

特定用途のマルチスペクトルカメラ設計のための 重要フィルタの選択

櫻井 俊^{1,a)} 青砥 隆仁¹ 船富 卓哉¹ 久保 尋之¹ 向川 康博¹

概要: 分光分布に対し百数十の分解能をもつハイパースペクトルカメラは、土壌や森林のモニタリング、地質の分類など広い分野で多目的に活用されている。しかしカメラが高価、撮影に時間がかかるという問題から産業利用は限定的であった。一方、計量化学分野では物質が固有の特定少数の波長の光を吸収することを利用して、物質の計量が行えるという知見がある。そこで本研究では、物質を限定すれば計量に必要な波長は離散的で少数となる性質を利用し、少数の光学フィルタを通して撮影を行うマルチスペクトルカメラを用いることで、安価かつ高速に特定物質の分布のイメージングを実現し、これを産業応用することを目的とする。本論文は、(1) 市販されている光学フィルタがもつ有限幅の分光透過率を考慮して多変量校正モデルを再定義し、(2) 候補となるフィルタから特定物質の計量に用いる最適なフィルタの組み合わせを Adaptive lasso により選択する手法を提案する。こうして選択された少数の光学フィルタを用いて構成された (3) 簡易なマルチスペクトルカメラにより物質分布のイメージングが行えることを示す。実環境下での提案手法の有効性を確認するため、最適なフィルタの選択とこれを備えるマルチスペクトルカメラを作成し、その応用例として果実における糖度分布の推定精度を検証する。

キーワード: 糖度分布, 近赤外分光法, ハイパースペクトルカメラ, マルチスペクトルカメラ, 計量化学

1. はじめに

コンピュータビジョンの分野では、ヒトと同等、もしくはそれ以上にシーンを理解することを目指す研究が行われており、近年のセンシング技術の発展により、ヒトの目を超えた知覚の情報処理も行われている。例えば、暗視野センサ、サーモセンサ、距離センサなどはヒトの目が知覚できない光の情報を手掛かりとする。特に、ヒトが知覚する赤緑青の3バンドを拡張した分光イメージングは、農作物、土壌、森林のモニタリング、生物医工学、バイオメトリクスなどを目的とした高度な分析を可能とする非常に重要な技術である。

分光情報は、プリズムまたは回折格子を備える分光器によって光をスペクトルへ分解(分光)することで獲得できる。このような機器によって得た分光情報は一般的に波長に対して高い分解能を有するが、単一点でのみの計測しかできず、空間分解能が十分ではなかった。そこで、分光器を拡張し空間的な分光情報を獲得する分光イメージングの研究が行われている。ハイパースペクトルカメラ

(HSC:HyperSpectral Camera) は分光イメージングの手段の一つであり、百数十バンドのスペクトル画像が取得できる。

また、従来の3バンドカメラの拡張として、複数の光学フィルタにより離散的な分光情報を得るマルチスペクトルカメラ(MSC:MultiSpectral Camera)も分光イメージングの手段の一つである。一般的に分光感度の固定されたフィルタを備え、十数バンドのスペクトル画像を撮影できる。

バンド数の観点から、HSCはMSCに対して大きな利点があるが、市販のHSCの多くは撮影時間が長く費用がかかる問題がある。これに対し、MSCはこの問題を比較的緩和しているが、特定のタスクに特化した光学設計とはなっていないため、HSCに比べて十分な精度を達成することは困難である。用途に応じて光学フィルタの組を設定可能な製品[1]もあるが、数や組み合わせを選定する方法は明らかではない。

そこで本研究では果実に含まれる糖分分布のイメージングを応用先として想定し、特定物質の計量を目的とした場合にMSCに用いる光学フィルタの組をどう設計すれば良いかの戦略を立てる方法について検討を行う。特定物質の計量に有用な光学フィルタを少数選択することで、測定におけるシステムの実用性やコストはもちろん、精度について

¹ 奈良先端科学技術大学院大学 〒630-0192 奈良県生駒市高山町8916 番地 5

^{a)} sakurai.shun.sk4@is.naist.jp

ても HSC と同等の結果を得ることが可能になると考えられる。本稿では、市販の光学フィルタから物質分布のイメージングに有効な少数の組を選択する最適化手法について述べる。また、提案手法によるフィルタ選択に基づき MSC を設計し、イチゴの糖度分布のイメージングを行った結果について述べる。

2. 関連研究

分光イメージングとは空間的な分光情報を取得することであり、例えば、地球観測やリモートセンシング [17], [21], 食品の安全や品質の検査 [7], [14], 化学薬品工場での品質モニタリング [5], [20] など、その他様々な用途で重要な役割を果たす。この分光イメージングは取得するバンドの数や各バンドの狭さに応じて、HSC および MSC の 2 つのタイプに分類することができる。

HSC はプリズムなどの分光装置を用いることで、百数十のバンドを連続的に撮影することができる。しかし、このようなカメラは一般的に高価であり、また撮影に時間がかかるため、産業に導入することは現実的でない。HSC で取得されるスペクトル画像は 1 画素あたり百数十の分光情報を持つ高次元なデータと捉えることができる。様々な目的に対し、このような高次元データからクラス分類や回帰を行う予測モデルを構築する際には、波長数 (次元数) の増加に由来する“ヒューズの現象”が避けられない [10]。例えばリモートセンシングでは、現地調査数の限界から標本の数が限られるため、通常回帰モデルの適用は困難である [13], [24]。そこで、混合ガウス分布モデル (GMM: Gaussian Mixture Model) や縮小推定を用いた波長選択 (次元圧縮) が行われてきた [19], [20], [27]。

一方で MSC は、光学フィルタによって 3 から 10 程度のバンドのスペクトル画像を取得するものである。ほとんどの MSC は、一般的に分光感度の固定されたフィルタを備え、十数バンドのスペクトル画像を撮影する。汎用的な目的のため各バンドの分光感度は等間隔に設計されている。一部の製品では、光学フィルタを組み合わせることによって取得するバンドの分光感度を自由に選択でき、用途に応じてフィルタの選択を設計することができる [1]。このフィルタの組み合わせの選択手法に注目した研究が存在する。Preece らは遺伝的アルゴリズムによって最適化された複数のフィルタを用い、人間の皮膚を表現するパラメータを復元する方法を提案した [18]。Styles らは組合せ最適化によって用いるフィルタを選択し、眼底のマルチスペクトル画像から病理診断に有効なパラメータの推定を行った [22]。これらの例は、用途を限定すればハイパースペクトル画像のような多数のバンドを持つスペクトル画像は必要とならないことを示している。しかし、これらの方法は眼底の病理診断や皮膚パラメータの推定に特化したモデルに基づいてフィルタを選択しているため、他の目的に適用すること

ができない。

3. 分光情報と計量化学

1 章で述べたように、本研究では少数の光学フィルタを備える MSC による特定物質の計量を目指す。化学物質は光を照射すると、それを構成する分子が構造に応じたエネルギーを吸収し、熱や化学エネルギーに変化する。この時波長 λ に応じたエネルギー $h\nu$ (h はプランク定数; 振動数 $\nu = \frac{c}{\lambda}$; c は光速) を吸収する。この性質により化学物質はその構造固有の吸収波長を少数持つ。これを利用し、特定の化学物質の計量を行う研究が行われている。

例えば、クロロフィル (葉緑素) と糖度の含有量の相関を利用したメロンの糖度分布の可視化が提案されている [23]。この方法はクロロフィルが 675nm の光を吸収するという事前知識を利用した簡易な装置であるが、一般的な果実がクロロフィルとの相関を持つとは限らないため、汎用的な手法ではないと言える。また、4 つのレーザーダイオードを用いた幼児の脳血液の酸素変化の観察装置が提案されている [15]。この方法は、ヘモグロビンのもつ近赤外領域における吸収特性に注目している。これらの研究は、少数の波長を選択することによって簡易な分光法を実現している。

3.1 多変量校正モデル

分光情報を用いた化学物質の計量は、ランベルト-ベールの法則と呼ばれる特定の波長の吸光度と化学物質の量との間にある線形性を重要な前提とする。 n 番目の波長における吸光度 a_n は、以下の式によって与えられる。

$$a_n = -\log\left(\frac{x_n}{\tilde{x}_n}\right) \quad (1)$$

ここで \tilde{x}_n と x_n は対象に入射する前の光強度と、対象を透過した後の光強度を表す。

y を化学物質の量とし、これが計量化学における多変量校正モデルによって吸光度と関連づけられる。すなわち、

$$y = \sum_n a_n c_n + c_0. \quad (2)$$

ここで、 c_n は $n = 1, \dots, N$ 番目の波長における化学物質の吸収係数を表す。多変量校正においては、未知のサンプル \mathbf{a} に対し、 $\hat{y} = \mathbf{a}^T \mathbf{c} + c_0$ による妥当な物理量の予測モデルを与える係数 \mathbf{c} を推定することが目的となり、これを回帰と呼ぶ。また、サンプル数が M の場合、式 (2) は次のような行列表現に書き換えられる。

$$\mathbf{y} = \mathbf{A}\mathbf{c} + c_0 \mathbf{1} \quad (3)$$

最も基本的な係数の推定には最小二乗 (OLS: Ordinary Least Squares) 回帰が用いられ、以下の式により得る。ただし、多変量校正モデルに含まれる切片 c_0 は省略し、これ以降も断りなく省く。

$$\arg \min_{\mathbf{c}} \|\mathbf{Ac} - \mathbf{y}\|^2 \quad (4)$$

ここで \mathbf{A} は計測された吸光度 a_n を縦に並べた $M \times N$ 行列, 係数 \mathbf{c} は $\mathbf{c} = (c_1, \dots, c_N)^T$ の係数ベクトル, \mathbf{y} は M 個のサンプルの物質量を格納したベクトルである.

3.2 変数の選択

一般に回帰においては変数の数を増やしすぎると過学習により推定精度が悪くなることが知られる. これは, 回帰モデルの表現力の高さ (変数の多さ) によって訓練データに対し過剰に適合してしまうことによる. 推定の汎化性能を高めるには現実のデータの持つ潜在変数の少なさに注目し, 変数の選択・削減を行う必要がある. 最も原始的には, OLS 回帰で決定された係数を手掛かりとし手動により得ることができる. これに対し, 半自動による変数選択の方法としてステップワイズ法, Moving Window PLS [11], 人工ニューラルネットワーク [4] が提案されているが, 局所解に陥りやすいという問題がある [12], [26].

そこで, ダーウィンの自然選択説に着想を得た遺伝的アルゴリズム (GA:Genetic Algorithm)[6] が, 非局所的探索を行う手法として提案されている. Genetic Algorithm and PLS, Genetic Algorithm-based WaveLength Selection [2], [9] は遺伝的アルゴリズムを用いた変数選択法, 変数領域の選択法であるが, 確率的な収束を待つため極端に遅くなることや, パラメータの設定が膨大となる問題がある [3].

近年, 凸コスト関数の最適化により変数選択に対する大域解を得る縮小推定と呼ばれる方法が計量化学の分野において導入されている. まず, 変数の選択問題は, 係数 \mathbf{c} の非ゼロ要素, すなわち ℓ_0 ノルム $\|\mathbf{c}\|_0$ を含めて最小化する問題として捉えることもできる. 具体的には OLS 回帰のコスト関数式 (4) に正則化項を加え, 次のように書き替えられる.

$$\arg \min_{\mathbf{c}} \|\mathbf{Ac} - \mathbf{y}\|_2^2 + \alpha \|\mathbf{c}\|_0 \quad (5)$$

第二項目の正則化項は, 係数肥大による罰則を与えるもので, 非負の正則化パラメータ α により罰則の強さを調節する. α が大きくなると変数は縮小し, 十分大きい時いくつかの変数は 0 に一致する. 正則化項は係数ベクトルに対する事前情報を与える事に相当し, 所望の性質を得るために設計される. しかしこの問題は NP 困難 [16] のため, ℓ_0 ノルム罰則を ℓ_1 ノルム罰則に凸緩和し近似解を得る. すなわち以下を解き, lasso とも呼ばれる [25].

$$\arg \min_{\mathbf{c}} \|\mathbf{Ac} - \mathbf{y}\|_2^2 + \alpha \|\mathbf{c}\|_1 \quad (6)$$

4. MSC の光学フィルタ最適化

4.1 提案手法の概要

計量化学におけるアプローチを HSC を用いて撮影され

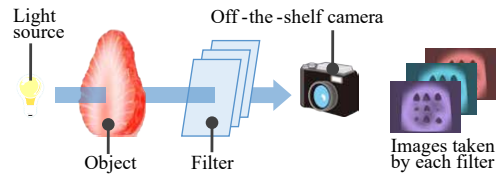


図 1 市販のカメラと光学フィルタを複数用いスペクトル画像を撮影する MSC システム. 計量に有用な少数のバンドのみを撮影する.

Type	The number of types
3 nm	10
10nm	69
25nm	16
50nm	24

表 1 各半値幅のフィルタの種類数.

たハイパースペクトル画像に適用することによって, 化学物質の分布定量を考えることができる. しかし, HSC は上に述べた理由より高価で撮影に時間がかかるため, その利用は実用における問題となると考えられる. ハイパースペクトル画像は汎用的な目的に利用できるが, 特定の化学物質の定量という目的に対しては物質に吸収される少数の波長のみが有用であり, ほとんどの波長は不要である. もし対象とする化学物質に関する事前の知識があれば, その分布定量には重要な波長のみを観測で十分である. 複数のバンドパスフィルタとカメラからなる MSC を利用し, それらの波長における吸光度を計測できれば, 化学物質の定量に関する撮影コストを大幅に減らすことが可能である. 図 1 に提案手法の概念図を示す.

このシステム実現のため, HSC によって効率的に収集された高次元スペクトル画像を用い, 計量化学に基づく解析を行うことによって実用可能な MSC を作成する方法を提案する. 提案手法は以下のステップからなる.

- (1) ハイパースペクトル画像による, 回帰分析に基づいた最適な光学フィルタの組み合わせの選定 (4.2 節, 4.3 節)
- (2) MSC の各フィルタの係数の決定 (4.4 節)

3.2 節で述べたように, 多くの先行研究が分光情報に対して重要な波長の選択を行なっている. しかし, 波長選択と MSC の光学フィルタの選択では異なる点が 3 つ存在する. 1 つ目に, 市販の光学フィルタは多く存在するが, その透過特性は有限幅の波長領域を通過する特性をもつことである. 図 2 に本研究で候補とする市販の光学フィルタの分光透過率, そのうちのバンドパスフィルタの仕様例を表 1 に示す. 狙った波長のみを観測できる狭バンドのフィルタは種類が非常に少なく, 一方で種類が多いフィルタは比較的広バンドな透過特性を持つため前後の波長も一緒に観測してしまうため, 狙った波長のみを観測ができるわけではない. 2 つ目に, 提案する MSC とフィルタ選択を行う際に用いる HSC とのカメラの分光感度の違いにより, 正

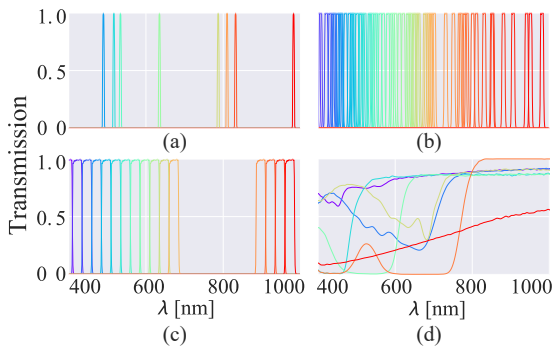


図 2 光学フィルタの透過特性の例. (a)-(c) は半幅幅が各 3nm, 10nm, 25nm のバンドパスフィルタ (Edmund optics corp.), (d) はカラーフィルタ (Roscolux, Rosco Laboratories, Inc.) の例を示す.

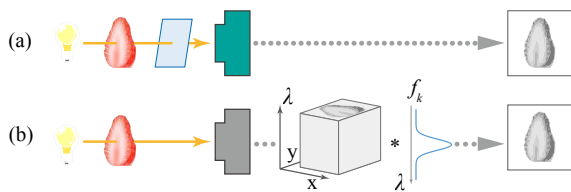


図 3 光学フィルタによる撮影の仮想化. (a) が光学フィルタを通しカメラで観測されるのに対し, (b) 光学フィルタの分光透過率とハイパースペクトル画像から, 光学フィルタによって撮影される画像を仮想的に計算可能.

しいフィルタの選択が行えない可能性がある. 市販のカメラの分光感度が分かっていると仮定することは現実的でなく, 4.4 節でこれらの問題について述べる. 3つ目に, 式 (1) の式で仮定していた線形性がフィルタを導入した場合に非線形となる点である. したがって, 従来の波長選択を拡張することにより, 光学フィルタの透過特性を考慮するようなフィルタの選択の枠組みを提案する.

4.2 ハイパースペクトル画像を用いた光学フィルタを用いた観測のシミュレーション

はじめにハイパースペクトル画像から光学フィルタを透過させた場合の観測を仮想的に再現する方法について述べる. まず候補となる市販のフィルタのうち, k 番目のフィルタの分光透過率を $\mathbf{F} = \{f_{nk}\}$ とし, これをフィルタの数だけ縦に配列した $N \times K$ 行列を作成する. 次に, HSC と同じ分光感度を持つカメラを想定すれば, スペクトルデータとフィルタ行列の積によって, フィルタを用いた場合の観測を仮想的に計算することができる (図 4 参照). なお, ここでは図 3(a) に示す光学フィルタを用いた MSC によって観測される画像と, 図 3(b) に示す HSC のスペクトル画像を用いて仮想的に MSC による観測を再現した画像が同じであると仮定している.

式 (1) から, k 番目のフィルタで観測される光強度 $\mathbf{Z} = \{z_{mk}\}$ はその透過特性 f_{nk} を用いた以下の式によって計算される.

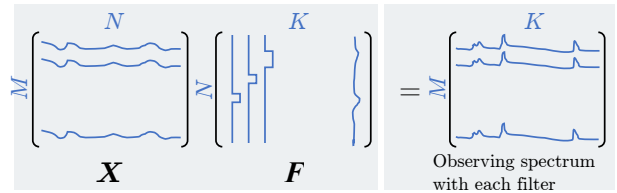


図 4 行列積により仮想的な光学フィルタを用いた場合の観測が計算される.

$$z_{mk} = -\log\left(\frac{\sum_n x_{mn} f_{nk}}{\sum_n \tilde{x}_{mn} f_{nk}}\right). \quad (7)$$

したがって, 式 (3) の \mathbf{A} を単に $\mathbf{Z} = \{z_{mk}\}$ に置き換えることによってフィルタを用いた物質量の多変量校正モデルが得られる.

$$\mathbf{y} = \mathbf{Zd} \quad (8)$$

ここで, 係数 $\mathbf{d} = (d_1, \dots, d_K)^\top$ は各フィルタの重みを与える係数ベクトルを表す. この定式化によって, 従来の波長を用いた多変量校正モデルの単なる拡張として, フィルタを用いた計量を実現することができる. 重要なフィルタの選択は, 係数 \mathbf{d} をできるだけ疎に推定することによって達成することができる.

4.3 Adaptive lasso を用いたフィルタの選択方法

縮小推定の方法は 3.2 節で紹介した lasso の他, 解析するデータの形や性質に応じて様々な派生が提案されている. その背景として, lasso の適用できる範囲の制限があげられる [8]. Fan らは lasso のような縮小推定が持つべき望ましい性質として以下のオラクル性を挙げている [12].

- 変数選択の一致: サンプル数 M を増加させると, 係数が 0 でない説明変数が正しく選択される確率が 1 に収束する性質.
- 漸近正規性: 非ゼロ係数を用いた推定量の分布が正規分布に近似できる性質.

本研究ではオラクル性を満たす adaptive lasso [28] を用いることでより正しいフィルタ選択を行うことを考える. Adaptive lasso は, 式 (6) の lasso における正則化項に重み係数 \mathbf{w} を加えたものとして解釈することができる. つまり, フィルタ特性を考慮した最適化問題として以下を解き, フィルタ選択を行う.

$$\arg \min_{\mathbf{d}} \|\mathbf{Zd} - \mathbf{y}\|_2^2 + \alpha \|\mathbf{w}^\top \mathbf{d}\|_1 \quad (9)$$

ここで,

$$w_k = \frac{1}{|d_k|^\gamma} \quad (10)$$

Adaptive lasso では, 重み \mathbf{w} の更新と \mathbf{d} の推定を繰り返すことによって最終的な推定量を得る. 重み w_k は, 0 に近い推定量に対しては大きな重みを与え 0 にし, 大きな推定量に対しては小さな重みを適用することで係数 \mathbf{d} のばらつきを抑える. なお, 初期の重み w_k は, 共線性が問題で

ない限り OLS 回帰によって求めることが推奨されている。また ridge 回帰によっても推定に安定性を与えることができる [28]。本稿では、パラメータ $\gamma > 0$ を $\gamma = 1.0$ とする。また、正則化項の強さを与えるパラメータ α によって選択される係数の数、すなわちフィルタの数の制御を行う。

4.4 選択されたフィルタの係数推定

特定物質の分布定量を行うには、必要なフィルタを備える MSC と、それらフィルタによって撮影される画像を合成するための係数が必要となる。4.3 節によってフィルタの選択と係数が求まるが、4.2 節で述べたように、図 3 において (a) と (b) が等価、つまり MSC と HSC の分光感度が同じであることを仮定していた。この仮定が成り立つ場合は、推定した係数が MSC においても有効である。しかし一般的には異なるカメラの分光感度が同じであるとは限らない。そこで本節ではそれらが異なる場合を想定し、対処について述べる。

4.4.1 各カメラの分光感度が既知の場合

はじめに MSC と HSC の分光感度が既知である場合を考える。この場合あらかじめカメラの分光感度を含めた最適化を行えばよい。つまり式 (7) に MSC と HSC の分光感度の商 $s_n = S_n^{MSC} / S_n^{HSC}$ を加え、 z を計算する。

$$z_k = \sum_n f_{nk} x_n s_n \quad (11)$$

この z_k を用いることで、フィルタの選択と係数の推定を同時に行うことができる。しかし、カメラレンズや計測環境の微小な違いなどが係数の推定に影響を与え、設計した MSC で期待の性能が得られない可能性もある。

4.4.2 各カメラの分光感度が未知の場合

各カメラの分光感度が未知の場合、 z には式 (7) による算出を用いる。ここで推定される係数は、MSC の分光感度が考慮されていないため、そのまま用いることができない。そこで、光学フィルタの選択と係数決定のステップを分離する。つまり、HSC を用いた 4.3 節の手法により光学フィルタを選択し、これを用いて構成した MSC でスペクトル画像を取得した後、改めて分布定量のための係数を推定する。この方法は、カメラの分光感度が未知の場合の 1 つの解法ではあるが、分布定量に適したフィルタが正しく選択されるためには、以下で述べる理由により用いることのできるフィルタの制限がある。

説明変数と目的変数の間に線形性を仮定する多変量校正モデルに立ち返る。一般に、カメラの分光感度特性は波長に対して急激に変化することは多くない。しかし、カラーフィルタ (図 2(d)) のような広帯域な光学フィルタについては、図 5(a) に示すように線形性自体を仮定できなくなることが想定される。この結果、分光感度の異なるカメラ間では、選ばれる最適なフィルタはそれぞれで異なると考え

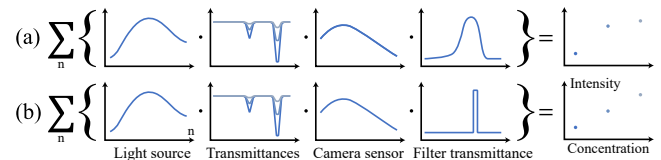


図 5 用いるフィルタによる応答の違い。ドット (\cdot) は要素同士の掛け算を表す。濃度の異なる媒質の透過特性は濃度と線形の関係にあるため、透過させるフィルタの形状によってカメラで観測される光強度は変わる。狭帯域の場合は線形の応答になると近似できるが、広帯域のフィルタを用いた場合は非線形の応答となる。

るべきである。したがって、MSC の分光感度にとって最適なフィルタ選択とならない可能性がある。一方、図 5(b) に示すように、フィルタがバンドパスフィルタのような狭帯域の透過特性を持つ場合、その範囲での感度特性はほぼ一定であるとみなすことができるため、応答の線形性は保たれ、フィルタは正しく選択されると予想される。任意形状のフィルタを用いることができない点は提案手法における制限の 1 つである。

光学フィルタの選択と係数決定のステップを分離した手法は以下の通りとなる。まず、式 (9) にしたがって適当な Z を用いることで候補となるフィルタの係数 d を推定し、 d の非ゼロ要素に対応するフィルタを使い MSC を作成する。ここで、推定された係数 d のうちゼロの要素を削除した係数ベクトルを d' とする。また、係数の非ゼロ要素に基づくフィルタ選択の後、作成する MSC によって取得される行列を Z' とする。これらを用いて、 $y = Z'd'$ のフィルタ係数 d' を OLS 回帰などによって推定する。

5. 実験

本章では、イチゴの糖分を特定物質と設定し、この分布定量を行う MSC を作成することで実環境下における提案手法の有効性を検証する。はじめに最適なフィルタの組み合わせを Adaptive lasso を用い選択する (5.1 節)。次に、選択したフィルタを用いた MSC を作成し、これを用いた糖度可視化を行う。 (5.2 節)

5.1 最適な光学フィルタの選択

イチゴのうち、品種エラン 22 個を対象とした光学フィルタ選択の実験を行う。スペクトル画像の撮影には、400~1,000nm を 128 バンドの分解能で測定可能な HSC (SOC710-VP, Surface optics corp.) を使用する。図 6 に示すように、対象の厚みの違いによる影響を取り除くため 5mm でスライスし、ハロゲンランプから光学フィルタを透過した光をイチゴに照射し、HSC によって観測する。ハイパースペクトル画像の各ピクセルを 1 サンプルとして扱うこととし、本実験では計 $M = 47,596$ サンプルから観測 x_{mn}, \tilde{x}_{mn} を構成した。本実験で候補とする市販の光学フィルタは

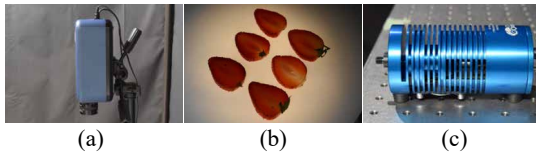


図 6 HSC を用いたイチゴの透過光の撮影環境. (b) 対象は拡散板の上に乗せられ, (c) 底面のハロゲンランプによって照らされる. 光源は拡散板で拡散し, 対象を通過した後 (a)HSC によって観測される

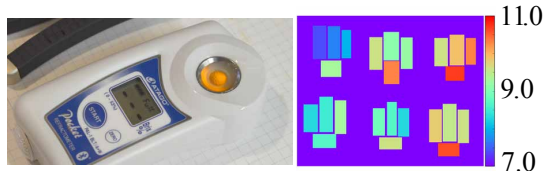


図 7 実測した糖度値の例. それぞれのイチゴを小片に切り, 左に示す屈折糖度計の金属部のポケットに搾汁する. 機器に表示された Brix 値を糖度とする.

Edmund optics 社のバンドパスフィルタとし, 公開されているカタログ値を用いてフィルタ行列 F を設定する. これを用いて仮想的にフィルタ画像 Z を生成する.

また, スライスしたイチゴは透過光を計測した後に適当な大きさにカットし, 各切片から搾汁した果汁の糖度を屈折糖度計 (PR-101 α , Atago Co. Ltd.) によって計測し, この値を計量対象とする. 図 7 に計測した糖度値の例を示す. \mathbf{y} は, Z の各ピクセルに対応する糖度を縦に配列したベクトルとする.

波長の選択数は, 式 (10) の adaptive lasso の持つ重み γ を 1.0 で固定し, α を変えることで調節する. 推定精度を以下の決定係数と呼ばれるスコア S で評価する.

$$S = 1 - \frac{\|\mathbf{y} - \hat{\mathbf{y}}\|_2^2}{\|\mathbf{y} - \bar{\mathbf{y}}\|_2^2}, \quad (12)$$

ここで, $\bar{\mathbf{y}}$ は \mathbf{y} の平均値を要素とするベクトル, \mathbf{y} は真値, $\hat{\mathbf{y}}$ は推定値を表す. 決定係数は, 推定量に対し全て平均値 $\bar{\mathbf{y}}$ を出力すると $S = 0$, 全て正答した場合には $S = 1.0$ となる. また評価スコアには 10-分割交差検証をおこなった結果を採用する.

実験結果について述べる. 図 8 は選択されたフィルタ数 $|d|_0$ と lasso および adaptive lasso を用いた際のスコアの関係を表す. adaptive lasso は全てのフィルタを用いた場合に $S = 0.469$, フィルタを 4 枚用いた場合に $S = 0.407$ となった. 大幅なフィルタ数の削減にもかかわらず, スコアを維持できていることが確認できる. 一方, lasso を用いた場合は適切なフィルタを選択することができなかった. 図 9 にフィルタを 4 枚使った場合の糖度分布の推定結果を示す. 見た目での結果からも, フィルタ数 4 枚程度で十分イチゴの糖度推定が可能であることが示された. 以上から, これら 4 枚を備える MSC の作成によって計測がより

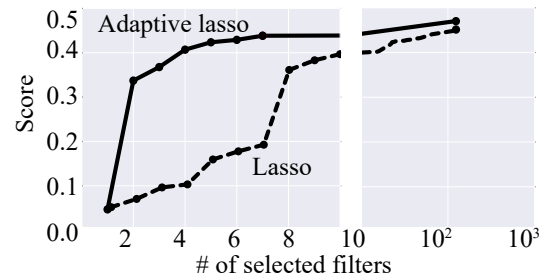


図 8 選択されたフィルタ数とスコアの関係. 実線は adaptive lasso の結果, 破線は lasso を用いた場合の結果を示す. Adaptive lasso を用いた場合のスコアは $|d|_0=119$ のとき 0.469, $|d|_0=4$ のとき 0.407 となった.

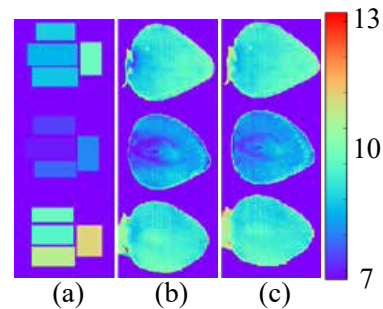


図 9 選択したフィルタを用いた糖度可視化のシミュレーション. (a) 糖度の実測値 (b) 光学フィルタ 119 枚を用いた糖度推定結果 (c) 提案するフィルタの選択モデルに対し, Adaptive lasso の変数選択法を適用した結果.

簡易になると期待される.

5.2 設計した MSC を用いた糖度可視化

本節では, 上の実験で選択したフィルタを用い, 簡易な MSC を作成する. また, 作成した MSC を用いた糖度の可視化実験を行う. それに際し, 4.4 節で述べたように各フィルタの係数の再推定をおこなう.

まず作成する MSC について述べる. 対象を拡散板の上に乗せ, 底面に設置した光源によって照らす. 光源は, フィルタを設置しその透過光を 90% の ND フィルタによって減光したものを用いる. 本実験では, フィルタはカメラの前ではなく光源の前に配置するが, 撮影される結果に違いは生じない. 対象を透過した光を冷却 CCD カメラ (Aspen CG6 LowNoise Apogee) によって観測する. 今回用いるフィルタは回転式の機構に取り付けられており, 手動で回転させることで各フィルタの切り替えを行い, 4 バンドのスペクトル画像の撮影を行う. 実際に使用したフィルタを図 10 に, 作成した MSC のシステムを図 11 に示す.

次にこの MSC によって取得される 4 バンドの画像 Z' から糖度を推定する. 本実験では, 光学フィルタ選択時に用いた HSC と, 作成した MSC のセンサの分光感度が異なるため, 4.4.2 節に基づく係数の再推定を行う. 係数の推定に用いる対象は 6 個のイチゴ (品種:エラン) とする. 上の実験と同様にスライスし, 作成した MSC によって撮



図 10 購入したフィルタ。中心波長-半値幅 の表記によって左から順に 600nm-50, 625nm-50, 675nm-25, 690nm-10 のフィルタを示す。

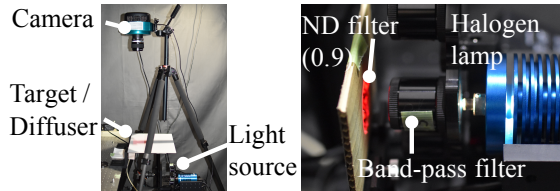


図 11 作成した MSC を用いたイチゴの撮影環境。光学フィルタを取り付けたりボルバーを回転させ、手動によって 4 チャンネルの撮影を行う。

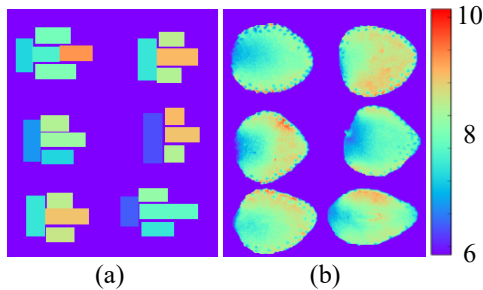


図 12 提案する MSC による糖分分布の推定結果。各 (a) 糖度の実測値 (b) 設計された MSC を用いた糖度推定の結果を示す。

影を行う。これまでと同様に多変量構成モデルを仮定し、最小二乗法により係数の推定を行う。なお、6 個体に対し Leave-one out Cross validation を用い計 6 回の係数推定を行い、最も高いスコア (式 (12)) を達成した係数を採用する。

次に新たなイチゴ 6 個体に対し図 11 と同様の実験環境によって 4 バンドの画像 Z' を取得し、これと推定された係数を用いた糖度推定を行なう。図 12(b) に推定された係数を用いた糖度分布の推定結果を示す。各ピクセルにおける糖度は \hat{y} は $z^T d'$ によって計算されたものである。また、図 12(a) に屈折糖度計を用いて計測された糖度を示す。式 (12) の評価方法によりスコア $S = 0.23$ を得た。スコアとしては低い値であったが、いちごの先端部の糖度が高い様子が可視化されており、いちごの一般的な特徴と傾向が一致していることから、結果の妥当性が示唆されていると考えられる。

6. 結論

本論文では、特定物質の分布イメージングを安価かつ高速に実現することを目的とし、市販の光学フィルタから最適な組を選択して MSC を構成する手法を提案した。提案手法は従来の研究とは異なり、市販のフィルタのみを用い

るという制約のもとで MSC の光学フィルタの選択を行う。これを実現するために、市販の光学フィルタがもつ分光透過率を考慮した多変量校正モデルを再定義した。また、HSC を利用して分光解析のためのサンプルを効率的に収集し、市販されている数百の光学フィルタによる撮影を仮想的に行うことで、特定物質の計量に最適な光学フィルタの組み合わせを選択する手法について述べた。実験ではイチゴの糖度推定というタスクに対し最適な光学フィルタの選定を行い、それらフィルタを用いた MSC の作成した。これを用いた糖度分布の可視化結果を示し、提案手法の有効性を検証した。

今後の課題として、他の物質の分布定量に対する提案手法の評価を行い、一般性を検証することが挙げられる。

謝辞

本研究は、ヤンマー株式会社およびヤンマーグリーンシステム株式会社との共同研究の成果である。

参考文献

- [1] : SpectroCam. <http://pixelteq.com/spectrocam/>.
- [2] Arakawa, M., Yamashita, Y. and Funatsu, K.: Genetic algorithm-based wavelength selection method for spectral calibration, *Journal of Chemometrics*, Vol. 25, No. 1, pp. 10–19 (online), DOI: 10.1002/cem.1339 (2011).
- [3] Balabin, R. M. and Smirnov, S. V.: Variable selection in near-infrared spectroscopy: Benchmarking of feature selection methods on biodiesel data, *Analytica Chimica Acta*, Vol. 692, No. 1-2, pp. 63–72 (online), DOI: 10.1016/j.aca.2011.03.006 (2011).
- [4] Bishop, C. M.: *Neural networks for pattern recognition*, Oxford university press (1995).
- [5] Gendrin, C., Roggo, Y. and Collet, C.: Pharmaceutical applications of vibrational chemical imaging and chemometrics: a review, *Journal of pharmaceutical and biomedical analysis*, Vol. 48, No. 3, pp. 533–553 (2008).
- [6] Goldberg, D. E.: *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1st edition (1989).
- [7] Gowen, A., O'Donnell, C., Cullen, P., Downey, G. and Frias, J.: Hyperspectral imaging—an emerging process analytical tool for food quality and safety control, *Trends in Food Science & Technology*, Vol. 18, No. 12, pp. 590–598 (2007).
- [8] H. Zou, T. H.: Regularization and variable selection via the elastic net, *Journal of the Royal Statistical Society B*, Vol. 67, pp. 301–320 (2005).
- [9] Hasegawa, K., Miyashita, Y. and Funatsu, K.: GA Strategy for Variable Selection in QSAR Studies: GA-Based PLS Analysis of Calcium Channel Antagonists, *Journal of Chemical Information and Computer Sciences*, Vol. 37, No. 2, pp. 306–310 (1997).
- [10] Hughes, G.: On the mean accuracy of statistical pattern recognizers, *IEEE transactions on information theory*, Vol. 14, No. 1, pp. 55–63 (1968).
- [11] Jiang, J.-H., Berry, R. J., Siesler, H. W. and Ozaki, Y.: Wavelength interval selection in multicomponent

- spectral analysis by moving window partial least-squares regression with applications to mid-infrared and near-infrared spectroscopic data, *Analytical chemistry*, Vol. 74, No. 14, pp. 3555–3565 (2002).
- [12] Jianqing, F. and Runze, L.: Variable Selection via Non-concave Penalized, Vol. 96, No. 456, pp. 1348–1360 (2001).
- [13] Lee, K. S., Cohen, W. B., Kennedy, R. E., Maiersperger, T. K. and Gower, S. T.: Hyperspectral versus multispectral data for estimating leaf area index in four different biomes, *Remote Sensing of Environment*, Vol. 91, No. 3–4, pp. 508–520 (online), DOI: 10.1016/j.rse.2004.04.010 (2004).
- [14] Mahesh, S., Manickavasagan, A., Jayas, D., Paliwal, J. and White, N.: Feasibility of near-infrared hyperspectral imaging to differentiate Canadian wheat classes, *Biosystems Engineering*, Vol. 101, No. 1, pp. 50–57 (2008).
- [15] M.Cope, D.: System for long-term measurement of cerebral blood and tissue oxygenation on newborn infants by near infra-red transillumination, *Medical and Biological Engineering and Computing*, Vol. 26, No. 3, pp. 289–294 (1988).
- [16] Natarajan, B. K.: Sparse approximate solutions to linear systems, *SIAM journal on computing*, Vol. 24, No. 2, pp. 227–234 (1995).
- [17] Plaza, A., Benediktsson, J. A., Boardman, J. W., Brazile, J., Bruzzone, L., Camps-Valls, G., Chanussot, J., Fauvel, M., Gamba, P., Gualtieri, A. et al.: Recent advances in techniques for hyperspectral image processing, *Remote sensing of environment*, Vol. 113, pp. S110–S122 (2009).
- [18] Preece, S. J. and Claridge, E.: Spectral filter optimization for the recovery of parameters which describe human skin, *IEEE transactions on pattern analysis and machine intelligence*, Vol. 26, No. 7, pp. 913–922 (2004).
- [19] Qian, Y., Zhou, J., Ye, M. and Wang, Q.: Structured sparse model based feature selection and classification for hyperspectral imagery, *Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International*, pp. 1771–1774 (2011).
- [20] Rodionova, O. Y., Houmøller, L. P., Pomerantsev, A. L., Geladi, P., Burger, J., Dorofeyev, V. L. and Arzamastsev, A. P.: NIR spectrometry for counterfeit drug detection: a feasibility study, *Analytica Chimica Acta*, Vol. 549, No. 1, pp. 151–158 (2005).
- [21] Schaepman, M. E., Ustin, S. L., Plaza, A. J., Painter, T. H., Verrelst, J. and Liang, S.: Earth system science related imaging spectroscopy—An assessment, *Remote Sensing of Environment*, Vol. 113, pp. S123–S137 (2009).
- [22] Styles, I. B., Calcagni, A., Claridge, E., Orihuela-Espina, F. and Gibson, J.: Quantitative analysis of multi-spectral fundus images, *Medical Image Analysis*, Vol. 10, No. 4, pp. 578–597 (2006).
- [23] Sugiyama, J.: Visualization of sugar content in the flesh of a melon by near-infrared imaging, *Journal of Agricultural and Food Chemistry*, Vol. 47, No. 7, pp. 2715–2718 (1999).
- [24] Thenkabail, P. S., Smith, R. B. and Pauw, E. D.: Hyperspectral Vegetation Indices and Their Relationships with Agricultural Crop Characteristics, *Remote Sensing of Environment*, Vol. 71, No. 2, pp. 158 – 182 (online), DOI: [http://dx.doi.org/10.1016/S0034-4257\(99\)00067-X](http://dx.doi.org/10.1016/S0034-4257(99)00067-X) (2000).
- [25] Tibshirani, R.: Regression Selection and Shrinkage via the Lasso (1996).
- [26] Xiaotong Shen, J. Y.: Adaptive Model Selection, *Journal of the American Statistical Association*, Vol. 97, No. 457, pp. 210–221 (online), available from <http://www.jstor.org/stable/3085776> (2002).
- [27] Zhang, X., Wen, J. and Zhao, D.: Band selection method for retrieving soil lead content with hyperspectral remote sensing data, *Remote Sensing*, International Society for Optics and Photonics, pp. 78311K–78311K (2010).
- [28] Zou, H.: The Adaptive Lasso and Its Oracle Properties, *Journal of the American Statistical Association*, Vol. 101, pp. 1418–1429 (2006).