

# 構造色パターンから光源方向を推定可能なARマーカ

伊藤 澄美<sup>\*1</sup> 浦西 友樹<sup>\*1</sup> ラサミー ポチャラ<sup>\*1</sup> 間下 以大<sup>\*1</sup> 竹村 治雄<sup>\*1</sup>

Kiyomi Ito<sup>\*1</sup>, Yuki Uranishi<sup>\*1</sup>, Photchara Ratsamee<sup>\*1</sup>, Tomohiro Mashita<sup>\*1</sup> and Haruo Takemura<sup>\*1</sup>

Abstract – 拡張現実 (Augmented Reality; AR) において実世界に仮想オブジェクトを合成する とき、実世界と仮想オブジェクトとの整合性が維持されていることで、より違和感のない描画が可能となる. 仮想オブジェクトと実世界との整合性を維持するために重要な要素の一つが、実世界と仮想世界の光源環境 を一致させることによる光学的整合性である.光学的整合性を保つためには、実世界における光源環境の情 報を取得し、仮想オブジェクトに反映させる必要がある.本研究では、光の当たり方と視点の方向によって 色の見え方が異なる構造色を用いた AR マーカを提案する.本稿では、カメラで観測された AR マーカの 位置および姿勢と構造色シートの色情報から、機械学習により AR マーカからの光源方向を推定する手法 について報告する.

Keywords : 拡張現実,光学的整合性,構造色,機械学習

# 1 はじめに

拡張現実 (Augmented Reality; AR), 複合現実 (Mixed Reality; MR) 技術において実世界に重畳される仮想 オブジェクトは、コンピュータグラフィクス技術の発 展もあり、より高精細になりつつある.一方で、実世 界に存在する物体は周囲の環境の影響を受けているた め、AR を高品質にするためには、実物体と同様に、 仮想オブジェクトを重畳する実世界の環境に合わせて 描画し、仮想オブジェクトと実世界の間の整合性を保 つ必要がある [1].

これらの整合性として、物体同士の物理的な位置関 係における整合性である幾何学的整合性、時間経過に よる物体の変化における整合性である時間的整合性, 光源の色や位置姿勢など、光源環境における整合性で ある光学的整合性の3つが挙げられる.この中で,幾 何学的整合性と時間的整合性については、AR におい て一般的な手法の1つである AR マーカを用い,マー カの位置、姿勢、大きさを基準として仮想オブジェク トを重畳することで、比較的容易に維持できる、しか し、光学的整合性を保つために必要な光源の情報、特 に光源の方向の情報は、一般的な平面形状の AR マー カでは取得が困難である.形状や材質が事前に分かっ ている立体物を画面に映し、その見え方から光源情報 を取得する手法がいくつか存在している [2][3] が,平 面物に比べ立体物は利用や配布における制約が大きく, 利用できる場面が限られる.

本研究では,構造色シートを貼り付けた AR マー カを用いて,カメラにより観測されたマーカの姿勢お よび構造色パターンから,光源方向を推定できる AR マーカを提案する.

構造色とは、光の波長よりも微細な物体表面の物理 的構造によって光路差が発生して観測される色であり、 日常生活の中ではコンパクトディスクやシャボン玉な どの表面で観察できる.構造色には、入射する光に含 まれている波長、光源方向および視点の方向に依存し て観測される色が変化する特徴があるため、構造色の 発色パターンは光源情報の取得に利用できると考えら れる.また、構造色は薄い平面形状でも発色するため、 立体形状を用いて光源情報取得する手法より幅広い場 面での利用が期待できる.このARマーカを用いて光 源方向の情報を取得し、マーカ上部に重ね合わせる仮 想オブジェクトのライティングに使用することで、仮 想オブジェクトを描画する際の光学的恒常性の維持に 寄与することが期待される.

構造色のパターンから周辺環境を推定するには,発 色する構造色パターンをモデル化する必要がある.構 造色のパターンを実時間で描画する手法は井村らによ り提案されている [5] ものの,それらは物理的な正し さを考慮していない.逆に,物理的な正確性を指向し た構造色モデルは計算量が増大する傾向がある.実時 間での動作には考慮範囲を狭める必要があり,また, ルックアップテーブルやミップマップの使用,GPU上 での処理を要求される [6].したがって,これらのモデ ルを用いて,ユーザの端末上で構造色パターンのレン ダリングおよび光源環境推定を実時間で実行すること は困難である.よって,AR での利用を想定している 本研究では,実時間で光源方向を推定し,その情報を 用いて仮想オブジェクトを実世界に重畳するため,事

<sup>\*1</sup>大阪大学

<sup>&</sup>lt;sup>\*1</sup>Osaka University

前の機械学習により構造色パターンをモデル化するこ とを試みる.

以下,2節では従来の実世界の光源環境を推定する ための技術について述べる.3節で,構造色パターン と機械学習を用いて光源方向を推定する手法を提案 手法として述べる.4節で,暗室で構造色パターンを 撮影したデータを提案手法に適用し,光の方向の推定 を行った結果を述べる.最後に,5節で本研究をまと める.

#### **2** 関連研究

実環境における光源環境を取得するため,様々な手 法が提案されている.以下ではその一部を紹介する.

# **2.1** 立体物を用いた手法

Debevec ら [2] は、鏡面球と拡散板を組み合わせて 作成したライトプローブを環境の中に置いて撮影し, 鏡面上の映り込みと映った拡散板の色から,ライトプ ローブに当たっている複数の明るい光源の強度を推 定する手法を提案している.また, Calian ら [3] は, 複数の仕切りを組み合わせて形成された立体形状の シェーディングプローブを撮影することで、光ではな くシェーディングを推定し、AR における仮想オブジェ クトのシェーディングに利用する手法を提案している. Knorr らの手法 [4] では、画面内の人間の顔の特定の 点について事前に顔の画像のデータセットで学習した 結果を適用することで,専用のプローブを用意せずに 顔に当たっている実環境の光の方向と色の情報を取得 できる.仮想オブジェクトをレンダリングする際,取 得した照明条件を適用することで、顔の陰影と一貫し たライティングや、仮想オブジェクトの影の顔への重 畳などを行う.

これらの手法では、高い精度で実世界の光源環境の 情報を取得することが可能である.しかし、実用にあ たっては、Debevec らの手法と Calian らの手法につ いては専用の立体物をユーザの手元に用意する必要が あり、Knorr らの手法はユーザの顔が画面に映る場面 でしか利用できない.幅広い場面での利用を考えるな らば、背景となる実環境に依らず、準備が容易な平面 形状のマーカで光源環境を取得できることが望ましい.

# 2.2 構造色シートを用いた手法

浦西ら [7] は、構造色シートを AR マーカに貼り付 けて用い、平面形状のマーカのみで、あらかじめ決め られた一定の角度からマーカを撮影し、光源方向を推 定する手法を提案した.この手法では、事前に1方向 から光を当てた構造色シートの色のパターンを撮影し、 その時のマーカの姿勢と光源方向を記録しておく.そ の後、撮影時と同じ姿勢で画面に映っているマーカの 色を、撮影したデータと比較し、最近傍のデータの撮 影時の光源方向を用いて仮想オブジェクトをライティ ングする.

### 3 提案手法

提案手法では,平面形状のマーカで実世界の光源情報を推定するため,構造色シートを貼付した AR マーカを用いる.本手法は AR での利用を想定しているため,オフラインでの事前計算の結果を利用し,実時間で推定することを目指す.事前に用意できる構造色のパターンの画像の量には限界があるため,先述した浦西らの手法による最近傍のデータを参照する推定では,入力に近いデータが用意したものの中にない場合の推定が適切にできないという問題がある.提案手法では,より広い範囲の入力に対応するため,用意したデータを用いて機械学習を行い,構造色パターンとマーカ姿勢から光源方向を推定するモデルを構成して利用する.

## 3.1 処理の流れ

まず,構造色マーカに,ある方向  $(L_{\theta}, L_{\phi})$ から光を 当て,その光の方向を変えていく様子を,角度と位置 を固定したビデオカメラで撮影する.撮影時と異なる 視点からマーカを観測している場合に対応するため, マーカの角度  $(M_{\theta}, M_{\phi})$ を変更した動画を複数撮影す る.動画は1フレームごとの画像に分割して,学習に 用いる.

次に、画像を撮影した時の光源方向の情報から、ラ イトマップを生成する. このライトマップのピクセル 数を縦横どちらも *s* とすると、角度 ( $\phi$ , $\theta$ ) から入射 する光を、座標 ( $x = (1 + (\theta/\pi)\cos\phi) \times (s/2), y =$  $(1 - (\theta/\pi)\sin\phi) \times (s/2)$ )のピクセルで表す. 光源方向  $(L_{\theta}, L_{\phi})$ に対して、座標 ( $x = (1 + (L_{\theta}/\pi)\cos L_{\phi}) \times$  $(s/2), y = (1 - (L_{\theta}/\pi)\sin L_{\phi}) \times (s/2)$ )のピクセル と、そのピクセルと一定距離以内のピクセルのみが輝 度値 1、それ以外は輝度値 0 となるライトマップを表 現する球面調和関数を学習の正解値に用いる. 本研究 では、Ramamoorthi らの手法 [8] を参考にライトマッ プから球面調和関数の係数を求めた.

その後、生成した学習データを用いて機械学習を行う. 縦 h,横 w ピクセルの RGB の 3 チャンネルの 画像に, $M_{\theta}$  で埋めた 4 チャンネル目, $M_{\phi}$  で埋め た 5 チャンネル目を加えた, $h \times w \times 5$  のデータ群を 入力として与える.学習には,図 1 に示すネットワー クを使用する.活性化関数には、出力層では正解値が 負である場合もあるためハイパボリックタンジェント を、それ以外では ReLU (Rectified Linear Unit)を使 用した.この出力値を、その学習データ撮影時の光源 方向 ( $L_{\theta}, L_{\phi}$ )から生成した光源環境を表現する球面 調和関数 f の係数に近づけるように学習を行う.



図 1 ネットワークの構造 Fig. 1 structure of the network

## 4 実験および考察

本節では、実際に撮影したデータを用いて深層学習 によりモデルを構成し、そのモデルによる別の撮影 データの光源方向推定の精度を調べた結果について述 べる.

#### 4.1 実験内容

構造色シートは、光源の方向と分光分布、観測する 視点の位置姿勢によって見える色が変わるが、構造色 パターンが観測されない光源と視点の組み合わせが存 在する. そのような状況になり光源方向の推定を正し く行うことが難しい場合を少なくするため、4枚の構 造色シートを溝の角度が 45 度ずつ異なるように並べ て,構造色 AR マーカを作成する. このマーカを構成 する4つの構造色シートと、各構造色シートに彫ら れた溝の向きを図2に示す.本実験では、1ミリメー トルあたり 1000 本の溝が彫られた構造色シートを 1 辺5センチメートルの正方形に切り取って使用した. ビデオカメラ (SONY NEX-VG10) を用いて 1920 × 1080 画素のマーカの動画像を撮影した. 懐中電灯に ディフューザーを取り付けて面光源としたもの1つを 照明として用い, 暗室で撮影を行うことで, 光源方向 を1方向に限定して構造色パターンのデータを取得す る. 図5にこのセットの外観を示す.

撮影した動画をフレームごとに分割し,射影変換で 構造色シートの部分のみを切り出したのち,その時の 光源方向  $(L_{\theta}, L_{\phi})$  およびカメラから見たマーカの角 度  $(M_{\theta}, M_{\phi})$  と紐づける.また,切り出した画像を L\*a\*b\*色空間に変換したのち L\*平滑化を行い輝度値 を一定にして,画像の色相と彩度の情報のみを学習で きるようにする.射影変換のみを行った状態のマーカ の例を図3に,L\*平滑化を行ったマーカの例を図4に 示す.

懐中電灯には再帰性反射素材でできたマーカを複数 個取り付けた.ビデオカメラでの撮影と並行して,暗 室内を複数個の赤外線カメラで撮影する.赤外線カメ ラの情報からある瞬間での懐中電灯の角度を求め,こ れを用いて,ビデオカメラで録画した各フレームにお



図 2 構造色シートの配置 Fig. 2 arrangement of structural color sheets



図 3 射影変換を行ったマーカの例 Fig. 3 example of markers done perspective transformation

けるマーカに当たっている光の角度  $(L_{\theta}, L_{\phi})$  を求める. この光源方向の情報からライトマップを生成し, さらにこのライトマップを表現する球面調和関数を求める.本実験では簡単化のため,球面調和関数の三次までの係数9つを求めた.

学習データと同条件で撮影を行ったテストデータに ついて、学習したモデルが出力した球面調和関数の9 つの係数を、正解となる光源方向に対応する球面調和 関数の係数と比較した.このときの評価値は、9つの 係数それぞれについての平均二乗誤差を合計した値と した.

## 4.2 実験結果

323,740 個の学習データを用いて,40,000 ステップ 学習させたモデルについて,80,932 個のテストデー タについて 9 つの係数を推測させた結果,9 つの係数 それぞれについての正解の値との平均二乗誤差の和は  $1.1869 \times 10^{-4}$ となった.テストデータの一部につい て,正解の値とモデルによる推測値の例を表1に示す.



図 4 L\*平滑化を行ったマーカの例 Fig. 4 example of markers done L\* smoothing



図 5 構造色パターン撮影セット外観 Fig. 5 Appearance of the set for taking photos of structural color pattern

# 4.3 考察

表1より,推測値に置ける係数1や係数3など,入 カデータと相関のない値を出力している場合が存在す ることがわかる.また,他の係数の推測値を正解値と 比較しても,二乗誤差の値が大きくなっているものが 多い.これらは,学習データの内容に起因するものと 考えられる.すなわち,構造色が観測されていない画 像や,撮影に使用した懐中電灯の一部が画面内に映り こみ,影になっている画像,構造色が観測される配置 であっても,構造色シートの光の反射が強く構造色を 撮影できなかった画像が多く含まれていたことが原因 だと推測される.これらの画像は,正解値は大きく異 なっても入力画像に大きな差がない学習データとなる ため,学習がうまくいかない原因になると考えられる.

#### 5 おわりに

本研究では,周囲の光源環境を推定できる平面形状のARマーカを作ることを目的として,構造色パター

ンを機械学習によってモデル化し,そのモデルを用い て構造色 AR マーカから光源方向を推定する手法を提 案した.実験の結果,現状では未だ十分な精度を得る ことができなかった.今後は,精度の高い光源方向推 定を行うため,学習データの精度向上が必要と考えら れる.加えて,ネットワークについてもさらなる検討 が必要であろう.

## 謝辞

本研究の一部は JSPS 科研費 JP16H02858 の助成を 受けた.

#### 参考文献

- P. Kán, H. Kaufmann: High-Quality Reflections, Refractions, and Caustics in Augmented Reality and their Contribution to Visual Coherence; IEEE International Symposium on Mixed and Augmented Reality (ISMAR), (2012.Nov)
- [2] P. Debevec et al.: A Single-Shot Light Probe; ACM SIGGRAPH, (2012.Aug)
- [3] D. A. Calian et al.: The Shading Probe: Fast Appearance Acquisition for Mobile AR; SIGGRAPH Asia, (2013.Nov)
- [4] B. Knorr et al.: Real-time illumination estimation from faces for coherent rendering; IEEE International Symposium on Mixed and Augmented Reality (ISMAR), (2014.Sep)
- [5] M. Imura et al.: A Generic Real-time Rendering Approach for Structural Colors; Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology, 95-102 (2009.Nov)
- [6] T. Cuypers et al.: Reflectance model for diffraction; ACM Transactions on Graphics (TOG), (2012.Aug)
- [7] Y. Uranishi et al.: The Rainbow Marker: An AR Marker with Planar Light Probe based on Structural Color Pattern Matching; IEEE Virtual Reality (VR), 303-304 (2016.Jul)
- [8] R. Ramamoorthii et al.: An Efficient Representation for Irradiance Environment Maps; ACM SIG-GRAPH, 497-500 (2001.Aug)

 $\odot~2018$  by the Virtual Reality Society of Japan (  $\rm VRSJ$  )

example						
		coefficient 1	coefficient 2	coefficient 3	coefficient 4	coefficient 5
	ground truth	0.0051	-0.0025	0.0083	0.0001	-0.0001
	estimation	0.0045	-0.0015	0.0053	0.0009	0.0007
$\mathbf{e}\mathbf{x}$	squared error	2.8786e-7	1.0940e-6	9.2879e-6	6.2208e-7	5.2958e-7
1		coefficient 6	coefficient 7	coefficient 8	coefficient 9	
	ground truth	-0.0054	0.0098	0.0002	-0.0008	
	estimation	-0.0016	0.0029	0.0010	0.0001	
	squared error	1.4537e-5	4.7672e-5	6.5878e-7	7.8119e-7	
		coefficient 1	coefficient 2	coefficient 3	coefficient 4	coefficient 5
	ground truth	0.0047	-0.0012	0.0061	0.0053	-0.0017
	estimation	0.0045	-0.0001	0.0053	0.0010	0.0004
ex	squared error	3.2993e-8	1.1150e-6	6.1632e-7	1.8264e-5	4.3817e-6
2		coefficient 6	coefficient 7	coefficient 8	coefficient 9	
	ground truth	-0.0020	0.0035	0.0088	0.0036	
	estimation	-0.0001	0.0028	0.0011	0.0001	
	squared error	3.3856e-6	4.3832e-7	5.9424e-5	1.2275e-5	
		coefficient 1	coefficient 2	coefficient 3	coefficient 4	coefficient 5
	ground truth	0.0039	0.0064	0.0020	-0.0004	-0.0008
	estimation	0.0045	0.0010	0.0053	0.0008	-0.0002
ex	squared error	4.0336e-7	2.8933e-5	1.0508e-5	1.4542e-6	8.8430e-7
3		coefficient 6	coefficient 7	coefficient 8	coefficient 9	
	ground truth	0.0043	-0.0031	-0.0002	-0.0068	
	estimation	0.0011	0.0027	0.0009	0.0002	
	squared error	1.0168e-5	3.4646e-5	1.2894e-6	4.8615e-5	

表 1 正解値,推測値,二乗誤差の例 Table 1 examples of ground truth, estimated values and squared errors