

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6153134号
(P6153134)

(45) 発行日 平成29年6月28日 (2017. 6. 28)

(24) 登録日 平成29年6月9日 (2017. 6. 9)

(51) Int. Cl.

G 0 6 T 7/00 (2017.01)

F I

G 0 6 T 7/00

Z

請求項の数 19 外国語出願 (全 16 頁)

(21) 出願番号	特願2014-71893 (P2014-71893)	(73) 特許権者	000006013
(22) 出願日	平成26年3月31日 (2014. 3. 31)		三菱電機株式会社
(65) 公開番号	特開2014-203459 (P2014-203459A)		東京都千代田区丸の内二丁目7番3号
(43) 公開日	平成26年10月27日 (2014. 10. 27)	(74) 代理人	100110423
審査請求日	平成28年12月7日 (2016. 12. 7)		弁理士 曾我 道治
(31) 優先権主張番号	13/855, 800	(74) 代理人	100111648
(32) 優先日	平成25年4月3日 (2013. 4. 3)		弁理士 梶並 順
(33) 優先権主張国	米国 (US)	(74) 代理人	100122437
早期審査対象出願			弁理士 大宅 一宏
		(74) 代理人	100147566
			弁理士 上田 俊一
		(74) 代理人	100161171
			弁理士 吉田 潤一郎
		(74) 代理人	100161115
			弁理士 飯野 智史

最終頁に続く

(54) 【発明の名称】 2つ以上の光源の組によって照光されたシーンの2つ以上の入力画像の組から照明基底画像の組を復元する方法

(57) 【特許請求の範囲】

【請求項 1】

2つ以上の光源の組によって照光されたシーンの2つ以上の入力画像の組から照明基底画像の組を復元する方法であって、

各光源は固定であり、

該方法は、前記入力画像の組を前記照明基底画像と指標係数の組とに因数分解することを含み、

各照明基底画像は、前記光源のうちの1つのみによって照光された前記シーンの外観に対応する非負の画像であり、

各指標係数は、バイナリ照明指標行列の要素であり、

各指標係数は、前記入力画像のうちの1つに対する前記光源のうちの1つの寄与を示し、

前記入力画像の組は固定カメラによって取得され、

各入力画像において、前記シーンは前記光源の組合せによって照光され、

前記入力画像の組内の少なくとも2つの画像が前記光源の異なる組合せによって照光され、

前記光源のうちの1つ又は複数から前記シーンへの距離は、該シーンの大きさよりも1桁又は複数桁未満だけ大きく、

前記復元すること及び前記因数分解することはプロセッサにおいて実行される、方法。

【請求項 2】

前記シーンは屋内環境にある、請求項 1 に記載の方法。

【請求項 3】

各指標係数は 0 又は 1 である、請求項 1 に記載の方法。

【請求項 4】

前記指標係数は非負の実数である、請求項 1 に記載の方法。

【請求項 5】

前記分解することは因数分解することを含み、

前記因数分解することは、準バイナリ非負行列因数分解 (SBNMF) を用いて、非負の照明基底画像の行列と指標係数のバイナリ行列とを生成する、請求項 1 に記載の方法。

10

【請求項 6】

制約付き最適化：

【数 1】

$$\min_{V, W} \|Y - VW\|_F^2, \text{ such that } V_{ij} \geq 0, W_{jk} \in \{0, 1\}, \forall i, j, k$$

を用いて前記準バイナリ非負行列因数分解を解くことを更に含み、式中、min は最小値を返す関数であり、Y は前記入力画像の組を表す行列であり、V は前記非負の照明基底画像の行列であり、W は前記指標係数のバイナリ行列であり、i はピクセルロケーションをインデックス付けし、j は前記光源をインデックス付けし、k は前記入力画像をインデックス付けする、請求項 5 に記載の方法。

20

【請求項 7】

連続緩和：

【数 2】

$$\min_{V, W} \|Y - VW\|_F^2, \text{ such that } V_{ij} \geq 0, 0 \leq W_{jk} \leq 1, \forall i, j, k$$

によって前記最適化を解くことを更に含み、式中、 W_{jk} に対するバイナリ制約はボックス制約に置き換えられている、請求項 6 に記載の方法。

30

【請求項 8】

交互方向乗数法によって前記連続緩和を解くことを更に含む、請求項 7 に記載の方法。

【請求項 9】

前記照明基底画像の組を編集することであって、それにより編集された基底画像の組を生成することと、

前記編集された基底画像の組をマージすることであって、それにより出力画像を生成することと、

を更に含む、請求項 1 に記載の方法。

【請求項 10】

前記編集することは線形関数を用いる、請求項 9 に記載の方法。

【請求項 11】

前記編集することは非線形関数を用いる、請求項 9 に記載の方法。

40

【請求項 12】

前記マージすることは、前記編集された基底画像の組の線形結合である、請求項 9 に記載の方法。

【請求項 13】

特定の照明基底画像内のシャドウ領域及びハイライト領域を検出することと、

前記シャドウ領域及び前記ハイライト領域のそれぞれにおけるピクセルの輝度を、異なる照明基底画像内の同一ロケーションにある領域内の前記ピクセルの前記輝度から導出された値と置き換えることと、

を更に含む、請求項 9 に記載の方法。

50

【請求項 1 4】

特定の照明基底画像内の領域のテクスチャを、対応する前記編集された基底画像内の異なるテクスチャに置き換えることと、

前記特定の照明基底画像において影になっている、前記領域内のピクセルの組を検出することと、

前記特定の照明基底画像内の前記ピクセルの組と同一のロケーションにある、前記編集された基底画像内の前記ピクセルを暗くすることと、

を更に含む、請求項 9 に記載の方法。

【請求項 1 5】

前記カメラから前記シーンの一部分への距離は、該シーンの大きさよりも 1 桁又は複数桁未満だけ大きい、請求項 1 に記載の方法。

【請求項 1 6】

前記画像の組は、低速度撮影写真を用いて取得される、請求項 1 に記載の方法。

【請求項 1 7】

前記カメラはビデオカメラであり、前記画像の組は、前記ビデオカメラによって取得されたビデオからの画像を含む、請求項 1 に記載の方法。

【請求項 1 8】

前記入力画像の組における 3 次元幾何学境界を検出することを更に含む、請求項 1 に記載の方法。

【請求項 1 9】

前記シーンは屋外環境にある、請求項 1 に記載の方法。

【発明の詳細な説明】**【技術分野】****【0001】**

本発明は、包括的にはコンピュータビジョンに関し、より詳細には、照明変動の影響下にあるシーンの画像を基底画像に因数分解することに関する。

【背景技術】**【0002】**

エッジ検出は、コンピュータビジョンにおける基本的な問題である。エッジ検出は、多くの用途にとって重要な低レベルの特徴を提供する。シーンの画像内のエッジは、深さの不連続性、面の向きの差異、面テクスチャ、材料特性の変化及び変動する照明を含む様々な原因から生じる可能性がある。

【0003】

多くの方法は、エッジを、個々の画像内の明度、色及びテクスチャ等の低レベル画像特性の変化としてモデル化する。しかし、面の深さ又は向きにおける別個の変化である 3 D 幾何学境界に対応する画像ピクセルを識別する問題は、あまり注目されてこなかった。

【0004】

Raskar は、特許文献 1 において、マルチフラッシュカメラを用いてシャドウに基づく技法を適用することによって深さエッジを検出する。その方法は、深さの不連続性のみに適用され、面法線における変化には適用されず、カメラのレンズを取り囲む制御された 1 組のライト (light) を必要とする。

【0005】

3 D 幾何学境界は、セグメンテーション、シーン分類、3 D 再構成及びシーンレイアウト復元を含む多岐にわたるタスクの有用な手がかりを提供することができるシーンの特性を正確に表す。

【0006】

非特許文献 1 において、Weiss 他は、照光変化を受けるシーンの画像シーケンスについて記載している。シーケンス内の各画像は、一定の反射率の単一画像と、画像固有の照光画像との積に因数分解される。

【0007】

10

20

30

40

50

特許文献 2 は、屋外シーンの低速度撮影の写真シーケンスを、シャドウ成分と、照光成分と、反射率成分とに因数分解することを記載しており、これによってシーンのモデル化及び編集への適用を容易にすることができる。その方法は、経時的に平滑的に移動する無限遠の単一点光源（太陽）と、周囲照明成分とを仮定している。

【 0 0 0 8 】

非特許文献 2 において、Koppal 他は、シーンの回りで遠方にある光源を振り動かすことによって取得される画像シーケンスについて記載している。これらの画像は、その後、同様の面法線を有する領域にクラスタリングされる。その作業も、位置が経時的に平滑的に変動する単一の遠方にある光源と、正投影カメラモデルとを仮定している。

【 先行技術文献 】

【 特許文献 】

【 0 0 0 9 】

【 特許文献 1 】 米国特許第 7 , 2 9 5 , 7 2 0 号明細書

【 特許文献 2 】 米国特許第 7 , 7 5 6 , 3 5 6 号明細書

【 非特許文献 】

【 0 0 1 0 】

【 非特許文献 1 】 Weiss 他, 「Deriving intrinsic images from image sequences」, ICCV 2001, Volume: 2, Page(s): 68 - 75, vol.2

【 非特許文献 2 】 Koppal 他, 「Appearance derivatives for isonormal clustering of scenes」, IEEE TPAMI, 31(8):1375-1385, 2009

【 発明の概要 】

【 発明が解決しようとする課題 】

【 0 0 1 1 】

本発明は、変動する照光条件下で固定カメラによって取得される固定シーン（屋内環境又は屋外環境にあることができる）の画像の組を検討する。本発明の 1 つの目的は、画像の組から 3 D 幾何学境界を検出することである。

【 0 0 1 2 】

本発明の別の目的は、これらの画像を基底画像の組に因数分解することである。これらの用途において、光源の位置は知られておらず、ライトは必ずしも点光源であるとは限らず、ライト（及びカメラ）からシーンまでの距離は、必ずしもシーンの大きさよりもはるかに大きい（すなわち 1 桁又は複数桁大きい）とは限らないので、無限遠であると仮定することができない。これによって、測光立体視、構造化光及び等法線クラスタリング等の変動する照明下で 2 D 画像から 3 D 構造を復元する既存の方法、並びに因数分解された低速度撮影ビデオ及び固有画像等の画像の組から照光の効果を因数分解する既存の方法の仮定が崩れる。

【 0 0 1 3 】

本発明の実施形態は、未知の変動する照光条件下で、固定のカメラ視点を用いて捕捉された静止シーンの画像の組から照明基底画像を復元する方法を提供する。いくつかの実施形態では、3 D 幾何学境界が検出される。

【 0 0 1 4 】

いくつかの実施形態の 1 つの目的は、未知の変化する照光の影響下にある静的シーン（屋内環境にあることができる）の 2 D 画像の組において 3 D 幾何学境界を特定することである。本明細書において厳密に定義されるように、シーンの画像内で観察されるような 3 D 画像幾何学境界は、シーン内の 2 つの面を分離する輪郭であり、この輪郭において、3 D 深さ不連続性又は面の向きの変化が存在する。これらの境界は、シーンの 3 D レイアウトを理解するのに効果的に用いることができる。3 D 幾何学境界は、テクスチャエッジ又はシャドウエッジ等の 2 D エッジと異なる。

【 0 0 1 5 】

3 D 幾何学境界を 2 D テクスチャエッジと区別するために、本発明のいくつかの実施形態は、各画像ロケーションにおける局所外観の照光部分空間を解析する。これは、非鏡面

10

20

30

40

50

、例えばランバート面の場合、同じ平滑な3D面上の近傍ピクセルが、それらのピクセルが異なる色、すなわち異なるアルベド又は反射係数を有する場合があるにもかかわらず、照明に対し同じ相対感度を有する傾向があるという認識に基づく。その理由は、小さな近傍領域において、3D面は局所的に平面であり、隣接するピクセルに対応する面上の2つの点は、概ね同じ法線を有するためである。これらの2つの点間の距離は通常、光源及びカメラのいずれに対する距離よりもはるかに小さい。

【0016】

この認識に基づいて、3D幾何学境界検出方法は、3D幾何学境界に関連付けられたピクセル、すなわち、画像内のその直接近傍が面法線又は深さにおいて不連続性を含むピクセルを、その近傍がシャープなテクスチャ又は輝度境界を含む場合があるが単一の面に対応するピクセルと区別することができる。

10

【0017】

本方法は、各ピクセルロケーションにおいて局所外観の照光部分空間を解析することによって、ピクセルごとの分類問題として3D幾何学境界検出を定式化する。特に、本方法は、照光部分空間の次元を用いて3D幾何学境界の存在を示す。

【0018】

本発明の1つの目的は、固定光源の組の組合せに起因して未知で変化する照光の影響下にある静的シーンの画像の組から照明基底画像の組を求めることである。照明基底画像とは、シーンが個々の光源のうちの1つによって照光されているときに形成される画像である。光源は点光源である必要はない。基底画像は、従来技術と比較して質的及び量的改善を伴ってシーンの自然で簡潔な表現を提供し、シーン編集（再照明等）並びにシャドウエッジの識別及び除去を可能にする。

20

【0019】

本発明のいくつかの実施形態では、照明基底画像を復元する方法は、準バイナリ非負行列因数分解（SBNMF：semi-binary nonnegative matrix factorization）を用いる。SBNMFは非負行列因数分解（NMF：nonnegative matrix factorization）に関する。NMFは、非負データ行列を、2つの非負行列の積に因数分解する。このための多くの技法が知られている。

【0020】

NMFと異なり、SBNMFは非負データ行列を因数分解して非負行列とバイナリ行列との積にする。ここで、バイナリ行列は、各要素が0又は1のいずれかである行列である。すなわち、本方法は、画像を含む非負行列を、照明基底画像の非負行列と、画像ごとにいずれの光源がオン又はオフであることを示すバイナリ重み行列とに因数分解する。照明基底画像の復元された組は、変動する照明下のシーンのコンパクトな表現を提供する。

30

【0021】

いくつかの実施形態では、基底画像を、3D幾何学境界検出方法と併せて用いて、シャドウエッジを真の3D幾何学エッジと区別することができる。

【図面の簡単な説明】

【0022】

【図1】本発明のいくつかの実施形態による、1組の画像を1組の照明基底画像に因数分解する方法の概略図である。

40

【図2】準バイナリ非負行列因数分解を用いた図1の方法のブロック図である。

【図3】図2の方法の概略図である。

【図4】準バイナリ非負行列因数分解（SBNMF）の擬似コードのブロック図である。

【図5】本発明のいくつかの実施形態によって用いられるランバート面の概略図である。

【図6】本発明のいくつかの実施形態による幾何学境界検出方法の擬似コードのブロック図である。

【図7】本発明のいくつかの実施形態による、シャドウエッジが除去される幾何学境界検出方法の擬似コードのブロック図である。

【図8】本発明の一実施形態による、シーン編集用途のブロック図である。

50

【発明を実施するための形態】

【0023】

図1に示すように、本発明のいくつかの実施形態は、変動する照明条件の影響下にあるシーンから取得された画像の組 Y 101を因数分解して照明基底画像の組 V 102にする方法110を提供する。シーンは複数の光源によって照明され、それらの光源のそれぞれは、特定の画像についてオン又はオフにすることができる。組 V 内の各基底画像は、単一の光源（例えば、ライト1、ライト2、．．．、又はライト5）のみがオンであるときのシーンの外観に対応する。 Y 内の画像は、任意の時点に又はシーケンスでビデオカメラによって取得することができることを理解すべきである。また、「照明」は、赤外線（IR：infrared）、紫外線（UV：ultraviolet）及び電磁スペクトル内の他の周波数等の他の形式の加法性の「照光」様式を含むことができることも理解すべきである。

10

【0024】

図2に示すように、本発明のいくつかの実施形態の場合、画像の組101は、準バイナリ非負行列因数分解（SBNMF）を用いて、非負の照明基底画像 V 201と、バイナリ照明指標行列 W 202とに因数分解される（210）。図3は因数分解を概略的に示している。

【0025】

生成的画像モデル

シーンを照光する1個の光源が存在し、各光源は独立したオン/オフスイッチによって制御される。一緒に全てがオンであるか又は全てがオフである2つ以上のライトからなるグループが存在する（グループ内のライトのうちの1つがオンである一方、そのグループ内の別のライトがオフである入力画像が存在しないようになっている）場合、ライトのグループは単一の光源であるとみなされる。例えば、同じスイッチによって制御される2つのライトは単一の光源とみなされる。

20

【0026】

各光源 i の状態を示すバイナリ変数 w_i を割り当てる。次に、基底画像

【数1】

$$v_i \in \mathbb{R}^{+n}$$

を、 i 番目のライトのみがオンであるときに形成される画像として定義する。式中、

30

【数2】

$$\mathbb{R}^{+}$$

は非負の実数を表す。1個の光源の任意の組合せによって照光される画像 y を所与とすると、画像は個々の基底画像の重ね合わせとして表すことができる。

【0027】

【数3】

$$y = \sum_{i=1}^l w_i v_i, w_i \in \{0,1\} \quad (1)$$

【0028】

ここで、画像内の全てのピクセルの値を単一の列にスタックすることによって形成される列ベクトルとして全ての画像を表す。

40

【0029】

1個の光源の様々なバイナリ組合せによって照明される m 個の画像の組 Y を取得し、画像データを配列して以下の行列にする。

【0030】

【数4】

$$Y = [y_1, y_2, \dots, y_m] \in \mathbb{R}^{+n \times m}$$

【0031】

50

式(1)に従って、このデータ行列は以下のように因数分解することができる。

【0032】

【数5】

$$Y = VW \quad (2)$$

【0033】

式中、

【数6】

$$V \in \mathbb{R}^{+n \times l}$$

の列は1個の基底画像 v_i に対応し、 $W \in \{0, 1\}^{l \times m}$ はバイナリ指標行列であり、このバイナリ指標行列において、各エントリ W_{ij} は、図2及び図3に示すように、j番目の入力画像に対するi番目の光源の寄与を表す指標係数である。すなわち、全ての入力画像はm個の照明基底画像のバイナリの組合せである。

10

【0034】

シーン内に周囲照明(全ての画像内に存在する光)が存在する場合、本発明によるモデルにおいて、これは追加の基底画像(行列V内の追加の列)と、要素が全て1である指標行列Wの対応する追加の行とによってモデル化することができることに留意されたい。

【0035】

いくつかの実施形態において、2つの入力画像が取得される時点間に単一の光源が異なる位置に移される場合、その光源を2つの別個の光源とみなす。例えば、数時間離れて取得された2つの入力画像がともに太陽によって照光されているとき、第1の画像における太陽を第1の光源とみなし、第2の画像における太陽を第2の光源とみなす。

20

【0036】

SBNMFによる基底画像の復元

いくつかの実施形態では、SBNMFを用いて入力画像の組から照明基底画像及び指標行列を復元する。真の照明基底画像が線形に独立しており、かつ十分な照光変動性を観測する、すなわち真の指標行列Wの階数がライト数以上である場合、シーン内のライト数1は、データ行列Yの階数によって与えられる。

【0037】

基底画像及び指標行列の復元を、制約された最適化問題として定式化する。

30

【0038】

【数7】

$$\min_{V, W} \|Y - VW\|_F^2, \text{ such that } V_{ij} \geq 0, W_{jk} \in \{0, 1\}, \forall i, j, k \quad (3)$$

【0039】

これをSBNMFと呼ぶ。これは、非凸目的関数とWに対するバイナリ制約とに起因した困難な問題である。したがって、最初に以下の連続緩和を解く。

【0040】

【数8】

$$\min_{V, W} \|Y - VW\|_F^2, \text{ such that } V_{ij} \geq 0, 0 \leq W_{jk} \leq 1, \forall i, j, k \quad (4)$$

40

【0041】

ここで、 W_{jk} に対するバイナリ制約は、上界及び下界の単純なボックス制約に置き換えられる。これは、交互方向乗数法(ADMM: alternating direction method of multipliers)を用いて解く両凸問題である。ADMMは従来の拡張ラグランジュ法の一変形である。以下を参照されたい。

【0042】

補助変数Xを用いて式(4)を書き換え、正值性制約及びボックス制約を指標関数と置

50

き換える。

【 0 0 4 3 】

【 数 9 】

$$\min_{\mathbf{X}, \mathbf{V}, \mathbf{W}} \quad \|\mathbf{Y} - \mathbf{X}\|_F^2 + \mathbf{I}_{[0, \infty)}(\mathbf{V}) + \mathbf{I}_{[0, 1]}(\mathbf{W}) \quad (5)$$

such that $\mathbf{X} - \mathbf{V}\mathbf{W} = \mathbf{0}$

【 0 0 4 4 】

式中、指標関数 $\mathbf{I}_S(x)$ は、 $x \in S$ の場合に値 0 をとり、他の全ての場合に 1 になる。

10

【 0 0 4 5 】

次に、以下の拡張ラグランジュを形成する。

【 0 0 4 6 】

【 数 1 0 】

$$L(\mathbf{X}, \mathbf{V}, \mathbf{W}, \mathbf{U}) = \|\mathbf{Y} - \mathbf{X}\|_F^2 + \mathbf{I}_{[0, \infty)}(\mathbf{V}) + \mathbf{I}_{[0, 1]}(\mathbf{W}) \quad (6)$$

$$+ (\mu/2) \|\mathbf{X} - \mathbf{V}\mathbf{W} + \mathbf{U}\|_F^2 - (\mu/2) \|\mathbf{U}\|_F^2$$

【 0 0 4 7 】

式中、 \mathbf{U} はスケーリングされた双対変数であり、 μ は拡張ラグランジュパラメーターである。ここでは、拡張ラグランジュ関数のスケーリングされた形式を用いる。また、ここではスケーリングされたラグランジュ乗数が $\mathbf{U} = \mathbf{Z} / \mu$ として再定義される。式中、 \mathbf{Z} は元のラグランジュ乗数である。

20

【 0 0 4 8 】

ADMM は、凸部分問題の組によって拡張ラグランジュ双対関数を解く。ここで、両凸関数は以下に分離される。

【 0 0 4 9 】

【 数 1 1 】

$$(\mathbf{X}^{k+1}, \mathbf{V}^{k+1}) = \underset{\mathbf{X}, \mathbf{V} \geq 0}{\operatorname{argmin}} \|\mathbf{X} - \mathbf{Y}\|_F^2 + (\mu/2) \|\mathbf{X} - \mathbf{V}\mathbf{W}^k + \mathbf{U}^k\|_F^2, \quad (7)$$

30

$$\mathbf{W}^{k+1} = \underset{0 \leq \mathbf{W} \leq 1}{\operatorname{argmin}} \|\mathbf{X}^{k+1} - \mathbf{V}^{k+1}\mathbf{W} + \mathbf{U}\|_F^2 \quad (8)$$

$$\mathbf{U}^{k+1} = \mathbf{U}^k + \mathbf{X}^{k+1} - \mathbf{V}^{k+1}\mathbf{W}^{k+1} \quad (9)$$

【 0 0 5 0 】

これらの部分問題は、主残余及び双対残余の収束まで反復して解かれる。

【 0 0 5 1 】

それに続いて、行列 \mathbf{W} の各エントリを $\{0, 1\}$ に丸め、非負最小二乗を用いてバイナリ指標行列に基づいて基底画像 \mathbf{V} を求める。

40

【 0 0 5 2 】

【 数 1 2 】

$$\min_{\mathbf{V}} \|\mathbf{Y} - \mathbf{V}\mathbf{W}\|_F^2 \quad \text{such that } \mathbf{V}_{ij} \geq 0, \forall i, j. \quad (10)$$

【 0 0 5 3 】

最適化 (10) において \mathbf{W} は定数であるので、問題は凸であることに留意されたい。

【 0 0 5 4 】

図 4 は、本発明による因数分解手順の疑似コードを示している。疑似コードとして示さ

50

れるこの手順及び他の手順における擬似コードにおいて参照される変数及び式は、全て本明細書において詳述される。

【 0 0 5 5 】

他の実施形態において、個々の光源はオン又はオフのみに制約されず、光源の輝度は連続的に変化することができる（例えばディマースイッチが用いられる）ことに留意されたい。この場合、指標行列 W における指標係数は、バイナリ値 $\{0, 1\}$ に制約されるのではなく、任意の非負の実数とすることができる。この場合、全ての入力画像は m 個の照明基底画像の非負の線形結合である。そのような場合、因数分解は従来の非負行列因数分解を用いて行うことができる。

【 0 0 5 6 】

3 D 幾何学境界の検出

いくつかの実施形態において、シーンから取得された画像の組において 3 D 幾何学境界を検出する。シーンの画像において観察されるように、3 D 幾何学境界は、シーン内の 2 つの面を分離する輪郭であり、この輪郭において、3 D 深さの不連続性が存在するか、又は面法線の大きな変化が存在する。通常の屋内シーンの場合、遠方にある照明の仮定は有効でない。近傍の照明を考慮するには、一度に 1 つの小さな画像パッチを検討し、複数の照明条件とともにそのパッチの局所外観がどのように変動するかを解析する。本方法は、多岐にわたる形状及び大きさを有するパッチを用いることができる。例えば、各画像ピクセルを中心とする固定の直径（例えば 7 ピクセル）を有する正方形又は円形のパッチを検討することができる。

【 0 0 5 7 】

パッチ内の全てのピクセルがシーン内の単一の平滑な面から生じる場合、変動する照明にわたるパッチ外観は 1 次元部分空間を形成する。パッチが 3 D 幾何学境界を含む場合、パッチの外観部分空間は、概して 2 以上の次元を有する。

【 0 0 5 8 】

局所外観の照光部分空間

簡単にするために、本方法を、直接照明成分のみを有するランバート面について証明するが、類似した議論が、より広い種類（クラス）の反射率関数及び間接照明、例えば複数の反射に適用される。説明を簡単にするために、点光源についてのみ説明する。なぜなら、拡張等方性光源は、複数の点光源の重ね合わせとして適宜良好に近似することができるためである。

【 0 0 5 9 】

図 5 (a)、図 5 (b) 及び図 5 (c) は、カメラ 5 0 0 によって観察されるような 3 つのランバート面 5 0 1 ~ 5 0 3 を概略的に示している。これらのランバート面はそれぞれ 2 つの点光源 A 及び B によって照光される。3 つ全ての面について、カメラによって取得される画像内の点 1 及び 2 に対応するピクセルは互いに近くにある。図 5 (a) において、面 5 0 1 内に 3 D 幾何学境界は存在しない。図 5 (b) において、面 5 0 2 の面法線において不連続性が存在する。図 5 (c) において、カメラの視点から面 5 0 3 の深さの不連続性が存在する。

【 0 0 6 0 】

光源 A の場合のここでの表記を説明する。なお、光源 B の場合の表記は、それに類似する。点 i における面法線は

【 数 1 3 】

$$\hat{\mathbf{n}}_i$$

であり、点 i からライト A へのベクトルは \mathbf{r}_i^a （対応する単位ベクトルは

【 数 1 4 】

$$\hat{\mathbf{r}}_i^a$$

）である。面の点 i に対応する画像平面上の点の輝度は I_i^a （光源 A の場合）又は I_i

10

20

30

40

50

^b（光源 B の場合）である。

【 0 0 6 1 】

【 数 1 5 】

$$I_i^a = \gamma_i^a \frac{\hat{\mathbf{n}}_i^T \hat{\mathbf{r}}_i^a}{\|\mathbf{r}_i^a\|^2} E^a \rho_i, \quad I_i^b = \gamma_i^b \frac{\hat{\mathbf{n}}_i^T \hat{\mathbf{r}}_i^b}{\|\mathbf{r}_i^b\|^2} E^b \rho_i. \quad (11)$$

【 0 0 6 2 】

ここで、

【 数 1 6 】

$$\hat{\mathbf{n}}_i^T \hat{\mathbf{r}}_i^a$$

は

【 数 1 7 】

$$\hat{\mathbf{n}}_i$$

と \mathbf{r}_i^a との間の角度のコサインである。 E^a は光源 A の放射輝度であり、 i は点 i における面アルベドである。点 i が光源 A によって照光されている場合、バイナリ値 $i^a = 1$ であるのに対し、点 i が陰 (attached shadow) 又は影 (cast shadow) に起因して光源 A によって照光されていない場合、 $i^a = 0$ である。

【 0 0 6 3 】

3 つの面のそれぞれについて、点 1 及び点 2 はカメラの観点から互いの近くにあり、このため、これらの点はともに同じ小さな画像パッチ内に含まれる。図 5 (a) において、パッチは法線の急な変化を含まず、深さの不連続性を含まない。このため、点 1 と点 2 との間の 3 D 距離は、各点から各ライトへの距離と比較して小さく、このため、以下の近似等式が得られる。

【 0 0 6 4 】

【 数 1 8 】

$$\hat{\mathbf{n}}_1 \approx \hat{\mathbf{n}}_2, \quad \mathbf{r}_1^a \approx \mathbf{r}_2^a, \quad \mathbf{r}_1^b \approx \mathbf{r}_2^b \quad (12)$$

【 0 0 6 5 】

パッチ内の全ての点がほぼ同じ法線及び各光源への同じベクトルを共有するので、式 (1 1) における上付き文字 i を消去し、パッチ内の全ての点について、

【 数 1 9 】

$$\hat{\mathbf{n}}$$

、 \mathbf{r}^a 及び \mathbf{r}^b を用いることができる。今のところ、パッチ内の全ての点 i が i^a の単一の値（ここでは a と呼ぶ）を共有し、 i^b の単一の値 b を共有すると仮定する。これは、光源ごとに、パッチ全体が照光されるか又はその光の影になる (shadowed)、すなわちパッチがシャドウエッジを含まないことを意味する。以下においてシャドウエッジを検討する。

【 0 0 6 6 】

P^a 及び P^b が、それぞれライト A のみ及びライト B のみの下で撮像されたパッチのピクセル輝度のベクトルを表すものとする。図 5 (a) の場合、下記の近似等式 $P^a = k^a$ を有する。

【 0 0 6 7 】

10

20

30

40

【数 2 0】

$$\begin{bmatrix} I_1^a \\ I_2^a \\ \vdots \\ \mathbf{P}^a \end{bmatrix} \approx \underbrace{\frac{\gamma^a E^a \hat{\mathbf{n}}^T \hat{\mathbf{r}}^a}{\|\mathbf{r}^a\|^2}}_{k^a} \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \boldsymbol{\rho} \end{bmatrix} \quad (13)$$

【0068】

式中、スカラー k^a はパッチ内の全てのピクセルの定数であり、 $\boldsymbol{\rho}$ はパッチ内の全てのピクセル面アルベドのベクトルである。光源 B 下の同じパッチの場合、類似した式 $\mathbf{P}^b = k^b \boldsymbol{\rho}$ を有する。

10

【0069】

このため、パッチが法線においても深さにおいても急な変化を含まない（そしてシャドウエッジを含まない）場合、任意の光源下のピクセル輝度は、 k^a のスカラー倍に等しい。換言すれば、全ての光源下でその局所パッチの外観によって広がる部分空間（局所外観の照光部分空間と呼ぶ）は 1 次元（1D: one-dimensional）である。これは、面テクスチャ（アルベド）に関わらず真であることに留意されたい。パッチの面アルベドが高コントラストのテクスチャエッジを含む場合であっても、局所外観のその照光部分空間は依然として 1D である。

20

【0070】

この実現は、幾何学エッジを見つける本発明による方法の核心となっている。なぜなら、パッチが 3D 幾何学エッジを含む場合、同じことは概ね当てはまらないためである。

【0071】

例えば、図 5（b）に示すように、パッチが法線において急激な変化を含む場合、式（12）における第 1 の概算は成り立たず、パッチ内の各点の輝度は、その面法線と、光源へのその方向との間の角度のコサインに依拠する。パッチが異なるタイプの 3D 幾何学境界、図 5（c）におけるような深さの不連続性を含む場合、（12）における最後の 2 つの近似は成り立たず（光が無遠慮でないため）、パッチ内の各点の輝度は、その面法線と、光源へのその方向との間の角度のコサインに依拠する。概して、パッチが 3D 幾何学エッジを含む場合、局所外観のその照光部分空間は 2 以上の次元を有する。

30

【0072】

3D 幾何学境界の信頼度マップ

いくつかの実施形態では、局所外観の照光部分空間が 2 以上の次元を有するパッチを特定することによって幾何学境界を検出する。ピクセルロケーションごとに、 m 個全ての入力画像（ m 個のライト組合せ）から、そのロケーションを中心とした $n \times m$ の行列 Z 内の列ベクトルとして配列する。

【0073】

【数 2 1】

$$\mathbf{Z} = [\mathbf{P}^{(1)}, \mathbf{P}^{(2)}, \dots, \mathbf{P}^{(m)}] \quad (14)$$

40

【0074】

式中、ベクトル $\mathbf{P}^{(i)}$ は、そのピクセルロケーションにおいて画像 i から抽出したパッチの n 個全てのピクセル（色又は輝度）値を含む。そのパッチロケーションの局所外観の照光部分空間の階数を求めるために、 Z に特異値分解（SVD: singular value decomposition）を適用し、特異値 $\{\sigma_i^P\}$ （降順に順序付けされる）を得る。雑音がない場合、1 次元照光部分空間は 1 つのみの非ゼロ特異値 σ_1^P をもたらし、ここで $\sigma_2^P = 0$ である。画像内の雑音に起因して、 σ_2^P は正確に 0 でないが、概ね 0 である。局所外観の照光部分空間が階数 1 を有するか否かを判断するために、雑音の存在下で正確であるこ

50

との信頼値を用いる。

【 0 0 7 5 】

いくつかの実施形態では、ピクセルロケーションごとに、対応するパッチが 3 D 幾何学境界を含むことの信頼値を、そのロケーションを中心とするパッチについて、次式のように、第 1 の特異値に対する第 2 の特異値の割合として求める。

【 0 0 7 6 】

【 数 2 2 】

$$c(P) = \sigma_2^P / \sigma_1^P \quad (15)$$

【 0 0 7 7 】

式 (1 5) を用いて信頼度マップを得る。信頼度マップとは、各ピクセルの輝度が、そのピクセルロケーションについて求められた信頼値である画像である。

【 0 0 7 8 】

他の実施形態では、局所外観の照光部分空間が 2 以上の階数を有することの信頼値を、式 (1 5) 以外の方法で計算することができる。例えば、 $c(P)$ に、 $c(P) = \sigma_2^P / k$ 等の特異値の何らかの他の関数を定義することができ、式中、 k は全てのパッチの照光部分空間の特異値から求められた正規化係数である。ここでの 3 D 幾何学境界検出手順の擬似コードは図 6 に示される。

【 0 0 7 9 】

1 つの実施形態では、 m 個の元の入力画像からパッチを抽出するのではなく、上記で説明した 1 個の非負の照明基底画像からパッチを抽出する。これは、入力画像の組が不平衡である場合、例えば、多数の入力画像が単一の照明条件から生じ、他の照明条件から生じる入力画像が僅かしかない場合、より安定した手法となることができる。

【 0 0 8 0 】

シャドウエッジの除去

本発明による方法は、双方のタイプの 3 D 幾何学境界、すなわち法線における不連続性及び深さにおける不連続性の検出に成功する。本明細書において、双方のタイプが「境界」として特徴付けられる。加えて、本発明による方法は、テクスチャエッジによって混乱しない。しかしながら、図 6 に概説した方法によって、シャドウエッジが誤検出 3 D 幾何学境界として検出される可能性がある。光源のうちの 1 つについて、パッチのいくつかのピクセルが照光され、他のピクセルが影になっている場合、パッチはシャドウエッジを含む。

【 0 0 8 1 】

ほとんどの場合に、各シャドウエッジは単一の光源にのみによって生じる。この観測に基づいて、シーンの画像の組を単一光源照明基底画像に因数分解する本発明による機能を用いて、シャドウエッジによって生じた誤検出のほとんどを取り除くことができる。

【 0 0 8 2 】

次式のように、画像の組 Y から基底画像 v_i を減算することによって、光源 i によって生成されるシャドウを取り除くことができる。

【 0 0 8 3 】

【 数 2 3 】

$$Y^{(i)} = Y - v_i w^i \quad (16)$$

【 0 0 8 4 】

式中、 w^i は照明指標行列 W の i 番目の行であり、 $Y^{(i)}$ は、ライト i がオフにされた状態で再レンダリングされるシーン画像を表す。

【 0 0 8 5 】

本発明による境界検出技法を $Y^{(i)}$ に適用すると、結果として境界信頼度マップ $C^{(i)}$ が得られる。境界信頼度マップ $C^{(i)}$ では、 i 番目の光源の結果得られるシャドウエッジが取り除かれている。最終的な信頼度マップは、

【数 2 4】

$$\{C^{(i)}\}_{i=1}^l$$

の全ての信頼度マップの中で各ピクセルロケーションにおいて最小値をとることによって集約され、それによって、光源のうちの任意の 1 つが除去されたときにシャドウエッジが消える場合、そのエッジは最終信頼度マップ内に存在しないことになる。

【0086】

シャドウエッジが除去された本発明による境界検出手順の擬似コードが図 7 に示される。

【0087】

10

一実施形態では、式 (16) を用いて $Y^{(i)}$ を元の画像の再レンダリングされたバージョンに等しく設定するのではなく、代わりに、 i ごとに、 $Y^{(i)}$ を照明基底画像 i 以外の全ての照明基底画像の組に等しく設定する。この実施形態では、 i ごとに、照明基底画像の低減された組 $Y^{(i)}$ は $1 - 1$ 個の照明基底画像を含む。これは、入力画像の組が不平衡である場合、例えば、多数の入力画像が単一の照明条件から生じ、他の照明条件から生じる入力画像が僅かしかない場合、より安定した手法となることができる。

【0088】

シーン編集

図 8 に示すように、本発明の実施形態を用いることができる 1 つの用途はシーン編集である。照明基底画像 102 は、取得 (入力) 画像ではなく、個々の照明基底画像を直接変更することによるシーン編集を可能にする。

20

【0089】

個々の照明基底画像を編集して、スケーリング等の線形関数、又はヒストグラム等化、補正、トーンマッピング、若しくは明度及びコントラスト調整等の非線形関数とすることができる編集関数 710 を適用することによって編集された基底画像 712 を生成することができる。加えて、オブジェクトの挿入又はテクスチャの変更等、照明基底画像の領域を編集することが可能である。出力画像 720 は、線形結合等のマージ関数 715 を編集された基底画像に適用することによって構築される。編集関数は、照明基底画像の全て又は一部に適用することができる。

【0090】

30

別の用途において、照明基底画像内のシャドウ領域及びハイライト領域は、照明基底画像内でほとんど黒色の領域及び飽和領域をそれぞれ見つけることによって検出することができる。そのような領域の輝度 (色) 値は、値が黒でなく、飽和もしていない、異なる照明基底画像内の同一のロケーションからコピーすることができる。次に、値は、画像値が元のシャドウ境界及びハイライト境界をまたがって連続するように変更される。この手法は、画像全体を飽和させることも暗くすることもなく、シーン内の暗いシャドウ領域及びハイライト領域を取り除くことができる。

【0091】

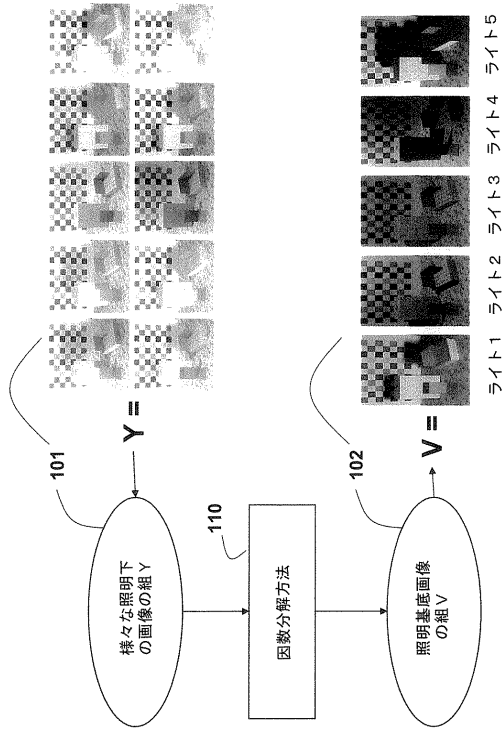
検出されたシャドウ領域を用いて、照光情報に従いながら、面のテクスチャを異なるテクスチャと置き換えることができる。新たなテクスチャのうち、影になっている (under shadow) 部分の明度はシャドウ情報に合致するように暗くされる。

40

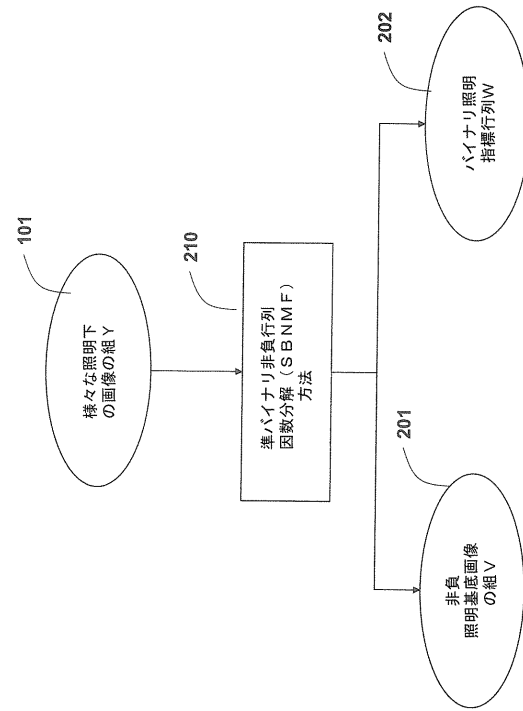
【0092】

本明細書において説明し示した方法のステップは、当該技術分野において既知のメモリ及び入 / 出力インタフェースに接続されたプロセッサにおいて実行することができる。通常のデジタル画像は、何百万ものピクセルを含み、この多大な量の輝度又は色値を頭の中で処理することは不可能であることが理解される。

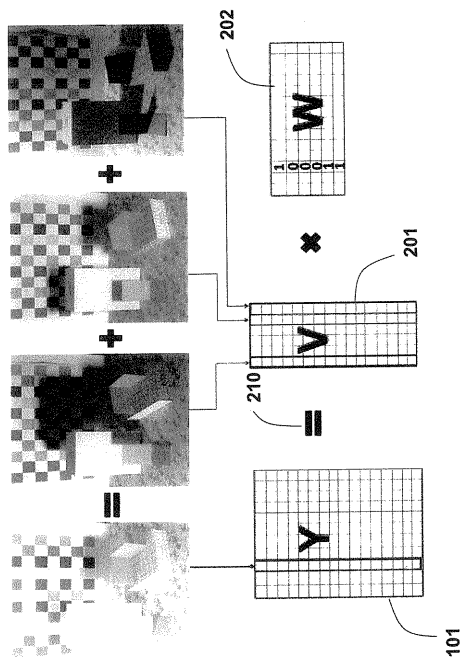
【図 1】



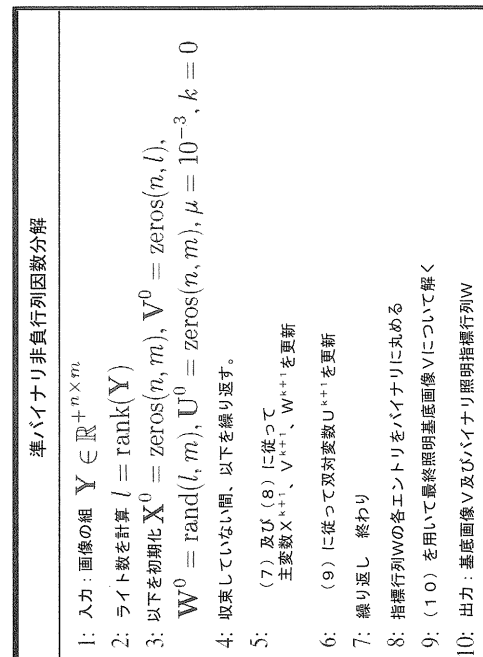
【図 2】



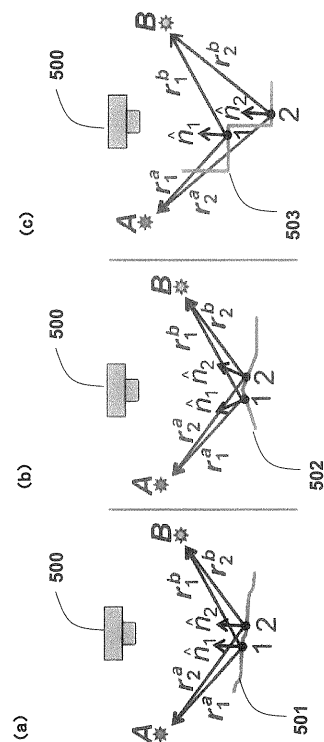
【図 3】



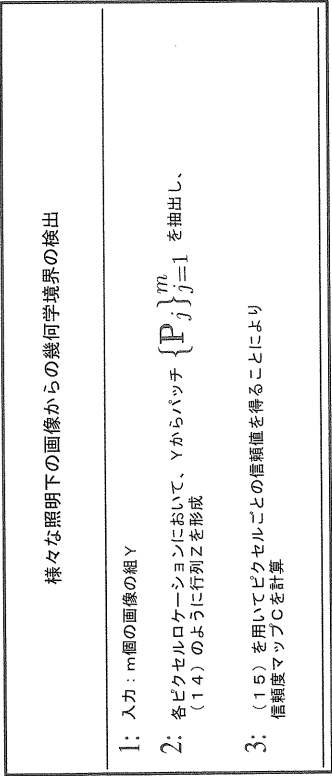
【図 4】



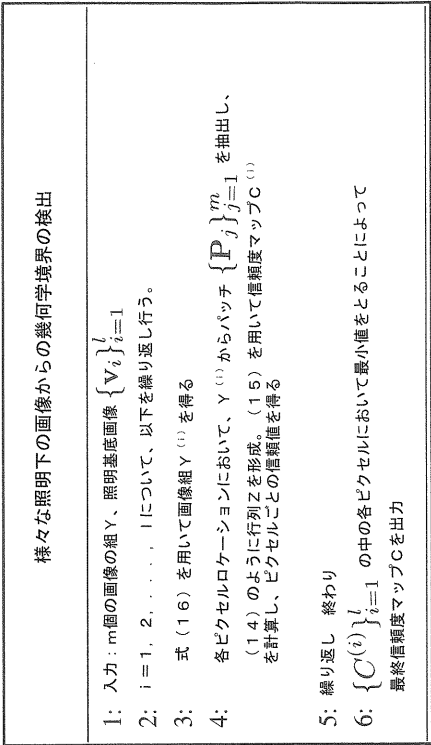
【図 5】



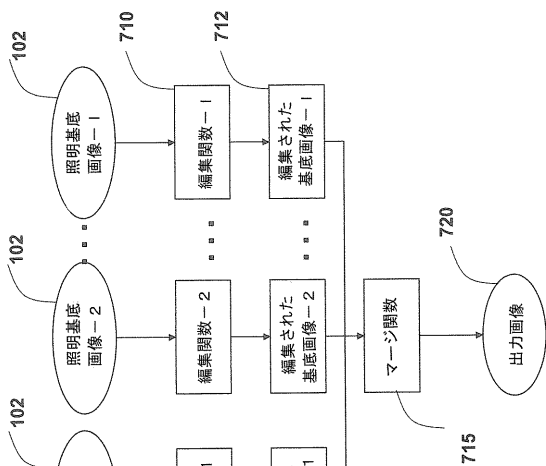
【図 6】



【図 7】



【図 8】



フロントページの続き

- (72)発明者 オンセル・チュゼル
アメリカ合衆国、マサチューセッツ州、ケンブリッジ、ブロードウェイ 201、ミツビシ・エレクトリック・リサーチ・ラボラトリーズ・インコーポレイテッド
- (72)発明者 ティム・ケイ・マークス
アメリカ合衆国、マサチューセッツ州、ケンブリッジ、ブロードウェイ 201、ミツビシ・エレクトリック・リサーチ・ラボラトリーズ・インコーポレイテッド
- (72)発明者 ファティー・ポリクリ
アメリカ合衆国、マサチューセッツ州、ケンブリッジ、ブロードウェイ 201、ミツビシ・エレクトリック・リサーチ・ラボラトリーズ・インコーポレイテッド
- (72)発明者 ジエ・ニ
アメリカ合衆国、マサチューセッツ州、ケンブリッジ、ブロードウェイ 201、ミツビシ・エレクトリック・リサーチ・ラボラトリーズ・インコーポレイテッド

審査官 広 島 明芳

- (56)参考文献 特開2007-281937(JP,A)
特開2004-164571(JP,A)
米国特許出願公開第2004/0001059(US,A1)
特開2011-103082(JP,A)
特開2012-022120(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06T 7/00 - 7/90