

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6902811号  
(P6902811)

(45) 発行日 令和3年7月14日(2021.7.14)

(24) 登録日 令和3年6月24日(2021.6.24)

(51) Int.Cl. F I  
**G06T 7/00 (2017.01)** G O 6 T 7/00 3 5 0 C  
**G06T 7/593 (2017.01)** G O 6 T 7/593

請求項の数 10 外国語出願 (全 33 頁)

<p>(21) 出願番号 特願2020-205055 (P2020-205055)                  (22) 出願日 令和2年12月10日 (2020.12.10)                  (65) 公開番号 特開2021-96850 (P2021-96850A)                  (43) 公開日 令和3年6月24日 (2021.6.24)                  審査請求日 令和2年12月10日 (2020.12.10)                  (31) 優先権主張番号 201911281475.1                  (32) 優先日 令和1年12月13日 (2019.12.13)                  (33) 優先権主張国・地域又は機関                  中国 (CN)</p> <p>早期審査対象出願</p>	<p>(73) 特許権者 519155675                  ネクストヴィピーユー (シャンハイ) カン                  パニー リミテッド                  NEXTVPU (SHANGHAI) CO                  ., LTD.                  中華人民共和国 201203 シャンハ                  イ, プドオン ニュー ディストリクト,                  ロンドン ロード, #3000, ビル                  ディング 1, セクション エー, ルーム                  607                  RM607, SECTION A, BLD                  G 1, #3000 LONGDONG                  RD, PUDONG NEW DISTR                  ICT, SHANGHAI, 201203                  P. R. CHINA</p> <p style="text-align: right;">最終頁に続く</p>
---	---

(54) 【発明の名称】 視差推定システムと方法、電子機器及びコンピュータ可読記憶媒体

(57) 【特許請求の範囲】

【請求項1】

特徴抽出ネットワーク及び視差生成ネットワークを含む視差推定システムであって、  
 前記特徴抽出ネットワークは、画像対における各画像に対して特徴抽出を行い、抽出した画像特徴を前記視差生成ネットワークに出力するように構成され、

前記視差生成ネットワークは、抽出された前記画像特徴に基づいて、カスケード接続される多段視差処理を行い、サイズが順次増大する複数の視差図を取得するように構成され、

前記多段視差処理のうちの第一段の視差処理の入力は、該段の視差処理に対応するサイズを有する複数の画像特徴を含み、

前記多段視差処理のうちの前記第一段の視差処理以外の各段の視差処理の入力は、該段の視差処理に対応するサイズを有する1つ又は複数の画像特徴、及び1つ前の段の視差処理により生成される視差図を含み、

抽出した前記画像特徴は、N種類のサイズの画像特徴を含み、前記Nは、2以上の正の整数であり、

前記多段視差処理は、N+1段の視差処理を含み、

前記視差生成ネットワークは、

前記多段視差処理のうちの前記第一段の視差処理において、前記N種類のサイズの画像特徴のうち最小サイズの画像特徴の少なくとも一部に基づいて、前記最小サイズを有する初期視差図を生成し；及び

前記第一段の視差処理以外のN段の視差処理において、サイズの从小から大への順序に従って、順に、前記N種類のサイズの画像特徴のうちの、対応するサイズを有する画像特徴の少なくとも一部に基づいて、1つ前の段の視差処理により生成される視差図に対して視差最適化処理を行い、サイズが順次増大するN個の最適化視差図を取得し、そして、取得したN個の最適化視差図を、サイズが順次増大する前記複数の視差図とするように構成され、

前記N個の最適化視差図のサイズは、それぞれ、前記N種類のサイズに対応する、システム。

【請求項2】

視差推定方法であって、

画像対における各画像に対して特徴抽出を行い；及び

抽出した画像特徴に基づいて、カスケード接続される多段視差処理を行い、サイズが順次増大する複数の視差図を取得することを含み、

前記多段視差処理のうちの第一段の視差処理の入力は、該段の視差処理に対応するサイズを有する複数の画像特徴を含み、

前記多段視差処理のうちの前記第一段の視差処理以外の各段の視差処理の入力は、該段の視差処理に対応するサイズを有する1つ又は複数の画像特徴、及び1つ前の段の視差処理により生成される視差図を含み、

抽出した前記画像特徴は、N種類のサイズの画像特徴を含み、前記Nは、2以上の正の整数であり、

前記多段視差処理は、N+1段の視差処理を含み、

前記の、抽出した画像特徴に基づいて、カスケード接続される多段視差処理を行い、サイズが順次増大する複数の視差図を取得することは、

前記多段視差処理のうちの前記第一段の視差処理において、前記N種類のサイズの画像特徴のうちの最小サイズの画像特徴の少なくとも一部に基づいて、前記最小サイズを有する初期視差図を生成し；及び

前記第一段の視差処理以外のN段の視差処理において、サイズの从小から大への順序に従って、順に、前記N種類のサイズの画像特徴のうちの、対応するサイズを有する画像特徴の少なくとも一部に基づいて、1つ前の段の視差処理により生成される視差図に対して視差最適化処理を行い、サイズが順次増大するN個の最適化視差図を取得し、そして、取得したN個の最適化視差図を、サイズが順次増大する前記複数の視差図とすることを含み

前記N個の最適化視差図のサイズは、それぞれ、前記N種類のサイズに対応する、方法

【請求項3】

請求項2に記載の視差推定方法であって、

前記第一段の視差処理以外の前記N段の視差処理のうちの各段の視差処理において、以下の方式により、前記N種類のサイズの画像特徴のうちの、対応するサイズを有する画像特徴の少なくとも一部に基づいて、1つ前の段の視差処理により生成される視差図に対して視差最適化処理を行い、前記対応するサイズを有する最適化視差図を生成し、

即ち、対応するサイズを有する画像特徴の少なくとも一部に基づいて、1つ前の段の視差処理により生成される視差図に対して残差計算処理を行い、前記対応するサイズを有する残差図を取得し、そして、前記対応するサイズを有する残差図と、前記1つ前の段の視差処理により生成される視差図との組み合わせを行って前記対応するサイズを有する最適化視差図を得る、方法。

【請求項4】

請求項2に記載の視差推定方法であって、

前記多段視差処理のうちの前記第一段の視差処理以外の各段の視差処理において、1つ前の段の視差処理により生成される視差図に対して視差最適化処理を行う前に、

前記1つ前の段の視差処理により生成される視差図のサイズが現在の段の視差処理に対応するサイズよりも小さいことに応じて、前記1つ前の段の視差処理により生成される視

10

20

30

40

50

差図を、現在の段の視差処理に対応するサイズにアップサンプリングすることをさらに含む、方法。

【請求項 5】

請求項 2 に記載の視差推定方法であって、

前記 N 種類のサイズの画像特徴のうちの最小サイズの画像特徴は、前記画像対における第一画像の少なくとも 1 種類の画像特徴及び第二画像の少なくとも 1 種類の画像特徴を含み、

前記 N 種類のサイズの画像特徴のうちの各種類の最小サイズでない画像特徴は、前記画像対における前記第一画像の少なくとも 1 種類の画像特徴及び / 又は前記第二画像の少なくとも 1 種類の画像特徴を含む、方法。

10

【請求項 6】

請求項 2 に記載の視差推定方法であって、

各最適化視差図の生成が基づく画像特徴は、前記画像対における少なくとも 1 つの画像のエッジ特徴、及び / 又は、前記画像対における少なくとも 1 つの画像の画像自身に基づく特徴を含む、方法。

【請求項 7】

請求項 2 ~ 6 のうちの任意の 1 項に記載の視差推定方法であって、

ターゲット装置のパフォーマンスに基づいて、サイズが順次増大する前記複数の視差図のうちから、サイズが前記ターゲット装置のパフォーマンスとマッチした視差図を、前記ターゲット装置に提供する視差図として選択することをさらに含む、方法。

20

【請求項 8】

電子装置であって、

処理器；及び

プログラムを記憶している記憶器を含み、

前記プログラムは指令を含み、前記指令は、前記処理器により実行されるときに、前記処理器に、請求項 2 ~ 7 のうちの任意の 1 項に記載の方法を実行させる、電子装置。

【請求項 9】

プログラムを記憶しているコンピュータ可読記憶媒体であって、

前記プログラムは指令を含み、前記指令は、電子装置の処理器により実行されるときに、前記電子装置に、請求項 2 ~ 7 のうちの任意の 1 項に記載の方法を実行させる、コンピュータ可読記憶媒体。

30

【請求項 10】

コンピュータに、請求項 2 ~ 7 のうちの任意の 1 項に記載の方法を実行させるためのプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本開示は、コンピュータビジョン技術分野に関し、特に、視差推定システムと方法、電子機器及びコンピュータ可読記憶媒体に関する。

【背景技術】

40

【0002】

関連技術では、コンピュータビジョン技術を採用して、同一シーンの 2 つの異なる画角 (angle of view) の画像における各対のマッチした画素点の間の視差を得ることで視差図 (disparity map) を取得し、また、前記視差図に基づいて前記シーンの深度情報を取得することができる。ここで、前記深度情報は、3次元再構成、自動運転、障害物検知などの様々な分野で用いることができる。例示として、コンピュータビジョン技術を採用して視差を得る方法は、ローカルエリアマッチング法、グローバル最適化法、セミグローバルマッチング法、畳み込みニューラルネットワークに基づく方法などを含んでも良い。

【0003】

なお、この部分で説明した方法は、必ずしも以前に考案又は採用された方法ではない。

50

特に明記されない限り、この部分に記載されている方法のいずれかが、この部分に含まれているという理由だけで従来技術と見なされると想定すべきではない。同様に、特に明記されない限り、この部分で言及されている問題が、従来技術で認識されていると見なされるべきではない。

【発明の概要】

【発明が解決しようとする課題】

【0004】

本開示の目的は、視差推定システム、方法、電子機器及びコンピュータ可読記憶媒体を提供することにある。

【課題を解決するための手段】

【0005】

本開示の一側面によれば、視差推定システムが提供され、それは、特徴抽出ネットワーク及び視差生成ネットワークを含み、前記特徴抽出ネットワークは、画像対における各画像に対して特徴抽出を行い、抽出した画像特徴を前記視差生成ネットワークに出力するように構成され、前記視差生成ネットワークは、抽出された前記画像特徴に基づいて、カスケード接続される多段視差処理を行い、サイズが順次増大する複数の視差図を取得し、ここで、前記多段視差処理のうちの第一段の視差処理の入力は、該段の視差処理に対応するサイズを有する複数の画像特徴を含み、前記多段視差処理のうちの前記第一段の視差処理以外の各段の視差処理の入力は、該段の視差処理に対応するサイズを有する1つ又は複数の画像特徴、及び1つ前の段の視差処理により生成される視差図を含む。

【0006】

本開示の他の側面によれば、視差推定方法が提供され、それは、画像対における各画像に対して特徴抽出を行い；及び、抽出した画像特徴に基づいて、カスケード接続される多段視差処理を行い、サイズが順次増大する複数の視差図を取得することを含み、ここで、前記多段視差処理のうちの第一段の視差処理の入力は、該段の視差処理に対応するサイズを有する複数の画像特徴を含み、前記多段視差処理のうちの前記第一段の視差処理以外の各段の視差処理の入力は、該段の視差処理に対応するサイズを有する1つ又は複数の画像特徴、及び1つ前の段の視差処理により生成される視差図を含む。

【0007】

本開示の他の側面によれば、電子機器が提供され、それは、処理器；及びプログラムを記憶している記憶器を含み、前記プログラムは指令を含み、前記指令は、前記処理器により実行されるときに、前記処理器に、本開示に記載の方法を実現させる。

【0008】

本開示の他の側面によれば、プログラムを記憶しているコンピュータ可読記憶媒体が提供され、前記プログラムは指令を含み、前記指令は、電子機器の処理器により実行されるときに、前記電子機器に、本開示に記載の方法を実現させる。

【0009】

本開示の上述の特徴及び利点をより明らかにするために、以下、実施例を挙げて添付した図面を参照することにより詳細に説明する。

【図面の簡単な説明】

【0010】

添付した図面は、実施例を例示的に示し且つ明細書の一部を構成し、明細書の文字説明とともに実施例の例示的な実施方式を説明するために使用される。また、図示した実施例は、例示のみを目的としており、特許請求の範囲を限定するものではない。なお、すべての図面では、同じ参照記号は、類似しているが必ずしも同じでない要素を指す。

【図1】本開示の例示的な実施例による視差推定システムの構成図である。

【図2】本開示の例示的な実施例による画像の基礎構造特徴を示す図である。

【図3】本開示の例示的な実施例による画像の語義的特徴を示す図である。

【図4】本開示の例示的な実施例による画像のエッジ特徴を示す図である。

【図5】本開示の例示的な実施例による視差推定システムの1つの可能な全体構成図であ

10

20

30

40

50

る。

【図6】本開示の例示的な実施例による視差推定システムのもう1つの可能な全体構成図である。

【図7A】本開示の例示的な実施例によるネットワーク訓練に基づく参照画像を示す図である。

【図7B】本開示の例示的な実施例によるネットワーク訓練に基づく参照画像に対応する真値視差図を示す図である。

【図8】本開示の例示的な実施例による訓練後の視差推定システムを採用して、図7Aに示す参照画像に対して、カスケード接続される多段視差処理を行うことにより得られた、右から左へサイズが順次増大する複数の視差図を示す図である。

【図9】本開示の例示的な実施例による視差推定方法のフローチャートである。

【図10】本開示の例示的な実施例による視差推定方法における多段視差処理のフローチャートである。

【図11】本開示の例示的な実施例に適用され得る例示的な計算装置の構成図である。

【発明を実施するための形態】

【0011】

本開示では、特に明記されない限り、用語「第一」、「第二」などが、各種の要素を説明するために使用されるが、これらの要素の位置関係、タイミング関係又は重要性関係を限定することを意図するものではない。このような用語は、1つの要素をもう1つの要素と区別するためだけに用いられる。幾つかの例示において、第一要素及び第二要素が該要素の同一実例に指向することができるが、幾つかの場合、コンテキストに基づいて、それらは、異なる実例を指すこともできる。

【0012】

本開示では、様々な前記例示についての説明に使用する用語は、特定の例示を説明することのみを目的としており、限定を行うことを目的としない。コンテキストに明記されない限り、要素の数が特に限定されない場合、該要素は1つであっても良く、複数であっても良い。また、本開示に使用する用語「及び/又は」は、リストされている項目のうちの任意の1つ及びすべての可能な組み合わせをカバーする。

【0013】

関連技術では、コンピュータビジョン技術を採用して、同一シーンの2つの異なる画角の画像における各対(ペア)のマッチした画素点の間の視差を得ることで視差図を取得し、また、前記視差図に基づいて前記シーンの深度情報を取得することができる。ここで、前記深度情報は、3次元再構成、自動運転、障害物検知などの様々な分野で用いることができる。例示として、コンピュータビジョン技術を採用して視差を得る方法は、ローカルエリアマッチング法、グローバル最適化法、セミグローバルマッチング法、畳み込みニューラルネットワークなどに基づくニューラルネットワークの方法などを含んでも良い。

【0014】

ローカルエリアマッチング法は、主に、マッチングコスト計算、コストアグリゲーション、視差計算、視差最適化などのステップを含み、比較的高い速度及び比較的低いエネルギー消費を有するが、そのアルゴリズムの効果及びアルゴリズムのパラメータ(例えば、マッチングウィンドウのサイズなど)と関係があり、複雑シーンのニーズを満足することが困難である。ローカルエリアマッチング法に比較して、グローバル最適化法は、より良いマッチング精度(accuracy)を有し、それは、平滑項に対して仮定を行い、そして、視差計算というステレオマッチング問題をエネルギー最適化問題に変え、また、大部分のグローバル最適化法は、コストアグリゲーションステップをスキップし、マッチングコスト及び平滑項を考慮することで、グローバルな点についてエネルギー関数を提供し、そして、エネルギー関数を最小化することで、視差を取得する。しかし、ローカルエリアマッチング法に比べて、グローバル最適化法の計算量はより大きく、エネルギー消費はより高い。セミグローバルマッチング法は、マッチング精度と演算速度とのある程度のバランスをとることができ、グローバルアルゴリズムがグローバルな点に対して最適化を行うことと

10

20

30

40

50

は異なり、それは、各点のエネルギー関数を複数の方向の経路に分け、各経路に対して値を求める必要があり、その後、すべての経路の値の和を算出することで、該点のエネルギーを得ることができ、ここで、各経路に対して値を求めるには、動的計画法を採用することができる。しかし、セミグローバルマッチング法は、ローカルエリアマッチング法に比較して、計算量及びエネルギー消費もともに比較的高い。CNN (Convolutional Neural Network、畳み込みニューラルネットワーク) などに基づくニューラルネットワークの方法は、視差ネットワークを構築することでより大きい感知領域を得ることができるため、画像のテキスチャー無し領域でより良い視差予測能力を有する。しかし、その計算量は、CNNなどのニューラルネットワークのパラメータ及び画像のサイズと関係があり、ネットワークパラメータが複雑であり及び画像のサイズが大きい場合、メモリ消費は大きくなり、実行速度は低い。

10

**【 0 0 1 5 】**

本開示は、新しい視差推定システムを提供し、それは、抽出した画像対における各画像の画像特徴に基づいて、カスケード接続される多段視差処理を行い、サイズが順次増大する複数の視差図を取得することができる。ここで、前記多段視差処理のうちの第一段の視差処理の入力は、該段の視差処理に対応するサイズを有する複数の画像特徴を含んでも良く、前記多段視差処理のうちの第一段の視差処理以外の各段の視差処理の入力は、該段の視差処理に対応するサイズを有する1つ又は複数の画像特徴、及び1つ前の段の視差処理により生成される視差図を含んでも良い。換言すると、抽出した画像特徴に対して、カスケード接続される多段視差処理を行い、ここで、各段の視差処理の入力は、該段の視差処理

20

**【 0 0 1 6 】**

図1は本開示の例示的な実施例における視差推定システムの構成図である。図1に示すように、前記視差推定システム100は例えば、特徴抽出ネットワーク200及び視差生成ネットワーク300を含み、特徴抽出ネットワーク200は、画像対における各画像に対して特徴抽出

30

**【 0 0 1 7 】**

図1の構成図に示す視差推定システムによれば、抽出した画像対における各画像の画像特徴に基づいて、カスケード接続される多段視差処理を行い、サイズが順次増大する複数の視差図を取得することができ、ここで、各段の視差処理の入力は、該段の視差処理に対応するサイズを有する画像特徴を含んでも良い。これにより、一回だけで異なるパフォーマンスを有する又は異なる精度を要する複数のターゲット装置の使用に供する複数の異なるサイズの視差図を得ることができる。このようにして、異なるターゲット装置の精度及び速度への要求を満足することができるだけでなく、視差推定システムの柔軟性及び適用性を向上させることもできる。

40

**【 0 0 1 8 】**

本開示では、前記画像対は、マルチカメラにより採取した、同一シーンに対する画像対であっても良い。前記画像対における各画像は、サイズが一致しているが、対応する画角が異なる。もちろん、前記画像対は、他の方法で取得した(例えば、他の装置から取得し

50

た)、要求を満たす画像であっても良い。また、前記画像対における各画像は、グレースケール画像又はカラー画像であっても良い。

【0019】

本開示では、前記マルチカメラとは、2つ、3つ、ひいてはより多くのカメラヘッドが配置される、静的又は動的な画像キャプチャーを行うことができるカメラを指す。それは、配置される複数のカメラヘッドによって異なる画角又は範囲のシーンをカバーすることで、カメラの、シーンにおける物体への検知能力を向上させることができる。2つのカメラヘッド(例えば、左、右カメラヘッド)が配置される双目カメラを例にとれば、任意のシーンについて、前記双目カメラは、配置される2つのカメラヘッドによって該シーンの2つの、サイズが一致しているが、対応する画角が異なる画像(例えば、左画像及び右画像)を得ることができる。ここで、前記2つの画像が構成する画像対は、該シーンにおける物体の、前記2つの画像において対応する画素点の間の変位(例えば、水平変位)即ち視差を確定することで、物体の距離などの深度情報を確定するために用いられ得る。

10

【0020】

また、本開示では、視差推定システム100及びマルチカメラは、互いに独立したものであっても良い。換言すると、視差推定システム100は、その中に含まれる特徴抽出ネットワーク200により、マルチカメラが採取した、同一シーンの画像対における各画像に対して特徴抽出を行い、また、その中に含まれる視差生成ネットワーク300により、抽出された画像特徴に対して、カスケード接続される多段視差処理を行い、サイズが順次増大する複数の視差図を取得することができる。代替案として、マルチカメラは、前記視差推定システム100の一部とされても良い。換言すれば、前記視差推定システム100は、特徴抽出ネットワーク200及び視差生成ネットワーク300以外に、前記マルチカメラをさらに含んでも良い。

20

【0021】

幾つかの実施例によれば、前記視差推定システム100の特徴抽出ネットワーク200が抽出した前記画像対における各画像の画像特徴は、少なくとも次のような特徴のうちの1つ又は複数を含んでも良く、即ち、基礎構造特徴、語義的特徴、エッジ特徴、テキストチャー特徴、色特徴、物体形状特徴、及び画像自身に基づく特徴である。

【0022】

図2は、(a)、(b)及び(c)の3つの画像(例えば、グレースケール画像であるが、カラー画像であっても良い)をもって、本開示の例示的な実施例において抽出し得る1種類の画像の基礎構造特徴を示す図である。図2から分かるように、前記基礎構造特徴とは、画像を反映するための各種類の微細構造の特徴を指す。

30

【0023】

図3は、(a)、(b)、(c)及び(d)の4つの画像(例えば、グレースケール画像であるが、カラー画像であっても良い)をもって、本開示の例示的な実施例において抽出し得る1種類の画像の語義的特徴を示す図である。図3から分かるように、前記語義的特徴とは、画像の中の異なる対象又は異なるタイプの対象を区別し得る特徴を指す。また、前記語義的特徴に基づいて、画像の曖昧領域(例えば、大面積平坦領域)の視差確定の正確性を向上させることができる。

40

【0024】

図4は、(a)及び(b)の2つの画像(例えば、グレースケール画像であるが、カラー画像であっても良い)をもって、本開示の例示的な実施例において抽出し得る1種類の画像のエッジ特徴を示す図である。図4から分かるように、前記エッジ特徴とは、画像の中の物体又は領域の境界情報を反映し得る特徴を指す。

【0025】

また、図示しないが、前記テキストチャー特徴、色特徴及び物体形状特徴とは、それぞれ、画像を反映し得るテキストチャー、色、及び画像に含まれる物体の形状の特徴を指す。前記の画像自身に基づく特徴とは、前記画像自身を指しても良く、前記画像自身に対して所定の係数又は比率のアップサンプリング又はダウンサンプリングを行うことで得られた画

50

像を指しても良い。なお、前記アップサンプリング又はダウンサンプリングの係数又は比率は例えば、2、3又は他の1以上の数値であっても良い。

【0026】

幾つかの実施例によれば、本開示では、前記画像自身に基づく特徴以外に、他の各種類の画像特徴は、すべて、対応する特徴抽出サブネットワークにより、対応する画像に対して抽出を行うことにより得ることができる。これにより、画像特徴抽出の効率を向上させることができるため、視差推定の効率を向上させることができる。また、視差推定の正確性を向上させるために、少なくとも、基礎構造特徴、語義的特徴及びエッジ特徴の3つの異なる面から、画像に対して特徴抽出を行っても良い。

【0027】

例えば、幾つかの実施例によれば、図5又は図6を参照するに、図5及び図6は、本開示の例示的な実施例による視差推定システム100の可能な全体構成図であり、前記特徴抽出ネットワーク200は、複数の、画像の異なる特徴をそれぞれ抽出するための特徴抽出サブネットワークを含んでも良く、前記複数の特徴抽出サブネットワークは、少なくとも、画像の基礎構造特徴を抽出するための基礎構造特徴サブネットワーク201、画像の語義的特徴を抽出するための語義的特徴サブネットワーク202、及び画像のエッジ特徴を抽出するためのエッジ特徴サブネットワーク203を含んでも良い。

【0028】

幾つかの実施例によれば、前記基礎構造特徴サブネットワーク201は、VGG (very deep convolutional networks for large scale image recognition、大規模の画像識別のための超深層畳み込みネットワーク)、ResNet (Residual Network、残差ネットワーク)などの任意の、画像の基礎構造特徴を抽出し得るネットワークを採用しても良い。前記語義的特徴サブネットワーク202は、DeepLabV3+ (encoder-decoder with atrous separable convolution for semantic image segmentation、語義的な画像分割のための分離可能な畳み込みエンコーダー・デコーダー)などの任意の、画像の語義的特徴を抽出し得るネットワークを採用しても良い。前記エッジ特徴サブネットワーク203は、HED (holistically-nested edge detection) ネットワークなどの任意の、画像のエッジ特徴を抽出し得るネットワークを採用しても良い。幾つかの実施方式によれば、HEDネットワークは、VGGを主幹ネットワークとして採用しても良く、また、前記エッジ特徴サブネットワーク203がHEDネットワークを採用するときに、前記基礎構造特徴サブネットワーク201は、前記エッジ特徴サブネットワーク203と同じVGGネットワークを採用することで、前記特徴抽出ネットワークの構造を簡略化する効果を達成することもできる。

【0029】

幾つかの実施例によれば、前記特徴抽出ネットワーク200、又は前記特徴抽出ネットワーク200に含まれる各特徴抽出サブネットワークは、訓練サンプル集に基づいて事前訓練された抽出ネットワークであっても良い。このようにして、画像特徴抽出の効率を向上させることができるため、視差推定の効率を向上させることができる。もちろん、実際のニーズに応じて、前記特徴抽出ネットワーク200、又は前記特徴抽出ネットワーク200に含まれる各特徴抽出サブネットワークは、訓練サンプル集に基づいてリアルタイム訓練を行うことにより得られても良く、或いは、更新後の訓練サンプル集に基づいて事前訓練済みの抽出ネットワークに対してリアルタイムな又は定時の最適化を行うことにより得られても良い。これにより、特徴抽出ネットワークが特徴を抽出する正確性を向上させることができる。

【0030】

幾つかの実施例によれば、前記特徴抽出ネットワーク200、又は前記特徴抽出ネットワーク200に含まれる各特徴抽出サブネットワークの訓練過程は、教師有り訓練又は教師無し訓練を採用しても良く、実際のニーズに応じて柔軟に選択することができる。教師有り訓練では、通常、既存の訓練サンプル (例えば、ラベル付きデータ) を利用して入力から出力までのマッピングを学習し、その後、このようなマッピング関係を未知のデータに適用することで、分類又は回帰の目的を達成する。教師有り訓練のアルゴリズムは、例えば

10

20

30

40

50

、論理回帰アルゴリズム、SVM (Support Vector Machine) アルゴリズム、決定木アルゴリズムなどを含んでも良い。教師無し訓練と教師有り訓練の相違点は、教師無し訓練ではサンプルを訓練する必要がなく、直接、ラベル無しのデータに対してモデリングを行うことで、そのうちの規律を見つけることにあり、その典型的なアルゴリズムは、例えば、クラスタリングアルゴリズム、ランダムフォレスト (Random forests) アルゴリズムなどを含んでも良い。

【 0 0 3 1 】

幾つかの実施例によれば、前述のように、前記多段視差処理のうちの第一段の視差処理の入力は、該段の視差処理に対応するサイズを有する複数の画像特徴を含んでも良く、前記多段視差処理のうちの第一段の視差処理以外の各段の視差処理の入力は、該段の視差処理に対応するサイズを有する1つ又は複数の画像特徴を含んでも良い。前記視差生成ネットワーク300により得られた前記複数の視差図が、サイズが順次増大するN個の視差図であることを例にとれば、前記特徴抽出ネットワーク200が抽出した画像特徴は、N種類のサイズの画像特徴を含んでも良く、Nは、2以上の正の整数である。各種類のサイズの画像特徴のうち少なくとも一部が、対応するサイズの視差図の生成を助けるために用いられ得るため、前記視差推定システムが得る前記複数の視差図の正確性を向上させることができる。幾つかの実施例によれば、前記Nの値は(図5又は図6に示すように)4であっても良く、もちろん、実際のニーズに応じて2、3、5などの値に設定することもできる。また、必ずしも、Nの値が大きければ大きいほど良いとは限られず、ターゲット装置の精度のニーズと視差推定システムの速度とのバランスをとることを前提として、適切な数値を選択しても良い。

10

20

【 0 0 3 2 】

本開示では、各画像(前記画像対における各画像、視差図などを含む)のサイズとは、各画像の単一チャンネルのサイズを指しても良く、画像の高さ及び幅により表すことができ、例えば、 $H \times W$ と表すことができる。ここで、Hは画像の高さを示し、Wは画像の幅を示し、また、両者の単位はともに画素であっても良い。もちろん、これは例示に過ぎず、画像のサイズは、画像を反映し得る画素数、データ量、記憶量、クラリティ (Clarity) などのうちの1つ又は複数の他のパラメータにより表されても良い。また、留意されたいのは、グレースケール画像について言えば、そのチャンネル数は1であり、カラー画像について言えば、R、G、Bの3つの色チャンネルがあるので、そのチャンネル数は3であっても良く、即ち、カラー画像の実際のサイズは $H \times W \times 3$ と表され得るということである。また、本開示では、前記画像対における各画像のサイズ(即ち、ダウンサンプリング、アップサンプリングなどの処理が行われる前のオリジナル画像のサイズ)は、例えば、前記画像対を採取するためのマルチカメラのセンサーのサイズ、画素数などのパラメータに基づいて確定することができる。

30

【 0 0 3 3 】

本開示では、各段の視差処理に対応するサイズとは、各段の視差処理が得る必要のある視差図のサイズと一致したサイズを指しても良い。また、画像特徴のサイズとは、画像特徴自身により構成される画像の単一チャンネルのサイズを指しても良く、又は、必要なサイズの画像特徴の抽出が基づく被抽出画像(即ち、該被抽出画像に基づいて、必要なサイズの画像特徴の抽出を行う)のサイズを指しても良く、前記被抽出画像とは、前記画像対における各画像自身を指しても良く、各画像自身に対して所定の係数又は比率のアップサンプリング又はダウンサンプリングを行うことにより得られた画像であっても良い。例えば、前記画像対における画像のサイズが $H \times W$ (フルサイズと言う)であることを例にとると、該画像について抽出したフルサイズの画像特徴は、該画像自身に対して特徴抽出を行うことにより得られた画像特徴であっても良く; 該画像について抽出した $(H/2) \times (W/2)$ のサイズ(1/2サイズと言う)の画像特徴は、該画像に対して2倍のダウンサンプリングを行うことにより1/2サイズの画像を取得し、そして、該1/2サイズの画像に対して特徴抽出を行うことにより得られた画像特徴であっても良い。

40

【 0 0 3 4 】

50

幾つかの実施例によれば、前記多段視差処理のうちの第一段の視差処理以外の各段の視差処理の入力は、該段の視差処理に対応するサイズを有する1つ又は複数の画像特徴以外に、1つ前の段の視差処理により生成される視差図をさらに含んでも良い。換言すると、抽出した画像特徴のうちの1つ又は複数の、対応するサイズを有する画像特徴に基づいて、第一段の視差処理が生成した視差図に対して逐段最適化処理を行うことで、対応するサイズの視差図を得ることができる。これにより、後続に得られた視差図の精度が次第に向上し、各精度について最初から視差図を計算する必要がない。このようにして、複数の視差図の全体の生成効率を向上させることができる。

【0035】

幾つかの実施例によれば、前記特徴抽出ネットワーク200が抽出した前記N種類のサイズの画像特徴のうちの最小サイズの画像特徴は、例えば、前記画像対における第一画像の少なくとも1種類の画像特徴及び第二画像の少なくとも1種類の画像特徴を含んでも良く、前記N種類のサイズの画像特徴のうちの各種類の非最小サイズの（最小サイズでない）画像特徴は、例えば、前記画像対における前記第一画像の少なくとも1種類の画像特徴及び/又は前記第二画像の少なくとも1種類の画像特徴を含んでも良い。

【0036】

例えば、図5又は図6に示すように、前記特徴抽出ネットワーク200が抽出した前記N種類のサイズの画像特徴のうちの最小サイズの画像特徴は、前記画像対における第一画像（例えば、左画像）の基礎構造特徴、語義的特徴、エッジ特徴、及び前記画像対における第二画像（例えば、右画像）の基礎構造特徴を含んでも良く、前記特徴抽出ネットワーク200が抽出した前記N種類のサイズの画像特徴のうちの各非最小サイズの画像特徴は、前記画像対における第一画像のエッジ特徴又は前記第一画像自身に基づく特徴を含んでも良い。

【0037】

幾つかの実施例によれば、前記特徴抽出ネットワーク200が抽出した各画像の各種類の画像特徴は1つ又は複数サイズを有しても良く、前記複数の種類のサイズの数は、N以下であっても良い。例えば、図5又は図6に示すように、前記Nの値は4であっても良く、前記特徴抽出ネットワーク200が抽出した前記第一画像のエッジ特徴及び前記第一画像自身に基づく特徴はそれぞれ2種類のサイズを有しても良く、抽出した前記第一画像の基礎構造特徴及び語義的特徴はそれぞれ1種類のサイズを有しても良く、抽出した前記第二画像の基礎構造特徴は1種類のサイズを有しても良い。また、図5又は図6は例示に過ぎず、前記特徴抽出ネットワーク200が抽出した各画像の各種類の画像特徴は、示された1種類又は2種類のサイズ以外に、より多くの種類のサイズをさらに有しても良い。例えば、前記Nの値が4であることを例にとれば、前記特徴抽出ネットワーク200が抽出した前記第一画像のエッジ特徴はさらに2種類又は4種類のサイズを有しても良いが、これについては限定しない。

【0038】

幾つかの実施例によれば、前記特徴抽出ネットワーク200は、前記画像対における各画像の画像特徴を抽出した後に、それを記憶装置又は記憶媒体に記憶（例えば、バッファリング）して後続の読み出し及び使用に供しても良い。また、前記特徴抽出ネットワーク200は、前記画像対における各画像に対して画像特徴抽出を行う前に、前記画像対における画像に対してエピポーラ修正（epipolar rectification）を行うことで、前記画像対における画像が1つのみの方向（例えば、水平方向又は垂直方向）において視差を有するようにさせることもできる。これにより、画像の視差サーチ範囲が1つのみの方向に限定されるから、後続の特徴抽出及び視差生成の効率を向上させることができる。代替案として、前記画像対における画像のエピポーラ修正の操作は、マルチカメラ又は他の装置により行われても良い。例えば、マルチカメラは、前記画像対を採取した後に、前記画像対における画像に対してエピポーラ修正を行い、そして、エピポーラ修正後の画像対を前記視差推定システムに送っても良い。或いは、マルチカメラが前記画像対を採取した後に、それを他の装置に送り、この他の装置は、前記画像対における画像に対してエピポーラ修正を行い、そして、エピポーラ修正後の画像対を前記視差推定システムに送っても良い。

10

20

30

40

50

## 【0039】

幾つかの実施例によれば、前記視差生成ネットワーク300が得た前記複数の視差図のうちの最大サイズを有する視差図は、前記画像対における各画像のサイズ（即ち、各画像のオリジナルサイズ）と一致しても良い。これによって、カスケード接続される多段視差処理により、少なくとも1つの、対応するサイズが前記画像対における各画像のサイズと一致しており、且つ精度が比較的高い視差図、及び他の精度を有する視差図を得ることができる。このようにして、視差推定システムの柔軟性及び可適用性が向上したうえで、高パフォーマンスのターゲット装置の、視差推定システムが生成する視差図の精度への要求をより良く満たすこともできる。代替案として、前記複数の視差図のうちの各視差図のサイズは、すべて、前記画像対における各画像のサイズよりも小さくても良い。

10

## 【0040】

幾つかの実施例によれば、前記複数の視差図のうちの任意の2つの隣接視差図のうち、後ろの1つの視差図の高さ及び幅は、それぞれ、前の1つの視差図の高さ及び幅の2倍であっても良い。例示として、前記複数の視差図が4つあり、且つ前記複数の視差図のうちの最後の1つの視差図のサイズが $H \times W$ である（前記画像対における各画像のサイズと一致しても良い）ことを例にとれば、前記複数の視差図のうちのその前の他の各視差図のサイズは、順次、 $(H/2) \times (W/2)$ （ $H \times W$ がフルサイズと称されれば、 $(H/2) \times (W/2)$ は1/2サイズと称される）、 $(H/4) \times (W/4)$ （1/4サイズと称される）、及び $(H/8) \times (W/8)$ （1/8サイズと称される）であっても良い。換言すると、本開示では、数値2を隣接視差図の高さ及び幅の拡張ステップ長（又は、隣接視差図のアップサンプリング又はダウンサンプリングの係数又は比率）とする。代替案として、前記複数の視差図のうちの任意の2つの隣接視差図のうち、後ろの1つの視差図の高さ及び幅は、それぞれ、前の1つの視差図の高さ及び幅の3倍、4倍又は他の1よりも大きい正の整数倍であっても良く、実際に必要な精度に基づいて適切な数値を選択することができる。

20

## 【0041】

幾つかの実施例によれば、前述のように、前記特徴抽出ネットワーク200が抽出した画像特徴はN種類のサイズの画像特徴を含んでも良く、前記Nは、2以上の正の整数である。それ相応に、前記視差生成ネットワークは次のように構成されても良く、即ち、前記多段視差処理の第一段の視差処理において、前記N種類のサイズの画像特徴のうちの最小サイズの画像特徴の少なくとも一部に基づいて、前記最小サイズを有する初期視差図を生成し；及び、前記多段視差処理のうちの後続の各段の視差処理において、前記N種類のサイズの画像特徴のうちの、対応するサイズを有する画像特徴の少なくとも一部に基づいて、1つ前の段の視差処理により生成される視差図に対して視差最適化処理を行い、前記対応するサイズを有する最適化視差図を生成し、ここで、前記複数の視差図は、少なくとも、各最適化視差図を含んでも良い。

30

## 【0042】

幾つかの実施例によれば、前記多段視差処理は $N+1$ 段の視差処理を含んでも良い。前記視差生成ネットワークは次のように構成されても良く、即ち、第一段の視差処理以外のN段の視差処理において、サイズの小さいから大への順序に従って、順に、前記N種類のサイズの画像特徴のうちの、対応するサイズを有する画像特徴の少なくとも一部に基づいて、1つ前の段の視差処理により生成される視差図に対して視差最適化処理を行い、サイズが順次増大するN個の最適化視差図を取得し、そして、前記N個の最適化視差図を前記複数の視差図とし、ここで、前記N個の最適化視差図のサイズは、それぞれ、前記N種類のサイズに対応する。

40

## 【0043】

例えば、図5に示すように、前記多段視差処理は $4+1$ 段の視差処理を含んでも良い。前記抽出した画像特徴は4種類のサイズの画像特徴を含んでも良く、ここで、図5では、抽出した前記4種類のサイズの画像特徴がそれぞれ1/8サイズ（ $(H/8) \times (W/8)$ ）、1/4サイズ（ $(H/4) \times (W/4)$ ）、1/2サイズ（ $(H/2) \times (W/2)$ ）、及びフルサイズ（ $H \times W$ であり、前記フルサイズとは、前記画像対におけるオリジナル画像のサイズと一致したサイズを指しも良い）で

50

あることを例にとって例示的に説明する。前記視差生成ネットワークは次のように構成されても良く、即ち、前記多段視差処理の第一段の視差処理において、前記4種類のサイズの画像特徴のうちの最小サイズ（即ち、1/8サイズ）の画像特徴の少なくとも一部（例えば、抽出した第一画像の1/8サイズの基礎構造特徴、1/8サイズの語義的特徴、1/8サイズのエッジ特徴、及び第二画像の1/8サイズの基礎構造特徴の一部又は全部）に基づいて、前記最小サイズ（即ち、1/8サイズ）を有する初期視差図を生成する。第一段の視差処理以外の4段の視差処理において、サイズの小さいから大への順序に従って、順に、前記4種類のサイズの画像特徴のうちの、対応するサイズを有する画像特徴の少なくとも一部（例えば、順に、抽出した第一画像の1/8サイズのエッジ特徴の一部又は全部、抽出した第一画像の1/4サイズのエッジ特徴の一部又は全部、抽出した1/2サイズの第一画像自身に基づく特徴、及び抽出したフルサイズの第一画像自身に基づく特徴）に基づいて、1つ前の段の視差処理により生成される視差図に対して視差最適化処理を行い、4個のサイズが順次増大する最適化視差図（例えば、1/8サイズを有する最適化視差図、1/4サイズを有する最適化視差図、1/2サイズを有する最適化視差図、及びフルサイズを有する最適化視差図）を取得し、そして、前記4個の最適化視差図を前記複数の視差図とする。

#### 【0044】

上述の説明から分かるように、該実施例では、前記視差推定システム100が得た前記複数の視差図は、前記多段視差処理のうちの第一段の視差処理が生成した初期視差図を含まず、前記第一段の視差処理が生成した初期視差図に対しての逐段最適化後の各最適化視差図を含んでも良い。これにより、前記視差推定システムが得る前記複数の視差図の精度を向上させることができる。

#### 【0045】

他の幾つかの実施例によれば、前記多段視差処理はN段の視差処理を含んでも良い。前記視差生成ネットワークは次のように構成されても良く、即ち、第一段の視差処理以外のN-1段の視差処理において、サイズの小さいから大への順序に従って、順に、前記N種類のサイズの画像特徴のうちのN-1種類の非最小サイズの画像特徴のうちの、対応するサイズを有する画像特徴の少なくとも一部に基づいて、1つ前の段の視差処理により生成される視差図に対して視差最適化処理を行い、N-1個のサイズが順次増大する最適化視差図を取得し、そして、前記初期視差図及び前記N-1個の最適化視差図を前記複数の視差図とし、ここで、前記初期視差図及び前記N-1個の最適化視差図のサイズは、それぞれ、前記N種類のサイズに対応する。

#### 【0046】

例えば、図6に示すように、前記多段視差処理は4段の視差処理を含んでも良い。前記抽出した画像特徴は4種類のサイズの画像特徴を含んでも良い、ここで、図6では、抽出した前記4種類のサイズの画像特徴がそれぞれ1/8サイズ（ $(H/8) \times (W/8)$ ）、1/4サイズ（ $(H/4) \times (W/4)$ ）、1/2サイズ（ $(H/2) \times (W/2)$ ）、及びフルサイズ（ $H \times W$ であり、前記フルサイズとは、前記画像対における各画像のサイズと一致したサイズを指しても良い）であることを例にとって例示的に説明する。前記視差生成ネットワークは次のように構成されても良く、即ち、前記多段視差処理の第一段の視差処理において、前記4種類のサイズの画像特徴のうちの最小サイズ（即ち、1/8サイズ）の画像特徴の少なくとも一部（例えば抽出した第一画像の1/8サイズの基礎構造特徴、1/8サイズの語義的特徴、1/8サイズのエッジ特徴及び第二画像の1/8サイズの基礎構造特徴の一部又は全部）に基づいて、前記最小サイズ（即ち、1/8サイズ）を有する初期視差図を生成する。第一段の視差処理以外の他の3段の視差処理において、サイズの小さいから大への順序に従って、順に、他の3種類の非最小サイズの画像特徴のうちの、対応するサイズを有する画像特徴の少なくとも一部（例えば、順に、抽出した第一画像の1/4サイズのエッジ特徴の一部又は全部、抽出した1/2サイズの第一画像自身に基づく特徴、及び抽出したフルサイズの第一画像自身に基づく特徴）に基づいて、1つ前の段の視差処理により生成される視差図に対して視差最適化処理を行い、3個のサイズが順次増大する最適化視差図（例えば、1/4サイズを有する最適化視差図、1/2サイズを有する最適化視差図、及びフルサイズを有する最適化視差図）を取得し、そして

、前記初期視差図及び前記3個の最適化視差図を前記複数の視差図とする。

【0047】

上述の説明から分かるように、該実施例では、前記視差推定システム100が得た前記複数の視差図は、前記多段視差処理のうちの第一段の視差処理が生成した初期視差図をさらに含んでも良い。これにより、前記視差推定システムの処理効率を向上させることができる。

【0048】

幾つかの実施例によれば、前記視差生成ネットワーク300は次のように構成されて良く、即ち、前記多段視差処理のうちの第一段の視差処理以外の各段の視差処理において、対応するサイズを有する画像特徴の少なくとも一部に基づいて、1つ前の段の視差処理により生成される視差図に対して残差(residual)計算処理を行い、前記対応するサイズを有する残差図を取得し、そして、前記対応するサイズを有する残差図と、前記1つ前の段の視差処理により生成される視差図との組み合わせを行って前記対応するサイズを有する最適化視差図を得る。

【0049】

例えば、図5に示すように、第一段の視差処理以外の他の4段の視差処理のうちの1/8サイズに対応する視差処理において、抽出した第一画像の1/8サイズのエッジ特徴の一部又は全部、及び前の段の視差処理が生成した初期視差図(1/8サイズ)に基づいて、1/8サイズの第一残差図を算出し、そして、前記第一残差図と、前記初期視差図との組み合わせ(例えば、重ね合わせ)を行って1/8サイズの第一最適化視差図を得て該段の視差処理の出力とすることができる。次の段の1/4サイズに対応する視差処理において、抽出した第一画像の1/4サイズのエッジ特徴の一部又は全部、及び前の段の視差処理が生成した第一最適化視差図(1/8サイズ)に基づいて1/4サイズの第二残差図を算出し、そして、前記第二残差図と、前記第一最適化視差図との組み合わせ(例えば、前記第二残差と、前記第一最適化視差図の1/4サイズのアップサンプリング後のものとの和(重ね合わせ))を行って1/4サイズの第二最適化視差図を得て該段の視差処理の出力とすることができる。その他は、これに基づいて類推することができる。なお、より具体的な例示については、後述する。

【0050】

幾つかの実施例によれば、前記視差生成ネットワーク300はさらに次のように構成されても良く、即ち、前記多段視差処理のうちの第一段の視差処理以外の各段の視差処理において、1つ前の段の視差処理により生成される視差図に対して視差最適化処理を行う前に、前記1つ前の段の視差処理により生成される視差図のサイズが現在の段の視差処理に対応するサイズよりも小さいことに応じて、前記1つ前の段の視差処理により生成される視差図を現在の段の視差処理に対応するサイズにアップサンプリングする。前記アップサンプリングに採用するアルゴリズムは、例えば、最近隣補間アルゴリズム、双線形補間アルゴリズム、デコンボリューションアルゴリズムなどを含んでも良い。このようにして、各段の視差最適化処理の対象となる視差図は、すべて、該段の視差処理に対応するサイズを有する視差図であるようになる。

【0051】

例示として、図5に示すように、第一段の視差処理以外の他の4段の視差処理のうちの1/8サイズに対応する視差処理において、1つ前の段の視差処理が生成した初期視差図のサイズが現在の段の視差処理以上であるから、直接、抽出した第一画像の1/8サイズのエッジ特徴の一部又は全部に基づいて、1つ前の段の視差処理が生成した初期視差図に対して視差最適化処理を行って1/8サイズの第一最適化視差図を得ることができる。次の段の1/4サイズに対応する視差処理において、まず、1つ前の段の視差処理が生成した1/8サイズの第一最適化視差図を現在の段の視差処理に対応する1/4サイズにアップサンプリングし、その後、抽出した第一画像の1/4サイズのエッジ特徴の一部又は全部に基づいて、アップサンプリング後の1/4サイズの第一最適化視差図に対して視差最適化処理を行って1/4サイズの第二最適化視差図を得ることができる。例えば、前述のように、この次の段の1/4サイ

10

20

30

40

50

ズに対応する視差処理において、抽出した第一画像の1/4サイズのエッジ特徴の一部又は全部、及びアップサンプリング後の1/4サイズの第一最適化視差図に基づいて、1/4サイズの残差図を算出し、そして、前記1/4サイズの残差図と、前記アップサンプリング後の1/4サイズの第一最適化視差図との和を求めて1/4サイズの第二最適化視差図を得ることができる。その他は、これに基づいて類推することができる。

【0052】

幾つかの実施例によれば、異なる最適化視差図の生成に基づく画像特徴は、同じ種類の画像特徴又は異なる種類の画像特徴であっても良く、及び/又は、異なる最適化視差図の生成に基づく画像特徴は、前記画像対における同じ画像又は異なる画像の画像特徴であっても良い。例えば、図5に示すように、前の2つの異なる最適化視差図の生成に基づく画像特徴は、前記画像対における同じ画像（例えば、第一画像）の同じ種類の画像特徴（例えば、エッジ特徴）であっても良く、中間の2つの異なる最適化視差図の生成に基づく画像特徴は、同じ画像（例えば、第一画像）の異なる種類の画像特徴（例えば、1つは第一画像のエッジ特徴であり、もう1つは第一画像の画像自身に基づく特徴である）であっても良い。これにより、各段の視差最適化に基づく画像特徴を柔軟に選択することで、視差推定システムの柔軟性及び可適用性をさらに向上させることができる。

10

【0053】

幾つかの実施例によれば、各最適化視差図の生成に基づく画像特徴は、例えば、前記画像対における少なくとも1つの画像のエッジ特徴、及び/又は、前記画像対における少なくとも1つの画像の画像自身に基づく特徴を含んでも良い。

20

【0054】

例示として、図5に示すように、前記複数の視差処理のうちの第一段の視差処理以外の前の2段に対応するサイズが比較的小さい視差最適化処理において、第一画像のエッジ特徴に基づいて対応するサイズの最適化視差図を生成することができ、後ろの2段に対応するサイズが比較的大きい視差最適化処理において、第一画像自身に基づく特徴をエッジ特徴の代わりに使用して1つ前の段の視差処理が生成した視差図に対して視差最適化処理を行うことで、大サイズの画像特徴の抽出に必要な計算量を減少させ、視差推定システムの処理効率を向上させることができる。もちろん、図5は例示に過ぎず、各視差最適化処理について言えば、対応する最適化視差図の生成に基づく画像特徴は、エッジ特徴又は画像自身に基づく特徴以外に、両者の組み合わせであっても良く、又は、その他の抽出した1つ又は複数の画像特徴の組み合わせであっても良い。

30

【0055】

幾つかの実施例によれば、前記画像対における少なくとも1つの画像の画像自身に基づく特徴は、例えば、前記少なくとも1つの画像の画像自身を含んでも良く、又は、生成する必要のある最適化視差図のサイズに基づいて前記少なくとも1つの画像の画像自身に対してダウンサンプリングを行うことにより得られた画像であっても良い。前記ダウンサンプリングの過程は、例えば、1つのサイズが $H \times W$ の画像について言えば、採用するダウンサンプリング係数又は比率が $K$ である場合、前記画像の原図における各行各列において $K$ 個の点毎に1つの点を選択して1つの画像を構成しても良い。前記ダウンサンプリング係数又は比率は、前記のように、2、3又は他の1よりも大きい数値であっても良い。もちろん、これは例示に過ぎず、ダウンサンプリングは他の方式で実現されても良く、例えば、 $K$ 個の点の平均を取っても良い。

40

【0056】

図5に示すように、前記複数の視差処理のうちの後ろの2段に対応する1/2サイズ及びフルサイズの視差最適化処理において、それぞれ、ダウンサンプリング係数2に基づいて第一画像自身に対してダウンサンプリングを行うことにより得られた1/2サイズの第一画像及び第一画像自身を対応するサイズのエッジ特徴の代わりに使用して、1つ前の段の視差処理が生成した視差図に対して視差最適化処理を行うことで、大サイズの画像特徴の抽出に必要な計算量を減少させ、視差推定システムの処理効率を向上させることができる。

【0057】

50

幾つかの実施例によれば、図5又は図6に示すように、前記視差生成ネットワーク300は、初期視差生成サブネットワーク301及び少なくとも1つの視差最適化サブネットワーク302を含んでも良く、前記初期視差生成サブネットワーク301及び前記少なくとも1つの視差最適化サブネットワーク302のうちの各視差最適化サブネットワーク302は、順に、カスケード接続され、前記初期視差生成サブネットワーク301は、第一段の視差処理を行うように構成され、前記少なくとも1つの視差最適化サブネットワーク302は、第一段の視差処理以外の各段の視差処理を行うように構成される。

【0058】

前述の実施例から分かるように、前記視差推定システム100の画像特徴抽出、及び多段視差処理のうちの各段の視差処理は、すべて、対応するサブネットワークにより実現することができる。

10

【0059】

例示として、以下、図5を例にとり、複数の特徴抽出サブネットワーク（例えば、基礎構造特徴サブネットワーク201、語義的特徴サブネットワーク202、及びエッジ特徴サブネットワーク203）、初期視差生成サブネットワーク301、及び複数の（例えば、4つの）視差最適化サブネットワーク302を含む視差推定システム100の全体ワーキング過程について例示的に説明する。

【0060】

図5から分かるように、前記視差推定システム100に入力される画像対における第一画像I1及び第二画像I2（サイズはそれぞれ $H \times W$ であっても良い）について、前記視差推定システム100は、その中に含まれる特徴抽出ネットワーク200における複数の特徴抽出サブネットワークにより、その後続の多段視差処理の実行に必要なサイズの画像特徴の抽出を行うことができる。例えば、基礎構造特徴サブネットワーク201により第一画像I1の1/8サイズ（ $(H/8) \times (W/8)$ ）の基礎構造特徴及び第二画像I2の1/8サイズの基礎構造特徴を抽出し、語義的特徴サブネットワーク202により第一画像I1の1/8サイズの語義的特徴を抽出し、エッジ特徴サブネットワーク203により第一画像I1の1/8サイズのエッジ特徴及び1/4サイズ（ $(H/4) \times (W/4)$ ）のエッジ特徴を抽出しても良い。また、これらの画像特徴以外に、前記視差推定システム100の特徴抽出ネットワーク200はさらに、1/2サイズ（ $(H/2) \times (W/2)$ ）の第一画像I1自身に基づく特徴、及びフルサイズ（ $H \times W$ ）の第一画像I1自身に基づく特徴、即ち第一画像I1自身を抽出することもできる。

20

30

【0061】

前記第一画像I1の1/8サイズの基礎構造特徴、第二画像I2の1/8サイズの基礎構造特徴、第一画像I1の1/8サイズの語義的特徴、及び第一画像I1の1/8サイズのエッジ特徴は、対応する特徴抽出サブネットワークから初期視差生成サブネットワーク301に出力して第一段の視差処理を行って貰うことで、1/8サイズを有する初期視差図dispS1を得ることができる。その後、前記初期視差生成サブネットワーク301と順にカスケード接続される4つの視差最適化サブネットワーク302は、それぞれ、特徴抽出ネットワーク200が抽出した、対応するサイズを有する画像特徴に基づいて、前記1/8サイズの初期視差図dispS1に対して順に異なる段の視差最適化処理を行うことで、サイズが順次増大する複数の最適化視差図を得ることができる。

40

【0062】

例えば、1番目の視差最適化サブネットワークは、エッジ特徴サブネットワーク203からの第一画像I1の1/8サイズのエッジ特徴（一部又は全部）に基づいて、初期視差生成サブネットワーク301が出力した1/8サイズの初期視差図dispS1に対して視差最適化処理を行い、1/8サイズの第一最適化視差図dispS1\_refineを得ることができる。幾つかの実施例によれば、1番目の視差最適化サブネットワークは、第一画像I1の1/8サイズのエッジ特徴（一部又は全部）及び1/8サイズの初期視差図dispS1に基づいて、1/8サイズの第一残差図を算出し、そして、1/8サイズの第一残差図と1/8サイズの初期視差図dispS1との和を求めて1/8サイズの第一最適化視差図dispS1\_refineを得ることができる。

【0063】

50

2番目の視差最適化サブネットワークは、エッジ特徴サブネットワーク203からの第一画像I1の1/4サイズのエッジ特徴（一部又は全部）に基づいて、1番目の視差最適化サブネットワークが出力した1/8サイズの第一最適化視差図dispS1\_refineに対して視差最適化処理を行い、1/4サイズの第二最適化視差図dispS2\_refineを得ることができる。幾つかの実施例によれば、2番目の視差最適化サブネットワークは、1番目の視差最適化サブネットワークが出力した1/8サイズの第一最適化視差図を現在の段の視差処理に対応する1/4サイズにアップサンプリングし、その後、第一画像I1の1/4サイズのエッジ特徴（一部又は全部）及びアップサンプリング後の1/4サイズの第一最適化視差図に基づいて、1/4サイズの第二残差図を算出し、そして、1/4サイズの第二残差図と、アップサンプリング後の1/4サイズの第一最適化視差図との和を求めて1/4サイズの第二最適化視差図dispS2\_refineを得ることができる。

10

【0064】

3番目の視差最適化サブネットワークは、特徴抽出ネットワーク200が抽出した1/2サイズの第一画像I1自身に基づく特徴（一部又は全部）に基づいて、2番目の視差最適化サブネットワークが出力した1/4サイズの第二最適化視差図dispS2\_refineに対して視差最適化処理を行い、1/2サイズの第三最適化視差図dispS3\_refineを得ることができる。幾つかの実施例によれば、3番目の視差最適化サブネットワークは、2番目の視差最適化サブネットワークが出力した1/4サイズの第二最適化視差図を現在の段の視差処理に対応する1/2サイズにアップサンプリングし、その後、1/2サイズの第一画像自身に基づく特徴（一部又は全部）及びアップサンプリング後の1/2サイズの第二最適化視差図に基づいて1/2サイズの第三残差図を算出し、そして、1/2サイズの第三残差図と、アップサンプリング後の1/2サイズの第二最適化視差図との和を求めて1/2サイズの第三最適化視差図dispS3\_refineを得ることができる。

20

【0065】

4番目の視差最適化サブネットワークは、特徴抽出ネットワーク200が抽出したフルサイズの第一画像I1自身に基づく特徴（一部又は全部）に基づいて、3番目の視差最適化サブネットワークが出力した1/2サイズの第三最適化視差図dispS3\_refineに対して視差最適化処理を行い、フルサイズの第四最適化視差図dispS4\_refineを得ることができる。幾つかの実施例によれば、4番目の視差最適化サブネットワークは、3番目の視差最適化サブネットワークが出力した1/2サイズの第三最適化視差図を現在の段の視差処理に対応するフルサイズにアップサンプリングし、その後、フルサイズの第一画像自身に基づく特徴（一部又は全部）及びアップサンプリング後のフルサイズの第三最適化視差図に基づいてフルサイズの第四残差図を算出し、そして、フルサイズの第四残差図と、アップサンプリング後のフルサイズの第三最適化視差図との和を求めてフルサイズの第四最適化視差図dispS4\_refineを得ることができる。なお、この例では、3番目及び4番目の視差最適化サブネットワークは、第一画像自身に基づく特徴を使用して視差最適化処理を行うことで計算量を減少させるが、それらのうちの1つ又は2つは、第一画像のエッジ特徴又は他の特徴を使用しても良い。同様に、計算量を更に減少させる必要がある場合、1番目及び/又は2番目の視差最適化サブネットワークは、第一画像自身に基づく特徴を、抽出したエッジ特徴の代わりに使用しても良い。

30

40

【0066】

前記1/8サイズの第一最適化視差図dispS1\_refine、1/4サイズの第二最適化視差図dispS2\_refine、1/2サイズの第三最適化視差図dispS3\_refine、及びフルサイズの第四最適化視差図dispS4\_refineは、図5に示す視差推定システム100が得る、サイズが順次増大する複数の視差図とすることができる。

【0067】

また、図6に示す視差推定システム100の全体ワーキング過程は、図5に示す視差推定システム100の全体ワーキング過程と類似しているが、相違点は、初期視差生成サブネットワーク301が生成した初期視差図のサイズが、1番目の視差最適化サブネットワークが生成した最適化視差図のサイズよりも小さく、及び、初期視差生成サブネットワーク301が生

50

成した初期視差図を、視差推定システム100が得る、サイズが順次増大する複数の視差図のうちの一つとすることができることのみにあるため、ここでは、その詳しい説明を割愛する。

【0068】

幾つかの実施例によれば、前記初期視差生成サブネットワーク301及び前記少なくとも一つの視差最適化サブネットワーク302のうちの各視差最適化サブネットワーク302は、2DCNN (2次元深層畳み込みニューラルネットワーク)、3DCNN (3次元深層畳み込みニューラルネットワーク)などの任意の、対応する視差処理機能を実現し得る畳み込みニューラルネットワークであっても良い。畳み込みニューラルネットワークを視差処理サブネットワークとして採用することにより、比較的大きい感知領域を得ることができるので、視差推定システム100が得る視差図の精度を向上させることができる。

10

【0069】

幾つかの実施例によれば、前記初期視差生成サブネットワーク301について言えば、それが2DCNN構造を採用して視差を得る場合、前記初期視差生成サブネットワーク301は、第一の数(例えば、5であるが、もちろん、実際のニーズに応じて対応する数値を柔軟に選択しても良い)の順次カスケード接続される畳み込み層(convolution layer)を含んでも良い。各畳み込み層の畳み込み方式は、例えば、深層分離可能畳み込み(depthwise separable convolution)などを採用しても良い。

【0070】

例示として、以下、表1に基づいて、1種類の、図5に示す視差推定システムに適用される、5つの順次カスケード接続される畳み込み層(例えば、表1におけるconv1-conv5)を含む2DCNN構造の初期視差生成サブネットワーク301について例示的に説明する。例示として、該サブネットワークは、MobileNetV2ネットワークアーキテクチャを採用する。

20

【0071】

表1は、初期視差生成サブネットワーク301の1種類の可能な2DCNNネットワーク構造に関する説明である。

【0072】

【表1】

名称 (Name)	レイヤ説明 (Layer Description)	出力テンソルサイズ (Output Tensor Dim)
corr1d	相関層 (correlation layer)	1/8H × 1/8W × 1/8D
semanS1_conv	語義的特徴に由来する : conv 3×3, 1/8D features (特徴)	1/8H × 1/8W × 1/8D
edgeS1_conv	エッジ特徴に由来する : conv 3×3, 1/8D features	1/8H × 1/8W × 1/8D
concat	concat (corr1d, semanS1_conv, edgeS1_conv)	1/8H × 1/8W × 3/8D
conv1	MB_conv, 3/8D features	1/8H × 1/8W × 3/8D
conv2	MB_conv, 2/8D features	1/8H × 1/8W × 2/8D
conv3	MB_conv_res, 2/8D features	1/8H × 1/8W × 2/8D
conv4	MB_conv, 1/8D features	1/8H × 1/8W × 1/8D
conv5	MB_conv_res, 1/8D features	1/8H × 1/8W × 1/8D
dispS1	soft argmin	1/8H × 1/8W × 1

30

40

表1及び図5から分かるように、corr1d層は、図5における特徴抽出ネットワーク200が抽出した第一画像の1/8サイズの基礎構造特徴及び第二画像の1/8サイズの基礎構造特徴に対して関連操作を行うために用いられ得る。semanS1\_conv層は、3×3の畳み込みカーネルに基づいて、抽出した第一画像の1/8サイズの語義的特徴に対して畳み込み処理を行うため

50

に用いられ得る。edgeS1\_conv層は、 $3 \times 3$ の畳み込みカーネルに基づいて、抽出した第一画像の1/8サイズのエッジ特徴に対して畳み込み処理を行うために用いられ得る。concat層は、corr1d、semanS1\_conv及びedgeS1\_convが出力した特徴に対して合併処理を行うために用いられ得る。

【0073】

また、conv1-conv5層に係るMB\_conv操作とは、MobileNetV2における深層分離可能畳み込み(depthwise separable convolution)操作を指しても良く、MB\_conv\_res操作とは、MobileNetV2における残差深層分離可能畳み込み(residual depthwise separable convolution)操作を指しても良い。換言すれば、conv1層、conv2層及びconv4層は、それぞれ、1つ前の層が出力した特徴に対して深層分離可能畳み込み操作を行うために用いられ、conv3層及びconv5層は、それぞれ、1つ前の層が出力した特徴に対して残差深層分離可能畳み込み操作を行うために用いられ得る。また、dispS1は、1つ前の層が出力した特徴に対してsoft argmin計算を行い、対応するサイズ(即ち、1/8サイズ)の初期視差図dispS1を得るために用いられ得る。

10

【0074】

なお、表1に言及されているH及びWは、それぞれ、視差推定システム100に入力される画像対における画像の高さ及び幅を示し、Dは、画像の最大視差範囲を示し、三者の単位はすべて画素であっても良く、ここで、Dの値は、前記画像対を採取するためのマルチカメラにおける各カメラヘッドの焦点距離及び/又は各カメラヘッド間の間隔と関係がある。また、2DCNN構造を採用する初期視差生成サブネットワーク301の畳み込み層の数は、concat層が得た特徴の数に基づいて決定されても良い。例えば、concat層が得た特徴の数が比較的多い場合、前記初期視差生成サブネットワーク301に含まれる畳み込み層の数を更に増やしても良い。

20

【0075】

代替案として、前記初期視差生成サブネットワーク301はさらに、3DCNN構造を採用して視差を得ても良い。3DCNN構造を採用する前記初期視差生成サブネットワーク301は、第二の数(例えば、7であるが、もちろん、実際のニーズに応じて対応する数値を柔軟に選択しても良い)の順次カスケード接続される畳み込み層を含んでも良い。

【0076】

例示として、以下、表2をもとに、1種類の、図5に示す視差推定システムに適用され得る、7つの順次カスケード接続される畳み込み層(例えば、表2におけるconv1-conv7)を含む3DCNN構造の初期視差生成サブネットワーク301について例示的に説明する。

30

【0077】

表2は、初期視差生成サブネットワーク301の1種類の可能な3DCNNネットワーク構造に関する説明である。

【0078】

【表2】

名称 (Name)	レイヤ説明 (Layer Description)	出力テンソルサイズ (Output Tensor Dim)
edgeS1_conv	エッジ特徴に由来する : conv 3×3, F features (特徴)	1/8H × 1/8W × F
semanS1_conv	語義的特徴に由来する : conv 3×3, F features	1/8H × 1/8W × F
concat	concat (featS1, semanS1_conv, edgeS1_conv)	1/8H × 1/8W × 3F
cost	シフト連結層 (shift concatenate layer)	1/8D × 1/8H × 1/8W × 6F
conv1	3DCNN, 3×3×3, 4F features	1/8D × 1/8H × 1/8W × 4F
conv2	3DCNN, 3×3×3, 4F features, add conv1	1/8D × 1/8H × 1/8W × 4F
conv3	3DCNN, 3×3×3, 2F features	1/8D × 1/8H × 1/8W × 2F
conv4	3DCNN, 3×3×3, 2F features, add conv3	1/8D × 1/8H × 1/8W × 2F
conv5	3DCNN, 3×3×3, 1F features	1/8D × 1/8H × 1/8W × 1F
conv6	3DCNN, 3×3×3, 2F features, add conv5	1/8D × 1/8H × 1/8W × 1F
conv7	3DCNN, 3×3×3, 1F features	1/8D × 1/8H × 1/8W × 1
dispS1	soft argmin	1/8H × 1/8W × 1

表2及び図5から分かるように、edgeS1\_conv層は、3×3の畳み込みカーネルに基づいて、抽出した第一画像の1/8サイズのエッジ特徴に対して畳み込み処理を行うために用いることができる。semanS3\_convは、3×3の畳み込みカーネルに基づいて、抽出した第一画像の1/8サイズの語義的特徴に対して畳み込み処理を行うために用いることができる。concat層は、featS1、semanS1\_conv及びedgeS1\_convが出力した特徴に対して合併処理を行うために用いることができる。ここで、表2に示されていないが、featS1とは、抽出した第一画像の1/8サイズの基礎構造特徴及び第二画像の1/8サイズの基礎構造特徴を指しても良い。

## 【0079】

また、cost層は、concat層が出力した特徴に対してシフト操作を行うために用いることができる。conv1層乃至conv7層は、それぞれ、3×3×3の畳み込みカーネルに基づいて、1つ前の層が出力した特徴に対して畳み込み操作を行うために用いることができ、ここで、conv2層、conv4層及びconv6層は、3DCNNネットワークの残差モジュールに相当し、さらに、それぞれ、1つ前の層が出力した特徴に対して畳み込み操作を行った後に、畳み込み結果と、1つ前の層が出力した結果との和の処理を行うために用いることもできる。dispS1層は、1つ前の層が出力した特徴に対してsoft argmin計算を行い、対応するサイズ(即ち、1/8サイズ)の初期視差図dispS1を得るために用いることができる。

## 【0080】

表1と同様に、表2に言及されているH及びWは、それぞれ、視差推定システム100に入力される画像対における画像の高さ及び幅を示す。また、Fは、特徴のチャンネル数を示し、1Fは、チャンネル数がFであることを示し、3Fは、チャンネル数が3×Fであることを示す。また、3DCNN構造を採用する初期視差生成サブネットワーク301の畳み込み層の数も、concat層が得た特徴の数に基づいて決定されても良い。例えば、concat層が得た特徴の数が比較的多い場合、前記初期視差生成サブネットワーク301に含まれる畳み込み層の数をさらに増やしても良い。

## 【0081】

幾つかの実施例によれば、前記少なくとも1つの視差最適化サブネットワーク302のうちの各視差最適化サブネットワーク302に含まれる畳み込み層の数は、前記初期視差生成サ

10

20

30

40

50

ブネットワーク301に含まれる畳み込み層の数未満であっても良い。例えば、各視差最適化サブネットワーク302が2DCNN構造を採用することを例にとると、各視差最適化サブネットワーク302に含まれる畳み込み層の数は3つであっても良いが、もちろん、実際のニーズに応じて他の数値を設定しても良い。また、前述の実施例の関連説明に記載のように、各視差最適化サブネットワーク302はさらに3DCNN構造を採用することもできるが、これについては限定しない。

【0082】

以下、表3乃至表6に基づいて、図5に示す視差推定システムに適用され得る複数の視差最適化サブネットワーク302の構造について例示的に説明する。ここで、表3乃至表6は、順に、図5に示す視差推定システムの1番目乃至4番目の視差最適化サブネットワークの可能な2DCNNネットワーク構造について例示的に説明する。

10

【0083】

表3は、1番目の視差最適化サブネットワークの1種類の可能な2DCNNネットワーク構造に関する説明である。

【0084】

【表3】

名称 (Name)	レイヤ説明 (Layer Description)	出力テンソルサイズ (Output Tensor Dim)
edgeS1_conv	エッジ特徴に由来する : conv 3×3, 8 features (特徴)	1/8H × 1/8W × 8
concat	concat (dispS1, edgeS1_conv)	1/8H × 1/8W × 9
conv1	MB_conv, 4 features	1/8H × 1/8W × 4
conv2	MB_conv_res, 4 features	1/8H × 1/8W × 4
conv3	MB_conv, 1 features	1/8H × 1/8W × 1
dispS1_refine	add(dispS1, conv3)	1/8H × 1/8W × 1

20

30

前述の表1などと同様に、edgeS1\_conv層は、3×3の畳み込みカーネルに基づいて、抽出した第一画像の1/8サイズのエッジ特徴に対して畳み込み処理を行うために用いることができる。Concatは、1つ前の段の視差処理（即ち、初期視差生成処理）が生成した1/8サイズの初期視差図dispS1、及びedgeS1\_convが出力した特徴に対して合併処理を行うために用いることができる。

【0085】

また、conv1層及びconv3層は、それぞれ、1つ前の層が出力した特徴に対して深層分離可能畳み込み操作を行うために用いることができ、conv2層は、1つ前の層が出力した特徴に対して残差深層分離可能畳み込み操作を行うために用いることができる。また、dispS1\_refine層は、1つ前の層conv3が出力した特徴、及び1つ前の段の視差処理が生成した1/8サイズの初期視差図dispS1に対して重ね合わせ演算を行い、対応するサイズ（即ち、1/8サイズ）の第一最適化視差図dispS1\_refineを得るために用いることができる。

40

【0086】

表4は、2番目の視差最適化サブネットワークの1種類の可能な2DCNNネットワーク構造に関する説明である。

【0087】

【表 4】

名称 (Name)	レイヤ説明 (Layer Description)	出力テンソルサイズ (Output Tensor Dim)
dispS1_up	upsample(dispS1_refine)	1/4H × 1/4W × 1
edgeS2_conv	エッジ特徴に由来する : conv 3×3, 8 features (特徴)	1/4H × 1/4W × 8
concat	concat (dispS1_up, edgeS2_conv)	1/4H × 1/4W × 9
conv1	MB_conv, 4 features	1/4H × 1/4W × 4
conv2	MB_conv_res, 4 features	1/4H × 1/4W × 4
conv3	MB_conv, 1 features	1/4H × 1/4W × 1
dispS2_refine	add(dispS1_up, conv3)	1/4H × 1/4W × 1

表4及び図5から分かるように、dispS1\_upは、1つ前の段の視差処理（即ち、第一段の視差最適化処理）が生成した1/8サイズの第一最適化視差図dispS1\_refineに対してアップサンプリング処理を行い、1/4サイズの最適化視差図dispS1\_upを得るために用いることができる。edgeS2\_conv層は、3×3の畳み込みカーネルに基づいて、抽出した第一画像の1/4サイズのエッジ特徴に対して畳み込み処理を行うために用いることができる。concat層は、アップサンプリング処理後の1/4サイズの最適化視差図dispS1\_up及びedgeS2\_convが出力した特徴に対して合併処理を行うために用いることができる。

【 0 0 8 8 】

また、conv1層及びconv3層は、それぞれ、1つ前の層が出力した特徴に対して深層分離可能畳み込み操作を行うために用いることができ、conv2層は、1つ前の層が出力した特徴に対して残差深層分離可能畳み込み操作を行うために用いることができる。また、dispS2\_refine層は、1つ前の層conv3が出力した特徴及びアップサンプリング処理後の1/4サイズの最適化視差図dispS1\_upに対して重ね合わせ演算を行い、対応するサイズ（即ち、1/4サイズ）の第二最適化視差図dispS2\_refineを得るために用いることができる。

【 0 0 8 9 】

表5は、3番目の視差最適化サブネットワークの1種類の可能な2DCNNネットワーク構造に関する説明である。

【 0 0 9 0 】

【表 5】

名称 (Name)	レイヤ説明 (Layer Description)	出力テンソルサイズ (Output Tensor Dim)
dispS2_up	upsample(dispS2_refine)	1/2H × 1/2W × 1
imgS3	downsample(I1)	1/2H × 1/2W × 3
concat	concat (dispS2_up, imgS3)	1/2H × 1/2W × 4
conv1	conv 3×3, 4 features (特徴)	1/2H × 1/2W × 4
conv2	conv 3×3, 2 features	1/2H × 1/2W × 2
conv3	conv 3×3, 1 features	1/2H × 1/2W × 1
dispS3_refine	add(dispS2_up, conv3)	1/2H × 1/2W × 1

表5及び図5から分かるように、dispS2\_up層は、1つ前の段の視差処理（即ち、第二段の

視差最適化処理)が生成した1/4サイズの第二最適化視差図dispS2\_refineに対してアップサンプリング処理を行い、1/2サイズの最適化視差図dispS2\_upを得るために用いることができる。imgS3層は、第一画像自身に対してダウンサンプリング処理を行い、1/2サイズの第一画像自身に基づく特徴を得るために用いることができ、ここで、表5におけるI1は第一画像を示す。concat層は、アップサンプリング処理後の1/2サイズの最適化視差図dispS2\_up、及びimgS3が出力した特徴に対して合併処理を行うために用いることができる。

【0091】

また、conv1層、conv2層及びconv3層は、それぞれ、1つ前の層が出力した特徴に対して畳み込み操作を行うために用いることができ、dispS3\_refine層は、1つ前の層conv3が出力した特徴及びアップサンプリング処理後の1/2サイズの最適化視差図dispS2\_upに対して重ね合わせ演算を行い、対応するサイズ(即ち、1/2サイズ)の第三最適化視差図dispS3\_refineを得るために用いることができる。

10

【0092】

表6は、4番目の視差最適化サブネットワークの1種類の可能な2DCNNネットワーク構造に関する説明である。

【0093】

【表6】

名称 (Name)	レイヤ説明 (Layer Description)	出力テンソルサイズ (Output Tensor Dim)
dispS3_up	upsample(dispS3_refine)	H × W × 1
Concat	concat (dispS3_up, I1)	H × W × 4
conv1	conv 3×3, 4 features (特徴)	H × W × 4
conv2	conv 3×3, 2 features	H × W × 2
conv3	conv 3×3, 1 features	H × W × 1
dispS4_refine	add(dispS3_up, conv3)	H × W × 1

20

表6及び図5から分かるように、dispS3\_upは、1つ前の段の視差処理(即ち、第三段の視差最適化処理)が生成した1/2サイズの第三最適化視差図dispS3\_refineに対してアップサンプリング処理を行い、フルサイズの最適化視差図dispS3\_upを得るために用いることができる。concat層は、アップサンプリング処理後のフルサイズの最適化視差図dispS3\_up、及び第一画像自身に対して合併処理を行うために用いることができ、ここで、表6におけるI1は第一画像を表す。

30

【0094】

また、conv1層、conv2層及びconv3層は、それぞれ、1つ前の層が出力した特徴に対して畳み込み操作を行うために用いることができ、dispS4\_refine層は、1つ前の層conv3が出力した特徴及びアップサンプリング処理後のフルサイズの最適化視差図dispS3\_upに対して重ね合わせ演算を行い、対応するサイズ(即ちフルサイズ)の第四最適化視差図dispS4\_refineを得るために用いることができる。

40

【0095】

なお、前述の実施例と同様に、表3乃至表6に言及されているH及びWは、それぞれ、視差推定システム100に入力される画像対における画像の高さ及び幅を示す。また、2DCNN構造を採用する各視差最適化サブネットワーク302の畳み込み層の数も、concat層が得た特徴の数に基づいて決定されても良い。例えば、concat層が得た特徴の数が比較的多い場合、各視差最適化サブネットワーク302に含まれる畳み込み層の数をさらに増やしても良い。

【0096】

幾つかの実施例によれば、前記初期視差生成サブネットワーク301及び前記少なくとも1つの視差最適化サブネットワーク302のうちの各サブネットワークは、訓練サンプル集に

50

基づいて事前訓練されたネットワークであっても良い。このようにして、視差処理の効率を向上させることができる。もちろん、実際のニーズに応じて、前記初期視差生成サブネットワーク301及び前記少なくとも1つの視差最適化サブネットワーク302のうちの各サブネットワークは、訓練サンプル集に基づいてリアルタイム訓練を行うことにより得られても良く、又は、更新後の訓練サンプル集に基づいて事前訓練済みのネットワークに対してリアルタイムな又は定時の最適化を行うことにより得られても良い。これにより、視差生成の正確性を向上させることができる。

【0097】

幾つかの実施例によれば、前記初期視差生成サブネットワーク301及び前記少なくとも1つの視差最適化サブネットワーク302のうちの各サブネットワークの訓練過程は、教師有り訓練又は教師無し訓練を採用しても良く、実際のニーズに応じて柔軟に選択することができる。教師有り訓練及び教師無し訓練についての説明は、前述の関連実施例における関連記載を参照することができ、ここではその詳しい説明を省略する。

10

【0098】

幾つかの実施例によれば、前記初期視差生成サブネットワーク301及び前記少なくとも1つの視差最適化サブネットワーク302のうちの各サブネットワークはさらに、損失関数を計算するように構成されても良く、前記損失関数は、前記サブネットワークが生成する視差図における視差と、対応する真の視差との間の誤差を表すために用いることができる。このようにして、損失関数を計算することで、視差推定システムが生成する各視差図の正確度を明確にすることができる。また、さらに損失関数に基づいて対応するシステムに対して最適化を行うこともできる。

20

【0099】

幾つかの実施例によれば、前記初期視差生成サブネットワーク301及び前記少なくとも1つの視差最適化サブネットワーク302のうちの各サブネットワークが教師有り訓練を採用することを例にとれば、各視差処理サブネットワーク又は各段の視差処理が出力した損失関数は、 $L_n = f(\text{Disp}^{\text{GT}^n} - \text{Disp}^{\text{S}^n}) + g(\text{Disp}^{\text{S}^n})$ と定義することができる。nの値は、1~N(図5に示す視差推定システムに対応する)、又は、0~N(図6に示す視差推定システムに対応する)であっても良い。関数fは、予測の視差( $\text{Disp}^{\text{S}^n}$ )と真の視差( $\text{Disp}^{\text{GT}^n}$ )との差を示し、gは、視差連続性制約を示す。

【0100】

【数1】

$$f(x) = \sqrt{\left(\frac{x}{2}\right)^2 + 1} - 1$$

であり、 $g(x) = |x_x| + |x_y|$ である。また、エッジ特徴を損失関数の正則化項とすることをさらに考慮しても良いが、これについて限定しない。それ相応に、前記視差推定システム100の最終の損失関数は、各視差処理サブネットワーク又は各段の視差処理が出力した損失関数の和であっても良い。

30

【0101】

他の幾つかの実施例によれば、前記初期視差生成サブネットワーク301、及び前記少なくとも1つの視差最適化サブネットワーク302のうちの各サブネットワークが教師無し訓練を採用する場合、画像を再構成し、再構成誤差を計算することで、各視差処理サブネットワーク又は各段の視差処理の損失関数を得ることができる。例えば、そのうちの一段の視差処理サブネットワークの損失関数を計算することを例にとると、算出された損失関数は、

40

【0102】

【数2】

$$\text{Loss} = \|I_1 - \text{warp}I_1\|_2^2$$

と表すことができ、ここで、 $\text{warp}I_1 = \text{warp}(I_2, \text{Disp}_1)$ であり、warp関数は、該段の視差処理サブネットワークが計算した視差に基づいて、第二画像I2から1つのI1画像を再構成することを表す。

50

## 【 0 1 0 3 】

以下、前記初期視差生成サブネットワーク301及び前記少なくとも1つの視差最適化サブネットワーク302のうちの各サブネットワークが教師有り訓練を採用し、訓練集がScene Flow（シーン流）を使用し、且つ視差推定システムの構造が図5に示すようなものであることを例にとり、図7A、図7B及び図8を参照しながら、訓練に基づく参照画像、対応する真値視差図、及び訓練後のパラメータをMiddleburyデータ集（画像）に適用してテストを行うことにより得られた結果について例示的に説明する。

## 【 0 1 0 4 】

図7A及び図7Bは、それぞれ、本開示の例示的な実施例においてネットワーク訓練に基づく参照画像及び対応する真値視差図（ground truth）を示す図であり、図8は、本開示の例示的な実施例において、訓練後の視差推定システムを採用して、図7Aに示す参照画像に対して、カスケード接続される多段視差処理を行うことにより得られた、右から左へサイズが順次増大する複数の視差図を示す図である（即ち、訓練後のパラメータをMiddleburyデータ集（画像）に適用してテストを行うことにより得られた結果である）。これらの図面から分かるように、得られた複数の視差図のサイズは順次増大し、精度は順次高くなり、且つ最大サイズの視差図の精度は真値視差図に近い。また、図7A、図7B及び図8は、それぞれ、グレースケール画像の方式で参照画像、真値視差図及び生成した複数の視差図を示したが、理解すべきは、図7Aに示す参照画像がカラー画像であるときに、図7B及び図8に示す各視差図も対応するカラー画像であるということである。

## 【 0 1 0 5 】

幾つかの実施例によれば、前記視差生成ネットワーク300はさらに、ターゲット装置のパフォーマンスに基づいて、前記複数の視差図のうちから、サイズが前記ターゲット装置のパフォーマンスとマッチした視差図を、前記ターゲット装置に提供する視差図として選択するように構成されても良い。例えば、前記ターゲット装置のパフォーマンスが比較的高く、及び/又は、その要する視差図の精度が比較的高い場合、前記複数の視差図のうちからサイズが比較的大きい視差図を選んで前記ターゲット装置に提供しても良い。また、前記ターゲット装置は、自身のパフォーマンスに基づいて、前記視差推定システムが得た前記複数の視差図のうちから、その要する視差図を自発的に選択しても良いが、これについて限定しない。

## 【 0 1 0 6 】

また、示されていないが、前記視差推定システムが得た前記複数の視差図はさらに、対応するターゲット装置に提供して更なる処理をしても貰っても良い。例えば、対応するターゲット装置に提供することで、前記ターゲット装置は、前記視差図に基づいて深度図を算出し、シーンの深度情報を得ることができる。このような深度情報は、例えば、3次元再構成、自動運転、障害物検知などの様々な応用シーンに適用することができる。

## 【 0 1 0 7 】

以上、図1～図8をもとに本開示の視差推定の例示的なシステムを説明した。以下、図9、図10及び図11に基づいて本開示の視差推定の例示的な方法及び例示的な電子装置の例示的な実施例についてさらに説明する。なお、前文に図1～図8を参照して説明した各種の定義、実施例、実施方式、実例などは、以下に説明する例示的な実施例に適用し又はそれと組み合わせることもできる。

## 【 0 1 0 8 】

幾つかの実施例によれば、図9は、本開示の例示的な実施例の視差推定方法のフローチャートを示している。図9に示すように、本開示の視差推定方法は次のようなステップを含んでも良く、即ち、画像対における各画像に対して特徴抽出を行い（ステップS901）；及び抽出した画像特徴に基づいて、カスケード接続される多段視差処理を行い、サイズが順次増大する複数の視差図を取得し（ステップS902）、ここで、前記多段視差処理のうちの第一段の視差処理の入力は、該段の視差処理に対応するサイズを有する複数の画像特徴を含み；前記多段視差処理のうちの前記第一段の視差処理以外の各段の視差処理の入力は、該段の視差処理に対応するサイズを有する1つ又は複数の画像特徴、及び1つ前の段の視

10

20

30

40

50

差処理により生成される視差図を含む。

【0109】

幾つかの実施例によれば、前記画像対は、マルチカメラが採取した同一シーンに対しての画像対であっても良い。前記画像対における各画像のサイズは一致しているが、対応する視角は異なる。また、前記画像対における各画像はグレースケール画像又はカラー画像であっても良い。

【0110】

幾つかの実施例によれば、抽出した前記画像対における各画像の画像特徴は少なくとも次のような特徴のうちの1つ又は複数を含んでも良く、即ち、基礎構造特徴、語義的特徴、エッジ特徴、テキストチャー特徴、色特徴、物体形状特徴、又は画像自身に基づく特徴である。例えば、抽出した前記画像対における第一画像（例えば、左画像）の画像特徴は、基礎構造特徴、語義的特徴及びエッジ特徴を含んでも良く、抽出した前記画像対における第二画像（例えば、右画像）の画像特徴は、基礎構造特徴を含んでも良い。或いは、抽出した前記画像対における第一画像及び第二画像の画像特徴はすべて、基礎構造特徴、語義的特徴及びエッジ特徴を含んでも良い。

【0111】

幾つかの実施例によれば、前記複数の視差図のうちの最大サイズを有する視差図は、前記画像対における各画像のサイズと一致しても良く、もちろん、前記複数の視差図のうちの各視差図のサイズは、すべて、前記画像対における各画像のサイズよりも小さくても良い。また、前記複数の視差図のうち、任意の2つの隣接視差図のうちの後ろの1つの視差図の高さ及び幅は、それぞれ、前の1つの視差図の高さ及び幅の2倍であっても良く、もちろん、実際に必要な精度に応じて、それぞれ、前の1つの視差図の高さ及び幅の3倍、4倍又は他の1よりも大きい正の整数倍などに設定しても良い。例示として、前記複数の視差図が4つあり、且つ前記複数の視差図のうちの最後の1つの視差図のサイズが $H \times W$ である（前記画像対における各画像のサイズと一致しても良い）ことを例にとると、その前の前記複数の視差図のうちの他の各視差図のサイズは、順に、 $(H/2) \times (W/2)$ （ $H \times W$ がフルサイズと称されれば、 $(H/2) \times (W/2)$ は1/2サイズと称される）、 $(H/4) \times (W/4)$ （1/4サイズと称される）、及び $(H/8) \times (W/8)$ （1/8サイズと称される）であっても良い。

【0112】

幾つかの実施例によれば、前記抽出した画像特徴は、N種類のサイズの画像特徴を含んでも良く、前記Nは、2以上の正の整数である。それ相応に、図10（図10は、本開示の例示的な実施例の視差推定方法における多段視差処理のフローチャートを示す図である）に示すように、抽出した画像特徴に基づいて、カスケード接続される多段視差処理を行い、サイズが順次増大する複数の視差図を取得することは、以下のステップを含んでも良い。

【0113】

ステップS1001：前記多段視差処理の第一段の視差処理において、前記N種類のサイズの画像特徴のうちの最小サイズの画像特徴の少なくとも一部に基づいて、前記最小サイズを有する初期視差図を生成する。

【0114】

例示として、抽出した前記N種類のサイズの画像特徴が1/8サイズ、1/4サイズ、1/2サイズ、及びフルサイズの4種類のサイズの画像特徴を含むことを例にとれば、前記多段視差処理の第一段の視差処理において、前記4種類のサイズの画像特徴のうちの最小サイズ（即ち、1/8サイズ）の画像特徴の少なくとも一部に基づいて、前記最小サイズ（即ち、1/8サイズ）を有する初期視差図を生成することができる。

【0115】

また、前述の表1及び表2に示すように、対応するサイズを有する対応する画像特徴に対して視差シフト重ね合わせを行い、3DCNNを用いて初期視差図を取得し、又は、シフト後の対応するサイズを有する画像特徴の差を計算し、2DCNNを用いて初期視差図を取得しても良い。

【0116】

ステップS1002：前記多段視差処理の後続の各段の視差処理において、前記N種類のサイズの画像特徴のうちの対応するサイズを有する画像特徴の少なくとも一部に基づいて、1つ前の段の視差処理により生成される視差図に対して視差最適化処理を行い、前記対応するサイズを有する最適化視差図を生成し、ここで、前記複数の視差図は少なくとも各最適化視差図を含む。

【0117】

幾つかの実施例によれば、前記多段視差処理はN+1段の視差処理を含んでも良い。それ相応に、前記の、前記多段視差処理の後続の各段の視差処理において、前記N種類のサイズの画像特徴のうちの対応するサイズを有する画像特徴の少なくとも一部に基づいて、1つ前の段の視差処理により生成される視差図に対して視差最適化処理を行い、前記対応するサイズを有する最適化視差図を生成し、ここで、前記複数の視差図が少なくとも各最適化視差図を含むことは、以下のことを含んでも良い。

10

【0118】

第一段の視差処理以外のN段の視差処理において、サイズの小さいから大への順序に従って、順に、前記N種類のサイズの画像特徴のうちの対応するサイズを有する画像特徴の少なくとも一部に基づいて、1つ前の段の視差処理により生成される視差図に対して視差最適化処理を行い、サイズが順次増大するN個の最適化視差図を取得し、そして、前記N個の最適化視差図を前記複数の視差図とし、ここで、前記N個の最適化視差図のサイズは、それぞれ、前記N種類のサイズに対応する。

【0119】

20

例示として、抽出した前記N種類のサイズの画像特徴が1/8サイズ、1/4サイズ、1/2サイズ、及びフルサイズの4種類のサイズの画像特徴を含み、且つ前記多段視差処理が4+1段の視差処理を含むことを例にとれば、第一段の視差処理以外の4段の視差処理において、サイズの小さいから大への順序に従って、順に、前記4種類のサイズの画像特徴のうちの対応するサイズを有する画像特徴の少なくとも一部に基づいて、1つ前の段の視差処理により生成される視差図に対して視差最適化処理を行い、4つのサイズが順次増大する最適化視差図（例えば、1/8サイズを有する最適化視差図、1/4サイズを有する最適化視差図、1/2サイズを有する最適化視差図、及びフルサイズを有する最適化視差図）を取得し、そして、前記4つの最適化視差図を前記複数の視差図とする。

【0120】

30

他の幾つかの実施例によれば、前記多段視差処理はN段の視差処理を含んでも良い。それ相応に、前記の、前記多段視差処理の後続の各段の視差処理において、前記N種類のサイズの画像特徴のうちの対応するサイズを有する画像特徴の少なくとも一部に基づいて、1つ前の段の視差処理により生成される視差図に対して視差最適化処理を行い、前記対応するサイズを有する最適化視差図を生成し、ここで、前記複数の視差図が少なくとも各最適化視差図を含むことは、以下のことを含んでも良い。

【0121】

第一段の視差処理以外のN-1段の視差処理において、サイズの小さいから大への順序に従って、順に、前記N種類のサイズの画像特徴のうちのN-1種類の非最小サイズの画像特徴のうちの対応するサイズを有する画像特徴の少なくとも一部に基づいて、1つ前の段の視差処理により生成される視差図に対して視差最適化処理を行い、N-1個のサイズが順次増大する最適化視差図を取得し、そして、前記初期視差図及び前記N-1個の最適化視差図を前記複数の視差図とし、ここで、前記初期視差図及び前記N-1個の最適化視差図のサイズは、それぞれ、前記N種類のサイズに対応する。

40

【0122】

例示として、抽出した前記N種類のサイズの画像特徴が1/8サイズ、1/4サイズ、1/2サイズ、及びフルサイズの4種類のサイズの画像特徴を含み、且つ前記多段視差処理が4段の視差処理を含むことを例にとると、第一段の視差処理以外の他の3段の視差処理において、サイズの小さいから大への順序に従って、順に、他の3種類の非最小サイズの画像特徴のうちの対応するサイズを有する画像特徴の少なくとも一部に基づいて、1つ前の段の視差処理

50

により生成される視差図に対して視差最適化処理を行い、3つのサイズが順次増大する最適化視差図（例えば、1/4サイズを有する最適化視差図、1/2サイズを有する最適化視差図、及びフルサイズを有する最適化視差図）を取得し、前記初期視差図及び前記3つの最適化視差図を前記複数の視差図とする。

【0123】

これにより、得られた前記複数の視差図は、前記多段視差処理のうちの第一段の視差処理が生成した初期視差図を含んでも良く含まなくても良いので、視差生成の柔軟性を向上させることができる。

【0124】

幾つかの実施例によれば、前記多段視差処理の後続の各段の視差処理において、前記N種類のサイズの画像特徴のうちの対応するサイズを有する画像特徴の少なくとも一部に基づいて、1つ前の段の視差処理により生成される視差図に対して視差最適化処理を行い、前記対応するサイズを有する最適化視差図を生成することは、以下のことを含んでも良い。

【0125】

前記多段視差処理のうちの前記第一段の視差処理以外の各段の視差処理において、対応するサイズを有する画像特徴の少なくとも一部に基づいて、1つ前の段の視差処理により生成される視差図に対して残差計算処理を行い、前記対応するサイズを有する残差図を取得し、そして、前記対応するサイズを有する残差図と、前記1つ前の段の視差処理により生成される視差図との組み合わせを行って前記対応するサイズを有する最適化視差図を取得する。

【0126】

例えば、抽出した前記N種類のサイズの画像特徴が1/8サイズ、1/4サイズ、1/2サイズ、及びフルサイズの4種類のサイズの画像特徴を含み、且つ前記多段視差処理が4+1段の視差処理を含むことを例にとれば、第一段の視差処理以外の他の4段の視差処理のうちの1/8サイズに対応する視差処理（即ち、1/8サイズに対応する視差最適化処理）において、抽出した1/8サイズの画像特徴の一部又は全部、及び1つ前の段の視差処理が生成した初期視差図に基づいて1/8サイズの第一残差図を算出し、そして、前記第一残差図及び前記初期視差図に基づいて1/8サイズの第一最適化視差図を計算することができる。次の段の1/4サイズに対応する視差最適化処理において、抽出した1/4サイズの画像特徴の一部又は全部、及び1つ前の段の視差処理が生成した第一最適化視差図に基づいて1/4サイズの第二残差図を算出し、そして、前記第二残差図及び前記第一最適化視差図に基づいて1/4サイズの第二最適化視差図を計算することができる。その他は、これに基づいて類推することができる。

【0127】

幾つかの実施例によれば、前記多段視差処理のうちの前記第一段の視差処理以外の各段の視差処理において、1つ前の段の視差処理により生成される視差図に対して視差最適化処理を行う前、前記方法はさらに、前記1つ前の段の視差処理により生成される視差図のサイズが現在の段の視差処理対応するサイズよりも小さいことに応じて、前記1つ前の段の視差処理により生成される視差図を現在の段の視差処理に対応するサイズにアップサンプリングすることを含んでも良い。

【0128】

例えば、依然として、抽出した前記N種類のサイズの画像特徴が1/8サイズ、1/4サイズ、1/2サイズ、及びフルサイズの4種類のサイズの画像特徴を含み、且つ前記多段視差処理が4+1段の視差処理を含むことを例にとれば、第一段の視差処理以外の他の4段の視差処理のうちの1/4サイズに対応する視差処理（即ち、1/4サイズに対応する視差最適化処理）において、前の段の視差処理が生成した1/8サイズの第一最適化視差図を現在の段の視差処理に対応する1/4サイズにアップサンプリングし、その後、抽出した1/4サイズの画像特徴の一部又は全部に基づいて、アップサンプリング後の1/4サイズの第一最適化視差図に対して視差最適化処理を行い、1/4サイズの第二最適化視差図を得ることができる。

## 【0129】

幾つかの実施例によれば、前記N種類のサイズの画像特徴のうちの最小サイズの画像特徴は、例えば、前記画像対における第一画像の少なくとも1種類の画像特徴及び第二画像の少なくとも1種類の画像特徴を含んでも良い。例えば、前記N種類のサイズの画像特徴のうちの最小サイズの画像特徴は、前記画像対における第一画像（例えば、左画像）の基礎構造特徴、語義的特徴、エッジ特徴及び前記画像対における第二画像（例えば、右画像）の基礎構造特徴を含んでも良い。

## 【0130】

前記N種類のサイズの画像特徴のうちの各種類の非最小サイズの画像特徴は、例えば、前記画像対における前記第一画像の少なくとも1種類の画像特徴及び/又は前記第二画像の少なくとも1種類の画像特徴を含んでも良い。例えば、前記N種類のサイズの画像特徴のうちの各非最小サイズの画像特徴は、前記画像対における第一画像のエッジ特徴又は前記第一画像自身に基づく特徴を含んでも良い。

10

## 【0131】

また、前述のシステムの実施例に記載のように、異なる最適化視差図の生成が基づく画像特徴は、同じ種類の画像特徴又は異なる種類の画像特徴であっても良く；及び/又は、異なる最適化視差図の生成が基づく画像特徴は、前記画像対における同じ画像又は異なる画像の画像特徴であっても良い。

## 【0132】

各最適化視差図の生成が基づく画像特徴は、例えば、前記画像対における少なくとも1つの画像のエッジ特徴、及び/又は、前記画像対における少なくとも1つの画像の画像自身に基づく特徴を含んでも良い。前記の、前記画像対における少なくとも1つの画像の画像自身に基づく特徴は、例えば、前記少なくとも1つの画像の画像自身、又は、生成する必要のある最適化視差図のサイズに基づいて前記少なくとも1つの画像の画像自身に対してダウンサンプリングを行うことにより得られた画像を含んでも良い。

20

## 【0133】

幾つかの実施例によれば、前記視差推定方法はさらに、前記多段視差処理のうちの各段の視差処理の損失関数を計算することを含んでも良く、前記損失関数は、該段の視差処理が生成した視差図の中の視差と、対応する真の視差との間の誤差を表すために用いることができる。このようにして、損失関数を計算することで、生成した各視差図の正確度を明確にすることができ、また、損失関数に基づいて視差推定方法に対して最適化を行うこともできる。

30

## 【0134】

幾つかの実施例によれば、前記視差推定方法はさらに、ターゲット装置のパフォーマンスに基づいて、前記複数の視差図のうちから、サイズが前記ターゲット装置のパフォーマンスとマッチした視差図を選んで前記ターゲット装置に提供するための視差図とすることを含んでも良い。例えば、前記ターゲット装置のパフォーマンスが比較的高く、及び/又は、その要する視差図の精度が比較的高い場合、前記複数の視差図のうちから、サイズが比較的大きい視差図を選択して前記ターゲット装置に提供することができる。また、前記ターゲット装置は、自身のパフォーマンスに基づいて、前記視差推定システムが得た前記複数の視差図のうちから、その要する視差図を自発的に選択しても良い。

40

## 【0135】

また、前記視差推定方法はさらに、前記画像対における各画像に対して画像特徴抽出を行う前に、前記画像対における画像に対してエピポーラ修正(epipolar rectification)を行うことで、前記画像対における画像に、1つのみの方向（例えば、水平方向）で視差を持たせることを含んでも良い。これにより、画像の視差サーチ範囲を1つのみの方向に限定することで、後続の特徴抽出及び視差生成の効率を向上させることができる。

## 【0136】

本開示の他の側面によれば、電子装置が提供され、該電子装置は、処理器；及びプログラムを記憶している記憶器を含み、前記プログラムは指令を含み、前記指令は、前記処理

50

器により実行されるときに、前記処理器に、前述の任意の方法を実行させることができる。

【0137】

本開示の他の側面によれば、プログラムを記憶している計算機可読記憶媒体が提供され、前記プログラムは指令を含み、前記指令は、電子装置の処理器により実行されるときに、前記電子装置に、前述の任意の方法を実行させることができる。

【0138】

以下、図11を参照しながら計算装置2000を説明する。それは、本開示の各側面に適用され得るハードウェア装置の例示である。計算装置2000は、処理及び/又は計算を実行するように構成される任意の装置であっても良く、例えば、ワークステーション、サーバ、デスクトップコンピュータ、ノートコンピュータ、タブレットコンピュータ、PDA、スマートフォン、車載コンピュータ又はその任意の組み合わせであっても良いが、これに限定されない。上述の電子機器は、計算装置2000又は類似装置もしくは類似システムにより全部又は部分的に実現することができる

計算装置2000は、バス2002に接続される又はバス2002と通信される素子を含んでも良い（なお、1つ又は複数のインターフェースを経由する可能性がある）。例えば、計算装置2000は、バス2002、1つ又は複数の処理器2004、1つ又は複数の入力装置2006、及び1つ又は複数の出力装置2008を含んでも良い。1つ又は複数の処理器2004は、任意の種類の処理器であっても良く、例えば、1つ又は複数の汎用処理器及び/又は1つ又は複数の専用処理器（例えば、特殊処理チップ）を含んでも良いが、これに限定されない。入力装置2006は、計算装置2000に情報を入力し得る任意の種類の装置であっても良く、例えば、マウス、キーボード、タッチパネル、マイク、及び/又はリモコンであっても良いが、これに限定されない。出力装置2008は、情報を表示し得る任意の種類の装置であっても良いが、例えば、表示器、スピーカー、ビデオ/音声出力端子、振動器、及び/又はプリンタであっても良いが、これに限定されない。計算装置2000はさらに、記憶装置2010を含んでも良く、又は記憶装置2010に接続されても良く、記憶装置は、非一時的で且つデータの記憶を実現し得る任意の記憶装置であっても良く、例えば、ディスクドライブ、光ストレージデバイス、ソリッドステートメモリ、フロッピーディスク、フレキシブルディスク、ハードディスク、テープ又はその他の磁気メディア、光ディスク又はその他の光メディア、ROM（読み取り専用メモリ）、RAM（ランダムアクセスメモリ）、キャッシュメモリ、及び/又はその他のメモリチップ又はカートリッジ、及び/又はコンピュータがデータ、命令、及び/又はコードを読み取ることができるその他の媒体であっても良いが、これに限定されない。記憶装置2010は、インターフェースに対して着脱可能である。記憶装置2010は、上述の方法及びステップを実現するためのデータ/プログラム（指令を含む）/コードを有しても良い。計算装置2000はさらに、通信装置2012を含んでも良い。通信装置2012は、外部装置及び/又はネットワークと通信し得る任意の種類の装置又はシステムであっても良く、例えば、モデム、ネットワークカード、赤外線通信装置、無線通信装置、及び/又はチップアセンブリ、例えば、ブルートゥース（登録商標）装置、1302.11装置、WiFi装置、Wimax装置、セルラー通信装置及び/又は類似物を含んでも良いが、これに限定されない。

【0139】

計算装置2000はさらに、ワーキング記憶器2014を含んでも良く、それは、処理器2004のワーキングに有用なプログラム（指令を含む）及び/又はデータを記憶する任意の種類のワーキング記憶器であっても良く、例えば、ランダムアクセスメモリ及び/又は読み取り専用メモリを含んでも良いが、これに限定されない。

【0140】

ソフトウェア要素（プログラム）がワーキング記憶器2014にあっても良く、それは、オペレーティングシステム2016、1つ又は複数のアプリ（アプリケーション、即ち、応用プログラム）2018、駆動プログラム、及び/又はその他のデータ及びコードを含んでも良いが、これに限定されない。上述の方法及びステップを実行するための指令は、1つ又は複数のアプリ2018に含まれても良く、また、上述のネットワーク200、300は、処理器2004が

10

20

30

40

50

該1つ又は複数のアプリ2018の指令を読み取り及び実行することにより実現することができる。より具体的には、上述のネットワーク200は、ステップS901を実行する指令を有するアプリ2018を実行することにより実現することができる。また、上述のネットワーク300は、ステップS902を実行する指令を有するアプリ2018を実行することにより実現することができる。ソフトウェア要素（プログラム）の指令の実行可能なコード又はソースコードは、非一時的コンピュータ可読記憶媒体（例えば、上述の記憶装置2010）に記憶され、且つ実行時にワーキング記憶器2014に格納され得る（コンパイル及び/又はインストールされ得る）。また、ソフトウェア要素（プログラム）の指令の実行可能なコード又はソースコードは、リモート位置からダウンロードすることもできる。

**【0141】**

なお、具体的な要求に応じて様々な変形を行っても良い。例えば、カスタム化ハードウェアを用いても良く、及び/又は、ハードウェア、ソフトウェア、ファームウェア、ミッドウェア、マイクロコード、ハードウェア記述言語又はその任意の組み合わせにより特定の素子を実現することもできる。例えば、開示される方法及び装置における幾つか又は全部は、本開示の論理及びアルゴリズムに基づいて、アセンブリ言語又はハードウェアプログラミング言語（例えば、VERILOG、VHDL、C++）を用いて、ハードウェア（例えば、FPGA及び/又はPLAを含むプログラム可能な論理回路）に対してプログラミングを行うことにより実現されても良い。

**【0142】**

また、理解すべきは、前述の方法がサーバ-クライアントの方式により実現されても良いということである。例えば、クライアントは、ユーザが入力したデータを受信し、前記データをサーバに送信することができる。クライアントは、ユーザが入力したデータを受信し、前述の方法における一部の処理を行い、そして、処理により得られたデータをサーバに送信することもできる。サーバは、クライアントからのデータを受信し、前述の方法又は前述の方法における一部を実行し、そして、実行の結果をクライアントに返すことができる。クライアントは、サーバから方法の実行の結果を受信し、また、例えば、出力装置によりユーザに示すことができる。

**【0143】**

さらに、理解すべきは、計算装置2000の部品がネットワークに分散しても良いということである。例えば、1つの処理器が幾つかの処理を実行し、また、該1つ処理器を離れる他の処理器が他の処理を実行しても良い。計算装置2000の他の部品も同様に分散しても良い。このように、計算装置2000は、複数の位置で処理を実行する分散型計算システムと見なされても良い。

**【0144】**

なお、図面を参照しながら本開示の実施例又は実施方式を説明したが、理解すべきは、上述の方法、システム及び装置は例示的な実施例又は実施方式に過ぎず、本開示の範囲はこれらの実施例又は実施方式により限定されず、特許請求の範囲及びその均などな範囲により限定されるということである。実施例又は実施方式における各種類の要素は、省略されても良く又は他の均などな要素により置換されても良い。また、本開示に述べた順序とは異なる順序に従って各ステップを実行しても良い。さらに、各種類の方式で実施例又は実施方式における各種類の要素を組み合わせても良い。重要なのは、技術が進化するにつれて、ここで記載されている多くの要素を、本開示の後に現れる均などな要素により置き換えることができるということである。

10

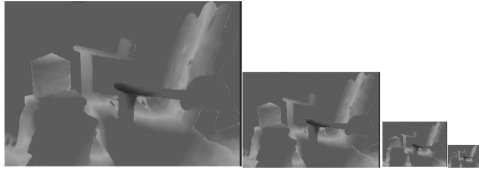
20

30

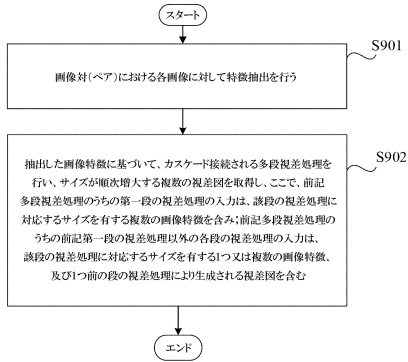
40



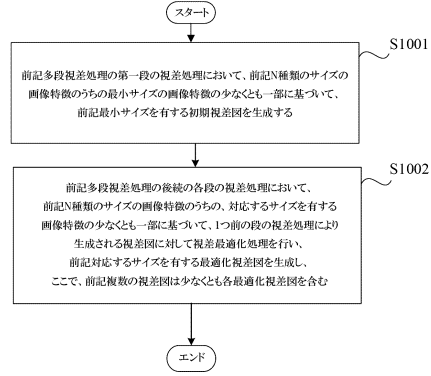
【図8】



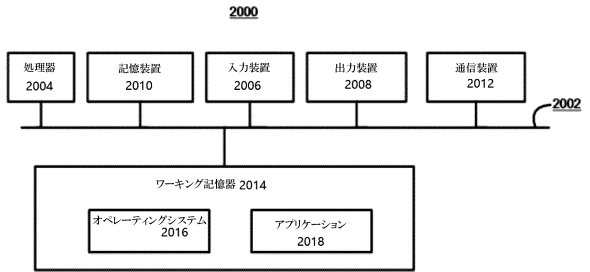
【図9】



【図10】



【図11】



## フロントページの続き

(74)代理人 100107766

弁理士 伊東 忠重

(74)代理人 100070150

弁理士 伊東 忠彦

(74)代理人 100135079

弁理士 宮崎 修

(72)発明者 シュウ ファーン

中華人民共和国 201203 シャンハイ, プドオン ニュー ディストリクト, ロンドン  
ロード, #3000, ビルディング 1, セクション エー, ルーム607

(72)発明者 ジィ ジョウ

中華人民共和国 201203 シャンハイ, プドオン ニュー ディストリクト, ロンドン  
ロード, #3000, ビルディング 1, セクション エー, ルーム607

(72)発明者 シンボン フォン

中華人民共和国 201203 シャンハイ, プドオン ニュー ディストリクト, ロンドン  
ロード, #3000, ビルディング 1, セクション エー, ルーム607

審査官 藤原 敬利

(56)参考文献 特表2017-520852(JP, A)

特開2019-121349(JP, A)

特開2019-096294(JP, A)

(58)調査した分野(Int.Cl., DB名)

G01B 11/00 - 11/30

G06N 3/00 - 3/12

G06N 7/08 - 99/00

G06T 7/00 - 7/90