

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5654127号  
(P5654127)

(45) 発行日 平成27年1月14日(2015. 1. 14)

(24) 登録日 平成26年11月28日(2014. 11. 28)

(51) Int.Cl.

F I

G 0 6 T 7/00 (2006.01)

G 0 6 T 7/00 3 0 0 B

G 0 6 T 7/00 3 0 0 F

請求項の数 50 (全 48 頁)

(21) 出願番号	特願2013-523231 (P2013-523231)	(73) 特許権者	595020643
(86) (22) 出願日	平成23年7月29日 (2011. 7. 29)		クゥアルコム・インコーポレイテッド
(65) 公表番号	特表2013-534342 (P2013-534342A)		Q U A L C O M M I N C O R P O R A T E D
(43) 公表日	平成25年9月2日 (2013. 9. 2)		アメリカ合衆国、カリフォルニア州 9 2
(86) 国際出願番号	PCT/US2011/045942		1 2 1 - 1 7 1 4、サン・ディエゴ、モア
(87) 国際公開番号	W02012/016168		ハウス・ドライブ 5 7 7 5
(87) 国際公開日	平成24年2月2日 (2012. 2. 2)	(74) 代理人	100108855
審査請求日	平成25年4月1日 (2013. 4. 1)		弁理士 蔵田 昌俊
(31) 優先権主張番号	13/193, 294	(74) 代理人	100109830
(32) 優先日	平成23年7月28日 (2011. 7. 28)		弁理士 福原 淑弘
(33) 優先権主張国	米国 (US)	(74) 代理人	100088683
(31) 優先権主張番号	61/369, 228		弁理士 中村 誠
(32) 優先日	平成22年7月30日 (2010. 7. 30)	(74) 代理人	100103034
(33) 優先権主張国	米国 (US)		弁理士 野河 信久

最終頁に続く

(54) 【発明の名称】 インクリメントな特徴抽出を使用するオブジェクト認識

(57) 【特許請求の範囲】

【請求項 1】

画像中のオブジェクトに対するアイデンティティを決定する方法において、  
受信した画像の第1のオクターブのぼかし画像の第1のセットから、1つ以上のキーポイントの第1のセットを抽出することと、

前記キーポイントの第1のセットに対する1つ以上の記述子の第1のセットを計算することと、

前記記述子の第1のセットにより特徴記述子データベースに問い合わせることにより生じた結果に対する信頼値を受信し、前記結果は、前記受信した画像中のオブジェクトのアイデンティティを記述する情報を含むことと、

前記信頼値が信頼しきい値を超えないときに、前記受信した画像の第2のオクターブのぼかし画像の第2のセットから、1つ以上のキーポイントの第2のセットを抽出することを含む方法。

【請求項 2】

前記信頼値が前記信頼しきい値を超えていないとの判定に応じて、前記判定に基づいて、前記第2のオクターブのぼかし画像の第2のセットから、前記キーポイントの第2のセットを抽出することと、前記キーポイントの第2のセットから、1つ以上の記述子の第2のセットを計算することと、前記記述子の第1のセットおよび前記記述子の第2のセットを含む記述子のセットにより、前記特徴記述子データベースに問い合わせることとをさらに含む請求項1記載の方法。

**【請求項 3】**

前記受信した画像中の前記オブジェクトに対するスケール値を推定することと、  
前記推定したスケールに基づいて、前記特徴記述子データベースの最低オクターブを選択することと、

前記特徴記述子データベース中の前記選択した最低オクターブにおいて、または、前記特徴記述子データベース中の前記選択した最低オクターブより上で、特徴記述子をサーチするための前記特徴記述子データベースのサーチを起こすように、前記特徴記述子データベース中の最低オクターブの指示を提供することとをさらに含む請求項 1 記載の方法。

**【請求項 4】**

前記スケール値を推定することは、  
スケール空間にわたる前記キーポイントの第 1 のセットの分布を解析することと、  
前記キーポイントのうちのおおよそ 90 % がスケールレベルより下に属するように、前記分布中の前記スケールレベルに対応するカットオフポイントを決することと、  
前記決定したカットオフポイントの関数として、前記スケール値を推定することを含む請求項 3 記載の方法。

**【請求項 5】**

前記オブジェクトに対する深度情報を取得することをさらに含み、  
前記スケール値を推定することは、前記オブジェクトに対する深度情報の少なくとも一部に基づいて、前記スケール値を推定することを含む請求項 3 記載の方法。

**【請求項 6】**

前記受信した画像を取り込んだカメラに係する 1 つ以上のセンサからのセンサ情報を解析することと、  
前記センサ情報の解析に基づいて、前記第 1 のオクターブおよび前記第 2 のオクターブのうち少なくとも 1 つに対するオクターブレベルを決することとをさらに含む請求項 1 記載の方法。

**【請求項 7】**

前記センサ情報を解析することは、グローバルポジショニングシステム (GPS) 情報を解析して、前記受信した画像が取り込まれたときに前記カメラが屋外環境に位置付けられていたか否かを決定することを含む請求項 6 記載の方法。

**【請求項 8】**

前記センサ情報を解析することは、GPS 情報を解析して、前記受信した画像が取り込まれたときの、前記カメラに比較的近いオブジェクトのロケーションを決することと、  
前記オブジェクトに対する記述データから、前記オブジェクトに対するサイズを決することを含む請求項 6 記載の方法。

**【請求項 9】**

前記センサ情報を解析することは、前記受信した画像が取り込まれたときに、前記カメラを備えているデバイスがワイヤレスネットワークに通信可能に結合されていたか否かを示すネットワークデータに基づいて、前記カメラが屋内環境に位置付けられていたか否かを決定することを含む請求項 6 記載の方法。

**【請求項 10】**

前記センサ情報を解析することは、前記受信した画像が取り込まれたときの、前記受信した画像中の 1 つ以上のオブジェクトと前記カメラとの間の距離を示す深度情報を計算することを含む請求項 6 記載の方法。

**【請求項 11】**

前記センサ情報を解析することは、アクティブプロービングセンサにより提供されるデータを使用して、前記オブジェクトに対する深度値を推定することを含む請求項 6 記載の方法。

**【請求項 12】**

前記 1 つ以上の記述子をサーバに送って、前記サーバに、前記 1 つ以上の記述子を使用して前記特徴記述子データベースに問い合わせさせることをさらに含み、

10

20

30

40

50

前記信頼値を受信することは、前記問い合わせに応答して前記サーバから前記信頼値を受信することを含む請求項 1 記載の方法。

【請求項 1 3】

画像中のオブジェクトに対するアイデンティティを決定するための装置において、  
プロセッサを具備し、  
前記プロセッサは、  
受信した画像の第 1 のオクターブのぼかし画像の第 1 のセットから、1 つ以上のキーポイントの第 1 のセットを抽出し、

前記キーポイントの第 1 のセットに対する 1 つ以上の記述子の第 1 のセットを計算し、  
前記記述子の第 1 のセットにより特徴記述子データベースに問い合わせることにより生じた結果に対する信頼値を受信し、前記結果は、前記受信した画像中のオブジェクトのアイデンティティを記述する情報を含み、

前記信頼値が信頼しきい値を超えないときに、前記受信した画像の第 2 のオクターブのぼかし画像の第 2 のセットから、1 つ以上のキーポイントの第 2 のセットを抽出するように構成されている装置。

【請求項 1 4】

前記画像を取り込み、前記画像を前記プロセッサに提供するように構成されているカメラをさらに具備する請求項 1 3 記載の装置。

【請求項 1 5】

前記プロセッサが、前記信頼値が前記信頼しきい値を超えないと判定するときに、前記プロセッサは、

前記判定に基づいて、前記第 2 のオクターブのぼかし画像の第 2 のセットから、前記キーポイントの第 2 のセットを抽出し、

前記キーポイントの第 2 のセットから、1 つ以上の記述子の第 2 のセットを計算し、  
前記記述子の第 1 のセットおよび前記記述子の第 2 のセットを含む記述子のセットにより、前記特徴記述子データベースに問い合わせるようにさらに構成されている請求項 1 3 記載の装置。

【請求項 1 6】

前記プロセッサは、  
前記受信した画像中の前記オブジェクトに対するスケール値を推定し、  
前記推定したスケールに基づいて、前記特徴記述子データベースの最低オクターブを選択し、

前記特徴記述子データベース中の記述子をサーチするようにさらに構成され、  
前記サーチした記述子は、前記選択した最低オクターブにおける、または、前記選択した最低オクターブより上の、前記特徴記述子データベース中のオクターブに対応する請求項 1 5 記載の装置。

【請求項 1 7】

前記スケール値を推定するために、前記プロセッサは、  
スケール空間にわたる前記キーポイントの第 1 のセットの分布を解析し、  
前記キーポイントのうちのおおよそ 90 % がスケールレベルより下に属するように、前記分布中の前記スケールレベルに対応するカットオフポイントを決し、  
前記決定したカットオフポイントの関数として、前記スケール値を計算するように構成されている請求項 1 6 記載の装置。

【請求項 1 8】

前記オブジェクトに対する深度情報を取得するように構成されている深度推定ユニットをさらに具備し、

前記プロセッサは、前記オブジェクトに対する深度情報の少なくとも一部に基づいて、前記スケール値を推定するように構成されている請求項 1 6 記載の装置。

【請求項 1 9】

前記プロセッサは、

10

20

30

40

50

前記受信した画像を取り込んだカメラに係する１つ以上のセンサからのセンサ情報を解析し、

前記センサ情報の解析に基づいて、前記第１のオクターブおよび前記第２のオクターブのうちの少なくとも１つに対するオクターブレベルを決定するようにさらに構成されている請求項１３記載の装置。

【請求項２０】

前記装置に対するロケーション情報を決定するように構成されているグローバルポジショニングシステム（ＧＰＳ）ユニットをさらに具備し、

前記センサ情報を解析するために、前記プロセッサは、前記ＧＰＳユニットにより決定されたロケーション情報に基づいて、前記受信した画像が取り込まれたときに前記カメラが屋外環境に位置付けられていたか否かを決定するように構成されている請求項１９記載の装置。

10

【請求項２１】

前記装置に関連するロケーション情報を決定するように構成されているＧＰＳユニットをさらに具備し、

前記センサ情報を解析するために、前記プロセッサは、前記ＧＰＳユニットにより決定されたロケーション情報を解析して、前記受信した画像が取り込まれたときの、前記カメラに比較的近いオブジェクトのロケーションを決定し、前記オブジェクトに対する記述データから、前記オブジェクトに対するサイズを決定するように構成されている請求項１９記載の装置。

20

【請求項２２】

ワイヤレスネットワークインターフェースをさらに具備し、

前記センサ情報を解析するために、前記プロセッサは、前記受信した画像が取り込まれたときに、前記ワイヤレスネットワークインターフェースがワイヤレスネットワークに通信可能に結合されていたか否かを示すネットワークデータに基づいて、前記カメラが屋内環境に位置付けられていたか否かを決定するように構成されている請求項１９記載の装置。

【請求項２３】

前記受信した画像を取り込んだカメラを含む少なくとも２つのカメラを備えるカメラアレイをさらに具備し、

30

前記センサ情報を解析するために、前記プロセッサは、前記受信した画像が取り込まれたときの、前記受信した画像中の１つ以上のオブジェクトと前記カメラとの間の距離を示す、前記カメラアレイにより取り込まれた画像からの深度情報を計算するように構成されている請求項１９記載の装置。

【請求項２４】

前記１つ以上のセンサは、前記オブジェクトに対する深度値を推定するように構成されているアクティブプロービングセンサを備える請求項１９記載の装置。

【請求項２５】

前記プロセッサは、前記１つ以上の記述子をサーバに送って、前記サーバに、前記１つ以上の記述子を使用して前記特徴記述子データベースに問い合わせさせ、前記問い合わせに回答して前記サーバから前記信頼値を受信するように構成されている請求項１３記載の装置。

40

【請求項２６】

前記装置は、

集積回路と、

マイクロプロセッサと、

前記プロセッサを含むワイヤレス通信デバイスとのうちの少なくとも１つを具備する請求項１３記載の装置。

【請求項２７】

画像中のオブジェクトに対するアイデンティティを決定する装置において、

50

受信した画像の第1のオクターブのぼかし画像の第1のセットから、1つ以上のキーポイントの第1のセットを抽出する手段と、

前記キーポイントの第1のセットに対する1つ以上の記述子の第1のセットを計算する手段と、

前記記述子の第1のセットにより特徴記述子データベースに問い合わせることにより生じた結果に対する信頼値を受信し、前記結果は、前記受信した画像中のオブジェクトのアイデンティティを記述する情報を含む手段と、

前記信頼値が信頼しきい値を超えないときに、前記受信した画像の第2のオクターブのぼかし画像の第2のセットから、1つ以上のキーポイントの第2のセットを抽出する手段とを具備する装置。

10

【請求項28】

前記信頼値が前記信頼しきい値を超えていないと判定する手段と、

前記判定に基づいて、前記第2のオクターブのぼかし画像の第2のセットから、前記キーポイントの第2のセットを抽出する手段と、

前記キーポイントの第2のセットから、1つ以上の記述子の第2のセットを計算する手段と、

前記記述子の第1のセットおよび前記記述子の第2のセットを含む記述子のセットにより、前記特徴記述子データベースに問い合わせる手段とをさらに具備する請求項27記載の装置。

【請求項29】

20

前記受信した画像中の前記オブジェクトに対するスケール値を推定する手段と、

前記推定したスケールに基づいて、前記特徴記述子データベースの最低オクターブを選択する手段と、

前記特徴記述子データベース中の前記選択した最低オクターブにおいて、または、前記特徴記述子データベース中の前記選択した最低オクターブより上で、特徴記述子をサーチするための前記特徴記述子データベースのサーチを起こすように、前記特徴記述子データベース中の最低オクターブの指示を提供する手段とをさらに具備する請求項28記載の装置。

【請求項30】

前記スケール値を推定する手段は、

スケール空間にわたる前記キーポイントの第1のセットの分布を解析する手段と、

前記キーポイントのうちのおおよそ90%がスケールレベルより下に属するように、前記分布中の前記スケールレベルに対応するカットオフポイントを決定する手段と、

前記決定したカットオフポイントの関数として、前記スケール値を推定する手段とを具備する請求項29記載の装置。

30

【請求項31】

前記オブジェクトに対する深度情報を取得する手段をさらに具備し、

前記スケール値を推定する手段は、前記オブジェクトに対する深度情報の少なくとも一部に基づいて、前記スケール値を推定する手段を備える請求項29記載の装置。

【請求項32】

40

前記受信した画像を取り込んだカメラに関係する1つ以上のセンサからのセンサ情報を解析する手段と、

前記センサ情報の解析に基づいて、前記第1のオクターブおよび前記第2のオクターブのうち少なくとも1つに対するオクターブレベルを決定する手段とをさらに具備する請求項27記載の装置。

【請求項33】

グローバルポジショニングシステム(GPS)情報を受信する手段をさらに具備し、

前記センサ情報を解析する手段は、前記GPS情報を解析して、前記受信した画像が取り込まれたときに前記カメラが屋外環境に位置付けられていたか否かを決定する手段を備える請求項32記載の装置。

50

## 【請求項 3 4】

G P S 情報を受信する手段をさらに具備し、  
前記センサ情報を解析する手段は、  
前記 G P S 情報を解析して、前記受信した画像が取り込まれたときの、前記カメラに比較的近いオブジェクトのロケーションを決定する手段と、  
前記オブジェクトに対する記述データから、前記オブジェクトに対するサイズを決定する手段とを備える請求項 3 2 記載の装置。

## 【請求項 3 5】

少なくとも 1 つのワイヤレスネットワークプロトコルを介して通信する手段をさらに具備し、

10

前記センサ情報を解析する手段は、前記受信した画像が取り込まれたときに、前記ワイヤレスネットワークプロトコルを介して通信する手段がワイヤレスネットワークに通信可能に結合されていたか否かを示すネットワークデータに基づいて、前記カメラが屋内環境に位置付けられていたか否かを決定する手段を備える請求項 3 2 記載の装置。

## 【請求項 3 6】

シーンのうちの 2 つ以上の画像を取り込む手段をさらに具備し、  
前記 2 つ以上の画像のうちの 1 つは、前記受信した画像を含み、  
前記センサ情報を解析する手段は、前記受信した画像が取り込まれたときの、前記受信した画像中の 1 つ以上のオブジェクトと前記カメラとの間の距離を示す、前記シーンのうちの 2 つの画像からの深度情報を計算する手段を備える請求項 3 2 記載の装置。

20

## 【請求項 3 7】

前記センサ情報を解析する手段は、アクティブプロービングセンサにより提供されるデータを使用して、前記オブジェクトに対する深度値を推定する手段を備える請求項 3 2 記載の装置。

## 【請求項 3 8】

前記 1 つ以上の記述子をサーバに送って、前記サーバに、前記 1 つ以上の記述子を使用して前記特徴記述子データベースに問い合わせさせる手段をさらに具備し、

前記信頼値を受信する手段は、前記問い合わせに回答して前記サーバから前記信頼値を受信する手段を備える請求項 2 7 記載の装置。

## 【請求項 3 9】

30

その上に記憶されている命令を有するコンピュータ読取可能記録媒体において、  
前記命令は、実行されるとき、  
受信した画像の第 1 のオクターブのぼかし画像の第 1 のセットから、1 つ以上のキーポイントの第 1 のセットをプロセッサに抽出させ、

前記キーポイントの第 1 のセットに対する 1 つ以上の記述子の第 1 のセットを前記プロセッサに計算させ、

前記記述子の第 1 のセットにより特徴記述子データベースに問い合わせることにより生じた結果に対する信頼値を前記プロセッサに受信させ、前記結果は、前記受信した画像中のオブジェクトのアイデンティティを記述する情報を含み、

前記信頼値が信頼しきい値を超えないときに、前記受信した画像の第 2 のオクターブのぼかし画像の第 2 のセットから、1 つ以上のキーポイントの第 2 のセットを前記プロセッサに抽出させるコンピュータ読取可能記録媒体。

40

## 【請求項 4 0】

前記信頼値が前記信頼しきい値を超えていないことを前記プロセッサに判定させ、  
前記判定に基づいて、前記第 2 のオクターブのぼかし画像の第 2 のセットから、前記キーポイントの第 2 のセットを前記プロセッサに抽出させ、

前記キーポイントの第 2 のセットから、1 つ以上の記述子の第 2 のセットを前記プロセッサに計算させ、

前記記述子の第 1 のセットおよび前記記述子の第 2 のセットを含む記述子のセットにより、前記特徴記述子データベースに前記プロセッサに問い合わせさせる命令をさらに含む

50

請求項 3 9 記載のコンピュータ読取可能記録媒体。

【請求項 4 1】

前記受信した画像中の前記オブジェクトに対するスケール値を前記プロセッサに推定させ、

前記推定したスケールに基づいて、前記特徴記述子データベースの最低オクターブを前記プロセッサに選択させ、

前記特徴記述子データベース中の前記選択した最低オクターブにおいて、または、前記特徴記述子データベース中の前記選択した最低オクターブより上で、特徴記述子をサーチするための前記特徴記述子データベースのサーチを起こすように、前記特徴記述子データベース中の最低オクターブの指示を前記プロセッサに提供させる命令をさらに含む請求項 3 9 記載のコンピュータ読取可能記録媒体。

10

【請求項 4 2】

前記スケール値を前記プロセッサに推定させる命令は、

スケール空間にわたる前記キーポイントの第 1 のセットの分布を前記プロセッサに解析させ、

前記キーポイントのうちのおおよそ 9 0 % がスケールレベルより下に属するように、前記分布中の前記スケールレベルに対応するカットオフポイントを前記プロセッサに決定させ、

前記決定したカットオフポイントの関数として、前記スケール値を前記プロセッサに推定させる請求項 4 1 記載のコンピュータ読取可能記録媒体。

20

【請求項 4 3】

前記オブジェクトに対する深度情報を前記プロセッサに取得させる命令をさらに含み、

前記スケール値を前記プロセッサに推定させる命令は、前記オブジェクトに対する深度情報の少なくとも一部に基づいて、前記スケール値を前記プロセッサに推定させる命令を有する請求項 4 1 記載のコンピュータ読取可能記録媒体。

【請求項 4 4】

前記受信した画像を取り込んだカメラに関係する 1 つ以上のセンサからのセンサ情報を前記プロセッサに解析させ、

前記センサ情報の解析に基づいて、前記第 1 のオクターブおよび前記第 2 のオクターブのうち少なくとも 1 つに対するオクターブレベルを前記プロセッサに決定させる命令をさらに含む請求項 3 9 記載のコンピュータ読取可能記録媒体。

30

【請求項 4 5】

前記センサ情報を前記プロセッサに解析させる命令は、グローバルポジショニングシステム (GPS) 情報を前記プロセッサに解析させ、前記受信した画像が取り込まれたときに前記カメラが屋外環境に位置付けられていたか否かを決定させる命令を含む請求項 4 4 記載のコンピュータ読取可能記録媒体。

【請求項 4 6】

前記センサ情報を前記プロセッサに解析させる命令は、GPS 情報を前記プロセッサに解析させ、前記受信した画像が取り込まれたときの、前記カメラに比較的近いオブジェクトのロケーションを決定させ、前記オブジェクトに対する記述データから、前記オブジェクトに対するサイズを前記プロセッサに決定させる命令を含む請求項 4 4 記載のコンピュータ読取可能記録媒体。

40

【請求項 4 7】

前記センサ情報を前記プロセッサに解析させる命令は、前記受信した画像が取り込まれたときに、前記カメラを備えているデバイスがワイヤレスネットワークに通信可能に結合されていたか否かを示すネットワークデータに基づいて、前記カメラが屋内環境に位置付けられていたか否かを前記プロセッサに決定させる命令を含む請求項 4 4 記載のコンピュータ読取可能記録媒体。

【請求項 4 8】

前記センサ情報を前記プロセッサに解析させる命令は、前記受信した画像が取り込まれ

50

たときの、前記受信した画像中の１つ以上のオブジェクトと前記カメラとの間の距離を示す深度情報を前記プロセッサに計算させる命令を含む請求項４４記載のコンピュータ読取可能記録媒体。

【請求項４９】

前記センサ情報を前記プロセッサに解析させる命令は、アクティブブローピングセンサにより提供されるデータを使用して、前記オブジェクトに対する深度値を前記プロセッサに推定させる命令を含む請求項４４記載のコンピュータ読取可能記録媒体。

【請求項５０】

前記１つ以上の記述子をサーバに送って、前記サーバに、前記１つ以上の記述子を使用して前記特徴記述子データベースに問い合わせさせるように前記プロセッサにさせる命令をさらに含み、

10

前記信頼値を前記プロセッサに受信させる命令は、前記問い合わせに応答して前記サーバから前記信頼値を前記プロセッサに受信させる命令を含む請求項３９記載のコンピュータ読取可能記録媒体。

【発明の詳細な説明】

【技術分野】

【０００１】

本開示は、画像処理システムに関連し、さらに詳細には、画像中のオブジェクト認識に関連する。

【背景技術】

20

【０００２】

コンピューティングデバイスまたはコンピュータのコンテキスト中の視覚探索は、コンピュータまたは他のデバイスが、１つ以上の画像中のオブジェクトに対するアイデンティティ情報を提供することを可能にする技術のことを指す。視覚サーチにおける最近の関心は、部分的に隠蔽されたオブジェクトを、および／または、画像スケール、ノイズ、照度および局所的な幾何学的な歪みにおける変化を含む、幅広いさまざまな可変画像条件中の特徴を、コンピュータが識別することを可能にするアルゴリズムを結果として生じている。これと同期間中に、カメラを備える移動デバイスが登場しているが、このような移動デバイスは、テキストを入力するための限定されたユーザインターフェース、または、そうでなければ、他の移動デバイスとのインターフェースを有しているかもしれない。移動デバイスおよび移動デバイスアプリケーションの開発者は、移動デバイスとのユーザ対話を強化するために、移動デバイスのカメラを利用しようとしている。

30

【０００３】

１つの強化を説明すると、移動デバイスのユーザは、店舗で買い物をする間に、移動デバイスのカメラを使用して、何らかの所定の製品の画像を取り込んでもよい。移動デバイスはその後、さまざまな画像に対するアーカイブの特徴記述子のセット内で視覚サーチアルゴリズムを開始させ、マッチング画像に基づいて製品を識別してもよい。製品を識別した後に、移動デバイスはその後、インターネットベースのサーチを開始し、識別した製品についての情報を含むウェブページを提示する。ウェブページは、例えば、近くの売買業者および／またはオンライン売買業者から製品が入手可能な最低価格を含む。別の例では、オブジェクト認識は、拡張現実感と呼ばれるものを達成するように、移動デバイスのディスプレイ中にオーバーレイできる補足的な情報を発生させるために使用してもよい。

40

【概要】

【０００４】

一般的に、本開示は、デジタル画像中のオブジェクト識別に対する技術を記述している。これらの技術は、画像中のオブジェクトを識別する計算の複雑性を低減できる。予め定められている数のオクターブから特徴記述子を抽出することよりむしろ、本開示の技術は、オクターブを通して進行する、インクリメントな特徴記述子抽出を提供する。さらに、本開示は、開始オクターブを選択する技術と、特徴記述子を抽出する後続するオクターブを選択する技術とを提供する。

50

## 【0005】

1つの例では、方法は、受信した画像の第1のオクターブのぼかし画像の第1のセットから、1つ以上のキーポイントの第1のセットを抽出することと、キーポイントの第1のセットに対する1つ以上の記述子の第1のセットを計算することと、記述子の第1のセットにより特徴記述子データベースに問い合わせることにより生じた結果に対する信頼値を受信し、結果は、受信した画像中のオブジェクトのアイデンティティを記述する情報を含むことと、信頼値が信頼しきい値を超えないときに、受信した画像の第2のオクターブのぼかし画像の第2のセットから、1つ以上のキーポイントの第2のセットを抽出することを含む。

## 【0006】

10

別の例では、装置はプロセッサを具備し、プロセッサは、受信した画像の第1のオクターブのぼかし画像の第1のセットから、1つ以上のキーポイントの第1のセットを抽出し、キーポイントの第1のセットに対する1つ以上の記述子の第1のセットを計算し、記述子の第1のセットにより特徴記述子データベースに問い合わせることにより生じた結果に対する信頼値を受信し、結果は、受信した画像中のオブジェクトのアイデンティティを記述する情報を含み、信頼値が信頼しきい値を超えないときに、受信した画像の第2のオクターブのぼかし画像の第2のセットから、1つ以上のキーポイントの第2のセットを抽出するように構成されている。

## 【0007】

20

別の例では、装置は、受信した画像の第1のオクターブのぼかし画像の第1のセットから、1つ以上のキーポイントの第1のセットを抽出する手段と、キーポイントの第1のセットに対する1つ以上の記述子の第1のセットを計算する手段と、記述子の第1のセットにより特徴記述子データベースに問い合わせることにより生じた結果に対する信頼値を受信し、結果は、受信した画像中のオブジェクトのアイデンティティを記述する情報を含む手段と、信頼値が信頼しきい値を超えないときに、受信した画像の第2のオクターブのぼかし画像の第2のセットから、1つ以上のキーポイントの第2のセットを抽出する手段とを具備する。

## 【0008】

30

別の例では、コンピュータプログラムプロダクトは、その上に記憶されている命令を有するコンピュータ読取可能媒体を含む。命令は、実行されるとき、受信した画像の第1のオクターブのぼかし画像の第1のセットから、1つ以上のキーポイントの第1のセットをプロセッサに抽出させ、キーポイントの第1のセットに対する1つ以上の記述子の第1のセットをプロセッサに計算させ、記述子の第1のセットにより特徴記述子データベースに問い合わせることにより生じた結果に対する信頼値をプロセッサに受信させ、結果は、受信した画像中のオブジェクトのアイデンティティを記述する情報を含み、信頼値が信頼しきい値を超えないときに、受信した画像の第2のオクターブのぼかし画像の第2のセットから、1つ以上のキーポイントの第2のセットをプロセッサに抽出させる。

## 【図面の簡単な説明】

## 【0009】

40

【図1】図1は、インクリメントな特徴記述子抽出に対する技術が適用される例示的なシステムを図示しているブロックダイヤグラムである。

【図2】図2は、キーポイント抽出において使用するために決定されているガウシアン差分(DoG)ピラミッドを図示している概念図である。

【図3】図3は、キーポイントの検出をより詳細に図示している概念図である。

【図4】図4は、特徴抽出ユニットが勾配分布および方向ヒストグラムを取得できる例示的な技術を図示している概念図である。

【図5】図5は、さまざまなガウシアンスケール空間にわたる例示的なキーポイント分布を図示しているチャートの線図である。

【図6】図6は、インクリメントな特徴記述子抽出を実行するための例示的な方法を図示しているフローチャートである。

50

## 【詳細な説明】

## 【0010】

一般的に、本開示は、デジタル画像中のオブジェクト識別に対する技術を記述している。これらの技術は、画像中のオブジェクトを識別する計算の複雑性を低減できる。予め定められている数のオクターブから特徴記述子を抽出することよりむしろ、本開示の技術は、オクターブを通して進行する、インクリメントな特徴記述子抽出を提供する。特徴記述子は、一般的に1つ以上のオクターブから抽出され、各オクターブは、画像の特定のサイズ（例えば、ピクセル解像度）に対応する。所定のオクターブに対して、入力画像は、解像度に関係するサイズにスケーリングされてもよい。さらに、次第にガウシアンぼけする複数の画像が形成されてもよく、各々は、オクターブに関係する解像度を有する。オクターブに対して次第にぼけるこれらの画像は、特徴記述子を抽出するために解析されてもよい。

10

## 【0011】

予め定められている数のオクターブに対する特徴記述子を抽出することよりむしろ、本開示の技術は、第1のオクターブに対する特徴記述子を抽出して、この特徴記述子の第1のセットに基づいて、画像中のオブジェクトを検出するように試行することに向けられている。これらの特徴記述子を使用してオブジェクトを検出できる場合に、それ以上のオクターブが解析される必要はない。しかしながら、第1のオクターブからオブジェクトを決定できない場合に、後続するオクターブから特徴記述子が検出されてもよい。

20

## 【0012】

さらに、本開示は、開始オクターブを選択する技術と、特徴記述子を抽出する後続するオクターブを選択する技術とを提供する。さまざまなファクタを使用して、開始オクターブを選択してもよい。例えば、画像に対する倍率（スケーリングファクタ）は初めに、例えば、画像が屋内または屋外で取り込まれたか、および/または、その画像を取り込むために使用したカメラに対する、画像中のオブジェクトの深度（すなわち、カメラから画像中のオブジェクトまでの距離）に基づいて、推定されてもよい。

## 【0013】

カメラを含むデバイス、例えば、スマートフォン、タブレットコンピューティングデバイス、ラップトップコンピュータまたは他の移動デバイスのような、移動デバイスは、画像が取り込まれたときに、デバイスが屋内または屋外にあったかを推定するために使用できるセンサを備えていてもよい。例えば、グローバルポジショニングシステム（GPS）ユニットは、画像が取り込まれたときのカメラのロケーションを示す情報を提供してもよい。このロケーションは、カメラが屋内または屋外にあったかを決定するために使用されてもよい。別の例として、デバイスが、画像が取り込まれたときにワイヤレスアクセスポイントに接続されていたワイヤレスネットワークインターフェースを備えていた場合、このことは、カメラが取り込まれたときに、デバイスが屋内にあったか否かを推定するためのファクタとして使用されてもよい。さらに別の例として、画像中のオブジェクトに対する深度情報は、例えば、デバイスがステレオカメラ配置を有するときに決定されてもよい。

30

## 【0014】

いくつかの例では、デバイスは、ステレオカメラ配置を使用して実行される深度推定の他に、または、ステレオカメラ配置を使用して実行される深度推定の代わりに、画像中のオブジェクトに対する深度を推定するように構成されている深度推定ユニットを備えていてもよい。例えば、デバイスは、シーン中のオブジェクトの相対的な深度を推定するために、1つ以上の赤外線センサから測定データを収集する深度推定ユニットを備えていてもよい。別の例として、デバイスは、画像中のオブジェクトの深度を推定する、光検出および測距（LIDAR）ユニットのような、アクティブプロービングセンサを備えていてもよい。LIDARユニットは、オブジェクトを投影するために紫外線、可視光線または赤外線を使用して、戻ってきた信号を処理し、オブジェクトに対する深度を推定してもよい。これらの深度推定ユニットは、画像中のオブジェクトの深度を推定するために、単独で

40

50

、または、互いに任意の組み合わせで使用されてもよい。深度推定ユニットは、ユニットが、カメラからの画像データを実際に処理することなく、カメラによって取り込まれた画像中のオブジェクトに対する深度を決定できるように、画像を取り込むカメラと実質的に同時に動作してもよい。

【0015】

デバイスは、推定したオブジェクトスケールとともに、例えば、画像、および、デバイスの追加のセンサに基づいて、さまざまなファクタに関係付けるコンフィグレーションデータを備えていてもよい。例えば、コンフィグレーションデータは、特定の開始オクターブおよびスケールをオブジェクトに対するさまざまなロケーションセル（または、“ロクセル（loxel）”）（すなわち、オブジェクトが生じるロクセル）にマッピングしてもよい。例として、使用量統計は、所定のロクセル中でオブジェクトが生じるときに、開始スケールおよび／またはオクターブとともに、後続するスケールおよび／またはオクターブを決定するために、最も可能性の高いスケールおよび／またはオクターブに対する推定を提供してもよい。言い換えると、いくつかの例では、本開示の技術は、オブジェクトが生じるロクセルに基づいて、オブジェクトに対する開始オクターブおよび／またはスケール（とともに、後続するオクターブ）を選択してもよい。

【0016】

いくつかの例では、キーポイント分布を使用して、オブジェクトに対するスケールを推定してもよい。キーポイントは、一般的に、特徴記述子を生成させるために使用できる画像のエリアである。一般的に、比較的大きなスケールにおいて検出されたキーポイントは、オブジェクトの比較的大きな特徴に対応する。同様に、画像解像度が減少するにつれて、より細かいディテールを認識することがより困難になる。それゆえ、より大きな解像度に対してより大きなスケールで検出されるディテールは、典型的に、より小さな解像度に対してより低いスケールで検出される（したがって、より細かいディテールになる）。スケール空間にわたるキーポイントの分布を示す統計を収集することができ、統計は、画像の倍率を推定するために使用されてもよい。前のオクターブに対して検出されたキーポイントを使用して、制御ユニットは、問い合わせ（クエリ）画像のオブジェクトに対するスケールを推定し、推定したスケールを使用して、サーチする特徴記述子のデータベースの最低オクターブを選択してもよい。すなわち、制御ユニットは、最低オクターブ以上であって、最低オクターブよりも低くないオクターブにおける特徴記述子のみをデータベースに問い合わせさせてもよい。

【0017】

いくつかのケースでは、デバイスのカメラは継続的に画像を取り込んでもよく、プロセッサ、または、オブジェクト認識のための他のユニットは、取り込んだ画像のすべて、または、取り込んだ画像のうちのサブセットの中でオブジェクトを検出しようと試行してもよい。いくつかのケースでは、プロセッサは、先の画像のうちの1つの中でオブジェクトスケールの近似に到達してもよい。本開示の技術は、例えば、ユーザが移動して、画像トラッカーが前もって獲得したターゲットを見失った後に、前のピクチャに対する前もって決定されたオブジェクトスケール情報を使用して、現在の画像に対するオブジェクトスケールの推定を初期化してもよい。

【0018】

一般的に、オブジェクト認識に対して、特徴記述子のセットおよび他の差別的な情報のデータベースが、トレーニング画像から導出される。その後、特徴記述子が、ターゲット画像から抽出され、データベースに問い合わせ、所定の問い合わせ画像のコンテンツを評価するために使用される。拡張現実感または視覚探索アプリケーションに対して、クライアント（例えば、セル電話機）は、対象のオブジェクトの画像を取り込み、画像、特徴およびメタデータ情報のデータベースとそれを比較する。このデータベースは、ネットワーク上のサーバ上に記憶されることができ、局所的な処理のために、または、代替的に、クライアントによって取り出すことができ、問い合わせは、サーバに送信されて、ネットワークリソースを使用して処理されることができ、本開示の技術は、一般的に、キーポ

イントのローカリゼーションおよび特徴記述子の抽出を実行するスケール不変特徴変換 (SIFT) アルゴリズムに関して記述されている。各 SIFT 特徴は以下の属性を有しているてもよい: 2D (x, y) 特徴ロケーション、特徴が検出されるスケール、特徴の近隣 (feature neighborhood) における最も強い画像勾配の特徴により与えられる特徴の回転方向、および、判別法 (discriminating way) で局所的なピクセル変化を、本質的には局所的な画像勾配のヒストグラムを記述するベクトル。

#### 【0019】

キーポイント識別および記述子抽出は、計算的に要求が厳しい。例えば、典型的なハーフサイズビデオグラフィックアレイ (HVG A) 画像は、数千の SIFT 特徴を生成させることができる。反復マッチングプロセスを含む SIFT アルゴリズムの実現は、いくつかの移動体プラットフォーム上で利用可能な計算のリソースを上回ることになる。しかしながら、最も現実的なケースでは、データベース中の特徴記述子とのマッチにつながる実際の特徴記述子の数は、実際に計算される特徴記述子の数よりもはるかに少ない傾向がある。つまり、計算された特徴記述子のセットのうちの、小さなサブセットがオブジェクト識別結果をもたらしてもよい。これは、最終的に少数のみがデータベース中の特徴とマッチするように特徴記述子に影響する、問い合わせ画像中の異なる画像化条件 (例えば、照度、遠近感等) によっていくぶん引き起こされるかもしれない。本開示の技術は、スケール空間中のキーポイント間での相互関係を活用し、記述子抽出およびマッチングプロセスにより課せられる計算的な負荷を減少させるように支援してもよい。また、所定のシーン内に複数のオブジェクトを認識/追跡するとき、偽陽性は、特にいくつかのオブジェクトが類似している場合に問題を引き起こすことになる。いくつかの例では、これらの技術は、偽陽性を低減させるためにマッチングを制約する、シーンの大まかで相対的な深度マップの計算を含んでいてもよい。

#### 【0020】

図1は、インクリメントな特徴記述子抽出に対する技術を適用できる例示的なシステム2を図示しているブロックダイアグラムである。この例では、システム2は、ネットワーク40を介して通信する、クライアントデバイス10およびサーバデバイス50を備える。クライアントデバイス10は、この例では、ラップトップ、ネットブックと呼ばれるもの、タブレットコンピュータ、パーソナルデジタルアシスタント (PDA)、(“スマートフォン”と呼ばれるものを含む) セルラもしくは携帯電話機もしくはハンドセット、グローバルポジショニングシステム (GPS) デバイス、デジタルカメラ、デジタルメディアプレイヤー、ゲームデバイス、または、サーバデバイス50と通信することが可能な他の任意の移動デバイスのような、移動デバイスを表す。本開示では、移動クライアントデバイス10に関して記述されているが、本開示中で記述されている技術は、この点に関して移動クライアントデバイスに限定されるべきではない。代わりに、技術はローカルオブジェクトデータベースを記憶することが可能な、および/または、ネットワーク40または他の任意の通信媒体を介してサーバデバイス50と通信することが可能な任意のデバイスにより実現されてもよい。

#### 【0021】

サーバデバイス50は、接続、例えば、伝送制御プロトコル (TCP) 接続を受け入れ、問い合わせデータを受信して、識別データを提供する TCP セッションを形成するために、それ自体の TCP 接続により応答するサーバデバイスを表す。サーバデバイス50は、サーバデバイス50が、画像内の1つ以上の特徴またはオブジェクトを識別する視覚探索アルゴリズムを実行、または、そうでなければ実現する、視覚探索サーバデバイスを表してもよい。

#### 【0022】

ネットワーク40は、クライアント10およびサーバデバイス50を相互接続する、インターネットのような、パブリックネットワークを表す。通常、ネットワーク40は、オープンシステム相互接続 (OSI) モデルのさまざまなレイヤを実現して、クライアントデバイス10とサーバデバイス50との間の通信またはデータの転送を促進する。ネット

ワーク４０は、典型的に、スイッチ、ハブ、ルータ、サーバのような、クライアントデバイス１０とサーバデバイス５０との間のデータの転送を可能にする任意の数のネットワークデバイスを含む。単一のネットワークとして示されているが、ネットワーク４０は、ネットワーク４０を形成するために相互接続されている１つ以上のサブネットワークを備えていてもよい。これらのサブネットワークは、ネットワーク４０全体を通してデータの転送を提供するための、サービスプロバイダネットワーク、アクセスネットワーク、バックエンドネットワーク、または、通常パブリックネットワーク中で用いられる他の任意のタイプのネットワークを含む。この例では、パブリックネットワークとして記述されているが、ネットワーク４０は、一般的に公共でアクセス可能ではないプライベートネットワークを含んでもよい。

10

#### 【００２３】

図１の例の中で示されているように、クライアントデバイス１０は、カメラ１２、センサ１４、制御ユニット２０、ローカルオブジェクトデータベース３０およびネットワークインターフェース３２を備える。制御ユニット２０は、この例では、特徴抽出ユニット２２、スケール推定ユニット２４、および、深度情報ユニット２６を備える。いくつかの例では、制御ユニット２０はまた、サーバデバイス５０の特徴マッチングユニット６２および信頼評価ユニット６４に類似するユニットを備える。例示の目的のために、これらのユニットは、サーバデバイス５０に関して論じられるが、類似したユニットがクライアントデバイス１０の制御ユニット２０内に含まれていてもよいことを理解すべきである。

#### 【００２４】

20

制御ユニット２０および制御ユニット６０ならびにこれらのサブユニットに帰する機能性は、ハードウェア、ソフトウェアおよび／またはファームウェアの任意の組み合わせにより実現されてもよい。制御ユニット２０に帰する１つ以上の機能が、ソフトウェアまたはファームウェアで実現されるとき、制御ユニット２０は、ソフトウェアに対する命令を記憶する１つ以上のコンピュータ読取可能記憶媒体とともに、ソフトウェアを実行する１つ以上のハードウェアベースの処理ユニットを備えていてもよい。処理ユニットは、１つ以上の汎用プロセッサを備えていてもよい。制御ユニット２０および制御ユニット６０は、１つ以上のデジタル信号プロセッサ（DSP）、汎用マイクロプロセッサ、特定用途向け集積回路（ASIC）、フィールドプログラマブル論理アレイ（FPGA）、または、他の同等な集積または離散論理回路のような、１つ以上のハードウェアユニットを追加的にまたは代替的に備えていてもよい。制御ユニット２０および制御ユニット６０に帰する機能性のうちのいくつかまたはすべては、それらの、任意のそれぞれのハードウェア、ソフトウェアまたはファームウェアにより実行されてもよい。したがって、図１の例では、制御ユニット２０および６０内に別々のユニットが図示されているが、実際には、これらのような別々のユニットは、任意の妥当な方法で組み合わせられてもよい。

30

#### 【００２５】

カメラ１２は、例えば、行および列で構成された、個々の画像センサエレメントの２次元アレイを含んでもよい。カメラ１２は、例えば、相補的金属酸化膜半導体（CMOS）センサのような、ソリッドステートセンサエレメントまたは他のセンサエレメントのアレイを備えていてもよい。カメラ１２内の画像センサエレメントは、シーンから光情報を取得し、シーンのデジタル画像を取り込むためにシーンにさらされている。デジタル画像は、ピクセル値の２次元行列を含んでもよく、ピクセル値の２次元行列は、例えば、赤 - 緑 - 青の値を有する、または、ピクセル輝度（強度もしくは明るさ）値ならびにピクセルクロミナンス（カラー）値を有する、特定のカラー空間中に規定されてもよい。

40

#### 【００２６】

センサ１４は、情報を制御ユニット２０に提供する１つ以上の環境センサを備える。制御ユニット２０は、この情報を使用して、カメラ１２により取り込まれた画像中のオブジェクトのスケールを予想してもよい。いくつかの例では、センサ１４は、グローバルポジショニングシステム（GPS）衛星と対話して、クライアントデバイス１０の地理的ロケーションを決定する、GPSユニットを備えていてもよい。GPSユニットはまた、画像

50

が取り込まれたときにカメラ 1 2 が向けられていた方位の指示を提供する。制御ユニット 2 0 は、このロケーションおよび方位情報を使用して、カメラ 1 2 の照準線上にある可能性がある潜在的なオブジェクトにサーチを限定してもよい。

【 0 0 2 7 】

例えば、遠隔オブジェクトデータベース 5 2 は、遠隔オブジェクトデータベース 5 2 のオブジェクトに対する地理的ロケーションを記憶してもよい。センサ 1 4 の GPS ユニットがカメラ 1 2 の地理的ロケーションおよび方位を決定するとき、制御ユニット 2 0 は、カメラ 1 2 の照準線上にある地理的ロケーションを有するデータのサブセットを遠隔オブジェクトデータベース 5 2 から取り出して、取り出したデータをローカルオブジェクトデータ 3 0 として記憶させる。特徴抽出ユニット 2 2 はその後、画像から特徴記述子を抽出して、抽出した特徴記述子によりローカルオブジェクトデータベース 3 0 に問い合わせてもよい。

10

【 0 0 2 8 】

別の例として、特徴抽出ユニット 2 2 は、地理的ロケーション方位を使用して、カメラ 1 2 の照準線上にあるオブジェクトの最初のスケールを予想してもよい。例えば、遠隔オブジェクトデータベース 5 2 は、センサ 1 4 の GPS ユニットによって決定されたクライアントデバイス 1 0 のロケーションに基づいて、クライアントデバイス 1 0 の近くの地理的ロケーションにおけるオブジェクト、および、カメラ 1 2 の照準線上にあるオブジェクトに対する最も可能性の高いスケールを記憶してもよい。特徴抽出ユニット 2 2 はその後、遠隔オブジェクトデータ 5 2 に問い合わせるために、この最初のスケールを使用して、特徴記述子を初めに抽出し、ネットワーク 4 0 を介してサーバデバイス 5 0 に問い合わせを送ってもよい。問い合わせは、いくつかの例では、特徴記述子に加えて、ロケーション情報を含んでもよい。

20

【 0 0 2 9 】

センサの他の例は、例えば、温度センサ、圧力センサ（例えば、気圧センサ）、光センサ、赤外線センサ、紫外線センサ、湿度センサ、二酸化炭素（CO<sub>2</sub>）センサ、あるいは、これらのような他の環境センサを含んでもよく、これらのうちのいくつかまたはすべては、それぞれの感知した値に対して値を提供してもよい。制御ユニット 2 0 は、屋内および屋外環境を示す、これらのセンサのうちのいくつかまたはすべてに対するデータ領域を含んでもよい。制御ユニット 2 0 は、センサのうちのいくつかまたはすべてに対する現在の値を対応するデータ領域と比較して、カメラ 1 2 が問い合わせ画像を取り込んだときに、クライアントデバイス 1 0 が屋内または屋外にあったかを決定してもよい。

30

【 0 0 3 0 】

特徴抽出ユニット 2 2 は、スケール不変特徴変換（SIFT）アルゴリズム、勾配の圧縮ヒストグラム（CHoG）アルゴリズム、または、他のキーポイントもしくは特徴記述子抽出アルゴリズムのような、特徴抽出アルゴリズムにしたがって、特徴抽出を実行するユニットを表す。一般的に、特徴抽出ユニット 2 2 は、カメラ 1 2、または、クライアントデバイス 1 0 内に備えられている他の画像取り込みデバイスを使用して局所的に取り込まれてもよい画像データについて動作する。代替的に、クライアントデバイス 1 0 は、例えば、ネットワーク 4 0 を介して、別のコンピューティングデバイスとのワイヤード接続を介して局所的に、または、接続の他の任意のワイヤードまたはワイヤレス形態を介して、画像データをダウンロードすることにより、この画像データ自体を取り込まずに、画像データを記憶させてもよい。

40

【 0 0 3 1 】

スケール推定ユニット 2 4 は、例えば、カメラ 1 2 により取り込まれた問い合わせ画像中のオブジェクトに対するスケールを推定するユニットを表す。例えば、スケール推定ユニット 2 4 は、画像に対するさまざまなスケールからの実際のキーポイント分布と比較して、（例えば、トレーニング画像のセットから導き出された）さまざまなスケールから予測されるキーポイントの分布に基づいて、問い合わせ画像中のオブジェクトに対するスケールを推定してもよい。特徴抽出ユニット 2 2 は、画像に対する前のオクターブに対して

50

のキーポイントを抽出し、スケール推定ユニット24にこれらのキーポイントを提供し、スケール推定ユニット24は、これらのキーポイントを使用し、コンフィグレーションデータを使用して、問い合わせ画像中のオブジェクトに対するスケールを予想してもよい。推定したスケールを使用して、スケール推定ユニット24は、サーチするデータベース（例えば、ローカルオブジェクトデータベース30および/または遠隔オブジェクトデータベース52のうちのいずれか、または、双方）の最低オクターブを計算してもよい。

【0032】

スケール推定ユニット24は、さまざまなスケールに対するキーポイント分布を記述するコンフィグレーションデータを記憶してもよい。スケール推定ユニット24はその後、問い合わせ画像に対するキーポイントの分布をコンフィグレーションデータの分布と比較してもよい。スケール推定ユニット24はその後、問い合わせ画像中のオブジェクトに対するおおよそのスケールを、最良のマッチングキーポイント分布を有するスケールとして決定する。すなわち、スケール推定ユニット24は、さまざまなスケールにおける、問い合わせ画像に対するキーポイント分布を表す値を決定し、それらの値を、コンフィグレーションデータのさまざまなスケールのキーポイント分布に対して予期される値とを比較してもよい。スケール推定ユニット24は、問い合わせ画像からのキーポイント分布値が、対応するスケールに対してコンフィグレーションデータから推定されたキーポイント分布値に最もよくマッチするスケールとして、オブジェクトのスケールを予想してもよい。

【0033】

本開示の技術にしたがって、特徴抽出ユニット22は、問い合わせ画像に対する推定されたスケールに対応するオクターブにおいて開始してもよい。すなわち、特徴抽出ユニット22は、推定されたスケールに対応するオクターブに対する特徴記述子を、そのオクターブから抽出したキーポイントを使用して抽出してもよい。その後、制御ユニット20は、問い合わせ画像中のオブジェクトのアイデンティティが、それらの特徴記述子のみを使用して決定できるか否かを決定し、決定できない場合は、後続するオクターブに対する特徴記述子を抽出するように進めてもよい。

【0034】

深度推定ユニット26は、一般的に、カメラ12により取り込まれた画像中の1つ以上のオブジェクトに対する深度を推定するように構成されている。深度推定ユニット26は、飛行時間センサ、構造光センサまたはキネクト（Kinect）のような、深度センサを備えていてもよい。いくつかの例では、クライアントデバイス10は、例えば、3次元の画像および/またはビデオデータを取り込むために、2つ以上のカメラを有するカメラアレイを備えていてもよい。このような例では、深度推定ユニット26は、カメラアレイにより取り込まれた画像中のオブジェクトに対する深度情報を計算するように構成されていてもよい。例えば、深度推定ユニット26は、カメラアレイにより取り込まれたシーンのうちの2つ以上の画像を受け取ってもよい。画像の類似したピクセルの水平ロケーションを比較することにより、深度推定ユニット26は、画像間の類似したピクセルの水平格差（すなわち、視差）を計算し、この格差に基づいて、ピクセルに対応するオブジェクトの深度を決定してもよい。深度推定ユニット26は、1つ以上の深度値（例えば、深度マップ）を特徴抽出ユニット22に送ってもよく、特徴抽出ユニット22は、特定記述子を抽出する開始オクターブを決定するときに、深度値を使用してもよい。

【0035】

いくつかの例では、デバイスは、深度推定ユニットを備えている必要はなく、本開示の特定の技術は依然として動作可能であってもよい。しかしながら、デバイスが、深度推定ユニット26のような深度推定ユニットを備えるとき、デバイスは、特徴が計算されるオクターブと、オクターブが処理される順序とを選択するために、深度情報を使用するように構成されていてもよい。例えば、カメラ12が、比較的大きなオブジェクトに直接（すなわち、正面から）向かい合っている場合に、オブジェクトに対応するピクセルは、一般的に、クライアントデバイス10からの同じ距離において生じてよい。このような情報は、いくつかの例では、最も価値あるキーポイントが見つかることになるだろうスケールを決

10

20

30

40

50

定するために使用されてもよい。加えて、または、代替として、深度情報は、最も価値あるキーポイントが見つかるスケールを決定するために、信頼値に加えて使用できる。同様に、いくつかの例では、オブジェクトが、カメラに対して斜めに置かれている場合に、オブジェクト中の、最も近いポイントと最も遠いポイントとの間の深度差は、最も価値あるキーポイントが見つかるスケールの領域を識別するために使用することが可能である。この方法では、深度情報は、画像中のオブジェクトに対するスケール値を推定するために、少なくとも一部使用されてもよい。

#### 【0036】

以下で説明するように、ネットワークインターフェース32は、ワイヤレスネットワークインターフェースを含んでもよい。ネットワークインターフェース32が、ワイヤレスネットワークインターフェースを含むとき、ワイヤレスネットワークインターフェースは、IEEE802.11標準規格（例えば、802.11a/b/g/n）、第3世代移動体電気通信（3G）標準規格、第4世代電気通信標準規格（モバイルワールドワイドインターオペラビリティフォーマイクロウェーブアクセス（モバイルWiMAX（登録商標））、3Gロングタームエボリューション（LTE）、LTEアドバンスド、ワイヤレスMAN-アドバンスド（IEEE802.16）を含んでもよい）、または、他のワイヤレスネットワークプロトコルのうちの1つ以上のもののような、1つ以上のワイヤレスネットワークプロトコルを実現してもよい。一般的に、ワイヤレスネットワークが利用可能な場合、制御ユニット20は、クライアントデバイス10が屋内にある可能性が非常に高いことを決定してもよい。クライアントデバイス10が屋内にあるとき、制御ユニット20は、カメラ12により取り込まれた画像中のオブジェクトのサイズが比較的小さい可能性が高いことを決定してもよい。したがって、特徴抽出ユニット22は、問い合わせ画像に対する特徴記述子を抽出する開始オクターブを決定する際のファクタとして、ワイヤレスネットワークが利用可能であるか否かを取り扱うように構成されていてもよい。

#### 【0037】

図1の例では、クライアントデバイス10がローカルオブジェクトデータベース30を記憶する一方で、サーバデバイス50は遠隔オブジェクトデータベース52を記憶する。クライアントデバイス10は、ネットワーク40を介して、サーバデバイス50から、ローカルオブジェクトデータベース30に対するデータを取り出してもよい。したがって、ローカルオブジェクトデータベース30は、遠隔オブジェクトデータベース52のすべてまたは一部を表してもよい。いくつかの例では、ローカルオブジェクトデータベース30は、遠隔オブジェクトデータベース52からのデータとともに、（図1中には示されていない）他のサーバデバイスの他の遠隔オブジェクトデータベースからのデータを含んでもよい。一般的に、本開示の技術は、画像中のオブジェクトのアイデンティティを決定するために、ローカルオブジェクトデータベース30および/または遠隔オブジェクトデータベース52のうちのいずれかまたは双方を問い合わせる。

#### 【0038】

以下でより詳細に記述されるように、特徴抽出ユニット22は、一般的に、オリジナル画像に基づいてフィルタリングされた画像のさまざまなセットからキーポイントを抽出してもよい。いくつかの例では、オリジナル画像はガウシアンフィルタを使用してフィルタリングされ、ガウシアンぼかし画像のセットが生成されてもよい。さまざまなフィルタリング技術が適用されてもよいことを理解すべきである一方、本開示は、主に、例としてガウシアンフィルタに的を絞っている。しかしながら、他の類似したフィルタリング技術も、本開示の技術から逸脱することなく適用されてもよい。一般的に、1つのセット中のフィルタリングされた画像のそれぞれは、（例えば、ピクセル解像度の点から）類似したサイズを有していてもよく、ここで、セット中の各画像は、漸進的にさらにフィルタリングされてもよい。セットのそれぞれは、“オクターブ”と呼ばれてもよい。ガウシアンぼかしは、一般的に、規定されたスケールにおいて、特定のオクターブに対する画像データとガウシアンぼかし関数の畳み込みをとることを含む。特徴抽出ユニット22は、インクリメントに画像データの畳み込みをとってもよく、ここで、可変シグマ（ ）により参照さ

れる固定乗法的因子 (fixed multiplicative factor) は、オクターブ内の結果としてのガウシアンフィルタをインクリメントさせる。特徴抽出ユニット 22 は、特定のオクターブに対して、ガウシアンぼかし画像のそれぞれを有する、“ガウシアンピラミッド”と呼ばれるものを形成してもよい。特徴抽出ユニット 22 は、その後、ピラミッド中の連続してスタックされた 2 つのガウシアンぼかし画像を比較して、ガウシアンの差分 (D o G) 画像を発生させる。D o G 画像は、“D o G 空間”と呼ばれるものを形成してもよい。

#### 【0039】

この D o G 空間に基づいて、特徴抽出ユニット 22 は、キーポイントを検出してもよく、ここで、キーポイントは、幾何学的な観点から潜在的に関心がある画像データ中の、特定のサンプルポイントまたはピクセルの周りの、ピクセルの領域またはパッチのことを指す。一般的に、特徴抽出ユニット 22 は、構築された D o G 空間中の局所的な最大値および/または局所的な最小値として、キーポイントを識別する。特徴抽出ユニット 22 は、その後、キーポイントが検出されたパッチに対する局所的な画像勾配の方位に基づいて、1 つ以上の方向、すなわち方位をこれらのキーポイントに割り振ってもよい。これらの方向を特徴付けるために、特徴抽出ユニット 22 は、勾配方向ヒストグラムに関して方向を規定してもよい。特徴抽出ユニット 22 は、その後、ロケーションおよび (例えば、勾配方向ヒストグラムによる) 方向として特徴記述子を規定してもよい。特徴記述子を規定した後、特徴抽出ユニット 22 は、例えば、特徴記述子によりローカルオブジェクトデータベース 30 または遠隔データベース 52 に問い合わせることにより、特徴記述子を出力してもよい。特徴抽出ユニット 22 は、このプロセスを使用して、特徴記述子のセットを出力してもよい。

#### 【0040】

ネットワークインターフェース 32 は、ワイヤレスインターフェースおよびワイヤードインターフェースを含む、ネットワーク 40 を介してサーバデバイス 50 と通信することができる任意のタイプのインターフェースを表す。ネットワークインターフェース 32 は、ワイヤレスセルラインターフェースを表してもよく、ワイヤレスセルラネットワークを介してネットワーク 40 と、ネットワーク 40 を介してサーバデバイス 50 と通信するための、アンテナ、変調器およびこれらに類するもののような、必要なハードウェアまたは他のコンポーネントを含む。この例では、図 1 の例中では示されていないが、ネットワーク 40 は、ワイヤレスセルラネットワークインターフェース 32 がネットワーク 40 と通信するワイヤレスセルラアクセスネットワークを含む。図 1 中には図示されていないが、クライアントデバイス 10 はさらに、ディスプレイ、例えば、オブジェクトアイデンティティが決定されている画像データ、または、他の任意のタイプのデータのような、画像を表示することができる任意のタイプのディスプレイユニットを備えていてもよい。ディスプレイは、例えば、発光ダイオード (LED) ディスプレイデバイス、オーガニック LED (OLED) ディスプレイデバイス、液晶ディスプレイ (LCD) デバイス、プラズマディスプレイデバイス、または、他の任意のタイプのディスプレイデバイスを含んでもよい。

#### 【0041】

サーバデバイス 50 は、ネットワークインターフェース 54、遠隔オブジェクトデータベース 52 および制御ユニット 60 を備える。制御ユニット 60 は、この例では、特徴マッチングユニット 62 および信頼評価ユニット 64 を備える。ネットワークインターフェース 54 は、ネットワーク 40 のようなネットワークと通信することができる任意のタイプのインターフェースをネットワークインターフェース 54 が表している点で、クライアントデバイス 10 のネットワークインターフェース 32 に類似しているかもしれない。特徴マッチングユニット 62 は、クライアントデバイス 10 から受信した特徴記述子に基づいて、画像データ中の 1 つ以上の特徴またはオブジェクトを識別するために、特徴マッチングを実行するユニットは表す。

#### 【0042】

特徴マッチングユニット 62 は、この特徴識別を実行するために、遠隔オブジェクトデ

10

20

30

40

50

ータベース 52 にアクセスしてもよく、ここで、遠隔オブジェクトデータベース 52 は、特徴記述子を規定するデータを記憶し、受信した特徴記述子のうちの少なくともいくつかを、画像データから抽出された対応する特徴またはオブジェクトを識別する識別データに関係付ける。信頼評価ユニット 64 は、受信した特徴識別子に対応するとして識別されたオブジェクトが、画像データ中のオブジェクトの実際のアイデンティティであるという信頼を表す信頼値を決定する。

#### 【0043】

遠隔オブジェクトデータベース 52 中に記憶されているオブジェクトは、複数の特徴記述子を含んでいてもよく、クライアントデバイス 10 から受信した特徴記述子は、遠隔オブジェクトデータベース 52 中のオブジェクトの特徴記述子のサブセットのみにマッチしてもよい。一般的に、信頼値は、受信した特徴記述子と、対応するオブジェクトに関する特徴記述子との間のマッチ間の対応を表す。したがって、より大きい信頼値は、受信した特徴記述子が、遠隔オブジェクトデータベース 52 により記憶されているオブジェクトの相対的に多数の特徴記述子にマッチすることを反映してもよい一方で、より小さい信頼値は、受信した特徴記述子が、遠隔オブジェクトデータベース 52 により記憶されているオブジェクトの相対的に少数の特徴記述子にマッチすることを反映してもよい。

#### 【0044】

クライアントデバイス 10 から受信した特徴記述子に対するオブジェクトのアイデンティティを決定した後に、特徴マッチングユニット 62 は、決定したアイデンティティを表す識別データを提供する。同様に、信頼評価ユニット 64 は、決定したアイデンティティが、受信した特徴記述子に適切にマッチするという信頼を評価し、対応する信頼値を発生させ、信頼値をクライアントデバイス 10 に提供する。クライアントデバイス 10 は、信頼値がしきい値を超えているか否かを決定し、超えていない場合には、画像データに対するさらなるオクタブからの追加の特徴記述子を抽出し、さらなる解析のために、それらの追加の特徴記述子をサーバデバイス 50 に送ってもよい。

#### 【0045】

初めに、クライアントデバイス 10 のユーザは、クライアントデバイス 10 とインターフェースして、視覚探索を開始させてもよい。ユーザは、ユーザインターフェースと、または、クライアントデバイス 10 のディスプレイにより示されている他のタイプのインターフェースとインターフェースして、画像データを選択し、その後、1つ以上の特徴を、または、画像データとして記憶されている画像の焦点であるオブジェクトを識別する視覚探索を開始する。例えば、画像データは、一枚の著名な芸術作品の画像に対応していてもよい。ユーザは、クライアントデバイス 10 のカメラ 12 を使用して、この画像を取り込んだとしてもよく、ネットワーク 40 からこの画像をダウンロードしたとしてもよく、または、別のコンピュータデバイスとのワイヤードもしくはワイヤレス通信を介して、画像を局所的に取り込んだとしてもよい。何らかのイベントにおいて、画像データを選択した後に、ユーザは、視覚探索を開始して、この例では、例えば、名称、芸術家および完成日により、一枚の著名な芸術作品を識別してもよい。

#### 【0046】

視覚探索を開始することに応答して、クライアントデバイス 10 は、特徴抽出ユニット 22 を読み出して、画像データの解析を通して見つかった“キーポイント”と呼ばれるもののうちの1つを記述する少なくとも1つの特徴記述子を抽出する。制御ユニット 20 は、特徴記述子を使用してローカルオブジェクトデータベース 30 に問い合わせてもよく、および/または、ネットワーク 40 を介してサーバデバイス 50 に特徴記述子を送って、特徴記述子を使用して遠隔オブジェクトデータベース 52 に問い合わせてもよい。いくつかの例では、特徴抽出ユニット 22 は、特徴記述子の特徴圧縮ユニット（この例には示されていない）に転送し、特徴圧縮ユニットは、ネットワーク 40 を介しての特徴記述子の送信の前に特徴記述子を圧縮してもよい。サーバデバイス 50 が送られるとき、制御ユニット 20 は、例えば、TCP/IP パケットのような、ネットワークパケットとして、特徴記述子（カプセル化の前に圧縮されていてもよい）をカプセル化してもよい。

## 【 0 0 4 7 】

本開示では、開示した技術を実行するように構成されているデバイスの機能的な態様を強調するために、さまざまなコンポーネント、モジュールまたはユニットが記載されているが、これらのユニットは、必ずしも異なるハードウェアユニットによる実現を必要とするわけではない。むしろ、さまざまなユニットは、ハードウェアユニット中で組み合わせられていてもよく、あるいは、コンピュータ読取可能媒体に記憶されている適切なソフトウェアおよび/またはファームウェアとともに、上述した1つ以上のプロセッサを含む相互作用可能なハードウェアユニットの集合によって提供されてもよい。この態様では、本開示中のユニットへの参照は、別々のハードウェアユニット、ならびに/あるいは、ハードウェアおよびソフトウェアユニットとして、実現されてもよく、または、実現されなくてもよい異なる機能ユニットを提案することが意図されている。

10

## 【 0 0 4 8 】

上記で論じたように、本開示の技術にしたがった特徴抽出ユニット22は、カメラ12により取り込まれた画像から、特徴記述子をシーケンシャルに抽出するように構成されていてもよい。つまり、特徴抽出ユニット22は、第1のオクターブから特徴記述子を抽出し、それらの特徴記述子を使用して、画像中のオブジェクトのアイデンティティを正確に決定できるか否かを決定し、決定できない場合に、1つ以上の後続するオクターブから特徴記述子を抽出してもよい。さらに、特徴抽出ユニット22にデータを提供してもよい、センサ14、ネットワークインターフェース32、ローカルオブジェクトデータベース30中に記憶されているデータ、スケール推定ユニット24、および/または、深度推定ユニット26のような、クライアントデバイス10の要素は、第1および/または後続のオクターブを選択するために使用してもよい。一般的に、上位オクターブは、比較的良好な平滑度であり、下位オクターブよりも比較的安定していてもよい。したがって、いくつかの例では、特徴抽出ユニット22は、第1のオクターブが結果を出さなかった場合に、最後のオクターブを試してもよく、最後のオクターブが結果を出さない場合には、より高いオクターブにより後続される。いくつかの例では、オクターブゼロ(0)を試すときにマッチが存在しない場合、データベース中の参照画像は、より高いスケールにおいて存在しているかもしれない、したがって、特徴抽出ユニット22は、特徴抽出のために、より高いオクターブに移動してもよい。

20

## 【 0 0 4 9 】

トレーニング画像またはテスト画像からの特徴属性は、特徴生成およびマッチングにおける複雑性を低減させるために使用されてもよい。これらの統計は、特徴抽出プロセス中に含まれているパラメータを初期化するために使用されてもよい。いったん初期化されると、特徴抽出ユニット22は、特徴をインクリメントに抽出し、認識および/または推測を実行し、認識が、画像中のオブジェクトのアイデンティティに関する信頼できる決定に対して十分ではなかった場合に、特徴抽出の次のセットのために、特徴抽出パラメータを更新してもよい。トレーニングデータ(ローカルオブジェクトデータベース30中に記憶されていてもよく、および/または、遠隔オブジェクトデータベース52から取り出されてもよい)と、過去の問い合わせ統計とを使用して、問い合わせ画像中のオブジェクトの認識は、より少ないステップで達成でき、結果として計算を省くことができる。

30

40

## 【 0 0 5 0 】

スケール推定ユニット24は、特徴属性に基づいて、オブジェクトのスケールを識別するように試行してもよい。オブジェクトのスケールは、問い合わせる、データベースの最低オクターブを選択するために使用されてもよい。つまり、特徴記述子をデータベースにサブミットするとき、データベースは、最低オクターブが特定されているときに、最低オクターブ以上であり、最低オクターブより少なくないオクターブをサーチしてもよい。したがって、キーポイントの各セットが抽出された後に、スケール推定ユニット24は、キーポイントの分布を使用して、画像中のオブジェクトに対するスケールを予想してもよく、スケールを使用して、最低オクターブを推定してもよい。以下の図5は、キーポイント分布と対応する最低オクターブとの間の例示的な関係を図示している。データベースの最

50

低オクターブを推定するためのキーポイント分布の使用に関連するさらなる詳細が、図5に関して以下に論じられる。任意のケースでは、特徴記述子のインクリメントな抽出とともに、最低オクターブを使用することは、初めに、より少ない特徴記述子を抽出することと、実際にサーチされるデータベースの量を低減させることとの双方により、相対的により速いサーチを提供してもよい。

#### 【0051】

深度推定ユニット26は、問い合わせ画像に対する深度値を決定してもよい。例えば、深度推定ユニット26は、問い合わせ画像に対する深度マップを発生させてもよく、ここで、深度マップは、問い合わせ画像中のピクセルまたはオブジェクトに対する深度値を含んでいてもよい。深度マップを使用して、特徴抽出ユニット22は、マッチングまたはトラッキングを制約してもよく、このことは、結果として偽陽性を少なくすることができる。さらに、特徴抽出ユニット22は、特定の深度レベルへのマッチングまたはトラッキングを制限してもよく、このことは、結果として複雑性を低減できる。

#### 【0052】

この方法で、本開示の技術は、特徴記述子抽出における特定の観測の利益を得てもよい。それらの観測は、初めに、問い合わせ画像中のオブジェクトのスケールが、一般的に、そのキーポイントがスケール空間中でどのように分散されるかを決定してもよいことを含む。より小さいオブジェクトは、スケール空間中でより低いスケールへ圧縮される分布を有する傾向がある。また、画像中の特定の領域が他の領域よりもざらつきがある場合に、結果としてのキーポイントは、それらの安定性の点で異なる傾向がある。同様に、キーポイントにおいてくられる特徴記述子に対する品質も、それらの安定性の点で異なる。さらに、画像品質は、一般的に、より良い造影条件（例えば、良好な照度）により向上し、このことは、結果として画像からのキーポイントがより多くなってもよい。しかしながら、解像度が相対的に高いときに、それに対するデータが、例えば、ローカルオブジェクトデータベース30および/または遠隔オブジェクトデータベース52中に記憶されているトレーニング画像と比べて、より大きな変化、つまり、より大きなディテールであるために、実際の特徴記述子の数はより少なくなる傾向がある。これらの差分は、異なる照度、遠近感、または、他の写真技術パラメータからの結果であってもよい。

#### 【0053】

上記で論じたローカルオブジェクトデータベース30および/または遠隔オブジェクトデータベース52は、一般的に、トレーニング画像のセットに対するさまざまなデータを記憶する。このデータは、さまざまなオブジェクトに対する特徴記述子とともに、スケール推定、深度マッピング、および/または、画像中のオブジェクト認識の複雑性を低減する他の技術を支援するデータを含む。例えば、データベースは、画像の所定のロケーションセル（すなわち、“ロクセル”）における特徴記述子をいずれのオクターブおよびスケールから抽出するかを示すデータを記憶してもよい。画像は、ロクセルと呼ばれる2次元行列の領域に分割されてもよい。データベースデータは、オクターブおよびスケールの順序、および/または、マッチング特徴がいずれのオクターブおよびスケールに典型的に属するかについてのサイド情報を含んでもよい。データベースはまた、所定のロクセル内にうまくマッチする可能性が最も高いスケールまたはオクターブを記述する使用量統計を記憶してもよい。

#### 【0054】

いくつかの例では、データベースは、画像スケール推定に基づくキーポイント分布を記憶してもよい。データベースはまた、いくつかの例では、センサベースのロケーション情報に対応するデータを記憶してもよく、センサベースのロケーション情報は、特徴抽出において供給するスケールおよびオクターブパラメータを初期化するために使用されてもよい。例えば、制御ユニット20は、センサ14からGPS情報を使用して、問い合わせ画像が屋外環境のものであることを決定してもよく、データベースは、問い合わせ画像が取り込まれるときに、クライアントデバイス10の地理的ロケーションに対するオブジェクトのロケーションについての情報を記憶してもよい。制御ユニット20は、ワイヤレスネ

ットワークが利用可能であるか否かを決定することのような屋内ポジショニング技術に基づいて、問い合わせ画像が屋内環境であることを決定してもよい。ステレオカメラは、問い合わせ画像中のオブジェクトの深度を推定するために使用されてもよく、オブジェクトの深度は、オクターブ選択を初期化する際に支援してもよい。カメラ内部パラメータ（例えば、焦点距離、、主点、画像フォーマット）を使用して、オブジェクトに対する深度を推定してもよい。いくつかの例では、内部パラメータは、希薄な深度マップに対してアプリアリに構成される必要はない。いくつかの例では、制御ユニット20は、例えば、8点アルゴリズムを使用して、特徴対応から内部パラメータおよび外部パラメータを計算してもよい。一般的に、8点アルゴリズムは、ステレオ画像対に対する本質的な行列または基礎となる行列を計算するために、ステレオ画像対中の8個の対応するポイントを使用することを含む。

10

#### 【0055】

いくつかの例では、制御ユニット20は、特徴抽出を実行する最初のスケールまたはオクターブを決定するために、履歴データを使用するように構成されていてもよい。例えば、制御ユニット20により実行される（示されていない）トラッカーは、カメラ12により取り込まれた画像中のオブジェクトに対する深度またはスケールを決定する。トラッカーがターゲットを見失った場合、制御ユニット20は、より新しい画像中のオブジェクトに対するスケールを決定するときに、追跡したターゲットのスケールについての1番最近の“良好な”情報を使用してもよい。

#### 【0056】

20

上述したデータのうちのいくつかまたはすべてを使用して、制御ユニット20は、問い合わせ画像中のオブジェクトのアイデンティティを決定するように構成されていてもよい。例えば、特徴抽出ユニット22は、初めに、好ましいオクターブおよびスケールのセットを選択し、好ましいオクターブおよびスケールのセットに対する特徴記述子を抽出してもよい。特徴抽出ユニット22は、必要とされるときに、スケール空間中に次のオクターブを漸進的に追加してもよく、このことは、プロセスを早まって終了させるリスクなしで、オブジェクト認識プロセスの早い終了を可能にしてもよい。オクターブおよびスケールのセットに対するプリファレンスは、ローカルオブジェクトデータベース30および/または遠隔オブジェクトデータベース52からのデータから、ならびに/あるいは、トラッカーからの問い合わせ統計から、取り出すことができる。スケール推定ユニット24は、キーポイント抽出の各反復の後に、累積のしきい値統計を使用して、オブジェクトのスケールを推定してもよく、これはまた、データベースのデータを使用して精練されてもよい。さらに、推定されたスケールは、センサ14からのデータ、深度推定ユニット26により決定された深度情報、または、他のデータのような、追加の入力に基づいて精練されてもよい。

30

#### 【0057】

本開示の技術は、1つ以上の利点を提供してもよい。例えば、これらの技術は、例えば、SIFTに対して、SIFTが別の方法で達成するだろうものよりも、より速い特徴記述子の局所的な抽出を提供してもよい。これらの技術は、処理に対するスケール空間レベル（例えば、オクターブレベル）の最も適切なサブセットの推定を可能にしてもよく、このことは、オブジェクト認識のスピードおよび複雑性を向上させてもよい。これらの技術はさらに、クライアントデバイス10（例えば、移動電話機）のローカルオブジェクトデータベース30中をより速くサーチすることを可能にしてもよい。つまり、これらの技術は、初めに、上位レベルからの特徴記述子抽出を、そして、十分な正確度に到達するときにアルゴリズムを終了させることを可能にする。この方法では、図1のシステム2のような、分配された視覚探索システムにおける比較的短い応答とすることができる。特徴記述子は、初めに、上位レベルで開始して、ネットワーク40を介して、サーバデバイス50に漸進的に送信されてもよく、サーバデバイス50は、受信したデータについてのサーチのいくつかの反復を実行して、いったん十分な正確なマッチがみつかり、結果をまたは終了信号を送り返してもよい。クライアントデバイス10は、サーチの結果または終了信号

40

50

を受信した後に、サーバデバイス 50 に特徴記述子を送ることを終了してもよい。

#### 【0058】

本開示の技術はまた、トラッキングの複雑性を低減させて、スピードを向上させてもよい。いったん画像中のオブジェクトが認識されると、その関連するスケールも同様に既知になってもよい。トラッカーがオブジェクトを見失った場合、次のサーチ動作は、以前に追跡されていたオブジェクトのスケールを使用して簡略化されてもよい。これらの技術は、推定されたスケールを使用して、キーポイントのクラスタ化に対するカーネルサイズを初期化し、その後、関連する深度マップ推定に基づいて、セグメント化を実行してもよい。

#### 【0059】

この方法で、クライアントデバイス 10 は、プロセッサを備えるデバイスの例を表す。プロセッサは、受信した画像の第 1 のオクターブのぼかし画像の第 1 のセットから、1 つ以上のキーポイントの第 1 のセットを抽出し、キーポイントの第 1 のセットに対する 1 つ以上の記述子の第 1 のセットを計算し、記述子の第 1 のセットにより特徴記述子データベースに問い合わせることにより生じた結果に対する信頼値を受信し、結果は、受信した画像中のオブジェクトのアイデンティティを記述する情報を含み、信頼値が信頼しきい値を超えないときに、受信した画像の第 2 のオクターブのぼかし画像の第 2 のセットから、1 つ以上のキーポイントの第 2 のセットを抽出するように構成されている。

#### 【0060】

図 2 は、キーポイント抽出において使用するために決定されているガウシアン差分 (DOG) ピラミッド 104 を図示している概念図である。図 2 の例は、ガウシアンピラミッド 102 および対応する DOG ピラミッド 104 中の画像のセット 100 を図示しており、ここで、セット 100 は、第 1 の選択されたオクターブ、ならびに、ガウシアンピラミッド 122 および対応する DOG ピラミッド 124 中の画像の第 2 のセット 120 に対応し、セット 120 は、第 2 のオクターブに対応する。この開示の技術にしたがうと、制御ユニット 20 は、セット 100 に対するデータを使用して識別されたオブジェクトが、しきい値を下回る信頼値を生じさせるときに、セット 120 に対してデータを発生させるように構成されていてもよい。

#### 【0061】

図 1 の特徴抽出ユニット 22 は、ガウシアンピラミッド 102 中の任意の 2 つの連続するガウシアンぼかし画像の差分を計算することにより、DOG ピラミッド 104 を構築してもよい。図 1 の例の中の特徴抽出ユニット 22 によってカメラ 12 から受け取られたかもしれない入力画像  $I(x, y)$  は、ガウシアンピラミッド 102 を構築するために、段階的にガウシアンぼかしされる。ガウシアンぼかしは、一般的に、ガウシアンぼかし関数  $L(x, y, c)$  が  $L(x, y, c) = G(x, y, c) * I(x, y)$  として規定されるように、スケール  $c$  においてオリジナル画像  $I(x, y)$  とガウシアンぼかし関数  $G(x, y, c)$  の畳み込みをとるプロセスを含む。ここで、 $G$  は、ガウシアンカーネルであり、 $c$  は、画像  $I(x, y)$  をぼかすために使用される、ガウシアン関数の標準偏差を示す。 $c$  が変化するにつれて ( $c_0 < c_1 < c_2 < c_3 < c_4$ )、標準偏差  $c$  は変化し、段階的なぼかしが取得される。シグマは、ベーススケール変数 (本質的には、ガウシアンカーネルの幅) である。最初の画像  $I(x, y)$  がインクリメントにガウシアン  $G$  と畳み込まれ、ぼかし画像  $L$  を生成させるとき、ぼかし画像  $L$  は、スケール空間中で定数因子  $c$  により分離される。SIFT の例では、ガウシアンカーネルは、スケール空間を発生させるために使用されてもよい。いくつかの例では、代替的な、ローパスカーネルを使用して、スケール空間、例えば、直方体関数、三角関数または他のこのような関数を発生させてもよい。

#### 【0062】

DOG 空間またはピラミッド 104 では、 $D(x, y, a) = L(x, y, c_n) - L(x, y, c_{n-1})$  である。DOG 画像  $D(x, y, )$  は、スケール  $c_n$  および  $c_{n-1}$  における 2 つの近接するガウシアンぼかし画像  $L$  間での差分である。 $D(x, y, )$  のスケー

10

20

30

40

50

ルは、 $c_n$ と $c_{n-1}$ との間のどこかに存在する。ガウシアンぼかし画像 $L$ の数が増加し、ガウシアンピラミッド102に対して提供された近似が連続空間に近づくにつれて、2つのスケールも1つのスケールに近づく。畳み込まれた画像 $L$ は、オクターブによりグループ化されてもよく、ここで、オクターブは、標準偏差の値の倍加に対応する。さらに、畳み込まれた画像 $L$ の固定数がオクターブ毎に取得されるように、乗数 $k$ の値（例えば、 $c_0 < c_1 < c_2 < c_3 < c_4$ ）が選択される。その後、DoG画像 $D$ は、オクターブ毎の近似ガウシアンぼかし画像 $L$ から取得されてもよい。

#### 【0063】

本開示の技術にしたがうと、所定のオクターブに対して画像 $D$ が取得された後に、特徴抽出ユニット22は、オクターブに対するキーポイントを抽出し、それらの抽出したキーポイントに対する特徴記述子を決定してもよい。現在のオクターブ（および任意の以前のオクターブ）に対する特徴記述子は、画像中のオブジェクトのアイデンティティを決定するように試行するために使用されてもよい。オブジェクトのアイデンティティが決定された後に、制御ユニット20は、決定されたアイデンティティに対する信頼値を取得してもよい。信頼値がしきい値より小さい場合に、特徴抽出ユニット22は、後続のオクターブに進んでもよく、後続のオクターブに対する追加の特徴記述子を決定し、再び、画像中のオブジェクトのアイデンティティを決定するように試行する。その一方で、信頼値がしきい値を超えるときは、特徴抽出ユニット22は、後続するオクターブへ進む必要はない。下記でより詳細に記述するように、制御ユニット20は、例えば、画像中のオブジェクトに対して予想されたスケール、問い合わせ画像が屋内または屋外環境で取り込まれたかを示すセンサデータ、および/または、問い合わせ画像中のオブジェクトに対する深度値を示す、問い合わせ画像に対する深度マップのような、さまざまなファクタに基づいて、第1のオクターブおよび後続するオクターブを選択してもよい。

#### 【0064】

特徴抽出ユニット22は、その後、DoGピラミッド104を使用して、画像 $I(x, y)$ に対するキーポイントを識別してもよい。キーポイント抽出を実行する際に、特徴抽出ユニット22は、画像中の特定のサンプルポイントまたはピクセルの周りの局所的な領域またはパッチが、潜在的に（幾何学的に言って）対象のパッチであるか否かを決定する。一般的に、特徴抽出ユニット22は、DoG空間104中の局所的な最大値および/または局所的な最小値を識別し、DoG空間104中のキーポイントロケーションとして、これらの最大値および最小値のロケーションを使用する。図2中で図示されている例では、特徴抽出ユニット22は、パッチ106内のキーポイント108を識別する。（局所的極値検出としても知られている）局所的な最大値および最小値の発見は、DoG空間104中の各ピクセル（例えば、キーポイント108に対するピクセル）を、同一のスケールにおけるその8個の隣接するピクセルと、両側の隣接スケールのそれぞれにおける（近接パッチ110および112中の）9個の隣接するピクセルとの合計26ピクセル（ $9 \times 2 + 8 = 26$ ）とを比較することにより達成されてもよい。キーポイント106に対するピクセル値が、パッチ106、110および108における、26個のすべての比較されたピクセルの中で最大値または最小値である場合に、特徴抽出ユニット22はキーポイントとしてこれを選択する。特徴抽出ユニット22は、これらロケーションがより正確に識別されるように、キーポイントをさらに処理してもよい。特徴抽出ユニット22は、いくつかの例では、低コントラストのキーポイントおよびエッジのキーポイントのような、キーポイントのうちのいくつかを廃棄してもよい。

#### 【0065】

図3は、キーポイントの検出をより詳細に図示している概念図である。図3の例では、パッチ106、110および112のそれぞれが $3 \times 3$ ピクセル領域を含む。特徴抽出ユニット22は、初め、対象（例えば、キーポイント108）のピクセルを、同一スケール（例えば、パッチ106）におけるその8個の隣接するピクセル132、ならびに、キーポイント108の両側の隣接するスケールのそれぞれにおける隣接するパッチ110および112中の9個の隣接するピクセル134および136と比較する。

## 【 0 0 6 6 】

特徴抽出ユニット 2 2 は、局所的な画像勾配の方位に基づいて、1 つ以上の方向、すなわち方位を各キーポイントに割り振ってもよい。局所的な画像プロパティに基づいて、各キーポイントに矛盾のない方向を割り振ることにより、特徴抽出ユニット 2 2 は、この方向に関連するキーポイント記述子を表し、それゆえ、画像回転に対する不変性を達成してもよい。特徴抽出ユニット 2 2 は、その後、ガウシアンぼかし画像  $L$  中の、および / または、キーポイントスケールにおける、キーポイント 1 0 8 の周りの隣接する領域中のすべてのピクセルに対する大きさおよび方位を計算する。 $(x, y)$  に位置付けられているキーポイント 1 0 8 に対する勾配の大きさは、 $m(x, y)$  として表されてもよく、 $(x, y)$  におけるキーポイントに対する勾配の方向または方位は、 $\theta(x, y)$  として表されてもよい。

10

## 【 0 0 6 7 】

特徴抽出ユニット 2 2 は、その後、すべての計算がスケールが変化しない方法で実行されるように、キーポイントのスケールを使用して、キーポイント 1 0 8 のスケールに最も近いスケールを持つガウシアン平滑化画像  $L$  を選択する。このスケールにおける各画像サンプル  $L(x, y)$  に対して、特徴抽出ユニット 2 2 は、ピクセル差分を使用して、勾配の大きさ  $m(x, y)$  および方向  $\theta(x, y)$  を計算する。例えば、大きさ  $m(x, y)$  は、以下の方程式 (1) にしたがって計算されてもよい。

## 【 数 1 】

20

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (1)$$

## 【 0 0 6 8 】

特徴抽出ユニット 2 2 は、以下の方程式 (2) にしたがって、方位または方向  $\theta(x, y)$  を計算してもよい。

## 【 数 2 】

30

$$\Gamma(x, y) = \arctan \left[ \frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \right] \quad (2)$$

## 【 0 0 6 9 】

方程式 (2) では、 $L(x, y)$  は、キーポイントのスケールでもあるスケールにおける、ガウシアンぼかし画像  $L(x, y)$  のサンプルを表す。

## 【 0 0 7 0 】

特徴抽出ユニット 2 2 は、D o G 空間中のキーポイントの面よりも高いスケールにおける、上側にあるガウシアンピラミッド中の面に対するキーポイント、または、キーポイントよりも低いスケールにおける、下側にあるガウシアンピラミッドの面中のキーポイントのうちのいずれかに対する勾配を矛盾なく計算してもよい。いずれにしても、各キーポイントに対して、特徴抽出ユニット 2 2 は、キーポイントを囲む長方形エリア (例えば、パッチ) 中の同一のスケールにおける勾配を計算する。さらに、画像信号の周波数は、ガウシアンぼかし画像のスケールにおいて反映される。しかしながら、S I F T、および、勾配の圧縮ヒストグラム (C H o G) アルゴリズムのような他のアルゴリズムは単に、パッチ (例えば、長方形エリア) 中のすべてのピクセルにおける勾配値を使用する。パッチは、キーポイントの周りで規定され、サブブロックはブロック内で規定され、サンプルはサブブロック内で規定され、この構造は、たとえキーポイントのスケールが異なるときでも

40

50

すべてのキーポイントに対して同じものが維持される。それゆえ、画像信号の周波数は、同一のオクターブにおける、ガウシアン平滑フィルタの連続のアプリケーションによって変化するが、異なるスケールにおいて識別されたキーポイントは、スケールにより表される、画像信号の周波数における変化に関わらず、同数のサンプルによりサンプリングされてもよい。

#### 【 0 0 7 1 】

キーポイント方向を特徴付けるために、特徴抽出ユニット 2 2 は、例えば、勾配の圧縮ヒストグラム (CHoG) を使用して、勾配方向ヒストグラム (図 4 参照) を生成されてもよい。各隣接するピクセルのコントリビューションは、勾配の大きさおよびガウシアンウィンドウにより重み付けされてもよい。ヒストグラム中のピークは、優位方向に対応する。特徴抽出ユニット 2 2 は、キーポイント方向に関連するキーポイントのすべてのプロパティを測定してもよく、これは、回転に対する不変性を提供してもよい。

10

#### 【 0 0 7 2 】

1つの例では、特徴抽出ユニット 2 2 は、各ブロックに対するガウシアン重み付け勾配の分布を計算し、ここでは、各ブロックは、2サブブロック×2サブブロックの合計4サブブロックである。ガウシアン重み付け勾配の分布を計算するために、特徴抽出ユニット 2 2 は、いくつかのピンを持つ方向ヒストグラムを形成し、各ピンは、キーポイントの周りのエリアの一部をカバーする。例えば、方向ヒストグラムは、36個のピンを有していてもよく、各ピンは、方向の360度の範囲のうちの10度をカバーする。代替的に、ヒストグラムは8個のピンを有していてもよく、それぞれは、360度の範囲のうちの45度をカバーする。ここに記述するヒストグラムコーディング技術は、任意の数のピンのヒストグラムに適応可能であってもよいことは明白であるべきである。

20

#### 【 0 0 7 3 】

図 4 は、特徴抽出ユニット 2 2 が勾配分布および方向ヒストグラムを取得できる例示的な技術を図示している概念図である。ここで、2次元勾配分布 ( $dx, dy$ ) (例えば、ブロック 1 5 6) は、1次元分布 (例えば、ヒストグラム 1 6 4) に変換される。キーポイント 1 0 8 は、キーポイント 1 0 8 を囲む (セルまたは領域とも呼ばれる) パッチ 1 5 6 の中心に位置付けられている。ピラミッドの各レベルに対して予め計算されている勾配が、各サンプルロケーション 1 5 8 における小さな矢印として示されている。示されているように、サンプル 1 5 8 の  $4 \times 4$  グループは、サブブロック 1 6 0 を形成し、サブブロックの  $2 \times 2$  グループは、ブロック 1 5 6 を形成する。ブロック 1 5 6 を、記述子ウィンドウと呼ぶこともある。

30

#### 【 0 0 7 4 】

ガウシアン重み付け関数は、円 1 5 2 により示され、サンプルポイント 1 5 8 のそれぞれの大きさに重みを割り振るために使用されてもよい。円形ウィンドウ 1 5 2 中の重みはスムーズに減少する。ガウシアンウィンドウ 1 5 2 の目的は、ウィンドウのポジションにおける小さな変化により記述子における突然の変化を回避し、記述子の中心から遠い勾配に少ない強調を与えることである。方向ヒストグラム 1 6 2 の  $2 \times 2 = 4$  アレイは、ヒストグラムの各ピン中に8方向を持つ  $2 \times 2$  サブブロック、結果として、 $(2 \times 2) \times 8 = 32$  次元特徴記述子ベクトルから取得される。例えば、方向ヒストグラム 1 6 3 および 1 6 5 は、サブブロック 1 6 0 に対する勾配分布に対応してもよい。しかしながら、各ヒストグラム中の8方向を持つヒストグラム (8ピンヒストグラム) の  $4 \times 4$  アレイ、結果として、各キーポイントに対する  $(4 \times 4) \times 8 = 128$  次元特徴記述子ベクトルを使用することは、より良好な結果をもたらす場合がある。(例えば、異なるボロノイセル構造を持つ) 他のタイプの量子化ピン配列もまた、勾配分布を取得するために使用されてもよいことに留意すべきである。

40

#### 【 0 0 7 5 】

ここで使用されるヒストグラムはマッピング  $k_i$  であり、マッピング  $k_i$  は、ピンとして知られている、さまざまな互いに素なカテゴリに分類される、観測、サンプルまた実現値 (例えば、勾配) の数をカウントする。ヒストグラムのグラフは、ヒストグラムを表す 1

50

つの方法に過ぎない。したがって、 $k$  が、観測、サンプルまたは実現値の総数であり、 $m$  がピンの総数である場合、ヒストグラム  $k_i$  中の周波数は、以下の条件を満たす。

【数 3】

$$n = \sum_{i=1}^m k_i \quad (3)$$

10

【0076】

ここで、 $+$  は、加算演算子である。

【0077】

ヒストグラム 162 に追加される各サンプルは、標準偏差を持つ、ガウシアン重み付け円状ウィンドウ 152 内のその勾配の大きさ、つまり、キーポイントのスケールの 1.5 倍により重み付けされてもよい。結果としての方向ヒストグラム 164 中のピークは、局所的な勾配の優位方位に対応する。ヒストグラム中の最も高いピークが検出され、その後、最も高いピークのうちの、80% のような、特定のパーセンテージ内である、他の何らかの局所的なピークも使用されて、その方向を持つキーポイントを作成する。それゆえ、類似した大きさの複数のピークを持つロケーションに対して、同一のロケーションおよびスケールではあるが、異なる方向において、複数のキーポイントを作成できる。

20

【0078】

サブブロックからのヒストグラムは、キーポイントに対する特徴記述子ベクトルを取得するために連結されてもよい。16 個のサブブロックからの 8 個のピンのヒストグラム中の勾配が使用される場合、128 次元特徴記述子ベクトルが結果として生じてもよい。

【0079】

この方法では、各キーポイントに対して、記述子が取得されてもよく、ここで、このような記述子は、ガウシアン重み付け勾配の分布の、ロケーション ( $x, y$ )、方向および記述子により特徴付けられてもよい。画像は、1 つ以上のキーポイント記述子 (画像記述子とも呼ばれる) により特徴付けられてもよいことに留意すべきである。

30

【0080】

いくつかの例示的なアプリケーションでは、画像は、移動デバイスにより取得および/または取り込まれてもよく、オブジェクト認識は、取り込まれた画像上、または、取り込まれた画像の一部について実行されてもよい。第 1 のオプションにしたがうと、取り込まれた画像は移動デバイスによりサーバへ送られてもよく、そこで、画像は (例えば、1 つ以上の記述子を取得するために) 処理され、および/または、複数の画像 (例えば、複数の画像に対する 1 つ以上の記述子) と比較されて、マッチ (例えば、取り込まれた画像またはその中のオブジェクトの識別) を取得してもよい。しかしながら、このオプションでは、取り込まれた画像全体が送られ、このことは、そのサイズにより望ましくないかもしれない。第 2 のオプションでは、移動デバイスは、画像を処理 (例えば、画像について特徴抽出を実行) して、1 つ以上の画像記述子を取得し、画像および/またはオブジェクト識別のために、記述子をサーバに送る。画像よりもむしろ、画像に対するキーポイント記述子が送られることから、これには、画像に対するキーポイント記述子が画像自体よりも小さい限り、より少ない送信時間を要してもよい。したがって、キーポイント記述子のサイズを圧縮することは、極めて望ましいことかもしれない。

40

【0081】

キーポイント記述子のサイズを最小化するために、勾配の分布の記述子を圧縮することが有益であるかもしれない。勾配の分布の記述子はヒストグラムにより表されるので、ヒストグラムに対する効率的なコーディング技術をここに記述する。

【0082】

50

特徴記述子を効率的に表すおよび／または圧縮するために、分布の記述子（例えば、方向ヒストグラム）は、より効果的に表されてもよい。したがって、ヒストグラムの効率的なコーディングに対する１つ以上の方法または技術をここに記述する。これらの方法または技術は、圧縮された形態でヒストグラムを効率的に（または、もっと言えば最適に）コード化するために、任意のタイプのヒストグラムインプリメンテーションにより実現されてもよい。ヒストグラムの効率的なコーディングは、従来のエンコーディング技術により取り扱われない特異な問題である。従来のエンコーディング技術は、値のシーケンスの効率的なエンコーディングに的を絞っていた。ヒストグラムでは、シーケンス情報が使用されないことから、ヒストグラムの効率的なエンコーディングは、従来の技術により取り扱われた問題とは異なる問題である。

10

## 【 0 0 8 3 】

第１のステップにおいて、ヒストグラムの最適な（最小サイズまたは長さの）コーディングが考慮される。情報理論上の原理を適用して、ヒストグラムの無損失および／または有損失のエンコーディングに対する最大の長さを取得してもよい。

## 【 0 0 8 4 】

上記で述べたように、（例えば、セルまたは領域とも呼ばれることが多い）特定のパッチに対して、パッチ中の勾配の分布はヒストグラムで表されてもよい。ヒストグラムは、 $m$ 個のシンボル（ $2^m$ ）の長さを持つアルファベット  $A$  で表されてもよく、ここで、各シンボルは、ヒストグラム中のピンに関係する。それゆえ、ヒストグラムは、 $m$ 個のピンの総数を有する。例えば、アルファベット  $A$  中の各シンボル（ピン）は、規定されている勾配／方向のセットからの勾配／方向に対応してもよい。ここで、

20

## 【 数 4 】

$$n = \sum_{i=1 \dots m} k_i$$

## 【 0 0 8 5 】

になるように、 $n$  は、観測、サンプルまたは実現値（セル、パッチまたは領域中の勾配サンプル）の総数を表してもよく、 $k$  は、特定のピン中の観測、サンプルまたは実現値の数を表す（例えば、 $k_1$  は、第１のピン中の勾配サンプルの数であり、 $\dots$   $k_m$  は、 $m$ 番目のピン中の勾配サンプルの数である）。つまり、ヒストグラムのピン中のすべての勾配サンプルの合計は、パッチ中の勾配サンプルの総数に等しい。ヒストグラムは、セル、マッチまたは領域内の勾配サンプルの第１の分布に対する確率分布を表してもよいことから、勾配サンプルの（第１の分布とは異なる）第２の分布を有する異なるセル、パッチまたは領域がそれでもなお同一のヒストグラムを有する可能性がある。

30

## 【 0 0 8 6 】

$P$  が、 $m$ に関する確率分布  $[p_1, \dots, p_m]$  を示す場合、この分布のエントロピー  $H(P)$  は、

## 【 数 5 】

40

$$H(P) = - \sum_{i=1}^m p_i \log p_i \quad (4)$$

## 【 0 0 8 7 】

のように規定できる。

## 【 0 0 8 8 】

このケースでは、２つの既知の分布  $P$  および  $Q$  の間の相対エントロピー  $D(P||Q)$  は、

【数 6】

$$D(P \parallel Q) = \sum_{i=1}^m p_i \log \frac{p_i}{q_i} \quad (5)$$

【 0 0 8 9 】

により与えられる。

【 0 0 9 0 】

勾配分布の所定のサンプル  $w$  に対して、各勾配値が現れる回数は、 $(i = 1, \dots, m$  10  
 に対して)  $k_i$  により与えられると仮定する。サンプル  $w$  の確率  $P(w)$  は、したがって、

【数 7】

$$P(w) = \prod_{i=1}^m p_i k_i^{k_i} \quad (6)$$

【 0 0 9 1 】

20

により与えられる。

【 0 0 9 2 】

ここで、 $\prod$  は積演算子である。

【 0 0 9 3 】

例えば、セルまたはパッチのケースでは、確率  $P(w)$  は、特定のセルまたはパッチの  
 確率である。

【 0 0 9 4 】

しかしながら、方程式 6 は、分布  $P$  が既知であると仮定する。ソース分布が既知ではな  
 いケースでは、パッチ中に典型的な勾配を持つケースであるとして、サンプル  $w$  の確率は  
 、Krichcvsky-Trofimov (KT) 推定により与えられてもよい。 30

【数 8】

$$P_{KT}(w) = \Gamma\left(\frac{m}{2}\right) \frac{\prod_{i=1}^m \Gamma(k_i + \frac{1}{2})}{\pi^{\frac{m}{2}} \Gamma(n + \frac{m}{2})} \quad (7)$$

【 0 0 9 5 】

ここで、

40

【数 9】

$$\Gamma(n) = (n-1)!$$

【 0 0 9 6 】

であるように、 $\Gamma$  はガンマ関数である。

【 0 0 9 7 】

サンプル  $w$  は、その確率の KT 推定を使用してエンコードされることになる場合に、( 50

実際の分布  $P$  の下での ) このようなエンコーディングの長さ  $L$  は、  
【数 10】

$$L_{KT}(w, P) = - \sum_w P(m) \log P_{KT}(w) \sim nH(P) + \frac{m-1}{2} \log n \quad (8)$$

【0098】

を満たす。

10

【0099】

方程式 8 は、ヒストグラムの無損失エンコーディングに対する最大コード長を提供する。  
。  $KT$  推定量ベースのコードの冗長は、

【数 11】

$$R_{KT}(n) \sim \frac{m-1}{2} \log n \quad (9)$$

20

【0100】

により与えられてもよい。

【0101】

これは、実際のソース分布に依存しない。このことは、このようなコードが普遍的であることを意味する。したがって、 $KT$  推定量は、使用されるサンプル  $w$  が十分長い限りは、実際の確率  $P$  に近い近似を提供する。

【0102】

$KT$  推定量は、分布に対する確率を計算する唯一の方法であることに留意すべきである。  
。例えば、最尤 ( $ML$ ) 推定量も使用されてもよい。

30

【0103】

また、ヒストグラムをコード化するとき、エンコーダおよびデコーダの双方は、ヒストグラム中のサンプルの総数  $n$  およびヒストグラムに対するビンの数  $m$  を決定できる、または、ヒストグラム中のサンプルの総数  $n$  およびヒストグラムに対するビンの数  $m$  により構成されていることが仮定されてもよい。したがって、このケースでは、この情報はエンコードされる必要がない。それゆえ、エンコーディングは、 $m$  個のビンのそれぞれに対するサンプルの数にフォーカスされる。

【0104】

キーポイント (または画像) 記述子の一部としてヒストグラム自体を送信するよりもむしろ、ヒストグラムの圧縮形態を使用してもよい。これを達成するために、ヒストグラムをタイプにより表してもよい。一般的に、タイプは、ヒストグラムの圧縮表示である (例えば、タイプは、ヒストグラム全体よりもむしろ、ヒストグラムの形状を表す)。サンプル  $w$  のタイプ  $t$  は、タイプ  $t(w)$  が、そのシンボルの周波数のセット (例えば、勾配分布  $k_i$  の周波数) を表すように、

40

【数 1 2】

$$t(w) = \left[ \frac{k_1}{n}, \dots, \frac{k_n}{n} \right] \quad (10)$$

【0 1 0 5】

として規定されてもよい。タイプはまた、サンプル  $w$  を生成させたソースの真の分布の推定として理解できる。したがって、タイプ  $t(w)$  のエンコーディングおよび送信は、それが特定のサンプル  $w$  に基づいて推定できることから、分布の形状のエンコーディングおよび送信と同等なものである。

10

【0 1 0 6】

しかしながら、従来のエンコーディング技術は、値のシーケンスの効率的なエンコーディングにフォーカスされていた。ヒストグラムではシーケンス情報が使用されないことから、ヒストグラムの効率的なエンコーディングは、従来のエンコーディング技術により取り扱われた問題とは異なる問題である。ピン数がエンコーダおよびデコーダに対して既知であると仮定すると、ヒストグラムのエンコーディングは、ポイント（例えば、勾配）の総数、および、ピン毎のポイントの総数のエンコーディングを含む。

【0 1 0 7】

20

以下、1つの目標は、タイプ  $t(w)$  をどのように効率的にエンコードするかを見つける出すことである。何らかの所定のタイプ  $t$  は、

【数 1 3】

$$t = \left[ \frac{k_1}{n}, \dots, \frac{k_n}{n} : \sum_{i=1}^m k_i = n \right] \quad (11)$$

【0 1 0 8】

30

として規定できることに注目されたい。

【0 1 0 9】

ここで、 $K_1$  から  $K_m$  は、サンプルの総数  $n$  を与えられた、可能性あるタイプ  $t$  の数を示す。それゆえ、タイプ  $t$  を持つ可能性あるシーケンスの総数は、

【数 1 4】

$$\xi(t) = \binom{n}{k_1, \dots, k_m} \quad (12)$$

40

【0 1 1 0】

により与えられることができる。

【0 1 1 1】

ここで、 $(t)$  は、母集団  $t$  を持つシンボルの可能性ある配置の総数である。

【0 1 1 2】

可能性あるタイプの総数は、本質的に、 $k_1 + \dots + k_m = n$  になるように、すべての整数  $k_1, \dots, k_m$  の数であり、これは、マルチセット係数により与えられる。

【数 1 5】

$$M(m, n) = \binom{n+m-1}{m-1} \quad (13)$$

【0 1 1 3】

タイプ  $t$  の任意のサンプル  $w$  の実現値の確率は、 $P(t)$  により示されてもよい。可能性あるサンプルのような  $(t)$  が存在し、それらがすべて同一の確率を有していることから、それなら：

10

【数 1 6】

$$\begin{aligned} P(t) &= \xi(t) P(w : t(w) = t) \\ &= \binom{n}{k_1, \dots, k_m} p_1^{k_1} \dots p_m^{k_m} \end{aligned} \quad (14)$$

20

【0 1 1 4】

である。

【0 1 1 5】

この密度  $P(t)$  は、タイプの分布として参照されてもよい。これは、明らかに、

【数 1 7】

$$P(t^*) = P(t : k_i = np_i) = \binom{n}{np_1, \dots, np_m} p_1^{np_1} \dots p_m^{np_m} \quad (15)$$

30

【0 1 1 6】

における最大値（モード）を持つ多項分布である。

【0 1 1 7】

タイプの分布のエントロピーが、（集合プロパティにより）後続する。

【数 1 8】

40

$$H(P(t)) = -\sum_t P(t) \log P(t) \sim \log(P(t^*)) = \frac{m-1}{2} \log n + O(1) \quad (16)$$

【0 1 1 8】

長さ  $n$  のサンプル  $w$  が与えられると、ユニバーサルエンコーダのタスクは、その最も悪いケースの平均冗長：

【数 1 9】

$$R^*(n) = \sup_P \left[ \sum_{|w|=n} P(w) |f(w)| - nH(P) \right] \quad (17)$$

$$\geq \sup_P \sum_{|w|=n} P(w) \log \frac{P(w)}{P_f(w)} = n \sup_P D(P//P_f) \quad (18)$$

10

【0 1 1 9】

が最小になるように、コード  $f(w)$  (または、同等に、その誘導分布  $P_f(w)$ ) を設計することである。方程式 17 および 18 は、ユニバーサルコーディングにより取り扱われる問題を描いており、シーケンスが与えられ、平均コード長と  $n * H(P)$  との間の差分が、すべての可能性ある入力分布に対して最小であるコード長が求められる。つまり、最小の最も悪いケースのコード長は、予め分布を知らずに求められる。

【0 1 2 0】

同一のタイプのサンプルの確率が同一であることから、コード誘導分布  $P_f(w)$  は、このプロパティを保つことが予期され、 $P_f(w)$  は、

20

【数 2 0】

$$P_f(w) = \frac{P_f(w : t(w) = t)}{\xi(t)} \quad (19)$$

【0 1 2 1】

として規定できる。ここで、 $P_f(t)$  は、タイプ  $t(w)$  の確率であり、 $\xi(t)$  は、同一タイプ  $t(w)$  内のシーケンスの総数である。タイプ  $t(w)$  に割り振られたコードの確率  $P_f$  は、したがって、

30

【数 2 1】

$$P_f(t) = \xi(t) P_f(w : t(w) = t) \quad (20)$$

【0 1 2 2】

は、タイプのコード誘導分布である、として規定できる。

40

【0 1 2 3】

このような分解を方程式 18 に当てはめて、タイプを超えて (個々のサンプルの代わりに) 加法を変更することにより、平均冗長  $R^*(n)$  は、

【数 2 2】

$$R^*(n) \geq \sup_P \sum_{|w \in A^n} P(w) \log \frac{P(w)}{P_f(w)} \quad (21.1)$$

$$= \sup_P \left[ \sum_t \sum_{w:t(w)=t} P(w) \log \frac{P(t)}{P_f(t)} \right] \quad (21.2)$$

$$= \sup_P \left[ \sum_t P(t) \log \frac{P(t)}{P_f(t)} \right] \quad (21.3)$$

$$= \sup_P D(P(t) \| P_f(t)) \quad (21.4)$$

【0 1 2 4】

として、規定できる。ここで、“sup”は、上限（supremum）演算子であり、値は、それが少なくともそのセットのうちの任意の要素と同じくらいの大きさである場合、セットに関する上限である。これらの方程式は、タイプのコーディングの問題は、最小冗長ユニバーサル（minimum redundancy universal）コーディングの問題と同等であることを意味する。

【0 1 2 5】

結果として、タイプの無損失コーディングの問題は、タイプのKT推定分布を使用することにより、漸近的に最適に解決できる。

【数 2 3】

$$P_{KT}(t) = \xi(t) P_{KT}(w : t(w) = t) \quad (22.1)$$

$$= \binom{n}{k_1, \dots, k_m} \Gamma\left(\frac{m}{2}\right) \frac{\prod_{i=1}^m \Gamma(k_i + \frac{1}{2})}{\pi^{\frac{m}{2}} \Gamma(n + \frac{m}{2})} \quad (22.2)$$

【0 1 2 6】

この方程式 22.2 に基づいて、単一の母集団を持つタイプ（0 回のカウントを持つもの）がそのピークになる一方で、近くの均一な母集団を持つタイプが推定密度の谷に落ちることは明白になる。

【0 1 2 7】

図 5 は、さまざまなガウシアンスケール空間にわたる例示的なキーポイント分布を図示しているチャート 180 の線図である。グラフ 180 に対するデータは、トレーニング画像のセット、すなわち、700 × 1000 ピクセルの解像度を持つ 105 個の画像から発生された。各画像に対して、データは、1.0（グラフ 186）、0.75（グラフ 188）、0.5（グラフ 190）、および、0.25（グラフ 192）のスケールに対して

10

20

30

40

50

収集された。チャート 180 は、スケール空間にわたる、キーポイントの典型的な分布を示す。この観測に基づいて、制御ユニット 20 (図 1) は、画像の倍率の推定を可能にする統計を計算してもよい。

【0128】

グラフ 180 の X 軸 184 は、データのこの特定の例示的なセットに対するガウシアンスケール空間 (GSS) 中の、可能性あるスケールインデックスおよび対応するスケールを表す。Y 軸 182 は、この例に対するキーポイントの数を表す。

【0129】

自然な画像に対して、示されているガウシアンスケール空間にしたがう、スケールパラメータに関して検出されたキーポイントの分布が存在してもよい。より高いスケールにおいて検出されたキーポイントは、一般的により大きな特徴に対応する。画像解像度が増加するにつれて (例えば、ユーザがターゲットから遠くに離れていくにつれて)、より細かいディテールは低減され、より高いスケール (粗いディテールに対応する) において検出されたディテールは、より低いスケール (より細かなディテールに対応する) に移動する。キーポイントのうちのほとんどは、それゆえ、より低いスケールに属する。ピン中心としてスケール空間レベルにより構築されたキーポイントのヒストグラムは、画像サイズが増加するにつれて、ヒストグラムのエリアのうちのほとんどは、より低いピン中心に集中してもよいという特性を有していてもよい。実質上、これは、スケール空間におけるキーポイントの分布はテクスチャに依存しており、そして、オブジェクトからのカメラの距離にしたがって変化することを意味している。

【0130】

所定のシーンにおいて、多数のオブジェクトが現れる可能性があり、それぞれは、カメラからのそれ独自の距離において、それゆえ、一意的なスケールにある。ユーザが対話することを望むオブジェクトに関係するスケールを適切に評価するために、いくつかの粗いセグメント化を実行して、画像中の他のものからオブジェクトを分離して、オブジェクトのスケールについての局所的な推論を行ってもよい。これは、例えば、ダイナミックウィンドウ、または、オブジェクト領域間の境界をおおよそ描く他のアルゴリズムを使用することにより達成できる。

【0131】

サーチするデータベースの最低オクターブを決定するのに使用できるカットオフスケールを計算するために、スケール推定ユニット 24 は、キーポイント  $P(s)$  のヒストグラムにより、スケール空間レベル  $s$  にわたるキーポイントの分布を推定してもよい。次に、スケール推定ユニット 24 は、キーポイントのうちのおよそ 90% が、カットオフスケール  $s_c$  までのスケールレベル内に保たれるように、 $s_c$  を推定してもよい。形式的には、スケール推定ユニット 24 は、以下の公式 (23) を計算してもよい。

【数 24】

$$\sum_{s=1}^{s_c} P(s) = 0.9 * \sum_{s=1}^S P(s) \quad (23)$$

【0132】

その後、スケール推定ユニット 24 は、追跡したスケール空間カットオフレベルを画像倍率  $Y$  にマッピングする関数を推定する。すなわち、スケール推定ユニット 24 は、

10

20

30

40

【数 2 5】

$$Y = f(s_c) \quad (24)$$

【0 1 3 3】

を計算してもよい。

【0 1 3 4】

ここで、 $Y$  は画像倍率であり、 $s_c$  はカットオフスケール空間レベルである。関数を知るために、スケール推定ユニット 2 4 は、最も反復されたパス、例えば、グラフ 1 8 6、1 8 8、1 9 0 または 1 9 2 の 1 つから、 $Y$  および  $s_c$  データポイントを拾ってもよい。

10

【0 1 3 5】

一般的に、画像解像度が増加するにつれて、 $s_c$  はスケール空間においてさらに上に移動し、逆もまた同じである。グラフ 1 8 0 はまた、この例示的なデータセットにおける、1 . 0 (ポイント 1 9 8)、0 . 7 5 (ポイント 1 9 6)、および、0 . 5 (ポイント 1 9 4) のスケールに対する、例示的なカットオフスケール (最低オクターブに対応する) を図示している。この例から、解像度が減少するにつれて、カットオフスケールは、左に向かって移動することが分かる。つまり、解像度が減少するにつれて、カットオフスケールが減少する。

【0 1 3 6】

20

図 6 は、インクリメントな特徴記述子抽出を実行するための例示的な方法を図示しているフローチャートである。説明の目的のために、図 6 の方法は、クライアントデバイス 1 0 のコンポーネントに関して記述されている。しかしながら、図 6 の方法は、他のデバイスにより実行されてもよく、または、さまざまなデバイスにより共同で実行されてもよいことを理解すべきである。例えば、サーバデバイス 5 0 (図 1) の制御ユニット 6 0 は、図 6 の方法を実行するように構成されていてもよい。同様に、追加のステップが実行されてもよく、または、特定のステップが省略され、方法のステップが、本開示の技術から逸脱することなく、異なる順序で (または、並行して) 実行されてもよい。

【0 1 3 7】

図 6 の例では、初めに、制御ユニット 2 0 は、オブジェクトの (問い合わせ画像とも呼ばれる) 画像を受信する (2 0 0)。例えば、制御ユニット 2 0 は、カメラ 1 2 から画像を受信してもよい。代替的に、制御ユニット 2 0 は、クライアントデバイス 1 0 のメモリから、記憶された画像を取り出してもよい。制御ユニット 2 0 の特徴抽出ユニット 2 2 はその後、特徴記述子を抽出する画像に対する第 1 のオクターブを選択してもよい (2 0 2)。

30

【0 1 3 8】

特徴抽出ユニット 2 2 は、本開示において論じた、技術の組み合わせを含む、さまざまな技術のうちの任意のものを使用して、第 1 のオクターブを選択してもよい。例えば、特徴抽出ユニット 2 2 は、特徴抽出ユニット 2 2 が第 1 のオクターブを選択するために使用してもよいセンサ 1 4 から、データを受信してもよい。特徴抽出ユニット 2 2 は、例えば、クライアントデバイス 1 0 のロケーションを示す、センサ 1 4 の GPS ユニットからのロケーションデータを受信してもよい。特徴抽出ユニット 2 2 は、ロケーションデータを使用して、クライアントデバイス 1 0 の近くのオブジェクトのロケーションを決定してもよい。特徴抽出ユニット 2 2 は、画像中のオブジェクトに対するスケールを予想するために、クライアントデバイス 1 0 の近くのオブジェクトを使用してもよい。

40

【0 1 3 9】

例えば、GPS データが、クライアントデバイス 1 0 がフランスのパリのエッフェル塔付近にあると示す場合に、特徴抽出ユニット 2 2 は、画像中のオブジェクトの可能性あるスケールは比較的大きいと決定してもよい。この決定は、画像がエッフェル塔のものであるという (コンフィグレーションデータを通しての) 仮定に基づいていてもよく、この仮

50

定は、特徴記述子抽出および比較を通して確定されてもよい。一方で、GPSデータが、クライアントデバイス10がフランスのパリのルーブル美術館付近もしくは中にあると示す場合に、特徴抽出ユニット22は、画像中のオブジェクトのスケールが比較的小さいと決定してもよい。この決定は、例えば、画像がルーブル美術館内に保存されている作品のものであるという（コンフィグレーションデータを通しての）仮定に基づいていてもよい。

#### 【0140】

GPSユニットからのロケーションデータ、または、他のタイプの感知されたデータに加えて、または、代替実施形態において、制御ユニット20は、深度推定ユニット26によって計算された深度マップに基づいて第1のオクターブを選択してもよい。深度推定ユニット26は、深度マップを計算するさまざまな技術のうちの任意のものを使用してもよい。例えば、深度推定ユニット26は、ステレオカメラによって取り込まれた2つ以上の画像のピクセル間での、水平差異（すなわち視差）を解析してもよい。これらの画像のうちの1つは、問い合わせ画像に対応してもよい。問い合わせ画像中のオブジェクトに対する深度に基づいて、制御ユニット20は、オブジェクトに対するスケールを決定し、このスケールは第1のオクターブに対応してもよい。第1のオクターブに影響し得る他のファクタは、例えば、画像中でオブジェクトが存在するロクセル、ワイヤレスネットワークが利用可能であるか否か（例えば、ワイヤレスアクセスポイントが領域内にあるか否か）、トラッカーが見失って以来の以前の新しい問い合わせ画像に対してスケールが決まっていたか否か、または、他のファクタを含む。

#### 【0141】

第1のオクターブを決定した後に、特徴抽出ユニット22は、第1のオクターブに対するキーポイントを抽出してもよい（204）。上記で論じたように、特徴抽出ユニット22は、上で選択した第1のオクターブに対応するサイズに、問い合わせ画像の解像度をスケールリングすることにより、問い合わせ画像から、第1のオクターブに対する画像を計算してもよい。特徴抽出ユニット22は、その後、種々の第1のオクターブ画像にガウシアンぼかしフィルタを適用して、キーポイントをそこから抽出するガウシアン差分ピラミッドを計算してもよい。キーポイントを使用して、特徴抽出ユニット22は、問い合わせ画像の第1のオクターブに対して、1つ以上の特徴記述子を抽出してもよい（206）。

#### 【0142】

特徴抽出ユニット22は、その後、特徴記述子を使用して、問い合わせ画像中のオブジェクトのアイデンティティを決定するように試行してもよい（208）。いくつかの例では、制御ユニット20は、遠隔オブジェクトデータベース52のすべてまたは一部を取り出して、データをローカルオブジェクトデータベース30として局所的に記憶してもよい。他の例では、制御ユニット20は、抽出された特徴記述子のうちの1つ以上をネットワーク40を介してサーバデバイス50に送ってもよい。特徴マッチングユニット62は、クライアントデバイス10から受信した特徴記述子のいくつかまたはすべてが、遠隔オブジェクトデータベース52のオブジェクトにマッチするか否かを決定してもよい。特徴マッチングユニット62は、特徴記述子に対する1つ以上のマッチを決定してもよい。特徴マッチングユニット62が、サーチする最低オクターブの指示を受信していた場合に、特徴マッチングユニット62は、最低オクターブにおける、または、最低オクターブを上回るオクターブであるが、特定された最低オクターブを下回るオクターブに対する特徴記述子のみをサーチしてもよい。信頼評価ユニット64は、その後、マッチがどのくらいうまく特徴記述子に適合するかを示す、マッチのそれぞれに対する信頼値を決定してもよい。制御ユニット60は、その後、決定されたマッチのそれぞれに対する識別子およびそれぞれの信頼値を含む情報を、ネットワーク40を介してクライアントデバイス10送ってもよい。

#### 【0143】

サーバデバイス50から情報を受信した後に、クライアントデバイス10の制御ユニット20は、オブジェクトのさまざまな決定されたアイデンティティに対する信頼値のうち

のいずれかのものがしきい値を超えるか否かを決定してもよい(210)。例えば、制御ユニット20は、最も大きな信頼値がしきい値を超えるか否かを決定してもよい。最も大きい信頼値がしきい値を超えない場合(210の“いいえ”の枝)、制御ユニット20は、下記に記述するように、特徴記述子を抽出するための次のオクターブとともに、いくつかの例において問い合わせられているデータベースの(最低オクターブとも呼ばれる)カットオフスケールを決定してもよい。

#### 【0144】

スケール推定ユニット24は、いずれのスケールが最も厳密にキーポイント分布に適するかを決定し、その後、最良に適合しているスケールが、オブジェクトのスケールである可能性が高いことを決定してもよい(212)。スケールのこの近似に基づいて、特徴抽出ユニット22は、抽出された特徴記述子と比較するための、データベース中の最低オクターブを選択してもよい(214)。特徴抽出ユニット22はまた、キーポイントを抽出する次のオクターブを選択してもよい(216)。特徴抽出ユニット22は、その後、この次に決定されたオクターブに対するキーポイントを抽出し(218)、抽出したキーポイントを使用して、現在のオクターブに対する特徴記述子を抽出してもよい(206)。特徴抽出ユニット22は、その後、再び、抽出された特徴記述子を使用して、オブジェクトアイデンティティを決定してもよい(208)。特に、特徴抽出ユニット22は、抽出された特徴記述子を使用して選択した最低オクターブにおける、または、選択した最低オクターブを上回るデータベースのオクターブにおいて、データベースにオブジェクトをサーチさせ、オブジェクトの決定されたアイデンティティにおける新たな信頼値を受け取らせてもよい。

#### 【0145】

決定されたオブジェクトのアイデンティティ中の信頼値がしきい値を超える場合(210の“はい”の枝)、特徴抽出ユニット22は、画像中のオブジェクトに対するアイデンティティ情報を出力してもよい(218)。いくつかの例では、特徴抽出ユニット22は、制御ユニット20により実行されるウェブブラウザと対話して、ウェブブラウザに、サーチ用語またはキーワードとして決定されたアイデンティティを使用して、インターネットベースのサーチエンジンにサーチ問い合わせをサブミットさせてもよい。サーチは、例えば、オブジェクトの別の画像、(例えば、オブジェクトがランドマークである場合に)オブジェクトのロケーション情報、(例えば、オブジェクトまたはそのレプリカが購入可能である場合に)プライス情報、オブジェクトに対するより多くの情報を持つウェブアドレスへのリンク(例えば、ユニフォームリソースロケータ(URL))、または、これらのような他の情報のような、画像中のオブジェクトに対するさまざまな情報をウェブブラウザに表示させてもよい。

#### 【0146】

この方法で、図6は、受信した画像の第1のオクターブのぼかし画像の第1のセットから、1つ以上のキーポイントの第1のセットを抽出することと、キーポイントの第1のセットに対する1つ以上の記述子の第1のセットを計算することと、記述子の第1のセットにより特徴記述子データベースに問い合わせることにより生じた結果に対する信頼値を受信し、結果は、受信した画像中のオブジェクトのアイデンティティを記述する情報を含むことと、信頼値が信頼しきい値を超えないときに、受信した画像の第2のオクターブのぼかし画像の第2のセットから、1つ以上のキーポイントの第2のセットを抽出することとを含む方法の例を表す。

#### 【0147】

1つ以上の例において、記述される機能は、ハードウェア、ソフトウェア、ファームウェア、あるいはそれらの任意の組み合わせで実現されてもよい。ソフトウェアで実現された場合、機能は、1つ以上の命令またはコードとしてコンピュータ読取可能媒体上に記憶されてもよく、あるいは、1つ以上の命令またはコードとしてコンピュータ読取可能媒体へ送信されてもよく、ハードウェアベースの処理ユニットにより実行されてもよい。コンピュータ読取可能媒体は、データ記憶媒体のような有形の媒体に対応するコンピュータ読

取可能記憶媒体、あるいは、例えば、通信プロトコルにしたがって、1つの場所から別の場所へのコンピュータプログラムの伝送を促進する任意の媒体を含む通信媒体を含んでもよい。この方法では、コンピュータ読取可能媒体は、一般的に、(1)一時的ではない有形のコンピュータ読取可能記憶媒体、あるいは、(2)信号または搬送波のような通信媒体に対応してもよい。データ記憶媒体は、本開示中で説明した技術のインプリメンテーションのための命令、コードおよび/またはデータ構造を取り出すために、1つ以上のコンピュータ、または、1つ以上のプロセッサによってアクセスできる任意の利用可能な媒体であってもよい。コンピュータプログラムプロダクトは、コンピュータ読取可能媒体を含んでもよい。

【0148】

限定ではなく例として、有形のコンピュータ読取可能記憶媒体は、RAM、ROM、EEPROM、CD-ROMまたは他の光ディスク記憶装置、磁気ディスク記憶装置または他の磁気記憶デバイス、フラッシュメモリ、あるいは、コンピュータによってアクセスでき、命令もしくはデータ構造の形式で所望のプログラムコードを記憶するために使用可能である他の任意の媒体を備えることができる。また、任意の接続は、コンピュータ読取可能媒体と適切に呼ばれる。例えば、命令が、ウェブサイトから、サーバから、あるいは、同軸ケーブル、光ファイバケーブル、撚り対、デジタル加入者線(DSL)、または、赤外線、無線、マイクロ波のようなワイヤレス技術を使用する他の遠隔ソースから送信された場合、同軸ケーブル、光ファイバケーブル、撚り対、DSL、あるいは、赤外線、無線およびマイクロ波のようなワイヤレス技術は、媒体の定義に含まれる。しかしながら、コンピュータ読取可能記憶媒体およびデータ記憶媒体は、接続、搬送波、信号または他の一時的な媒体を含まず、代わりに、一時的ではない、有形の記憶媒体に向けられていることを理解すべきである。ここで使用するようなディスク(diskおよびdisc)は、コンパクトディスク(CD)、レーザディスク(登録商標)、光ディスク、デジタル多用途ディスク(DVD)、フロッピー(登録商標)ディスク、および、ブルーレイ(登録商標)ディスクを含んでいる。ここで、ディスク(disk)が通常、データを磁氣的に再生する一方で、ディスク(disc)はデータをレーザによって光学的に再生する。上記のものの組み合わせも、コンピュータ読取可能媒体の範囲内に含められるべきである。

【0149】

命令は、1つ以上のデジタル信号プロセッサ(DSP)、汎用マイクロプロセッサ、特定用途向け集積回路(ASIC)、フィールドプログラマブル論理アレイ(FPGA)、または、他の同等な集積または離散論理回路のような、1つ以上のプロセッサにより実行されてもよい。したがって、ここで使用される“プロセッサ”という用語は、前述の構造のうちの任意のもの、または、ここで記述した技術のインプリメンテーションに適切な他の何らかの構造のことを指してもよい。加えて、いくつかの態様では、ここで記述した機能性は、エンコーディングおよびデコーディングのために構成されている、あるいは、組み合わせたコーデック中に組み込まれている、専用のハードウェアおよび/またはソフトウェアモジュール内で提供されてもよい。また、技術は、1つ以上の回路または論理エレメント中で完全に実現できる。

【0150】

本開示の技術は、ワイヤレスハンドセット、集積回路(IC)、または、ICのセット(例えば、チップセット)を含む、幅広いさまざまなデバイスまたは装置で実現してもよい。開示した技術を実行するように構成されているデバイスの機能的な態様を強調するために、本開示中では、さまざまなコンポーネント、モジュールまたはユニットが記述されているが、必ずしも異なるハードウェアユニットによる実現を必要とするわけではない。むしろ、上述したように、さまざまなユニットは、適切なソフトウェアおよび/またはファームウェアと併用して、コーデックハードウェアユニット中で組み合わされてもよく、または、上述したような1つ以上のプロセッサを含む、相互動作可能なハードウェアユニットを集めたものによって提供されてもよい。

以下に、本願出願時の特許請求の範囲に記載された発明を付記する。

[ 1 ] 画像中のオブジェクトに対するアイデンティティを決定する方法において、受信した画像の第 1 のオクターブのぼかし画像の第 1 のセットから、1 つ以上のキーポイントの第 1 のセットを抽出することと、

前記キーポイントの第 1 のセットに対する 1 つ以上の記述子の第 1 のセットを計算することと、

前記記述子の第 1 のセットにより特徴記述子データベースに問い合わせることにより生じた結果に対する信頼値を受信し、前記結果は、前記受信した画像中のオブジェクトのアイデンティティを記述する情報を含むことと、

前記信頼値が信頼しきい値を超えないときに、前記受信した画像の第 2 のオクターブのぼかし画像の第 2 のセットから、1 つ以上のキーポイントの第 2 のセットを抽出することを含む方法。

10

[ 2 ] 前記信頼値が前記信頼しきい値を超えていないことを決定することと、

前記決定に基づいて、前記第 2 のオクターブのぼかし画像の第 2 のセットから、前記キーポイントの第 2 のセットを抽出することと、

前記キーポイントの第 2 のセットから、1 つ以上の記述子の第 2 のセットを計算することと、

前記記述子の第 1 のセットおよび前記記述子の第 2 のセットを含む記述子のセットにより、前記特徴記述子データベースに問い合わせることとをさらに含む [ 1 ] 記載の方法。

[ 3 ] 前記受信した画像中の前記オブジェクトに対するスケール値を推定することと、

前記推定したスケールに基づいて、前記特徴記述子データベースの最低オクターブを選択することと、

20

前記特徴記述子データベース中の前記選択した最低オクターブにおいて、または、前記特徴記述子データベース中の前記選択した最低オクターブより上で、特徴記述子をサーチするための前記特徴記述子データベースのサーチを起こすように、前記特徴記述子データベース中の最低オクターブの指示を提供することとをさらに含む [ 1 ] 記載の方法。

[ 4 ] 前記スケール値を推定することは、

スケール空間にわたる前記キーポイントの第 1 のセットの分布を解析することと、

前記キーポイントのうちのおおよそ 90 % がスケールレベルより下に属するように、前記分布中の前記スケールレベルに対応するカットオフポイントを決定することと、

前記決定したカットオフポイントの関数として、前記スケール値を推定することを含む [ 3 ] 記載の方法。

30

[ 5 ] 前記オブジェクトに対する深度情報を取得することをさらに含み、

前記スケール値を推定することは、前記オブジェクトに対する深度情報の少なくとも一部に基づいて、前記スケール値を推定することを含む [ 3 ] 記載の方法。

[ 6 ] 前記受信した画像を取り込んだカメラに関係する 1 つ以上のセンサからのセンサ情報を解析することと、

前記センサ情報の解析に基づいて、前記第 1 のオクターブおよび前記第 2 のオクターブのうちの少なくとも 1 つに対するオクターブレベルを決定することとをさらに含む [ 1 ] 記載の方法。

[ 7 ] 前記センサ情報を解析することは、グローバルポジショニングシステム (GPS) 情報を解析して、前記受信した画像が取り込まれたときに前記カメラが屋外環境に位置付けられていたか否かを決定することを含む [ 6 ] 記載の方法。

40

[ 8 ] 前記センサ情報を解析することは、GPS 情報を解析して、前記受信した画像が取り込まれたときの、前記カメラに比較的近いオブジェクトのロケーションを決定することと、前記オブジェクトに対する記述データから、前記オブジェクトに対するサイズを決定することを含む [ 6 ] 記載の方法。

[ 9 ] 前記センサ情報を解析することは、前記受信した画像が取り込まれたときに、前記カメラを備えているデバイスがワイヤレスネットワークに通信可能に結合されていたか否かを示すネットワークデータに基づいて、前記カメラが屋内環境に位置付けられていたか否かを決定することを含む [ 6 ] 記載の方法。

50

[ 1 0 ] 前記センサ情報を解析することは、前記受信した画像が取り込まれたときの、前記受信した画像中の 1 つ以上のオブジェクトと前記カメラとの間の距離を示す深度情報を計算することを含む [ 6 ] 記載の方法。

[ 1 1 ] 前記センサ情報を解析することは、アクティブプロービングセンサにより提供されるデータを使用して、前記オブジェクトに対する深度値を推定することを含む [ 6 ] 記載の方法。

[ 1 2 ] 前記 1 つ以上の記述子をサーバに送って、前記サーバに、前記 1 つ以上の記述子を使用して前記特徴記述子データベースに問い合わせさせることをさらに含み、

前記信頼値を受信することは、前記問い合わせに 응답して前記サーバから前記信頼値を受信することを含む [ 1 ] 記載の方法。

[ 1 3 ] 画像中のオブジェクトに対するアイデンティティを決定するための装置において、

プロセッサを具備し、

前記プロセッサは、

受信した画像の第 1 のオクターブのぼかし画像の第 1 のセットから、1 つ以上のキーポイントの第 1 のセットを抽出し、

前記キーポイントの第 1 のセットに対する 1 つ以上の記述子の第 1 のセットを計算し、

前記記述子の第 1 のセットにより特徴記述子データベースに問い合わせることにより生じた結果に対する信頼値を受信し、前記結果は、前記受信した画像中のオブジェクトのアイデンティティを記述する情報を含み、

前記信頼値が信頼しきい値を超えないときに、前記受信した画像の第 2 のオクターブのぼかし画像の第 2 のセットから、1 つ以上のキーポイントの第 2 のセットを抽出するように構成されている装置。

[ 1 4 ] 前記画像を取り込み、前記画像を前記プロセッサに提供するように構成されているカメラをさらに具備する [ 1 3 ] 記載の装置。

[ 1 5 ] 前記プロセッサが、前記信頼値が前記信頼しきい値を超えないと決定するときに、前記プロセッサは、

前記決定に基づいて、前記第 2 のオクターブのぼかし画像の第 2 のセットから、前記キーポイントの第 2 のセットを抽出し、

前記キーポイントの第 2 のセットから、1 つ以上の記述子の第 2 のセットを計算し、

前記記述子の第 1 のセットおよび前記記述子の第 2 のセットを含む記述子のセットにより、前記特徴記述子データベースに問い合わせるようにさらに構成されている [ 1 3 ] 記載の装置。

[ 1 6 ] 前記プロセッサは、

前記受信した画像中の前記オブジェクトに対するスケール値を推定し、

前記推定したスケールに基づいて、前記特徴記述子データベースの最低オクターブを選択し、

前記特徴記述子データベース中の記述子をサーチするようにさらに構成され、

前記サーチした記述子は、前記選択した最低オクターブにおける、または、前記選択した最低オクターブより上の、前記特徴記述子データベース中のオクターブに対応する [ 1 5 ] 記載の装置。

[ 1 7 ] 前記スケール値を推定するために、前記プロセッサは、

スケール空間にわたる前記キーポイントの第 1 のセットの分布を解析し、

前記キーポイントのうちのおおよそ 9 0 % がスケールレベルより下に属するように、前記分布中の前記スケールレベルに対応するカットオフポイントを決定し、

前記決定したカットオフポイントの関数として、前記スケール値を計算するように構成されている [ 1 6 ] 記載の装置。

[ 1 8 ] 前記オブジェクトに対する深度情報を取得するように構成されている深度推定ユニットをさらに具備し、

前記プロセッサは、前記オブジェクトに対する深度情報の少なくとも一部に基づいて、

10

20

30

40

50

前記スケール値を推定するように構成されている〔１６〕記載の装置。

〔１９〕前記プロセッサは、

前記受信した画像を取り込んだカメラに関係する１つ以上のセンサからのセンサ情報を解析し、

前記センサ情報の解析に基づいて、前記第１のオクターブおよび前記第２のオクターブのうちの少なくとも１つに対するオクターブレベルを決定するようにさらに構成されている〔１３〕記載の装置。

〔２０〕前記装置に対するロケーション情報を決定するように構成されているグローバルポジショニングシステム（ＧＰＳ）ユニットをさらに具備し、

前記センサ情報を解析するために、前記プロセッサは、前記ＧＰＳユニットにより決定されたロケーション情報に基づいて、前記受信した画像が取り込まれたときに前記カメラが屋外環境に位置付けられていたか否かを決定するように構成されている〔１９〕記載の装置。

〔２１〕前記装置に関連するロケーション情報を決定するように構成されているＧＰＳユニットをさらに具備し、

前記センサ情報を解析するために、前記プロセッサは、前記ＧＰＳユニットにより決定されたロケーション情報を解析して、前記受信した画像が取り込まれたときの、前記カメラに比較的近いオブジェクトのロケーションを決定し、前記オブジェクトに対する記述データから、前記オブジェクトに対するサイズを決定するように構成されている〔１９〕記載の装置。

〔２２〕ワイヤレスネットワークインターフェースをさらに具備し、

前記センサ情報を解析するために、前記プロセッサは、前記受信した画像が取り込まれたときに、前記ワイヤレスネットワークインターフェースがワイヤレスネットワークに通信可能に結合されていたか否かを示すネットワークデータに基づいて、前記カメラが屋内環境に位置付けられていたか否かを決定するように構成されている〔１９〕記載の装置。

〔２３〕前記受信した画像を取り込んだカメラを含む少なくとも２つのカメラを備えるカメラアレイをさらに具備し、

前記センサ情報を解析するために、前記プロセッサは、前記受信した画像が取り込まれたときの、前記受信した画像中の１つ以上のオブジェクトと前記カメラとの間の距離を示す、前記カメラアレイにより取り込まれた画像からの深度情報を計算するように構成されている〔１９〕記載の装置。

〔２４〕前記１つ以上のセンサは、前記オブジェクトに対する深度値を推定するように構成されているアクティブブロービングセンサを備える〔１９〕記載の装置。

〔２５〕前記プロセッサは、前記１つ以上の記述子をサーバに送って、前記サーバに、前記１つ以上の記述子を使用して前記特徴記述子データベースに問い合わせさせ、前記問い合わせにตอบสนองして前記サーバから前記信頼値を受信するように構成されている〔１３〕記載の装置。

〔２６〕前記装置は、

集積回路と、

マイクロプロセッサと、

前記プロセッサを含むワイヤレス通信デバイスとのうちの少なくとも１つを具備する〔１３〕記載の装置。

〔２７〕画像中のオブジェクトに対するアイデンティティを決定する装置において、

受信した画像の第１のオクターブのぼかし画像の第１のセットから、１つ以上のキーポイントの第１のセットを抽出する手段と、

前記キーポイントの第１のセットに対する１つ以上の記述子の第１のセットを計算する手段と、

前記記述子の第１のセットにより特徴記述子データベースに問い合わせることにより生じた結果に対する信頼値を受信し、前記結果は、前記受信した画像中のオブジェクトのアイデンティティを記述する情報を含む手段と、

10

20

30

40

50

前記信頼値が信頼しきい値を超えないときに、前記受信した画像の第2のオクターブのぼかし画像の第2のセットから、1つ以上のキーポイントの第2のセットを抽出する手段とを具備する装置。

[ 2 8 ] 前記信頼値が前記信頼しきい値を超えていないことを決定する手段と、  
前記決定に基づいて、前記第2のオクターブのぼかし画像の第2のセットから、前記  
キーポイントの第2のセットを抽出する手段と、  
前記キーポイントの第2のセットから、1つ以上の記述子の第2のセットを計算する手  
段と、  
前記記述子の第1のセットおよび前記記述子の第2のセットを含む記述子のセットによ  
り、前記特徴記述子データベースに問い合わせる手段とをさらに具備する [ 2 7 ] 記載の  
装置。

10

[ 2 9 ] 前記受信した画像中の前記オブジェクトに対するスケール値を推定する手段と  
、  
前記推定したスケールに基づいて、前記特徴記述子データベースの最低オクターブを選  
択する手段と、

前記特徴記述子データベース中の前記選択した最低オクターブにおいて、または、前記  
特徴記述子データベース中の前記選択した最低オクターブより上で、特徴記述子をサー  
チするための前記特徴記述子データベースのサーチを起こすように、前記特徴記述子デー  
タベース中の最低オクターブの指示を提供する手段とをさらに具備する [ 2 8 ] 記載の装  
置。

20

[ 3 0 ] 前記スケール値を推定する手段は、  
スケール空間にわたる前記キーポイントの第1のセットの分布を解析する手段と、  
前記キーポイントのうちのおおよそ90%がスケールレベルより下に属するように、前  
記分布中の前記スケールレベルに対応するカットオフポイントを決定する手段と、  
前記決定したカットオフポイントの関数として、前記スケール値を推定する手段とを具  
備する [ 2 9 ] 記載の装置。

[ 3 1 ] 前記オブジェクトに対する深度情報を取得する手段をさらに具備し、  
前記スケール値を推定する手段は、前記オブジェクトに対する深度情報の少なくとも一  
部に基づいて、前記スケール値を推定する手段を備える [ 2 9 ] 記載の装置。

[ 3 2 ] 前記受信した画像を取り込んだカメラに関係する1つ以上のセンサからのセン  
サ情報を解析する手段と、

30

前記センサ情報の解析に基づいて、前記第1のオクターブおよび前記第2のオクターブ  
のうちの少なくとも1つに対するオクターブレベルを決定する手段とをさらに具備する [  
2 7 ] 記載の装置。

[ 3 3 ] グローバルポジショニングシステム ( G P S ) 情報を受信する手段をさらに具  
備し、

前記センサ情報を解析する手段は、前記 G P S 情報を解析して、前記受信した画像が取  
り込まれたときに前記カメラが屋外環境に位置付けられていたか否かを決定する手段を備  
える [ 3 2 ] 記載の装置。

[ 3 4 ] G P S 情報を受信する手段をさらに具備し、

40

前記センサ情報を解析する手段は、  
前記 G P S 情報を解析して、前記受信した画像が取り込まれたときの、前記カメラに比  
較的近いオブジェクトのロケーションを決定する手段と、

前記オブジェクトに対する記述データから、前記オブジェクトに対するサイズを決定す  
る手段とを備える [ 3 2 ] 記載の装置。

[ 3 5 ] 少なくとも1つのワイヤレスネットワークプロトコルを介して通信する手段を  
さらに具備し、

前記センサ情報を解析する手段は、前記受信した画像が取り込まれたときに、前記ワイ  
ヤレスネットワークプロトコルを介して通信する手段がワイヤレスネットワークに通信可  
能に結合されていたか否かを示すネットワークデータに基づいて、前記カメラが屋内環境

50

に位置付けられていたか否かを決定する手段を備える [ 3 2 ] 記載の装置。

[ 3 6 ] シーンのうちの 2 つ以上の画像を取り込む手段をさらに具備し、

前記 2 つ以上の画像のうちの 1 つは、前記受信した画像を含み、

前記センサ情報を解析する手段は、前記受信した画像が取り込まれたときの、前記受信した画像中の 1 つ以上のオブジェクトと前記カメラとの間の距離を示す、前記シーンのうちの 2 つの画像からの深度情報を計算する手段を備える [ 3 2 ] 記載の装置。

[ 3 7 ] 前記センサ情報を解析する手段は、アクティブブローピングセンサにより提供されるデータを使用して、前記オブジェクトに対する深度値を推定する手段を備える [ 3 2 ] 記載の装置。

[ 3 8 ] 前記 1 つ以上の記述子をサーバに送って、前記サーバに、前記 1 つ以上の記述子を使用して前記特徴記述子データベースに問い合わせさせる手段をさらに具備し、

前記信頼値を受信する手段は、前記問い合わせに回答して前記サーバから前記信頼値を受信する手段を備える [ 2 7 ] 記載の装置。

[ 3 9 ] その上に記憶されている命令を有するコンピュータ読取可能媒体を含むコンピュータプログラムプロダクトにおいて、

前記命令は、実行されるとき、

受信した画像の第 1 のオクターブのぼかし画像の第 1 のセットから、1 つ以上のキーポイントの第 1 のセットをプロセッサに抽出させ、

前記キーポイントの第 1 のセットに対する 1 つ以上の記述子の第 1 のセットを前記プロセッサに計算させ、

前記記述子の第 1 のセットにより特徴記述子データベースに問い合わせることにより生じた結果に対する信頼値を前記プロセッサに受信させ、前記結果は、前記受信した画像中のオブジェクトのアイデンティティを記述する情報を含み、

前記信頼値が信頼しきい値を超えないときに、前記受信した画像の第 2 のオクターブのぼかし画像の第 2 のセットから、1 つ以上のキーポイントの第 2 のセットを前記プロセッサに抽出させるコンピュータプログラムプロダクト。

[ 4 0 ] 前記信頼値が前記信頼しきい値を超えていないことを前記プロセッサに決定させ、

前記決定に基づいて、前記第 2 のオクターブのぼかし画像の第 2 のセットから、前記キーポイントの第 2 のセットを前記プロセッサに抽出させ、

前記キーポイントの第 2 のセットから、1 つ以上の記述子の第 2 のセットを前記プロセッサに計算させ、

前記記述子の第 1 のセットおよび前記記述子の第 2 のセットを含む記述子のセットにより、前記特徴記述子データベースに前記プロセッサに問い合わせさせる命令をさらに含む [ 3 9 ] 記載のコンピュータプログラムプロダクト。

[ 4 1 ] 前記受信した画像中の前記オブジェクトに対するスケール値を前記プロセッサに推定させ、

前記推定したスケールに基づいて、前記特徴記述子データベースの最低オクターブを前記プロセッサに選択させ、

前記特徴記述子データベース中の前記選択した最低オクターブにおいて、または、前記特徴記述子データベース中の前記選択した最低オクターブより上で、特徴記述子をサーチするための前記特徴記述子データベースのサーチを起こすように、前記特徴記述子データベース中の最低オクターブの指示を前記プロセッサに提供させる命令をさらに含む [ 3 9 ] 記載のコンピュータプログラムプロダクト。

[ 4 2 ] 前記スケール値を前記プロセッサに推定させる命令は、

スケール空間にわたる前記キーポイントの第 1 のセットの分布を前記プロセッサに解析させ、

前記キーポイントのうちのおおよそ 9 0 % がスケールレベルより下に属するように、前記分布中の前記スケールレベルに対応するカットオフポイントを前記プロセッサに決定させ、

10

20

30

40

50

前記決定したカットオフポイントの関数として、前記スケール値を前記プロセッサに推定させる [ 4 1 ] 記載のコンピュータプログラムプロダクト。

[ 4 3 ] 前記オブジェクトに対する深度情報を前記プロセッサに取得させる命令をさらに含み、

前記スケール値を前記プロセッサに推定させる命令は、前記オブジェクトに対する深度情報の少なくとも一部に基づいて、前記スケール値を前記プロセッサに推定させる命令を有する [ 4 1 ] 記載のコンピュータプログラムプロダクト。

[ 4 4 ] 前記受信した画像を取り込んだカメラに関係する 1 つ以上のセンサからのセンサ情報を前記プロセッサに解析させ、

前記センサ情報の解析に基づいて、前記第 1 のオクタブおよび前記第 2 のオクタブのうちの少なくとも 1 つに対するオクタブレベルを前記プロセッサに決定させる命令をさらに含む [ 3 9 ] 記載のコンピュータプログラムプロダクト。

[ 4 5 ] 前記センサ情報を前記プロセッサに解析させる命令は、グローバルポジショニングシステム ( G P S ) 情報を前記プロセッサに解析させ、前記受信した画像が取り込まれたときに前記カメラが屋外環境に位置付けられていたか否かを決定させる命令を含む [ 4 4 ] 記載のコンピュータプログラムプロダクト。

[ 4 6 ] 前記センサ情報を前記プロセッサに解析させる命令は、G P S 情報を前記プロセッサに解析させ、前記受信した画像が取り込まれたときの、前記カメラに比較的近いオブジェクトのロケーションを決定させ、前記オブジェクトに対する記述データから、前記オブジェクトに対するサイズを前記プロセッサに決定させる命令を含む [ 4 4 ] 記載のコンピュータプログラムプロダクト。

[ 4 7 ] 前記センサ情報を前記プロセッサに解析させる命令は、前記受信した画像が取り込まれたときに、前記カメラを備えているデバイスがワイヤレスネットワークに通信可能に結合されていたか否かを示すネットワークデータに基づいて、前記カメラが屋内環境に位置付けられていたか否かを前記プロセッサに決定させる命令を含む [ 4 4 ] 記載のコンピュータプログラムプロダクト。

[ 4 8 ] 前記センサ情報を前記プロセッサに解析させる命令は、前記受信した画像が取り込まれたときの、前記受信した画像中の 1 つ以上のオブジェクトと前記カメラとの間の距離を示す深度情報を前記プロセッサに計算させる命令を含む [ 4 4 ] 記載のコンピュータプログラムプロダクト。

[ 4 9 ] 前記センサ情報を前記プロセッサに解析させる命令は、アクティブプロービングセンサにより提供されるデータを使用して、前記オブジェクトに対する深度値を前記プロセッサに推定させる命令を含む [ 4 4 ] 記載のコンピュータプログラムプロダクト。

[ 5 0 ] 前記 1 つ以上の記述子をサーバに送って、前記サーバに、前記 1 つ以上の記述子を使用して前記特徴記述子データベースに問い合わせさせるように前記プロセッサにさせる命令をさらに含み、

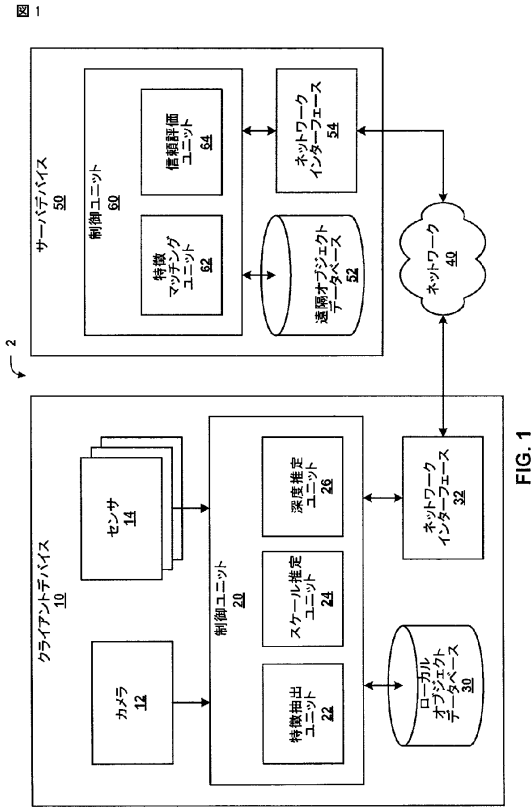
前記信頼値を前記プロセッサに受信させる命令は、前記問い合わせに応答して前記サーバから前記信頼値を前記プロセッサに受信させる命令を含む [ 3 9 ] 記載のコンピュータプログラムプロダクト。

10

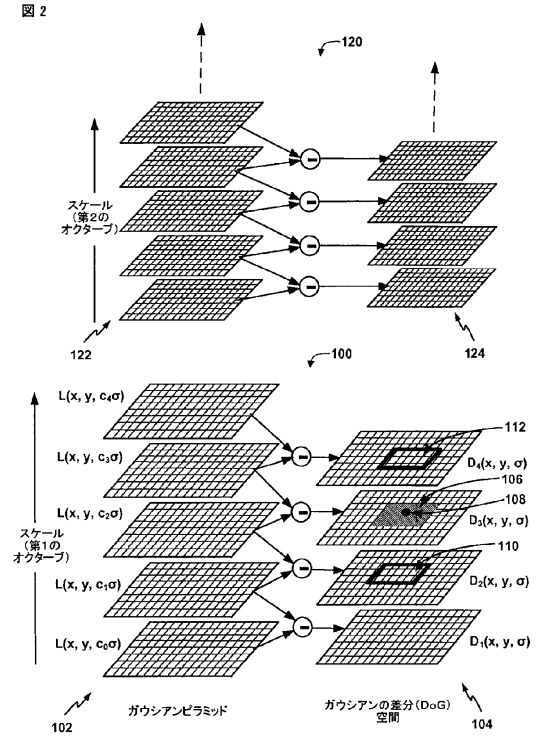
20

30

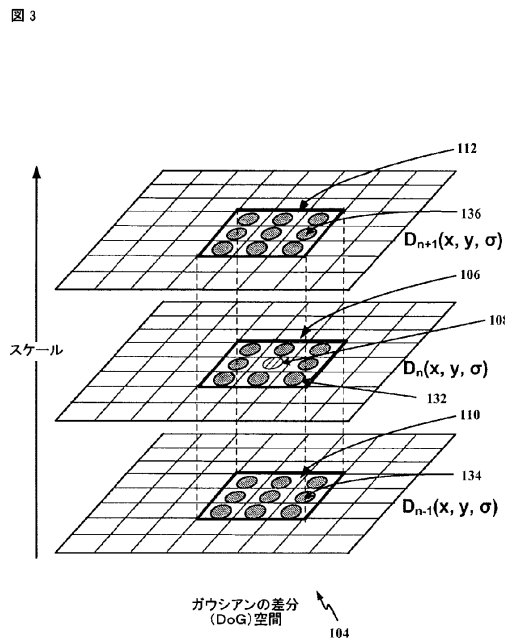
【図 1】



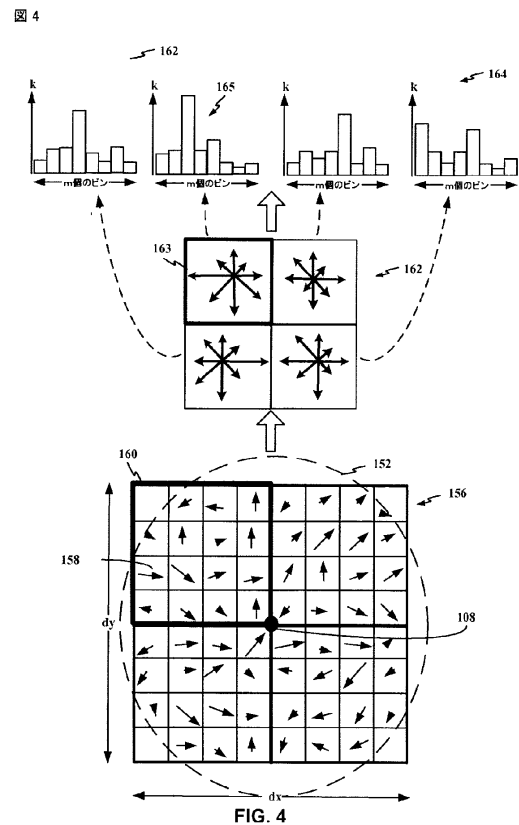
【図 2】



【図 3】



【図 4】



【図 5】

図 5

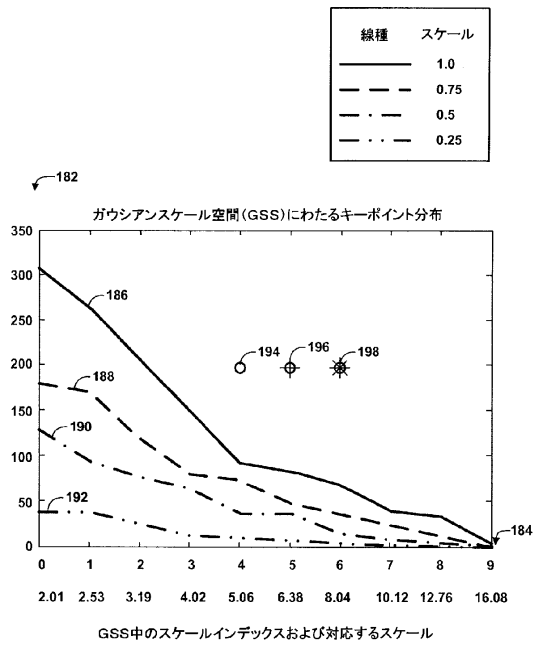


FIG. 5

【図 6】

図 6

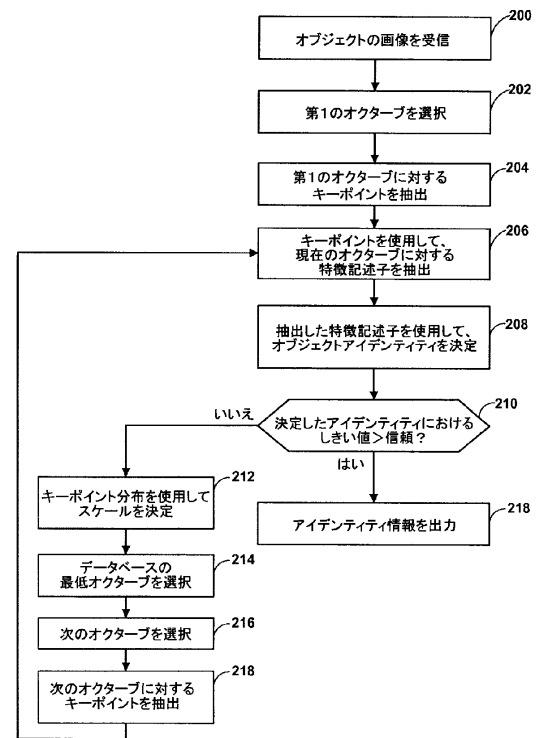


FIG. 6

## フロントページの続き

- (74)代理人 100095441  
弁理士 白根 俊郎
- (74)代理人 100075672  
弁理士 峰 隆司
- (74)代理人 100119976  
弁理士 幸長 保次郎
- (74)代理人 100153051  
弁理士 河野 直樹
- (74)代理人 100140176  
弁理士 砂川 克
- (74)代理人 100158805  
弁理士 井関 守三
- (74)代理人 100172580  
弁理士 赤穂 隆雄
- (74)代理人 100179062  
弁理士 井上 正
- (74)代理人 100124394  
弁理士 佐藤 立志
- (74)代理人 100112807  
弁理士 岡田 貴志
- (74)代理人 100111073  
弁理士 堀内 美保子
- (74)代理人 100134290  
弁理士 竹内 将訓
- (72)発明者 バヘティ、パワン・クマー  
アメリカ合衆国、カリフォルニア州 9 2 1 2 1、サン・ディエゴ、モアハウス・ドライブ 5 7  
7 5
- (72)発明者 バドダディー、サンディーブ  
アメリカ合衆国、カリフォルニア州 9 2 1 2 1、サン・ディエゴ、モアハウス・ドライブ 5 7  
7 5
- (72)発明者 スワミナサン、アシュウィン  
アメリカ合衆国、カリフォルニア州 9 2 1 2 1、サン・ディエゴ、モアハウス・ドライブ 5 7  
7 5
- (72)発明者 レズニク、ユリー  
アメリカ合衆国、カリフォルニア州 9 2 1 2 1、サン・ディエゴ、モアハウス・ドライブ 5 7  
7 5
- (72)発明者 ハムシシー、オヌル・シー  
アメリカ合衆国、カリフォルニア州 9 2 1 2 1、サン・ディエゴ、モアハウス・ドライブ 5 7  
7 5
- (72)発明者 チャリ、ムラリ・ラマスワミイ  
アメリカ合衆国、カリフォルニア州 9 2 1 2 1、サン・ディエゴ、モアハウス・ドライブ 5 7  
7 5
- (72)発明者 ホン、ジョン・エイチ  
アメリカ合衆国、カリフォルニア州 9 2 1 2 1、サン・ディエゴ、モアハウス・ドライブ 5 7  
7 5
- (72)発明者 リー、チョン・ウク  
アメリカ合衆国、カリフォルニア州 9 2 1 2 1、サン・ディエゴ、モアハウス・ドライブ 5 7  
7 5

審査官 新井 則和

(56)参考文献 特開2010-009599(JP,A)  
特開2009-223527(JP,A)  
特開2009-157821(JP,A)  
特開2001-238177(JP,A)  
特開2009-266003(JP,A)  
特表2011-512520(JP,A)  
特開平08-048199(JP,A)  
米国特許出願公開第2007/0092143(US,A1)

(58)調査した分野(Int.Cl., DB名)  
G06T 1/00 - 7/60