

# RGB 画像に基づく高次元分光画像推定のための次元拡張アルゴリズムの提案 —農作業者の高齢化に伴う収穫時期予測システムの構築—

佐藤遼史<sup>†1</sup> 高尾郁也<sup>†2</sup> 浜田百合<sup>†1</sup> 鏑木崇史<sup>†3</sup> 栗原陽介<sup>†1</sup>  
 青山学院大学<sup>†1</sup> SAP ジャパン株式会社 ICA<sup>†2</sup> 国際基督教大学<sup>†3</sup>

## 1. はじめに

農場において、野菜や果物といった農作物の収穫時期は、一般的にそれらの着色度合いによって判別され、長期間の作業により習熟した農作業者の経験や勘により決定される。しかし、近年の農作業者の高齢化に伴い、そのような収穫時期を正確に予測するための技術を持つ農作業者が減少傾向にある[1]。それに伴い、農作物の収穫時期決定の基準が不明確になり、正しい時期に収穫されないという問題点が挙げられる。

高次元分光画像を用いることで、農作物を解析し、損傷の度合いを特定するシステムの開発が進んでいる[2]。RGB画像が3つの波長情報を持っているのに対し、高次元分光画像は数百の波長情報を持っているため、農作物の成熟による僅かな色の変化や、成分の変化を捉え、高精度に損傷の被害を予測できると考えられている。この技術を用いることで、収穫物を高次元分光画像によって解析し、適切な収穫時期を予測することを考える。高次元分光画像を撮影するためにはハイパースペクトルカメラが用いられるが、RGB画像取得する際と比較してシステムの規模が大きいことや、カメラが非常に高価である点が問題点として挙げられ、実際の農場への導入は難しいと考えられる。そこで、我々の身の回りに普及しているRGB画像から高次元分光画像を取得できれば、より安価に高次元分光画像を取得でき、普及への一助となる。

RGB画像から高次元分光画像を推定することは、3波長の波長情報から数百波長の波長情報への次元拡張と捉えることができるが、次元を拡張して画像を推定する研究としては、1次元から3次元へ拡張するカラー化の研究が挙げられる。このカラー化の研究は主にNeural Network[3,4]やAutoencoder[5]を手法として高精度な推定が行われている。

そこで本研究では、Autoencoderを用いた次元拡張アルゴリズムにより、RGB画像から高次元分光画像を推定するための手法を提案する。

## 2. 提案手法

### 2.1 高次元分光画像の仕組み

図1に高次元分光画像の取得のためのハイパースペクトル

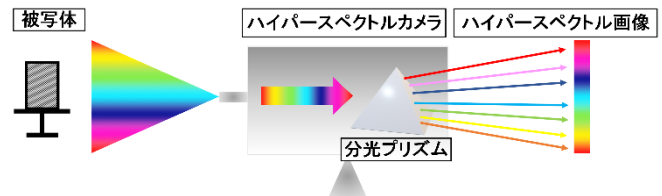


図1 ハイパースペクトルカメラの原理

ルカメラの原理の概要図を示す。光源による光から被写体に反射した白色光を、カメラ内の分光プリズムによって分光する。白色光に含まれる波長が広範囲であるほど、より多くの波長情報を表現できる。自然界に含まれる波長は本来連続的であるが、それらを画像として取得し出力する場合、カメラによって特定の波長情報に限定して取得する。RGB画像では3波長に限定して取得し画像として出力している一方、高次元分光画像の場合、数百波長のスペクトル値を取得することができる。

### 2.2 次元拡張アルゴリズム

本提案手法では、ハイパースペクトルカメラの「レンズから取り込まれた白色光を数百次元の波長情報に分光し高次元分光画像を取得する」という機能をソフトウェア上で疑似的に構築し、RGB画像とRGBの3値 $R_{i,j}, G_{i,j}, B_{i,j}$ から算出される色相、彩度、明度を表すHSVの3値 $H_{i,j}, S_{i,j}, V_{i,j}$ から高次元分光画像の推定を行う。なお、 $i$ は縦方向の画素、 $j$ は横方向の画素を表す。図2に示すように、次元拡張アルゴリズムでは、Neural Networkを用いたチューニング機能とAutoencoderのDecoder部分を用いた次元拡張機能を組み合わせ構築される。

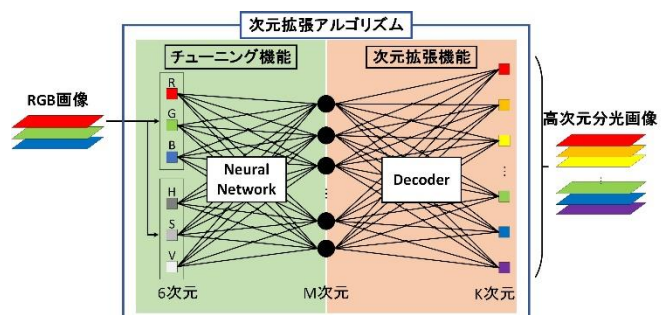


図2 次元拡張アルゴリズムのアーキテクチャー

How to Prepare Your National Convention of IPSJ Reports in MS-Word  
 †1 RYOJI SATO, YURI HAMADA, YOSUKE KURIHARA, Aoyama Gakuin University

†2 IKUYA TAKAO, SAP Japan ICA  
 †3 TAKASHI KABURAGI, International Christian University

### 2.2.1 次元拡張機能のための Autoencoder の学習

図 3 に示すような入力層, 出力層に高次元分光画像を与えた Autoencoder を学習させる. 中間層の次元数を入出力層の次元数よりも小さく設定させることで, 低次元から高次元の波長情報へと分光させる機能を Decoder に与える. 高次元分光画像の  $K$  次元の波長情報を入力層に与える. Autoencoder の中間層は  $M$  次元とする.

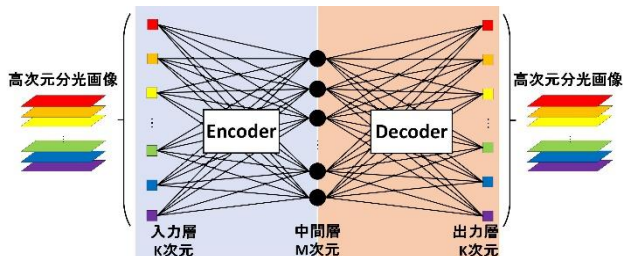


図 3 Autoencoder の学習

### 2.2.2 チューニング機能を持つ Neural Network の学習

$K$  次元の波長情報を  $M$  次元に圧縮する Encoder,  $M$  次元の入力値を  $K$  次元の波長情報に分光する Decoder の実装が Autoencoder の学習により完了した後, 高次元分光画像の推定の際に用いる Decoder への入力値を変換する Neural Network の学習を行う. 図 4 に示すように, Neural Network の入力層には推定される高次元分光画像の説明変数にあたる RGB 画像の RGB の値, そしてそれらから算出される HSV の値を与える. 出力層には高次元分光画像を Encoder に入力することで算出される Autoencoder の  $M$  次元の中間層の値を与える. HSV の値は, RGB の画素値から以下の(1)~(3)式を

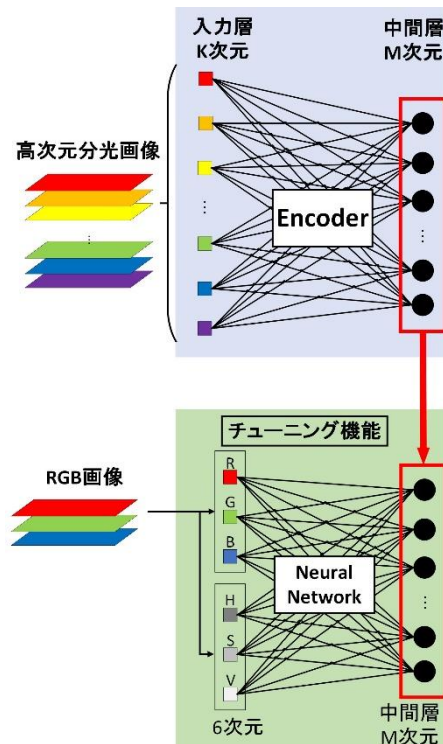


図 4 Neural Network の学習

用いて算出する. 各画素における画素の RGB のうち最大値を  $MAX_{i,j}$ , 最小値を  $MIN_{i,j}$  とする.

$$h_{i,j} = \begin{cases} 60 \times \frac{G_{i,j} - R_{i,j}}{MAX_{i,j} - MIN_{i,j}} & (MIN_{i,j} = R_{i,j}) \\ 60 \times \frac{G_{i,j} - R_{i,j}}{MAX_{i,j} - MIN_{i,j}} + 120 & (MIN_{i,j} = R_{i,j}) \\ 60 \times \frac{G_{i,j} - R_{i,j}}{MAX_{i,j} - MIN_{i,j}} + 240 & (MIN_{i,j} = R_{i,j}) \\ 0 & (MIN_{i,j} = MAX_{i,j}) \end{cases} \quad (1)$$

$$s_{i,j} = \frac{MAX_{i,j} - MIN_{i,j}}{MAX_{i,j}} \quad (2)$$

$$v_{i,j} = MAX_{i,j} \quad (3)$$

ニューラルネットワークの出力層の値を, Autoencoder の Decoder に入力することで高次元分光画像を出力する.

## 3. 検証実験

### 3.1 実験環境

本検証では, ハイパースペクトルカメラとして PikaXC2(RESONON)を使用する. カメラの撮影可能波長範囲は 398.78nm~1016.78nm であり, 1.34nm の分解能を持つ. したがって, 462 次元の波長情報を持つ. 画像サイズは 1 画像あたり 1600×1000 画素であり, 計 1600000 画素を持つ. なお, RGB 画像も同じハイパースペクトルカメラより取得した.

### 3.2 実験手順

大学キャンパス内屋外駐車場にて, 太陽光を光源として画像の取得を行った. 撮影対象はトマトであり, 本検証では学習と推定に同じ画像を用いた. 検証に用いた RGB 画像を図 5 に示す.



図 5 取得した RGB 画像

本検証では, 462 の波長帯域をもつ高次元分光画像のうち, ノイズが多く含まれている赤外線の帯域の波長を除き, 画像において目視でトマトと判別できる 390 波長分 (398.67nm~918.76nm) を検証に用いた. それらの 390 波長分

の高次元分光画像を用いて、入力層と出力層の次元数  $K=390$ 、中間層の数  $M=6$  の Autoencoder を学習させる。そして、入力層に、取得した RGB 画像の RGB と HSV の 6 次元、出力層に Encoder によって圧縮される中間層 6 次元の値を与え、Neural Network を学習させる。

推定では、図 5 に示す RGB 画像の RGB から HSV を算出し、それらの 6 次元のデータを次元拡張アルゴリズムに入力することで、各画素に 390 波長分の高次元分光画像を出力する。

### 3.3 評価方法

本検証では、ハイパースペクトルカメラによって取得された高次元分光画像である正解画像と、推定された高次元分光画像の対応する画素間の画素値を用い、次式に定義する類似度を算出し、評価を行う。なお、類似度は 0 から 1 の範囲で算出され、各波長に 1 つの値が算出される。したがって本検証では 390 の類似度が算出される。それらの 390 の値を、以下の(a)から(h)に示す 8 つの帯域ごとの平均類似度を算出し、評価を行う。

$$Similarity_h = \sum_{x=1}^{1600} \sum_{y=1}^{1110} \frac{Image_{correct_h}^{x,y}}{\sqrt{\sum_{i=1}^{1600} \sum_{j=1}^{1110} Image_{correct_h}^{i,j^2}}} \times \frac{Image_{predicted_h}^{x,y}}{\sqrt{\sum_{i=1}^{1600} \sum_{j=1}^{1110} Image_{predicted_h}^{i,j^2}}} \quad (4)$$

- (a)紫色帯域, 398.67nm~428.97nm
- (b)藍色帯域, 430.29nm~459.39nm
- (c)青色帯域, 460.66nm~499.02nm
- (d)緑色帯域, 500.35nm~569.37nm
- (e)黄色帯域, 570.70nm~589.33nm
- (f)橙色帯域, 590.66nm~609.32nm
- (g)赤色帯域, 610.65nm~779.49nm
- (h)赤外線帯域, 780.89nm~918.76nm

## 4. 実験結果

### 4.1 推定画像

図 6 と図 7 に、撮影された正解画像と、提案手法によって出力された推定画像を示す。全 390 波長のうち、例として 602.65nm と 810.49nm の 2 波長分を示す。

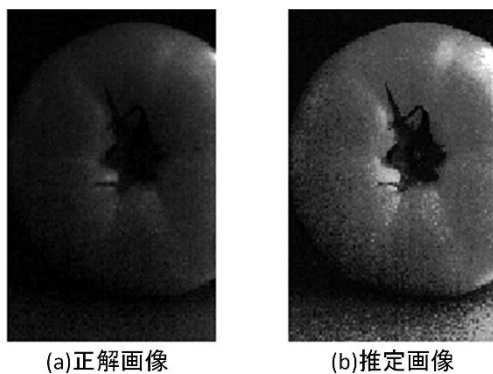


図 6 602.65nm における結果

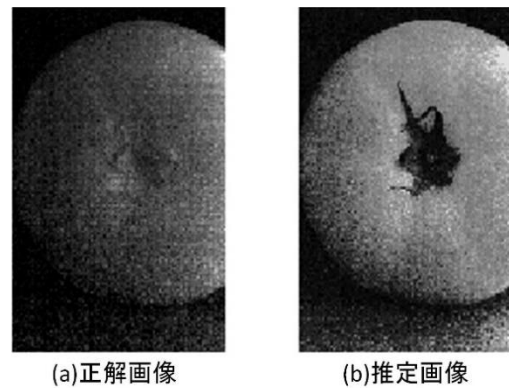


図 7 810.49nm における結果

図 6 に示す 602.65nm における結果を比較すると、正解画像に比べ推定画像における全他の色味はやや明るいものの、トマトの実とヘタの形の特徴を適切に捉えられている。なお、図 6 における 2 枚の画像の類似度は 0.955 となった。一方、図 7 に示す赤外線帯域の波長における結果を比較すると、正解画像においてトマトのヘタが白っぽく写り、消えているように見えるのに対し、推定画像においてはヘタが黒くはっきりと写っており、次元拡張アルゴリズムにおける学習の過程で、赤外線帯域の波長の特徴を捉え切れなかったといえる。なお、図 7 における 2 枚の画像の類似度は 0.912 となった。

### 4.2 類似度

表 1 に、各波長帯域における類似度評価の結果を示す。

表 1 各波長帯域における類似度

| 波長帯域  | 類似度   |
|-------|-------|
| 紫色帯域  | 0.901 |
| 藍色帯域  | 0.945 |
| 青色帯域  | 0.949 |
| 緑色帯域  | 0.950 |
| 黄色帯域  | 0.945 |
| 橙色帯域  | 0.953 |
| 赤色帯域  | 0.945 |
| 赤外線帯域 | 0.853 |
| 平均    | 0.930 |

390 波長全体での平均類似度は 0.930 と、高精度で推定できていることが分かる。各波長帯域の類似度を見ると、紫色帯域から赤色帯域までの可視光の帯域においてはいずれも 0.90 を超えていることが分かる。一方赤外線帯域においては 0.853 と、可視光帯域と比較してやや精度が落ちる結果となった。推定画像におけるトマトのヘタに表れた正解画像と推定画像の特徴の違いが、類似度にも表れたと考えられる。

## 5. 考察

次元拡張アルゴリズムにおける次元拡張機能、つまり

Decoder での圧縮・復元の処理が適切に行われているか検証するため、図 8 と図 9 に、圧縮前の正解画像と Autoencoder の学習後の復元された画像を示す。全 390 波長のうち、4.1 と同じ 602.65nm と 810.49nm の 2 波長分における結果を示す。

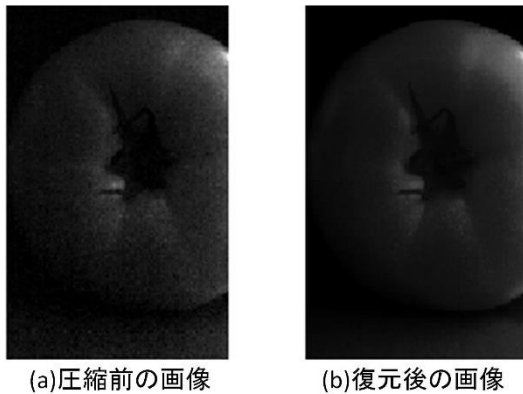


図 8 602.65nm における結果

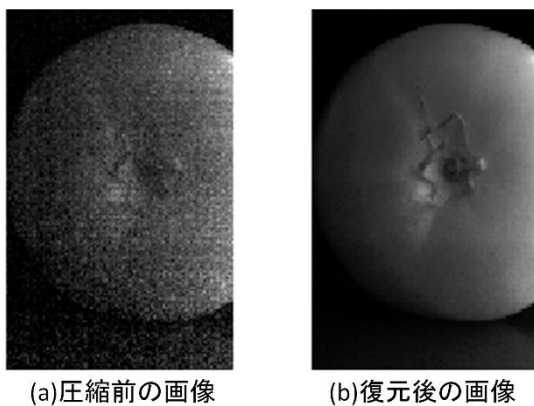


図 9 810.49nm における結果

まず図 8 における画像を比較すると、図 6 の疑似プリズムにおける推定結果においてやや異なる結果となった全体の色味も含め、非常に類似している。赤外線波長における画像においては、図 7 における推定結果においてトマトのヘタの特徴に違いが見られたのに対し、図 9 では復元後の画像についてもトマトのヘタが白っぽく写り、2 枚の画像は類似しているといえる。

これらの結果より、Autoencoder における Encoder および Decoder の学習は適切に行われているといえる。したがって、チューニング機能をもつ Neural Network の学習が適切に行われていないといえる。

## 6. まとめと今後の展望

本研究では、Autoencoder と Neural Network を組み合わせた次元拡張アルゴリズムによって、RGB 画像から高次元分光画像の推定を行った。

推定画像と類似度の結果より、概ね高精度で高次元分光画像の推定を行うことができ、次元拡張アルゴリズムの高次元分光画像を推定する手法としての有用性が示された。

しかし、赤外線波長における画像にて、トマトのヘタの特徴が学習しきれず、正解画像と推定画像の特徴に差がある結果となった。しかし次元拡張機能のための Autoencoder の学習は適切に行われていたため、チューニング機能を持つ Neural Network の学習を改善し、特に赤外線帯域の画像推定をより高精度に行う必要があると考える。

今後の展望として、本検証でやや精度の低かった赤外線帯域の学習を、可視光帯域と分けて行うことや、Neural Network における学習において、入力層の次元や中間層を異なるパラメータで試行する必要があると考える。

今後、更に高精度で高次元分光画像を推定することができれば、実際の農業分野においてアプリケーションとしての普及が進み、収穫時期などの判別を RGB 画像の取得のみで行うことができ、農業分野への一助となるのではないかと考える。

## 参考文献

- [1] 農林水産省, 農業就業人口及び基幹的農業従事者数について, 2020, (参照 2021-08-04)
- [2] W. Yang et al, "Diagnosis of Plant Cold Damage Based on Hyperspectral Imaging and Convolutional Neural Network," IEEE Access, vol. 7, pp. 118239-118248, 2019.
- [3] B. Karlik and M. Sariöz, "Coloring gray-scale image using artificial neural networks," in - 2009 2nd International Conference on Adaptive Science & Technology (ICAST), 2009, pp. 366-371.
- [4] Z. Cheng, Q. Yang and B. Sheng, "Colorization Using Neural Network Ensemble," IEEE Transactions on Image Processing, vol. 26, (11), pp. 5491-5505, 2017.
- [5] S. Lee et al, "Colorization-Based Compression Using Optimization," IEEE Transactions on Image Processing, vol. 22, (7), pp. 2627-2636, 2013.