

直接・大域成分への分解のための撮影条件と分解処理の同時最適化 Image Acquisition and Processing for Direct-Global Separation

上田 宇起[†] 川原 僚[‡] 岡部 孝弘[‡]
Takaaki Ueda Ryo Kawahara Takahiro Okabe

1. はじめに

光源に照らされたシーンの画像は、直接成分と大域成分の 2 つの成分で構成されている。直接成分は、光源から出た光が直接照らすことで生じる成分であり、鏡面反射や拡散反射が含まれる。一方、大域成分は、光源から出た光がシーンの他の点を介して間接的に照らすことで生じる成分であり、相互反射や表面下散乱が含まれる。シーンの画像をこれらの成分に分解することは、3 次元形状復元[1]、シーンの光学的解析[2]、質感編集[3]などへの応用に重要である。

従来手法では、大域成分が低周波数成分であることに基づいて、白黒 2 値のチェッカーパターンなどの高周波数パターンをシーンに投影して撮影した画像から、直接成分と大域成分を求めている[1]。この手法では、理想的には 2 枚の画像から成分分解が可能である。しかしながら、実際にはプロジェクタの被写界深度やカメラの解像度の限界により、パターン境界ボケや焦点ボケといった多くの空間アーティファクトが発生するため理想的な撮影画像が得られない。そのため、理想的な撮影画像を仮定した物理モデルに基づくアプローチでは、少数の画像からの分解精度が悪化するという問題点がある。

成分分解手法ではシーンに投影するパターンと撮影画像を直接・大域成分に分解する分解処理が肝要である。投影パターンについては、物理モデルや信号処理理論に基づいた従来の白黒 2 値のチェッカーパターン[4]や多値パターン[5]が用いられている。分解処理に関しては、スパース性と平滑化に基づく手法[4]と機械学習を用いた手法[5]が提案されている。これらの研究は、投影パターンと分解処理の一方、または、各々を独立に最適化していること、および、理想的な物理モデルや信号処理理論に基づいて投影パターンを最適化していることに限界がある。

また、従来までは図 1(a), (b)のような同一露光時間での撮影画像数の増減のみしか考えられていなかった[1], [4], [5]。しかし、一定の撮影時間の中に、露光時間の配分を変えることで複数回の撮影を行うことができる。つまり、所望の撮影時間内に撮影する画像の枚数および各画像の露光時間を最適に設計することで、より高精度な成分分解が期待できる。

そこで、本研究では、データ駆動のアプローチで、分解処理だけでなく撮影条件も含めて、撮影から直接・大域成分への分解までの全体を最適化する手法を新たに提案する。具体的には、 1×1 の畳み込みカーネルを用いた投影パターンと画像枚数の表現・Softmax 関数を用いた露光時間の表現

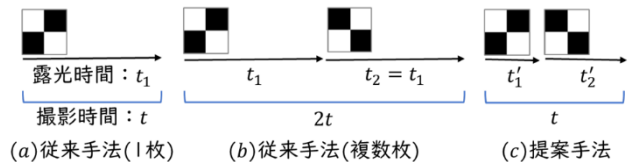


図 1 撮影時間と露光時間の関係

ができることに着目して、撮影条件（撮影枚数・投影パターン・露光時間）と分解処理の両方を、畳み込みニューラルネットワーク (CNN) の枠組みで同時に最適化する。

本研究の鍵は、無数に存在する投影パターン・撮影枚数・露光時間から CNN を用いて最適な投影パターン・撮影枚数・露光時間を学習している点である。従来までは[1], [4], [5], SN 比の最大化による高周波数の投影パターンの設計しか考慮されていなかった。これは、プロジェクタやカメラのハードウェア固有の空間アーティファクトにより撮影箇所ごとに SN 比が異なるため分解精度が悪化する。これに対し、提案手法では CNN における 1×1 の畳み込み処理と Softmax 関数を用いることで、無数の投影パターンから分解処理に最適な投影パターン・撮影枚数・露光時間の学習を実現する。実画像を用いた実験を行い、提案手法の有効性を示す。

本研究の貢献は以下の通りである。

- ・撮影条件（撮影枚数・投影パターン・露光時間）と分解処理を同時に最適化する新たなアプローチの提案
- ・既存の単一画像、複数画像による成分分解手法よりも高精度な成分分解の実現

2. 関連研究

2.1 直接・大域成分への分解

Nayar ら[1]は、一般に大域成分が低周波数成分であることに基づいて、プロジェクタから高周波数パターンを投影してカメラで撮影した画像を用いて、直接・大域成分に分解する手法を提案している。具体的には、シーンに互いに明暗が反転したチェッカーパターン 2 枚を投影し、それぞれカメラで撮影する。2 枚の撮影画像から画素ごとに線形演算を行うことで成分分解が可能となる。しかし、プロジェクタの被写界深度の浅さによる焦点ボケや、カメラの解像度の限界によるパターン境界のボケによって理想的な高周波数パターンが投影された画像の獲得は困難である。そのため、2 枚の投影パターンから分解を行うと、ボケの影響を受けるパターン境界付近では、格子状のアーティファクトが含まれた分解結果となる。一方、投影パターンをシフトしながら多数の画像を撮影することで良好な分解結果が得られることが知られているが、撮影時間が増大してしまい、動的シーンへの適用が困難になってしまう。

Subpa-Asa ら[4]は、直接・大域成分の空間的な平滑性を仮定し、それらをフーリエ基底または PCA 基底の線形結合として表現することで、単一画像から成分分解を行う手法

[†]九州工業大学 大学院情報工学府 情報創成工学専攻
Department of Artificial Intelligence, Kyushu Institute of Technology

[‡]九州工業大学 大学院情報工学府 知能情報工学研究系
Department of Artificial Intelligence, Kyushu Institute of Technology

を提案している。しかし、分解処理のみを最適化しており、分解の重要な手掛かりとなる投影パタンの最適化は行われていない。そのため、使用するチェッカーパタンの空間アーティファクトが成分分解に誤差を生じさせる原因となる。

Duan ら[5]は、分解処理に機械学習を用いた成分分解手法を提案した。投影パターンについては、ガウスノイズによるコントラストの低下を考慮して信号処理理論に基づき 2 値ではなく多値で作成している。しかし、投影パターン撮影時、ガウスノイズで表現できないパターン境界ボケや焦点ボケが生じる問題が存在する。そのため、信号処理理論に基づいた投影パタンの最適化には限界がある。

提案手法では、撮影画像から投影パターンも分解処理と同様に機械学習を用いて最適化を行う。これにより、実シーンにおける最適な投影パターンを獲得する。

2.2 センサと画像処理の同時最適化

一般的に、ディープニューラルネットワークは、画像処理タスクのエンドツーエンドの最適化のためのツールとして使用される。しかし、Duan ら[5]のように撮影済みの画像を入力として画像処理のみを最適化しているに過ぎず、所望の画像処理タスク全体の最適化は行えていない。

近年、光学イメージングモデルをパラメータ化して撮像層とし、それらの層をアプリケーション層（画像認識、画像生成、再構成などを行う）に接続し、最終的に逆伝播を利用してデータセットを学習し、センサと画像処理の同時最適化を行う手法が提案されている。Chakrabarti[6]はカラー画像の獲得のためにセンサのカラーフィルタと画像再構成手法の同時最適化を行った。Wu ら[7]は深層学習を利用した深度推定のために、レンズに装着する位相マスクと画像再構成の同時最適化を行った。

提案手法では、上述のようなセンサではなく、成分分解の鍵となる照明環境に着目し、撮影条件（撮影枚数・投影パターン・露光時間）と画像の直接・大域成分への分解処理の同時最適化を行う。

3. 提案手法

3.1 投影パタンの構造

従来手法では、直接・大域成分への分解を行う際、信号処理理論に基づいた高周波数パターンであるチェッカーパターン[1]、[4]や多値パターン[5]といった繰り返しパターンをシーンに投影し、カメラで撮影した画像を用いる。

提案手法では、通常、図 2 に示すような A 画素×B 画素のサイズを持つ基本パタンの繰り返しにより構成される投影パタンの輝度値を学習により最適化する。

3.2 畳み込みカーネルによる投影パタンの表現

提案手法では、画像符号化の国際標準方式に採用されている 2 次元離散コサイン変換（2 次元 DCT）の基底画像の一部の組み合わせにより投影パターンを表現する。2 次元 DCT を用いる動機は、チェッカーパタンの画素値が不連続に変化するのに対して DCT 基底の画素値は滑らかに変化するためである。また、比較的少数の DCT 基底を用いてチェッカーパターンを近似的に表現できるため、DCT 基底画像の一部を用いる。

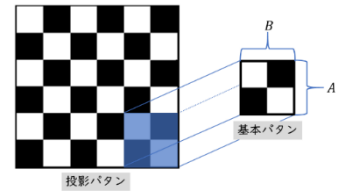


図 2 投影パターンと基本パターン

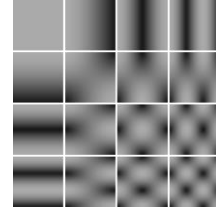


図 3 DCT 基底画像

重ね合わせの原理

$$\begin{aligned}
 & \text{重ね合わせの原理} \\
 & \text{重み} \times \text{基底画像} + \text{重み} \times \text{基底画像} + \dots \\
 & \begin{matrix} w_1 = -0.025 \\ w_2 = 0.185 \\ \vdots \\ w_{15} = -0.267 \\ w_{16} = 0.362 \end{matrix} \quad \begin{matrix} w_1: \text{結合係数}, N = 16 \end{matrix}
 \end{aligned}$$

1×1の畳み込み処理

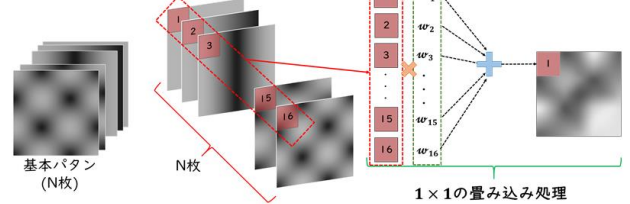


図 4 畳み込みカーネルによる投影パタンの表現

2 次元 DCT は、様々な大きさと周波数からなる正弦波である DCT の基底関数 $B_{u,v}(m,n)$ に重み $w(u,v)$ を掛けて加算することで $M \times N$ サイズの画像 $f(m,n)$ を表現する。

$$f(m,n) = \sum_u \sum_v w(u,v) B_{u,v}(m,n) \quad (1)$$

ここで、 $0 \leq u, m \leq M-1$, $0 \leq v, n \leq N-1$ である。

M 行 N 列の DCT の基底関数 $B_{u,v}(m,n)$ の式は以下の通りである。

$$B_{u,v}(m,n) = a_u a_v \cos \frac{\pi(2m+1)u}{2M} \cos \frac{\pi(2n+1)v}{2N} \quad (2)$$

$$a_u = \begin{cases} 1/\sqrt{M}, & (u=0) \\ \sqrt{2/M}, & (1 \leq u \leq M-1) \end{cases}$$

$$a_v = \begin{cases} 1/\sqrt{N}, & (v=0) \\ \sqrt{2/N}, & (1 \leq v \leq N-1) \end{cases}$$

このとき、式 2 は負の値を持つ。そこで、DCT の基底関数を画像化するため以下のように非負の値のみを持つような $B'_{u,v}(m,n)$ に変形する。

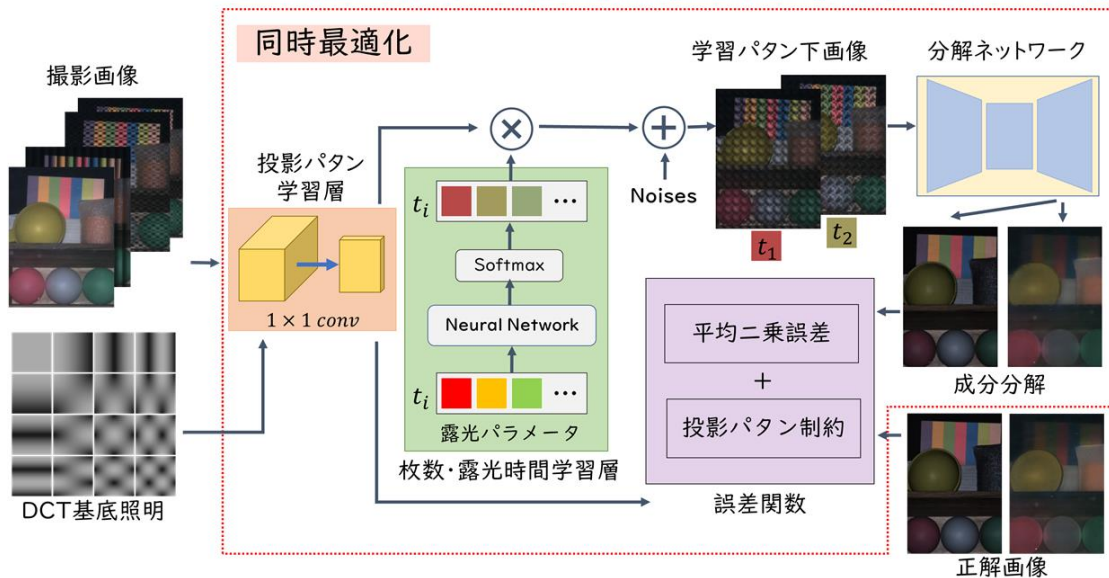


図 5 撮影条件と分解処理の同時最適化ネットワークの学習の流れ

$$B'_{u,v}(m, n) = B_{u,v}(m, n) + \left| \min(B_{u,v}(m, n)) \right| \quad (3)$$

式 3 により図 3 のような DCT 基底画像を計算できる。これにより得られた DCT 基底画像を投影パターンとしてシーンに投影することで、図 4 の上段に示すように、任意の投影パターンで照らされたシーンの画像を DCT 基底画像に照らされたシーンの画像の線形結合で表現できる (式 1)。DCT 基底照明下画像の結合係数 w を任意の値に設定することで、特定の照明パターン下の画像を表現できる。

また、図 4 の下段に示すように、CNN における 1×1 の畳み込み処理は、入力画像に対して画像ごとに重み w を付与し、画素ごとに合計する処理を行う。これら 2 つの処理は等価である。

提案手法では、基本パターンを 1×1 の畳み込みカーネルの重みで表せることに着目し、CNN の枠組みで学習を行う。具体的には、基本パターンの学習を 1×1 の畳み込みカーネルの重みの学習とみなす。これにより、最適な基本パターン下の画像は、入力である DCT 基底照明下の画像に対して、最適な基本パターンを表現する 1×1 の畳み込みカーネルを用いた畳み込み演算により得られる特徴マップとして表現できる。

このことから、図 5 に示すように、成分分解を行う分解ネットワークの前に 1×1 の畳み込み層を接続し、最適な投影パターンの学習を行う。これにより、学習された 1×1 の畳み込みカーネルの重みから最適な投影パターンを最少で 1 枚から取得可能となり、カーネルのフィルタ数 N_{filter} を変えることで任意枚数の投影パターンが取得できる。

3.3 撮影枚数と露光時間の学習

一般に任意光源下のシーンの撮影画像の画素値は、撮影するカメラの露光時間に比例する。提案手法では図 1(c) に示す一定の撮影時間下での最適な撮影枚数と露光時間の学習を実現するための層をネットワークに接続する。

図 5 における枚数・露光時間学習層では、 1×1 の畳み込みカーネルのフィルタ数と同数の露光時間パラメータ $t_i (1 \leq i \leq N_{filter})$ の値を全結合層からなるネットワークで

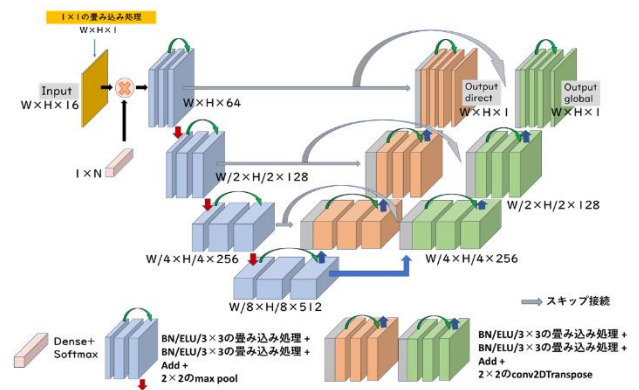


図 6 同時最適化ネットワークの詳細

学習を行う。最後に Softmax 関数を用いることで t_i は所望の撮影時間 t に対して、 $0 \leq t_i \leq t$ 、 $\sum_i t_i = t$ を満たす値を持つ。これにより、最適な撮影枚数と露光時間の実現を行う。

3.4 ノイズの付与

実際に投影パターン下で撮影した画像にはノイズが含まれる。そのため、学習時についても同様の条件にするため、線形結合画像に 2 種類のノイズを付与した画像を学習パターン下画像とする。具体的には、撮影画像にランダムに混入する観測ノイズとして、平均 0、標準偏差 1 のガウスノイズを付与する。また、画像を実数から整数に変換 (A/D 変換) する処理により生じる量子化ノイズについては、四捨五入と同等の丸め誤差であると言えるため、 -0.5 から 0.5 の範囲で一様分布するノイズを付与する。これらノイズの付与により、ネットワークにより作成する学習パターン下画像と実際の投影パターン下画像を同等として扱う。

3.5 撮影条件と画像分解の同時最適化ネットワーク

本研究では、図 6 に示すように、画像を直接成分と大域成分へ分解する画像分解ネットワークとして、エンコー

ダ・デコーダ構造にスキップ接続を加えた手法である U-Net[8]に基づいた構造を用いると同時に、学習時の勾配消失を避けるため ResNet[11]に基づいた残差ブロック構造を導入する。このとき、デコーダ部分を 2 つ用意し、エンコーダ部分は共通のものを使用する。この 2 つのデコーダによりそれぞれ直接成分と大域成分への分解を行う。

具体的には、投影パターンと画像分解の同時最適化ネットワークに対して、学習時に、C 枚の DCT 基底照明下画像から $W \times H$ 画素で切り出した $W \times H \times C$ のサイズの画素値を入力、真の直接・大域成分を入力に対応する $W \times H \times 1$ のサイズの画素値で切り出し正解としてネットワークに与える。また、学習データと同様に交差検証用データも同じサイズでネットワークに与える。最終的にネットワークから直接成分と大域成分の値をそれぞれ $W \times H \times 1$ のサイズで抽出する。テスト時は、出力された $W \times H$ サイズの画像を並べて、シーンの画像の直接・大域成分への分解結果とする。

3.6 損失関数

撮影条件と画像分解の同時最適化ネットワークでは次のような損失関数 \mathcal{L} を最小化することで学習される。

$$\mathcal{L} = \mathcal{L}_{mse} + \mathcal{L}_{light} \quad (4)$$

ここで、平均二乗誤差 \mathcal{L}_{mse} は以下の式で表される。

$$\mathcal{L}_{mse} = \mathcal{M}(\hat{I}_d, I_d) + \mathcal{M}(\hat{I}_g, I_g) \quad (5)$$

\hat{I}_d および \hat{I}_g は真の直接成分および大域成分の画素値であり、 I_d および I_g は同時最適化ネットワークで予測される直接成分および大域成分の画素値である。また、 \mathcal{M} は平均二乗誤差関数である。

また、投影パターン制約 \mathcal{L}_{light} は以下の式で表される。

$$\mathcal{L}_{light} = \begin{cases} 0, & (L_{min} \leq L \leq L_{max}) \\ \left(\frac{L_{max} - L_{min}}{2} - L \right)^2, & (\text{otherwise}) \end{cases} \quad (6)$$

L_{max} は画像化可能な画素値の上限、 L_{min} は画像化可能な画素値の下限であり、 L は 1×1 の畳み込み後に得られた学習パターンの画素値である。

4. 実験

4.1 実験環境

実験では、Crosstour 製の LED プロジェクタ P970 を用いて照明を行い、FLIR 社製のカメラ Blackfly S USB3 を用いて撮影した。このとき、図 7 のように、50-50 プレート型ビームスプリッターを使用して、プロジェクタ画素とカメラ画素の対応がシーンの深度に対して不変になるように、プロジェクタとカメラをビームスプリッターを介して同軸に配置するとともに、画素間の対応を事前に較正している*1。

4.2 学習データとパラメータ設定

図 7 のセットアップを用いて、26 シーンに対して撮影を行った。26 シーンのうち 22 シーンを学習データ、1 シーンを交差検証用データ、3 シーンをテストデータとした。被写体として、図 8 に示すようなアロマキャンドル、包装紙、

*1 提案手法では、白黒の 2 値パターンを用いた Nayar らの手法と異なり、撮影画像の画素値の大小から対応を求めることができないため、ビームスプリッターと事前の較正が必要である。

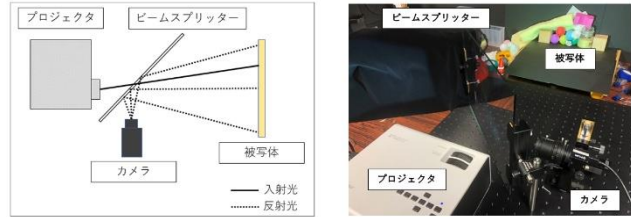


図 7 実験のセットアップ

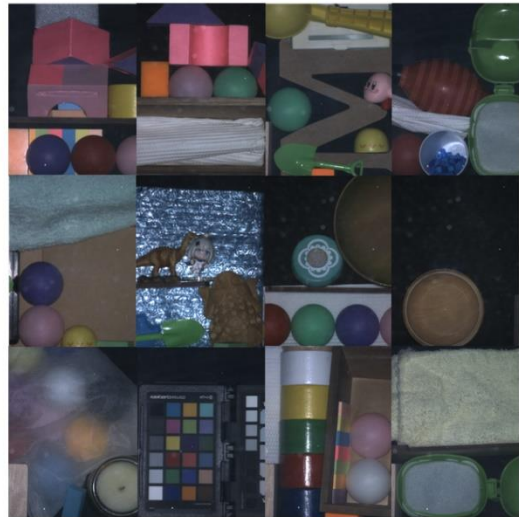


図 8 撮影対象の例

布、ピンポン玉などの表面下散乱や相互反射の生じやすい物体を用いた。

実験では、 20×20 のサイズからなる基本パターンを学習することを考えた、シーンの撮影時、図 3 に示す DCT 基底画像 16 枚を順に投影して撮影した。このとき、DCT 基底画像を 400 枚すべてではなく 16 枚のみを使用した理由は、基本パターンにおけるチェッカーパターンの 1 マスは 10×10 のサイズであったため、学習する投影パターンはそこまで高周波成分を使用する必要がないと判断したためである。また、投影パターンをプロジェクタで投影する際に、プロジェクタの γ 補正を考慮して画素値を調整して投影した。これにより得られた 22 シーンの撮影画像から、 $W=80$, $H=80$, $C=16$ として、 80×80 画素の領域を取り出すことで作成した 9,900 枚を学習データとした。また、Nayar ら[1]の手法に基づき、チェッカーパターンをシフトさせながら撮影した 25 枚の画像から真の直接・大域成分の正解画像を作成した。

ネットワークの学習では、最適化アルゴリズムとして Adam[9]を用いて初期の学習率を 0.01、 $\beta_1=0.9$ 、 $\beta_2=0.999$ とし、学習率は、学習が進むにつれて段階的に 0.001、0.0001 と下がるように設定した。ネットワークのすべての重みは He の正規分布[10]を用いて初期化した。また、提案手法では、 1×1 の畳み込みカーネルのフィルタ数を変えることで、最少 1 枚から任意の画像枚数で画像分解を行うことができる。例えば、フィルタ数を 1 とすることで単一画像からの成分分解が行え、2 とすると 2 枚の画像から成分分解を行う。実験では、フィルタ数を変化させることで、最適な撮影枚数と露光時間の導出を行った。露光時間の最適化の際には、実環境で発生する投影パターン画像の切り替えに要する時間であるオーバーヘッドを設定した。オーバ

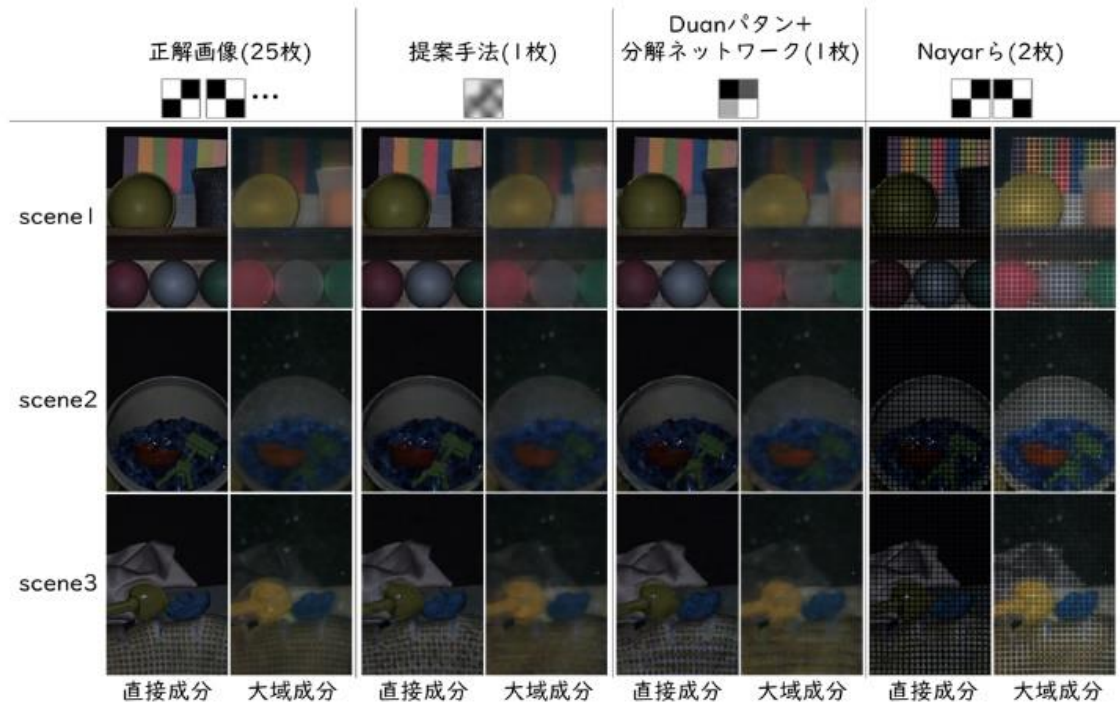


図 9 各手法との比較

ーヘッドとして 2 種類を考え、1 つ目が 60Hz で動作するプロジェクタの理想的な切り替え時間 0.0167s であり、2 つ目が図 7 のセットアップを用いて撮影時に要した 0.6s である。

また、投影パタンを学習することの有効性を示すため、投影パタンを Duan ら[5]が使用した 4 つの光源強度から構成される多値パタンを使用した場合の分解ネットワークの学習を行った。

4.3 従来手法との比較

図 9 は、正解画像、提案手法が単一画像から成分分解を行った結果、Duan らの 4 つの光源強度から構成される多値パタンを用いて成分分解を行った結果、Nayar らの 2 枚を用いて成分分解を行った結果を示している。Nayar らの手法では正解画像撮影時と同じ大きさである 20×20 の基本パタンから構成されるチェッカーパタンを使用し、Duan らの手法でも同様の大きさである 20×20 の基本パタンから構成される多値パタンを使用している。Nayar らの手法では、投影パタンボケの影響により、格子状のアーティファクトが非常に目立った結果となっている。また、Duan らの投影パタンを使用した場合でも、Nayar らの手法ほど顕著ではないが、投影時に目立つ境界パタンボケの影響から直線状のアーティファクトがあることが分かる、一方で、提案手法ではアーティファクトはほとんど発生せず、定性的に良好な結果が得られたことが確認できる。

また、表 1 に PSNR と SSIM による定量評価を示す。投影パタンと分解処理を同時に最適化した場合のほうが、Duan らの多値パタンを用いて分解ネットワークのみを最適化した成分分解よりも定量的に良好な結果が得られたことが確認できる。これより、画像分解だけではなく、投影パタンも最適化することが重要であるということが分かる。

表 1 定量評価による各手法との比較

		scene1		scene2		scene3	
		直接成分	大域成分	直接成分	大域成分	直接成分	大域成分
提案手法 (1 枚)	PSNR	34.066	35.328	33.966	36.196	31.346	34.883
	SSIM	0.9411	0.9390	0.9311	0.9381	0.9002	0.9249
Duan パタン (1 枚)	PSNR	32.552	34.458	33.154	35.458	29.309	33.378
	SSIM	0.9309	0.9341	0.9225	0.9339	0.8620	0.9154
Nayar ら (2 枚)	PSNR	20.205	19.636	24.269	23.748	21.532	20.928
	SSIM	0.6212	0.6972	0.6644	0.7800	0.6369	0.7063

4.4 撮影条件の違いによる分解結果の比較

図 10 は、一定の撮影時間内で露光時間などの撮影条件と画像分解の最適化を行った際に得られた投影パタンと成分分解結果である。また、表 2、3 に定量評価を行った結果を示す。結果より、理想的なオーバーヘッド時は、一定の撮影時間内で 2 枚の画像よりも 3 枚の画像を撮影した場合の方がより高精度な成分分解が行えているが、実環境では、オーバーヘッドの影響により、露光できる時間が少なくなってしまうため、3 枚の画像を使用するよりも 2 枚の画像を使用した方が高い精度が得られていることが分かる。

5. おわりに

本稿では、プロジェクタ-カメラシステムを用いた直接成分と大域成分の分解手法を提案した。具体的には、撮影条件(撮影枚数・投影パタン・露光時間)と分解処理の両方を CNN の枠組みで同時に最適化することで、最少 1 枚の画像から従来手法よりも高精度な成分分解を行った。今後の展望として、焦点ボケにも頑健な成分分解の学習や動的シーンへの適用に取り組みたい。

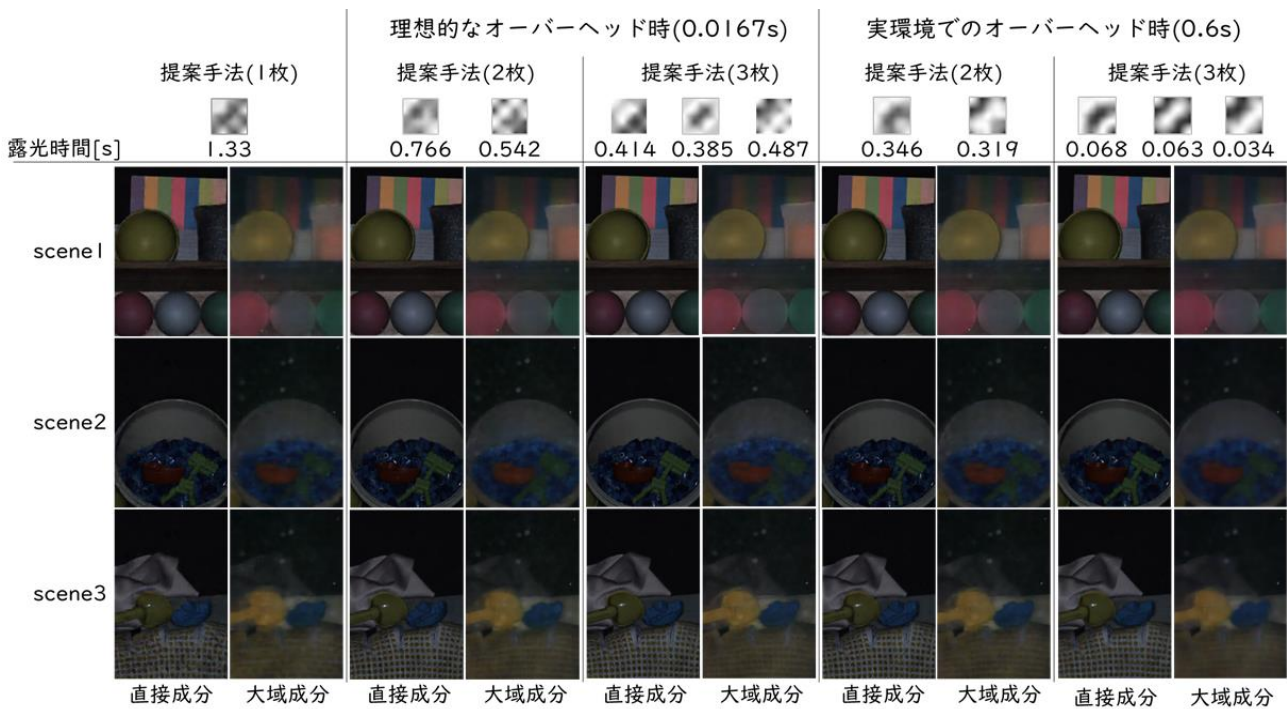


図 10 撮影条件と画像分解の同時最適化結果

謝辞

本研究の一部は、JSPS 科研費 JP23H04357, および JP20H00612 の助成を受けた。

参考文献

[1] S. K. Nayar, G. Krishnan, M. D. Grossberg, R. Raskar, "Fast Separation of Direct and Global Components of a Scene using High Frequency Illumination", In Proc. ACM SIGGRAPH2006, 25(3), pp.935--944, 2006.

[2] Y. Mukaigawa, Y. Yagi, and R. Raskar, "Analysis of light transport in scattering media", In Proc. IEEE CVPR2010, pp.153-160, 2010.

[3] M. Grossberg, H. Peri, S. Nayar, and P. Belhumeur, "Making one object look like another: controlling appearance using a projector-camera system", In Proc. IEEE CVPR2004, pp.1-452-459, 2004.

[4] A. Subpa-Asa, Y. Fu, Y. Zheng, T. Amano, I. Sato, "Separating the Direct and Global Components of a Single Image", Journal of Information Processing 26, pp.755--767, 2018.

[5] Z. Duan, J. Bieron, P. Peers, "Deep Separation of Direct and Global Components from a Single Photograph under Structured Lighting", Computer Graphics Forum Vol. 39, No. 7, pp.459--pp470, 2020.

[6] A. Chakrabarti, "Learning Sensor Multiplexing Design through Back-propagation", In Proc. NIPS2016, pp.3081--3089, 2016.

[7] Y. Wu, V. Boominathan, H. Chan, A. Sanjarnarayanan, A. Veeraraghavan, "Phasecam3dlearning phase masks for passive single view depth estimation", In Proc. IEEE ICCP2019, pp.1--12, 2019.

[8] O. Ronneberger, P. Fischer, T. Brox, "U-net: Convolutional networks for biomedical image. Segmentation", In Proc. MICCAI2015, pp234--241, 2015.

[9] D.P.Kingma, J.Ba, "Adam:A Method for Stochanic Optimization", In Proc. ICLR2015, 2015.

[10] K. He, X. Zhang, S. Ren, J. Sun, "Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification", In Proc. IEEE ICCV2015, pp.1026--1034, 2015.

[11] He, K., Zhang, X., Ren, S., Sun, J., "Deep residual learning for image recognition." In Proc. IEEE CVPR2016, pp.770--778, 2016.

表 2 理想的なオーバーヘッド時の定量評価

		scene1		scene2		scene3	
		直接成分	大域成分	直接成分	大域成分	直接成分	大域成分
提案手法 (1枚)	PSNR	34.066	35.328	33.966	36.196	31.346	34.883
	SSIM	0.9411	0.9390	0.9311	0.9381	0.9002	0.9249
提案手法 (2枚)	PSNR	36.388	37.120	37.637	37.742	35.703	37.102
	SSIM	0.9641	0.9457	0.9636	0.9446	0.9556	0.9361
提案手法 (3枚)	PSNR	37.381	37.693	38.456	37.852	36.385	37.324
	SSIM	0.9675	0.9439	0.9653	0.9400	0.9521	0.9328

表 3 実環境でのオーバーヘッド時の定量評価

		scene1		scene2		scene3	
		直接成分	大域成分	直接成分	大域成分	直接成分	大域成分
提案手法 (1枚)	PSNR	34.066	35.328	33.966	36.196	31.346	34.883
	SSIM	0.9411	0.9390	0.9311	0.9381	0.9002	0.9249
提案手法 (2枚)	PSNR	37.168	37.376	37.852	37.697	36.121	36.928
	SSIM	0.9627	0.9446	0.9626	0.9416	0.9476	0.9330
提案手法 (3枚)	PSNR	36.455	35.840	37.57	36.390	35.013	35.412
	SSIM	0.9522	0.9265	0.9544	0.9217	0.9315	0.9146