

氏 名 Nie Shijie

学位(専攻分野) 博士(情報学)

学位記番号 総研大甲第 2161 号

学位授与の日付 2020 年 3 月 24 日

学位授与の要件 複合科学研究科 情報学専攻
学位規則第6条第1項該当

学位論文題目 Physics-based Deep Learning for Optical Property Analysis

論文審査委員 主 査 教授 佐藤 いまり

教授 大山 敬三

教授 杉本 晃宏

准教授 ZHENG YIN QIANG

教授 佐藤 真一

国立情報学研究所 コンテンツ科学研究系

(Form 3)

Summary of Doctoral Thesis

Name in full Nie Shijie

Title Physics-based Deep Learning for Optical Property Analysis

One of the important goals in computational photography is to capture and represent how lights interact with real scenes. To overcome the limitations of ordinary cameras, techniques such as combining imaging and data-driven computation to simulate optical processes have attracted a great deal of attention. This thesis are going to explore two essential parts of optical properties: light spectral properties and light transport properties, and we aim to infer them with an image taken by ordinary cameras. The two representations of these optical properties is hyperspectral image and global/direct illumination.

Deep learning has been recently thriving in many research fields: from recognition objects in images, to accelerate medical research by predicting potential drug molecules, DNA mutation on gene, or even master the Go game. Nearly in every field, deep learning-based methods achieve state-of-the-art performance than conventional methods. The major limitation of deep learning models is it might be difficult to understand what is going on inside. Without relying on scientific knowledge, deep learning model is a "black-box", only from trial and test on data. However, scientific problems are always under physical constraints. This thesis aims to reveal the "black-box" of deep learning networks and embedding physical prior knowledge into deep learning networks to improve performance of analyzing two representations mentioned above.

Hyperspectral imaginary and global/direct illumination imaginary is costly and requires complex hardware setting. For direct and global illumination, capturing them requires modulated active light and multiple images. To capture a hyperspectral image, Hyperspectral Imaging Systems (HISs) with complex hardware setting is required. For a normal RGB three channels image taken by an ordinary camera, reconstruct these two optical properties by RGB is ill-posed. The three channels dimensions image is a many-to-three mapping: for direct/global separation is six to three, for spectral reflectance is the number of spectral bands to three. This image is a subspace of the two high dimension space, and back-projection mapping is highly ill-posed. Several machine learning and deep learning methods were proposed, but their performance is not satisfying. The major limitation of deep learning is transparency. We proposed a framework, by encoding the physical optical process into a deep learning network, the acquisition and analysis process is jointly learned. We carefully designed the loss and structure of deep learning architecture to get better results and hardware-implemented

the first layer of the network to capture three channels image for inference. By replacing complex hardware settings with our proposed deep learning network, the computation and acquire cost is reduced greatly.

This thesis propose the first method to analyze global and direct components from a single RGB image without any hardware restriction. Our method is a novel generative adversarial network (GAN) based networks which impose the prior physics knowledge to force a physics plausible component separation. In the experiments, our method has achieved satisfactory performance on images from our own testing set with global and direct component and hyperspectral analysis public datasets. Furthermore, to analyze the hyperspectral image, as existing RGB cameras are tuned to mimic human trichromatic perception, we optimize a new spectral response that is necessary for hyperspectral reconstruction. We learn the optimized camera spectral response functions (to be implemented in hardware) and a mapping for spectral reconstruction by using an end-to-end network. Our core idea is that since camera spectral filters act in effect like the convolution layer, their response functions could be optimized by training standard neural networks. We propose four types of designed filters: a three-chip setup without spatial mosaicing, a single-chip setup with a Bayer-style 2x2 filter array, a non-invasive filter learning approach combined with existing camera response functions, and a jointly learned camera sensor coded spatial pattern and response function. Numerical simulations verify the advantages of deeply learned spectral responses over existing RGB cameras. More interestingly, by considering physical restrictions in the design process, we are able to realize the deeply learned spectral response functions by using modern film filter production technologies and thus construct data-inspired multispectral cameras for snapshot hyperspectral imaging. Finally, we simultaneously learn the camera spectral response (CSS) functions and a material classification network. We show that the proposed method has higher overall accuracy than existing cameras due to CSS optimization.

博士論文審査結果

Name in Full
氏名 Nie Shijie

Title
論文題目 Physics-based Deep Learning for Optical Property Analysis

本学位論文は「Physics-based Deep Learning for Optical Property Analysis」と題し、全7章から構成されている。第1章「Introduction」では、本研究の背景と目的を述べている。近年、大規模なデータセットと強力な計算パワーを背景として、ディープラーニングに代表されるような効率的な機械学習技術による実世界解析技術の発展が目覚ましい。本学位論文は、実世界で観察される光の伝搬ならびに計測デバイスによる計測プロセスを考慮した機械学習手法を提案し、シーンの分光情報ならびに光伝搬の分離を1枚または少数枚の入力から実現することを目的とした研究である。

第2章「Related Work」では、本学位論文が対象とするシーンの分光情報ならびに光伝搬解析に対し、実世界の分光スペクトル情報の大規模なデータセットに基づく機械学習により RGB 画像からハイパースペクトル画像を復元する技術の紹介と実在シーンで観察される直接反射・間接反射（相互反射や散乱等）を分離する関連研究について紹介し、本論文が取り組む課題と当該分野への貢献について説明している。

第3章「Physics-Based Constraints in Network Loss for Global Direct Separation」では、プロジェクタ・カメラシステムを用いて撮像された画像列から実シーンの直接反射と間接反射を分離して、様々な物体の光伝搬データベースを整備し、観察光は直接反射と間接反射の総和となるという物理的な拘束を加えることで、1枚の RGB 画像から安定にシーンの直接反射・間接反射を分離に成功している。さらに、提案技術により分離された直接反射・間接成分を異なる割合で再構築することで、やわらかさ、金属感といった質感の操作が可能な応用技術について検討している。

第4章「Physics-Based Design in Network Architecture for Hyperspectral Recovery」では、実在シーンの分光情報の復元に関して、計測デバイスであるカメラの分光感度の最適化をネットワークの構成要素として取り込み、入力から分光情報の再構成までをつなぐ新たな推定手法を提案し、実験により効果を検証した。さらに、シングルイメージセンサのケースを想定し、ハイパースペクトル画像復元に最適なカラーフィルタの最適配置（ベイヤーパターン）についても検討し、実験によりその効果を検証している。

第5章「Physics-Based Constraints for Hyperspectral Classification」では、機械学習によるカメラの分光感度の最適化を素材識別という実応用問題へと発展させ、識別問題へと上記光情報解析ネットワークを拡張させ、最適化されたカメラの分光感度による素材識別の効果を CAVE fake/real データベースやリモートセンシングデータベースを用いて評価した。

第6章「Conclusion」では、本論文の成果と当該分野への貢献をまとめている。

以上を要するに本学位論文は，実世界で観察される光の伝搬ならびに計測デバイスによる計測プロセスを考慮した機械学習手法を提案することで，シーンの分光情報ならびに光伝搬の分離を1枚または少数枚の入力から実現し，提案手法の実装ならびに性能評価により，その有効性を示したものである。これらの成果は，査読付き国際会議論文（フルペーパー）として，コンピュータビジョン分野のトップ会議(CVPR)ならびにメジャー会議ACCVで発表されており，学術的な貢献も認められる。機械学習を用いた計測デバイスの最適化によるハイパースペクトル情報の復元では，論文発表当時最高の復元精度を実現していることは特筆すべきことである。以上の理由により，審査委員会は本学位論文が学位の授与に値すると判断した。