

(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2010-21638

(P2010-21638A)

(43) 公開日 平成22年1月28日(2010.1.28)

(51) Int.Cl.	F I	テーマコード (参考)
H04N 5/91 (2006.01)	H04N 5/91 Z	5C052
H04N 5/76 (2006.01)	H04N 5/76 B	5C053
G10L 15/00 (2006.01)	G10L 15/00 200T	5C122
H04N 5/225 (2006.01)	G10L 15/00 200G	5D015
	H04N 5/225 F	

審査請求 未請求 請求項の数 7 O L (全 9 頁)

(21) 出願番号	特願2008-178092 (P2008-178092)	(71) 出願人	502324066
(22) 出願日	平成20年7月8日 (2008.7.8)		株式会社デンソーアイティラボラトリ
			東京都渋谷区渋谷2-15-1 渋谷クロ
			スタワー28F
		(74) 代理人	100093067
			弁理士 二瓶 正敬
		(72) 発明者	佐々木 宏
			東京都渋谷区渋谷二丁目15番1号 渋谷
			クロスタワー28F株式会社デンソーアイ
			ティラボラトリ内
		(72) 発明者	岩崎 弘利
			東京都渋谷区渋谷二丁目15番1号 渋谷
			クロスタワー28F株式会社デンソーアイ
			ティラボラトリ内
		Fターム(参考)	5C052 AC08 DD02 DD04

最終頁に続く

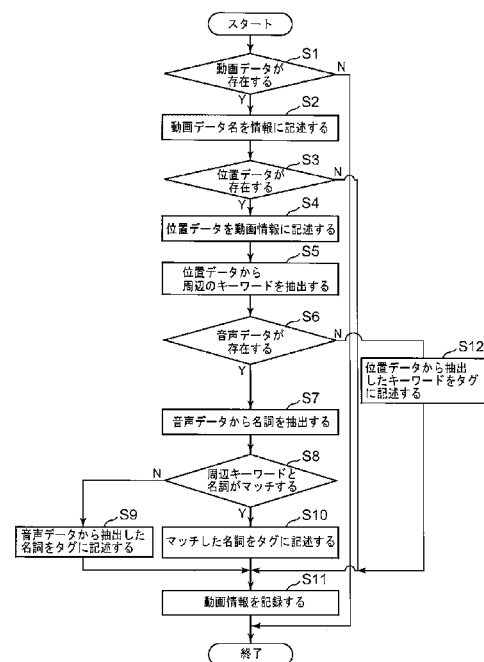
(54) 【発明の名称】 タグ情報付加装置、タグ情報付加方法及びコンピュータプログラム

(57) 【要約】

【課題】複数の画像フレームに対して、ユーザの所望の画像フレームを検索するためのタグ情報を自動的に生成して付加する。

【解決手段】動画処理装置16は、GPS部14により検出された現在位置と地図データに基づいて、現在位置近傍の地理的な名称を周辺キーワードとして抽出するとともに音声認識装置15により音声を認識し、音声認識されない場合に周辺キーワードを現在位置が検出された時点の画像フレームに対してタグ情報として付加し、音声認識された場合に、認識された音声から名詞を抽出して音声認識された時点の画像フレームに対してタグ情報として付加する。

【選択図】図2



【特許請求の範囲】**【請求項 1】**

撮像手段により撮像されて生成された複数の画像フレームに対して、ユーザの所望の画像フレームを検索するためのタグ情報を自動的に生成して付加するタグ情報付加装置であって、

前記撮像手段による撮像中に前記撮像手段の現在位置を検出する現在位置検出手段と、
前記現在位置検出手段により検出された現在位置に基づき、地理的な名称を有する地図データを用いて、現在位置近傍の地理的な名称を周辺キーワードとして抽出する周辺キーワード抽出手段と、

前記撮像手段による撮像中にユーザの音声を認識する音声認識手段と、

前記音声認識手段により音声が認識されない場合に、前記周辺キーワード抽出手段により抽出された周辺キーワードを、前記現在位置が検出された時点の画像フレームに対してタグ情報として付加する第 1 のタグ情報付加手段と、

前記音声認識手段により音声が認識された場合に、認識された音声から名詞を抽出して前記音声が認識された時点の画像フレームに対してタグ情報として付加する第 2 のタグ情報付加手段とを、

備えたタグ情報付加装置。

【請求項 2】

前記第 2 のタグ情報付加手段は、前記抽出した名詞と、前記周辺キーワード抽出手段により抽出された周辺キーワードを比較し、一致した名詞を前記音声が認識された時点の画像フレームに対してタグ情報として付加することを特徴とする請求項 1 に記載のタグ情報付加装置。

【請求項 3】

前記第 1、第 2 のタグ情報付加手段はそれぞれ、前記現在位置が検出された時点の画像フレーム、前記音声が認識された時点の画像フレームから所定数の画像フレームに対して同じタグ情報を付加することを特徴とする請求項 1 又は 2 に記載のタグ情報付加装置。

【請求項 4】

前記複数の画像フレームは、前記撮像手段として用いられるビデオカメラにより撮影された動画データであることを特徴とする請求項 1 から 3 までのいずれか 1 つに記載のタグ情報付加装置。

【請求項 5】

前記複数の画像フレームは、前記撮像手段として用いられるスチルカメラ又は携帯電話機に内蔵されたカメラにより撮影された静止画ファイルであることを特徴とする請求項 1 から 3 までのいずれか 1 つに記載のタグ情報付加装置。

【請求項 6】

撮像手段により撮像されて生成された複数の画像フレームに対して、ユーザの所望の画像フレームを検索するためのタグ情報を自動的に生成して付加するタグ情報付加方法であって、

前記撮像手段による撮像中に前記撮像手段の現在位置を検出する現在位置検出ステップと、

前記現在位置検出手段により検出された現在位置と地図データに基づいて、現在位置近傍の地理的な名称を周辺キーワードとして抽出する周辺キーワード抽出ステップと、

前記撮像手段による撮像中にユーザの音声を認識する音声認識ステップと、

前記音声認識手段により音声が認識されない場合に、前記周辺キーワード抽出手段により抽出された周辺キーワードを、前記現在位置が検出された時点の画像フレームに対してタグ情報として付加する第 1 のタグ情報付加ステップと、

前記音声認識手段により音声が認識された場合に、認識された音声から名詞を抽出して前記音声が認識された時点の画像フレームに対してタグ情報として付加する第 2 のタグ情報付加ステップとを、

備えたタグ情報付加方法。

10

20

30

40

50

【請求項 7】

撮像手段により撮像されて生成された複数の画像フレームに対して、ユーザの所望の画像フレームを検索するためのタグ情報をコンピュータが自動的に生成して付加するためのコンピュータプログラムであって、

前記撮像手段による撮像中に前記撮像手段の現在位置を検出する現在位置検出ステップと、

前記現在位置検出手段により検出された現在位置と地図データに基づいて、現在位置近傍の地理的な名称を周辺キーワードとして抽出する周辺キーワード抽出ステップと、

前記撮像手段による撮像中にユーザの音声を認識する音声認識ステップと、

前記音声認識手段により音声が認識されない場合に、前記周辺キーワード抽出手段により抽出された周辺キーワードを、前記現在位置が検出された時点の画像フレームに対してタグ情報として付加する第 1 のタグ情報付加ステップと、

前記音声認識手段により音声が認識された場合に、認識された音声から名詞を抽出して前記音声が認識された時点の画像フレームに対してタグ情報として付加する第 2 のタグ情報付加ステップとを、

備えたコンピュータプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、複数の画像フレームに対して、ユーザの所望の画像フレームを検索するためのタグ情報を付加するタグ情報付加装置、タグ情報付加方法及びコンピュータプログラムに関する。

本発明は特に、ビデオカメラにより撮影された動画像ファイルや、スチルカメラなどにより撮影された静止画ファイルを構成する複数の画像フレームに対して、ユーザの所望の画像フレームを検索するためのタグ情報を自動的に生成して付加するタグ情報付加装置、タグ情報付加方法及びコンピュータプログラムに関する。

【背景技術】

【0002】

従来技術として下記の特許文献 1 には、作業者が再生中の動画の所望のフレームなどを強調表示させるためにそのフレームにタグを付加するための入力操作を行う方法が記載されている。また、他の従来技術として下記の特許文献 2 には、作業者が画像フレーム内の所望の部分領域にキー画像を付加するための入力操作を行う方法が記載されている。

【特許文献 1】特開 2005 - 181599 号公報（図 13）

【特許文献 2】特開 2007 - 19768 号公報（図 10）

【発明の開示】

【発明が解決しようとする課題】

【0003】

しかしながら、上記従来技術では、作業者がタグを付加するための入力操作を行う必要があるため、入力操作作業に膨大な時間がかかるという問題点がある。

【0004】

本発明は上記従来技術の問題点に鑑み、動画像ファイルや静止画ファイルを構成する複数の画像フレームに対して、ユーザの所望の画像フレームを検索するためのタグ情報を自動的に生成して付加することができるタグ情報付加装置、タグ情報付加方法及びコンピュータプログラムを提供することを目的とする。

【課題を解決するための手段】

【0005】

本発明は上記目的を達成するために、撮像手段により撮像されて生成された複数の画像フレームに対して、ユーザの所望の画像フレームを検索するためのタグ情報を自動的に生成して付加するタグ情報付加装置であって、

前記撮像手段による撮像中に前記撮像手段の現在位置を検出する現在位置検出手段と、

前記現在位置検出手段により検出された現在位置に基づき、地理的な名称を有する地図データを用いて、現在位置近傍の地理的な名称を周辺キーワードとして抽出する周辺キーワード抽出手段と、

前記撮像手段による撮像中にユーザの音声を認識する音声認識手段と、

前記音声認識手段により音声が認識されない場合に、前記周辺キーワード抽出手段により抽出された周辺キーワードを、前記現在位置が検出された時点の画像フレームに対してタグ情報として付加する第1のタグ情報付加手段と、

前記音声認識手段により音声が認識された場合に、認識された音声から名詞を抽出して前記音声が認識された時点の画像フレームに対してタグ情報として付加する第2のタグ情報付加手段とを、

備えたことを特徴とする。

【0006】

また、前記第2のタグ情報付加手段は、前記抽出した名詞と、前記周辺キーワード抽出手段により抽出された周辺キーワードを比較し、一致した名詞を前記音声が認識された時点の画像フレームに対してタグ情報として付加することを特徴とする。

また、前記第1、第2のタグ情報付加手段はそれぞれ、前記現在位置が検出された時点の画像フレーム、前記音声が認識された時点の画像フレームから所定数の画像フレームに対して同じタグ情報を付加することを特徴とする。

また、前記複数の画像フレームは、前記撮像手段として用いられるビデオカメラにより撮影された動画データであることを特徴とする。

また、前記複数の画像フレームは、前記撮像手段として用いられるスチルカメラ又は携帯電話機に内蔵されたカメラにより撮影された静止画ファイルであることを特徴とする。

【0007】

本発明は上記目的を達成するために、撮像手段により撮像されて生成された複数の画像フレームに対して、ユーザの所望の画像フレームを検索するためのタグ情報を自動的に生成して付加するタグ情報付加方法であって、

前記撮像手段による撮像中に前記撮像手段の現在位置を検出する現在位置検出ステップと、

前記現在位置検出手段により検出された現在位置と地図データに基づいて、現在位置近傍の地理的な名称を周辺キーワードとして抽出する周辺キーワード抽出ステップと、

前記撮像手段による撮像中にユーザの音声を認識する音声認識ステップと、

前記音声認識手段により音声が認識されない場合に、前記周辺キーワード抽出手段により抽出された周辺キーワードを、前記現在位置が検出された時点の画像フレームに対してタグ情報として付加する第1のタグ情報付加ステップと、

前記音声認識手段により音声が認識された場合に、認識された音声から名詞を抽出して前記音声が認識された時点の画像フレームに対してタグ情報として付加する第2のタグ情報付加ステップとを、

備えたことを特徴とする。

【0008】

本発明は上記目的を達成するために、撮像手段により撮像されて生成された複数の画像フレームに対して、ユーザの所望の画像フレームを検索するためのタグ情報をコンピュータが自動的に生成して付加するためのコンピュータプログラムであって、

前記撮像手段による撮像中に前記撮像手段の現在位置を検出する現在位置検出ステップと、

前記現在位置検出手段により検出された現在位置と地図データに基づいて、現在位置近傍の地理的な名称を周辺キーワードとして抽出する周辺キーワード抽出ステップと、

前記撮像手段による撮像中にユーザの音声を認識する音声認識ステップと、

前記音声認識手段により音声が認識されない場合に、前記周辺キーワード抽出手段により抽出された周辺キーワードを、前記現在位置が検出された時点の画像フレームに対してタグ情報として付加する第1のタグ情報付加ステップと、

前記音声認識手段により音声が認識された場合に、認識された音声から名詞を抽出して前記音声認識された時点の画像フレームに対してタグ情報として付加する第2のタグ情報付加ステップとを、

備えたことを特徴とする。

【0009】

この構成により、動画像ファイルや静止画ファイルを構成する複数の画像フレームに対して、ユーザの所望の画像フレームを検索するためのタグ情報を自動的に生成して付加することができる。

【発明の効果】

【0010】

本発明によれば、撮像手段により撮像されて生成された動画像ファイルや静止画ファイルを構成する複数の画像フレームに対して、ユーザの所望の画像フレームを検索するためのタグ情報を自動的に生成して付加することができる。

【発明を実施するための最良の形態】

【0011】

以下、図面を参照して本発明の実施の形態について説明する。図1は本発明に係るタグ情報付加装置の一実施の形態を示すブロック図、図2は図1の動画処理装置の動作を説明するためのフローチャートである。

【0012】

図1において、撮像手段である車外カメラ11は不図示の車両の例えば前方などの周辺動画を撮影し、他の撮像手段である車内カメラ12は車両内の例えば運転者などの車内動画を撮影するように配置される。車内マイク13は車両内の運転者などの声を録音し、GPS(Global Positioning System)部14は車両の現在位置(緯度、経度)を取得する。音声認識装置15は車内マイク13により録音された音声を認識する。動画処理装置16は、車外カメラ11、車内カメラ12により撮影された各動画を構成する複数フレームのうちの任意のフレームに対して、車内マイク13及び音声認識装置15により得られた音声データと、GPS部14により取得された現在位置に基づいて検索用のタグなどを自動的に生成して付加し、タグ付き動画データとして動画記録装置17に記録する。ここで、動画処理装置16は地図データをあらかじめストアしており、現在位置近傍における地理的な名称(地名や、山、川、橋、ビルなどの名前)を地図データから周辺キーコードとして取得することができる。取得する周辺キーコードの数は、複数が望ましく、例えば20個程度とする。

【0013】

図2を参照して動画処理装置16の動作を説明する。ここで、動画データの動画情報の構成は、動画データ名(例えば撮影順を示す番号)と、ユーザの所望の画像フレームを検索するためのタグとなるキーワードと、同じタグが付与される開始フレーム番号及び終了フレーム番号と位置データ(緯度、経度)などを含む。ここで、検索表示時に見やすいように、同じタグが付与されるフレーム数を一義的に100フレーム(=1ブロック)とする。

【0014】

まず、車外カメラ11、車内カメラ12からの動画データが存在するか否かをチェックし(ステップS1)、存在しなければ終了する。ステップS1において動画データが存在する場合にはステップS2に進み、動画データ名を動画情報に記述する。次いでGPS部14からの位置データが存在するか否かをチェックし(ステップS3)、存在しなければステップS11に進み、動画情報を動画記録装置17に記録する。ステップS3において位置データが存在する場合にはステップS4に進み、その位置データを動画情報に記述する。次いでその位置データと地図データに基づいて現在位置近傍の1以上の地理的な名称を周辺キーワードとして抽出する(ステップS5)。

【0015】

次いで音声認識装置15からの音声データが存在するか否かをチェックする(ステップ

10

20

30

40

50

S 6)。音声データが存在しない場合にはステップ S 1 2 に進み、位置データを取得した時点の動画フレーム番号を開始フレーム番号としてその 1 0 0 フレーム後を終了フレーム番号とし、また、ステップ S 5 で位置データから抽出した 1 以上の周辺キーワードを動画情報のタグに記述し、次いでステップ S 1 1 に進んで動画情報を動画記録装置 1 7 に記録する。

【 0 0 1 6 】

ステップ S 6 において音声データが存在する場合には、ステップ S 7 に進んでその音声データから名詞を抽出し、次いでその抽出した名詞と、ステップ S 5 で位置データから抽出した 1 以上の周辺キーワードをマッチングする(ステップ S 8)。もしマッチングしない場合には、ステップ S 9 に分岐して位置データから抽出した 1 以上の周辺キーワードを動画情報のタグに記述し、他方、マッチングした場合にはステップ S 1 0 に進んでマッチングした名詞を動画情報のタグに記述する。また、ステップ S 9、S 1 0 では、音声データを取得した時点の動画フレーム番号を開始フレーム番号としてその 1 0 0 フレーム後を終了フレーム番号とし、次いでステップ S 1 1 に進んで動画情報を動画記録装置 1 7 に記録する。

10

【 0 0 1 7 】

図 3 は富士山の近傍を走行して撮影した動画データに対して付与した動画情報の例を示す。図 3 では、富士山の動画データの開始フレーム番号 = 1 から終了フレーム番号 = 1 0 0 までの 1 ブロックに対して、緯度データ (latitude) 及び経度データ (Longitude) と、キーワード = 御殿場を付与するとともに、開始フレーム番号 = 9 0 から終了フレーム番号 = 1 8 9 までの 1 ブロックに対して、緯度データ (latitude) 及び経度データ (Longitude) と、キーワード = 富士山を付与した例を示す。

20

【 0 0 1 8 】

図 4 は音声データが存在しない富士山の動画データに対して付与した動画情報の例を示す。図 4 では、開始フレーム番号 = 1 から終了フレーム番号 = 1 0 0 までの 1 ブロックに対して、緯度及び経度データと、図 3 においてステップ S 5 で位置データから抽出した周辺キーワード = 御殿場、富士山、足柄を付与した例を示す。

【 0 0 1 9 】

図 5 は動画データの途中 (フレーム番号 = 9 0) から音声データ「今日の富士山、きれいだね」が存在する富士山の動画データに対して付与した動画情報の例を示す。図 5 では、開始フレーム番号 = 1 から終了フレーム番号 = 1 0 0 までの 1 ブロックに対して、緯度及び経度データと、ステップ S 5 で位置データから抽出した周辺キーワード = 御殿場を付与するとともに、開始フレーム番号 = 9 0 から終了フレーム番号 = 1 8 9 までの 1 ブロックに対して、緯度及び経度データと、音声認識により抽出した名詞「富士山」を付与した例を示す。

30

【 0 0 2 0 】

< 本発明の適用例 >

図 6 は、本発明に係るタグ情報付加装置 1 0 が適用されたシステムを示す。図 6 に示すシステムでは、本発明に係るタグ情報付加装置 1 0 は、車両 V に搭載されて動画データベース (DB) 1 0 a として使用され、車両 V の走行中に撮影された周辺動画や車内動画を構成する複数フレームのうちの任意のフレームに対して、車両 V の走行中に録音された音声データと現在位置に基づいて検索用のタグを自動的に生成して付加し、タグ付き動画データとして記録する。タグ情報付加装置 1 0 (動画 DB 1 0 a) に記録された動画データは、サーバ 1 にアップロードしてサーバ 1 からユーザ Y の携帯電話機 2 や、PC (パーソナルコンピュータ) 3、車載情報端末 4 にダウンロードする。

40

【 0 0 2 1 】

また、本発明の撮像手段はビデオカメラに限定されず、通常の携帯型のビデオカメラ、デジタルスチルカメラ、携帯電話機内蔵のカメラにも適用することができる。また、タグ付加対象の映像は、動画に限定されず、複数枚の静止画により構成される画像ファイルにも適用することができる。

50

【産業上の利用可能性】

【0022】

本発明は、撮像手段により撮像されて生成される動画ファイルや静止画ファイルを構成する複数の画像フレームに対して、ユーザの所望の画像フレームを検索するためのタグ情報を自動的に生成して付加することができるという効果を有し、ビデオカメラ、スチルカメラ、携帯電話機などに利用することができる。

【図面の簡単な説明】

【0023】

【図1】本発明に係るタグ情報付加装置の一実施の形態を示すブロック図である。

【図2】図1の動画処理装置の動作を説明するためのフローチャートである。

10

【図3】動画情報を示す説明図である。

【図4】音声データがない場合の動画情報を示す説明図である。

【図5】音声データがある場合の動画情報を示す説明図である。

【図6】本発明に係るタグ情報付加装置を適用したシステムを示す説明図である。

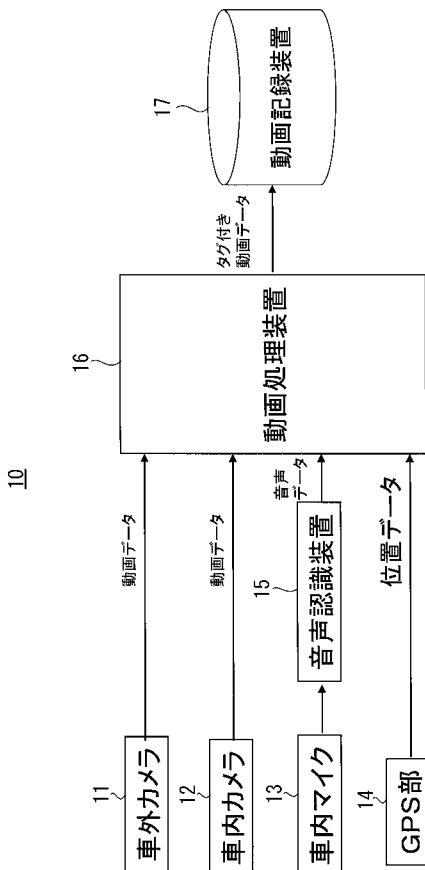
【符号の説明】

【0024】

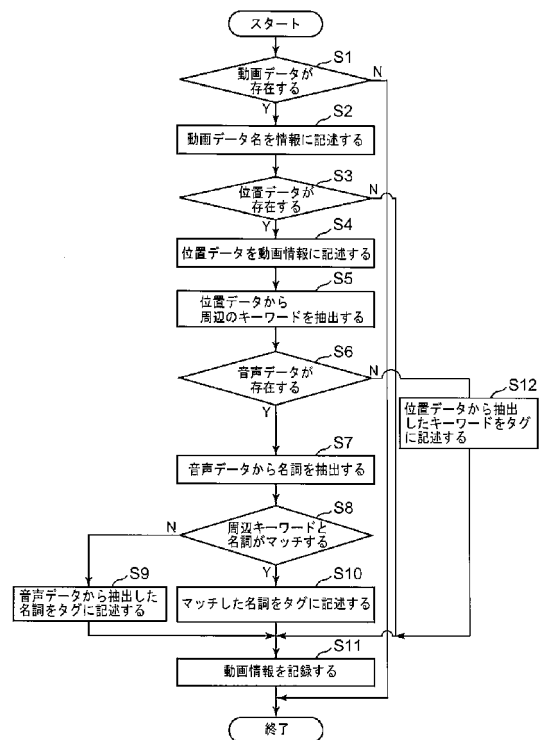
- 11 車外カメラ
- 12 車内カメラ
- 13 車内マイク
- 14 GPS部
- 15 音声認識装置
- 16 動画処理装置
- 17 動画記録装置

20

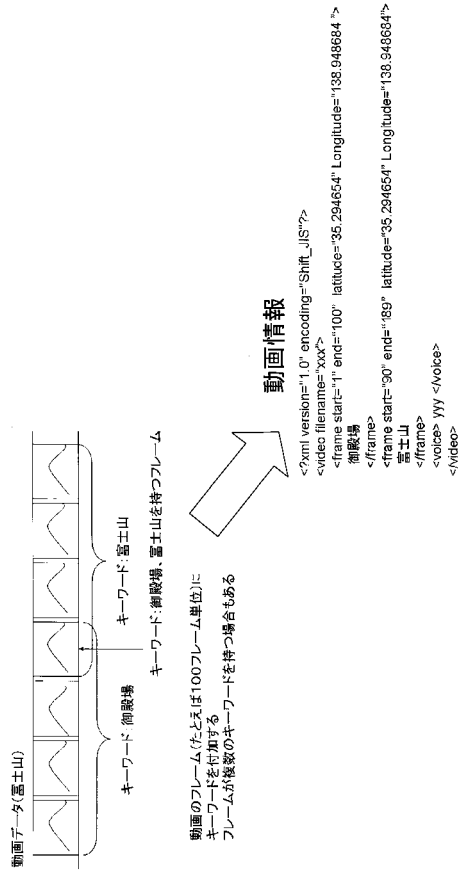
【図1】



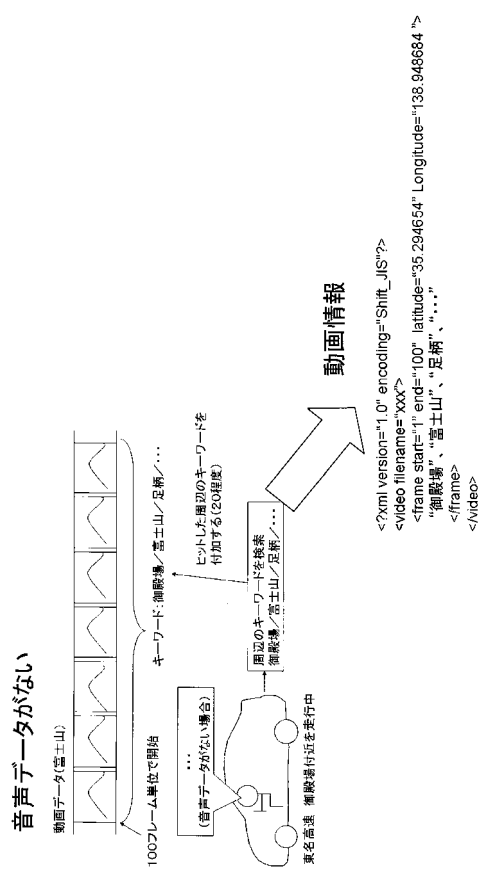
【図2】



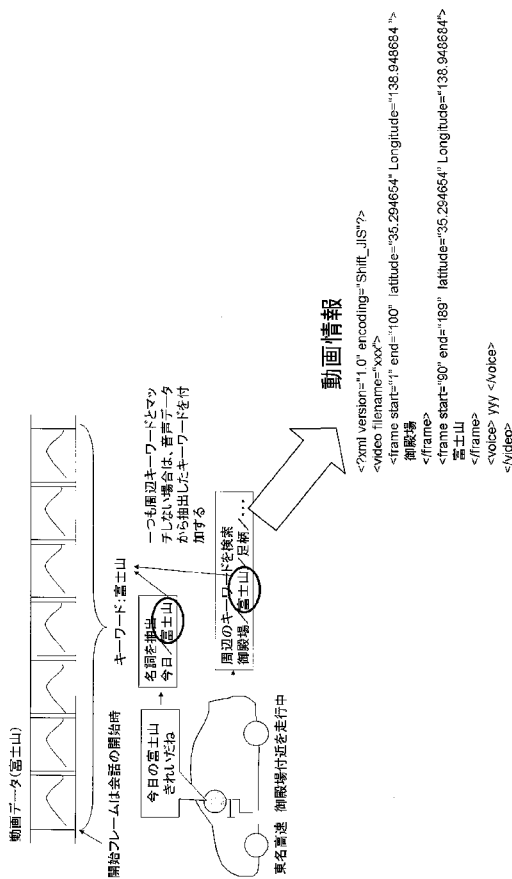
【図 3】



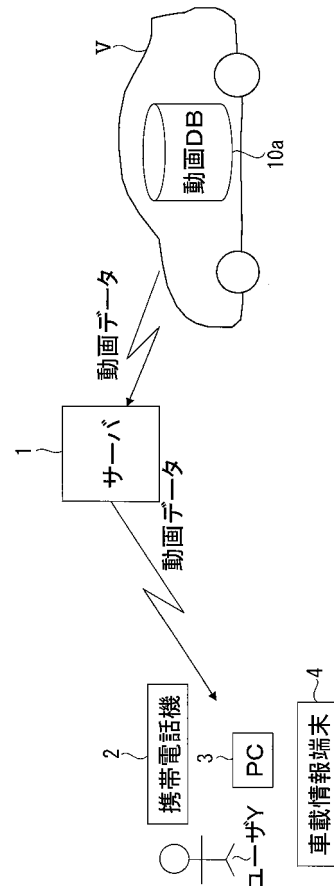
【図 4】



【図 5】



【図 6】



フロントページの続き

F ターム(参考) 5C053 FA08 GA06 GB06 JA01 JA21 LA01
5C122 DA03 DA04 DA09 EA61 EA68 GA21 GA24 GA31 GA34 GC06
HA03 HA04 HA90 HB01
5D015 KK01