24またはステップS26で生成されたアニメーションデータ270の時間長に、アニメーションデータ270における「主要部」の時間長を、台詞の時間長との差分が最も小さくなるように繰り返し加え、繰り返し回数を保持する。最後に、アニメーションデータ270の「収束部」の直前に、保持された繰り返し回数分だけ「主要部」のアニメーションデータを挿入する形で、アニメーションデータ270を再生成する。本明細書では、一例として、アニメーションデータにおける「主要部」を繰り返して時間長を調整するが、本発明の技術的思想は、「主要部」の繰り返しに限定されるわけではなく、アニメーションデータの速度制御や「開始部」「収束部」の切り捨て等でアニメーションデータ270の時間長を調整しても良い。

10

【0095】

次に、PC240において、ステップS23で判定された台詞の開始時刻および終了時刻260-2に応じて、音声データ250における音声信号ステップS27で生成されたアニメーションデータ270における画像信号を生成する(ステップS28)。まず、音声データ250の再生を開始して音声信号の生成を開始すると共に、キャラクタの描画を開始して画像信号の生成を開始する。次に、音声データ250の再生時刻がステップS23で判定された台詞の開始時刻に到達すると、ステップS27で生成されたアニメーションデータ270を反映したキャラクタの描画を開始して画像信号を生成する。

20

【0096】

一方、音声データ250の再生時刻がステップS23で判定された台詞の終了時刻に到達すると、アニメーションデータ270が反映されない標準のキャラクタの描画に戻り画像信号を生成する。時系列で記録された全ての台詞の開始時刻および終了時刻260-2に対して、同様の処理を繰り返し、アニメーションデータ270における画像信号を生成する。最後の台詞の開始時刻および終了時刻260-2に対しての処理が終了した後、マイクロフォン210により入力される音声データ250が続く場合は、ステップS21に戻り、次の音声データ250に対しての処理を開始する。

30

【0097】

最後に、音声データ250をスピーカ220から放射すると共に、生成されたキャラクタの画像信号をディスプレイ230に映写する(ステップS29)。

【0098】

図10は、第2の実施形態に係るデータフォーマットを示す図である。このように、第2の実施形態によれば、台詞として入力された音声信号を時系列で記録した音声データ250に基づいて台詞のテキストデータおよび台詞の開始時刻および終了時刻を時系列で生成する。また、生成された台詞のテキストデータに基づいて台詞の開始時刻および終了時刻毎にキャラクタのアニメーションデータおよび感情データを時系列で生成する。さらに、生成された感情データに基づいてアニメーションデータに感情を付与すると共に生成された台詞の開始時刻および終了時刻に基づいてアニメーションデータの時間長を調整する。そして、PC240は、生成された台詞の開始時刻および終了時刻に応じて、音声データにおける音声信号を随時スピーカ220に発信する。一方、PC240は、アニメーションデータにおける画像信号を随時ディスプレイ230に発信する。

40

【0099】

これにより、第2の実施形態では、キャラクタの台詞が動的に生成される場合であっても、画像信号と音声信号の同期ズレを解消して、キャラクタを生成することができる。また、アニメーションデータとしてキャラクタの任意のボーンまたはポリゴンの空間座標を制御した画像信号を生成するので、台詞に応じた複雑なアニメーションを持つキャラクタを生成することができる。

【0100】

50

以上説明したように、本発明によれば、任意の情報をWEBサイトから抽出して入力し、複数の入力データにおける感情または主観度を解析し、解析された複数の入力データに含まれる感情または主観度の動的特徴量に基づいて時間的アニメーションを生成し、解析された前記複数の入力データに含まれる感情または主観度の静的特徴量に基づいて空間的アニメーションを生成し、入力データを提示すると共に、生成されたキャラクタのアニメーションを再生するので、入力データがSNSのコメント群のような碎けたテキストデータであっても、適切なアニメーションを持つキャラクタにより提示することができる。また、複数の入力データを包括的に解析すれば、コメント群を反映したニュース本文の提示のように、集約されたデータを踏まえたアニメーションを持つキャラクタによる情報提示が可能となる。さらに、キャラクタのアニメーションとして、キャラクタが全身動作または表情変化する際の、任意の時刻におけるキャラクタの任意のボーンまたはポリゴンの空間座標を制御すれば、テレビ局が制作する本当のニュース番組に近いコンテンツを提供することができる。

10

【符号の説明】

【0101】

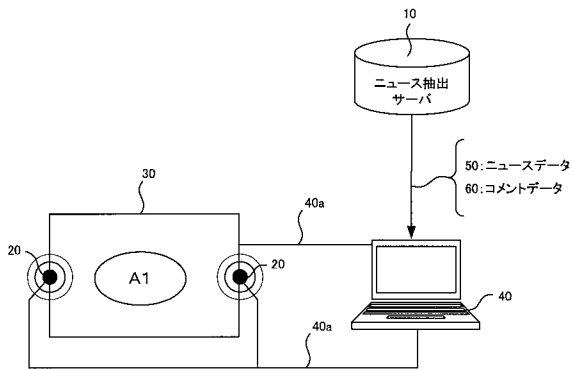
- 10 ニュース抽出サーバ
- 10-1 ニュース抽出サーバ
- 20 スピーカ
- 30 ディスプレイ
- 40 PC
- 40-1 入力データ解析部
- 40-2 時間的アニメーション生成部
- 40-3 空間的アニメーション生成部
- 40-4 キャラクタ再生部
- 40a ケーブル
- 50 ニュースデータ
- 60 コメントデータ
- 70-1 時間的アニメーションデータ
- 70-2 空間的アニメーションデータ
- 80 感情データ
- 90 主観度
- 210 マイクロフォン
- 220 スピーカ
- 230 ディスプレイ
- 240 PC
- 240-1 台詞テキスト生成部
- 240-2 台詞時刻判定部
- 240-3 アニメーション生成部
- 240-4 感情生成部
- 240-5 感情付与部
- 240-6 時間長調整部
- 240-7 キャラクタ再生部
- 240a ケーブル
- 250 音声データ
- 260-1 テキストデータ
- 260-2 開始時刻および終了時刻
- 270 アニメーションデータ
- 280 感情データ

20

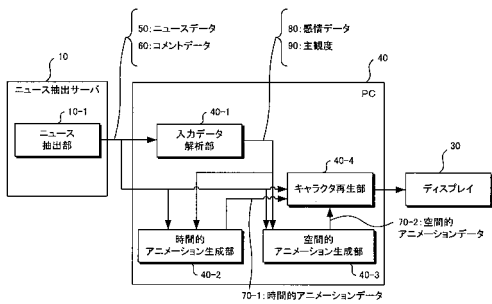
30

40

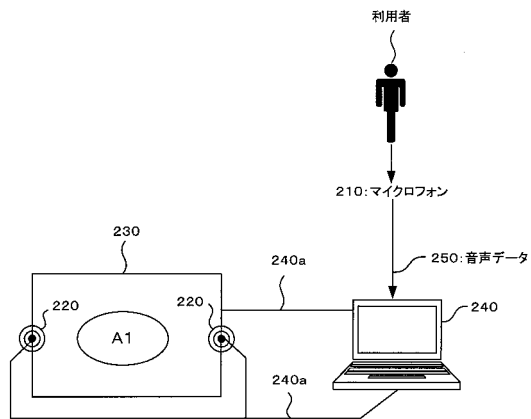
【 図 1 】



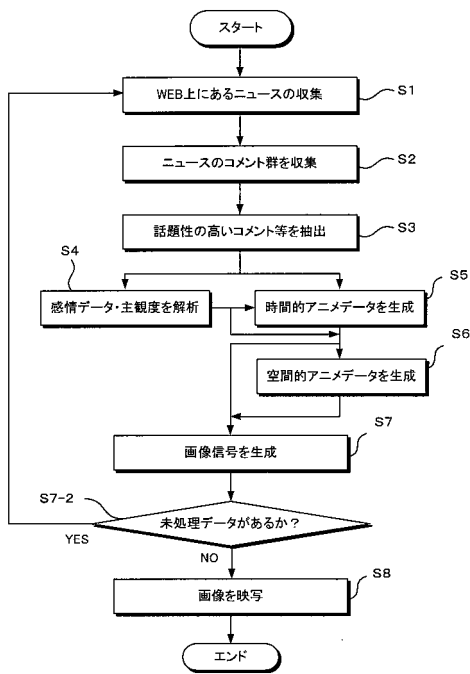
【 図 2 】



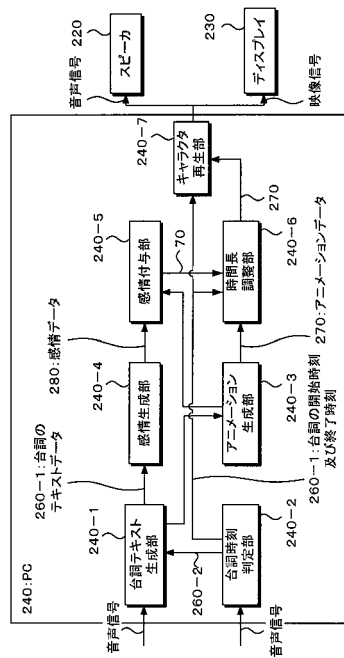
【 図 4 】



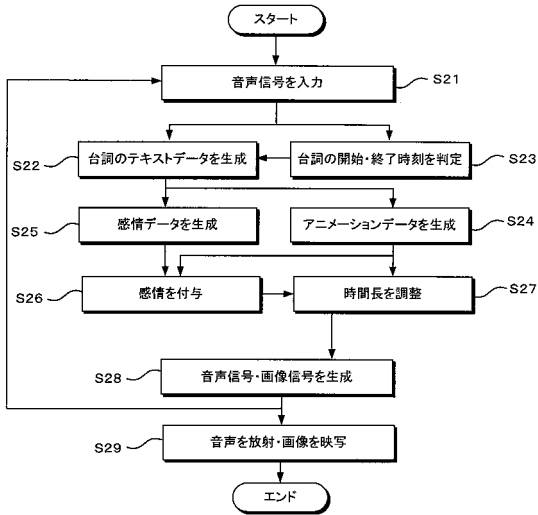
【 図 3 】



【 図 5 】



【 図 6 】



【 図 8 】

270				
台詞ID	アニメーションID	時刻	全ボーン空間座標	部区分
1	5(お辞儀)	T1	$(x_1^1 y_1^1 z_1^1) \dots (x_m^1 y_m^1 z_m^1)$	開始部
		T2	$(x_1^2 y_1^2 z_1^2) \dots (x_m^2 y_m^2 z_m^2)$	主要部
	
2	3(指差し)	T1	$(x_1^1 y_1^1 z_1^1) \dots (x_m^1 y_m^1 z_m^1)$	開始部
	
		Tn	$(x_1^n y_1^n z_1^n) \dots (x_m^n y_m^n z_m^n)$	収束部
...

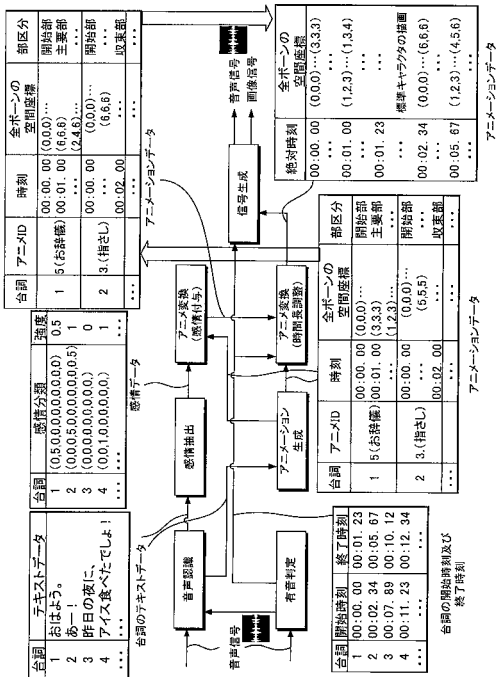
【 図 7 】

	260-1	260-2
台詞ID	テキストデータ	開始時刻 終了時刻
1	テキスト1	00:00.00 00:01.23
2	テキスト2	00:04.56 00:07.89
...
n	テキストn	01:23.45 01:26.78

【 図 9 】

280		
台詞ID	感情分類	全体強度
1	$(S_1^1 S_2^1 S_3^1 S_4^1 S_5^1 S_6^1 S_7^1 S_8^1 S_9^1)$	W^1
2	$(S_1^2 S_2^2 S_3^2 S_4^2 S_5^2 S_6^2 S_7^2 S_8^2 S_9^2)$	W^2
...
n	$(S_1^n S_2^n S_3^n S_4^n S_5^n S_6^n S_7^n S_8^n S_9^n)$	W^n

【 図 10 】



フロントページの続き

(72)発明者 酒澤 茂之

埼玉県ふじみ野市大原二丁目1番15号 株式会社K D D I 研究所内

Fターム(参考) 5B050 AA08 BA08 BA12 CA08 EA24 EA26 FA02 FA10