

(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2011-244058

(P2011-244058A)

(43) 公開日 平成23年12月1日(2011.12.1)

(51) Int.Cl.	F I	テーマコード (参考)
<b>H04N 5/225 (2006.01)</b>	H04N 5/225 F	2C032
<b>G06T 17/05 (2011.01)</b>	H04N 5/225 A	2F129
<b>G06T 1/00 (2006.01)</b>	G06T 17/50	5B050
<b>G09B 29/10 (2006.01)</b>	G06T 1/00 280	5B057
<b>G09B 29/00 (2006.01)</b>	G09B 29/10 A	5C122
審査請求 未請求 請求項の数 13 O L (全 28 頁) 最終頁に続く		

(21) 出願番号 特願2010-111944 (P2010-111944)  
 (22) 出願日 平成22年5月14日 (2010.5.14)

(71) 出願人 000002185  
 ソニー株式会社  
 東京都港区港南1丁目7番1号  
 (74) 代理人 100095957  
 弁理士 亀谷 美明  
 (74) 代理人 100096389  
 弁理士 金本 哲男  
 (74) 代理人 100101557  
 弁理士 萩原 康司  
 (74) 代理人 100128587  
 弁理士 松本 一騎  
 (72) 発明者 五十嵐 卓也  
 東京都港区港南1丁目7番1号 ソニー株式会社社内

最終頁に続く

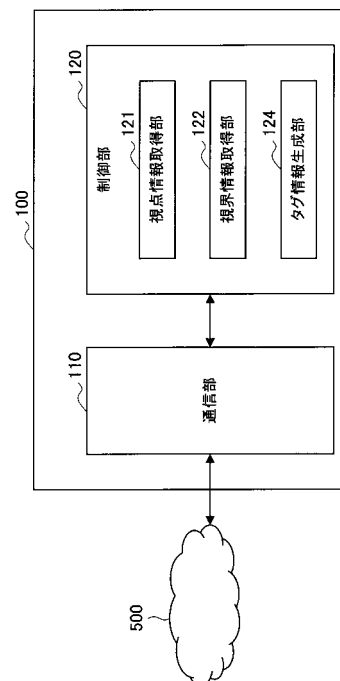
(54) 【発明の名称】 情報処理装置、情報処理システム、及びプログラム

## (57) 【要約】

【課題】撮影された映像の視界空間上の座標に対して関連づけられたタグを用いることのできる情報処理システムを提供する。

【解決手段】情報処理装置100は、コンテンツを撮影する撮像装置の位置である視点位置の情報を取得する視点情報取得部121と、上記視点位置から上記コンテンツを撮影したときの視界空間を示す視界情報を取得する視界情報取得部122と、上記視点情報、上記視界情報、及び奥行き情報により特定される視界空間上の座標位置と関連づけられたタグの情報を生成するタグ情報生成部124とを有する。

【選択図】図4



**【特許請求の範囲】****【請求項 1】**

コンテンツを撮影する撮像装置の位置である視点位置の情報を取得する視点情報取得部と、

前記視点位置から前記コンテンツを撮影したときの視界空間を示す視界情報を取得する視界情報取得部と、

前記視点情報、前記視界情報、及び奥行き情報により特定される視界空間上の座標位置と関連づけられたタグの情報を生成するタグ情報生成部と、  
を備える、情報処理装置。

**【請求項 2】**

前記視界情報は、前記視点位置からの撮影方向を示す方位と、視野角の情報とを含む、請求項 1 に記載の情報処理装置。

**【請求項 3】**

前記奥行き情報は、前記コンテンツにおいて撮影された被写体までの距離であり、

前記タグ情報生成部は、前記コンテンツの撮影情報に基づいた被写体距離又は過焦点距離を前記奥行き情報として用いる、請求項 2 に記載の情報処理装置。

**【請求項 4】**

前記タグ情報生成部は、前記視点情報及び前記視界情報を含むキーと、前記タグの文字列の値の情報及び前記視界空間上における前記タグの座標位置情報を含むバリューと、により示されるタグ情報を生成する、請求項 1 ～ 3 のいずれか 1 項に記載の情報処理装置。

**【請求項 5】**

前記タグ情報生成部は、前記コンテンツの撮影画角に応じたビット数で示される前記視界情報を含む前記タグ情報を生成する、請求項 4 に記載の情報処理装置。

**【請求項 6】**

前記タグ情報生成部は、前記タグを付与する対象物との距離に応じたビット数により示される前記視点情報を含む前記タグ情報を生成する、請求項 4 に記載の情報処理装置。

**【請求項 7】**

前記タグ情報生成部は、前記コンテンツの撮影画角に応じたビット数で示される前記視点情報を含む前記タグ情報を生成する、請求項 6 に記載の情報処理装置。

**【請求項 8】**

前記タグ情報生成部は、少なくとも前記視点情報及び前記視界情報を各情報の境界の情報を保持する形式により文字列に変換し、変換して得られた文字列の値を前記キーの値として用いる、請求項 7 に記載の情報処理装置。

**【請求項 9】**

前記タグ情報生成部は、前記各情報を所定の区切りビット列長で区切り、区切られることにより生成される各ビット列を、前記ビット列長内のビット列が取り得る全ての値を一意に特定する変換ルールに基づいて変換する、請求項 8 に記載の情報処理装置。

**【請求項 10】**

前記視点情報は、前記視点の位置情報の誤差精度の情報を含む、請求項 1 に記載の情報処理装置。

**【請求項 11】**

コンテンツデータに付帯する撮影情報から、前記コンテンツを撮影する撮像装置の位置である視点位置の情報を含む視点情報、及び、前記視点位置から前記コンテンツを撮影したときの視界空間を示す視界情報を取得する解析部と、

前記視点情報、前記視界情報、及び奥行き情報により特定される視界空間上の座標位置と関連づけられたタグ情報から、コンテンツに重畳させるタグの情報を検索する検索部と、  
を備える、情報処理装置。

**【請求項 12】**

コンテンツを撮影する撮像装置の位置である視点位置の情報を取得する視点情報取得部

10

20

30

40

50

と、

前記視点位置から前記コンテンツを撮影したときの視界空間を示す視界情報を取得する視界情報取得部と、

前記視点情報、前記視界情報、及び奥行き情報により特定される視界空間上の座標位置と関連づけられたタグの情報を生成するタグ情報生成部と、

を有するタグ情報生成装置と、

コンテンツデータに付帯する撮影情報から、前記視点情報及び前記視界情報を取得する解析部と、

前記視点情報および前記視界情報により特定される視界空間上の座標位置と関連づけられたタグ情報から、コンテンツに重畳させるタグの情報を検索する検索部と、

前記検索部により検索されたタグが重畳された映像を表示させる表示制御部と、

を有する再生装置と、

を備える、情報処理システム。

【請求項 13】

コンピュータを、

コンテンツを撮影する撮像装置の位置である視点位置の情報を取得する視点情報取得部と、

前記視点位置から前記コンテンツを撮影したときの視界空間を示す視界情報を取得する視界情報取得部と、

前記視点情報、前記視界情報、及び奥行き情報により特定される視界空間上の座標位置と関連づけられたタグの情報を生成するタグ情報生成部と、

を備える、情報処理装置として機能させるための、プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、情報処理装置、情報処理システム、及びプログラム

【背景技術】

【0002】

カメラで撮影している映像に、その撮影位置に関する情報を文字等にて映像に重畳して表示することで、現実社会の活動を補強する技術は、AR (Augmented Reality: 拡張現実) として 1990 年代初頭から急速に研究が進み、ヘッドマウントディスプレイを利用した拡張現実、既に一部軍事用の実用が始まっていた。

【0003】

また、昨今、小型電子機器の高性能化などに伴い、カメラ及び GPS (Global Positioning System) 機能を有する端末装置により拡張現実が実現されている。このような拡張現実の例においては、利用者は、実映像に文字を付加し、付加した文字情報をネットワーク経由で他の利用者に拡張現実技術を用いて提示することができる。このように、拡張現実、情報提供という実用的な目的のみならず、コミュニケーションのツール的一种として楽しむといった使い方がされている。拡張現実に関する技術としては、例えば特許文献 1 の景観ラベリングシステムが挙げられる。

【0004】

特許文献 1 に記載の景観ラベリングシステムは、地上の 3 次元空間座標上で示された構造物等の位置に関連付けられたタグを映像に重畳する。このため、タグに示される情報が、タグが関連付けられる構造物の名称等である場合には好適である。

【先行技術文献】

【特許文献】

【0005】

【特許文献 1】特許第 3225882 号

10

20

30

40

50

**【発明の概要】****【発明が解決しようとする課題】****【0006】**

ところが、利用者がタグとして共有したい情報の中には、所定の撮影条件により、映像を撮影した位置、及び映像を撮影した方位から見た映像中においてしか意味を成さない情報がある。このような情報を共有する方法は、従来考えられておらず、潜在的に求められていた。

**【0007】**

そこで、本発明は、上記問題に鑑みてなされたものであり、本発明の目的とするところは、撮影された映像の視界空間上の座標に対して関連づけられたタグを用いることが可能な、新規かつ改良された情報処理装置、情報処理システム、及びプログラムを提供することにある。

10

**【課題を解決するための手段】****【0008】**

上記課題を解決するために、本発明のある観点によれば、コンテンツを撮影する撮像装置の位置である視点位置の情報を取得する視点情報取得部と、上記視点位置から上記コンテンツを撮影したときの視界空間を示す視界情報を取得する視界情報取得部と、上記視点情報、上記視界情報、及び奥行き情報により特定される視界空間上の座標位置と関連づけられたタグの情報を生成するタグ情報生成部とを有する、情報処理装置が提供される。

**【0009】**

20

また、上記視界情報は、上記視点位置からの撮影方向を示す方位と、視野角の情報を含んでもよい。

**【0010】**

また、上記奥行き情報は、上記コンテンツにおいて撮影される被写体までの距離であり、上記タグ情報生成部は、上記コンテンツの撮影情報に基づいた被写体距離又は過焦点距離を上記奥行き情報として用いてもよい。

**【0011】**

また、上記タグ情報生成部は、上記視点情報及び上記視界情報を含むキーと、上記タグの文字列の値の情報及び上記視界空間上における上記タグの座標位置情報を含むバリューと、により示されるタグ情報を生成してもよい。

30

**【0012】**

また、上記タグ情報生成部は、上記コンテンツの撮影画角に応じたビット数で示される上記視界情報を含む上記タグ情報を生成してもよい。

**【0013】**

また、上記タグ情報生成部は、上記タグを付与する対象物との距離に応じたビット数により示される上記視点情報を含む上記タグ情報を生成してもよい。

**【0014】**

また、上記タグ情報生成部は、上記コンテンツの撮影画角に応じたビット数で示される上記視点情報を含む上記タグ情報を生成してもよい。

**【0015】**

40

また、上記タグ情報生成部は、少なくとも上記視点情報及び上記視界情報を各情報の境界の情報を保持する形式により文字列に変換し、変換して得られた文字列の値を上記キーの値として用いてもよい。

**【0016】**

また、上記タグ情報生成部は、上記各情報を所定の区切りビット列長で区切り、区切られることにより生成される各ビット列を、上記ビット列長内のビット列が取り得る全ての値を一意に特定することのできる変換ルールに基づいて変換してもよい。

**【0017】**

また、上記視点情報は、上記視点の位置情報の誤差精度の情報を含んでもよい。

**【0018】**

50

また、上記課題を解決するために、本発明の別の観点によれば、コンテンツデータに付帯する撮影情報から、上記コンテンツを撮影する撮像装置の位置である視点位置の情報を含む視点情報、及び、上記視点位置から上記コンテンツを撮影したときの視界空間を示す視界情報を取得する解析部と、上記視点情報、上記視界情報、及び奥行き情報により特定される視界空間上の座標位置と関連づけられたタグ情報から、コンテンツに重畳させるタグの情報を検索する検索部とを有する情報処理装置が提供される。

【 0 0 1 9 】

また、上記課題を解決するために、本発明の別の観点によれば、コンテンツを撮影する撮像装置の位置である視点位置の情報を取得する視点情報取得部と、上記視点位置から上記コンテンツを撮影したときの視界空間を示す視界情報を取得する視界情報取得部と、上記視点情報、上記視界情報、及び奥行き情報により特定される視界空間上の座標位置と関連づけられたタグの情報を生成するタグ情報生成部とを有するタグ情報生成装置と、コンテンツデータに付帯する撮影情報から、上記視点情報及び上記視界情報を取得する解析部と、上記視点情報および上記視界情報により特定される視界空間上の座標位置と関連づけられたタグ情報から、コンテンツに重畳させるタグの情報を検索する検索部と、上記検索部により検索されたタグが重畳された映像を表示させる表示制御部と、を有する再生装置とを有する、情報処理システムが提供される。

【 0 0 2 0 】

また、上記課題を解決するために、本発明の別の観点によれば、コンピュータを、コンテンツを撮影する撮像装置の位置である視点位置の情報を取得する視点情報取得部と、上記視点位置から上記コンテンツを撮影したときの視界空間を示す視界情報を取得する視界情報取得部と、上記視点情報、上記視界情報、及び奥行き情報により特定される視界空間上の座標位置と関連づけられたタグの情報を生成するタグ情報生成部とを有する情報処理装置として機能させるための、プログラムが提供される。

【 発明の効果 】

【 0 0 2 1 】

以上説明したように本発明によれば、撮影された映像の視界空間上の座標に対して関連づけられたタグを用いることが可能な、新規かつ改良された情報処理装置、情報処理システム、及びプログラムを提供することができる。

【 図面の簡単な説明 】

【 0 0 2 2 】

【 図 1 】 本発明の一実施形態に係る情報処理システムの概略構成を示す構成図である。

【 図 2 】 撮影情報の一例を示す表である。

【 図 3 】 情報処理システムの効果の一例を示すための説明図である。

【 図 4 】 タグ情報生成装置の機能構成を示すブロック図である。

【 図 5 】 リレーショナルデータベースを用いる場合のタグ情報の空間モデルを示す説明図である。

【 図 6 】 リレーショナルデータベースによるタグ情報のテーブル定義例である。

【 図 7 】 データセンターの構成図である。

【 図 8 】 キーバリュ型データストアの概要の説明図である。

【 図 9 】 キーバリュ型データストアを用いる場合のタグ情報の空間モデルを示す説明図である。

【 図 1 0 】 タグ情報のキーの構成の一例を示す説明図である。

【 図 1 1 】 撮影位置のグリッドの一例を示す説明図である。

【 図 1 2 】 視界情報ビットについての説明図である。

【 図 1 3 】 視界情報のビット割当ての一例を示す表である。

【 図 1 4 】 視点情報のビット割当ての一例を示す表である。

【 図 1 5 】 可変ビット列文字符号変換表の一例である。

【 図 1 6 】 可変ビット列の文字列変換の一例を示す説明図である。

【 図 1 7 】 タグ情報の J a v a (登録商標)を用いたクラス定義の一例を示す説明図であ

10

20

30

40

50

る。

【図 18】タグ情報の検索方法の一例を示す説明図である。(北緯 35 度)

【図 19】タグ情報の検索方法の一例を示す説明図である。(北緯 60 度)

【図 20】再生装置の機能構成を示すブロック図である。

【図 21A】タグ情報検索動作を示すフローチャートである。

【図 21B】タグ情報検索動作を示すフローチャートである。

【図 22】映像へのタグ情報の投影についての説明図である。

【図 23】情報処理システムを適用したサービスの一例である。

【図 24】情報処理システムを適用したサービスの他の一例である。

【発明を実施するための形態】

10

【0023】

以下に添付図面を参照しながら、本発明の好適な実施の形態について詳細に説明する。  
なお、本明細書及び図面において、実質的に同一の機能構成を有する構成要素については、同一の符号を付することにより重複説明を省略する。

【0024】

なお、説明は以下の順序で行うものとする。

1. 概要
2. 構成
3. リレーショナルデータベースによるタグ情報管理方法
4. キーバリュ型データストアによるタグ情報管理方法
5. 効果の例
6. 適用例

20

【0025】

< 1. 概要 >

まず、本発明の一実施形態に係る情報処理システム 10 の概要について図 1 ~ 図 3 を用いて説明する。図 1 は、本発明の一実施形態に係る情報処理システム 10 の概要構成を示す構成図である。また、図 2 は、撮像装置において記録される撮影付帯情報の一例を示す表である。図 3 は、情報処理システム 10 の効果の一例を示すための説明図である。

【0026】

本実施形態に係る情報処理システム 10 は、映像コンテンツを撮影する撮像装置 200 と、ユーザの操作に応じて、撮像装置 200 から取り込んだコンテンツ 3000 に対してタグを付与する操作等を受け付ける情報処理装置である端末装置 300 と、付与されたタグの情報に、コンテンツが撮影されたときの視界空間を特定するための視界情報を付加したタグ情報を生成する情報処理装置であるタグ情報生成サーバ 100 を有するデータセンタ 1000 と、データセンタ 1000 により管理されるタグ情報を利用して映像にタグを重畳して再生する再生装置 400 と、を有する。データセンタ 1000 内の各装置と、端末装置 300、及び再生装置 400 とは、それぞれネットワーク 500 を介して接続される。

30

【0027】

データセンタ 1000 には、動画共有サービスを提供するためのサーバ群が設置されている。端末装置 300 は、データセンタ 1000 内の図示しない映像配信サーバに対して映像コンテンツをアップロードすることができる。また、再生装置 400 は、データセンタ 1000 内の映像配信サーバから配信される映像コンテンツをダウンロードして再生することができる。

40

【0028】

ユーザは、撮像装置 200 を端末装置 300 に USB (Universal Serial Bus) 等を用いて接続し、端末装置 300 のアプリケーションソフトウェアにより映像 3000 のコンテンツデータを撮像装置 200 から取込む。このとき、端末装置 300 は、映像 3000 のコンテンツデータと共に、映像 3000 に付帯する撮影情報を取込む。そして、ユーザは、端末装置 300 により映像 3000 及び映像 3000 に付帯す

50

る撮影情報をデータセンタ１０００内の映像配信サーバにアップロードする。また、ユーザは、端末３００により、映像３０００に対してタグ６００を付与することもできる。そして、ユーザは、端末装置３００により映像３０００に付与したタグ６００の情報、例えば、映像上のタグ６００の位置及びタグ６００の文字列情報などを、データセンタ１０００内のタグ情報生成サーバ１００にアップロードする。

#### 【００２９】

ここで、撮像装置２００は、例えばＧＰＳ及び電子コンパスの機能を有し、撮影した映像に付帯する撮影情報として、撮影日時及び焦点距離などの撮影条件のみならず、位置情報及び方位情報を記録することができる。ＡＶＣＨＤ規格に基づくビデオカメラにおいては、ＭＰＥＧ２－ＴＳのエレメンタリストリーム（ＥＳ）の中のユーザデータとして記録されたモディファイドデジタルビデオパックに、例えば図２に示す撮影情報が記録されている。この撮影情報は、所定時間間隔で定期的に記録される。

#### 【００３０】

端末装置３００は、ここでは１台のみ示されているが、データセンタ１０００は、多数の端末装置３００からのアクセスを受付けることができる。また、再生装置４００に関しても同様に、多数の再生装置４００がデータセンタ１０００にアクセスすることが想定される。

#### 【００３１】

そして、再生装置４００は、映像３０００と同じ視点及び同じ視界で撮影された他の映像４０００にタグ６００が重畳された映像を再生する。図１の例においては、「Mt. Fuji over Lake. Motosu」という文字列情報を有するタグ６００が、映像４０００に重畳されて再生される。

#### 【００３２】

このようなタグ６００が示す情報は、映像３０００が撮影された視点から同じ視界空間を撮影した映像上で意味をなす情報である。つまり、本栖湖越しに富士山が見える映像上でなければ、たとえ富士山が被写体として写っている映像であっても、かかるタグ６００を表示することは望ましくない。例えば、図３に示す映像４００５は、御殿場から撮影された富士山の映像である。この場合、映像４００５には、本栖湖は映っていない。このため、タグ６００を表示させることは望ましくない。また、映像４０１０は、映像３０００の撮影位置よりも富士山頂に近い位置から撮影された映像である。この場合には、映像４０１０と映像３０００とが同じ方角から撮影した映像であっても、映像４０１０にタグ６００を表示させることは望ましくない。また、映像４０１５は、映像３０００と同じ撮影位置から撮影された映像である。ところが、映像４０１５は、映像３０００とは撮影条件が異なり、映像３０００よりも近景を撮影した映像である。この場合にも、本栖湖は映像４０１５中に映っていないため、タグ６００を表示させることは望ましくない。

#### 【００３３】

従来システムは、タグを被写体の地上の３次元空間座標にラベリングしていた。この場合、富士山の地上の３次元空間座標位置は同一であるため、図３に示される映像４００５、映像４０１０、映像４０１５のような映像に対しても、タグ６００が重畳されていた。しかし、上述の通り、タグ６００は、同じ視点からの同じ視界空間を撮影した映像上で意味をなす情報である。本実施形態に係る情報処理システム１０は、このタグを撮影映像の視界空間上の座標位置に関連付けて管理する。このため、視点及び方位の異なる映像４００５、視点の異なる映像４０１０、並びに、視点及び方位は同じであるが撮影条件が異なるため視界空間の異なる映像４０１５に対してはタグ６００が重畳されることはない。

#### 【００３４】

このため、単純に対象物の名称だけでなく、その視点からのある視界空間から見える複数の対象物の位置関係において意味をなす情報をタグとして共有することができる。

#### 【００３５】

なお、ここでは、撮像装置２００としてビデオカメラを示したが、本発明はこれに限られない。撮像装置２００は、例えば携帯電話、携帯用音楽再生装置、携帯用映像処理装置

10

20

30

40

50

、携帯用ゲーム機器など、撮像機能を有する装置であればあらゆる機器であってよい。また、端末装置 300 は、PC (Personal Computer) である場合が例示されているが、本発明はかかる例に限られない。例えば、端末装置 300 は、家庭用映像処理装置 (DVD レコーダ、ビデオデッキなど)、PDA (Personal Digital Assistants)、家庭用ゲーム機器、家電機器、携帯電話、携帯用音楽再生装置、携帯用映像処理装置、携帯用ゲーム機器など、通信機能及びユーザが操作入力する入力部を有する情報処理装置であってもよい。或いは、撮像装置 200 が通信機能を有し、データセンタ 1000 内の各装置と通信することができる場合には、撮像装置 200 自体が端末装置 300 の機能を併せ持ってもよい。

#### 【0036】

ここでいうコンテンツは、映像コンテンツであり、静止画像コンテンツであってもよく、動画像コンテンツであってもよい。なお、再生装置 400 は、データセンタ 1000 により提供される映像共有サービスにより配信される映像にタグを重畳して再生させることとしたが、本発明はかかる例に限られない。例えば、再生装置 400 は、ローカルな記憶部に保有する映像コンテンツに対してタグを重畳して再生させてもよい。或いは、再生装置 400 は、撮像機能も併せ持つ装置であって、リアルタイムに撮影中の映像に対してタグを重畳させてもよい。

#### 【0037】

以上説明したように、本実施形態に係る情報処理システム 10 は、タグをコンテンツが撮影された視点からの視界空間上の座標位置と関連付けて管理する。このため、再生装置 400 においては、同じ視点からの同じ視界空間上に関連付けられたタグしか重畳されない。このため、このタグ情報を用いる再生装置 400 は、所定の撮影条件により、映像を撮影した撮影位置、及び映像を撮影した方位から見た映像中においてしか意味を成さない情報を適切に表示することができる。かかる効果を奏するための、タグ情報生成装置 100 の機能構成について、次に説明する。

#### 【0038】

##### < 2. 構成 >

図 4 は、タグ情報生成装置 100 の機能構成を示すブロック図である。タグ情報生成装置 100 は、通信部 110 と制御部 120 とを主に有する情報処理装置である。タグ情報生成装置 100 は、映像の解析結果よりタグ情報を生成する。

#### 【0039】

ここで、タグ情報生成装置 100 において生成されるタグ情報は、撮影位置の情報を含む撮影情報、撮影位置からの視界に関する視界情報、撮影位置の情報と視界情報とに基づいて特定される視界空間における座標位置、及び、タグの情報の内容を示す文字列情報、を含む情報である。

#### 【0040】

このようなタグ情報を生成するために、タグ情報生成装置 100 の制御部 120 は、視点情報取得部 121、視界情報取得部 122、及びタグ情報生成部 124 の機能を主に有する。

#### 【0041】

視点情報取得部 121 は、コンテンツを撮影する撮像装置の位置 (以下、視点位置という。)を示す視点情報を取得する。視点情報取得部 121 は、例えば、映像のコンテンツデータに付帯した撮影条件を示す情報である撮影情報から視点情報を取得する。撮影情報は、上述の通り、撮像装置 200 の機能により記録される情報である。視点情報取得部 121 は、例えば、端末装置 300 からアップロードされた映像を記憶する映像配信サーバに保存された映像のコンテンツデータに付帯する撮影情報から視点情報を抽出して取得する。或いは、視点情報取得部 121 は、映像をアップロードする端末装置 300 から直接、映像のコンテンツデータを取得して視点情報を取得してもよい。

#### 【0042】

ここで視点情報は、例えば、世界測地系で表され、GPS などの位置取得装置から取得

10

20

30

40

50



した経度、緯度、高度の情報である。また、視点情報は、位置情報の誤差精度の情報を含んでいてもよい。

【 0 0 4 3 】

視界情報取得部 1 2 2 は、視点位置からコンテンツを撮影したときの視界空間を示す視界情報を取得する。視界情報取得部 1 2 2 は、例えば、映像のコンテンツデータに付帯した撮影情報から視界情報を取得する。

【 0 0 4 4 】

ここで視界情報は、例えば、視点位置からの方位、視野角の情報などを含む。例えば視点位置からの方位は、撮像装置の有する電子コンパスにより取得した方位情報が用いられる。視野角の情報は、撮像装置の設定に応じた値となる。方位及び視野角の情報は、映像のコンテンツデータに付帯した撮影情報から取得される。

10

【 0 0 4 5 】

タグ情報生成部 1 2 4 は、視点情報、視界情報、及び視界空間の奥行き情報により特定される視界空間上の座標位置と関連づけられたタグの情報を生成する。タグ情報生成部 1 2 4 は、例えば、通信部 1 1 0 を介して端末装置 3 0 0 から取得したタグの文字列情報、及び、映像中のタグ位置の情報と、視点情報及び視界情報と、視点からタグを付与する対象物までの距離の情報とを含むタグ情報を生成する。

【 0 0 4 6 】

ここで、視界空間の奥行き情報は、コンテンツにおいて撮影された被写体までの距離である。被写体距離は、例えばコンテンツに付帯する撮影情報から被写体距離の値が取得できる場合には、取得される被写体距離の値が用いられる。また、コンテンツの撮影情報に基づいて算出される過焦点距離が用いられてもよい。

20

【 0 0 4 7 】

以上説明したタグ情報生成装置 1 0 0 の各部の詳細な動作については、タグ情報の管理方法によって異なる。ここで、本実施形態においては、タグ情報の管理方法において、現在、最も一般的に用いられているリレーショナルデータベースによる管理方法と、より大規模なデータの管理や分散処理に適したキーバリュ型 of データストアによる管理方法の 2 つの管理方法について、それぞれ次に説明する。

【 0 0 4 8 】

< 3 . リレーショナルデータベースによるタグ情報管理方法 >

30

まずは、リレーショナルデータベースによるタグ情報管理方法について、図 5 及び図 6 を用いて説明する。図 5 は、タグ情報をリレーショナルデータベースにより管理する場合におけるタグの空間モデルの一例を示す説明図である。図 6 は、図 5 の空間モデルにより定義されたタグ情報を管理するリレーショナルデータベースのテーブル定義の一例を示す説明図である。

【 0 0 4 9 】

上述の通り、本実施形態においてタグは、タグを付加する映像を撮影した撮像装置の位置である視点を基準として映像の視界の 3 次元空間上にラベリングされる。視点は、世界測地系で表された経度、緯度、高度の情報であり、この場合緯度 N 1 , 経度 E 1 , 高度 A 1 である。視界の 3 次元空間上の位置は、撮影方向ベクターを軸に、撮影映像の midpoint を合わせて直交配置した 2 次元の視野平面を水平画角、及び視点からの距離により規定し、この視野平面上の位置で表される。

40

【 0 0 5 0 】

この視野平面の座標は、水平 X 軸、垂直 Y 軸で規定され、midpoint を ( 0 , 0 ) とし、画像の水平方向の解像度 W は、画像の横解像度となる。タグの視野平面上の位置は、( X , Y ) にて表される。

【 0 0 5 1 】

タグを他の映像と共有し、重畳して用いるためには、映像の視点位置、視野角、方位、視点からタグを付加する対象物までの距離が近いことを条件として、データベースから検索することにより実現することができる。位置情報、及び方位情報に加えて、視野角及び

50

視点からの距離によりタグ情報を検索できるようにすることは、視聴者の注目している被写体に対するタグ情報であるか否かを判断する上で重要である。

#### 【 0 0 5 2 】

次に、このような空間モデルにより定義されたタグ情報を管理するリレーショナルデータベースのテーブル定義の一例を図 6 を参照しながら説明する。図 6 の表の各行は、情報タグの 1 レコードを示す。そして、表の列は、視点情報、視界情報、情報タグの 3 つに分類される。

#### 【 0 0 5 3 】

視点情報は、経度、緯度、高度、及び誤差の列を有する。緯度は、浮動小数点の - 9 0 度から 9 0 度の範囲の値をとり、北緯は正の値で表され、南緯は負の値で表される。また、経度は、浮動小数点の - 1 8 0 度から 1 8 0 度の値をとり、東経は正の値で表され、西経は負の値で表される。高度は、整数で表されるメータ値である。そして誤差は、位置情報の精度をメータ値で表した値である。誤差は、GPS の場合には、DOP (Dilution of Precision) の値から算出する。誤差が不明な場合には、- 1 の値を入れる。

#### 【 0 0 5 4 】

視界情報は、方位、視野角、高さ、及び幅の列を有する。方位は、撮影方向の方角を示す。この方位の値は、浮動小数点により 0 - 3 6 0 度の値をとり、北は 0 度 (3 6 0 度) として、東は 9 0 度、南は 1 8 0 度、西は 2 7 0 度で表される。また、視野角は、浮動小数の度数で表される。高さ及び幅は、撮影映像の縦横の解像度であり、ピクセル値で表される。例えば、フルハイビジョンの映像であれば、解像度は縦 1 9 2 0、横 1 0 8 0 である。映像の解像度は、コンテンツデータに付帯する撮影情報から取得することができる。視野角を 2 とした場合、以下の関係から求まる。

#### 【 0 0 5 5 】

$$\tan = (\text{撮影素子体格サイズ}) / (\text{焦点距離})$$

#### 【 0 0 5 6 】

情報タグは、距離、X 座標値、Y 座標値、タグ名の列を有する。距離は、視野空間上の視点からの距離であり、整数によりメータ単位で表される。X 座標値および Y 座標値は、図 5 において説明した視野平面状の位置であり、ピクセル値で表される。タグ名は、タグの名前が文字列により表される。タグの名前は、タグによって共有したい情報である。

#### 【 0 0 5 7 】

なお、タグの距離は、タグ毎に計測することはできない。このため、簡易的に、撮影画像の被写体距離がタグの距離として用いられる。被写体距離は、動画の場合には図 2 に示す撮影付帯情報、また、静止画の場合には、EXIF ファイルに記録された画像情報の被写体距離タグの値が用いられてもよい。被写体距離がこれらの情報から取得できない場合には、以下の式を用いて算出した過焦点距離をタグの距離として用いることができる。

#### 【 0 0 5 8 】

$$\text{許容錯乱円径} = \text{画像素子対角サイズ} \div \text{許容錯乱円定数} (= 1300)$$

$$\text{過焦点距離} = (\text{焦点距離} \times \text{焦点距離}) \div (\text{許容錯乱円径} \times F \text{ 値}) \quad [\text{単位: mm}]$$

#### 【 0 0 5 9 】

EXIF 情報に 35 mm 換算の焦点距離 Focal Length In 35 mm タグが記録された場合は、35 mm Film において撮像素子の対角サイズは 43.266615 mm であることは既知である。このため、Aperture Value タグに記録された絞り F 値を用いて過焦点距離が算出される。

#### 【 0 0 6 0 】

或いは、タグの距離は、撮影映像を解析して各被写体のフォーカス距離を測定技術を用いて測定された距離が用いられてもよい。また、タグの距離は、例えば 3D カメラのように視差情報を記録できる装置が用いられた場合には、この情報を用いて解析により求められた値が用いられてもよい。

#### 【 0 0 6 1 】

以上のように、図 5 の空間モデルに従って、図 6 のテーブル定義のリレーショナルデータベースを利用することで画像に付加したタグを管理することができる。ところが、このようなリレーショナルデータベースは、大規模なデータの管理、および分散処理が行いづらいという特徴がある。そこで、管理する対象となる情報量が多くなった場合に好適なクラウドコンピューティングの技術を適用してタグ情報を管理する方法について、次に提案する。

#### 【 0 0 6 2 】

< 4 . キーバリュ型データストアによるタグ情報管理方法 >

図 7 は、本実施形態に係るコンテンツ共有サービスを提供するためのデータセンタ内の構成図である。映像コンテンツの共有サービスを提供するためには、アップロードされた多量の映像コンテンツデータの画像変換処理、記録管理、及び大勢の利用者からの映像コンテンツデータへのアクセスを処理する必要がある。このため、大規模な分散処理型のクラウドコンピューティング技術を利用したタグ情報管理方法を提案する。

10

#### 【 0 0 6 3 】

データセンタ 1 0 0 0 は、複数の Web サーバ 1 7 0、複数の映像配信サーバ 1 8 0、複数のタグ情報管理サーバ 1 9 0、及び複数のタグ情報生成サーバ 1 0 0 を有する。

#### 【 0 0 6 4 】

Web サーバ 1 7 0 は、端末装置 3 0 0 のブラウザに表示される HTML ドキュメント及び VR 情報を提供する。映像配信サーバ 1 8 0 は、映像を再生装置に配信する。また、タグ情報管理サーバ 1 9 0 は、タグ情報生成サーバ 1 0 0 が生成したタグ情報を Key - value 型データストアに管理する。タグ情報生成サーバ 1 0 0 は、Key - value 型データストアにて管理するのに適したタグ情報を生成する。大量のコンテンツデータ及び端末装置からのアクセスを拘束に処理するためにはクラウドコンピューティングの技術を利用することは有益であり、本実施形態においては、分散型のクラウドコンピューティングに適するといわれている Key - value 型のデータストアを利用してタグ情報を管理する方法を用いる。

20

#### 【 0 0 6 5 】

Key - value 型のデータストアは、図 8 に示されるように、キーと値の組を行としたシンプルなテーブルによりデータを管理する。Key - value 型のデータストアは、リレーショナルデータベースと比較して、複雑なデータ構造定義やリレーショナル機能を利用した検索をサポートしない代わりに、キーの値に基づいて、複数のサーバで大量なデータを分散処理することができる特長を有する。このような Key - value 型のデータストアは、近年、大規模な数のクライアントを対象としたインターネットサービスにおいて利用され始めている。

30

#### 【 0 0 6 6 】

Key - value 型のデータストアにおいては、テーブルは、キーの値によりソートされて管理される。このため、高速に、且つ、低負荷にて処理を行うためには、関連するデータは、テーブルの一部に局所化されるようにキーの値の形式を工夫することが非常に有効である。

#### 【 0 0 6 7 】

40

( 空間モデル )

図 9 は、本実施形態において、Key - value 型のデータストアによりタグ情報を管理する場合に用いるタグの空間座標について示す空間モデルである。本空間モデルにおいて、タグの位置情報は、視点位置情報 ( 緯度  $L A N_T$ 、経度  $L N G_T$ 、高度  $A T T_T$  )、及び視点位置からの方位情報 ( 水平方位角  $T_H$ 、垂直方位角  $T_V$  ) により示される。

#### 【 0 0 6 8 】

図 5 に示した空間モデル上の座標から、図 9 に示される空間座標への変換は、以下の式により行われる。なお、方位が真北を跨る場合には補正が必要である。

#### 【 0 0 6 9 】

50

タグの水平方位角： $T_H = F_H + F_W \times (X_T / W_F - 1 / 2)$

タグの垂直方位角： $T_V = F_V + (F_W \times H_F / W_F) \times (1 / 2 - Y_F / H_F)$

#### 【0070】

このとき、各記号は以下の意味を有する。

タグの画像面上の座標： $(X_T, Y_T)$

画像の撮影水平画角： $F_W$

画像の水平画素数： $W_F$

画像の垂直画素数： $H_F$

画素アスペクトレシオ： $1 : 1$

画像の撮影水平方位角： $F_H$

画像の撮影垂直方位角： $F_V$

10

#### 【0071】

なお、垂直方向の方位角情報は、真下を基点として0から180度までの値をとる。ここでは、説明の便宜上、水平を0度上空方面を正とし、地面方面を負として表記している。方位角情報が、コンテンツデータに付帯する撮影情報に含まれておらず不明な場合には、画角及び被写体距離から、垂直方位角を簡易的に割り当ててもよい。

#### 【0072】

(ViewHash)

タグ情報をKey-value型データストアにより管理する場合には、視点情報及び視界情報をキーの一部に用い、これらの情報によりデータを検索することができるようにする必要がある。このため、今回提案するViewhashを用いることにより、経度、緯度、位置精度、視野角、視野方向、及びタグの距離を効率よく文字列に符号化することができる。

20

#### 【0073】

Viewhashの構成が図10に示される。Viewhashのビット列は、位置情報、位置精度、視界情報、及びタグまでの距離としての被写体距離を示すビット列が順に並ぶ。位置情報のビット列は、経度(-90度から90度)、緯度(-180度から180度)、の座標を、分解精度にて表した値から生成する。分解精度のビット数R、ビットの経度n、緯度eとすると、分解精度にて表した経度Nおよび緯度Eの値は以下の式から求める。

30

#### 【0074】

$$N = 180 \times (n + 90) / 2^R - 1$$

$$E = 360 \times (e + 180) / 2^R$$

#### 【0075】

NおよびEの値は、浮動点小数であり、小数点以下は切り捨てて、n、およびeは整数値とする。なお、分解された地図上の区画を正方に近づけるために、経度の分解精度は、緯度の分解制度の1/2となるようにしてある。

#### 【0076】

位置情報のビット列は、図10に示したように、先の式で求められた経度のビット列( $E - b_n, E - b_{n-1}, \dots, E - b_1, E - b_0$ )および経度のビット列( $N - b_n, N - b_{n-1}, \dots, N - b_1, N - b_0$ )を交互に配置した形式である。ここで、最上位ビットはそれぞれ $E - b_n$ 及び $N - b_{n-1}$ である。

40

#### 【0077】

位置情報のビット列をこのような形式にすることによって、地球上の地理区画を特定することができ、少ないビット列により区画の位置を示すとともに、区画の長さも表すことができる。本実施形態においては、GPSの位置精度が1meterから100meterであることを考慮して、Rは22から25の値を用いる。このとき位置情報は、最小で43ビット、最大で49ビットとなる。

#### 【0078】

50

図 1 1 は、位置情報のビット数が 4 5 ビットである場合の、分割される地理区画の例を示す説明図である。この地理区画の南北方向の長さは、経度にかかわらず一定であり、約 4 . 7 5 メータとなる。一方、地理区画の東西方向の長さは、緯度に依存して、例えば、東京などの緯度が 3 5 度の付近においては 3 . 8 6 メータとなる。これに対してストックホルムなどの緯度が 6 0 度の付近においては、地理区画の東西方向の長さは、約 2 . 3 5 メータとなる。

#### 【 0 0 7 9 】

位置精度は、位置情報のビット列により特定された地理区画の南北方向の距離の何倍であるかをビット列により表し、D - b 7 が最上位ビット、D - b 0 が最下位ビットである。位置情報のビット列により特定された地理区画の南北方向の距離の 1 倍であった場合には、ビット列の値は 0、位置情報のビット列により特定された地理区画の南北方向の距離の 1 倍超 2 倍以下の場合には、ビット列の値は 1、位置情報のビット列により特定された地理区画の南北方向の距離の 2 倍超 4 倍以下であった場合には、ビット列の値は 3、位置情報のビット列により特定された地理区画の南北方向の距離の 4 倍超であった場合には、ビット列の値は 4 となる。位置誤差のビット数を 0 とする。

10

#### 【 0 0 8 0 】

視界情報は、タグの方位がどの視界に属するかを示す情報である。図 1 2 及び図 1 3 に示すように、視界情報は 4 - 6 ビット列であり、真東を起点として時計回りとした方位を 3 6 0 度の座標を分解精度で除した値をビット列で表す。真東を起点として時計回り表した方位の度数を d、分解精度のビットの精度を A としたとき、視界情報 D は以下の式にて算出される。なお、以下の説明では便宜のため、方位は真北を起点として表記しているが、実際にはすべて、真東を起点として変換した値を用いる。

20

#### 【 0 0 8 1 】

$$D = 360 \times d / 2^A$$

#### 【 0 0 8 2 】

本実施形態においては、分解精度のビット数は 4、5、6 の 3 通りとし、視野角に応じて異なるビット数が用いられる。ビット数により 3 6 0 度の空間が、それぞれ 2 2 . 5 度、1 1 . 2 5 度、8 . 1 2 5 度に分割される。本実施形態においては、タグが付加された画像の撮影画角に応じてビット数を変え、画角が 4 6 度以上の場合には 4 ビット、4 6 度未満 1 8 度以上の場合には 5 ビット、1 8 度未満は 6 ビットの精度が割り当てられ、3 6 0 度の空間がそれぞれ 2 2 . 5 度、1 1 . 2 5 度、8 . 1 2 5 度に分割された方位のどれに属するかを示す。図 1 2 は、4 ビットの場合の各方位のビット数が示されている。たとえば情報タグの水平方位が北北西は 3 1 5 度であった場合は、視野情報は、1 1 0 1 のビット値となる。なお、方位情報が不明の場合、ビット数は 0 とする。

30

#### 【 0 0 8 3 】

このように、視野角に応じて視界情報のビット数を異ならせ、このビットにより示される視界情報を Key - v a l u e 型データのキーの一部として用いることによって、キーに含まれる視界情報のビット数から視野角の情報を把握することができるようになる。視野角の情報を把握することができるようにより、より正確に各映像コンテンツ内に撮影されている範囲を知ることができるようになる。

40

#### 【 0 0 8 4 】

被写体距離は、メータ値の 8 ビットにて表され、上述のようにコンテンツデータに付帯する撮影情報に基づいて、被写体までの距離を求めた値として用いる。厳密には撮影画の過焦点距離とタグ情報の距離は一致しないが、もともと距離情報には誤差が含まれているものとして扱う。距離が 2 5 5 メーター以上のものは 2 5 5 とし、距離情報が不明の場合はビット数は 0 とする。

#### 【 0 0 8 5 】

図 1 4 は、位置情報のビット数の割り当て方式を示した説明図である。Key - v a l u e 型データストアの Key による検索を効率良くするために、タグ情報の検索方式に適するように Key 値を決めることは有効である。本実施形態においては、タグの視野のビ

50

ット数および距離情報によってタグのKey値を構成するViewHashの位置ビット数を変える。例えば、視界情報のビット数が3の場合は、タグが付加された映像は広角でとられたものであり、かつ、被写体距離が遠方(100m超)であった場合は、位置情報の精度が悪い場合でもタグ情報を利用することができるので、位置情報のビット数は43ビットを用いる。一方、視野情報のビット数が5、つまり、超望遠で撮影した映像に付加したタグ情報であれば、位置情報の精度も要求されるため、すべての距離において位置情報のビット数は49ビットを用いる。このように位置情報のビット数の割り当てを変えることで、有益なタグ情報の検索を効率よくおこなうことが可能である。前述のように、タグの位置誤差は位置情報のビット数に依存するので、比較対照の視点の位置精度によって、タグ情報の検索で用いるViewHashの位置情報のビット数を決定できる。なお、タグの検索については詳しくは後述する。

10

#### 【0086】

上記のように、ViewHashを構成する各ビット列は求めることができる。ところが、Key-value型のデータストアのKeyの値として用いる場合、又は、URLなどとして用いるためには文字列で表記した方が扱い易い。このため、ビット列を文字列に符号化する。ビット列を文字列に符号化するためには、例えばBASE64を利用する方法がある。しかし、本実施形態においては、各ビット列に依存させてKey値でデータがソートできるようにするために、可変ビット列文字符号変換方式を提案する。

#### 【0087】

ここで提案する可変ビット列文字符号変換方式は、可変長のビット列をつなげてキーを生成する場合に、各ビット列の区切りを判別することができるという特長を有する。このため、本方式を用いて符号文字に変換したビット列は、そのまま連結させた場合であっても、各情報の区切りを判別することができる。このため、Key-value型データストアのキーとして用いる場合に、各情報の値に基づいて検索することが可能となる。

20

#### 【0088】

図15に可変ビット列符号化変換表の一例を示す。本変換方式は、情報が複数の可変長ビット列(0ビット以上)から構成されており、それらビット列の境界を識別する場合に有効である。この変換方法について説明する。各可変長ビット列は最上位ビットから4ビット単位にてフラグメントを順次切り出し、その4ビット値に対応した文字コード(Uからz)にて置き換える。もし、切り出すフラグメントのビット数が4に満たない場合は、3ビット、2ビット、1ビット、それぞれに文字コード(EからT)、(6からD)、(2から5)、(0,1)にて置き換える。また、最後の切り出したフラグメント4ビット、もしくは該当のビット列のビット数が0であった場合は、文字コード-(ダッシュ)を追加する。なお、ここで示す変換表は一例であることは言うまでもない。ビット列と符号文字とが1対1に対応するように規定されるものであれば、各種の変形例が考えられる。

30

#### 【0089】

図16に変換の例を示す。変換前の情報が2つのビット列から構成され、それぞれ「01111010010010」と「10010」であった場合、変換後の文字列は"im8W-"となる。この変換方式により図10で示されるViewHashの値としては文字列に変換したものを用いる。

40

このViewHashによって、任意の精度で分割された地球座標上の区画の中心を視点として、任意の精度にて分割された視野の3次元空間を一意に特定することが可能になるが、このViewHashを利用してタグ情報をKey-value型データストアで管理する方法について説明する。

#### 【0090】

次に、本実施形態におけるKey-value型データストアのテーブルについて説明する。本実施形態において使用するKey-value型データストアのテーブルは、図8に示したようにKey値とValueの値の対をなす。図17には、タグ情報をKey-value型のデータとして管理するために用いるJava(登録商標)クラスの定義が示される。Tagクラスはタグ情報を表し、インナークラスであるKeyクラスとVa

50

l u e クラスから構成される。K e y クラスと V a l u e クラスとの対が K e y - v a l u e としてデータストアで管理される。

#### 【0091】

K e y クラスの実データは、変数 K e y であり、K e y はタグ情報から生成した V i e w H a s h と文字列 T y p e と文字列 T i m e S t a m p を連結した文字列である。T y p e は、同じ V i e w H a s h の値の T a g 情報が適切にソートされるために用いられ、本実施例ではタグ表示する優先度を T y p e 値とし 0 ビットの優先値を図 15 の変換表にて生成した文字列である。同じ V i e w H a s h 値のタグ情報にユニークな k e y を割り当てるのを主目的としており、タグが生成された日時を J a v a (登録商標) の D a t e クラスと同じ 8 バイトの値で求め、それを図 15 の変換表を用いて文字列に生成したものである。

10

#### 【0092】

V a l u e クラスの実データは以下の変数である。

#### 【0093】

n a m e : タグの名前を文字列

t y p e : タグ表示の優先度 ( K e y 生成に使用しものと同じ値 )

t i m e S t a m p : タグの生成日時 ( K e y 生成に使用しものと同じ値 )

t a g P o s i t i o n : タグの視点の位置 ( 例 : 図 10 の L A N <sub>T</sub> , L N G <sub>T</sub> , A T T <sub>T</sub> など )

t a g D i r e c t i o n : 視点からのタグの方位 ( 例 : 図 10 の T <sub>H</sub> , T <sub>V</sub> )

20

t a g S o u r c e : タグ情報を生成した基となる画像情報とタグの画面情報座標

#### 【0094】

また、T a g S o u r c e クラスの実データは以下の変数である。

#### 【0095】

i m a g e S o u r c e 画像情報

t a g X : 画像面でのタグの X 座標値 ( 例 : 図 10 の X <sub>T</sub> )

t a g Y : 画像面でのタグの Y 座標値 ( 例 : 図 10 の Y <sub>T</sub> )

#### 【0096】

また、I m a g e S o u r c e クラスの実データは以下の変数である。

#### 【0097】

u r l : 画像ファイル ( 動画、静止画 ) の U R L ( U n i q u e R e s o u r c e L o c a t o r )

i m a g e P o s i t i o n : 画像撮影した視点の座標 ( 例 : 図 9 では L A N <sub>T</sub> , L N G <sub>T</sub> , A T T <sub>T</sub> などに同じ )

i m a g e D i r e c t i o n : 視点から画像の中点への方位 ( 例 : 図 9 では F <sub>H</sub> , F <sub>V</sub> )

s c e n e : 画像ファイル ( 動画、静止画 ) の画像情報

#### 【0098】

そして、S c e n e クラスの実データは以下の変数である。

#### 【0099】

f r a m e ; 静止画の場合は 0、動画の場合は画像フレームを先頭からの時間 ( ミリ秒 ) で表記

a n g l e H o r i z o n t a l : 画像の撮影水平画角 ( 例 : 図 9 の F <sub>W</sub> )

w i d t h : 画像の水平画素数 ( 例 : 図 9 の W <sub>F</sub> )

h e i g h t : 画像の垂直画素数 ( 例 : 図 9 の H <sub>F</sub> )

#### 【0100】

V i e w P o s i t i o n の実データは以下の変数である。

#### 【0101】

l a t i t u d e ; 視点の緯度

l o n g i t u d e ; 視点の経度

40

50

altitude ; 視点の高度 (ミリメートル単位)

range : 視点の誤差 (ミリメートル単位)

【0102】

ViewDirectionの実データは以下の変数である。

【0103】

directionHorizontal : 視点から対象点への水平方位

directionVertical ; 視点から対象点への垂直方位

distance ; 視点から対象点への距離 (ミリメートル単位)

【0104】

次に、図18及び図19は、タグ情報の検索の例について示した説明図である。図18は、視点の位置 (東京と同じ北緯35度) から画角50度にて真北を撮影した場合を示しているが、この視界に含まれるタグ情報を検索するには、視野の方位、画角を図13に示す視界情報ビットによって特定される視野毎にタグを探す。この例では視野は25度に分割された4つの視野分割に対する検索を行う必要がある。また、視野位置の測定誤差は位置情報ビットにて特定される区画の南北方向の長さ以下であったとしても、区画の縁に視点位置がある場合には誤差によって、実際はその周辺の区画に位置している場合もある。このため、特定された区画の周辺の区画も検索する必要がある、誤差が3mの場合であれば、合計9個の区画を検索する必要がある。この場合、検索としては  $4 * 8 = 32$  の操作を行ことになる。

【0105】

図19の例では、視点の位置がストックホルム近郊の北緯60度の例を示してある。この場合は、区画の東西方向が先の場合より短くなるので、同じ位置誤差精度であったとしてもより、広い範囲の区画の検索が必要である。

【0106】

このような検索処理を行う再生装置の機能構成について、次に説明する。図20は、再生装置400の構成を示すブロック図である。再生装置400は、通信部410と、制御部420と、表示部430とを主に有する。制御部420は、解析部421の機能と、検索部422の機能と、重畳部423の機能と、表示制御部424の機能とを主に有する。

【0107】

通信部410は、インターネットなどのネットワーク500を介して他の装置と通信するためのネットワークインタフェースである。また、制御部420は、再生装置400全体の動作を制御する。表示部430は、例えば、CRT (Cathode Ray Tube) ディスプレイ装置、液晶ディスプレイ (LCD) 装置、OLED (Organic Light Emitting Diode) 装置およびランプなどの表示装置により構成される。

【0108】

解析部421は、再生装置400が再生するコンテンツを解析する機能を有する。解析部421は、例えば、検索部422がタグ情報を検索するために必要な情報を、コンテンツデータに付帯する撮影情報を解析することにより取得する機能を有する。また、必要に応じて、解析部421は、コンテンツの映像そのものを解析する機能を有していてもよい。

【0109】

検索部422は、再生するコンテンツデータに重畳するタグの情報を、データセンタ1000内のデータベースから検索する機能を有する。ここでは、検索部422は、Key-value型のデータストアに記憶されたタグ情報を検索するために適した検索条件を生成して、データセンタ1000内のタグ情報管理サーバ190に送信し、検索結果を取得する。タグ情報検索処理の詳細については、次に図21を用いて説明する。

【0110】

重畳部423は、検索部422が取得したタグ情報から、再生する映像コンテンツにタグを重畳させる処理を行う。この重畳処理の詳細については、後述される。また、表示制

10

20

30

40

50



御部 4 2 4 は、コンテンツデータを再生する機能を有し、重畳部 4 2 3 によりタグが重畳されたコンテンツデータを再生する。

【0 1 1 1】

図 2 1 は、タグ検索の動作を示すフローチャートである。Key - value 型のデータストアでは、Key の値の先頭と終わりを指定して、その範囲の Key 値の配列を得ることができる。このため、撮影された画像の位置情報、方位、および、撮影された画像の被写体距離の範囲から、ViewHashBegin と ViewHashEnd による Key - value 型データストアの検索を、視界と区画の分だけ検索を繰り返す。ただし、先に述べたように Key - value 型データストアでは分散処理が容易に行えるようになっている。このため、検索要求は並列に行えば、分散して処理することも可能である。

10

【0 1 1 2】

タグ情報検索処理は、まず、再生装置 4 0 0 が、対象とする映像の撮影情報を取得することにより始まる (S 1 0 3)。そして、検索部 4 2 2 は、取得された撮影情報の画角の情報から視野ビット数を決定する (S 1 0 6)。次に、検索部 4 2 2 は決定した視野ビット数と、撮影情報より得られた距離の情報から、検索するグリッドの位置ビット数を決定する (S 1 0 9)。

【0 1 1 3】

そして次に、検索部 4 2 2 は、測定誤差の値が、グリッドの南北の幅以下であるか否かを判断する (S 1 1 2)。ステップ S 1 1 2 における判断により、測定誤差がグリッドの南北幅より大きいと判断された場合には、検索対象のグリッドを特定することができないため、B に進んで処理を終える。

20

【0 1 1 4】

一方、ステップ S 1 1 2 の判断により、測定誤差がグリッドの南北幅以下であると判断された場合には、次に、検索部 4 2 2 は、検索するグリッドの位置情報を算出して配列にセットする (S 1 1 5)。そして、検索部 4 2 2 は、画角及び視野ビット数から対応する視界を算出して配列に入れる (S 1 1 8)。

【0 1 1 5】

そして、検索部 4 2 2 は、結果を保存する Tags [ ] 配列を初期化し (S 1 2 1)、距離精度の値を 0 にする (S 1 2 4)。そして、次に図 2 1 B に進み、検索部 4 2 2 は、最初の ViewPosition を選択し (S 1 2 7)、検索する距離範囲を決定する (S 1 3 0)。

30

【0 1 1 6】

そして、検索部 4 2 2 は、ViewHashBegin を算出し (S 1 3 3)、ViewHashEnd を算出する (S 1 3 6)。そして、検索部 4 2 2 は、ステップ S 1 3 3 及びステップ S 1 3 6 において算出された ViewHashBegin と ViewHashEnd の間の範囲のタグを検索する。

【0 1 1 7】

具体的には、検索部 4 2 2 は、検索条件を生成して、タグ情報管理サーバ 1 9 0 に送信する。そして、検索部 4 2 2 は、送信した検索条件に合致する検索結果を取得する。検索部 4 2 2 は、取得した検索結果を Tags [ ] に追加保存する。

40

【0 1 1 8】

そして、全ての ViewDirection を検索したか (S 1 4 2) 及び全ての ViewPosition を検索したか (S 1 4 4) が判断され、全てについて検索が完了するまでステップ S 1 3 0 からの処理が繰り返される。

【0 1 1 9】

次に、再生装置 4 0 0 の重畳部 4 2 3 により行われる重畳処理について図 2 2 を参照しながら説明する。図 2 2 は、検索により得られたタグ情報を、他の映像に重畳する方法について示した説明図である。

【0 1 2 0】

50

3 辺の距離が既知である場合は、余弦定理より夾角は求めることができ、タグの視点座標および方位から、撮影視点からみた視野に投影したタグの方位は以下の式より求めることができる。

【 0 1 2 1 】

投影されたタグの 水平方位角

$$U_H = T_H - \cos^{-1} (D_U^2 + D_S^2 - D_{DH}^2) / (2 * D_U * D_S) \quad \text{where } S_H > T_H$$

$$U_H = T_H + \cos^{-1} (D_U^2 + D_S^2 - D_{DH}^2) / (2 * D_U * D_S) \quad \text{where } S_H < T_H$$

投影されたタグの垂直方位角  $U_V = T_V + \cos^{-1} (D_U^2 + D_S^2 - D_{DV}^2) / (2 * D_U * D_S) \quad \text{where } S_V > T_V$  10

$$U_V = T_V - \cos^{-1} (D_U^2 + D_S^2 - D_{DV}^2) / (2 * D_U * D_S) \quad \text{where } S_V < T_V$$

【 0 1 2 2 】

ここで、情報は次の符号により示される。

タグの位置情報 ( L A N<sub>T</sub> , L N G<sub>T</sub> , A T T<sub>T</sub> )

撮影地点の位置情報 ( L A N<sub>S</sub> , L N G<sub>S</sub> , A T T<sub>S</sub> )

視点間の水平方向距離 D<sub>DH</sub>

視点間の垂直方向距離 D<sub>DV</sub>

タグの水平方位角 T<sub>H</sub> :

タグの垂直方位角 T<sub>V</sub>

タグの距離距離 D<sub>T</sub>

被写体距離の距離 D<sub>S</sub>

【 0 1 2 3 】

D<sub>DH</sub> は 2 つの視点の水平距離であり、視点間の標高差がなく垂直距離 D<sub>DV</sub> を 0 とすれば、D<sub>DHh</sub> は 2 点の経度、緯度から Hubeny の公式によって距離を求め D<sub>DHh</sub> とする。D<sub>Sh</sub> は撮影された画像の被写体距離であり、ここでは過焦点距離が用いられ、D<sub>Th</sub> はタグ情報の距離である。垂直距離 D<sub>DV</sub> 0 としているので、投影された垂直方位角はタグの垂直方位角と等しくなる。

【 0 1 2 4 】

検索されたタグを撮影された映像に投影する場合は、検索されたすべてのタグについて上記の座標変換を行い、タグの方位角が撮影の画角に収まるものだけを選択し投影する。また、タグ情報を元に、たとえば、前述の Type に指定されている優先度が高いもののみ表示するなどフィルタリングして表示をおこなう。

【 0 1 2 5 】

投影されたタグの水平方位角、垂直方位角が求めれば、撮影された映像の画像面での座標は以下の式で求めることができる。

【 0 1 2 6 】

$$\text{投影されたタグ位置 } X_S = ( (U_H - S_H) / (S_W + 1 / 2) ) * W_S$$

$$\text{投影されたタグ位置 } Y_S = ( 1 / 2 - (U_V - S_V) / (S_W * H_S / W_S) ) * H_S ; \quad 40$$

【 0 1 2 7 】

撮影映像の水平画角 : S<sub>W</sub>

撮影映像の水平画素数 W<sub>S</sub>

撮影映像の垂直画素数 H<sub>S</sub>

画素アスペクトレシオ 1 : 1

撮影方向水平方位角 S<sub>H</sub>

撮影方向垂直方位角 S<sub>V</sub> :

【 0 1 2 8 】

< 5 . 効果の例 >

以上のように、画像に付加されたタグ情報をKey-value型データストアで管理し、他の画像の撮影情報を投影可能なタグ情報を検索し、それらを重畳して表示する方法について示した。

【0129】

かかる構成によれば、ある視点位置からみた視界空間上の座標に関連づけられたタグ情報を生成し、共有して用いることができる。これにより、所定の撮影条件により、映像を撮影した位置、及び映像を撮影した方位から見た映像中においてしか意味を成さない情報を共有することができるようになる。

【0130】

また、このタグ情報を、視点の位置情報、視界情報、被写体距離、位置情報の誤差精度を簡潔に表現することのできるViewHashが新たに提案された。ViewHashを用いることにより、Key-value型のデータストアによりタグ情報を管理する場合において、タグ情報が大量となった場合であっても、高速に、軽い処理負荷で検索することができるようになる。

【0131】

< 6 . 適用例 >

次に、本発明の一実施形態形態に係るタグ情報の管理方法を適用することにより実現することのできるサービスの一例について説明する。

【0132】

例えば、図23は、ビデオカメラを用いて撮影された風景の映像に、リアルタイムにタグを重畳させて表示する例である。図1においては、重畳される映像は、予め撮影されて記憶された映像コンテンツであることとしたが、本発明はかかる例に限定されない。図23に示すように、撮影中の映像に対してリアルタイムにタグが表示されるようにしてもよい。

【0133】

また、上記実施形態においては、タグ情報は、インターネットなどのネットワークを介して取得することとしたが、本発明はかかる例に限定されない。例えば、撮像装置内の記憶部に予め訪問先の地域のタグ情報を記憶して出かければ、ネットワークに接続しなくてもタグを表示させることができる。

【0134】

或いは、図24は、地図情報と共に、予め撮影された映像にタグを表示させた映像4025を表示させる例である。タグ情報は、必ずしもリアルタイムの映像に重畳させることに意味があるわけではなく、見ている風景上の対象物が何であるかを知りたい場合には、過去の映像に重畳させて示してもよい。かかる構成によれば、撮像装置のレンズを被写体に向ける動作をしなくても、例えば携帯電話であればメールを打つ操作を行うときと同じ体勢で自然に所望の情報を得ることができる。リアルタイムの映像にタグを重畳させる場合には、必ず撮像装置のレンズを対象物に向けなくてはならないが、時と場所によっては、そのような動作をすることが困難な場合があるため、本構成は有効である。

【0135】

また、図24の装置は、写真を撮るのに適したスポットの情報をユーザに通知することも可能である。例えば、コンテンツデータを有する映像配信サーバ180において、記憶されたコンテンツデータの撮影位置を示す視点情報を解析することにより、このようなサービスを実現することができる。例えば、視点情報が多く集まるスポットは、写真を撮るのに適したスポットである可能性が高い。このため、撮影位置をお勧めする機能を提供するためにかかる情報が用いられてもよい。

【0136】

以上、添付図面を参照しながら本発明の好適な実施形態について詳細に説明したが、本発明はかかる例に限定されない。本発明の属する技術の分野における通常の知識を有する者であれば、特許請求の範囲に記載された技術的思想の範疇内において、各種の変更例または修正例に想到し得ることは明らかであり、これらについても、当然に本発明の技術的

10

20

30

40

50

範囲に属するものと了解される。

【 0 1 3 7 】

例えば、本実施形態においては、デジタルスチルカメラ、デジタルビデオカメラにGPS機能および方位機能を有して、その機能によってファイルに記録された撮影情報とその画像にユーザがPCにて付加したタグから、Key-value型データストアのデータを生成したが、対象物の座標位置およびそれに関わる情報、たとえば、山の名前などが分かっている場合には、コンピュータの処理によってある視点からのタグ情報を自動生成することも可能である。

【 0 1 3 8 】

また、従来は、GPSの測定誤差によってタグ情報の座標に誤差が大きいタグ情報を利用できなかったが、本実施例によれば、タグ情報の付近の区画にある他のタグも容易に検索でき、また、タグ情報にはタグ生成の基となった画像の情報も得ることができる。このため、タグ生成の基となった画像を解析し、被写体が近似し、かつ、位置精度が正確な画像を近傍の区画から検索することで、タグの位置情報を補正することが容易に実現可能である。

10

【 0 1 3 9 】

また、本実施形態においては、GPSの測定誤差によってタグ情報の座標に誤差が大きいタグ情報を利用できない場合が起こるが、本実施例によれば、タグ情報の付近の区画にある他のタグも容易に検索でき、また、タグ情報にはタグ生成の基となった画像の情報も得ることができるため、その画像を解析し、被写体が近似し、かつ、位置精度が正確な画像を近傍の区画から検索することで、タグの位置情報を補正することが容易に実現可能である。

20

【 0 1 4 0 】

また、タグ情報の検索を行う映像の撮影した座標情報もGPSの測定誤差が大きい場合はタグ情報を利用できない場合が起こるが、撮影した座標情報から、その周辺の区画のタグ情報は容易に検索できる。このため、撮影した画像の情報とその撮影座標周辺のタグ情報の基となった画像の特徴量を比較することで、特徴量が近似した画像の座標を元に、撮影した画像の座標情報を補正することが容易に実現可能である。この方式は、実時間で撮影しているビデオカメラの位置情報の補正にも適応できる。

【 0 1 4 1 】

また、上記実施形態においては、視界空間の奥行き情報は、コンテンツに付帯する撮影情報から取得する、又は過焦点距離が用いられることとしたが、本発明はかかる例に限定されない。例えば、タグの距離と同様に、あらゆる解析技術を用いて取得することのできる被写体までの距離を示すことのできる値が用いられてよい。

30

【 0 1 4 2 】

尚、本明細書において、フローチャートに記述されたステップは、記載された順序に沿って時系列的に行われる処理はもちろん、必ずしも時系列的に処理されなくとも、並列的に又は個別的に実行される処理をも含む。また時系列的に処理されるステップでも、場合によっては適宜順序を変更することが可能であることは言うまでもない。

【 符号の説明 】

40

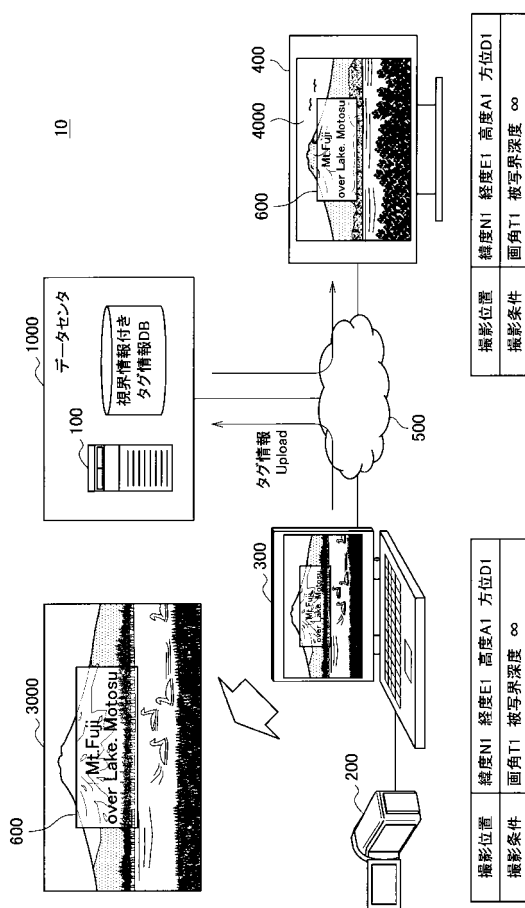
【 0 1 4 3 】

- 1 0 0 タグ情報生成装置
- 1 1 0 通信部
- 1 2 0 制御部
- 1 2 1 視点情報取得部
- 1 2 2 視界情報取得部
- 1 2 4 タグ情報生成部
- 2 0 0 撮像装置
- 3 0 0 端末装置
- 4 0 0 再生装置

50

- |       |        |
|-------|--------|
| 4 1 0 | 通信部    |
| 4 2 0 | 制御部    |
| 4 2 1 | 解析部    |
| 4 2 2 | 検索部    |
| 4 2 3 | 重畳部    |
| 4 2 4 | 表示制御部  |
| 4 3 0 | 表示部    |
| 5 0 0 | ネットワーク |

【 図 1 】

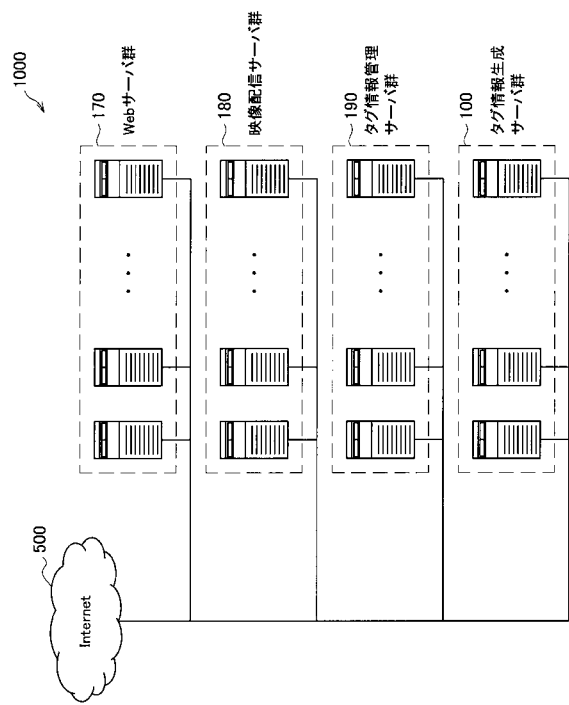


【圖 2】

mdp_id	名称	概要	mdp_id	名称	概要
Ox13	Time Code	タイムコード	OxB0	GPS Version ID	GPSバージョンID
Ox14	Binary Group	バイナリグループ	OxB1	GPS Latitude Ref	GPS緯度方向
Ox18	REC Time&Date1	撮影日時1	OxB2	GPS Latitude1	GPS緯度1
Ox19	REC Time&Date2	撮影日時2	OxB3	GPS Latitude2	GPS緯度2
Ox70	Consumer Camera1	コンシューマカメラ1	OxB4	GPS Latitude3	GPS緯度3
Ox71	Consumer Camera2	コンシューマカメラ2	OxB5	GPS Longitude Ref	GPS経度方向
Ox73	Lens	レンズ	OxB6	GPS Longitude1	GPS経度1
Ox74	Gain	ゲイン	OxB7	GPS Longitude2	GPS経度2
Ox75	Pedestal	ベDESTAL	OxB8	GPS Longitude3	GPS経度3
Ox76	Gamma	ガンマ	OxB9	GPS Altitude Ref	GPS高度方向
Ox77	Detail	輪郭強調	OxBa	GPS Altitude	GPS高度
Ox7B	Camera Preset	カメラプリセット	OxBB	GPS Timestamp1	GPSタイムスタンプ1
Ox7C	Flare	フレア	OxBC	GPS Timestamp2	GPSタイムスタンプ2
Ox7D	Shading	シェーディング	OxBD	GPS Timestamp3	GPSタイムスタンプ3
Ox7E	Knee	ニー	OxBE	GPS Status	GPS動作状態
Ox7F	Shutter	シャッター	OxBF	GPS Measure Mode	GPS計測モード
OxA0	Exposure Time	露光時間	OxC0	GPS DOP	GPS精度劣化係数
OxA1	F Number	F値	OxC1	GPS Speed Ref	GPS速度方向
OxA2	Exposure Program	露出プログラム	OxC2	GPS Speed	GPS速度
OxA3	Brightness Value	ブライトネス値	OxC3	GPS Track Ref	GPS軌跡方向
OxA4	Exposure Bias Value	露光バイアス値	OxC4	GPS Track	GPS軌跡
OxA5	Max Aperture Value	最大絞り	OxC5	GPS Img Direction R	GPS画像方向R
OxA6	Flash	フラッシュ	OxC6	GPS Img Direction	GPS画像方向
OxA7	Custon Rendered	特殊処理	OxC7	GPS Map Datum1	GPS地図データ1
OxA8	White Balance	ホワイトバランス	OxC8	GPS Map Datum2	GPS地図データ2
OxA9	Focal Length in 35mm Film	焦点距離 (35ミリフィルム)	OxC9   OxCF	GPS Option	GPSオプション
OxAA	Scene Capture Type	撮影シーンタイプ	OxCE0	Maker & Model ID	メーカー・モデルID
OxAB   OxAF	Exif Option	Exifオプション	OxE1   OxEF	Maker Option	メーカーオプション
			その他	(Reserved)	(予約)



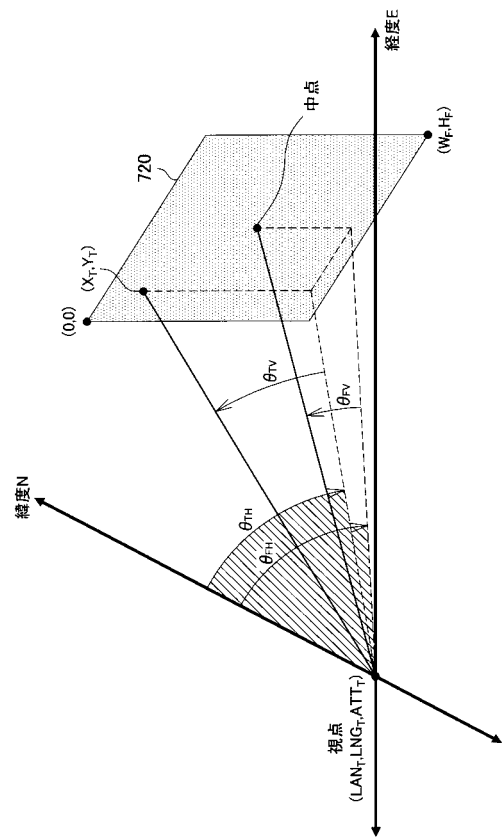
【 図 7 】



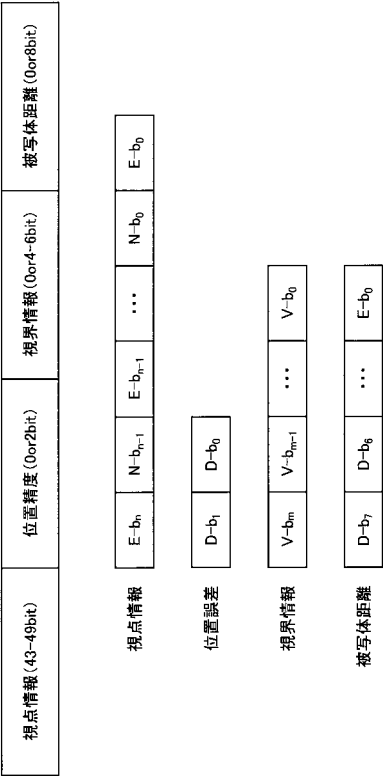
【 図 8 】

Key	Value
ABC	値1
AGF	値2
AGF8G	値3
AGF9G	値4
AGE	値5
AF	値6
AF72H	値7

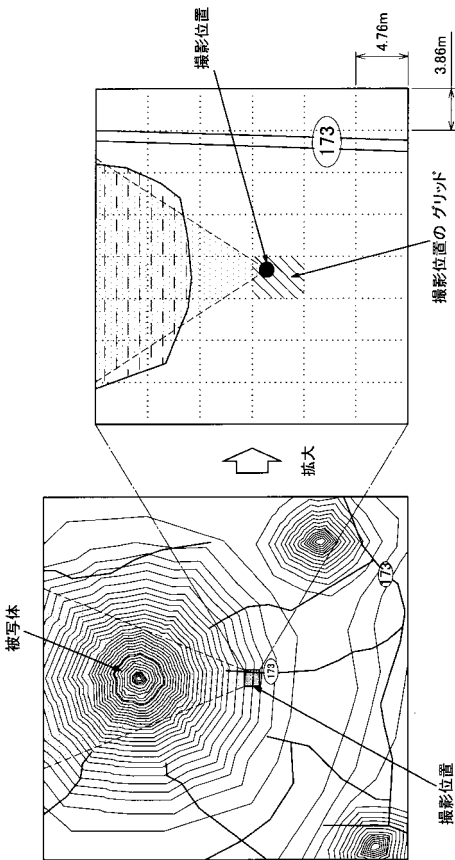
【 図 9 】



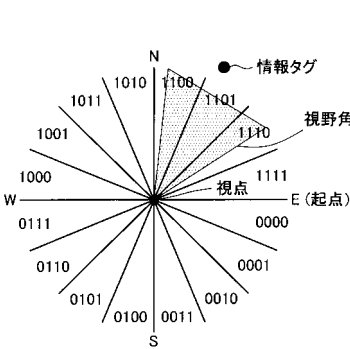
【 図 10 】



【図 1 1】



【図 1 2】



【図 1 3】

ビット数	方位分割 (度)	画角 (度)	参考 (35mm換算)	
			視野分類	焦点距離
4	22.5	> 74	超広角	< 28
		>= 46	広角	28-50
5	11.25	>= 18	ズーム	50-105
6	6.125	< 18	超ズーム	> 105

【図 1 4】

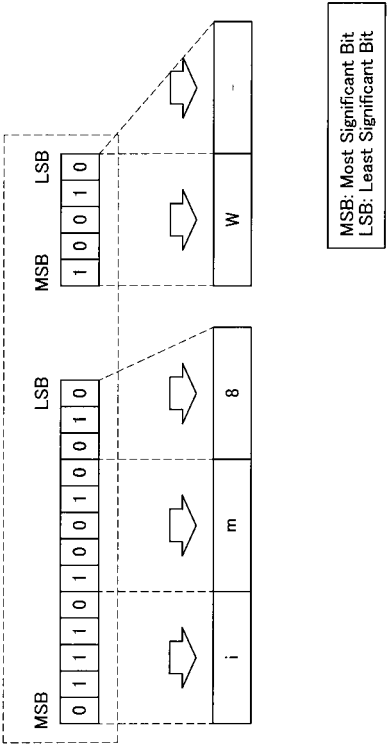
視点情報		視野情報のビット数(視野角)				距離			比較対象の 視点の位置誤差
ビット数	区画の長さ (南北)	3 (広角)	4 (ズーム)	5 (超ズーム)	100m超	100m未満 10m以上	10m未満 5m以上	5m以下	
43	約9.53m	○			○				9.53m以下
45	約4.76m	○				○			4.76m以下
47	約2.38m	○	○		○		○		2.39m以下
49	約1.19m	○	○					○	1.19m以下
				○	○	○	○	○	

【図 1 5】

ビット列	符号文字	ビット列	符号文字	ビット列	符号文字	ビット列	符号文字
なし(区点)	- (hyphen)	0000	E	00000	U	10000	k
0	0	0001	F	00001	V	10001	l
1	1	0010	G	00010	W	10010	m
00	2	0011	H	00011	X	10011	n
01	3	0100	I	00100	Y	10100	o
10	4	0101	J	00101	Z	10101	p
11	5	0110	K	00110	a	10110	q
000	6	0111	L	00111	b	10111	r
001	7	1000	M	01000	c	11000	s
010	8	1001	N	01001	d	11001	t
011	9	1010	O	01010	e	11010	u
100	A	1011	P	01100	f	11100	v
101	B	1100	Q	01101	g	11101	w
110	C	1101	R	01110	h	11110	x
111	D	1110	S	01111	i	11111	y
		1111	T	01111	j	11111	z



【 図 1 6 】



【 図 1 7 】

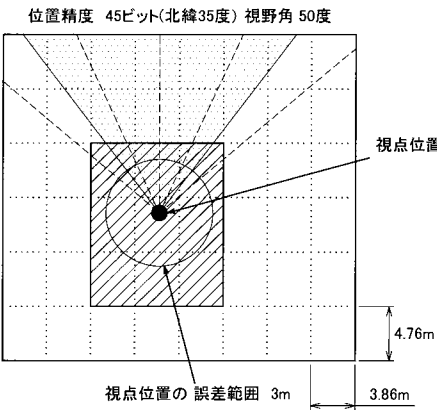
```
public class ImageSource {
    String uri;
    ViewPosition imagePosition;
    public class scene {
        int frame;
        double angleHorizontal;
        int width;
        int height;
    };
};

public class ViewPosition {
    double latitude;
    double longitude;
    double altitude;
    double range;
};

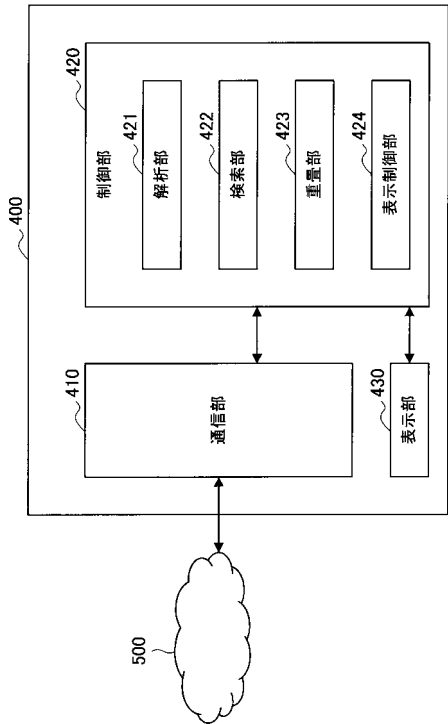
public class ViewDirection {
    double directionHorizontal;
    double directionVertical;
    int distance;
};

public class Tag {
    public class Key {
        String key;
        public void key(
            String viewHash,
            Unsigned char type,
            Date timeStamp) {
            key = viewHash
                + type
                + timeStamp;
        };
    };
    public class Value {
        String name;
        int type;
        Date timeStamp;
        ViewPosition tagPosition;
        ViewDirection tagDirection;
        class tagSource {
            ImageSource imageSource;
            int tagX;
            int tagY;
        };
    };
};
```

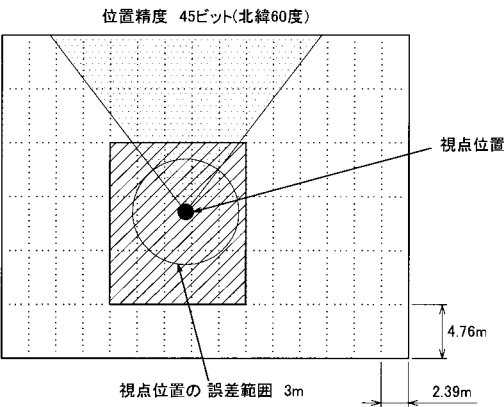
【 図 1 8 】



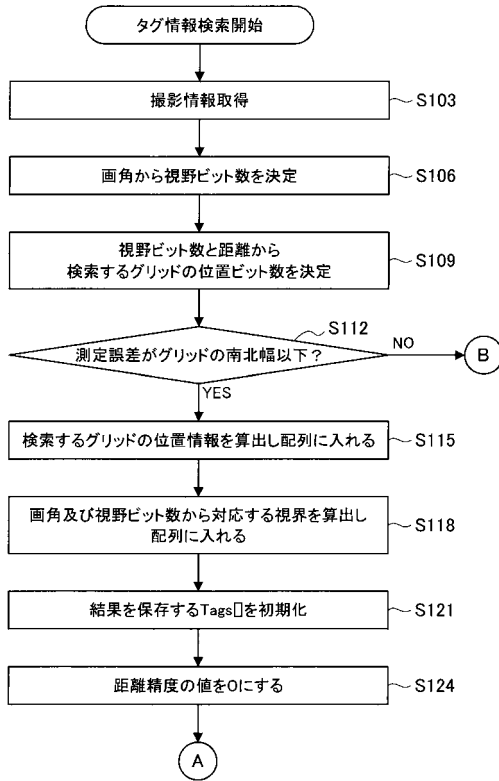
【 図 2 0 】



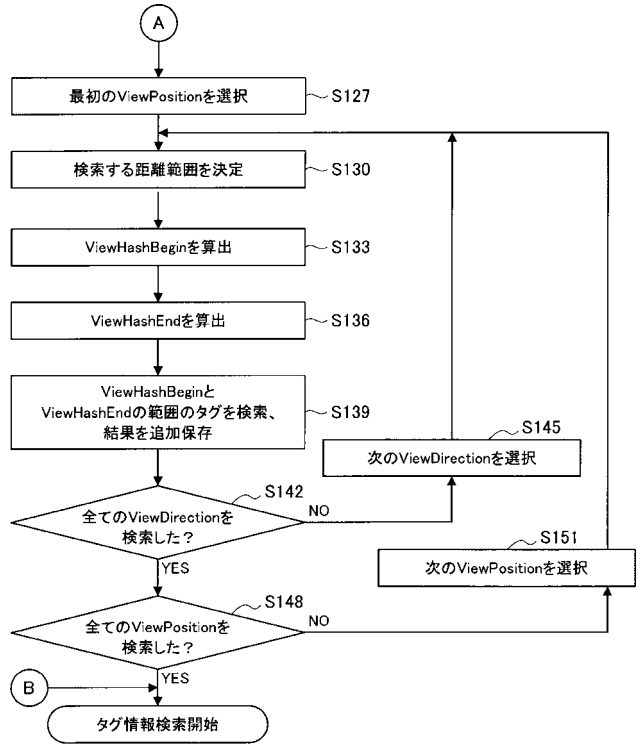
【 図 1 9 】



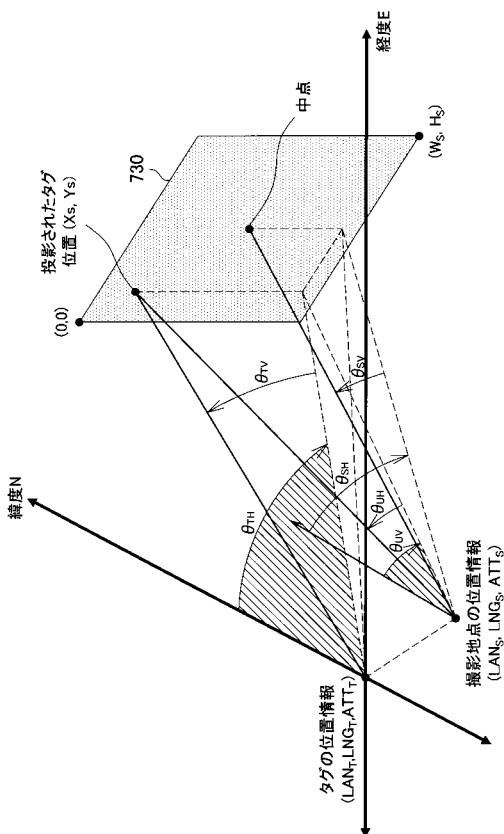
【 図 2 1 A 】



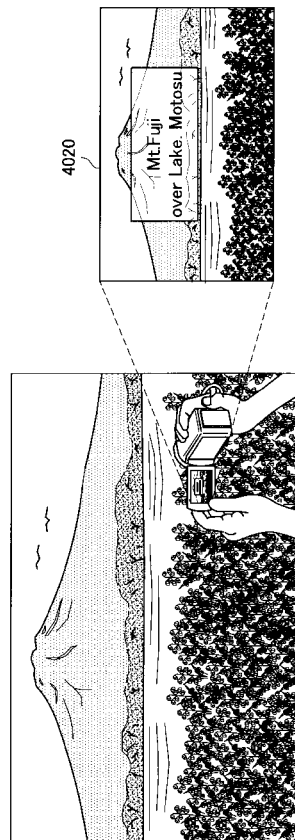
【 図 2 1 B 】



【 図 2 2 】



【 図 2 3 】



【図 24】

