

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2019-220176

(P2019-220176A)

(43) 公開日 令和1年12月26日(2019.12.26)

(51) Int.Cl.			F I			テーマコード (参考)	
G06T	1/00	(2006.01)	G06T	1/00	510	2G020	
G01J	3/51	(2006.01)	G01J	3/51		2H148	
H04N	9/04	(2006.01)	H04N	9/04	B	5B057	
G06T	1/40	(2006.01)	G06T	1/40		5C065	
G06N	3/04	(2006.01)	G06N	3/04	154		

審査請求 未請求 請求項の数 17 O L (全 35 頁) 最終頁に続く

(21) 出願番号 特願2019-110453 (P2019-110453)
 (22) 出願日 令和1年6月13日(2019.6.13)
 (31) 優先権主張番号 特願2018-114188 (P2018-114188)
 (32) 優先日 平成30年6月15日(2018.6.15)
 (33) 優先権主張国・地域又は機関 日本国(JP)

(71) 出願人 504202472
 大学共同利用機関法人情報・システム研究
 機構
 東京都立川市緑町10番3号
 (74) 代理人 100101454
 弁理士 山田 卓二
 (74) 代理人 100132241
 弁理士 岡部 博史
 (72) 発明者 ニエ・シジェ
 東京都千代田区一ツ橋二丁目1番2号 大
 学共同利用機関法人情報・システム研究機
 構 国立情報学研究所内

最終頁に続く

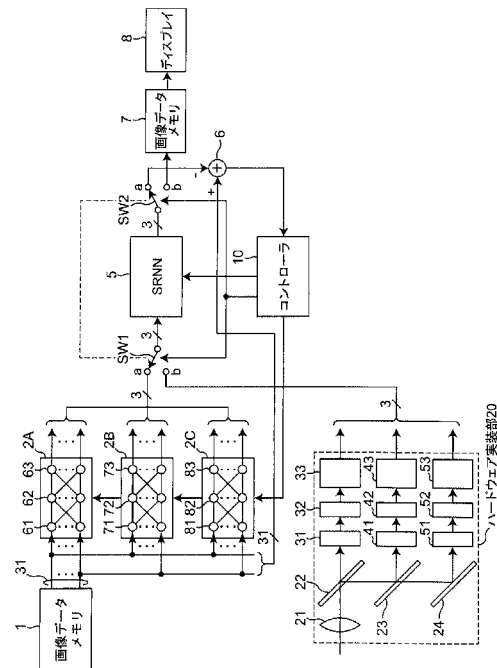
(54) 【発明の名称】 画像処理装置及び方法

(57) 【要約】

【課題】 スペクトル復元又はクラスタリングのための分光透過特性の最適化を行う。

【解決手段】 画像処理装置は、複数次元のハイパースペクトル画像データを、フィルタの分光透過特性に対応する応答関数を用いて、より低い複数次元のカラー空間の画像データに変換して出力する畳み込み層をそれぞれ有する複数の第1のニューラルネットワークと、複数の第1のニューラルネットワークからの画像データを複数次元のハイパースペクトル画像データに再構成するように変換して出力する第2のニューラルネットワークと、複数次元のハイパースペクトル画像データである学習画像データを用いて、第1及び第2のニューラルネットワークを、学習画像データと再構成された画像データとの誤差が最小になるように学習することで複数の第1のニューラルネットワークの各フィルタの分光透過特性に対応しかつ光学フィルタを形成するための応答関数を計算する制御手段とを備える。

【選択図】 図1



【特許請求の範囲】

【請求項 1】

複数次元のハイパースペクトル画像データを、少なくとも1つの光学フィルタの分光透過特性に対応する応答関数を用いて、前記ハイパースペクトル画像データの次元数より少ない複数次元のカラー空間の画像データに変換して出力する畳み込み層をそれぞれ有する複数の第1のニューラルネットワークと、

前記複数の第1のニューラルネットワークからの画像データを、前記複数の第1のニューラルネットワークの前記光学フィルタの分光透過特性に対応する応答関数を学習するための変換データに変換して出力する第2のニューラルネットワークと、

所定の学習データを用いて、前記第1及び第2のニューラルネットワークを、前記学習データと前記変換データとの誤差が最小になるように学習することにより、前記複数の第1のニューラルネットワークの前記光学フィルタの分光透過特性に対応しかつ前記光学フィルタを形成するための応答関数を計算する制御手段とを備えたことを特徴とする画像処理装置。

10

【請求項 2】

前記変換データは、複数次元のハイパースペクトル画像データであり、

前記学習データは、所定の複数次元のハイパースペクトル画像データであることを特徴とする請求項1記載の画像処理装置。

【請求項 3】

前記変換データは、前記複数の第1のニューラルネットワークからの画像データをクラスタリングしたクラスタリング結果データであり、

前記学習データは、前記クラスタリング結果データに対応し、前記クラスタリングを学習するための学習データであることを特徴とする請求項1記載の画像処理装置。

20

【請求項 4】

前記クラスタリングは、物体識別、産地識別、本人識別、個人認証、又は領域分割を含むことを特徴とする請求項3記載の画像処理装置。

【請求項 5】

前記第2のニューラルネットワークは、複数の畳み込み層と、複数の逆畳み込み層とを縦続接続して構成されたことを特徴とする請求項1～4のうちのいずれか1つに記載の画像処理装置。

30

【請求項 6】

前記制御手段は、前記光学フィルタの分光透過特性に対応する応答関数の関数値が0以上であって、かつ、L2ノルムのスムーズ正規化器を用いて前記光学フィルタの分光透過特性に対応する応答関数の変化が所定のスムーズ値以上のスムーズであるという制約条件にもとで、前記複数の第1のニューラルネットワークを学習することを特徴とする請求項1～5のうちのいずれか1つに記載の画像処理装置。

【請求項 7】

前記画像処理装置は複数の光学フィルタを備え、

前記画像処理装置は、前記計算された複数の光学フィルタの分光透過特性に対応する応答関数を有するように構成した複数の光学フィルタをそれぞれ前段に設けた複数の画像センサをさらに備え、

40

前記複数の画像センサは、撮像する画像光を前記各光学フィルタを介して受光して出力する画像データを、前記学習された第2のニューラルネットワークに入力し、

前記第2のニューラルネットワークから出力されるハイパースペクトル画像データを、前記撮像する画像光のハイパースペクトル画像データとして出力することを特徴とする請求項1～6のうちのいずれか1つに記載の画像処理装置。

【請求項 8】

前記画像処理装置は複数の光学フィルタを備え、

前記画像処理装置は、前記計算された複数の光学フィルタの分光透過特性に対応する応答関数を有するように構成した複数の光学フィルタを前段に設けた単一の画像センサをさ

50

らに備え、

前記単一の画像センサは、撮像する画像光を前記複数の光学フィルタを介して受光して出力する画像データを、前記学習された第2のニューラルネットワークに入力し、

前記第2のニューラルネットワークから出力されるハイパースペクトル画像データを、前記撮像する画像光のハイパースペクトル画像データとして出力することを特徴とする請求項1～6のうちのいずれか1つに記載の画像処理装置。

【請求項9】

前記画像処理装置は、前記計算された単一の光学フィルタの分光透過特性に対応する応答関数を有するように構成した単一の光学フィルタを前段に設けた複数の画像センサをさらに備え、

10

前記複数の画像センサは、撮像する画像光を前記単一の光学フィルタを介して受光して出力する画像データを、前記学習された第2のニューラルネットワークに入力し、

前記第2のニューラルネットワークから出力されるハイパースペクトル画像データを、前記撮像する画像光のハイパースペクトル画像データとして出力することを特徴とする請求項1～6のうちのいずれか1つに記載の画像処理装置。

【請求項10】

複数次元のハイパースペクトル画像データを、少なくとも1つの光学フィルタの分光透過特性に対応する応答関数を用いて複数次元のカラー空間の画像データに変換して出力する畳み込み層をそれぞれ有する複数の第1のニューラルネットワークと、

前記複数の第1のニューラルネットワークからの画像データを、前記複数の第1のニューラルネットワークの前記光学フィルタの分光透過特性に対応する応答関数を学習するための変換データに変換して出力する第2のニューラルネットワークとを備えた画像処理装置のための画像処理方法であって、

20

制御手段が、所定の学習データを用いて、前記第1及び第2のニューラルネットワークを、前記学習データと前記変換データとの誤差が最小になるように学習することにより、前記複数の第1のニューラルネットワークの前記光学フィルタの分光透過特性に対応しかつ前記光学フィルタを形成するための応答関数を計算するステップを含むことを特徴とする画像処理方法。

【請求項11】

前記変換データは、複数次元のハイパースペクトル画像データであり、

30

前記学習データは、所定の複数次元のハイパースペクトル画像データであることを特徴とする請求項10記載の画像処理方法。

【請求項12】

前記変換データは、前記複数の第1のニューラルネットワークからの画像データをクラスタリングしたクラスタリング結果データであり、

前記学習データは、前記クラスタリング結果データに対応し、前記クラスタリングを学習するための学習データであることを特徴とする請求項10記載の画像処理方法。

【請求項13】

前記クラスタリングは、物体識別、産地識別、本人識別、個人認証、又は領域分割を含むことを特徴とする請求項12記載の画像処理方法。

40

【請求項14】

前記制御手段が、前記光学フィルタの分光透過特性に対応する応答関数の関数値が0以上であって、かつ、L2ノルムのスムーズ正規化器を用いて前記光学フィルタの分光透過特性に対応する応答関数の変化が所定のスムーズ値以上のスムーズであるという制約条件にもとで、前記複数の第1のニューラルネットワークを学習するステップをさらに含むことを特徴とする請求項10～13のうちのいずれか1つに記載の画像処理方法。

【請求項15】

前記画像処理装置は、前記計算された少なくとも1つの光学フィルタの分光透過特性に対応する応答関数を有するように構成した光学フィルタを前段に設けた少なくとも1つの画像センサをさらに備え、

50

前記少なくとも1つの画像センサが、撮像する画像光を前記光学フィルタを介して受光して出力する画像データを、前記学習された第2のニューラルネットワークに入力するステップと、

前記制御手段が、前記第2のニューラルネットワークから出力されるハイパースペクトル画像データを、前記撮像する画像光のハイパースペクトル画像データとして出力するステップとを含むことを特徴とする請求項10～14のうちのいずれか1つに記載の画像処理方法。

【請求項16】

複数次元のハイパースペクトル画像データを、少なくとも1つの光学フィルタの分光透過特性に対応する応答関数を用いて、前記ハイパースペクトル画像データの次元数よりも少ない複数次元のカラー空間の画像データに変換して出力する畳み込み層をそれぞれ有する複数の第1のニューラルネットワークと、

前記複数の第1のニューラルネットワークからの画像データを、前記複数の第1のニューラルネットワークの光学フィルタの分光透過特性に対応する応答関数を学習するための変換データに変換して出力する第2のニューラルネットワークと、

所定の学習データを用いて、前記第1及び第2のニューラルネットワークを、前記学習データと前記変換データとの誤差が最小になるように学習することにより、前記複数の第1のニューラルネットワークの光学フィルタの分光透過特性に対応しかつ前記光学フィルタを形成するための応答関数を計算する制御手段とを備える画像処理装置のための画像センサ装置であって、

前記計算された少なくとも1つの光学フィルタの分光透過特性に対応する応答関数を有するように構成した光学フィルタが前段に設けられた少なくとも1つの画像センサを備え、

前記画像センサは、撮像する画像光を前記光学フィルタを介して受光して出力する画像データを、前記学習された第2のニューラルネットワークに入力し、

前記第2のニューラルネットワークから出力されるハイパースペクトル画像データを、前記撮像する画像光のハイパースペクトル画像データとして出力することを特徴とする画像センサ装置。

【請求項17】

複数次元のハイパースペクトル画像データを、少なくとも1つの光学フィルタの分光透過特性に対応する応答関数を用いて、前記ハイパースペクトル画像データの次元数よりも少ない複数次元のカラー空間の画像データに変換して出力する畳み込み層をそれぞれ有する複数の第1のニューラルネットワークと、

前記複数の第1のニューラルネットワークからの画像データを、前記複数の第1のニューラルネットワークの前記光学フィルタの分光透過特性に対応する応答関数を学習するための変換データに変換して出力する第2のニューラルネットワークと、

所定の学習データを用いて、前記第1及び第2のニューラルネットワークを、前記学習データと前記変換データとの誤差が最小になるように学習することにより、前記複数の第1のニューラルネットワークの前記光学フィルタの分光透過特性に対応しかつ前記光学フィルタを形成するための応答関数を計算する制御手段とを備える画像処理装置のための光学フィルタ装置であって、

撮像する画像光を前記光学フィルタを介して受光して出力する画像データを、前記学習された第2のニューラルネットワークに入力する少なくとも1つの画像センサの前段に設けられた光学フィルタであって、前記計算された光学フィルタの分光透過特性に対応する応答関数を有するように構成した少なくとも1つの光学フィルタを備えたことを特徴とする光学フィルタ装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、例えば撮像カメラなどの画像センサに用いられる、分光フィルタ等の光学フ

10

20

30

40

50

フィルタの分光透過特性の設計を支援する画像処理装置及び方法に関する。

【背景技術】

【0002】

ハイパースペクトルイメージング（画像化）は、波長軸に沿った詳細な光分布を捕捉し、リモートセンシング、医療診断、産業用検出などに有益であることが示されている。例えば、外科医の目に見えない腫瘍マージンは、ハイパースペクトル画像でよりよく視覚化することができる。また、漏洩した目に見えないガスの場合は、スペクトル信号を使用しても明らかである。ハイパースペクトル画像をキャプチャする既存のデバイスのほとんどは、一空間次元に沿ってラインスリットを駆動する（プッシュブローブスキャン）か、グレースケールカメラの前で狭帯域フィルタを連続的に変化させる（フィルタスキャン）のいずれかのスキャンベースである。

10

【先行技術文献】

【特許文献】

【0003】

【特許文献1】特許第5410088号公報

【特許文献2】特開2017-526910号公報

【特許文献3】特開2001-034752号公報

【非特許文献】

【0004】

【非特許文献1】B. Arad et al., "Sparse Recovery of Hyperspectral Signal from Natural RGB Images," ECCV, pp. 19-34, 2016

20

【非特許文献2】B. Arad et al., "Filter selection for hyperspectral estimation," ICCV, pp. 3172-3180, 2017

【非特許文献3】A. Chakrabarti et al., "Statistics of Real-World Hyperspectral Images," Proceedings of IEEE Conference on Computer Vision and Pattern recognition (CVPR), pp. 193-200, 2011

【非特許文献4】H. G. Chen et al., "Asp vision: Optically computing the first layer of convolutional neural networks using angle sensitive pixels," CVPR, June 2016

【非特許文献5】Y. Jia et al., "From RGB to spectrum for natural scenes via manifold-based mapping," ICCV, pp. 4715-4723, October 2017

30

【非特許文献6】J. Jiang et al., "What is the space of spectral sensitivity functions for digital color cameras," WACV, 2013

【非特許文献7】D. P. Kingma et al., "A Method for Stochastic Optimization," International Conference on Learning Representations 2015, pp. 1-15, 2015

【非特許文献8】R. M. H. Nguyen et al., "Training-based spectral reconstruction from a single RGB image," Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 8695 LNCS(PART 7): pp. 86-201, 2014

【非特許文献9】O. Ronneberger et al., "U-Net: Convolutional Networks for Biomedical Image Segmentation," Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2015

40

【非特許文献10】F. Yasuma et al., "Generalized assorted pixel camera: Post-capture control of resolution, dynamic range, and spectrum," IEEE Transactions on Image Processing, 19(9), pp. 2241-2253, September 2010

【発明の概要】

【発明が解決しようとする課題】

【0005】

これらの従来技術における重要な欠点は、スキャンが遅いため、ダイナミックシーンへの適用が妨げられることである。従って、例えば、ファイバ束及びランダムもしくは規則

50

的なアパーチャマスクを使用することにより、スキャンフリーのスナップショットハイパースペクトル装置が開発されている。しかしながら、残念ながら、これらのデバイスは空間分解能が極端に制限されているという問題点があった。

【0006】

近年、RGBからスペクトラムへの再構成の研究で示されているように、単一のRGB画像からの計算ハイパースペクトル再構成法は、上記の装置の欠点を克服するために有望である（例えば、非特許文献1, 5参照）しかしながら、既存のRGBカメラは、3チップビームスプリットプリズム技術又はシングルチップベイヤー型フィルタレイのいずれかを使用して、人間の色知覚を模倣するように設計されている（例えば、非特許文献6参照）。従って、それらのスペクトル応答関数は、コンピュータ視覚の課題、すなわちハイパースペクトル再構成に必ずしも最適ではない。

10

【0007】

ごく最近、非特許文献2では、カメラのスペクトル応答に対するハイパースペクトル再構成精度の依存性を確認したことが開示されている。この非特許文献2では、ブルートフォース探索とヒットアンドランの進化的最適化によって、有限集合の候補フィルタの中で最良のフィルタの組み合わせを見つけることが開示されている。

【0008】

例えば特許文献1では、透明基材上に多層膜を備えた透過型の分光フィルタの作成方法が開示されており、この分光フィルタを通して観察される2つの対象物の感度特性差パターンの差が大きくなるように分光フィルタを設計することが開示されている。当該特定対象の差がフィルタを通した時に大きく観察することを目的としている。従って、ハイパースペクトル画像データを画像化する場合にどのように最適化するかについては開示も示唆もない。

20

【0009】

また、例えば特許文献2では、複数のスペクトルフィルタアレイ素子（以下、SFA素子という。）をまとめて配置し、スペクトル撮像データを得るための手法が開示されている。この手法では、当該複数のSFC素子に対応する特定波長にそれぞれ中心を合わされた1つ又はそれ以上のスペクトルバンド幅をフィルタして所定のスペクトル範囲をカバーするように構成し、SFAでの記録値からデコンボリューション又は非線形スパーズ再構成法を利用してサンプリングされていないスペクトル情報を推量することが開示されている。ここで、複数のSFC素子により、スペクトル範囲をカバーするように分光透過特性が与えられおり、スペクトル復元のためのSFC分光透過特性の最適化は行われていない。

30

【0010】

さらに、特許文献3では、カラーフィルタリング処理された画像をカラー画像出力装置で再現するために、原画像の分光分布若しくは色彩値にカラーフィルタ処理する色再現装置が開示されている。この色再現装置は、原画像の色分解値の信号を対応する分光分布若しくは色彩値に変換する第1の変換手段と、前記分光分布若しくは色彩値にカラーフィルタ処理を行う第2の変換手段と、前記処理された分光分布若しくは色彩値をカラー画像出力装置で再現するための色分解値に変換する第3の変換手段とを具備したことを特徴としている。ここで、前記第1～第3の変換手段は、画像出力装置の種類毎に異なる複数のニューラルネットワークにより構成され、当該複数のニューラルネットワークは色分解値が既知の色票の分光分布若しくは色彩値と該色分解値の双方向の変換を行うように学習されることを特徴としている。すなわち、当該色再現装置は、色再現が最適化されるように、前記第1～第3の変換手段の複数のニューラルネットワークが学習されている。従って、画像センサの前段に配置するフィルタの分光透過特性をスペクトル復元又はクラスタリングのために分光透過特性を最適化することは行われていない。

40

【0011】

本発明の目的は以上の問題点を解決し、スペクトル復元又はクラスタリングのために、画像センサの前段に配置するフィルタの分光透過特性を最適化することができる画像処理装置及び方法を提供することにある。ここで、本発明は特に、学習により最適化された応

50

答関数に対応する分光透過特性を有する光学フィルタを使用して、スナップショットハイパースペクトルイメージング用のデータに基づいたマルチスペクトルカメラを構成することを目的とする。

【課題を解決するための手段】

【0012】

本発明の一態様に係る画像処理装置は、

複数次元のハイパースペクトル画像データを、少なくとも1つの光学フィルタの分光透過特性に対応する応答関数を用いて、前記ハイパースペクトル画像データの次元数より少ない複数次元のカラー空間の画像データに変換して出力する畳み込み層をそれぞれ有する複数の第1のニューラルネットワークと、

前記複数の第1のニューラルネットワークからの画像データを、前記複数の第1のニューラルネットワークの前記光学フィルタの分光透過特性に対応する応答関数を学習するための変換データに変換して出力する第2のニューラルネットワークと、

所定の学習データを用いて、前記第1及び第2のニューラルネットワークを、前記学習データと前記変換データとの誤差が最小になるように学習することにより、前記複数の第1のニューラルネットワークの前記光学フィルタの分光透過特性に対応しかつ前記光学フィルタを形成するための応答関数を計算する制御手段とを備えたことを特徴とする。

【発明の効果】

【0013】

従って、本発明に係る画像処理装置及び方法によれば、スペクトル復元のためのフィルタの分光透過特性の最適化を行うことができる。これにより、例えば撮像カメラなどの画像センサのための光学フィルタの分光透過特性の応答関数の設計を支援することができる。ここで、本発明は特に、学習により最適化された応答関数に基づく分光透過特性を有する光学フィルタを使用して、スナップショットハイパースペクトルイメージング用のデータに基づいたマルチスペクトルカメラを構成することができる。

【図面の簡単な説明】

【0014】

【図1】実施形態に係る画像処理装置の構成例を示すブロック図である。

【図2】図1のスペクトル再構成ニューラルネットワーク(SRNN)5の構成例を示すブロック図である。

【図3】変形例に係る1チップの画像処理装置のハードウェア実装部20Aとその周辺回路を示すブロック図である。

【図4A】3チャンネル3チップの場合の学習された分光透過特性の応答関数の一例を示すグラフである。

【図4B】2チャンネル2チップの場合の学習された分光透過特性の応答関数の一例を示すグラフである。

【図4C】3チャンネル1チップの場合の学習された分光透過特性の応答関数の一例を示すグラフである。

【図5】1×1の畳み込みとフィルタスペクトル応答との間の類似性を示す概略図である。

【図6A】典型的なベイヤー型フィルタの第1例(青色フィルタ)を示す正面図である。

【図6B】典型的なベイヤー型フィルタの第2例(緑色フィルタ)を示す正面図である。

【図6C】典型的なベイヤー型フィルタの第3例(赤色フィルタ)を示す正面図である。

【図7】図6A～図6Cのストライド2の2×2フィルタを用いたときの処理例を示す概略図である。

【図8】実施形態及び比較例のシミュレーション結果であって、エポック数に対する二乗平均平方根誤差(以下、RSMSEという。)の特性を示すグラフである。

【図9】実施形態及び比較例のシミュレーション結果であって、各波長におけるRSMSEを示す写真である。

【図10A】実施形態のシミュレーション結果であって、非特許文献10のCAVEデー

10

20

30

40

50

タセットを学習データとして用いたときの分光透過特性の応答関数を示すスペクトル図である。

【図10B】実施形態のシミュレーション結果であって、非特許文献3のNarvard Naturalデータセットを学習データとして用いたときの分光透過特性の応答関数を示すスペクトル図である。

【図10C】実施形態のシミュレーション結果であって、非特許文献3のNarverd Mixedデータセットを学習データとして用いたときの分光透過特性の応答関数を示すスペクトル図である。

【図11】実施例に係る学習された分光透過特性の応答関数の一例を示すスペクトル図である。

【図12】実施例に係る画像処理装置のハードウェア実装部20Bとその周辺回路を示すブロック図である。

【図13A】図12の画像センサ33によりキャプチャされた画像の一例を示す写真である。

【図13B】図12の画像センサ43によりキャプチャされた画像の一例を示す写真である。

【図14A】図12の光学フィルタ31の分光曲線（真値及び実施例）を示すスペクトル図である。

【図14B】図12の光学フィルタ41の分光曲線（真値及び実施例）を示すスペクトル図である。

【図15】図1の画像データメモリ1等に格納される画像データのデータ構造を模式的に示す図である。

【図16】変形例1に係る画像処理装置の構成例を示すブロック図である。

【図17】変形例2に係る画像処理装置の構成例を示すブロック図である。

【図18】変形例3に係る画像処理装置の構成例を示すブロック図である。

【図19】変形例2,3の実施例に係るパブリカの物体識別の画像例を示す写真画像である。

【図20】変形例2,3の実施例に係るパブリカの物体識別の別の画像例を示す写真画像である。

【図21】変形例2,3の実施例に係るいちごの物体識別の画像例を示す写真画像である。

【図22】変形例2,3の実施例に係る本人識別の画像例を示す写真画像である。

【図23A】変形例2,3の実施例1に係る領域分割の第1の学習入力画像例を示す写真画像である。

【図23B】変形例2,3の実施例2に係る領域分割の第2の学習入力画像例を示す写真画像である。

【図23C】変形例2,3の実施例3に係る領域分割の第3の学習入力画像例を示す写真画像である。

【図23D】変形例2,3の実施例4に係る領域分割の第4の学習入力画像例を示す写真画像である。

【図24A】変形例2,3の実施例1に係る領域分割の第1の学習目的画像例を示す写真画像である。

【図24B】変形例2,3の実施例2に係る領域分割の第2の学習目的画像例を示す写真画像である。

【図24C】変形例2,3の実施例3に係る領域分割の第3の学習目的画像例を示す写真画像である。

【図24D】変形例2,3の実施例4に係る領域分割の第4の学習目的画像例を示す写真画像である。

【発明を実施するための形態】

【0015】

10

20

30

40

50

(実施形態)

以下、本発明に係る実施形態について図面を参照して説明する。なお、以下の各実施形態において、同様の構成要素については同一の符号を付している。

【0016】

1. 実施形態の概要

近年、RGBイメージングからのハイパースペクトル再構成は、疎なコーディングと深い学習を経て著しい進歩を遂げている。しかしながら、ほとんど無視されている事実は、既存のRGBカメラが人間の三色知覚を模倣するように調整されているため、それらのスペクトル応答はハイパースペクトル再構成に必ずしも最適ではないということである。本発明に係る実施形態では、RGBスペクトル応答を使用するのではなく、ハードウェアで実施される最適化されたカメラスペクトル応答関数と、エンドツーエンドネットワークを使用するスペクトル再構成のためのマッピングとを同時に学習することを特徴としている。これにより、本発明者らのコアアイデアは、カメラのための光学フィルタの分光透過特性に対応する応答関数が最適化され、当該光学フィルタが畳み込み層のように効果的に作用するので、標準的なニューラルネットワークを学習することによって、それらの応答関数を最適化することができるということである。

10

【0017】

本実施形態では、空間モザイク処理を用いない3チップ構成と、ベイヤー(Bayer)形式の2×2フィルタアレイを用いた単一チップ構成を有する、2種類の設計支援されて形成されたフィルタについて説明する。実施例において、数値シミュレーションは、既存のRGBカメラと比較して深層学習されたスペクトル応答の利点を検証した。さらに興味深いことには、設計プロセスの物理的制約を考慮することによって、現代のフィルムフィルタ製造技術を使用して深層学習されたスペクトル応答機能を実現することができることを示す。また、スナップショットハイパースペクトルイメージングのためのデータに基づいたマルチスペクトルカメラを構成する。

20

【0018】

特に、最新のフィルムフィルタ製造技術により、非ネガティブかつスムーズなスペクトル応答機能を備えた画像センサを実装することができたことを示す。従って、本実施形態では、既存のフィルタからフィルタを選択するのではなく、非負関数の無限大空間における最適化されたスペクトル応答関数を直接学習する。次に、実施例では、スナップショットハイパースペクトルイメージング用のマルチスペクトルカメラを構成するためのこのデータ駆動アプローチに基づいて、学習された光学フィルタを作成した。

30

【0019】

本実施形態における特徴は、ニューラルネットワークの畳み込み層と同様に、カメラスペクトルフィルタが効果的に作用するという観察に基づいて、深層学習技術を用いてそれらを最適化することができることである。また、高分解能のエンドツーエンドネットワークを介してスペクトル再構成のためのマッピングと最適化されたフィルタ応答関数を同時に学習する。既存のRGBカメラに対するアプリケーションでは、空間モザイク処理なしの3チップセットアップと、ベイヤー形式の2×2光学フィルタアレイによるシングルチップセットアップを検討した。本発明者らは、公開されているデータセットに基づいて、数値シミュレーションを行って、既存のRGBカメラに比べて深層学習されたカメラのスペクトル応答の利点を検証した。

40

【0020】

2. 関連技術文献と本実施形態の位置づけ

以下、本発明に関連する技術文献と本実施形態の位置づけについて以下に説明する。

【0021】

走査型ハイパースペクトルカメラの速度のボトルネックを解決するために、ランダムに又は規則的に分布した光窓を有するファイバ及びアパーチャマスクを使用することにより、スキャンフリー装置が提案されている。しかし、このようなスナップショット装置の主な欠点は、限定された空間分解能にある。高解像度のグレースケール又はRGB画像デー

50

タを使用して空間分解能を向上させるための、いくつかの融合ベースの超解像度アルゴリズムも存在する。

【 0 0 2 2 】

ハイパースペクトル画像データを直接作成するのではなく、単一のRGBイメージのスペクトル分解能を高めるためのアプローチが、近年注目を集めている。ハイパースペクトル再構成における鍵は、RGB値と高次元スペクトル信号との間のマッピングを見つけることであり、このことは、明らかに不良設定問題であり、再構成のために適切な事前設定が必要となる。例えば、非特許文献8においては、ホワイトバランスアルゴリズムを介して照明効果を排除するために、ラジアル基底関数(RBF)ネットワークに基づいて、照明なしのRGB値から反射率スペクトルへのマッピングを学習することが開示されている。また、同様の問題点の解決のために、制約付きスパース符号化法を用いて代表的な辞書を学習することもなされている。さらに、例えば、非特許文献1では、自然シーンのハイパースペクトル画像に焦点を当て、スパース符号化を使用したRGBからスペクトルへのマッピング方法を開示している。上記のすべての研究は、市販のRGBカメラからの典型的な応答関数を使用して、RGB画像をシミュレーションしている。

10

【 0 0 2 3 】

最近では、非特許文献2では、ハイパースペクトル再構成の精度がフィルタ応答に依存することを認識しており、ブルートフォース検索とヒットアンドランの進化的最適化によって有限集合の候補フィルタの中で最良のフィルタの組み合わせを見つけることを試みている。

20

【 0 0 2 4 】

これらの従来技術文献に対して、本実施形態では、探索領域を非負の無限大の空間に拡張し、強力な深層学習技術を活用して、最適化されたフィルタ応答とスペクトル再構成マッピングを同時に学習することを特徴としている。ここで、最適化されたフィルタ応答のハードウェア実装は、ASPビジョン(例えば、非特許文献4参照)と並行して実行される。畳み込みニューラルネットワーク(CNN)の固定された第1の畳み込み層を直接計算してエネルギー、データ帯域幅、及びCNNのFLOPS(Floating-Point Operations Per Second)を節約するカスタムCMOS回折画像センサを使用する。しかし、ASPビジョンの場合、それらの目的は、CNN及び人間の一次視覚野のv1層に共通の予め定義されたエッジフィルタリング層をハードウェアで符号化することになる。

30

【 0 0 2 5 】

本実施形態の目的は、CNNと深層学習フレームワークを活用してカメラのフィルタ設計を最適化するように支援することにある。ここで、特に、チャンネル数の多い高次元スペクトル画像データを低次元の色空間での表現(例えばRGB3次元)の画像データに変換するためには、色空間変換のための分光感度曲線を定義する必要があるが、本実施形態では、スペクトル復元のための分光透過特性を最適化する。

【 0 0 2 6 】

3. 実施形態に係る画像処理装置の構成例

図1は実施形態に係る画像処理装置の構成例を示すブロック図である。図1において、実施形態に係る画像処理装置は、画像データメモリ1と、3個の畳み込みニューラルネットワーク(CNN)2A, 2B, 2Cと、スイッチSW1, SW2と、スペクトル再構成ニューラルネットワーク(SRNN)5と、誤差演算器6と、画像データメモリ7と、ディスプレイ8と、例えばデジタル計算機などのコンピュータで構成されるコントローラ10と、ハードウェア実装部20とを備えて構成される。本明細書において、画像データの表記において、X×YはX個の画素×Y個の画素を表し、X×Y×ZはX個の画素×Y個の画素×Z個のチャンネルを表す。

40

【 0 0 2 7 】

画像データメモリ1は、学習データである例えば256×256×31のハイパースペクトル画像データを格納する。畳み込みニューラルネットワーク(CNN)2Aは、入力

50

層 6 1 と、畳み込み層（動作層）である中間層 6 2 と、出力層 6 3 とを備えて構成される。畳み込みニューラルネットワーク（CNN）2 B は、入力層 7 1 と、畳み込み層（動作層）である中間層 7 2 と、出力層 7 3 とを備えて構成される。畳み込みニューラルネットワーク（CNN）2 C は、入力層 8 1 と、畳み込み層（動作層）である中間層 8 2 と、出力層 8 3 とを備えて構成される。各中間層 6 2 , 7 2 , 8 2 は、 $1 \times 1 \times 3 1$ でかつストライド 1 のカーネルによる畳み込み演算を行う。なお、各中間層 6 2 , 7 2 , 8 2 の複数の係数の初期値はランダム値であって、互いに異なるように設定される。また、ハードウェア実装部 2 0 は、対物レンズ 2 1 と、ビームスプリッタ 2 2 , 2 3 と、反射板 2 4 と、光学フィルタ 3 1 , 4 1 , 5 1 と、バンドパスフィルタ 3 2 , 4 2 , 5 2 と、画像センサ 3 3 , 4 3 , 5 3 とを備えて構成される。

10

【0028】

3 個の畳み込みニューラルネットワーク（CNN）2 A , 2 B , 2 C から出力される画像データは 3 次元のカラー空間の画像データである合計 $2 5 6 \times 2 5 6 \times 3$ の画像データはスイッチ SW 1 を介してスペクトル再構成ニューラルネットワーク（SRNN）5 に出力される。また、3 個の画像センサ 3 3 , 4 3 , 5 3 から出力される 3 次元のカラー空間の画像データである合計 $2 5 6 \times 2 5 6 \times 3$ の画像データはスイッチ SW 1 を介して学習済みのスペクトル再構成ニューラルネットワーク（SRNN）5 に出力される。

【0029】

図 2 は図 1 のスペクトル再構成ニューラルネットワーク（SRNN）5 の構成例を示すブロック図である。図 2 において、スペクトル再構成ニューラルネットワーク（SRNN）5 は、入力層 1 0 0 と、縦続接続された中間層 1 0 1 ~ 1 0 8 , 1 1 1 ~ 1 1 9 と、出力層 1 2 0 とを備えて構成される。

20

【0030】

図 2 において、入力層 1 0 0 は、 $2 5 6 \times 2 5 6 \times 3$ の画像データを入力した後、6 4 分岐した画像データを中間層 1 0 1 に出力する。

【0031】

(1) 中間層 1 0 1 はダウンサンプリングを行う最大プーリング処理部と 2 個の畳み込み処理部を備えて構成される。中間層 1 0 1 は、それぞれ 3×3 のサイズでストライド 3 の 6 4 個のフィルタを有し、入力される画像データに対して係数 2 のダウンサンプリング処理を行って、処理後の画像データを中間層 1 0 2 , 1 1 8 に出力する。

30

(2) 中間層 1 0 2 はダウンサンプリングを行う最大プーリング処理部と 2 個の畳み込み処理部を備えて構成される。中間層 1 0 2 は、それぞれ 3×3 のサイズでストライド 3 の 1 2 8 個のフィルタを有し、入力される画像データに対して係数 2 のダウンサンプリング処理を行って、処理後の画像データを中間層 1 0 3 , 1 1 7 に出力する。

(3) 中間層 1 0 3 はダウンサンプリングを行う最大プーリング処理部と 2 個の畳み込み処理部を備えて構成される。中間層 1 0 3 は、それぞれ 3×3 のサイズでストライド 3 の 2 5 6 個のフィルタを有し、入力される画像データに対して係数 2 のダウンサンプリング処理を行って、処理後の画像データを中間層 1 0 4 , 1 1 6 に出力する。

(4) 中間層 1 0 4 はダウンサンプリングを行う最大プーリング処理部と 2 個の畳み込み処理部を備えて構成される。中間層 1 0 4 は、それぞれ 3×3 のサイズでストライド 3 の 5 1 2 個のフィルタを有し、入力される画像データに対して係数 2 のダウンサンプリング処理を行って、処理後の画像データを中間層 1 0 5 , 1 1 5 に出力する。

40

(5) 中間層 1 0 5 はダウンサンプリングを行う最大プーリング処理部と 2 個の畳み込み処理部を備えて構成される。中間層 1 0 5 は、それぞれ 3×3 のサイズでストライド 3 の 5 1 2 個のフィルタを有し、入力される画像データに対して係数 2 のダウンサンプリング処理を行って、処理後の画像データを中間層 1 0 6 , 1 1 4 に出力する。

(6) 中間層 1 0 6 はダウンサンプリングを行う最大プーリング処理部と 2 個の畳み込み処理部を備えて構成される。中間層 1 0 6 は、それぞれ 3×3 のサイズでストライド 3 の 5 1 2 個のフィルタを有し、入力される画像データに対して係数 2 のダウンサンプリング処理を行って、処理後の画像データを中間層 1 0 7 , 1 1 3 に出力する。

50

(7) 中間層 107 はダウンサンプリングを行う最大プーリング処理部と 2 個の畳み込み処理部を備えて構成される。中間層 107 は、それぞれ 3×3 のサイズでストライド 3 の 512 個のフィルタを有し、入力される画像データに対して係数 2 のダウンサンプリング処理を行って、処理後の画像データを中間層 108, 112 に出力する。

(8) 中間層 108 はダウンサンプリングを行う最大プーリング処理部と 2 個の畳み込み処理部を備えて構成される。中間層 108 は、それぞれ 3×3 のサイズでストライド 3 の 512 個のフィルタを有し、入力される画像データに対して係数 2 のダウンサンプリング処理を行って、処理後の画像データを中間層 111, 111 に出力する。

【0032】

(9) 中間層 111 はアップサンプリング処理部と 2 個の畳み込み処理部を備えて構成される。中間層 111 は、それぞれ 3×3 のサイズでストライド 3 の 512 個のフィルタを有し、入力される 2 つの画像データを統合した画像データに対して係数 2 のアップサンプリング処理を行って、処理後の画像データを中間層 112 に出力する。

(10) 中間層 112 はアップサンプリング処理部と 2 個の畳み込み処理部を備えて構成される。中間層 112 は、それぞれ 3×3 のサイズでストライド 3 の 512 個のフィルタを有し、入力される 2 つの画像データを統合した画像データに対して係数 2 のアップサンプリング処理を行って、処理後の画像データを中間層 113 に出力する。

(11) 中間層 113 はアップサンプリング処理部と 2 個の畳み込み処理部を備えて構成される。中間層 113 は、それぞれ 3×3 のサイズでストライド 3 の 512 個のフィルタを有し、入力される 2 つの画像データを統合した画像データに対して係数 2 のアップサンプリング処理を行って、処理後の画像データを中間層 114 に出力する。

(12) 中間層 114 はアップサンプリング処理部と 2 個の畳み込み処理部を備えて構成される。中間層 114 は、それぞれ 3×3 のサイズでストライド 3 の 512 個のフィルタを有し、入力される 2 つの画像データを統合した画像データに対して係数 2 のアップサンプリング処理を行って、処理後の画像データを中間層 115 に出力する。

(13) 中間層 115 はアップサンプリング処理部と 2 個の畳み込み処理部を備えて構成される。中間層 115 は、それぞれ 3×3 のサイズでストライド 3 の 512 個のフィルタを有し、入力される 2 つの画像データを統合した画像データに対して係数 2 のアップサンプリング処理を行って、処理後の画像データを中間層 116 に出力する。

(14) 中間層 116 はアップサンプリング処理部と 2 個の畳み込み処理部を備えて構成される。中間層 116 は、それぞれ 3×3 のサイズでストライド 3 の 256 個のフィルタを有し、入力される 2 つの画像データを統合した画像データに対して係数 2 のアップサンプリング処理を行って、処理後の画像データを中間層 117 に出力する。

(15) 中間層 117 はアップサンプリング処理部と 2 個の畳み込み処理部を備えて構成される。中間層 117 は、それぞれ 3×3 のサイズでストライド 3 の 128 個のフィルタを有し、入力される 2 つの画像データを統合した画像データに対して係数 2 のアップサンプリング処理を行って、処理後の画像データを中間層 118 に出力する。

(16) 中間層 118 はアップサンプリング処理部と 2 個の畳み込み処理部を備えて構成される。中間層 118 は、それぞれ 3×3 のサイズでストライド 3 の 64 個のフィルタを有し、入力される 2 つの画像データを統合した画像データに対して係数 2 のアップサンプリング処理を行って、処理後の画像データを中間層 119 に出力する。

(17) 中間層 119 はアップサンプリング処理部と 2 個の畳み込み処理部を備えて構成される。中間層 119 は、それぞれ 3×3 のサイズでストライド 3 の 32 個のフィルタを有し、入力される 2 つの画像データを統合した画像データに対して係数 2 のアップサンプリング処理を行って、処理後の画像データを出力層 120 に出力する。

【0033】

出力層 120 は、中間層 119 から画像データを入力した後、 $256 \times 256 \times 31$ の画像データをスイッチ SW 2 の接点 a を介して誤差演算器 66 に出力し、また、スイッチ SW 2 の接点 b 及び画像データメモリ 7 を介してディスプレイ 8 に出力する。

【0034】

10

20

30

40

50

以上のように構成されたスペクトル再構成ニューラルネットワーク (SRNN) 5 は畳み込みニューラルネットワーク (CNN) 2A, 2B, 2C とともに学習され、学習後のスペクトル再構成ニューラルネットワーク (SRNN) 5 は各中間層 101 ~ 119 の係数が固定されて、 $256 \times 256 \times 3$ の画像データを、 $256 \times 256 \times 31$ の画像データに変換する画像変換回路として動作する。

【0035】

以上のように構成された画像処理装置において、画像データメモリ 1 は、学習データであるハイパースペクトル画像データを格納し、当該ハイパースペクトル画像データを畳み込みニューラルネットワーク (CNN) 2A, 2B, 2C の入力層 61, 71, 81 に出力する。コントローラ 10 は、図 1 の画像処理装置の全体を制御し、学習モードにおいて、スイッチ SW1, SW2 を接点 a 側に切り替えて、学習データであるハイパースペクトル画像データに基づいて、畳み込みニューラルネットワーク (CNN) 2A, 2B, 2C 及びスペクトル再構成ニューラルネットワーク (SRNN) 5 に対して、誤差演算器 6 からの RMSE 値が最小となるように学習させるように制御する。学習後の畳み込みニューラルネットワーク (CNN) 2A, 2B, 2C の変換関数に対応する分光曲線特性を有する光学フィルタ 31, 41, 51 を作成して、ハードウェア実装部 20 に実装する。

10

【0036】

一方、動作モードにおいては、コントローラ 10 はスイッチ SW1, SW2 を接点 b 側に切り替える。ハードウェア実装部 20 において、撮像する物体からの反射光は、対物レンズ 21、ビームスプリッタ 22 を介して光学フィルタ 31 に入射し、また、ビームスプリッタ 22 及び 23 により反射された後、光学フィルタ 41 に入射する。さらに、当該反射光はビームスプリッタ 23 を通過した後、反射板 24 により反射されて、光学フィルタ 51 に入射する。

20

【0037】

光学フィルタ 31 からの出力光は、バンドパスフィルタ 32 を介して画像センサ 33 に入射し、光学フィルタ 41 からの出力光は、バンドパスフィルタ 42 を介して画像センサ 43 に入射し、光学フィルタ 51 からの出力光は、バンドパスフィルタ 52 を介して画像センサ 53 に入射する。ここで、バンドパスフィルタ 32, 42, 52 は、ノイズ成分を除去するために、可視光のみを通過させる光学フィルタである。画像センサ 33 は入射する画像光を、電気信号である例えば $256 \times 256 \times 1$ の画像データに変換して出力する。画像センサ 43 は入射する画像光を、電気信号である例えば $256 \times 256 \times 1$ の画像データに変換して出力する。画像センサ 53 は入射する画像光を、電気信号である例えば $256 \times 256 \times 1$ の画像データに変換して出力する。

30

【0038】

学習後のスペクトル再構成ニューラルネットワーク (SRNN) 5 は各中間層 101 ~ 119 の係数が固定されて、 $256 \times 256 \times 3$ の画像データを、 $256 \times 256 \times 31$ の画像データに変換してスイッチ SW2 を介して画像データメモリ 7 に出力して格納する。画像データメモリ 7 は格納した画像データをディスプレイ 8 などの出力画像機器に出力する。

【0039】

以上のように構成された画像処理装置において、畳み込みニューラルネットワーク (CNN) 2A, 2B, 2C の個数はハードウェア実装部 20 での画像センサ 33, 43, 53 の個数に対応し、2 又は 4 以上の複数であってもよい。

40

【0040】

図 4A は 3 チャンネル 3 チップの場合の学習された分光透過特性の応答関数の一例を示すグラフである。また、図 4B は 2 チャンネル 2 チップの場合の学習された分光透過特性の応答関数の一例を示すグラフである。さらに、図 4C は 3 チャンネル 1 チップの場合の学習された分光透過特性の応答関数の一例を示すグラフである。なお、図 4C の場合において、各畳み込みニューラルネットワーク (CNN) 2A, 2B, 2C の中間層 62, 72, 82 は、例えば 2×2 のサイズであってストライド 2 の畳み込み処理を行うように構

50

成される。

【0041】

図3は変形例に係る1チップの画像処理装置のハードウェア実装部20Aとその周辺回路を示すブロック図である。図3に示すように、1チップの画像処理装置のハードウェア実装部20Aでは、画像センサ33の後段に、デモザイク処理を行うデモザイク処理回路9を設ける必要がある。

【0042】

以上説明したように、本実施形態では、学習により最適化された応答関数に対応する分光透過特性を有する光学フィルタ31, 41, 51を使用して、スナップショットハイパースペクトルイメージング用のデータに基づいたマルチスペクトルカメラを構成することができる。すなわち、製造販売前の学習モードでは、ハードウェア実装部20を備えず構成してもよい。この場合において、出力層63, 73, 83及び入力層100を設けず、中間層62, 72, 82の出力を中間層101に直接に接続してもよい。すなわち、学習モードにおける畳み込みニューラルネットワーク(CNN)2A, 2B, 2Cとスペクトル再構成ニューラルネットワーク(SRNN)5とを統合することで、1つのニューラルネットワークを構成することができる。このとき、畳み込みニューラルネットワーク(CNN)2A, 2B, 2Cは、統合されたニューラルネットワークの初段中間層である畳み込み層となる。

10

【0043】

また、製造販売後の動作モードでは、画像データメモリ1、畳み込みニューラルネットワーク(CNN)2A, 2B, 2C、及び誤差演算器6を備えず構成してもよい。後者の場合は、マルチスペクトルカメラ装置を構成できる。

20

【0044】

以上の実施形態に係る画像処理装置の特徴は以下の通りである。

【0045】

(1) カメラスペクトル応答関数とニューラルネットワークの畳み込み層との間の対応付けを行った。すなわち、カメラのスペクトル応答は、畳み込み層のハードウェア実装と見なすことができる。

(2) 畳み込み層としてのカメラ応答をシミュレーションし、スペクトル再構成ニューラルネットワークに追加することで、最適化された応答関数とハイパースペクトル再構成マッピングを同時に学習できる。

30

(3) 最適化されたフィルタ設計のための2つの設定を提案する。

図3のデモザイク処理回路9を使用しない3チップの設定と、ベイヤー形式の2×2フィルタアレイを使用した1チップの設定がある。深層学習された応答関数が、特定のコンピュータビジョンのタスク、スペクトル再構成において標準的なRGB応答より優れていることを詳細後述するように実証できた。

(4) 干渉フィルム製造技術を用いて、深層学習された分光曲線の変換応答関数に対応する分光透過特性を有する光学フィルタを形成し、スナップショットハイパースペクトル画像システムを構成することができる。

40

【0046】

4. 光学フィルタの設計及びアプリケーション

学習モードでは、スペクトル再構成ニューラルネットワーク(SRNN)5に、畳み込みニューラルネットワーク(CNN)2A, 2B, 2Cの畳み込み層を付加してエンドツーエンドネットワークを構成する。ここで、非負性及び滑らかさの制約を適切に組み込むことで、畳み込み層は効果的に、設計しようとするフィルタスペクトル応答関数として作用する。このことは、入力されるハイパースペクトル画像データを、最も適切な隠れた特徴マップ(マルチスペクトル画像)に符号化するので、後続のスペクトル再構成ニューラルネットワーク(SRNN)5は元の入力ハイパースペクトル画像データを可能な限り忠実に復元することができる。この意味で、エンドツーエンドのネットワークは自動エンコーダ/デコーダを構成する。

50

【0047】

動作モードでは、フィルム形式の光学フィルタの製造技術を用いて、深層学習された光学フィルタの応答機能を物理的に実現する。マルチチップ構成では、このカメラの出力が、学習された隠れた特徴マップに十分に近づくように、マルチスペクトルカメラを容易に構成することができる。また、前述のように、学習したフィルタアレイをシングルチップ構成で実現することもできる。

【0048】

動作モードにおいて、スペクトル再構成ニューラルネットワーク(SRNN)5を公知のオンラインアプリケーションで用いることもでき、この場合は、カスタマイズされたマルチスペクトルカメラを使用して画像をキャプチャし、既に学習されたスペクトル再構成ニューラルネットワーク(SRNN)5にオンラインで、直接インポートしてハイパースペクトル画像を生成することができる。

10

【0049】

光学フィルタの選択が必要とされる場合(非特許文献2)と比較して、本実施形態に係る画像処理装置は、より良い光学フィルタのために探索領域を拡張するだけでなく、再構成時間を節約することができ、オンラインでも適用可能である。スパースコードを計算する必要はないからである。また、再構成とは対照的に、設計した光学フィルタは、実際にはハイパースペクトル画像の保存スペースを節約するために原理的な損失圧縮方式を提供できる。

20

【0050】

5. 光学フィルタ設計とスペクトル復元

以下、同時のフィルタ応答設計とスペクトル再構成のためのエンドツーエンドネットワークの詳細について説明する。

20

【0051】

5.1 スペクトル再構成ネットワーク

本実施形態では、スペクトル再構成のために任意のエンドツーエンドネットワークを使用できる。スペクトル再構成ニューラルネットワーク(SRNN)5は、例えば、既知のU-net(例えば、非特許文献9参照)のアーキテクチャを適応させてスペクトル復元ネットワークを構成することができる。

30

【0052】

これらの動作層に情報を渡すことは、必然的に高解像度入力グリッドの低レベルの細部の多くを犠牲にすることになる。従って、画像対画像アプリケーションでは、スキップ接続構造によって低レベルの情報を層間で直接共有することができる。基本的に、スキップ接続は、より深い層に情報が到達することを可能にする。この構造は、モデルが非常に深い深層学習されるモデルである場合に、消滅/爆発勾配の問題を緩和することができる。さらに、U-netは小さなサイズの学習データセットでもうまく機能し、これは既存のハイパースペクトルデータセットの規模が依然として限定されているため、アプリケーションに特に適している。

40

【0053】

ところで、図1の畳み込みニューラルネットワーク(CNN)2A, 2B, 2C内の中間層62, 72, 82は例えば、2次元畳み込みBatchNorm-ReLU層を用いる。畳み込みニューラルネットワーク(CNN)2A, 2B, 2C及びスペクトル再構成ニューラルネットワーク(SRNN)5からなるニューラルネットワークは、サイズ256×256×3の画像データを入力とし、最終的にサイズ256×256×31の対応するスペクトル画像データを生成する。図2の中間層101~108のC_kは、複数k個のフィルタを有する1つの畳み込み層、1つの漏れReLU活性化層、1つのバッチ正規化層を含む畳み込みブロックを示す。各中間層101~108のC_kの畳み込み層は、ストライド2の33個のサイズのカーネルを有する。ここで、ダウンサンプリング係数は2で、エッジに適切なゼロパディングがある。漏れReLU活性化層のパラメータは0.2に設定されている。また、図2の中間層111~119のCD_kは、畳み込み層が逆畳み込

50

み層に置き換えられている点を除いて、Ckと同じブロックを示す。入力画像データに対するアップサンプリングも同様に2倍になる。各ブロックの後に50%のドロップアウト率を持つドロップアウト層が追加される。アーキテクチャ全体は図2に示すように構成される。標準的なU-netと比較して、U-netの最後の層を3チャンネルから31チャンネルに変更し、ロス関数をクロスエントロピーから平均二乗誤差(MSE)に変更する。

【0054】

5.2 フィルタスペクトル応答の設計

図1に示すように、本実施形態の1つの重要な新規性は、カメラカラーイメージングの定式化と畳み込み層との間の対応関係を描くことにある。これにより、既存のネットワーク学習アルゴリズムとツールを使用してスペクトルイメージングパラメータを最適化することができる。簡略化のために、CCD又はCMOS画像センサは理想的なフラットレスポンスを一時的に有すると仮定し、実際のシステムを構成するときはこの要素に対処する。

10

【0055】

位置(x, y)におけるスペクトル放射輝度L(x, y)が与えられると、カラーフィルタと結合された線形画像センサによる記録強度は、以下の式によって与えられる。

【0056】

【数1】

$$I_c(x, y) = \int_{\lambda} S_c(\lambda) L(x, y, \lambda) d\lambda \quad (1)$$

20

【0057】

ここで、 λ は波長であり、 $S_c(\lambda)$ はカラー光学フィルタのスペクトル応答関数である。ほとんどの商用カメラでは、赤-緑-青の三色フィルタがあり、すなわち、人間の色知覚を模倣するように(c ∈ {R, G, B})実現される。実際には、式(1)は次式のように離散的に近似することができる。

【0058】

【数2】

$$I_c(x, y) = \sum_{n=1}^N S_c(\lambda_n) L(x, y, \lambda_n) \quad (2)$$

30

【0059】

フィルタスペクトル応答関数 $S_c(\lambda)$ は、サンプリングされた波長でのベクトル $S_c = [S_c(\lambda_1), S_c(\lambda_2), \dots, S_c(\lambda_N)]$ の形式であり、Nはスペクトルチャンネルの数(波長数)である。興味深いのは、式(2)は、順方向伝搬における 1×1 の畳み込みカーネルの畳み込み演算と同一であることである。

【0060】

フィルタスペクトル応答関数 S_c を 1×1 の畳み込みカーネルの重みとして考えると、 1×1 の畳み込みとフィルタスペクトル応答との間の類似性を示す概略図を示す図5に示すように、強度 $I_c(x, y)$ は畳み込みの出力活性化マップとして解釈することができ、これは実際に畳み込みカーネル(カラーフィルタ)と入力(入射光) $L(x, y)$ のエントリ間のドット積の演算に対応する。

40

【0061】

この観察に従って、本実施形態では、図1に示すように、3チャンネルカメラで3つのカラーフィルタのように機能する3つの畳み込みカーネルを持つ 1×1 の畳み込み層(畳み込みニューラルネットワーク(CNN)2A, 2B, 2Cの中間層62, 72, 82)を追加した。追加した畳み込み層では、Nチャンネルのハイバースペクトル画像データを入力としてこのエンドツーエンドネットワークを学習する。この方法により、 1×1 の畳み込みカーネルの学習された重み係数から最適化されたスペクトル応答を得ることができ

50

る。

【0062】

5.2.1 モザイク処理なしのマルチチップ設定

市販のRGBカメラの中には、マルチチップの設定を採用しているものもある。すなわち、CCD又はCMOS画像センサごとに、特殊な三色プリズムアセンブリを使用する別個のカラーフィルタを有することが望ましい。また、空間的なモザイク処理がなければ、シングルチップ構成のベイヤー形式のフィルタアレイアセンブリよりも色精度と画像ノイズが優れている。1つの選択肢は、図1に示すように、ビームスプリッタ22、23及び反射板24とカラーの光学フィルタ31、41、51と一緒に組み合わせることであり、これはマルチチャンネルカメラプロトタイプを構成するのに適している。図1のマルチチップ構成では、上述のようにフィルタスペクトル応答関数を直接得ることができることは明らかである。

10

【0063】

5.2.2 2×2のフィルタアレイを使用したシングルチップセットアップ(図3の変形例)

市販のRGBカメラの大部分は、内部にCCD又はCMOS画像センサが1つあり、2×2のベイヤー形式のカラーフィルタアレイを使用して、空間モザイク処理を使用してRGBイメージをキャプチャする。フル解像度のRGB画像を得るためには、モザイク除去法が必要である。本実施形態では、この単一チップの変形例(図3)にも拡張することができる。ベイヤー配列の空間的構成に基づいて、3つの独立したチャンネルを持つ2×2のフィルタアレイを検討し、エンドツーエンドのネットワークを介してスペクトル応答関数を設計する。

20

【0064】

図6Aは典型的なベイヤー型フィルタの第1例(青色フィルタ)を示す正面図である。図6Bは典型的なベイヤー型フィルタの第2例(緑色フィルタ)を示す正面図である。図6Cは典型的なベイヤー型フィルタの第3例(赤色フィルタ)を示す正面図である。図7は図6A~図6Cのストライド2の2×2フィルタを用いたときの処理例を示す概略図である。

【0065】

図6A~図6Cに示すように、ベイヤー配列のパターンでは、各2×2セルには、青色画素が1つ、赤色画素が1つ、緑色画素が2つしかない。図7に示すように、ストライド2の2×2の畳み込みカーネルを使って直接シミュレートすることができる。これは、2×2の畳み込みカーネルを特定の位置の1×1の畳み込み処理に変換する。実施形態での実装では、赤色チャンネル及び青色チャンネルのために、畳み込みフィルタの重みの75%を手動でゼロに固定する。緑色チャンネルの場合、重みの半分をゼロに固定するだけである。ベイヤーパターンは、同じスペクトル応答関数を共有するために2つの緑色フィルタを必要とするので、共有スペクトル応答関数を畳み込みカーネルの平均対角重みで近似する。

30

【0066】

5.3 非ネガティブでかるスムーズな応答の拘束条件

物理的制約により、フィルタ応答関数は非負である必要がある。また、既存のフィルムフィルタ製造技術では、滑らかな(スムーズな)応答曲線を高精度にしか実現することができない。従って、数値設計プロセスではこれらの制約を考慮する必要がある。

40

【0067】

一般的な畳み込みニューラルネットワークには様々な正規化器があり、学習中に層パラメータをペナルティするように設定できる。これにより、スペクトル応答関数の非負性及び滑らかさの制約条件の実現は、それらの正規化器を利用することによって容易に実施することができる。

【0068】

非負のフィルタ応答を実現するために、本発明者らは、フィルタ設計のための畳み込み

50

層のカーネル（図1の畳み込みニューラルネットワーク（CNN）2A, 2B, 2Cの中間層62, 72, 82に対応する）において非負正則化器を適用して、フィルタスペクトル応答関数 $s_c(\lambda)$ が0以上になるようにする。また、滑らかさ（スムーズ）の制約条件に関しては、L2ノルム正則化器を使用する。これは、一般的に、深層学習ネットワークの学習でオーバーフィッティングを避けるために使用される。具体的には、平滑化を制御する次式の正則化項を導入する。ここで、実験を通して、0.02に設定されている。

【0069】

【数3】

$$\eta \sqrt{\sum_{n=1}^N (s_c(\lambda_n))^2}$$

10

【0070】

本実施形態では、図1のコントローラ10は、各フィルタの分光透過特性に対応する応答関数の関数値が0以上であって、かつ、L2ノルムのスムーズ正則化器を用いて前記各フィルタの分光透過特性に対応する応答関数の変化が所定のスムーズ値以上のスムーズであるという制約条件にもとで、畳み込みニューラルネットワーク（CNN）2A, 2B, 2Cを学習する。

【実施例】

20

【0071】

6. 合成データを用いた実験結果

ここでは、本実施形態に係る手法の有効性を示すために、合成データの実験を行う。本発明者らは、自然シーンと屋内シーンの両方からなるデータセットについて本手法を評価する（例えば、非特許文献10, 3参照）。

【0072】

6.1 学習データとパラメータ設定

いわゆる「CAVE」データセット（非特許文献10）は、10nmステップで400nmから700nmまでの31のチャンネルを持つ一般的な屋内シーンのハイパースペクトルデータセットである。各バンドは、サイズ512×512の16ビットグレースケール画像である。いわゆる「Harvard」データセット（非特許文献3）は、屋外と屋内両方のシナリオを含む実世界のハイパースペクトルデータセットである。画像データは、420nmから720nmまで10nmステップで取り込まれる。誤差を最小にするために、「Harvard Natural」データセットという自然照明下で50枚の画像にラベルを付ける。また、混合又は人工照明下の27枚の画像の残りの部分を「Harvard Mixed」データセットと呼ぶ。

30

【0073】

学習段階では、学習画像データから256×256の入力パッチデータをランダムにクロッピングすることでランダムジッタを適用する。本発明者らは、エポック数毎にバッチサイズ2及び50の繰り返し数でニューラルネットワーク（アルゴリズム）2A~2C, 5を学習した。初期学習率0.002、係数 $\lambda_1 = 0.5$ 、 $\lambda_2 = 0.999$ のアダム最適化器（例えば、非特許文献7参照）を用いてネットワークを学習した。すべての重み係数は、平均0及び標準偏差0.02のガウス分布から初期化された。

40

【0074】

提案されたニューラルネットワーク（アルゴリズム）2A~2Cは、NVIDIA GTX 1080 GPUで構成されたコントローラ10により実行された。ニューラルネットワーク5を構成するサーバー装置には、Intel(R) Core(TM) i7-6800K CPU@3.40GHzと128GBのメモリが搭載されている。CAVEデータセット（非特許文献10）、Harvard Natural and Mixedデータセット（非特許文献9）の学習時間はそれぞれ1.84時間、8.88時間、8.

50

5 2 時間であった。個々の画像データからフィルタ応答関数の分光曲線スペクトルを再構成する平均時間は約 5 . 8 3 秒であった。

【 0 0 7 5 】

実験を通して、本発明者らは評価指標として二乗平均平方根誤差 (R M S E) を選択し、図 1 に示すように、R M S E の誤差演算器 6 を用いた。各データセットについて、すべてのテストデータのハイパースペクトル画像データを再構成し、再構成されたハイパースペクトル画像データと、真値との間の R M S E の平均及び分散を計算した。一貫性のために、本発明者らはすべての分光曲線スペクトルを [0 , 2 5 5] の範囲に再スケールした。

【 0 0 7 6 】

5 . 2 3 チャンネルマルチチップ設定の結果

まず、前述したマルチチップの設定を評価する。ここでは、図 1 の 3 つの画像センサ 3 3 , 4 3 , 5 3 を使用してマルチチップ・セットアップのパフォーマンスを評価した。C A V E データセット (非特許文献 1 0) のための最適スペクトル応答関数を図 4 A ~ 図 4 C に示す。

【 0 0 7 7 】

図 4 A は 3 チャンネル 3 チップの場合の学習された分光透過特性の応答関数の一例を示すグラフである。また、図 4 B は 2 チャンネル 2 チップの場合の学習された分光透過特性の応答関数の一例を示すグラフである。さらに、図 4 C は 3 チャンネル 1 チップの場合の学習された分光透過特性の応答関数の一例を示すグラフである。なお、図 4 C の場合において、各畳み込みニューラルネットワーク (C N N) 2 A , 2 B , 2 C の中間層 6 2 , 7 2 , 8 2 は、例えば 2 x 2 のサイズであってストライド 2 の畳み込み処理を行うように構成される。

【 0 0 7 8 】

次いで、R M S E の平均と分散を表 1 に示し、3 つのベースラインの評価方法 (例えば、非特許文献 1 , 8 , 5 参照) と比較した。3 つのベースラインの評価方法の R G B 入力データは、C a n n o n 6 0 0 D のスペクトル応答関数から生成した。この表は、本発明者らの方法の R M S E が、3 つのデータセットのスペクトル再構成における代替方法より優れていることを示している。

【 0 0 7 9 】

[表 1]

ハイパースペクトルデータベースの再構成の R M S E の平均値と分散値 (非特許文献 1 0 , 5 , 1 0)

	C A V E (非特許文献 5)	H a r v a r d (非特許文献 3)	N a t u r a l	M i x e d (非特許文献 3)
実施形態	4 . 4 8 ± 2 . 9 7	7 . 5 7 ± 4 . 5 9	8 . 8 8 ± 4 . 2 5	
非特許文献 1	8 . 8 4 ± 7 . 2 3	1 4 . 8 9 ± 1 3 . 2 3	9 . 7 4 ± 7 . 4 5	
非特許文献 8	1 4 . 9 1 ± 1 1 . 0 9	9 . 0 6 ± 9 . 6 9	1 5 . 6 1 ± 8 . 7 6	
非特許文献 5	7 . 9 2 ± 3 . 3 3	8 . 7 2 ± 7 . 4 0	9 . 5 0 ± 6 . 3 2	

【 0 0 8 0 】

図 9 は実施形態及び比較例 1 (非特許文献 8) 及び比較例 2 (非特許文献 1) のシミュレーション結果であって、C A V E データセット (非特許文献 1 0) からの各波長におけるサンプル結果の R M S E を示す写真である。図 9 に示すように、本発明者らは、7 つの異なる波長で画像を示す C A V E データセットからの回収されたハイパースペクトル画像の空間的一貫性を実証した。

【 0 0 8 1 】

図 1 0 A は実施形態のシミュレーション結果であって、非特許文献 1 0 の C A V E デー

10

20

30

40

50

タセットを学習データとして用いたときの分光透過特性の応答関数を示すスペクトル図である。図10Bは実施形態のシミュレーション結果であって、非特許文献3のNarvard Naturalデータセットを学習データとして用いたときの分光透過特性の応答関数を示すスペクトル図である。図10Cは実施形態のシミュレーション結果であって、非特許文献3のNarverd Mixedデータセットを学習データとして用いたときの分光透過特性の応答関数を示すスペクトル図である。図10A～図10Cに示すように、3つのデータセットから得られたランダムな点のスペクトルを再現して示している。

【0082】

本発明者らのスペクトル応答関数の有効性を実証するために、既存のタイプのカメラによって生成されたRGB画像データに関するベクトル再構成ネットワークを学習し、テストを行った。

10

【0083】

図8は実施形態及び比較例のシミュレーション結果であって、エポック数に対する二乗平均平方根誤差（以下、RSM Eという。）の特性を示すグラフである。図8では、CAVEデータセット（非特許文献10）における本発明者らの設計された既存のスペクトル応答関数の各エポックのRSM Eを示しており、図8の各学習エポックのテストセットの平均RSM Eを比較した。

【0084】

図8に示すように、本発明者らの方法の再構成誤差は、既存のカメラタイプに基づく他のスペクトル再構成ネットワークと比較して、エポックが増加するにつれて急速に収束する。本発明者らの方法はまた、エポック数60で優れた性能を示す。

20

【0085】

5.3 シングルチップ設定用フィルタアレイ設計

本発明者らは、フィルタアレイの設計（セクション5.2.2）での性能を実証した。代替案と比較すると、入力画像データをベイヤー形式のパターンで符号化することによってシングルチップのデジタルカメラをシミュレーションした。その後、比較を行う前にベイヤー形式の符号化画像データをカラー画像データに変換するための標準的なデモザイク法である勾配補正線形補間を実行した。

【0086】

[表2]

CAVEデータセット（非特許文献10）に基づいて
フィルタアレイを用いた再構成のRSM Eの平均値と分散値

30

実施形態	非特許文献1	非特許文献8
4.73 ± 3.12	13.25 ± 13.88	18.13 ± 9.33

【0087】

表2においては、CAVEデータセットの3チャンネルシングルチップ設定の定量分析を示す。最適なスペクトル応答関数は1で与えられ、各スペクトル応答関数の対応する位置が図7に示されている。ベイヤー形式のパターンと同様に、緑色で着色されたスペクトル応答は、チップの50%をカバーすることに留意されたい。本発明者らの方法は、単一チップ設定におけるモザイク除去プロセスの下で既存の方法の性能が低下するアレイ設定の下で十分な精度を維持する。

40

【0088】

6. 学習データに基づいて生成されたマルチスペクトルカメラ

本実施形態では、画像キャプチャとハイパースペクトル再構成のためのマルチスペクトルカメラの構成を目指している。本発明者らは、FLIR製GS3-U3-15S5M型カメラを図12の画像センサ33, 43として使用して、300nmから1100nmまでのスペクトル範囲の光を集める画像を取り込んだ。紫外線（UV）と近赤外線（NIR

50

)の感度を遮断するために、可視光のみを通過するバンドパスフィルタ32, 42をそれぞれ、各光学フィルタ31, 41と、各画像センサ33, 43のカメラレンズとの間に追加した。マルチ画像センサの設定はフィルタアレイより簡単に実装できるので、セクション6.2に示すように設計作業を行った。

【0089】

CAVEデータセット(非特許文献10)で評価すると、2チャンネル最適化フィルタの平均RMSEは5.76であり、3チャンネル設定4.48よりもわずかに高い。両方の結果は、3チャンネル入力に基づく代替アルゴリズムよりもずっと優れていることに注意されたい。

【0090】

図11は実施例に係る学習された分光透過特性の応答関数の一例を示すスペクトル図である。図12は実施例に係る画像処理装置のハードウェア実装部20Bとその周辺回路を示すブロック図である。図12において、ハードウェア実装部20Bは上述のように、2つの画像センサ33, 43で構成される。

【0091】

図12の光学フィルタ31, 41のカスタマイズには高コストなため、ここでは、図11に応答関数を示す2つのチャンネルの場合に設計されたフィルタを実現することを選択した。本発明者らは、設計されたフィルタ応答関数を正確に実装するために、ある有力な光学企業に協力を求めて実現されたフィルム光学フィルタ31, 41は、50mm×50mm×1mmのサイズを有する。測定されたスペクトル応答関数を図11に示す(実線は設計された応答を示し、円は実際に測定された応答を示す)。フィルム光学フィルタ31, 41は、複数の薄いSiO₂層及びNb₂O₅層からなる干渉フィルタである。薄層境界での入射光と反射光との間の干渉効果により、設計されたフィルム光学フィルタ31, 41は、本発明者らの設計に非常に近いスペクトル応答関数を与える。本発明者らは50-50のビームスプリッタ22を使用して同軸バイスペクトルカメラを備えたハードウェア実装部20B(図12)を構成し、図12に示すように、2台のFLIR製GS3-U3-15S5M型カメラを画像センサ33, 43として適切に位置合わせした。

【0092】

図13Aは図12の画像センサ33によりキャプチャされた画像の一例を示す写真である。図13Bは図12の画像センサ43によりキャプチャされた画像の一例を示す写真である。図14Aは図12の光学フィルタ31の分光曲線(真値及び実施例)を示すスペクトル図である。図14Bは図12の光学フィルタ41の分光曲線(真値及び実施例)を示すスペクトル図である。

【0093】

図13A及び図13Bに、2つの光学フィルタ31, 41で取り込まれたサンプル画像を示す。本発明者らはまた、真値と比較して、再構成された分光曲線スペクトルを図14A及び図14Bに示す。図14A及び図14Bから明らかなように、以前のシミュレーション結果と一致して、再構成はかなり正確である。

【0094】

8. 結論

以上説明したように、本実施形態によれば、深層学習手法を用いて非負の無限大空間におけるフィルタ応答関数を学習する方法を示した。U-netベースの再構成ネットワークに特殊な畳み込み層を追加し、3つの独立した光学フィルタと、ベイヤー型2×2のフィルタアレイの形式で、標準のRGB応答より優れた応答関数を精度良く発見して実現した。実際のマルチスペクトルカメラを構成するために、本発明者らはまた、CCDカメラの応答を設計プロセスに組み込んだ。本発明者らは、2つの光学フィルタをうまく設計/実装し、スナップショットハイパースペクトル画像データに基づいたバイスペクトルカメラを構成した。

【0095】

上述のように、適切なデータセットが与えられれば、深く学んだ反応は、人間の目の色

10

20

30

40

50

合わせ機能に最終的に収束するはずであると推定した。なぜなら、後者は進化の長い歴史の中で「最適化」されていたからである。しかし、現在の実験では、学習されたフィルタ応答関数は、学習データセットごとに大きく異なる可能性があることがわかった。

【0096】

他の変形例

以上の実施形態では、ニューラルネットワークを、入力層に入力する画像データと、出力層から出力する画像データとが同一のハイパースペクトル画像データを用いて学習しているが、本発明はこれに限らず、後者を識別結果の画像データ（各カラーチャンネルが一つの識別対象の抽出結果を示す）としてもよく、この場合、図1の装置において識別に最適化したフィルタの分光透過特性に対応する応答関数を得ることができる。

10

【0097】

また、以上の実施形態では、単一チップ構成を有するフィルタアレイにおける各チャンネルに対応したフィルタの配置をベイヤー形式として説明したが、各チャンネルに対応したフィルタを単一チップに対応するよう適切に配置できればベイヤー形式に限定されず、例えば非周期的な配列など、所定の画素数のブロックを一単位とした任意の配置であってよい。

【0098】

また、配置も最適化するように機械学習させる場合であっても、本発明は同様に適用することができる。

【0099】

なお、学習時及び復元時においても、画像センサ33, 43, 53の分光感度特性を考慮して学習又は復元している。

20

【0100】

(実施形態の補足)

図15は図1の画像データメモリ1に格納される画像データのデータ構造を模式的に示す図である。図15の全面の四角は、画像表面の各画素に対応し、複数の画素が幅w方向（水平方向）及び高さh方向（垂直方向）に並置されている。ここで、幅w方向の位置と、高さh方向の位置で、画素位置が決定される。そして、図15の奥行c方向は複数のチャンネル（複数の波長に対応する）に対応する。従って、各画素は以下のデータフォーマットを有する。

30

【0101】

(w, h, channel 1, channel 2, ..., channel 31)

【0102】

ここで、例えば256×256×31の画像データのとき、各値は、w、h、各channelとも、0～255をとる。なお、画像データとして、圧縮形式や付随情報（撮影日時等）もある場合もあるが、本実施形態としては、圧縮なし、付随情報なし、上記のようなデータが得られるように画像センサ33, 43, 53を設定して、上記データフォーマットのデータを「画像データ」として画像データメモリ1に格納する。

【0103】

以下の変形例1～4においては、クラスタリング処理を実行する画像処理装置について以下に説明する。ここで、当該クラスタリングは、物体識別、産地識別、本人識別、個人認証、又は領域分割を含む。

40

【0104】

(変形例1)

図16は変形例1に係る画像処理装置の構成例を示すブロック図である。図16の画像処理装置は、図1の画像処理装置と比較して、以下の点異なる。

(1) 3個の光学フィルタ31, 41, 51を単一の光学フィルタ31Aで構成し、対物レンズ21の前段に設けた。この場合のハードウェア実装部に符号20Cを付す。

(2) 単一の光学フィルタ31Aのみで構成したので、当該光学フィルタ31Aに対応する1個の畳み込みニューラルネットワーク(CNN)2Aのみを備える。

50

【 0 1 0 5 】

以上のように構成された画像処理装置では、例えば3CCDを用いて、単一の光学フィルタ31Aのみで分光画像を生成することができる。また、形成された光学フィルタ31Aは当該処理に対して最適化された好適な光学フィルタとなる。

【 0 1 0 6 】

なお、ペイヤー型CCDなどの場合において、1枚のCCDに各色のCCDが配置された構成の場合、図1及び図16におけるビームスプリッタ22、23、反射板24を削除することができる。これについては、変形例3も同様である。

【 0 1 0 7 】

(変形例2)

図17は変形例2に係る、例えば物体識別等のクラスタリングのための画像処理装置の構成例を示すブロック図である。図17の画像処理装置は、図1の画像処理装置に比較して以下の点が異なる。

(1) 画像データメモリ1に代えて、学習入力画像データを格納する画像データメモリ1Aを備える。

(2) 学習目的画像データを格納する画像データメモリ1Bをさらに備え、学習目的画像データを誤差演算器6に入力する。すなわち、学習モード時において、これらの学習入力画像データ及び学習目的画像データを用いて、畳み込みニューラルネットワーク(CNN)2A~2C及びスペクトル再構成ニューラルネットワーク(SRNN)5を学習させる。

(3) 画像データメモリ7に代えて、画像変換テーブルメモリ7mを有する画像変換処理部7Aを備える。動作モード時において、スペクトル再構成ニューラルネットワーク(SRNN)5から出力されるデータに対して、例えば物体識別結果を色分けするなどの画像変換処理を行ってディスプレイ8に出力して表示する。

【 0 1 0 8 】

変形例2において、画像データメモリ1A、1Bに格納される画像データはセットとして扱われる。このため、下記のデータフォーマットに加えて、各画像データセットに対して通し番号を付与する等が必要となる。これは、コントローラ10での制御用に使用される。ただし、学習データとしてNN2A~2C、5に入力させるデータは下記のデータフォーマットとなる。当該変形例では、例えば対象を2つに識別する場合。すなわち、Real/Fake識別などのクラスタリング処理を実行できる。

【 0 1 0 9 】

画像データメモリ1Aに格納される画像データのデータフォーマットは実施形態と同様であるが、画像データメモリ1Bに格納されるデータフォーマットは例えば、以下によって与えられる。

【 0 1 1 0 】

(w, h, class1, class2)

【 0 1 1 1 】

ここで、各値は、w, hが0~255であり、class1, class2は、0又は1であり、例えば、Realに対応する画素部分を1、Fakeに対応する画素部分を0とする。また、対象を3つ以上に識別する場合の画像データメモリ1Bに格納される画像データのデータフォーマットは以下によって与えられる。

【 0 1 1 2 】

(w, h, class1, class2, class3, ...)

【 0 1 1 3 】

ここで、各値は上記と同様である。

【 0 1 1 4 】

以上のように構成された画像処理装置では、例えば物体識別等のクラスタリングを実行してその結果を表示できる。

【 0 1 1 5 】

10

20

30

40

50

なお、上記の変形例 2 及び以下の変形例 3 において、画像データメモリ 1 B に格納される学習目的画像データとしては、例えば赤と青等の各色で実施例として説明され、前記の `class 1` , `class 2` に、対応する 0 又は 1 の値が挿入される。後述する実施例において、赤と青で示しているのは、説明用に分かりやすくするためであって可視化して示しており、実際に当該画像データメモリ 1 B に格納するデータとは異なる。このことは、後述する領域分割についても同様である。

【0116】

(変形例 3)

図 18 は変形例 3 に係る画像処理装置の構成例を示すブロック図である。図 18 の画像処理装置は、図 17 の画像処理装置に比較して、以下の点が異なる。

(1) 3 個の光学フィルタ 31 , 41 , 51 を単一の光学フィルタ 31 A で構成し、対物レンズ 21 の前段に設けた。この場合のハードウェア実装部に符号 20 C を付す。

(2) 単一の光学フィルタ 31 A のみで構成したので、当該光学フィルタ 31 A に対応する 1 個の畳み込みニューラルネットワーク (CNN) 2 A のみを備える。

【0117】

以上のように構成された画像処理装置では、例えば 3 C C D を用いて、単一の光学フィルタ 31 A のみでクラスタリング処理を行うことができる。また、形成された光学フィルタ 31 A は当該処理に対して最適化された好適な光学フィルタとなる。

【0118】

(変形例 2 , 3 の実施例)

図 19 は変形例 2 , 3 の実施例 1 に係るパブリカの物体識別の画像例を示す写真画像である。図 19 の左側のパブリカは `Real` (真) のもので、右側のパブリカは `Fake` (偽) のものであり、画像データメモリ 1 A に格納するとききは、左右分割して格納される。ここで、`Real` / `Fake` 識別を赤色 / 青色に対応させて物体識別 (産地識別等を含む) を行うことができる。図 17 及び図 18 の画像データメモリ 1 A には図 19 の各画像を格納する。一方、画像データメモリ 1 B には以下の画像データが格納される。

(A) 左側のパブリカに対応する画素 (画像中の領域) を赤で塗りつぶした画像データ。

(B) 右側のパブリカに対応する画素 (画像中の領域) を青で塗りつぶした画像データ (なお、背景は黒色とする)。

【0119】

なお、以下の変形例 2 , 3 の実施例 1 ~ 5 において、画像データの各色は例示であって、本発明で限定されるものではない。

【0120】

画像変換処理部 7 A の画像変換テーブルメモリ 7 m に格納される画像変換テーブルは、赤の画素については、`Real` (真) に変換する一方、青の画素については、`Fake` (偽) に変換する。これにより、ディスプレイ 8 には、物体識別結果を示す画像内で、`Real` (真) 又は `Fake` (偽) に変換されて表示される。

【0121】

図 20 は変形例 2 , 3 の実施例 2 に係るパブリカの物体識別の別の画像例を示す写真画像である。この画像例では、`Real` / `Fake` 識別を赤色 / 青色に対応させる例を示す。画像データメモリ 1 A に格納される画像を図 20 に示す。画像データメモリ 1 B には以下の画像が格納される。

(A) 上段の 3 つのパブリカそれぞれに対応する画素 (画像中の領域) を赤で塗りつぶし他画像データ。

(B) 下段の 3 つのパブリカそれぞれに対応する画素 (画像中の領域) を青で塗りつぶした画像データ (なお、背景は黒とする)。

【0122】

画像変換処理部 7 A の画像変換テーブルメモリ 7 m に格納される画像変換テーブルは、赤の画素については、`Real` (真) に変換する一方、青の画素については、`Fake` () に変換する。これにより、ディスプレイ 8 には、物体識別結果を示す画像内で、`Real`

10

20

30

40

50

1 (真)又はFake (偽)に変換されて表示される。

【0123】

図21は変形例2, 3の実施例3に係るいちごの物体識別の画像例を示す写真画像である。図21の画像例は、Real/Fake識別を赤色/青色に対応させる例を示す。画像データメモリ1Aには図21の画像の画像データが格納される。また、画像データメモリ1Bには以下の画像データが格納される。

(A)左側の3つのイチゴそれぞれに対応する画素(画像中の領域)を赤で塗りつぶした画像データ。

(B)右側の3つのイチゴそれぞれに対応する画素(画像中の領域)を青で塗りつぶした画像データ(なお、背景は黒とする)。

【0124】

画像変換処理部7Aの画像変換テーブルメモリ7mに格納される画像変換テーブルは、赤の画素については、Real(真)に変換する一方、青の画素については、Fake(偽)に変換する。これにより、ディスプレイ8には、物体識別結果を示す画像内で、Real(真)又はFake(偽)に変換されて表示される。

【0125】

図22は変形例2, 3の実施例4に係る本人識別の画像例を示す写真画像である。図22は、Real/Fake識別を赤色/青色に対応させる例を示す。画像データメモリ1Aに格納される画像データは図22に示す。画像データメモリ1Bには以下の画像データが格納される。

(A)右側の本人に対応する画素(画像中の領域)を赤で塗りつぶした画像データ。

(B)左側の写真における本人部分に対応する画素(画像中の領域)を青で塗りつぶした画像データ(なお、背景は黒とする)。

【0126】

画像変換処理部7Aの画像変換テーブルメモリ7mに格納される画像変換テーブルは、赤の画素については、Real(真)に変換する一方、青の画素については、Fake(偽)に変換する。これにより、ディスプレイ8には、物体識別結果を示す画像内で、Real(真)又はFake(偽)に変換されて表示される。

【0127】

図22の変形例2, 3の実施例4のさらなる変形例として、図22の画像を、認証したい人物と、他の人物、とすることにより、個人認証にも応用可能である。認証したい人物の場合、赤の画像が出力され、「Real(本人)」と認証される。多数の学習用画像データを準備し、人物ごとに異なる色を割り当てると、多数の人物について認証が可能となる。

【0128】

図23Aは変形例2, 3の実施例5に係る領域分割の第1の学習入力画像例を示す写真画像である。図24Aは変形例2, 3の実施例5に係る領域分割の第1の学習目的画像例を示す写真画像である。図24Aは図23Aの学習目的画像である。

【0129】

図23Bは変形例2, 3の実施例5に係る領域分割の第2の学習入力画像例を示す写真画像である。図24Bは変形例2, 3の実施例に係る領域分割の第2の学習目的画像例を示す写真画像である。図24Bは図23Bの学習目的画像である。

【0130】

図23Cは変形例2, 3の実施例5に係る領域分割の第3の学習入力画像例を示す写真画像である。図24Cは変形例2, 3の実施例に係る領域分割の第3の学習目的画像例を示す写真画像である。図24Cは図23Cの学習目的画像である。

【0131】

図23Dは変形例2, 3の実施例5に係る領域分割の第4の学習入力画像例を示す写真画像である。図24Dは変形例2, 3の実施例に係る領域分割の第4の学習目的画像例を示す写真画像である。図24Dは図23Dの学習目的画像である。

10

20

30

40

50

【 0 1 3 2 】

変形例 2 , 3 の実施例 5 では、図 1 7 又は図 1 8 の画像処理装置を用いて、実施例 1 ~ 4 と同様に、領域分割（セグメンテーション）処理を行うことができる。具体的には、画像データメモリ 1 A には、図 2 3 A ~ 図 2 3 D の学習入力画像の画像データを格納し、画像データメモリ 1 B には図 2 4 A ~ 図 2 4 D の学習目的画像の画像データをそれぞれセットで格納した後学習し、学習して得た光学フィルタの分光透過特性を用いて光学フィルタを形成して動作させる。ただし、領域分割後の画像をそのまま出力させるときは、画像変換処理部 7 A（画像変換テーブルメモリ 7 m を備える）は必要がない。しかし、領域分割画像において、別の情報に変換するなどの用途があるときは、所定の画像変換テーブルを格納した画像変換テーブルメモリ 7 m を備える画像変換処理部 7 A により画像変換する必要がある。

10

【 0 1 3 3 】

以上の実施形態及び変形例においては、スペクトル復元（分光画像生成）又はクラスタリングについての学習及び識別処理について説明しているが、何れについても、下記例の画像データ及びデータテーブルのセットを大量に準備することが好ましい。ここで、好ましくは 1 0 0 0 セット以上、実験レベルでは 2 0 ~ 3 0 セットでも良好に識別できる場合もある。このため、各画像データに通し番号を付与したデータをそれぞれのメモリに格納するのが好ましく、学習時には各セット毎に学習させることとなる。また、画像データメモリ 1 A に格納される画像データは、通常の RGB 又は Y M C K でのカラー画像データである（分光画像生成（Hyperspectral reconstruction）の構成では例えば 2 5 6 × 2 5 6 × 3 1 の分光画像が格納されるが、これらの変形例 2 ~ 4（物体識別の構成）では分光画像データについて識別も可能だが、それに限定されない。

20

【 産業上の利用可能性 】

【 0 1 3 4 】

以上詳述したように、本発明に係る画像処理装置及び方法によれば、スペクトル復元のための分光透過特性の最適化を行うことができる。これにより、例えば撮像カメラなどの画像センサのための光学フィルタの分光透過特性の応答関数の設計を支援することができる。

【 符号の説明 】

【 0 1 3 5 】

- 1 , 1 A , 1 B 画像データメモリ
- 2 A , 2 B , 2 C 畳み込みニューラルネットワーク（CNN）
- 5 スペクトル再構成ニューラルネットワーク（SRNN）
- 6 誤差演算器
- 7 画像データメモリ
- 7 A 画像変換処理部
- 7 m 画像変換テーブルメモリ
- 8 ディスプレイ
- 9 デモザイク処理回路
- 1 0 コントローラ
- 2 0 , 2 0 A , 2 0 B , 2 0 C ハードウェア実装部
- 2 1 対物レンズ
- 2 2 , 2 3 ビームスプリッタ
- 2 4 反射板
- 3 1 , 4 1 , 5 1 , 3 1 A 光学フィルタ
- 3 2 , 4 2 , 5 2 バンドパスフィルタ
- 3 3 , 4 3 , 5 3 画像センサ
- 6 1 , 7 1 , 8 1 入力層
- 6 2 , 7 2 , 8 2 中間層
- 6 3 , 7 3 , 8 3 出力層

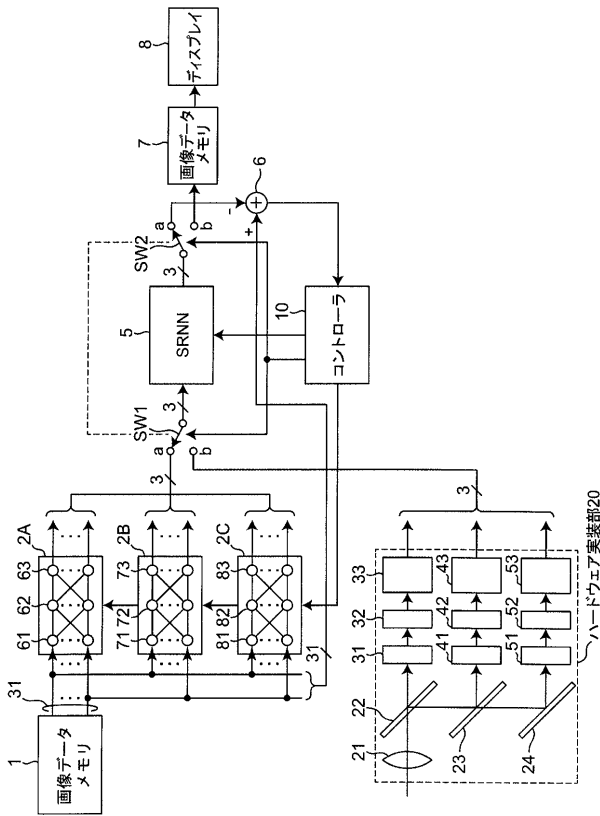
30

40

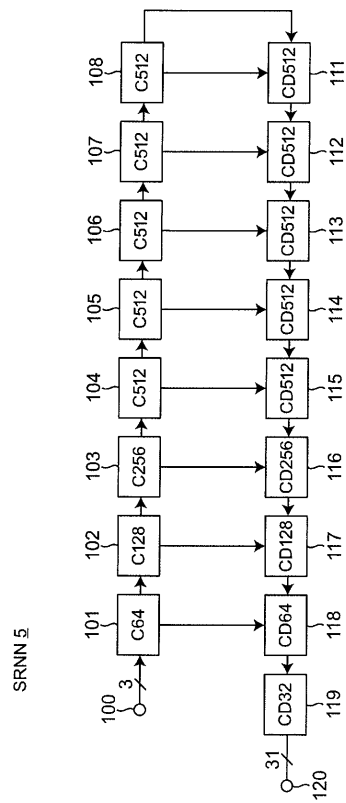
50

- 1 0 0 入力層
- 1 0 1 ~ 1 1 9 中間層
- 1 2 0 出力層
- S W 1 , S W 2 スイッチ

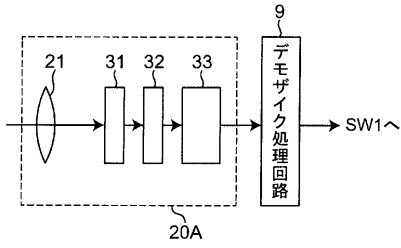
【 図 1 】



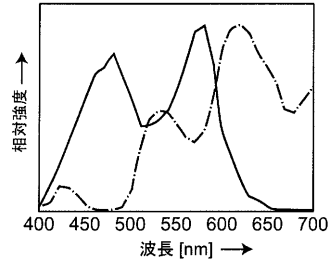
【 図 2 】



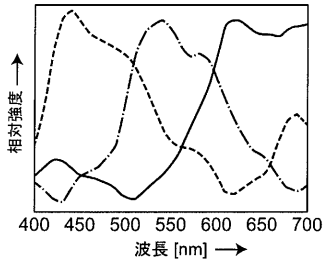
【 図 3 】



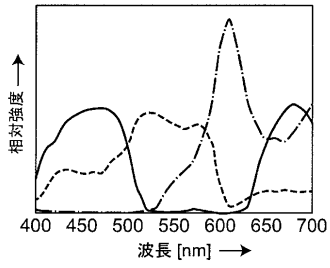
【 図 4 B 】



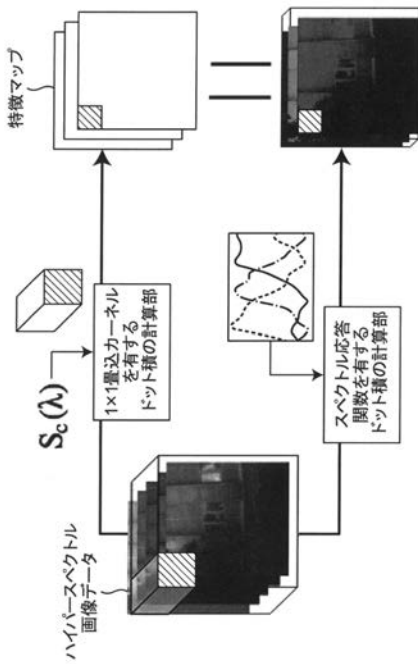
【 図 4 A 】



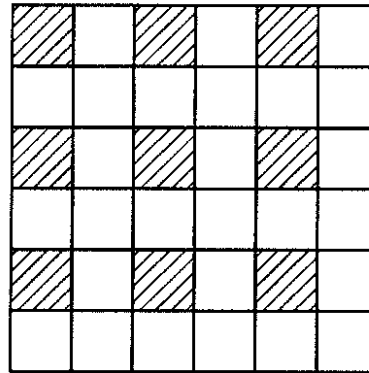
【 図 4 C 】



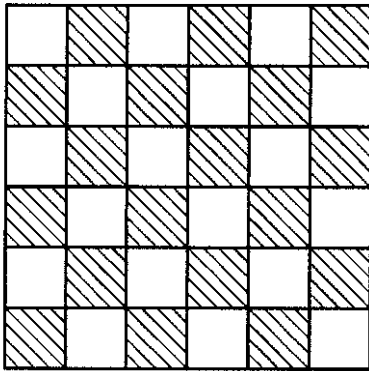
【 図 5 】



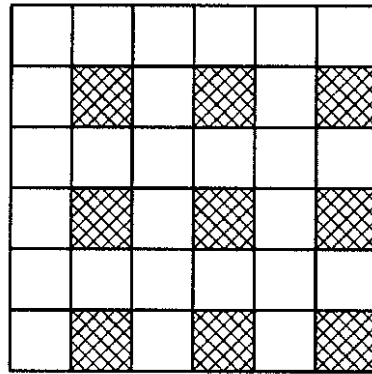
【 図 6 A 】



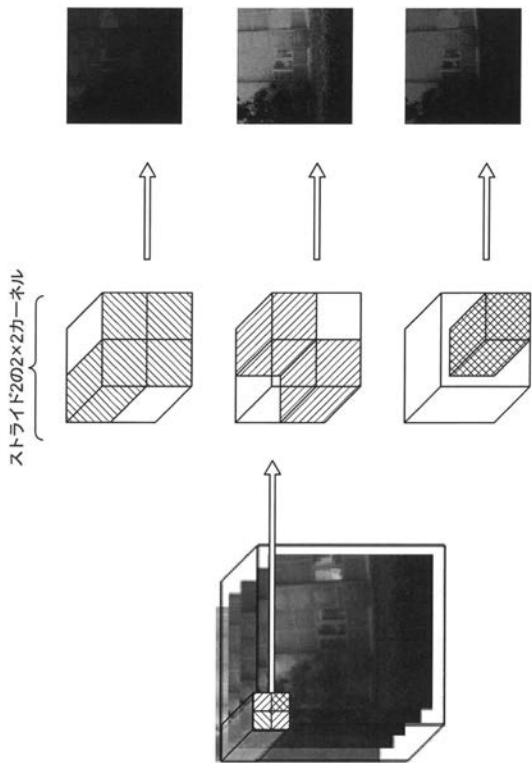
【 図 6 B 】



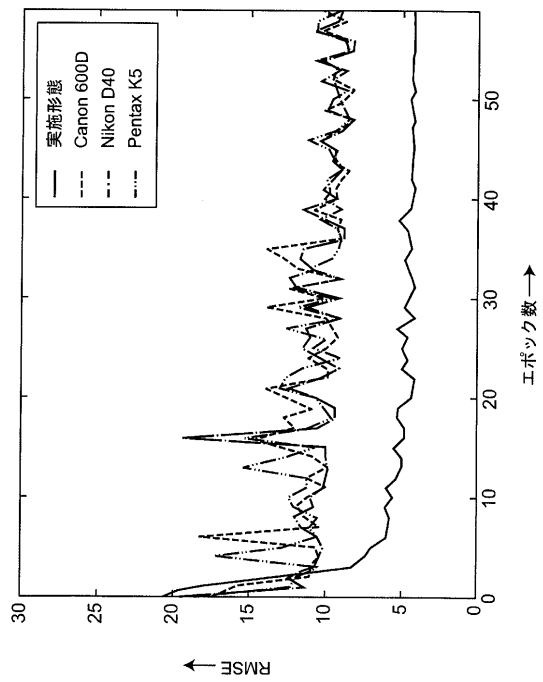
【 図 6 C 】



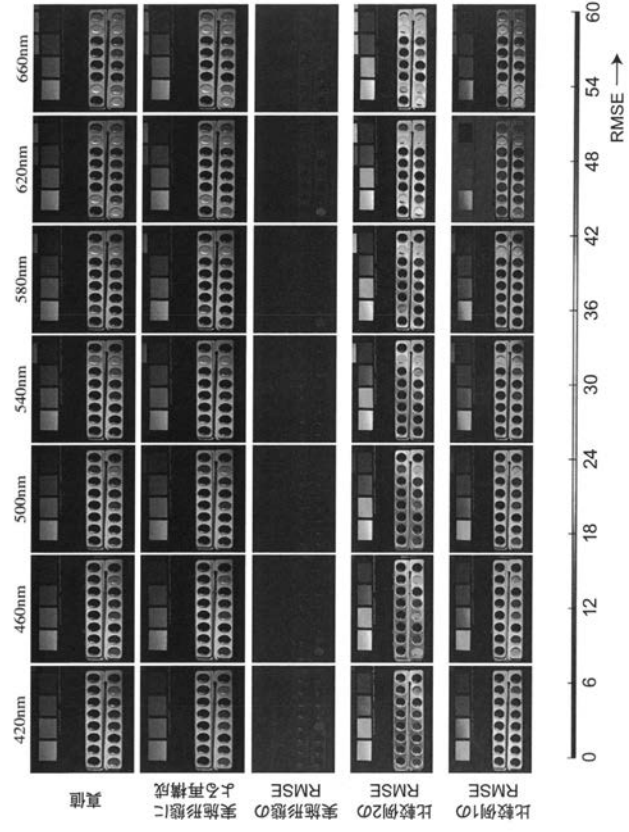
【 図 7 】



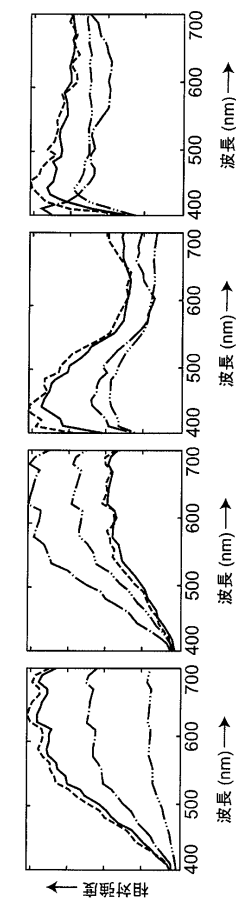
【 図 8 】



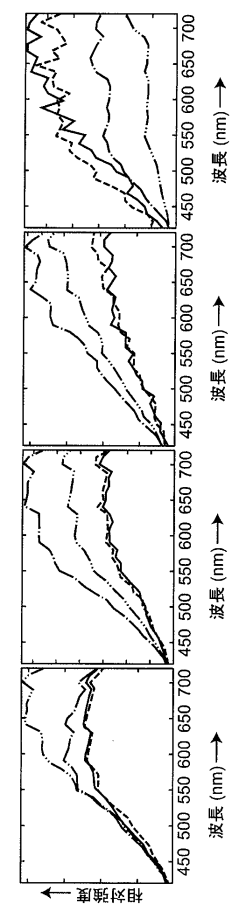
【図9】



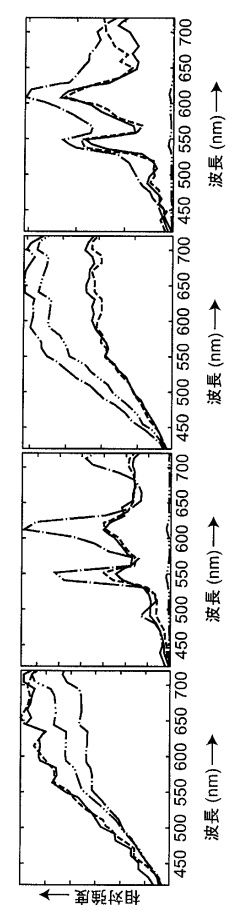
【図10A】



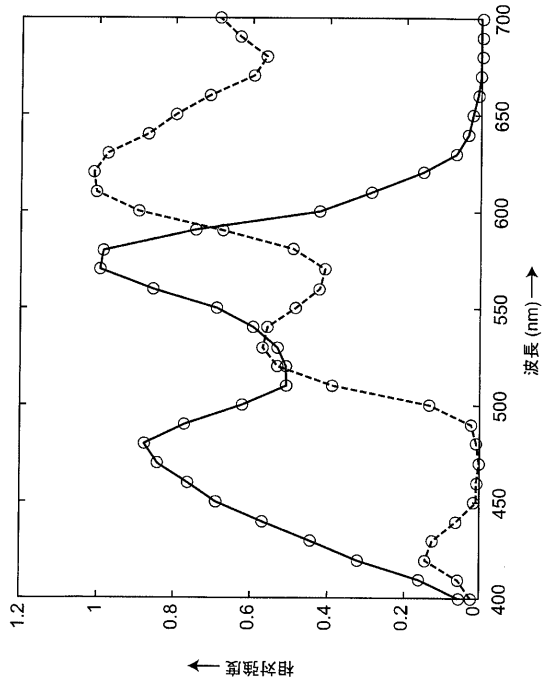
【図10B】



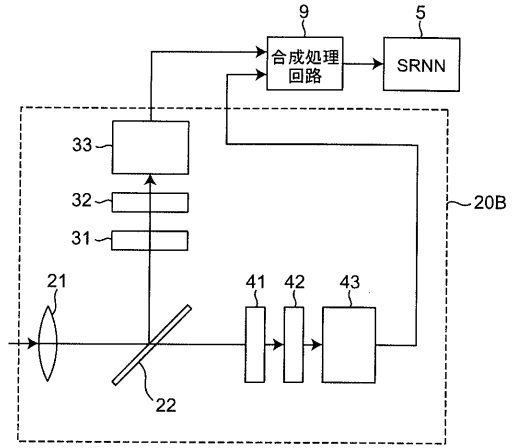
【図10C】



【图 1 1】



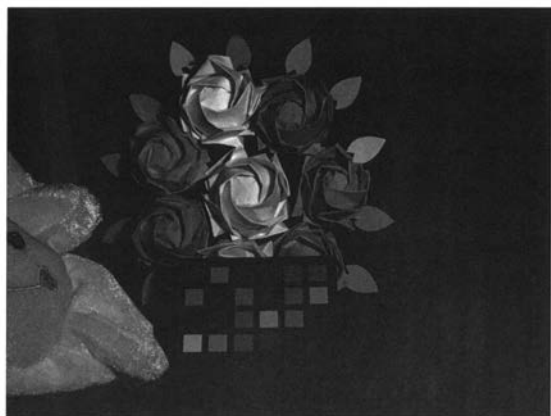
【图 1 2】



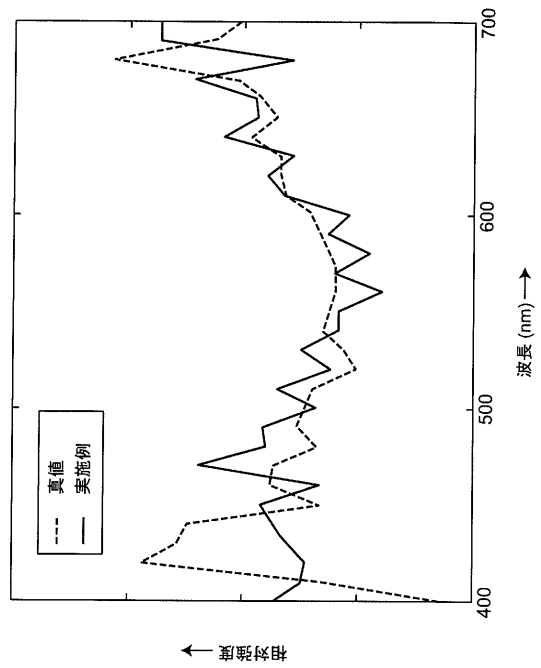
【图 1 3 A】



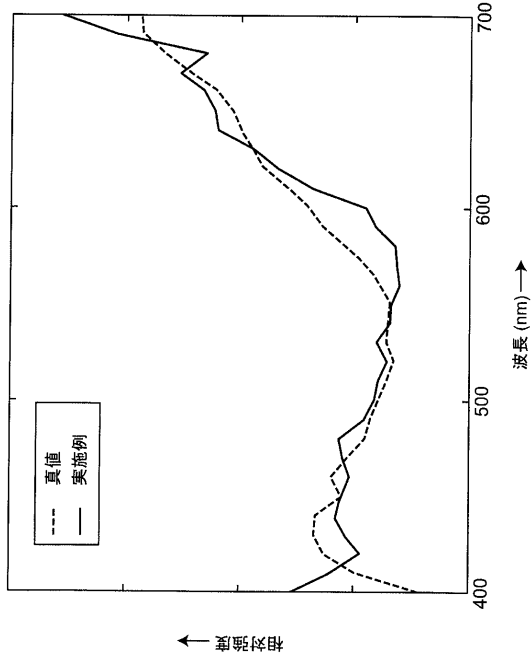
【图 1 3 B】



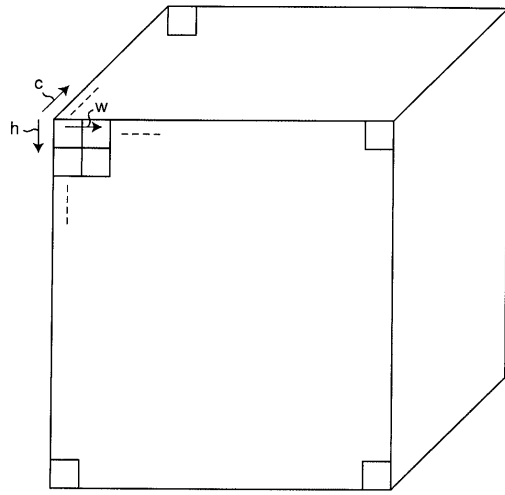
【图 1 4 A】



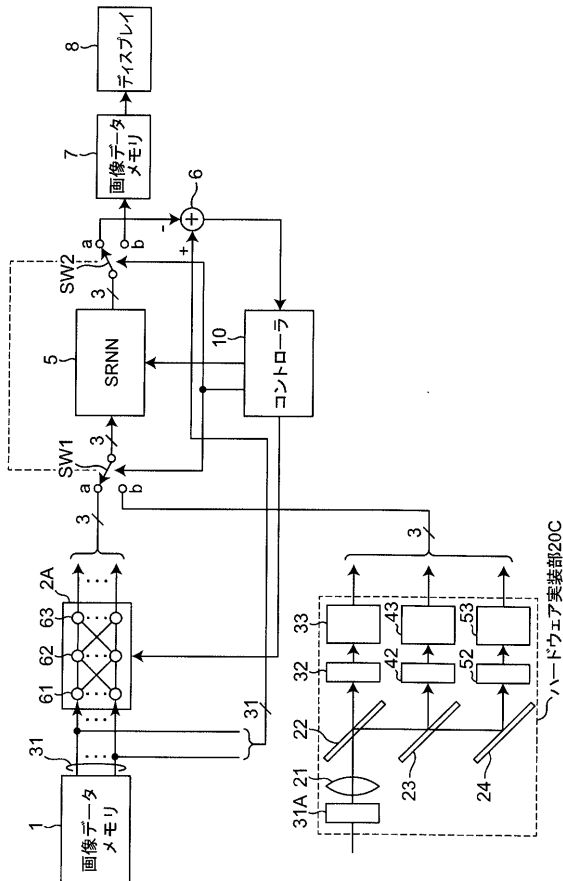
【図14B】



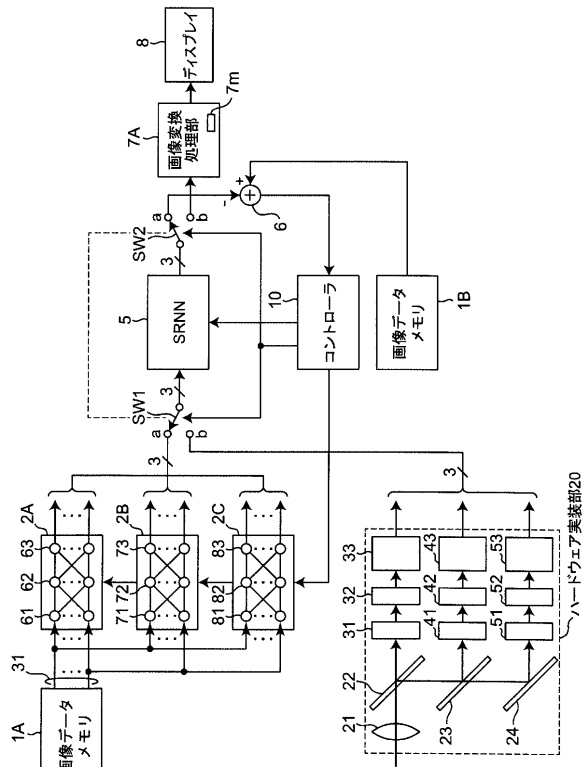
【図15】



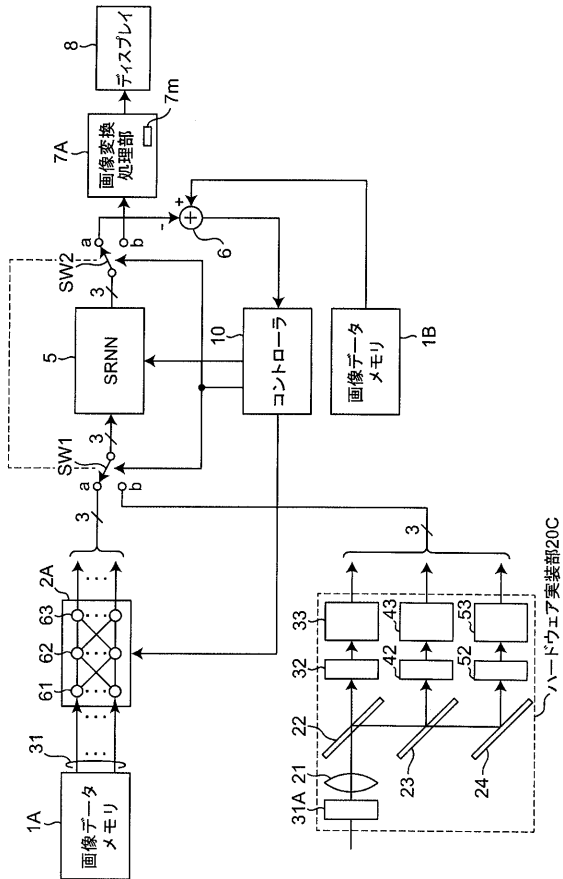
【図16】



【図17】



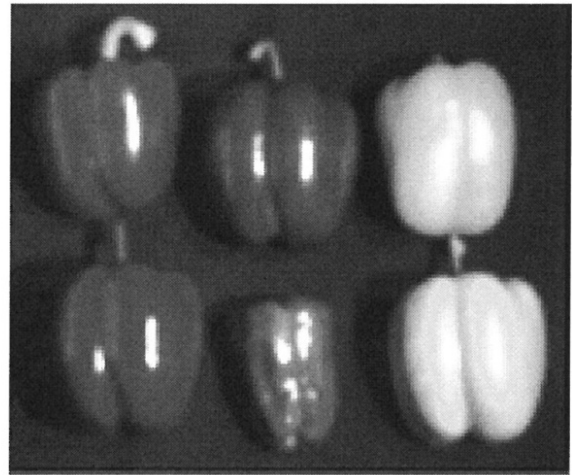
【 図 1 8 】



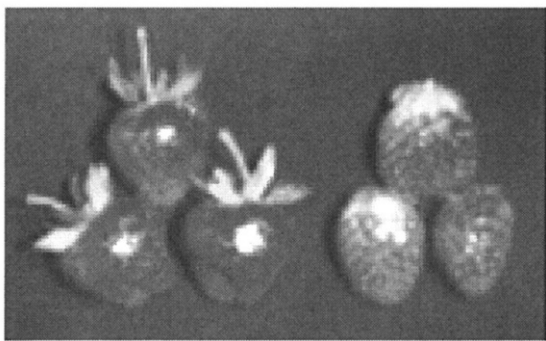
【 図 1 9 】



【 図 2 0 】



【 図 2 1 】



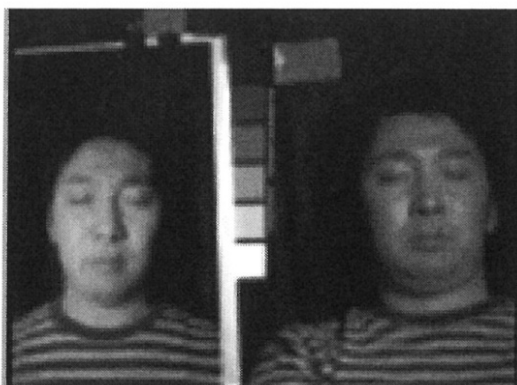
【 図 2 3 A 】



【 図 2 3 B 】



【 図 2 2 】



【 図 2 3 C 】



【 図 2 3 D 】



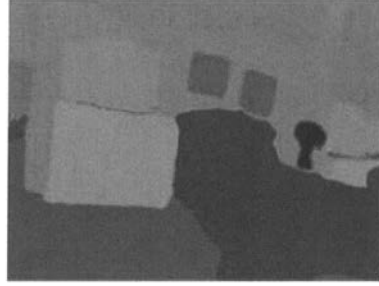
【 図 2 4 C 】



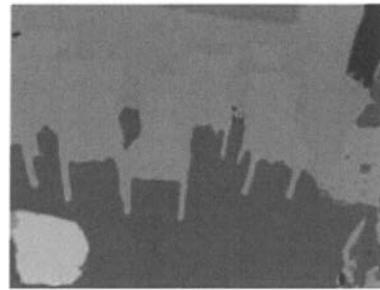
【 図 2 4 D 】



【 図 2 4 A 】



【 図 2 4 B 】



フロントページの続き

(51) Int.Cl.	F I	テーマコード(参考)
G 0 2 B 5/28 (2006.01)	G 0 2 B 5/28	
G 0 2 B 5/20 (2006.01)	G 0 2 B 5/20	

- (72)発明者 **ゲー・リン**
 東京都千代田区一ツ橋二丁目1番2号 大学共同利用機関法人情報・システム研究機構 国立情報学研究所内
- (72)発明者 **鄭 銀強**
 東京都千代田区一ツ橋二丁目1番2号 大学共同利用機関法人情報・システム研究機構 国立情報学研究所内
- (72)発明者 **佐藤 いまり**
 東京都千代田区一ツ橋二丁目1番2号 大学共同利用機関法人情報・システム研究機構 国立情報学研究所内
- (72)発明者 **アントニー・ラム**
 東京都千代田区一ツ橋二丁目1番2号 大学共同利用機関法人情報・システム研究機構 国立情報学研究所内
- (72)発明者 **小野 順貴**
 東京都千代田区一ツ橋二丁目1番2号 大学共同利用機関法人情報・システム研究機構 国立情報学研究所内

F ターム(参考) 2G020 AA04 AA08 CC28 CD38 CD60 DA13 DA33 DA34 DA65
 2H148 AA12 AA18 GA04 GA12 GA18 GA24 GA33 GA61
 5B057 CA01 CA08 CA12 CA16 CB01 CB08 CB12 CB16 CE17 CE18
 CH11 CH16 DA16 DC25 DC40
 5C065 BB30 EE03