

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7692224号
(P7692224)

(45)発行日 令和7年6月13日(2025.6.13)

(24)登録日 令和7年6月5日(2025.6.5)

(51)国際特許分類

F I

G 0 6 T 7/00 (2017.01)

G 0 6 T 7/00 C

G 0 6 T 7/00 3 5 0 C

請求項の数 7 (全30頁)

(21)出願番号	特願2023-558118(P2023-558118)	(73)特許権者	505072650
(86)(22)出願日	令和3年6月7日(2021.6.7)		浙江大学
(65)公表番号	特表2024-511083(P2024-511083 A)		Z H E J I A N G U N I V E R S I T Y
(43)公表日	令和6年3月12日(2024.3.12)		中国浙江省杭州市西湖区余杭塘路866号
(86)国際出願番号	PCT/CN2021/098576	(74)代理人	100128347
(87)国際公開番号	WO2022/256962		弁理士 西内 盛二
(87)国際公開日	令和4年12月15日(2022.12.15)	(72)発明者	▲呉▼ ▲鴻▼智
審査請求日	令和5年9月20日(2023.9.20)		中国 3 1 0 0 5 8 浙江省杭州市西湖区
前置審査			余杭▲塘▼路866号
		(72)発明者	周 昆
			中国 3 1 0 0 5 8 浙江省杭州市西湖区
			余杭▲塘▼路866号
		(72)発明者	▲馬▼ ▲暁▼▲鶴▼
			中国 3 1 0 0 5 8 浙江省杭州市西湖区
			最終頁に続く

(54)【発明の名称】 高次元マテリアルの自由式収集方法

(57)【特許請求の範囲】

【請求項1】

高次元マテリアルの自由式収集方法であって、トレーニング段階及び収集段階を含み、前記トレーニング段階は以下のステップ（1）～ステップ（2）を含み、前記トレーニング段階のステップ（1）では、収集装置のパラメータを取得して、実際のビデオカメラを模擬する収集結果をトレーニングデータとして生成し、前記トレーニング段階のステップ（2）では、生成されたトレーニングデータを用いてニューラルネットワークをトレーニングし、前記ニューラルネットワークの特徴は以下の特徴（2.1）～特徴（2.5）を含み、前記特徴（2.1）については、ニューラルネットワークの入力がk個の非構造化サンプリングにおけるL u m i t e x e lベクトルであり、kがサンプリング個数であり、L u m i t e x e lの各値が各光源からの入射光に対するサンプリング点のある観察方向に沿う反射強度を示し、L u m i t e x e lが光源の発光強度と線形関係を有し、線形完全接続層で模擬を行い、前記特徴（2.2）については、ニューラルネットワークの第1層が線形完全接続層を含み、実際に収集する際に用いた照明パターンを模擬して、前記k個のL u m i t e x e lをカメラ収集結果に変換することに用いられ、このk個の収集結果がそれぞれ対応するサンプリング点のポーズ情報と組み合わせて高次元点群を構成し、前記特徴（2.3）については、第2層からは特徴抽出ネットワークであり、前記高次元点群における各点から独立して特徴抽出して特徴ベクトルを取得し、前記特徴抽出ネッ

トワークの公式は、以下の式 9 に示され、

【数 9】

$$V_{feature}(j) = f(\text{concat}[B(I, P^j), x_p^j, \hat{n}_p^j, \hat{t}_p^j]), 1 \leq j \leq k$$

f が一次元畳み込み関数であり、畳み込みカーネルのサイズが 1×1 であり、 $B(I, P^j)$ が第 1 層のネットワークから出力される結果又は収集して取得された測定値を示し、
 x_p^j 、 $\hat{n}_p^j$ 、 $\hat{t}_p^j$

10

がそれぞれ j 回目にサンプリングするときのサンプリング点の空間位置、サンプリング点の幾何学的法線ベクトル及び幾何学的接ベクトルであり、

$\hat{n}_p$

が幾何学的モデルにより取得され、

$\hat{t}_p$

が

$\hat{n}_p$

20

に直交する任意の単位ベクトルであり、ビデオカメラのポーズを j 回目にサンプリングすることにより

$\hat{n}_p$ 、 $\hat{t}_p$

を変換して

$\hat{n}_p^j$ 、 $\hat{t}_p^j$

30

を取得することができ、 $V_{feature}(j)$ がネットワークから出力される j 回目にサンプリングする特徴ベクトルであり、

前記特徴 (2. 4) については、特徴抽出ネットワークの後では、k 個の非構造化画像から抽出された前記特徴ベクトルを集約して、グローバル特徴ベクトルを取得するための最大プーリング層であり、

前記特徴 (2. 5) については、最大プーリング層の後では、非線形マッピングネットワークであり、前記非線形マッピングネットワークは、前記グローバル特徴ベクトルに基づいて形骸化され、前記非線形マッピングネットワークの形骸化された表現は、以下の式 11 に示され、

40

【数 11】

$$y_{i+1}^d = f_{i+1}^d(y_i^d W_{i+1}^d + b_{i+1}^d), i \geq 1$$

$$y_{i+1}^s = f_{i+1}^s(y_i^s W_{i+1}^s + b_{i+1}^s), i \geq 1$$

f_{i+1} が第 i + 1 層のネットワークのマッピング関数であり、 W_{i+1} が第 i + 1 層のネットワークのパラメータ行列であり、 b_{i+1} が第 i + 1 層のネットワークのオフセットベクトルであり、 y_{i+1} が第 i + 1 層のネットワークの出力であり、d と s がそれぞれ乱

50

反射及び鏡面反射の2つの分岐を示し、入力 y_1^d と y_1^s が最大プーリング層から出力されるグローバル特徴ベクトルであり、

前記収集段階は、以下のステップ（１）及びステップ（２）を含み、

前記収集段階のステップ（１）では、マテリアル収集ステップであって、収集装置が前記照明パターンに応じて順に目標三次元物体とするサンプリング物体を照射し、ビデオカメラが１組の非構造化画像における写真を取得し、写真を入力とし、前記サンプリング物体にテクスチャ座標を持つ幾何学的モデル及び写真撮影時のビデオカメラのポーズを取得し、

前記収集段階のステップ（２）では、マテリアル回復ステップであって、前記マテリアル収集ステップにおける写真撮影時のビデオカメラのポーズに基づいて各枚の写真を撮影するときのサンプリング物体上の各有効テクスチャ座標の頂点に対するポーズを取得し、収集された前記写真及びポーズ情報に基づいて前記高次元点群をニューラルネットワークの第２層の特徴抽出ネットワークの入力として構成し、最終層の出力ベクトルとする乱反射特徴ベクトル及び鏡面反射特徴ベクトルを計算して取得し、

前記非構造化サンプリングは非固定視角の自由ランダムサンプリングであり、サンプリングデータが無秩序で、不規則で、不均一に分布しており、固定物体を用いて人の手で収集装置を持って収集し、又は物体を回転テーブルに置いて回転させて固定装置により収集してもよいことを特徴とする高次元マテリアルの自由式収集方法。

【請求項２】

トレーニングデータ生成過程において、光源がカラーである場合、光源、サンプリング物体及びビデオカメラのスペクトル応答関係を補正する必要があり、補正方法については、未知のカラー光源 L のスペクトル分布曲線を

$$S_{c_1}^L(\lambda)$$

として定義し、 λ が波長を示し、 c_1 が RGB の３つのチャンネルのうちの１つを示し、光強度 $\{I_R, I_G, I_B\}$ の光源のスペクトル分布曲線 $L(\lambda)$ は、以下の式２に示されることが可能であり、

【数２】

$$L(\lambda) = I_R S_R^L(\lambda) + I_G S_G^L(\lambda) + I_B S_B^L(\lambda)$$

任意のサンプリング点 p の反射スペクトル分布曲線 $p(\lambda)$ を、係数がそれぞれ p_R 、 p_G 、 p_B の３つの未知の基

$$S_{c_2}^p(\lambda)$$

の線形組合せに示し、 c_2 が RGB の３つのチャンネルのうちの１つを示し、具体的には、以下の式３に示され、

【数３】

$$p(\lambda) = p_R S_R^p(\lambda) + p_G S_G^p(\lambda) + p_B S_B^p(\lambda)$$

ビデオカメラ C のスペクトル分布曲線が

$$S_{c_3}^C(\lambda)$$

の線形組合せとして示され、光強度 $\{I_R, I_G, I_B\}$ の光源の照射下で、ビデオカメラによる反射係数 $\{p_R, p_G, p_B\}$ のサンプリング点のある特定チャンネル c_3 での測定値は、以下の式４に示され、

【数 4】

$$\int L(\lambda)p(\lambda)S_{c_3}^C(\lambda)d\lambda = \sum_{c_1, c_2} I_{c_1}p_{c_2} \delta(c_1, c_2, c_3)$$

$$\delta(c_1, c_2, c_3) = \int S_{c_1}^L(\lambda)S_{c_2}^p(\lambda)S_{c_3}^C(\lambda)d\lambda$$

10

照明条件 $\{I_R, I_G, I_B\} = \{1, 0, 0\} / \{0, 1, 0\} / \{0, 0, 1\}$ の下で、既知の反射係数 $\{p_R, p_G, p_B\}$ のカラーテストカードを撮影し、ビデオカメラにより収集された測定値に基づいて線形連立方程式を構築し、サイズ $3 \times 3 \times 3$ のカラー補正行列 $\delta(c_1, c_2, c_3)$ を求め、該カラー補正行列が光源、サンプリング物体及びビデオカメラのスペクトル応答関係を示すことを特徴とする請求項 1 に記載の高次元マテリアルの自由式収集方法。

【請求項 3】

前記トレーニング段階における特徴 (2. 1) において、物体表面のサンプリング点 p の写真における観測値 B 、反射関数 f_r 及び各光源の光強度の関係は、以下の式 5 に示され、

20

【数 5】

$$B(l, P) = \sum_l I(l) \int \frac{1}{||x_l - x_p||^2} \Psi(x_l, -\omega'_i) V(x_l, x_p) f_r(\omega'_i; \omega'_o, P) (\omega'_i \cdot n_p)^+ (-\omega'_i \cdot n_l)^+ dx_l$$

ここで、 I は各光源 l の発光情報を示し、光源 l の空間位置 x_l 、光源 l の法線ベクトル n_l 、光源 l の発光強度 $I(l)$ を含み、 P はサンプリング点 p のパラメータ情報を含み、サンプリング点の空間位置 x_p 、マテリアルパラメータ n 、 t 、 α_x 、 α_y 、 ρ_d 、 ρ_s を含み、

30

$$\Psi(x_l, \cdot)$$

が光源 l の異なる入射方向における光強度分布を説明し、 V が x_l の x_p に対する可視性の二値関数を示し、

$$(\cdot)^+$$

が 2 つのベクトルのドット積操作であり、 $f_r(\omega'_i; \omega'_o, P)$ は ω'_o が一定である場合に ω'_i に関する二次元反射関数であり、

40

ニューラルネットワークの入力が k 個の非構造化サンプリングにおける $Lumitex$ e_l ベクトルであって、 $m(l; P)$ と記され、

【数 6】

$$m(l; P) = B(\{I(l) = 1, \forall_{i \neq l} I(i) = 0\}, P)$$

上記式における B が単一チャンネルにおける表示であり、光源がカラー光源である場合、 B を下記形式に拡張し、

【数 7】

50

$$B(I, P; c_3) = \sum_{l, c_1, c_2} I(l; c_1) \int \frac{1}{||x_l - x_p||^2} \Psi(x_l, -\omega_i') V(x_l, x_p) f_r(\omega_i'; \omega_o', P, c_2) \\ (\omega_i' \cdot n_p)^+ (-\omega_i' \cdot n_l)^+ \delta(c_1, c_2, c_3) dx_l$$

ここで、 $f_r(\omega_i'; \omega_o', P, c_2)$ が $f_r(\omega_i'; \omega_o', P)$ における
 $\rho_d = \rho_{d_{c_2}}$ 、 $\rho_s = \rho_{s_{c_2}}$

10

の結果であることを特徴とする請求項2に記載の高次元マテリアルの自由式収集方法。

【請求項4】

前記ニューラルネットワークの損失関数は以下のステップ(1)～ステップ(4)のよ
 うに設計され、

当該ステップ(1)では、1つの *Lumitexel* 空間を仮定し、該空間は中心が
 サンプルング点の空間位置 x_p にある立方体であり、立方体の中心座標系は x 軸方向が

$\hat{t}_p$

であり、 z 軸方向が

$\hat{n}_p$

20

であり、

$\hat{n}_p$

が幾何学的法線ベクトルであり、

$\hat{t}_p$

30

が

$\hat{n}_p$

に直交する任意の単位ベクトルであり、

当該ステップ(2)では、1つのビデオカメラを仮定し、観察方向が立方体の z 軸の正
 方向であり、

当該ステップ(3)では、乱反射 *Lumitexel* の場合に立方体の分解能が $6 \times N_d^2$ であり、鏡面反射 *Lumitexel* の場合に立方体の分解能が $6 \times N_s^2$ であり、即
 ち各面から N_d^2 、 N_s^2 個の点を、光強度が単位光強度の仮想点光源として均一にサンプ
 リングし、以下のステップa～ステップcを含み、

40

前記ステップaでは、サンプルング点の鏡面反射率 ρ_s を0とし、この *Lumitexel*
el 空間における乱反射特徴ベクトル

$\tilde{m}_d$

を生成し、

前記ステップbでは、乱反射率 ρ_d を0とし、この *Lumitexel* 空間における鏡
 面反射特徴ベクトル

50

$\tilde{m}_s$

を生成し、

前記ステップ c では、ニューラルネットワークの出力をベクトル m_d 、 m_s とし、 m_d が $\tilde{m}_d$

の長さと同じであり、 m_s が

$\tilde{m}_s$

10

の長さと同じであり、ベクトル m_d 、 m_s がそれぞれ乱反射特徴ベクトル

$\tilde{m}_d$

、鏡面反射ベクトル

$\tilde{m}_s$

20

の予測であり、

当該ステップ (4) では、マテリアル特徴部分の損失関数は、以下の式 12 に示され、

【数 12】

$$L = \lambda_d \sum_l [m_d(l) - \tilde{m}_d(l)]^2 + \lambda_s \beta \sum_l [m_s(l) - \log(1 + \tilde{m}_s(l))]^2$$

ここで、 λ_d と λ_s がそれぞれ m_d 、 m_s の損失重みを示し、信頼度 β が鏡面反射 Limitexel の損失を評価することに用いられ、log がベクトルの各次元に作用し、

信頼度 β の決定は、以下の式 13 に示され、

【数 13】

$$\beta = \min\left(\frac{1}{\epsilon} \max_j \left[\frac{\max_l \log(1 + f_r(\omega_i^{j'}(l); \omega_o^{j'}, P))}{\max_{\omega_i'} \log(1 + f_r(\omega_i'; \omega_o^{j'}, P))} \right], 1\right), 1 \leq j \leq k$$

30

ここで、

$$\max_l \log(1 + f_r(\omega_i^{j'}(l); \omega_o^{j'}, P))$$

項が j 回目にサンプリングする全ての単一光源レンダリング値の最大値の対数を示し、

$$\max_{\omega_i'} \log(1 + f_r(\omega_i'; \omega_o^{j'}, P))$$

40

項が j 回目にサンプリングする際に理論的に取得できる単一光源レンダリング値の最大値の対数を示し、

ϵ

が比率調整係数であることを特徴とする請求項 1 に記載の高次元マテリアルの自由式収集

50

方法。

【請求項 5】

前記マテリアル収集ステップにおいて、マテリアル収集が終了した後に幾何学的位置合わせを行い、その後にマテリアル回復を行い、幾何学的位置合わせは具体的に、スキャナにより物体を走査して幾何学的モデルを取得し、それを三次元で再構築された幾何学的モデルに位置合わせした後に三次元で再構築された幾何学的モデルを代替することであることを特徴とする請求項 1 に記載の高次元マテリアルの自由式収集方法。

【請求項 6】

有効テクスチャ座標に対して、前記収集された写真及びサンプリング点のポーズ情報に基づいて写真における画素を順に取り出して、画素の有効性を判断し、頂点に対するポーズと組み合わせて高次元点群を構成し、ある有効テクスチャ座標に対して決定されたサンプリング物体の表面上の点 p について、 j 回目のサンプリングが頂点 p に対して有効である判断標準は以下の条件 (1) ~ (3) の 3 つの条件を含み、

前記条件 (1) では、頂点 p の位置 x_p^j が該サンプリングにおいてビデオカメラにとって可視となるものであり、且つ x_p^j がネットワークをトレーニングする際に定義したサンプリング空間内にあること、

前記条件 (2) では、

$$(\omega_o' \cdot \hat{n}_p^j) > \theta$$

であり、

($\cdot$)

がドット積操作であり、 θ が有効サンプリング角度の下限であり、 ω_o' がワールド座標系における出射光の方向を示し、

$\hat{n}_p^j$

が j 回目の頂点 p の法線ベクトルを示すこと、

前記条件 (3) では、写真における画素の各チャネル数値が区間 $[a, b]$ にあり、 a と b が有効サンプリング輝度の下限と上限であり、

前記 3 つの条件が全て満足される場合、 j 回目のサンプリングが頂点 p に対して有効であると見なされ、 j 回目のサンプリング結果を高次元点群に追加することであることを特徴とする請求項 1 に記載の高次元マテリアルの自由式収集方法。

【請求項 7】

マテリアル情報を回復した後にマテリアルパラメータをフィッティングすることは、ローカル座標系及び粗度をフィッティングするステップ及び反射率をフィッティングするステップを含み、

前記ローカル座標系及び粗度をフィッティングするステップでは、ある有効テクスチャ座標に対して決定されたサンプリング物体の表面上の点 p について、ネットワークから出力される単一チャネルの鏡面反射ベクトルに基づいて $L-BFGS-B$ 方法によってマテリアルパラメータにおけるローカル座標系及び粗度をフィッティングし、

前記反射率をフィッティングするステップでは、信頼領域アルゴリズムによって鏡面反射率及び乱反射率を求め、求める際に前の過程において取得されたローカル座標系及び粗度を一定にし、収集に用いる視角において観測値を合成して、それを収集して取得された観測値にできる限り近接することであることを特徴とする請求項 1 に記載の高次元マテリアルの自由式収集方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は高次元マテリアルの自由式収集方法に関し、コンピュータグラフィックス及びコンピュータビジョンの分野に属する。

【背景技術】

【0002】

現実世界における物体のデジタル化はコンピュータグラフィックス及びビジョンの中核問題の1つである。現在、デジタル化された真の物体は三次元グリッドモデル及び空間に従って変化する六次元の双方向反射率分布関数（SVBRDF）で表現されてもよく、デジタル化された真の物体は任意の視角及び照明条件下でその元の様子を明瞭に再現することができ、文化遺産、電子商取引、コンピュータゲーム及び映画作成などの分野に重要な応用を持っている。

10

【0003】

高精度の幾何学的モデルは商用モバイル3Dスキャナにより容易に取得できるが、下記原因により外観走査を自由に行うように軽量型装置を開発することも望まれている。第1として、ビデオカメラの姿勢を確実に推定できれば、異なるサイズの物体を走査することができる。第2として、装置の可搬性によって、輸送を許容しないオブジェクト例えば貴重な文化財に対してその場走査を実行できるようにする。また、軽量レベルの装置は製造に必要な時間が短く、コストが低いため、より多くの人に受け入れられる。それは更に幾何学的走査に類似するユーザーフレンドリーな体験を提供する。

【0004】

20

需要が急増しているが、有効な非平面外観走査は依然として解決の待たれる問題の1つである。一方、従来のほとんどの可動外観走査作業はいずれも単一点／平行光の場合に撮影するため、四次元の観察及び照明方向におけるサンプリング効率がより低く、空間分解能で角度精度を取り換えるために先験知識を必要としている（Giljoo Nam, Joo Ho Lee, Diego Gutierrez, and Min H Kim. 2018. Practical SVBRDF acquisition of 3D objects with unstructured flash photography. In SIGGRAPH Asia Technical Papers. 267.）。他方、固定収集システムは照明が変化する際に一定の画像条件に依存し、現在ではどのように携帯機器に拡張するかはまだ把握されておらず、携帯機器が非構造化されて絶えず変化している画像を有し、且つその外形寸法がより小さいため、照明分野を完全にカバーできない。

30

【発明の概要】

【発明が解決しようとする課題】

【0005】

本発明の目的は、従来技術の欠点に対して高次元マテリアルの自由式収集方法を提供することにあり、本方法は各画像の収集条件情報を効果的に利用して、無秩序で不均一に分布する収集結果から高品質の物体マテリアル属性を回復することができる。

【課題を解決するための手段】

【0006】

40

本発明は高次元マテリアルの自由式収集方法を開示し、本方法の主な要旨は、自由型の外観走査を非構造化点群における幾何学的学習の問題に変換することができ、点群における各点が1つの画像測定値及び画像撮影時の物体のポーズ（Pose）情報を示してもよい、ということである。この要旨に基づいて、本発明は1つのニューラルネットワークを設計し、異なる非構造化画像において情報を効果的に集約して、空間的に独立した反射属性を再構築することができ、且つ収集段階に使用される照明パターンを最適化して、最終的に高品質のマテリアル収集結果を取得することができる。本発明はある特定の収集装置に依存せず、固定物体を用いて人の手で装置を持って収集し、又は物体を回転テーブルに置いて回転させて固定装置により収集してもよいが、この2つの方式に限らない。

【0007】

50

本方法はマテリアル情報の学習を非構造化点群における幾何学的学習の問題に変換し、複数の異なる照明及び観察方向におけるサンプリング結果を高次元点群に構成し、点群における各点が画像測定値と画像撮影時の物体のポーズ情報とからなるベクトルであることを提案し、本方法は無秩序で、不規則で、不均一に分布し且つ精度が制限される高次元点群から非構造化画像の情報を効果的に集約して、高品質のマテリアル属性を回復することができる。形骸化された表示は、 $F(G(\text{high dimensional point cloud})) = m$ であり、

点群データ特徴抽出方法Gがある特定のネットワーク構造に限らず、点群から特徴を抽出できる他の方法も適用され、非線形マッピングネットワークFが完全接続ネットワークに限らず、物体マテリアル属性の表現がLumitexelベクトルmに限らない。

【0008】

本方法はトレーニング段階及び収集段階の2つの段階を含む。

【0009】

前記トレーニング段階は、

収集装置のパラメータを取得して、実際のビデオカメラを模擬する収集結果をトレーニングデータとして生成するステップ(1)と、

生成されたトレーニングデータを用いてニューラルネットワークをトレーニングするステップ(2)と、を含み、ニューラルネットワークの特徴は、

ニューラルネットワークの入力がk個の非構造化サンプリングにおけるLumitexelベクトルであり、kがサンプリング個数であり、Lumitexelの各値が各光源からの入射光に対するサンプリング点のある観察方向に沿う反射強度を説明し、Lumitexelが光源の発光強度と線形関係を有し、線形完全接続層で模擬する(2.1)と、

ニューラルネットワークの第1層が線形完全接続層を含み、実際に収集する際に用いる照明パターンを模擬して、前記k個のLumitexelをカメラ収集結果に変換することに用いられ、このk個の収集結果がそれぞれ対応するサンプリング点のポーズ情報と組み合わせ高次元点群を構成する(2.2)と、

第2層からは特徴抽出ネットワークであり、前記高次元点群における各点から独立して特徴抽出して特徴ベクトルを取得する(2.3)と、

特徴抽出ネットワークの後では、k個の非構造化画像から抽出された特徴ベクトルを集約して、グローバル特徴ベクトルを取得するための最大プーリング層である(2.4)と、

最大プーリング層の後では、前記グローバル特徴ベクトルに基づいて高次元マテリアル情報を回復するための非線形マッピングネットワークである(2.5)と、を含み、

前記収集段階は、

マテリアル収集ステップであって、収集装置が前記照明パターンに応じて順に目標三次元物体を照射し、ビデオカメラが1組の非構造化画像における写真を取得し、写真を入力とし、サンプリング物体にテクスチャ座標を持つ幾何学的モデル及び写真撮影時のビデオカメラのポーズを取得するステップ(1)と、

マテリアル回復ステップであって、マテリアル収集段階における写真撮影時のビデオカメラのポーズに基づいて各枚の写真を撮影するときのサンプリング物体上の各有効テクスチャ座標の頂点に対するポーズを取得し、前記収集された写真及びポーズ情報に基づいて前記高次元点群をニューラルネットワークの第2層特徴抽出ネットワークの入力として構成し、高次元マテリアル情報を計算して取得するステップ(2)と、を含む。

【0010】

更に、前記非構造化サンプリングは非固定視角の自由ランダムサンプリングであり、サンプリングデータが無秩序で、不規則で、不均一に分布しており、固定物体を用いて人の手で収集装置を持って収集し、又は物体を回転テーブルに置いて回転させて固定装置により収集してもよい。

【0011】

更に、トレーニングデータ生成過程において、光源がカラーである場合、光源、サンプリング物体及びビデオカメラのスペクトル応答関係を補正する必要があり、補正方法につ

10

20

30

40

50

いては、

未知のカラー光源 L のスペクトル分布曲線を

$$S_{c_1}^L(\lambda)$$

として定義し、 λ が波長を示し、 c_1 が RGB の 3 つのチャンネルのうちの 1 つを示し、光強度 $\{I_R, I_G, I_B\}$ の光源のスペクトル分布曲線 $L(\lambda)$ は、以下の式 2 に示される可能であり、

【数 2】

$$L(\lambda) = I_R S_R^L(\lambda) + I_G S_G^L(\lambda) + I_B S_B^L(\lambda)$$

10

任意のサンプリング点 p の反射スペクトル分布曲線 $p(\lambda)$ を、係数がそれぞれ p_R 、 p_G 、 p_B の 3 つの未知の基

$$S_{c_2}^p(\lambda)$$

の線形組合せに示し、 c_2 が RGB の 3 つのチャンネルのうちの 1 つを示し、具体的には、以下の式 3 に示され、

【数 3】

$$p(\lambda) = p_R S_R^p(\lambda) + p_G S_G^p(\lambda) + p_B S_B^p(\lambda)$$

20

ビデオカメラ C のスペクトル分布曲線が

$$S_{c_3}^C(\lambda)$$

の線形組合せとして示され、光強度 $\{I_R, I_G, I_B\}$ の光源の照射下で、ビデオカメラによる反射係数 $\{p_R, p_G, p_B\}$ のサンプリング点のある特定チャンネル c_3 での測定値は、以下の式 4 に示され、

【数 4】

$$\int L(\lambda) p(\lambda) S_{c_3}^C(\lambda) d\lambda = \sum_{c_1, c_2} I_{c_1} p_{c_2} \delta(c_1, c_2, c_3)$$

30

$$\delta(c_1, c_2, c_3) = \int S_{c_1}^L(\lambda) S_{c_2}^p(\lambda) S_{c_3}^C(\lambda) d\lambda$$

40

照明条件 $\{I_R, I_G, I_B\} = \{1, 0, 0\} / \{0, 1, 0\} / \{0, 0, 1\}$ の下で、既知の反射係数 $\{p_R, p_G, p_B\}$ のカラーテストカードを撮影し、ビデオカメラにより収集された測定値に基づいて線形連立方程式を構築し、サイズ $3 \times 3 \times 3$ のカラー補正行列 $\delta(c_1, c_2, c_3)$ を求め、該カラー補正行列が光源、サンプリング物体及びビデオカメラのスペクトル応答関係を示す。

【0 0 1 2】

更に、トレーニング段階ステップ (2. 1) において、物体表面のサンプリング点 p の写真における観測値 B 、反射関数 f_r 及び各光源の光強度の関係は、以下の式 5 に示され、

【数 5】

50

$$B(I, P) = \sum_l I(l) \int \frac{1}{||x_l - x_p||^2} \Psi(x_l, -\omega'_i) V(x_l, x_p) f_r(\omega'_i; \omega'_o, P) (\omega'_i \cdot n_p)^+ (-\omega'_i \cdot n_l)^+ dx_l$$

ここで、 I は各光源の発光情報を示し、光源 l の空間位置 x_l 、光源 l の法線ベクトル n_l 、光源 l の発光強度 $I(l)$ を含み、 P はサンプリング点 p のパラメータ情報を含み、サンプリング点の空間位置 x_p 、マテリアルパラメータ n 、 t 、 α_x 、 α_y 、 ρ_d 、 ρ_s を含む。

$$\Psi(x_l, \cdot)$$

10

が光源 l の異なる入射方向における光強度分布を説明し、 V が x_l の x_p に対する可視性の二値関数を示し、

$$(\cdot)^+$$

が2つのベクトルのドット積操作である $f_r(\omega'_i; \omega'_o, P)$ は ω'_o が一定である場合に ω'_i に関する二次元反射関数である。

【0013】

20

ニューラルネットワークの入力が k 個の非構造化サンプリングにおける $Lumitex$ ベクトルであって、 $m(l; P)$ と記され、

【数6】

$$m(l; P) = B(\{I(l) = 1, \forall_{i \neq l} I(i) = 0\}, P)$$

上記式における B が単一チャネルにおける表示であり、光源がカラー光源である場合、 B を下記形式に拡張し、

【数7】

$$B(I, P; c_3) = \sum_{l, c_1, c_2} I(l; c_1) \int \frac{1}{||x_l - x_p||^2} \Psi(x_l, -\omega'_i) V(x_l, x_p) f_r(\omega'_i; \omega'_o, P, c_2) (\omega'_i \cdot n_p)^+ (-\omega'_i \cdot n_l)^+ \delta(c_1, c_2, c_3) dx_l$$

30

ここで、 $f_r(\omega'_i; \omega'_o, P, c_2)$ が $f_r(\omega'_i; \omega'_o, P)$ における $\rho_d = \rho_{d_{c_2}}$ 、 $\rho_s = \rho_{s_{c_2}}$

の結果である。

40

【0014】

更に、トレーニング段階ステップ(2.3)において、特徴抽出ネットワークの公式は、以下の式9に示され、

【数9】

$$V_{feature}(j) = f(\text{concat}[B(I, P^j), x_p^j, \hat{n}_p^j, \hat{t}_p^j]), 1 \leq j \leq k$$

ここで、 f が一次元畳み込み関数であり、畳み込みカーネルのサイズが 1×1 であり、 $B(I, P^j)$ が第1層のネットワークから出力される結果又は収集して取得された測定値を示し、

50

$$x_p^j, \hat{n}_p^j, \hat{t}_p^j$$

がそれぞれ j 回目にサンプリングするときのサンプリング点の空間位置、サンプリング点の幾何学的法線ベクトル及び幾何学的接ベクトルであり、

$$\hat{n}_p$$

が幾何学的モデルにより取得され、

$$\hat{t}_p$$

が

$$\hat{n}_p$$

に直交する任意の単位ベクトルであり、ビデオカメラのポーズを j 回目にサンプリングすることにより

$$\hat{n}_p, \hat{t}_p$$

を変換して

$$\hat{n}_p^j, \hat{t}_p^j$$

を取得することができ、 $V_{feature}(j)$ がネットワークから出力される j 回目にサンプリングする特徴ベクトルである。

【0015】

更に、トレーニング段階のステップ（2.5）において、非線形マッピングネットワークの形骸化された表現は、以下の式11に示され、

【数11】

$$y_{i+1}^d = f_{i+1}^d(y_i^d W_{i+1}^d + b_{i+1}^d), i \geq 1$$

$$y_{i+1}^s = f_{i+1}^s(y_i^s W_{i+1}^s + b_{i+1}^s), i \geq 1$$

ここで、 f_{i+1} が第 $i+1$ 層のネットワークのマッピング関数であり、 W_{i+1} が第 $i+1$ 層のネットワークのパラメータ行列であり、 b_{i+1} が第 $i+1$ 層のネットワークのオフセットベクトルであり、 y_{i+1} が第 $i+1$ 層のネットワークの出力であり、 d と s がそれぞれ乱反射及び鏡面反射の2つの分岐を示し、入力 y_1^d と y_1^s が最大プーリング層から出力されるグローバル特徴ベクトルである。

【0016】

更に、前記ニューラルネットワークの損失関数は以下のように設計され、

（1）1つの $Lumitextel$ 空間を仮定し、該空間は中心がサンプリング点の空間位置 x_p にある立方体であり、立方体の中心座標系は x 軸方向が

$$\hat{t}_p$$

であり、 z 軸方向が

$\hat{n}_p$

であり、

 $\hat{n}_p$

が幾何学的法線ベクトルであり、

 $\hat{t}_p$

10

が

 $\hat{n}_p$

に直交する任意の単位ベクトルであり、

(2) 1つのビデオカメラを仮定し、観察方向が立方体の z 軸の正方向であり、

(3) 乱反射 $Lumitexel$ の場合に立方体の分解能が $6 \times N_d^2$ であり、鏡面反射 $Lumitexel$ の場合に立方体の分解能が $6 \times N_s^2$ であり、即ち各面から N_d^2 、 N_s^2 個の点を、光強度が単位光強度の仮想点光源として均一にサンプリングし、

20

a. サンプリング点の鏡面反射率 p_s を 0 とし、この $Lumitexel$ 空間における乱反射特徴ベクトル

 $\tilde{m}_d$

を生成し、

b. 乱反射率 p_d を 0 とし、この $Lumitexel$ 空間における鏡面反射特徴ベクトル

 $\tilde{m}_s$

30

を生成し、

c. ニューラルネットワークの出力をベクトル m_d 、 m_s とし、 m_d が

 $\tilde{m}_d$

の長さと同じであり、 m_s が

 $\tilde{m}_s$

の長さと同じであり、ベクトル m_d 、 m_s が乱反射特徴ベクトル

40

 $\tilde{m}_d$

、鏡面反射ベクトル

 $\tilde{m}_s$

の予測であり、

(4) マテリアル特徴部分の損失関数は、以下の式 12 に示され、

【数 12】

50

$$L = \lambda_d \sum_l [m_d(l) - \tilde{m}_d(l)]^2 + \lambda_s \beta \sum_l [m_s(l) - \log(1 + \tilde{m}_s(l))]^2$$

ここで、 λ_d と λ_s がそれぞれ m_d 、 m_s の損失重みを示し、信頼度 β が鏡面反射 Lum i t e x e l の損失を評価することに用いられ、 $\log$ がベクトルの各次元に作用し、信頼度 β の決定は、以下の式 13 に示され、

【数 13】

$$\beta = \min\left(\frac{1}{\epsilon} \max_j \left[\frac{\max_l \log(1 + f_r(\omega_i^{j'}(l); \omega_o^{j'}, P))}{\max_{\omega_i'} \log(1 + f_r(\omega_i'; \omega_o^{j'}, P))} \right], 1\right), 1 \leq j \leq k$$

10

ここで、

$$\max_l \log(1 + f_r(\omega_i^{j'}(l); \omega_o^{j'}, P))$$

項が j 回目にサンプリングする全ての単一光源レンダリング値の最大値の対数を示し、

$$\max_{\omega_i'} \log(1 + f_r(\omega_i'; \omega_o^{j'}, P))$$

20

項が j 回目にサンプリングする際に理論的に取得できる単一光源レンダリング値の最大値の対数を示し、

ϵ

が比率調整係数である。

【0017】

更に、前記収集段階において、マテリアル収集が終了した後に幾何学的位置合わせを行い、その後にマテリアル回復を行い、幾何学的位置合わせは具体的に、スキャナにより物体を走査して幾何学的モデルを取得し、それを三次元で再構築された幾何学的モデルに位置合わせした後に三次元で再構築された幾何学的モデルを代替することである。

30

【0018】

更に、有効テクスチャ座標に対して、前記収集された写真及びサンプリング点のポーズ情報に基づいて写真における画素を順に取り出して、画素の有効性を判断し、頂点に対するポーズと組み合わせて高次元点群を構成し、ある有効テクスチャ座標に対して決定されたサンプリング物体の表面上の点 p について、 j 回目のサンプリングが頂点 p に対して有効である判断標準は、

(1) 頂点 p の位置 x_p^j が該サンプリングにおいてビデオカメラにとって可視となるものであり、且つ x_p^j がネットワークをトレーニングする際に定義したサンプリング空間内にあること、

40

(2)

$$(\omega_o' \cdot \hat{n}_p^j) > \theta$$

であり、

($\cdot$)

50

がドット積操作であり、 θ が有効サンプリング角度の下限であり、 ω_o' がワールド座標系における出射光の方向を示し、

$$\hat{n}_p^j$$

が j 回目の頂点 p の法線ベクトルを示すこと、

(3) 写真における画素の各チャネル数値が区間 $[a, b]$ にあり、 a と b が有効サンプリング輝度の下限と上限であること、に表現され、

3つの条件が全て満足される場合、 j 回目のサンプリングが頂点 p に対して有効であると見なされ、 j 回目のサンプリング結果を高次元点群に追加する。

【0019】

更に、マテリアル情報を回復した後にマテリアルパラメータをフィッティングしてもよく、それは、

ローカル座標系及び粗度をフィッティングするステップであって、ある有効テクスチャ座標に対して決定されたサンプリング物体の表面上の点 p について、ネットワークから出力される単一チャネルの鏡面反射ベクトルに基づいて $L-BFGS-B$ 方法によってマテリアルパラメータにおけるローカル座標系及び粗度をフィッティングするステップ(1)、

反射率をフィッティングするステップであって、信頼領域アルゴリズムによって鏡面反射率及び乱反射率を求め、求める際に前の過程において取得されたローカル座標系及び粗度を一定にし、収集に用いる視角において観測値を合成して、それを収集して取得された観測値にできる限り近接するステップ(2)、の2つのステップに分けられる。

【発明の効果】

【0020】

本発明の有益な効果は以下のとおりである。本発明の方法はマテリアル情報の学習を非構造化点群における幾何学的学習の問題に変換し、複数の異なる照明及び観察方向におけるサンプリング結果を高次元点群に構成し、点群における各点が画像測定値と画像撮影時の物体のポーズ情報とからなるベクトルであることを提案し、本方法は無秩序で、不規則で、不均一に分布し且つ精度が制限される高次元点群から非構造化画像の情報を効果的に集約して、高品質のマテリアル属性を回復することができる。

【図面の簡単な説明】

【0021】

【図1】図1は本発明の実施形態における収集装置の三次元模式図である。

【図2】図2は本発明の実施形態における収集装置の正面図である。

【図3】図3は本発明の実施形態における収集装置の側面図である。

【図4】図4は本発明の実施形態における収集装置とサンプリング空間との関係の模式図である。

【図5】図5は本発明の実施形態における収集方法のフローチャートである。

【図6】図6は本発明の実施形態におけるニューラルネットワークの構造模式図である。

【図7】図7は本発明の実施形態において取得された照明パターンの単一チャネル表示であり、階調値で発光強度を代表する。

【図8】図8は本発明の実施形態に係るシステムにより回復された $Lumitexel$ ベクトル結果を示す図である。

【図9】図9は本発明の実施形態に係るシステムにより回復されたサンプリング物体のマテリアル属性結果を示す図である。

【発明を実施するための形態】

【0022】

本発明の目的、技術案及び利点をより明確にするために、以下に図面を参照しながら本発明を詳しく説明する。

【0023】

本発明に係る高次元マテリアルの自由式収集方法について、具体的な実施は下記のトレ

ーニング段階及び収集段階に分けられてもよい。

【0024】

一、トレーニング段階

1. トレーニングデータを生成して、収集装置のパラメータを取得し、該パラメータは光源からサンプリング空間原点までの距離及び角度、光源の特性曲線、ビデオカメラからサンプリング空間原点までの距離及び角度、ビデオカメラの内部パラメータ及び外部パラメータを含む。これらのパラメータを利用して実際のビデオカメラを模擬する収集結果をトレーニングデータとして生成する。トレーニングデータを生成する際に用いるレンダリングモデルがGGXモデルであり、生成公式は、以下の式1に示され、

【数1】

$$f_r(\omega'_i, \omega'_o; P) = \frac{\rho_d}{\pi} + \rho_s \frac{D_{GGX}(\omega_h; \alpha_x, \alpha_y) F(\omega_i, \omega_h) G_{GGX}(\omega_i, \omega_o; \alpha_x, \alpha_y)}{4(\omega'_i \cdot n)(\omega'_o \cdot n)}$$

ここで、 $f_r(\omega'_i, \omega'_o; P)$ が ω'_i, ω'_o に関する四次元反射関数であり、 ω'_i がワールド座標系における入射光の方向を示し、 ω'_o がワールド座標系における出射光の方向を示し、 ω_i がローカル座標系における入射方向であり、 ω_o がローカル座標系における出射方向であり、 ω_h がローカル座標系における途中ベクトルである。P はサンプリング点のパラメータ情報を含み、サンプリング点のマテリアルパラメータ $n, t, \alpha_x, \alpha_y, \rho_d, \rho_s$ を含み、ここで、 n がワールド座標系における法線ベクトルを示し、 t がワールド座標系においてサンプリング点のローカル座標系における x 軸方向を示し、 n と t が入射方向及び出射方向をワールド座標系からローカル座標系に変換するためのものである。 α_x と α_y が粗度係数を示し、 ρ_d が乱反射率を示し、 ρ_s が鏡面反射率を示し、 ρ_d と ρ_s が単一チャンネルにおいていずれも1つのスカラーであり、カラーの場合にそれぞれ3つのスカラー

$$(\rho_{dR}, \rho_{dG}, \rho_{dB})$$

及び

$$(\rho_{sR}, \rho_{sG}, \rho_{sB})$$

であり、 D_{GGX} が微小表面分布項であり、 F がフレネル項であり、 G_{GGX} が陰影係数関数を示す。

【0025】

収集に用いる光源がカラーである場合、まず、光源、サンプリング物体及びビデオカメラのスペクトル応答関係を補正する必要がある。補正方法については、未知のカラー光源 L のスペクトル分布曲線を

$$S_{c_1}^L(\lambda)$$

として定義し、 λ が波長を示し、 c_1 がRGBの3つのチャンネルのうちの1つを示し、光強度 $\{I_R, I_G, I_B\}$ の光源のスペクトル分布曲線 $L(\lambda)$ は、以下の式2に示される可能であり、

【数2】

$$L(\lambda) = I_R S_R^L(\lambda) + I_G S_G^L(\lambda) + I_B S_B^L(\lambda)$$

同様に、任意のサンプリング点 p の反射スペクトル分布曲線 $p(\lambda)$ を、係数がそれぞれ p_R, p_G, p_B の3つの未知の基

$$S_{c_2}^p(\lambda)$$

の線形組合せに示してもよく、 c_2 がRGBの3つのチャンネルのうちの1つを示し、具体的には、以下の式3に示され、

【数3】

$$p(\lambda) = p_R S_R^p(\lambda) + p_G S_G^p(\lambda) + p_B S_B^p(\lambda)$$

類似に、ビデオカメラCのスペクトル分布曲線が

$$S_{c_3}^C(\lambda)$$

の線形組合せとして示されてもよい。従って、光強度 $\{I_R, I_G, I_B\}$ の光源の照射下で、ビデオカメラによる反射係数 $\{p_R, p_G, p_B\}$ のサンプリング点のある特定チャンネル c_3 での測定値は、以下の式4に示され、

【数4】

$$\int L(\lambda) p(\lambda) S_{c_3}^C(\lambda) d\lambda = \sum_{c_1, c_2} I_{c_1} p_{c_2} \delta(c_1, c_2, c_3)$$

$$\delta(c_1, c_2, c_3) = \int S_{c_1}^L(\lambda) S_{c_2}^p(\lambda) S_{c_3}^C(\lambda) d\lambda$$

照明条件 $\{I_R, I_G, I_B\} = \{1, 0, 0\} / \{0, 1, 0\} / \{0, 0, 1\}$ の下で、既知の反射係数 $\{p_R, p_G, p_B\}$ のカラーテストカードを撮影し、ビデオカメラにより収集された測定値に基づいて線形連立方程式を構築し、サイズ $3 \times 3 \times 3$ のカラー補正行列 $\delta(c_1, c_2, c_3)$ を求めることができ、該カラー補正行列が光源、サンプリング物体及びビデオカメラのスペクトル応答関係を示す。

【0026】

2. 生成されたトレーニングデータを用いて図4に示されるニューラルネットワークをトレーニングする。ニューラルネットワークの特徴は下記(1)～(6)を含む。

【0027】

(1) 物体表面のサンプリング点 p の写真における観測値 B 、反射関数 f_r 及び各光源の光強度の関係は、以下の式5に示され、

【数5】

$$B(I, P) = \sum_l I(l) \int \frac{1}{\|x_l - x_p\|^2} \Psi(x_l, -\omega'_l) V(x_l, x_p) f_r(\omega'_l; \omega'_o, P) (\omega'_l \cdot n_p)^+ (-\omega'_l \cdot n_l)^+ dx_l$$

ここで、 I は各光源 l の発光情報を示し、光源 l の空間位置 x_l 、光源 l の法線ベクトル n_l 、光源 l の発光強度 $I(l)$ を含み、 P はサンプリング点 p のパラメータ情報を含み、サンプリング点の空間位置 x_p 、マテリアルパラメータ n 、 t 、 α_x 、 α_y 、 ρ_d 、 ρ_s を含む。

$$\Psi(x_l, \cdot)$$

10

20

30

40

50

が光源 l の異なる入射方向における光強度分布を説明し、 V が x_l の x_p に対する可視性の二値関数を示し、

$$(\cdot)^+$$

が2つのベクトルのドット積操作であり、負値が0にカットオフされることとなる。 $f_r(\omega_i'; \omega_o', P)$ は ω_o' が一定である場合に ω_i' に関する二次元反射関数である。

【0028】

ニューラルネットワークの入力が無秩序で不規則なサンプリングにおける k 個の $Lumitexel$ であり、 k がサンプリング個数であり、 $Lumitexel$ が1つのベクトルであって $m(l; P)$ と記され、その中の各値が各光源からの入射光に対するサンプリング点のある観察方向に沿う反射強度を説明し、

【数6】

$$m(l; P) = B(\{I(l) = 1, \forall_{i \neq l} I(i) = 0\}, P)$$

上記式における B が単一チャンネルにおける表示であり、光源がカラー光源である場合、 B を下記形式に拡張し、

【数7】

$$B(I, P; c_3) = \sum_{l, c_1, c_2} I(l; c_1) \int \frac{1}{\|x_l - x_p\|^2} \Psi(x_l, -\omega_i') V(x_l, x_p) f_r(\omega_i'; \omega_o', P, c_2) (\omega_i' \cdot n_p)^+ (-\omega_i' \cdot n_l)^+ \delta(c_1, c_2, c_3) dx_l$$

ここで、 $f_r(\omega_i'; \omega_o', P, c_2)$ が $f_r(\omega_i'; \omega_o', P)$ の公式における

$$\rho_d = \rho_{d_{c_2}}, \quad \rho_s = \rho_{s_{c_2}}$$

の結果である。 B が光源の発光強度と線形関係を有し、線形完全接続層で模擬してもよい。

【0029】

(2) ニューラルネットワークの第1層が線形完全接続層を含み、線形完全接続層のパラメータ行列は下記公式によりトレーニングして取得され、

【数8】

$$W_l = f_W(W_{raw})$$

ここで、 W_{raw} はトレーニング対象のパラメータであり、 W_l は照明行列であり、単一チャンネル光源の場合にサイズが $1 \times N$ であり、カラー光源の場合にサイズが $3 \times N$ であり、 N が $Lumitexel$ のベクトル長さであり、 f_W は1つのマッピングであり、 W_{raw} を変換することにより、生成された照明行列を光源の可能な発光強度に対応できるようにし、本実例ではマッピング f_W が $Sigmoid$ 関数を選択し、第1層のネットワークの照明行列 W_l の初期値を $(0, 1)$ に制限するが、 f_W が $Sigmoid$ 関数に限らない。

【0030】

W_l を光源の発光強度とし、上記(1)の前記関係に応じて k 個のサンプリング観測値 $B(I, P^1), B(I, P^2) \dots B(I, P^k)$

を計算して取得する。

【 0 0 3 1】

(3) 第2層からは特徴抽出ネットワークであり、k個のサンプリングは独立して特徴抽出して特徴ベクトルを取得し、公式は、以下の式9に示され、

【数9】

$$V_{feature}(j) = f(\text{concat}[B(I, P^j), x_p^j, \hat{n}_p^j, \hat{t}_p^j]), 1 \leq j \leq k$$

ここで、fが一次元畳み込み関数であり、畳み込みカーネルのサイズが 1×1 であり、
 x_p^j 、 $\hat{n}_p^j$ 、 $\hat{t}_p^j$

10

がそれぞれj回目にサンプリングするときのサンプリング点の空間位置、サンプリング点の幾何学的法線ベクトル及び幾何学的接ベクトルであり、

$\hat{n}_p$

が三次元で再構築され又はスキャナにより走査された後に取得した幾何学的モデルにより取得されてもよく、

$\hat{t}_p$

20

が

$\hat{n}_p$

に直交する任意の単位ベクトルであり、ビデオカメラのポーズをj回目にサンプリングすることにより

$\hat{n}_p$ 、 $\hat{t}_p$

30

を変換して

$\hat{n}_p^j$ 、 $\hat{t}_p^j$

を取得することができ、 $V_{feature}(j)$ がネットワークから出力されるj回目にサンプリングする特徴ベクトルである。

【 0 0 3 2】

(4) 特徴抽出ネットワークの後は最大プーリング層である。最大プーリング操作の公式は、以下の式10に示され、

【数10】

$$V_{feature} = \max(V_{feature}(1), V_{feature}(2), \dots, V_{feature}(k))$$

最大プーリング操作が $V_{feature}(1)$ 、 $V_{feature}(2)$ 、 $\dots$ 、 $V_{feature}(k)$ の各次元において行われる。

【 0 0 3 3】

(5) 最大プーリング層の後は非線形マッピングネットワークであり、

【数11】

50

$$y_{i+1}^d = f_{i+1}^d(y_i^d W_{i+1}^d + b_{i+1}^d), i \geq 1$$

$$y_{i+1}^s = f_{i+1}^s(y_i^s W_{i+1}^s + b_{i+1}^s), i \geq 1$$

ここで、 f_{i+1} が第 $i+1$ 層のネットワークのマッピング関数であり、 W_{i+1} が第 $i+1$ 層のネットワークのパラメータ行列であり、 b_{i+1} が第 $i+1$ 層のネットワークのオフセットベクトルであり、 y_{i+1} が第 $i+1$ 層のネットワークの出力であり、 d と s がそれぞれ乱反射及び鏡面反射の2つの分岐を示し、入力 y_1^d と y_1^s が $V_{feature}$ である。

10

【0034】

(6) ニューラルネットワークの損失関数は下記(6.1)～(6.4)を含む。

(6.1) 1つの $Lumitexel$ 空間を仮定し、該空間は中心がサンプリング点の空間位置 x_p にある立方体であり、立方体の中心座標系は x 軸方向が

$$\hat{t}_p$$

であり、 z 軸方向が

$$\hat{n}_p$$

20

であり、

(6.2) 1つのビデオカメラを仮定し、観察方向が立方体の z 軸の正方向であり、

(6.3) 乱反射 $Lumitexel$ の場合に立方体の分解能が $6 \times N_d^2$ であり、鏡面反射 $Lumitexel$ の場合に立方体の分解能が $6 \times N_s^2$ であり、即ち各面から N_d^2 、 N_s^2 個の点を、光強度が単位光強度の仮想点光源として均一にサンプリングし、本実施例では $N_d=8$ 、 $N_s=32$ であり、

a. サンプリング点の鏡面反射率 p_s を0とし、この $Lumitexel$ 空間における乱反射特徴ベクトル

$$\tilde{m}_d$$

30

を生成し、

b. 乱反射率 p_d を0とし、この $Lumitexel$ 空間における鏡面反射特徴ベクトル

$$\tilde{m}_s$$

を生成し、

c. ニューラルネットワークの出力をベクトル m_d 、 m_s とし、 m_d が

$$\tilde{m}_d$$

40

の長さと同じであり、 m_s が

$$\tilde{m}_s$$

の長さと同じであり、ベクトル m_d 、 m_s がそれぞれ乱反射特徴ベクトル

$$\tilde{m}_d$$

50

、鏡面反射ベクトル

$\tilde{m}_s$

の予測であり、

(6.4) マテリアル特徴部分の損失関数は、以下の式12に示され、

【数12】

$$L = \lambda_d \sum_l [m_d(l) - \tilde{m}_d(l)]^2 + \lambda_s \beta \sum_l [m_s(l) - \log(1 + \tilde{m}_s(l))]^2$$

10

ここで、 λ_d と λ_s がそれぞれ m_d 、 m_s の損失重みを示し、信頼度 β が鏡面反射Lum i t e x e lの損失を評価することに用いられ、logがベクトルの各次元に作用し、信頼度 β の決定は、以下の式13に示され、

【数13】

$$\beta = \min\left(\frac{1}{\epsilon} \max_j \left[\frac{\max_l \log(1 + f_r(\omega_i^{j'}(l); \omega_o^{j'}, P))}{\max_{\omega_i'} \log(1 + f_r(\omega_i'; \omega_o^{j'}, P))} \right], 1\right), 1 \leq j \leq k$$

ここで、

$$\max_l \log(1 + f_r(\omega_i^{j'}(l); \omega_o^{j'}, P))$$

20

項がj回目にサンプリングする全ての単一光源レンダリング値の最大値の対数を示し、

$$\max_{\omega_i'} \log(1 + f_r(\omega_i'; \omega_o^{j'}, P))$$

項がj回目にサンプリングする際に理論的に取得できる単一光源レンダリング値の最大値の対数を示し、

ϵ

30

が比率調整係数であり、本実施例では

ϵ

=50%である。

【0035】

40

3. トレーニングが終了した後、ネットワークの線形完全接続層のパラメータ W_{raw} を取り出して、公式 $W_l = f_W(W_{raw})$ により変換されたものを照明パターンとする。

【0036】

二、収集段階

収集段階は更にマテリアル収集段階、幾何学的位置合わせ段階（選択可能に）及びマテリアル回復段階に細分化されてもよい。

【0037】

1. マテリアル収集段階

収集装置は照明パターンに応じて順に目標三次元物体を照射し、ビデオカメラは1組の非構造化画像における写真を取得し、写真を入力とし、工業界に開示される三次元再構築

50

ツールによってサンプリング物体の幾何学的モデル及び写真撮影時のビデオカメラのポーズを取得することができる。

2. 幾何学的位置合わせ段階（選択可能に）

(1) 高精度スキャナにより物体を走査して幾何学的モデルを取得し、

(2) スキャナにより走査して取得された幾何学的モデルと三次元で再構築された幾何学的モデルとを位置合わせしたもので三次元で再構築された幾何学的モデルを代替し、位置合わせ方法は分野内に開示される方法CPDを用いてもよい（A. Myronenko and X. Song, 2010, Point Set Registration: Coherent Point Drift, IEEE PAMI 32, 12 (2010), 2262-2275.

<https://doi.org/10.1109/TPAMI.2010.46>

10

）。

3. マテリアル回復段階

(1) マテリアル収集ステップにおける写真撮影時のビデオカメラのポーズに基づいて第 j 枚の写真を撮影するときのサンプリング物体の各頂点のポーズ

$$x_p^j, \hat{n}_p^j, \hat{t}_p^j$$

20

を取得し、

(2) 三次元で再構築して取得されたサンプリング物体の幾何学的モデル又は位置合わせされた後のスキャナにより走査して取得されたサンプリング物体の幾何学的モデルを分野内に開示されるツール IsoCharts により処理して、テクスチャ座標付きの幾何学的モデルを取得し、

(3) 有効テクスチャ座標に対して、上記収集された1組の写真 $r_1, r_2, \dots, r_\pi$ 及びサンプリング点のポーズ情報に基づいて写真における画素を順に取り出して、画素の有効性を判断し、頂点に対するポーズ

$$x_p, \hat{n}_p, \hat{t}_p$$

30

と組み合わせて高次元点群をニューラルネットワークの第2層の特徴抽出ネットワークの入力ベクトルとして構成し、最終層の出力ベクトル m_d 及び m_s を計算して取得する。

【0038】

ある有効テクスチャ座標に対して決定されたサンプリング物体の表面上の点 p について、 j 回目のサンプリングが点 p に対して有効である判断標準は、

1) x_p^j が該サンプリングにおいてビデオカメラにとって可視となるものであり、且つ x_p^j がネットワークをトレーニングする際に定義したサンプリング空間内にあること、
2)

$$(\omega'_0 \cdot \hat{n}_p^j) > \theta$$

40

であり、

($\cdot$)

がドット積操作であり、 θ が有効サンプリング角度の下限であり、本実施例では $\theta = 0.3$ であること、

3) 写真における画素の各チャネル数値が区間 $[a, b]$ にあり、 a と b が有効サンプリング輝度の下限と上限であり、本実施例では $a = 32$ 、 $b = 224$ であること、に表現

50

され、

3つの条件が全て満足される場合、j回目のサンプリングが点pに対して有効であると見なされ、j回目のサンプリング結果を高次元点群に追加し、

(4) マテリアルパラメータをフィッティングし、それは下記1)及び2)の2つのステップに分けられる。

1) ローカル座標系及び粗度のフィッティング

ある有効テクスチャ座標に対して決定されたサンプリング物体の表面上の点pについて、ネットワークから出力される単一チャネルの鏡面反射ベクトルに基づいてL-BFGS-B方法によってマテリアルパラメータにおけるローカル座標系及び粗度をフィッティングし、最適化目標は、以下の式14に示され、

【数14】

$$\text{minimize} \left(\sum_i \|m_{s(i)} - f_r(\omega_i'; \omega_o', P)\|_2 \right)$$

ここで、iは(6.3)における仮想光源の番号であり、 $m_s(i)$ はネットワークにより予測された鏡面反射特徴ベクトルのi番目の次元における値を示し、 ω_i' は該番号iの仮想光源とサンプリング点とがなす入射角を示し、 ω_o' はサンプリング点から(6.2)における仮想ビデオカメラまでなす出射角を示し、Pは該サンプリング点のレンダリングのための法線ベクトル n' 、接ベクトル t' 及び他のマテリアルパラメータ p' を含み、選択されたモデルの相違によって異なる。この工程がGGXモデルを用いる場合、 p' は異方性粗度、鏡面反射率、乱反射率を含む。上記最適化目標において n' 、 t' 、 p' が最適化可能なパラメータである。

【0039】

2) 反射率のフィッティング

該過程は信頼領域アルゴリズムによって鏡面反射率及び乱反射率を求め、フィッティング目標は、以下の式15に示され、

【数15】

$$\text{minimize} \left(\sum_j \|B_j - \tilde{B}_j\|_2 \right)$$

ここで、

$\tilde{B}_j$

は該画素がj番目のフィッティングのための視角において、上記最適化して取得された照明パターンにより照射された後、ビデオカメラの観測値を示す。 B_j は該画素がj番目のフィッティングのための視角において、前の過程において取得された n' 及び t' 並びに粗度を用いて合成した観測値を示す。カラー光源の場合に合成パラメータは補正して取得されたカラー補正行列 $\delta(c_1, c_2, c_3)$ を更に含む。 B_j の計算過程は以下のとおりである。

まず、乱反射率を1、鏡面反射率を0とし、前のステップにおいて取得されたレンダリングのための座標系及び粗度によってj番目の視角における乱反射 $Lumitexel$ 、

$\dot{m}_j^d$

をレンダリングする。次に、乱反射率を 0、鏡面反射率を 1 とし、前のステップにおいて取得されたレンダリングのための座標系及び粗度によって j 番目の視角における鏡面反射 $Lumitexel$ 、

$$\dot{m}_j^s$$

をレンダリングする。2つの $Lumitexel$ を接続して行列

$$M_j = \{\dot{m}_j^d, \dot{m}_j^s\}$$

10

を形成し、サイズは $N \times 2$ であり、ここで、 N がサンプリング装置の光源数である。サンプリングする際に用いるサイズ $3 \times N$ の照明パターン行列 W_l を M_j に乗じて $W_l M_j$ を取得し、サイズが 3×2 である。更にサイズ 2×3 の最適化可能な変数 $\rho_{d,s}$ に乗じて、以下の式 16 を取得し、

【数 16】

$$T = W_l M_j \rho_{d,s}$$

T のサイズが 3×3 であり、そのコピーを接続してテンソル

$$\hat{T} = \{T, T, T\}$$

20

を形成し、下位二次元で和を求めて、最終的に三次元ベクトル B_j を取得する。カラー光源の場合、

$$\hat{T}$$

を、装置を校正して取得したカラー補正行列 δ (c_1, c_2, c_3) に要素ごとに乗じから下位二次元で和を求める必要がある。

【0040】

30

以下、1つの具体的な収集装置のシステム実例を挙げて、図 1 はシステム実例の三次元表示であり、図 2 は正面図であり、図 3 は側面図であり、該収集装置は 1つのランプパネルで構成され、上部には画像を収集するための 1つのカメラが固定される。ランプパネルには計 512 個の LED ランプビーズが密集して配列される。ランプビーズは FPG A により制御され、発光輝度及び発光時間を調整することができる。

【0041】

以下、本発明の方法を応用する 1つの取得システム実例を挙げて、システム全体は下記いくつかのモジュールに分けられる。

【0042】

用意モジュールについては、

40

ネットワークトレーニングにデータセットを提供し、該部分が G G X モデルを用い、1組のマテリアルパラメータ及び k 個のサンプリング点のポーズ情報、ビデオカメラの位置を入力すると、 k 個の反射状況で構成される高次元点群を取得することができる。ネットワークトレーニング部は Py t o r c h オープンソースフレームワークを用い、且つ A d a m 最適化装置によってトレーニングを行う。ネットワーク構造は図 6 に示されるとおりであり、各矩形が 1 層のニューロンを示し、矩形中の数字が該層のニューロンの個数を示す。最左側の層が入力層であり、最右側の層が出力層である。層と層との間の実線矢印が完全接続を示し、破線矢印が畳み込みを示す。

【0043】

収集モジュールについては、

50

装置は図 1、2、3 に示されるとおりであり、以上は具体的な構造を説明したが、このシステムにおいて定義されたサンプリング空間のサイズ、並びに収集装置とサンプリング空間との空間的な位置関係は図 4 に示される。

【0044】

回復モジュールについては、

三次元で再構築して取得されたサンプリング物体の幾何学的モデル又は位置合わせされた後のスキャナにより走査して取得されたサンプリング物体の幾何学的モデルによってテクスチャ座標付きの幾何学的モデルを計算して取得し、トレーニング済みのニューラルネットワークをロードし、テクスチャ座標付きの幾何学的モデルの各頂点に対してマテリアル特徴ベクトルを予測し、レンダリングのための座標系及びマテリアルパラメータをフィッティングする。

【0045】

図 5 は本実施例の作業プロセスである。まず、トレーニングデータを生成し、ランダムにサンプリングして 2 億個の *Lumitexel* を取得し、その 80% をトレーニングセットとし、残りを検証セットとする。ネットワークをトレーニングする際に *Xavier* 方法によってパラメータを初期化し、学習率を $1e-4$ とする。照明パターンがカラーであり、照明行列のサイズが $(3, 512)$ であり、行列の 3 行がそれぞれ赤、緑、青の 3 つのチャンネルの照明パターンを示す。トレーニングが終了した後、照明行列を取り出して照明パターンに変換し、各列のパラメータが該位置での光源の発光強度を指定しており、図 7 にはネットワークをトレーニングして取得された 1 つの赤、緑、青の 3 つのチャンネルの照明パターンを示す。次のプロセスは以下のとおりである。第 1 として、装置を手で持って、ランプパネルを照明パターンに応じて発光させ、ビデオカメラで物体を同時に撮影して、1 組のサンプリング結果を取得する。第 2 として、三次元で再構築して取得されたサンプリング物体の幾何学的モデル又は位置合わせされた後のスキャナにより走査して取得されたサンプリング物体の幾何学的モデルを *Isochart* によって処理して、テクスチャ座標付きの幾何学的モデルを取得する。第 3 として、テクスチャ座標付きの幾何学的モデルの各頂点に対して、サンプリング時のポーズ及びサンプリング写真の画素値に基づいて対応する有効な実際撮影データを見つけて、高次元点群を構成してネットワークに入力して、乱反射特徴ベクトル及び鏡面反射特徴ベクトルを回復する。第 4 として、ネットワークから出力される乱反射特徴ベクトル及び鏡面反射特徴ベクトルに基づいて、各頂点に対して *LBFGS-B* 方法によってレンダリングのための座標系及び粗度をフィッティングし、信頼領域アルゴリズムによって鏡面反射率及び乱反射率を求める。

【0046】

図 8 には上記システムにより回復された検証セットにおける 2 つの *Lumitexel* ベクトルを示し、左側の列が

$\tilde{m}_s$

であり、右側の列が対応する m_s である。

【0047】

図 9 には上記システムによりサンプリング物体に対してマテリアル外観走査を行って回復したマテリアル属性結果を示し、第 1 行がそれぞれサンプリング物体の

$(\rho_{dR}, \rho_{dG}, \rho_{dB})$

の 3 つの成分を示し、第 2 行がそれぞれサンプリング物体の

$(\rho_{sR}, \rho_{sG}, \rho_{sB})$

の 3 つの成分を示し、第 3 行がそれぞれサンプリング物体の粗度係数 α_x 、 α_y を示し、

階調値が数値の大きさを代表する。

【 0 0 4 8 】

以上は単に好適な実施例であり、本発明は上記実施形態に限らず、同じ手段で本発明の技術的効果を実現するものであれば、いずれも本発明の保護範囲に属すべきである。本発明の保護範囲内に、その技術案及び／又は実施形態は種々の異なる修正及び変化を行うことができる。

10

20

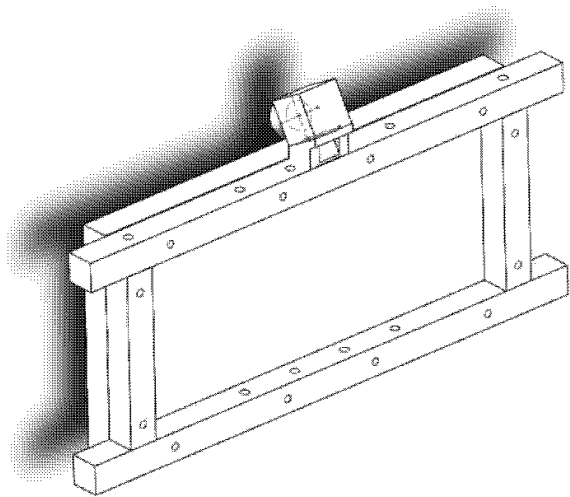
30

40

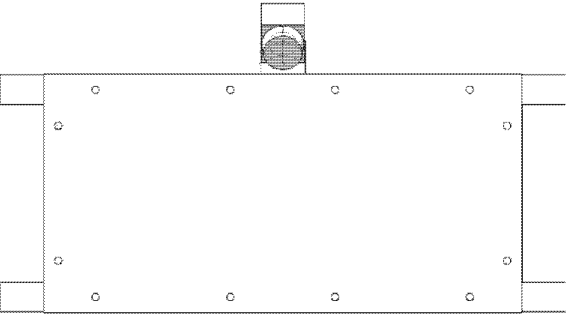
50

【図面】

【図 1】

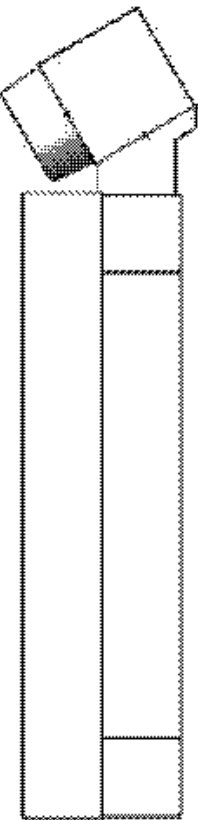


【図 2】

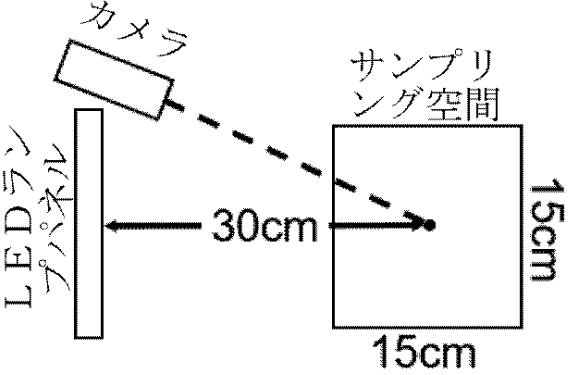


10

【図 3】



【図 4】



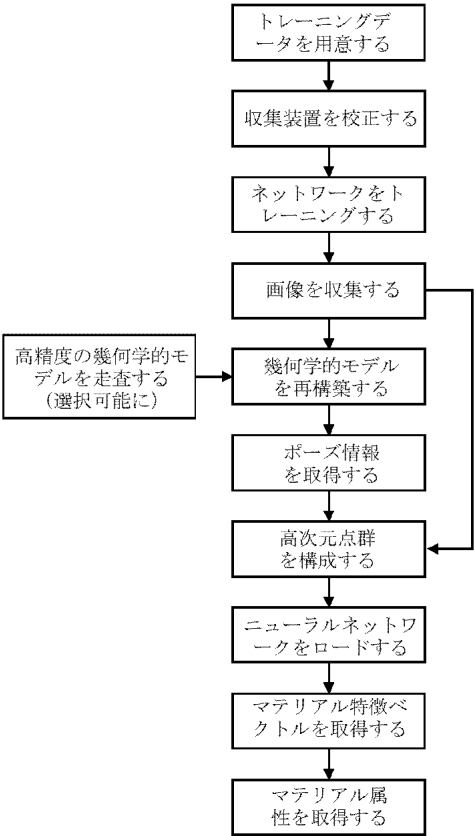
20

30

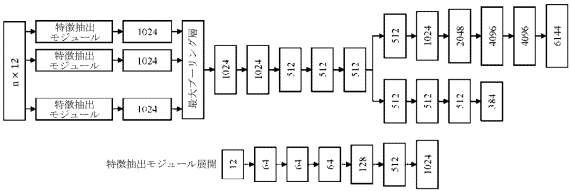
40

50

【図 5】



【図 6】



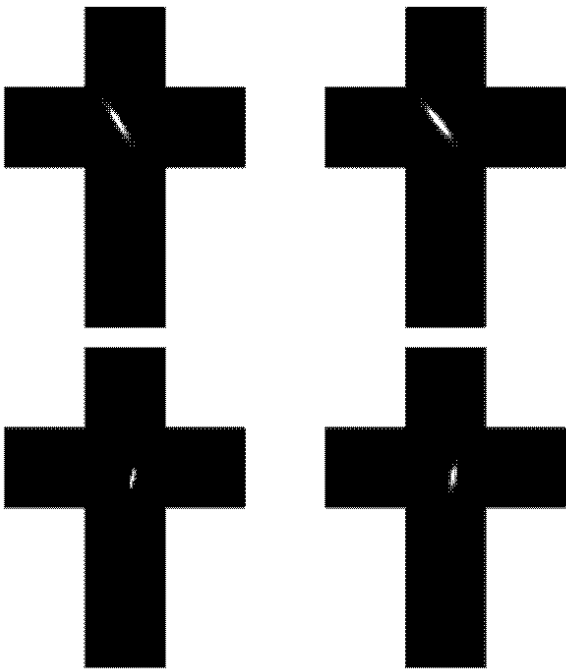
10

20

【図 7】



【図 8】

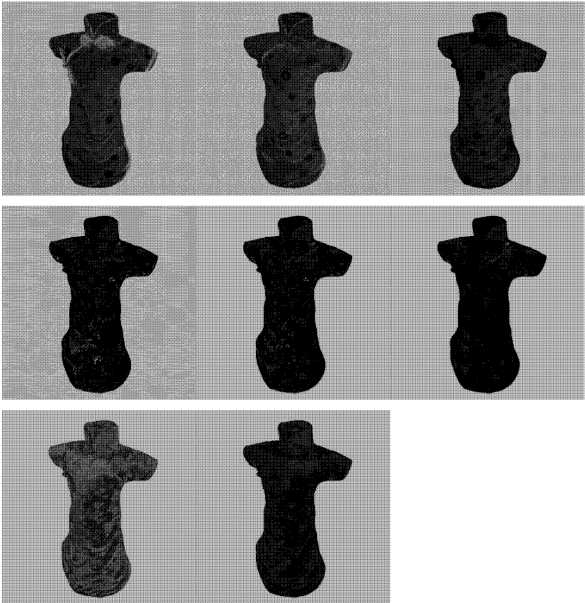


30

40

50

【図 9】



10

20

30

40

50

フロントページの続き

余杭▲塘▼路866号

審査官 菊池 伸郎

- (56)参考文献 特開2006-277748 (JP, A)
特開2019-219928 (JP, A)
中国特許出願公開第110570503 (CN, A)
- (58)調査した分野 (Int.Cl., DB名)
- G06T 7/00-7/90
G06T 1/00
G06T 15/00-19/20
G06T 11/00-11/80