

平成24年度 修士論文

画像を用いた野菜識別手法に関する研究
A Study on Vegetable Identification using Images

指導教授 渡辺 裕 教授

2013年3月15日

早稲田大学大学院 国際情報通信研究科
国際情報通信学専攻

4211A164-5

山 崎 太 一

Taichi YAMAZAKI

目次

第 1 章	序論	1
1.1	研究の背景	1
1.2	研究の目的	1
1.3	本論文の構成	2
第 2 章	食品管理システムの現状	3
2.1	はじめに	3
2.2	特殊なデバイスを用いた食品管理システム	3
2.3	モバイルデバイスを用いた画像による物体認識・識別手法	3
2.3.1	バーコードを用いた手法	3
2.3.2	局所特徴量を用いた手法	3
2.3.3	P型フーリエ記述子を用いた手法	4
2.4	おわりに	5
第 3 章	局所特徴量を用いた物体認識手法	6
3.1	はじめに	6
3.2	物体認識の要素技術	6
3.2.1	SIFT	6
3.2.2	SVM	11
3.3	Bag-of-Keypoints 法	17
3.4	おわりに	17
第 4 章	野菜識別手法	19
4.1	はじめに	19
4.2	形状情報	19
4.3	色情報	20
4.4	機械学習	20
4.5	提案手法の処理フロー	22
4.6	おわりに	22
第 5 章	提案手法の評価実験	24
5.1	はじめに	24
5.2	評価方法	24

5.3	提案手法の評価実験	24
5.3.1	P型フーリエ記述子のみを用いた野菜識別実験	24
5.3.2	P型フーリエ記述子と色情報を用いた野菜識別実験	25
5.4	おわりに	25
第6章	結論	30
6.1	総括	30
6.2	今後の課題	31
	謝辞	32
	参考文献	33
	図一覧	35
	表一覧	37
	研究業績	38

第1章

序論

1.1 研究の背景

近年、物品管理システムへの関心が高まっている。アメリカ最大小売り大手 Wal-Mart Stores, Inc. は、一部の商品に RFID(Radio Frequency IDentification)⁽²⁸⁾ タグを付け、物品管理システムの最適化を図っている。RFID 技術の導入により、従来のバーコードを用いた物品管理システムでは困難であった商品の棚在庫の最適化や盗難防止といった効果が見込まれている。

このような物品管理システムへの関心は物流だけではなく、家庭の冷蔵庫に対しても向けられており、物品および食品管理システムとして研究が進められている。既存の冷蔵庫では、冷蔵庫内に保管している食品を外部から確認できないという問題がある。そのため、スーパーマーケット等の外出先で、冷蔵庫内に保管してある食品まで購入してしまい、無駄な出費をしてしまう場合がある。メモを取り冷蔵庫内の食品を管理する方法が挙げられるが、メモを取る時間やメモそのものの共有方法に問題がある。

そのため、近年、RFID 技術を用いた冷蔵庫内の食品管理システムが提案されている。これは、冷蔵庫から食品を出し入れする際には、RFID リーダーから食品に付けた RFID タグの情報を取得し、食品管理を行うシステムである。しかし、RFID 技術を用いた物品管理システムでは、RFID タグや RFID リーダー等の特殊なデバイスを導入する必要がある、それらの導入コストが問題である。低コストな食品管理システムを実現するために、従来の RFID の代わりとして、身近にあるデバイスの利用が考えられる。近年、iPhone⁽¹⁾ や Android⁽²⁾ といったスマートフォンは、LTE(Long Term Evolution) などの高速通信規格によって普及が加速しており、一般に身近なものになりつつある。また性能の向上により、より複雑な処理が可能となってきており、また特に画像処理では、OpenCV⁽³⁾ をはじめ様々な開発環境が整いつつある。そこで、本研究では、モバイルデバイス上での画像による食品管理システムの実現を目指し、その核となる画像による認識・識別手法について検討を行う。

1.2 研究の目的

既存の画像による物体の認識・識別手法として、局所特徴量を用いた手法がある。この手法は、画像から局所特徴量を抽出し、代表ベクトルへ投票した結果を機械学習することでパ

ターン認識を行う。しかし、上記の手法は、特定物体の認識や識別の意味合いが強く、個体差のある物体に対しては有効ではない。食品管理においては野菜や魚といった個体差のある物体の識別を行う必要があり、上記の手法は適していないと考えられる。個体差のある物体の識別には、P型フーリエ記述子⁽²⁵⁾を用いた手法がある。森本ら⁽²⁷⁾は、個体差のある葉の形状に対してP型フーリエ記述子を適用し、その有効性を示した。ただし、識別が困難な例として、似たような形状の葉が挙げられている。以上から、個体差のある物体の識別には、P型フーリエ記述子による形状情報と似た形状比較の回避が有効であると考えられる。

そこで、本研究では、個体差のある食品の識別を可能とするために、P型フーリエ記述子と色情報によるクラス分けを適応させ、画像による野菜識別手法を実現する。ここで、食品には、野菜、魚、肉など様々なものが挙げられるが、まず購入頻度から野菜に限定した。また、識別手法は、モバイルデバイス上での使用を想定しているため、識別対象画像に対して、回転・拡大縮小・平行移動・照明変化にロバストである事を条件とした。識別対象画像は、予め手動で背景分離した画像を用意した。野菜の形状に対してP型フーリエ記述子を適用するためには、前景と背景の分離を行う必要である。手法としては、背景差分法やGraph Cut⁽⁷⁾が挙げられる。しかし、これらの手法は、影や画像のぶれの影響を受け、本来抽出すべき輪郭線を得られるとは限らない。そのため、それらの手法から得られた輪郭線では、P型フーリエ記述子の野菜形状に対する有効性を評価するのは難しいと考えられる。以上から、本研究では、予め手動で前景と背景を分離した画像を用意し、その画像に対してP型フーリエ記述子を適用した。

1.3 本論文の構成

以下に本章以降の構成を示す。

第1章 本章であり、研究背景およびその目的について述べている。

第2章 冷蔵庫内の食品管理システムの現状について述べる。

第3章 既存の画像を用いた物体認識手法について述べる。はじめに物体認識手法の要素技術である局所特徴量と機械学習について述べる。次にそれらを用いた物体認識手法について述べる。

第4章 提案手法である野菜識別手法について述べる。はじめに提案手法で用いる形状情報、色情報、機械学習について述べ、次にそれらを用いた提案手法の処理フローについて述べる。

第5章 第4章で提案した提案手法の評価実験方法と結果について述べる。

第6章 本研究の総括と今後の課題について述べる。

第2章

食品管理システムの現状

2.1 はじめに

本章では、食品管理システムの現状について、特殊なデバイスを用いた食品管理システムとモバイルデバイスを用いた画像による認識・識別手法について述べる。

2.2 特殊なデバイスを用いた食品管理システム

原田ら⁽²⁸⁾は、RFID技術を用いた食品管理システムを提案している。このシステムでは、RFIDタグは一つの食品につき一つ付属され、そのタグには商品番号が保存されている。ユーザの持つユーザタグからは、販売側DBサーバから購入履歴を取得し、商品の詳細情報を取得することができる。図2-1は、システムの全体を表している。

この手法では、RFIDタグやRFIDリーダーといったデバイスを導入する必要があり、それらのコストが問題である。

2.3 モバイルデバイスを用いた画像による物体認識・識別手法

2.3.1 バーコードを用いた手法

はてなモノリス⁽⁴⁾は、株式会社はてなが提供するスマートフォン向けの物品管理アプリケーションである。商品のバーコードを認識・識別し、そのバーコードの情報から間接的に物品情報を取得することで、物品管理を行う。しかし、この手法ではバーコードが汚れていたり、折れ曲がっている等、読み取り不可能な状態では使用できないという問題がある。

2.3.2 局所特徴量を用いた手法

画像に写っている物体を直接認識・識別する手法には、局所特徴量を用いた手法がある。Goggles⁽⁵⁾や望月ら⁽²⁶⁾の提案する物品管理システムでは、この技術が使われている。GogglesはGoogleの提供するスマートフォン向けの物体認識アプリケーションである。このアプリケーションは、カメラで撮影した物体を認識し、それに関連するWeb上の情報を表示する。しかし認識される物体は、文字、バーコード、ロゴ、絵画、書籍、ランドマークといっ

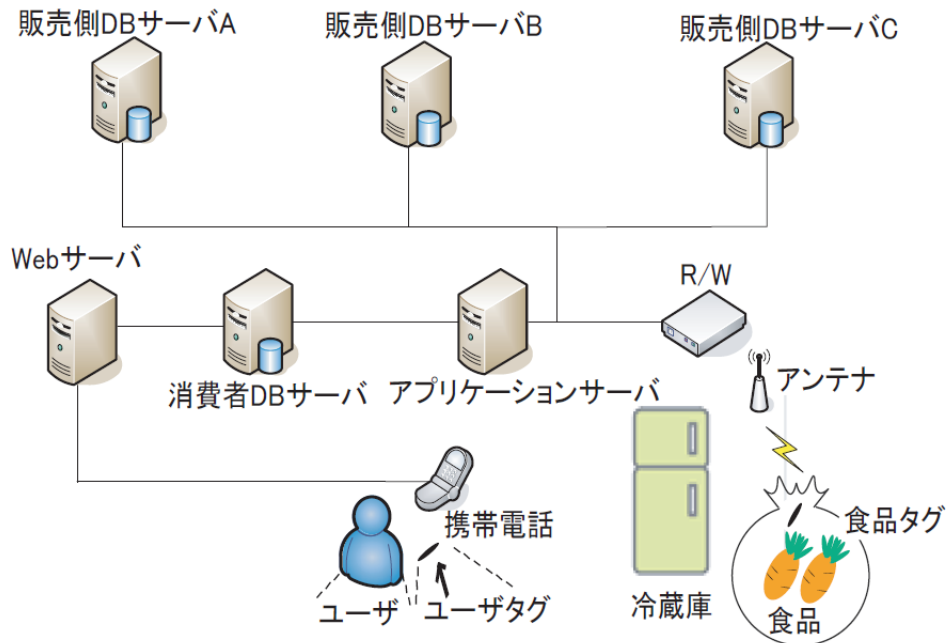


図 2-1: 冷蔵庫管理システム全体図
(文献⁽²⁹⁾より引用)

たものであり、特定物体認識の意味合いが強く、野菜などの個体差のある物体の認識は出来ない。

望月らは物体認識して得られた情報から物体を特定し、物品管理を行うシステムを提案している。望月らの認識手法では、動画像から局所特徴量を取得し、投票により物体を特定する。図2-2は、提案システムの動作例を表している。望月らの提案システムでは、図2-2のように商品の文字やパッケージ絵を認識することで、物品管理を行う。Gogglesと同様に、特定物体認識の意味合いが強く、個体差のある物体の管理は困難であると考えられる。

2.3.3 P型フーリエ記述子を用いた手法

森元ら⁽²⁷⁾は、P型フーリエ記述子を葉形状に適用し、植物を識別する手法を提案している。森元らの手法では、P型フーリエ記述子を葉形状に適用し、得られた特徴ベクトルから代表となるベクトルを辞書に登録する。そして、識別したい葉と辞書に登録した代表ベクトルのユークリッド距離が最小となるものを尤もらしい葉とする手法である。森元らは評価実験にて提案手法の有効性を示している。しかし、似たような形状は識別できないという問題がある。

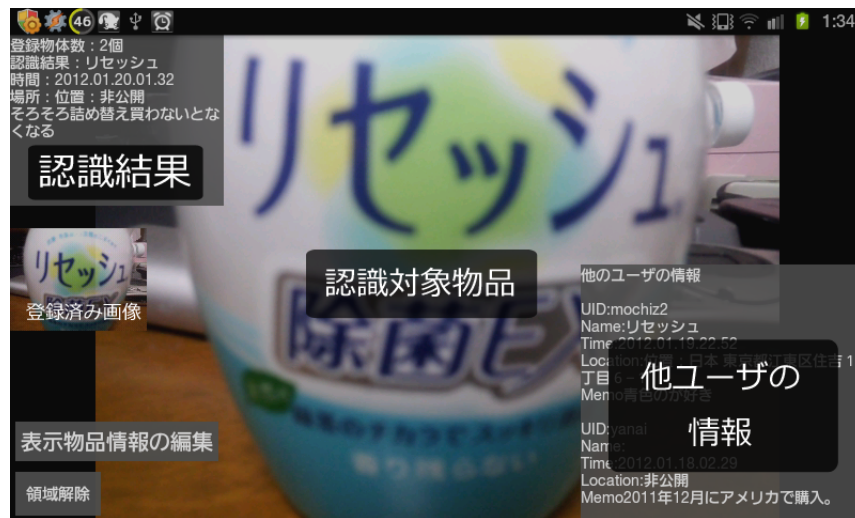


図 2-2: 望月らの提案システムの動作イメージ
(文献⁽²⁶⁾より引用)

2.4 おわりに

本章では、物品管理システムの現状について述べた。特殊なデバイスを用いたシステムと画像による物体認識・識別を用いた手法について述べ、各システムおよび手法の問題点を述べた。

第3章

局所特徴量を用いた物体認識手法

3.1 はじめに

本章では、まず物体認識手法の要素技術について述べる。次に、それらの要素技術を用いた物体認識手法について述べる。

3.2 物体認識の要素技術

局所特徴量を用いた物体認識手法は、大きく分けて局所特徴量の抽出とパターン認識の二つの技術からなる。局所特徴量は、画像への制約を取り払うために、回転・拡大・平均移動・照明変化に対してロバストであることが求められる。それらの特徴を持つ特徴量を記述する手法には、SIFT(Scale Invariant Feature Transform)⁽⁸⁾ や SIFT を積分画像で高速化した SURF(Speeded Up Robust Features)⁽⁹⁾, SIFT 特徴量に主成分分析を適用しより頑健で安定化した特徴量抽出を可能とする PCA-SIFT⁽¹⁰⁾ などがある。また、局所特徴量を学習する機械学習には、Adaboost⁽¹⁸⁾ や SVM(Support Vector Machine)⁽¹¹⁾⁽¹²⁾⁽¹⁴⁾⁽¹⁵⁾⁽¹⁶⁾⁽¹⁷⁾ がある。以下では、一般的に局所特徴量を用いた物体認識で使われる SIFT と SVM について、詳細を述べる。

3.2.1 SIFT

SIFT とは、スケール変化・回転変化・照明変化に頑強な特徴点(キーポイント)を求める手法である。SIFT の処理は大まかに以下の4つからなる。

STEP1. キーポイントとスケールの検出

STEP2. キーポイントのローカライズ

STEP3. オリエンテーションの算出

STEP4. 特徴量の記述

STEP1, STEP2 ではキーポイントの検出を行う。STEP1 では、スケールスペースを用いた極値検出を行うことで、スケール不変なキーポイントを検出する。STEP2 では、ロー



図 3-1: 画像から求められる SIFT 特徴量の例

カライズによるキーポイント候補の削除をすることで、STEP1 で検出したキーポイントの絞り込みを行う。

STEP3, STEP4 は、上記のキーポイントの検出後に行われる処理で、特徴量の記述を行う。STEP3 では、基準となるオリエンテーションを求める。STEP4 では、STEP3 で求めたオリエンテーションを用いることで、回転不変な特徴ベクトルを算出する。さらに、得られた特徴ベクトルを正規化することで、照明変化に対して影響の少ない特徴量を求める。以上の STEP1, STEP2, STEP3, STEP4 の処理を行うことで、SIFT はスケール変化・回転変化・照明変化に頑強な特徴量の記述を実現する。

図 3-1 は、画像に対し SIFT 特徴量を求め、重畳したものである。矢印の方向がオリエンテーションの向きを、矢印の長さがスケールを表す。

以下に SIFT の上記 STEP1, STEP2, STEP3, STEP4 の処理内容の詳細を述べる。

STEP1. キーポイントとスケールの検出

特徴点のスケール検索には Laplacian-of-Gaussian(LoG) を用いた手法がある。式 (3-1) で表される LoG オペレータ (図 3-2) を画像に適用し、オペレータの出力値が極大となる σ をキーポイントのスケールとする。 σ はガウシアンフィルタのスケール、 x と y は注目画素か

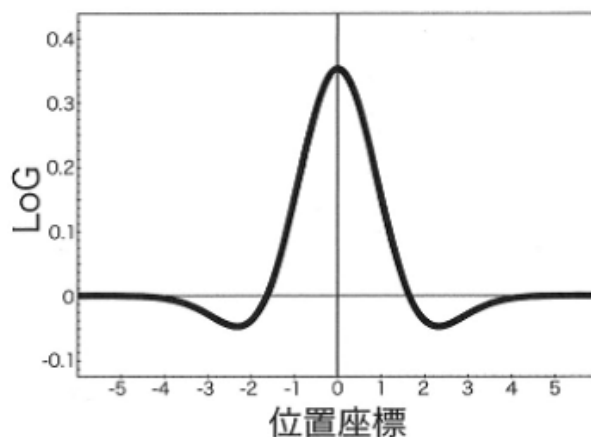


図 3-2: LoG オペレータ例
(文献⁽⁶⁾より引用)

らの距離である.

$$LoG = f(\sigma) = -\frac{x^2 + y^2 - 2\sigma^2}{2\pi\sigma^6} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (3.1)$$

しかし, LoG によるスケール探索は, 特徴点の検出とスケール探索を別々に行うため計算コストが高いという問題がある. そこで SIFT では特徴点の検出とスケール探索を同時に行う Difference-of-Gaussian(DoG) を用いることで, 計算コスト削減を実現する.

DoG によるキーポイント検出とスケール探索では, キーポイントの候補点は, スケールの異なるガウス関数 $G(x, y, \sigma)$ (式 (3.3)) と入力画像 $I(u, v)$ を畳み込んだ平滑化画像 $L(u, v, \sigma)$ (式 (3.2)) の差分 (DoG 画像) から求める.

$$L(u, v, \sigma) = G(x, y, \sigma) * I(u, v) \quad (3.2)$$

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (3.3)$$

DoG 画像は式 (3.4) より求められる.

$$\begin{aligned} D(u, v, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(u, v) \\ &= L(u, v, k\sigma) - L(u, v, \sigma) \end{aligned} \quad (3.4)$$

具体的には, まず初期スケール σ_0 を k 倍ずつ大きくした異なるスケールの平滑化画像を生成し, その差分から DoG 画像を求める. 次に DoG 画像三枚を用いて, 注目画素とその周辺 26 近傍から極値を求め, キーポイントの候補を検出する. 極値検出はスケールの小さい DoG 画像から順に行い, 一度極値が検出された注目画素は, より大きなスケールで極値が検出されてもキーポイント候補点から除外する. 周辺 26 近傍とは, DoG 画像 $D(u, v, k^\alpha\sigma_0)$ 上の注目画素に対して, 注目画素周辺の 8 近傍と DoG 画像 $D(u, v, k^{\alpha-1}\sigma_0)$ と $D(u, v, k^{\alpha+1}\sigma_0)$ の周辺 9 近傍である. 図 3-3 は注目画素とその周辺 26 近傍を表した図である. 上記の DoG

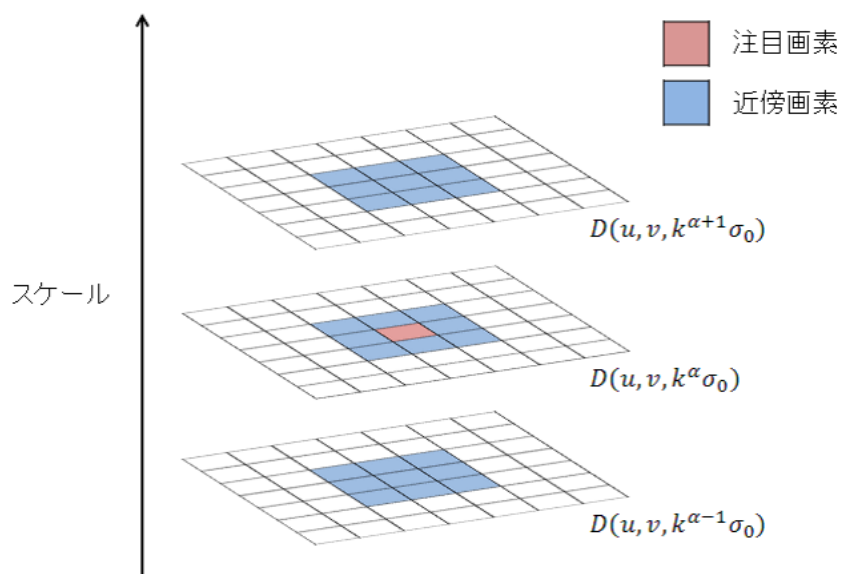


図 3-3: DoG 画像を用いた 26 近傍による極値の検出

画像による極値検出では、スケールの増加に伴うガウシアンフィルタの計算コストの増加という問題がある。この問題に対して SIFT では、ダウンサンプリングを行うことでガウシアンフィルタの計算コストを削減する。

STEP2. キーポイントのローカライズ

(1) で検出されたキーポイントの候補の中には、DoG 出力値が小さい点やエッジ上の点がノイズや開口問題の影響を受けやすいという問題がある。SIFT ではヘッセ行列によるエッジ検出を行い、適切でないキーポイントを削除することで、より安定したキーポイントを求める。ヘッセ行列 H は、式 (3.5) で求められる。

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad (3.5)$$

D_{xx} , D_{xy} , D_{yy} は DoG 出力値の二次微分から得られる。ヘッセ行列 H の二つの固有値 α , β ($\alpha > \beta$) から、キーポイントがフラットな領域上に存在するのか、エッジ上の点なのか、コーナ点なのかを判別する。 α , β の関係は以下の通りである。

a, b ともに小さい : フラット

a, b ともに大きい : コーナー

$a \gg b$ or $b \gg a$: エッジ

上記の判別処理では固有値を直接求める必要は無く、エッジ検出は α , β の比率から求めることが可能である。 α , β の関係は、ヘッセ行列の対角和 $Tr(H)$ と行列式 $det(H)$ から以下のように求まる。

$$Tr(H) = D_{xx} + D_{yy} = \alpha + \beta \quad (3.6)$$

$$Det(H) = D_{xx}D_{yy} - (D_{xy})^2 = \alpha\beta \quad (3.7)$$

α , β の比率 $\gamma (= \frac{\alpha}{\beta})$ は、以下の式から近似値として求められる。

$$\frac{Tr(H)^2}{Det(H)} = \frac{(\alpha + \beta)^2}{\alpha\beta} = \frac{(\gamma\beta + \beta)^2}{\gamma\beta^2} = \frac{(\gamma + 1)^2}{\gamma} (\approx \gamma) \quad (3.8)$$

式 (3.8) に閾値 γ_{th} を代入した判別式は、

$$\frac{Tr(H)^2}{Det(H)} < \frac{(\gamma_{th} + 1)^2}{\gamma_{th}} \quad (3.9)$$

である。式 (3.9) を満たさないキーポイントはエッジ上の点として候補から削除される。次にキーポイント候補点のサブピクセル推定による位置とスケールの補正について述べる。

SIFT では、DoG 画像の差分から検出したキーポイント候補点に対して、サブピクセル推定による高精度な推定を行う。推定には、ある点 $\mathbf{x} = (x, y, \sigma)^T$ での DoG 関数 $D(\mathbf{x})$ をテーラー展開した式 (3.10) を用いる。

$$D(\mathbf{x}) = D + \frac{\partial D}{\partial \mathbf{x}} \mathbf{x} + \frac{1}{2} \mathbf{x}^T \frac{\partial^2 D}{\partial \mathbf{x}^2} \mathbf{x} \quad (3.10)$$

式 (3.10) について \mathbf{x} に関する偏導関数を求め、0 とすると次式が得られる。

$$\frac{\partial D}{\partial \mathbf{x}} + \frac{\partial^2 D}{\partial \mathbf{x}^2} \hat{\mathbf{x}} = 0 \quad (3.11)$$

このとき $\hat{\mathbf{x}}$ は、サブピクセル推定により位置とスケールが補正されたキーポイント候補点を表している。式 (3.11) を変形した式 (3.12) を用いることで、 $\hat{\mathbf{x}}$ を算出する。

$$\hat{\mathbf{x}} = - \left(\frac{\partial^2 D}{\partial \mathbf{x}^2} \right)^{-1} \frac{\partial D}{\partial \mathbf{x}} \quad (3.12)$$

また、DoG 出力値の小さいキーポイントは、ノイズの影響を受けやすいという問題がある。そこで SIFT では、サブピクセル推定した候補点の DoG 出力値に閾値を設け、キーポイントの絞り込みを行う。

STEP3. オリエンテーションの算出

オリエンテーションとはキーポイントの方向を表し、このオリエンテーションを用いることで SIFT は回転に不変な特徴量を記述することが可能となる。オリエンテーションの算出には、まずキーポイント周辺の勾配強度 $m(u, v)$ と勾配方向 $\theta(u, v)$ を次式から求める。

$$m(u, v) = \sqrt{f_u(u, v)^2 + f_v(u, v)^2} \quad (3.13)$$

$$\theta(u, v) = \arctan\left(\frac{f_v(u, v)}{f_u(u, v)}\right) \quad (3.14)$$

$$f_u(u, v) = L(u + 1, v) - L(u - 1, v) \quad (3.15)$$

$$f_v(u, v) = L(u, v + 1) - L(u, v - 1) \quad (3.16)$$

上記の式から得られる勾配強度と勾配方向から重み付きヒストグラム $h(\theta')$ を次式により作成する。

$$h(\theta') = \sum_x \sum_y w(x, y) \cdot \delta[\theta', \theta(x, y)] \quad (3.17)$$

$$w(x, y) = G(x, y, \sigma) \cdot m(x, y) \quad (3.18)$$

$h(\theta')$ は、勾配方向を 36 方向に量子化したヒストグラムである。 δ は *Kronecker* のデルタ関数であり、 $\delta[\theta', \theta(x, y)]$ は、 θ' と $\theta(x, y)$ が同じ勾配方向ならば 1 を返す関数である。点 (x, y) にガウス窓かけた $w(x, y)$ を用いることで、キーポイントに近い点ほど重みの付いたヒストグラムが作成される。オリエンテーションは、ヒストグラム $h(\theta')$ の最大値の 80% 以上となる勾配方向に割り当てられる。

STEP4. 特徴量の記述

特徴量の記述では、キーポイント周辺の勾配情報を用いる。キーポイント周辺とは、DoG 処理で得られたスケールを半径とする円の領域内を指す。この領域を図 3-4 のように 4×4 の計 16 ブロックに分割し、STEP3 で求めたオリエンテーションを基軸に各ブロックの勾配方向を算出する。勾配情報は 8 方向のヒストグラムから構成される。オリエンテーションを基軸に勾配方向を算出するため、このようにして得られた特徴量は回転不変である。特徴量は、ブロックごとに 8 方向のヒストグラムの情報を記述するため、 $16 \times 8 = 128$ 次元の特徴ベクトルとなる。

さらに、上記の特徴ベクトルをベクトルの長さで正規化することで、SIFT は照明変化にも頑健な特徴量の記述を可能とする。

3.2.2 SVM

SVM は、データを 2 つのクラスに分類するパターン認識手法の一つである。パターン認識とは、既にクラス分けされたデータを用いて、観測された未知データのクラスを予測する

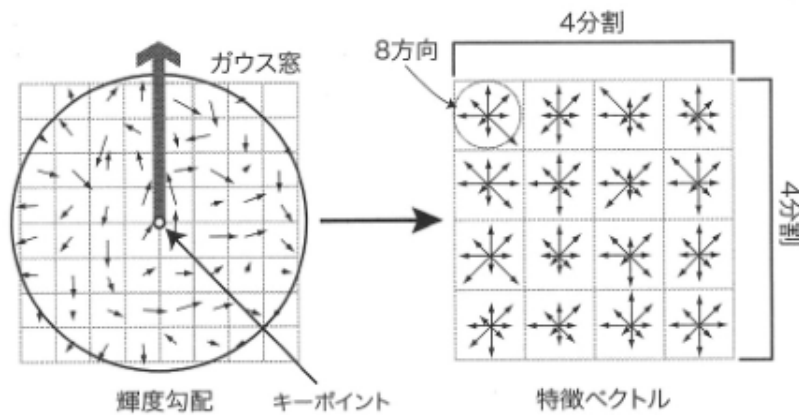


図 3-4: キーポイント周辺の各ブロックの勾配方向
(文献⁽⁶⁾より引用)

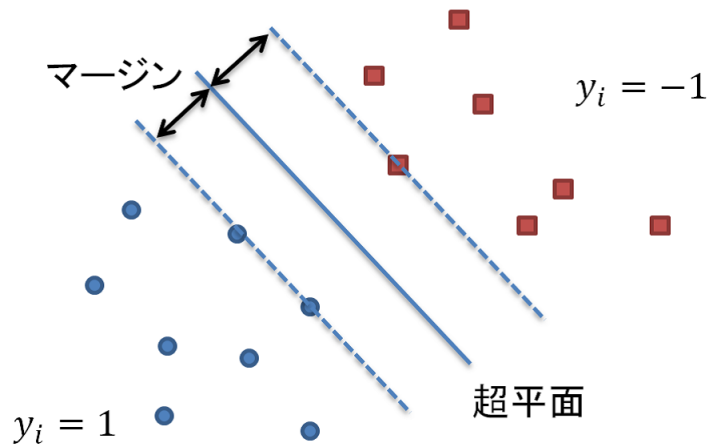


図 3-5: 学習データが線形分離可能な場合

技術である。SVM は、数字認識⁽¹¹⁾、テキスト分類⁽¹⁹⁾、顔検出⁽²⁰⁾ など幅広く用いられており、パターン認識性能の優れた学習モデルの一つである。SVM では、マージン最大化に基づいてクラス判別関数を決定し、学習データに対して識別性能が特化してしまう過学習を防いでいる。以下では SVM の数理について述べる。まず、学習データが線形分離可能な場合と線形分離不可能な場合について述べる。最後にカーネルトリックを用いた非線形分離について述べる。

学習データが線形分離可能な場合

SVM は、学習データからクラスの判別関数を算出する。ここで学習データは、次のように表される。

$$\text{入力ベクトル} : \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n \quad \mathbf{x}_i \in \mathbf{R}^d \tag{3-19}$$

$$\text{目標値} : y_i \in \{1, -1\} \quad (3.20)$$

$$i = 1, 2, \dots, N \quad (3.21)$$

目標値とは，入力データの属するクラスを数値で表したものである．例えば，図3-5では●は1，■は-1というように2値でクラス分けされる．上記の学習データから学習する判別関数は，

$$f(\mathbf{x}) = \text{sign}(\mathbf{w}^T \mathbf{x} + b) \quad (3.22)$$

である．ただし，

$$\text{sign}(u) = \begin{cases} 1 & u > 0 \\ -1 & u \leq 0 \end{cases} \quad (3.23)$$

である．また $\mathbf{w}^T \mathbf{x} + b$ は，超平面である．つまり，図3-5の●，■に当たるデータをそれぞれ1，-1と判別する超平面を求めることになる．しかし，その様な超平面は，一様に定まらない．そこでSVMでは，マージン最大化という基準に基づいて超平面を決定する．マージンとは，超平面と最も近い学習データとの距離である．

学習データが線形分離可能であると仮定すると，未知のデータ \mathbf{x} は，判別式 (3.22) の符号に応じてクラスに分類される．符号とクラスの対応を， $f(\mathbf{x}_i) > 0$ の時 $y=1$ ， $f(\mathbf{x}_i) < 0$ の時 $y = -1$ とすると以下の式が成り立つ．

$$y_i f(\mathbf{x}_i) > 0 \quad (3.24)$$

ここで，超平面と点の距離は次式で表される．

$$\frac{|\mathbf{w}^T \mathbf{x}_i + b|}{\|\mathbf{w}\|} \quad (3.25)$$

$\|\mathbf{w}\|$ は，2ノルムである．式 (3.24) と式 (3.25) より，点と超平面の距離は，

$$\frac{y_i (\mathbf{w}^T \mathbf{x}_i + b)}{\|\mathbf{w}\|} \quad (3.26)$$

となる．ここで，超平面に最も近い点について，

$$y_i (\mathbf{w}^T \mathbf{x}_i + b) = 1 \quad (3.27)$$

が成り立つことが分かっている．そのため，サポートベクターと超平面との距離は，

$$\frac{1}{\|\mathbf{w}\|} \quad (3.28)$$

と表される．また，全ての点に対して次式が成り立つ．

$$y_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1 \quad (3.29)$$

式(3.29)は、全ての点に対する制約と考えるとSVMは以下の最適化問題として表される。

$$\max_{\mathbf{w}, b} \quad \frac{1}{\|\mathbf{w}\|} \quad (3.30)$$

$$s.t. \quad y_i(\mathbf{w}^T \mathbf{x}_i) + b \geq 1 \quad (3.31)$$

また、この $1/\|\mathbf{w}\|$ の最大化問題は、 $\|\mathbf{w}\|^2$ の最小化問題と等価であることから、

$$\min_{\mathbf{w}, b} \quad \frac{1}{2} \|\mathbf{w}\|^2 \quad (3.32)$$

$$s.t. \quad y_i(\mathbf{w}^T \mathbf{x}_i) + b \geq 1 \quad (3.33)$$

と書き換えられる。ただし、 $1/2$ という係数は、後の計算を楽にするために導入したものである。ここで、この最適化問題はそのままでは複雑であるため、ラグランジュ乗数 a_n を導入し、解きやすい形に変形する。すると以下の式を \mathbf{a} について最大化すればよいことになる。

$$L(\mathbf{a}) = \sum_{n=1}^N a_n - \frac{1}{2} \sum_{n=1}^N \sum_{m=1}^M a_n a_m y_n y_m \mathbf{x}_n^T \mathbf{x}_m \quad (3.34)$$

ただし、 \mathbf{a} は以下の制約条件を満たす。

$$a_n \geq 0 \quad n = 1, 2, \dots, N \quad (3.35)$$

$$\sum_{n=1}^N a_n y_n = 0 \quad (3.36)$$

学習データが線形分離不可能な場合

学習データが線形分離不可能な場合では、上記の方法は使えない。そこで、データ \mathbf{x}_i に対してスラック変数 ξ_i を導入し、その分だけ超平面からはみ出すことを許す。その代わりはみ出した分に対してはペナルティを与える。図3-6の○で囲われた点は、はみ出しによるペナルティを受ける点である。

C をペナルティの重み定数とすると、SVMの最適化問題は、

$$\min_{\mathbf{w}, b} \quad \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^N \xi_i \quad (3.37)$$

$$s.t. \quad y_i(\mathbf{w}^T \mathbf{x}_i) + b \geq 1 - \xi_i \quad (3.38)$$

$$\xi_i \geq 0 \quad (3.39)$$

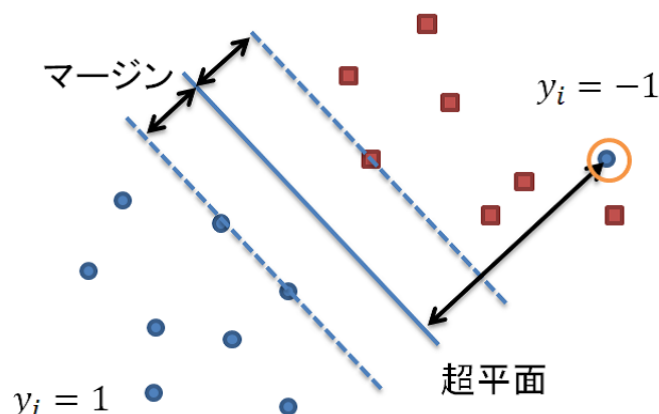


図 3-6: 学習データが線形分離不可能な場合

と表される． C を無限大にすると線形分離可能な場合の SVM と同じ意味を表す．線形分離可能な場合の時と同様にラグランジュ乗数を導入すると，以下の式を \mathbf{a} について最大化すればよいことになる．

$$L(\mathbf{a}) = \sum_{n=1}^N a_n - \frac{1}{2} \sum_{n=1}^N \sum_{m=1}^M a_n a_m y_n y_m \mathbf{x}_n^T \mathbf{x}_m \quad (3.40)$$

ただし，制約条件は，

$$0 \leq a_n \leq C \quad n = 1, 2, \dots, N \quad (3.41)$$

$$\sum_{n=1}^N a_n y_n = 0 \quad (3.42)$$

である．結果として，式 (3.34) と式 (3.40) から学習データが線形分離可能な場合と線形分離不可能な場合の目的関数は同一であることが分かる．

式 (3.40) と式 (3.40) は，最急降下法⁽¹⁶⁾ や SMO(Sequential Minimal Optimisation) 法⁽²³⁾ により効率的に最適解を求められる．また，線形分離可能な場合の SVM をハードマージン SVM，線形分離不可能な場合の SVM をソフトマージン SVM と言う．

カーネルトリック

ソフトマージン SVM を用いることで，線形分離不可能な学習データから識別平面を構成できたとしても，学習データが複雑で非線形な構造をしているとすると，必ずしも良い識別を行えるとは限らない．そこで，非線形なデータ構造に対応するための方法として，次式のようにデータ \mathbf{x} を非線形変換して，その特徴空間で線形判別をすることが考えられる．

$$\phi : \mathbf{R}^n \rightarrow \mathbf{R}^q \quad (n < q) \quad (3.43)$$

このような写像を行うことにより、写像先の空間 \mathcal{R}^q で線形識別を行うことが、元の空間では非線形識別を行っているのと同じことになる。また、この写像を SVM に導入し \mathbf{x} を $\phi(\mathbf{x})$ に変換すると、目的関数は

$$L(\mathbf{a}) = \sum_{n=1}^N a_n - \frac{1}{2} \sum_{n=1}^N \sum_{m=1}^M a_n a_m y_n y_m \phi(\mathbf{x}_n)^T \phi(\mathbf{x}_m) \quad (3.44)$$

となり、内積のみに依存した形になる。ここで、もし非線形に写像した空間での2つの要素 $\phi(\mathbf{x}_n)$ と $\phi(\mathbf{x}_m)$ の内積が、

$$\phi(\mathbf{x}_n)^T \phi(\mathbf{x}_m) = K(\mathbf{x}_n, \mathbf{x}_m) \quad (3.45)$$

のように \mathbf{x}_m と \mathbf{x}_n のみから計算できれば、 $\phi(\mathbf{x}_n)$ と $\phi(\mathbf{x}_m)$ を計算せずとも、 $K(\mathbf{x}_n, \mathbf{x}_m)$ から最適な非線形写像を構成出来る。K はカーネル関数である。このように、高次元に写像しながら、実際には写像された空間での計算を避け、カーネルの計算のみで最適な判別関数を構成するテクニックがカーネルトリックである。

ここで、式 (3.45) のように表せるカーネルを、半正定値カーネルと呼ぶ。 $K(\mathbf{x}_n, \mathbf{x}_m)$ が \mathbf{R}^n 上の半正定値カーネルであるとは、以下の二つの条件を満たすことである。

対称性:

$$K(\mathbf{x}_n, \mathbf{x}_m) = K(\mathbf{x}_m, \mathbf{x}_n) \quad (3.46)$$

正定値性:

任意の実数 a_1, \dots, a_N に対して、

$$\sum_{n=1}^N \sum_{m=1}^M a_n a_m K(\mathbf{x}_n, \mathbf{x}_m) \geq 0 \quad (3.47)$$

また、上記の条件を満たすカーネル関数には、計算が容易なものとして、

多項式カーネル:

$$K(\mathbf{x}_n, \mathbf{x}_m) = (1 + \mathbf{x}_n^T \mathbf{x}_m)^d \quad (3.48)$$

ガウシアンカーネル:

$$K(\mathbf{x}_n, \mathbf{x}_m) = \exp\left(\frac{-\|\mathbf{x}_n - \mathbf{x}_m\|^2}{2\delta^2}\right) \quad (3.49)$$

シグモイドカーネル:

$$K(\mathbf{x}_n, \mathbf{x}_m) = \tanh(a\mathbf{x}_n^T \mathbf{x}_m - b) \quad (3.50)$$

がある。

3.3 Bag-of-Keypoints 法

BoK(Bag-of-Keypoints) 法⁽²¹⁾は、言語処理で使われている bag-of-words⁽²²⁾の考え方を画像に適応させた手法である。bag-of-words では、文章は語順を無視した単語の集合として考える。このアイデアを画像に適応したものが BoK 法であり、画像は位置を無視した局所特徴量の集合として考える。BoK 法の処理フローは以下の通りである。

STEP1. 特徴量抽出

入力画像から全てのキーポイントの局所特徴量を抽出する。

STEP2. クラスタリング

抽出した局所特徴量をクラスタリングし、代表ベクトルを求める。クラスタリングには k-means 法が一般的に使われる。k-means 法は、クラスタの中心と各データとのユークリッド距離が最小になるようなクラスタ中心を逐次的に求めることで、クラスタリングを行う手法である。

STEP3. visual word histogram の作成

visual word は STEP2 で求めた代表ベクトルである。visual word の頻出頻度を visual word histogram と呼び、BoK 法では特徴ベクトルとなる。

STEP4. 学習と認識

STEP4 で求めた特徴ベクトルを機械学習により、学習または物体認識を行う。

図 3-7 は BoK 法の処理フローを図で表したものである。

3.4 おわりに

本章では、物体認識手法の要素技術である局所特徴量と機械学習について述べた。はじめに局所特徴量である SIFT について、キーポイントとスケールの検出、キーポイントのローカライズ、オリエンテーションの算出、特徴量の記述に分けて述べた。次に機械学習である SVM について、ハードマージン SVM、ソフトマージン SVM、カーネルトリックについて述べた。最後にそれらを組み合わせた物体認識手法である BoK 法について述べた。

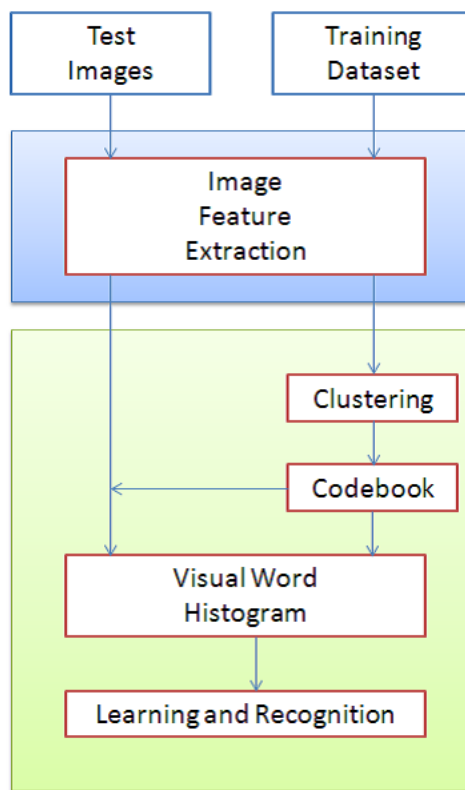


図 3-7: BoK 法の処理フロー

第4章

野菜識別手法

4.1 はじめに

本章では、提案手法である野菜識別手法について述べる。まず提案手法で用いる形状情報、色情報、機械学習について述べ、最後に提案手法の処理フローを述べる。

4.2 形状情報

平面上の曲線を周波数領域で記述する方法としてフーリエ記述子がある。曲線上の座標や偏角をある位置からの関数とみなし、その関数を離散フーリエ変換して得られるフーリエ係数の集合をフーリエ記述子と言う。フーリエ記述子にはZ型、G型をはじめ様々な種類があるが、本論文ではP型フーリエ記述子を採用した。P型フーリエ記述子では図4-1のように複素数平面上の曲線をN等分した折れ線で近似する。ここで、偏角 θ_i の指数関数と複素平面における座標 z_i との関係式は次式となる。

$$w_i = \exp(j\theta_i) = \cos \theta_i + j \sin \theta_i \quad (4.1)$$

$$= \frac{x_{i+1} - x_i}{\delta} + j \frac{y_{i+1} - y_i}{\delta} = \frac{z_{i+1} - z_i}{\delta} \quad (4.2)$$

δ は線分の長さであり、

$$\delta = |z_{i+1} - z_i| \quad (4.3)$$

である。

また w_i を周期関数として、離散フーリエ変換することでP型フーリエ記述子

$$c_K = \frac{1}{N} \sum_{i=0}^{N-1} w_i e^{-j \frac{2\pi K i}{N}} \quad (4.4)$$

を得る。

P型フーリエ記述子は拡大縮小、平行移動に対して不変であり、また絶対値は回転にも不変である。そのため、提案手法の形状情報は、P型フーリエ記述子の振幅スペクトルを採用した。振幅スペクトルは

$$p_K = \sqrt{a_K^2 + b_K^2} \quad (4.5)$$

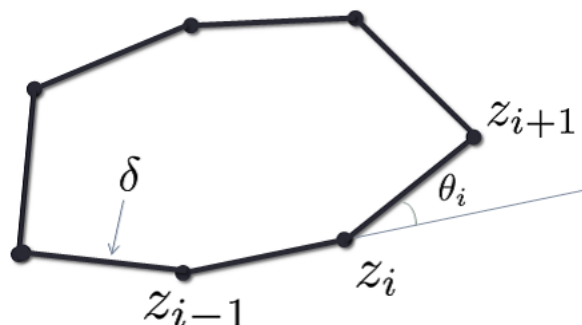


図 4-1: 曲線の折れ線近似

である。

ただし, a_k, b_k はそれぞれ c_k の実数項係数と虚数項係数である. 本研究では 8 近傍の輪郭追跡から野菜の輪郭線を抽出し, $N = 200$ で折れ線近似した. また P 型フーリエ記述子の項数 K は $0, 1, \dots, N$ である. P 型フーリエ記述子の再生曲線は,

$$\tilde{w}_i = c_0 + \sum_{K=1}^{K_c} \left\{ c_K e^{j \frac{2\pi K i}{N}} + c_{N-K} e^{-j \frac{2\pi K i}{N}} \right\} \quad (4.6)$$

$$\tilde{z}_i = \tilde{z}_{i-1} + \delta \tilde{w}_{i-1} \quad (4.7)$$

より得られる. 図 4-2 は P 型フーリエ記述子の項数 K を 1, 2, 5, 15 とした場合の再生曲線を表している. フーリエ記述子では低域成分が輪郭線の大まかな形状を, 高域成分が局所的な形状を表現する. 図 4-2 から, フーリエ記述子の項数 K の値が小さいときは大まかな形状を表し, K の値が大きくなるにつれ再生曲線が元の輪郭線に近づいていくことを確認できる.

4.3 色情報

本論文では, 野菜を白, 赤, 茶, 緑, 黒の 5 つの色クラスに分類するために, 色情報を用いる. P 型フーリエ記述子には似た形状の物体を識別できない欠点がある. そこで色情報による大まかな分類を行うことで, 似た形状同士の比較を防ぐ処理を行った. 色情報の記述には回転, 拡大縮小, 平行移動照明変化に対するロバスト性を考慮し RGB による平均値を用いた. 分類方法は, 入力画像の RGB と各色クラスの代表値とのユークリッド距離が最小値となる色クラスを野菜の属する色クラスとした. 各色クラスの代表値は, 予備実験にてあらかじめ最適な値を決定した. 表 4-1 は野菜をクラス分けした結果である.

4.4 機械学習

機械学習には SVM を用いた. 学習データは P 型フーリエ記述子から得られる野菜の形状情報である. 本研究では 17 品目の野菜の識別を行うため, one-vs.-one⁽¹⁵⁾ による多数決方

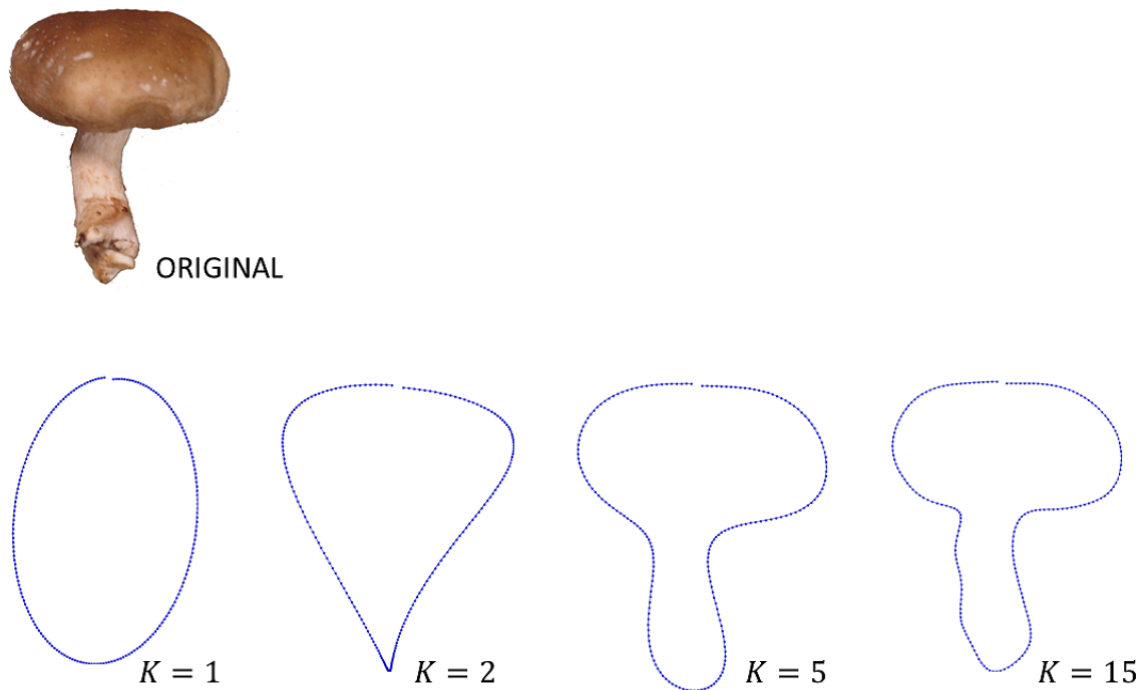


図 4-2: P 型フーリエ記述子から得られる再生曲線の例
(文献⁽³⁰⁾より引用)

式を用いることでマルチクラスに対応した。予備実験では，one-vs.-one と one-vs.-rest⁽¹⁵⁾ による判別器の生成と識別実験を行った。結果として，one-vs.-rest ではほとんどの判別器の出力が負の値となり正しく学習できなかった。one-vs.-rest による判別器生成の失敗の理由としては，正例と負例のサンプル数の偏りによるものが考えられる。この偏りの問題に対応するためには，正例以外の各野菜のサンプルから正例と同じ程度の負例を選択する必要がある。しかし本研究で集められた各野菜のサンプルが 10 枚と少なく，この方法では識別率がサンプルの選び方に大きく依存してしまう可能性がある。そのため，本研究ではマルチクラスの分類手法に one-vs.-one を用いた。

表 4-1: 色クラスと対応する野菜

色クラス	野菜
白	キャベツ, 大根
緑	ゴーヤ, ピーマン, 青梗菜, ブロッコリー, キャベツ, キュウリ, レタス, ほうれん草
赤	人参, 椎茸, さつまいも, トマト
茶	玉葱, ポテト, 椎茸
黒	アボガド, ナス

4.5 提案手法の処理フロー

図4-3は、提案手法の処理フローの概要を表している。SVM(V_i, V_j)は正例に V_i 、負例に V_j を学習させた判別器である。ただし $i < j$ とする。 V_i は野菜のクラスである。 X_i は4.1で述べた野菜を大まかに分ける色クラスである。 M は野菜の品種数で、本稿では $M = 17$ である。提案手法は、以下に示す処理フローに基づいて野菜の識別を行う。

STEP1. 情報取得

入力画像から形状と色情報を得る。入力画像内の前景領域は、輪郭線追跡を行い取得する。

STEP2. 色クラスの決定

色情報から入力画像の属するクラス X_i を決定する。

STEP3. マルチクラス SVM の選択

あらかじめ形状情報を学習させた判別器の中から X_i に関する判別器を全て選択する。

STEP4. 識別処理

STEP3で選択した判別器に入力画像の形状情報を判別させる。

STEP5. 投票

STEP4で最多識別回数となった野菜クラスを識別結果とする。また最多識別回数が重複するケースは、判別器の出力値の総和が最大となる野菜クラスを識別結果とする。

4.6 おわりに

本章では、提案する野菜識別手法について述べた。はじめに、P型フーリエ記述子を用いた形状情報の表現方法、RGBの平均値による野菜の分類方法、SVMを用いたマルチクラスの分類方法について述べ、最後に野菜識別手法の全体の処理フローについて述べた。上記の手法を用いることで、画像を用いた野菜の識別を実現する。

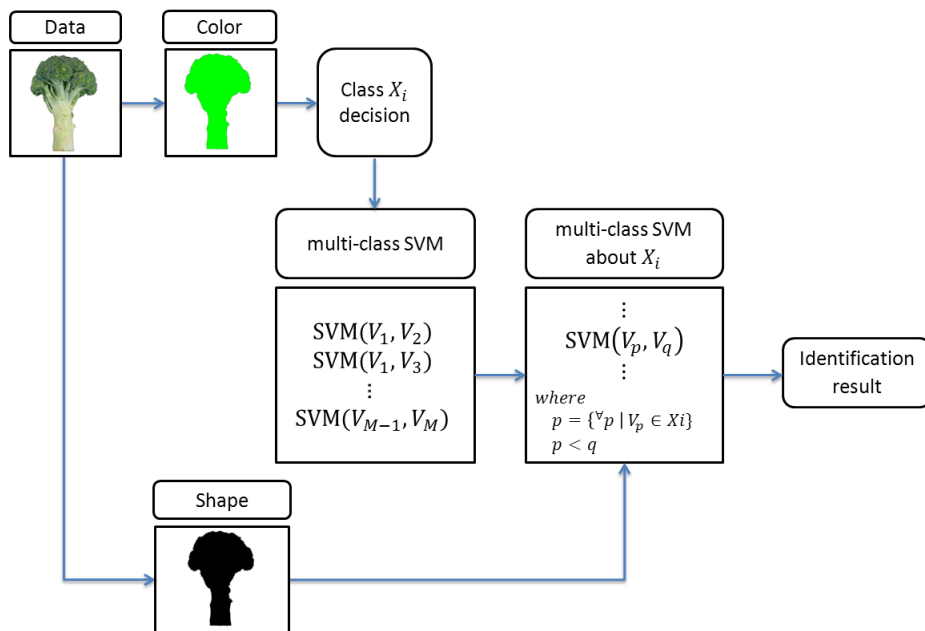


図 4-3: 提案手法の処理フロー
(文献⁽³⁰⁾より引用)

第5章

提案手法の評価実験

5.1 はじめに

本章では、前章で提案した野菜識別手法の評価実験の結果を示し、考察を述べる。

5.2 評価方法

本実験で用意できた各野菜のサンプル画像は10枚と少なく、識別結果はテストデータと学習データの分類の仕方に大きく依存する可能性がある。そこで本実験ではLOOCV(leave-one-out cross-validation)法⁽¹³⁾による平均識別率で性能を評価した。LOOCV法は、全サンプルの中から1サンプルだけをテストデータとして取り出し、残り全てのサンプルを学習データとする。そしてこれを全サンプルが一回ずつテストデータとなるよう検定を繰り返す手法である。各野菜10サンプルであるため、ある野菜のテストデータと学習データは図5-1のように切り替わる。

5.3 提案手法の評価実験

はじめに、P型フーリエ記述子の野菜形状の識別に対する有効性と有効な項数Kについて結果と考察を述べる。最後に、P型フーリエ記述子と色情報を用いた場合の識別率について結果と考察を述べる。

5.3.1 P型フーリエ記述子のみを用いた野菜識別実験

まず、P型フーリエ記述子のみを用いて野菜の識別を行った。図5-2は、実験で用いた野菜画像のサンプルである。野菜画像は、スーパーマーケットで購入した野菜を撮影した画像とネット上の著作権フリーの画像から17品目各10サンプル集めた。図5-3は、P型フーリエ記述子の項数Kを0から200まで変化させた際の平均識別率の結果である。

結果としてK=200の時に、最高平均識別率77%を得た。この結果からP型フーリエ記述子が野菜形状の識別に対して有効であると考えられる。次にK=200での各野菜の識別率と平均識別率を図5-4で示す。図5-4から、全体的に高い識別率を得られていることが分か

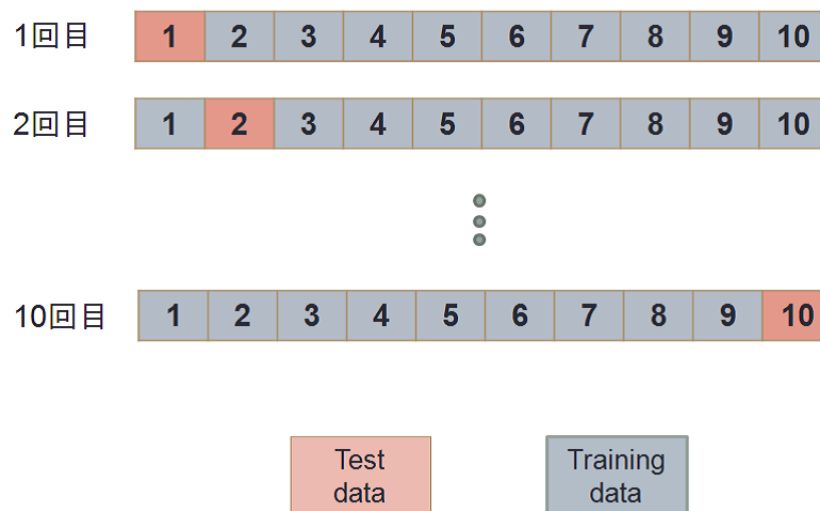


図 5-1: LOOCV 法

る。しかしその一方で青梗菜の識別率は極端に低く、ほとんど識別できていないという結果となった。原因としては、青梗菜の葉の柔らかい部分が様々な形状を取ってしまうためと考えられる。図 5-5 は、実際に実験で使用した青梗菜のサンプルである。

5.3.2 P 型フーリエ記述子と色情報を用いた野菜識別実験

P 型フーリエ記述子と色情報を用いた野菜識別手法の性能評価を行った。図 5-6 は、P 型フーリエ記述子の項数 $K=200$ での各野菜の識別率と平均識別率を表している。図 5-7 は、P 型フーリエ記述子のみの場合と P 型フーリエ記述子と色情報を用いた場合の識別率を比較したグラフである。結果として、平均識別率 89.4 % を得た。この結果から提案手法が野菜の識別に対して有効であると考えられる。また P 型フーリエ記述子のみによる識別よりも、各野菜の識別率とそれらの平均識別率が向上しており、色情報によるクラス分けの有効性も確認できる。

一方で、色情報を考慮することで識別率が低下した野菜もある。ニンジンでは、色情報を考慮することで、考慮しない場合と比べると 20 % 識別率が低下した。原因としては、色クラス内に似た形状の野菜があったためと考えられる。ニンジンは色情報を考慮すると赤クラスに属する。図 5-8 は、同じ赤クラスのサツマイモである。P 型フーリエ記述子では、図 5-8 のニンジンとサツマイモは、似たような形状となってしまう。そのため、色情報によるクラス分けによって似た形状の比較が行われてしまい、識別率が低下したと考えられる。

5.4 おわりに

本章では、提案手法である P 型フーリエ記述子と色情報によるクラス分けを用いた野菜識別手法の評価実験を行った。提案手法の有効性については、LOOCV 法による平均識別率



図 5-2: 野菜のリスト

で評価した。

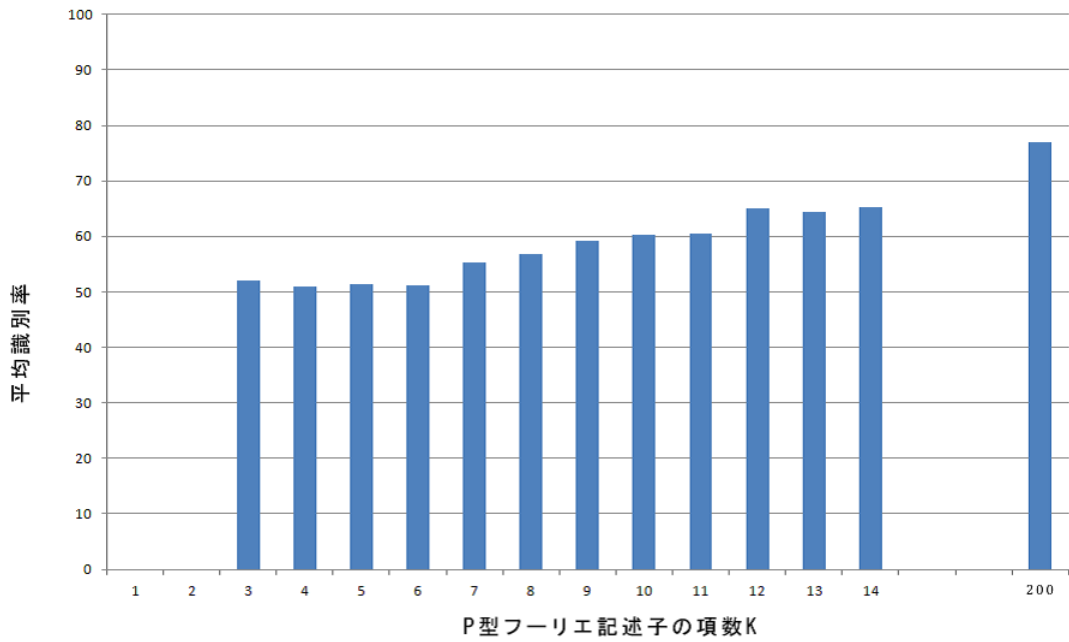


図 5-3: P 型フーリエ記述子の項数 K を変化させた際の平均識別率

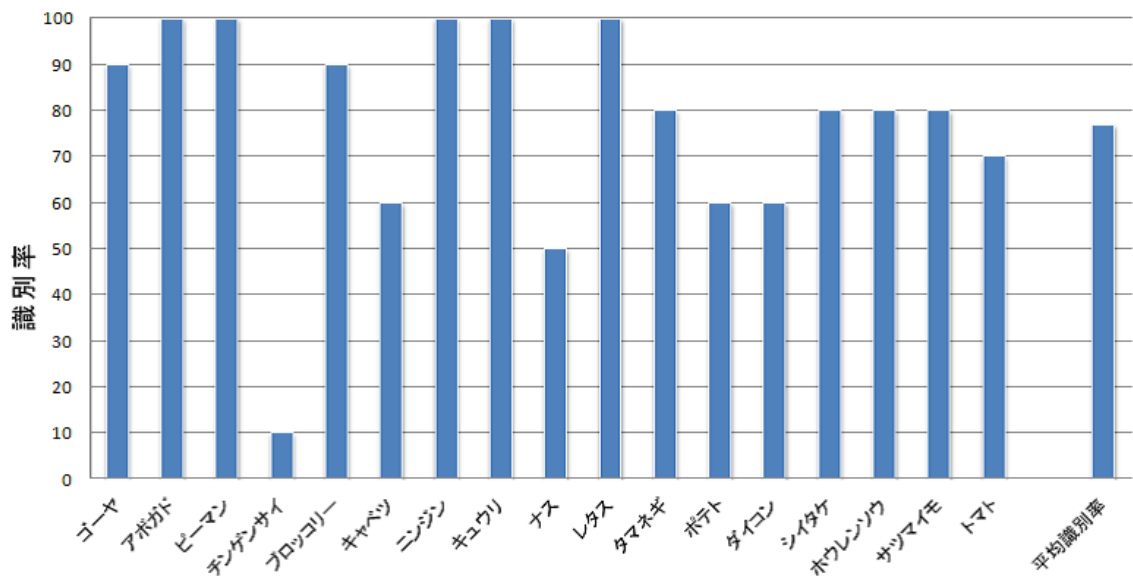


図 5-4: 形状情報のみの場合の識別率



図 5-5: 青梗菜のサンプル

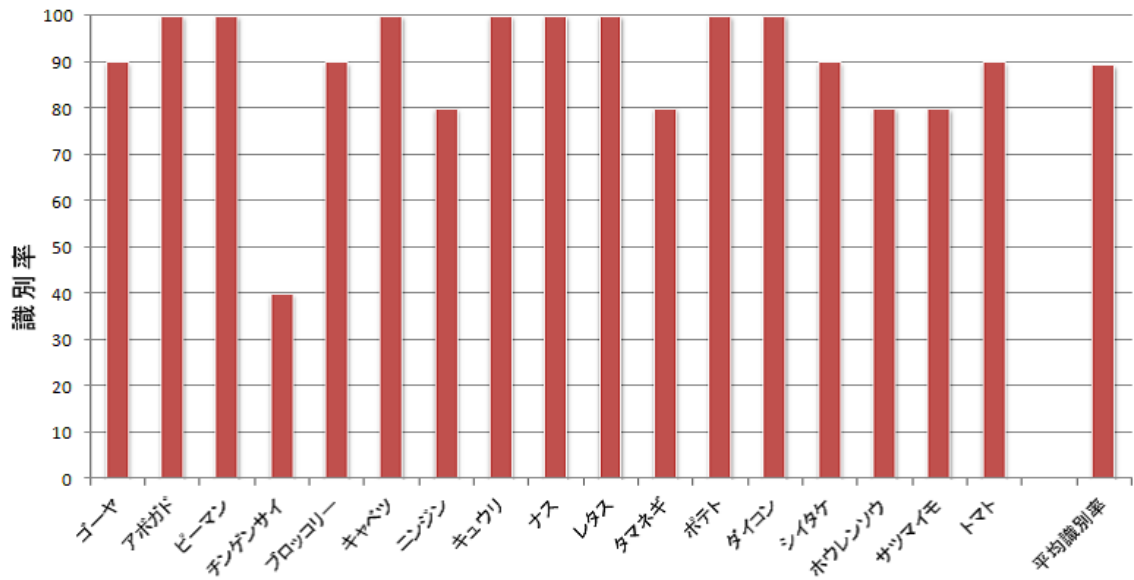


図 5-6: 形状情報と色情報を用いた場合の識別率

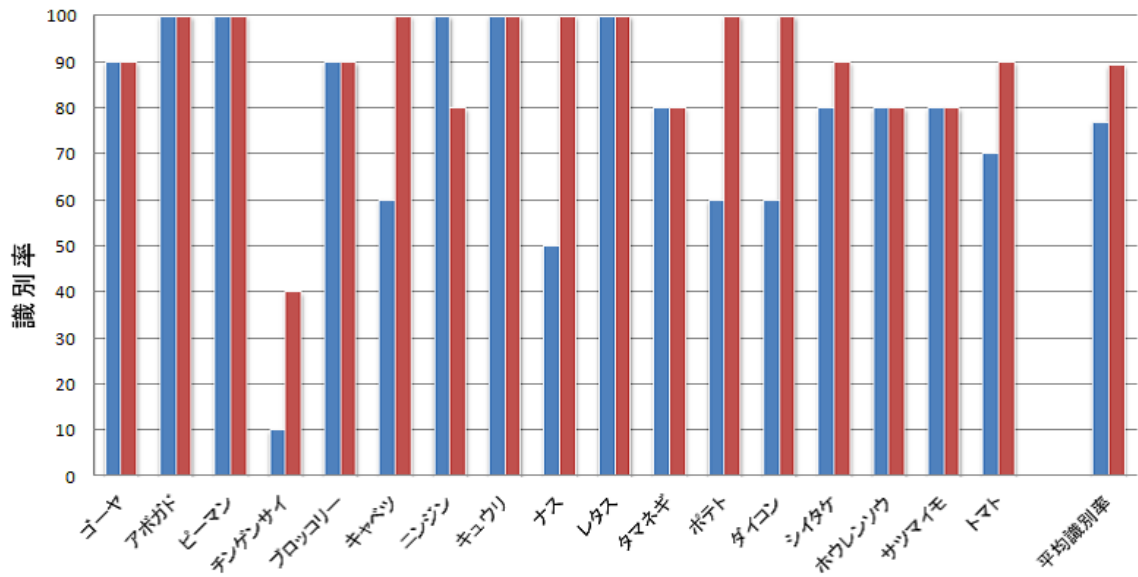


図 5-7: 形状情報のみの場合と形状情報と色情報を用いた場合の識別率の比較

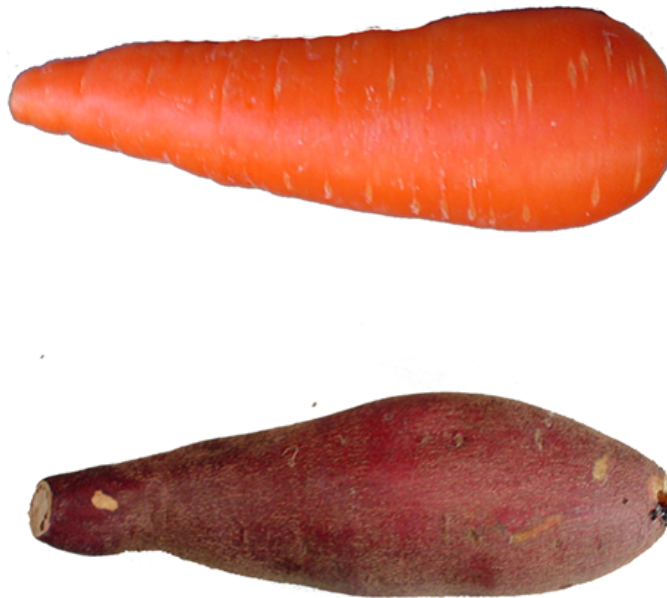


図 5-8: 似た形状となってしまうニンジンとサツマイモ

第6章

結論

本章では本論文を総括し、今後の検討課題を述べる。

6.1 総括

第1章では本研究の背景と目的、および本論文の構成について述べた。

近年、物品管理システムへの関心が高まっている。Wal-Mart Stores, Inc. は、従来のバーコードを用いた物品管理に RFID タグを導入し、物品管理システムの最適化を図っている。このような物品管理システムへの関心は物流だけではなく、家庭の冷蔵庫に対しても向けられており、物品および食品管理システムとして研究が進められている。既存の冷蔵庫では、保管している食品を外部から確認できないため、無駄に食品を購入してしまう問題がある。RFID 技術を導入するためには、RFID タグや RFID リーダー等の導入コストの問題がある。そのため、身近にあるモバイルデバイスを用いた物品管理システムが有効であると考えられる。

第2章では現状の食品管理システムについて概要と問題点を述べた。特殊なデバイスを用いたシステムと画像による物体認識・識別を用いた手法について述べ、各システムおよび手法の問題点を述べた。

第3章では、既存の画像を用いた物体認識・識別手法として局所特徴量を用いた手法の要素技術について述べた。局所特徴量として SIFT について、キーポイントとスケールの検出、キーポイントのローカライズ、オリエンテーションの算出、特徴量の記述に分けて述べた。機械学習手法として SVM について述べた。学習データが線形識別可能な場合と不可能な場合について述べた。また、非線形分離を行うカーネルトリックについて述べた。最後に BoK 法について述べた。

第4章では野菜の識別を実現するために、形状情報と色情報を用いた野菜識別手法を提案した。提案手法では形状情報に P 型フーリエ記述子を用いた。そして、似た形状比較を避けるために色によるクラス分けを行った。また、未知データに対する識別能力を考慮し、形状情報の学習および識別には SVM を採用した。

第5章では第4章で提案した野菜識別手法を、LOOCV 法による平均識別率を用いて評価実験を行った。まず評価方法である LOOCV 法について述べた。次に P 型フーリエ記述子

の有効な項数 K を算出した。結果として $K=200$ の時、最大平均識別率 77 % を得た。この結果から、P 型フーリエ記述子の野菜形状に対する有効性を確認した。最後に形状情報のみと形状情報と色情報を用いた場合について各野菜の識別率とそれらの平均識別率を求めた。結果として形状情報と色情報を用いた場合は平均識別率 89.4 % を得た。この結果から、提案手法の有効性を確認した。また、形状のみの場合と比較して、平均識別率が向上していることから、色情報によるクラス分けの有効性を確認した。

第 6 章は結論であり、本論文の総括および今後の課題について述べている。

6.2 今後の課題

今後検討すべき課題を以下に示す。

今後の課題としては、識別率が悪かった青梗菜への対応が挙げられる。青梗菜は、葉の部分が柔らかく様々な形状を取るため、P 型フーリエ記述子による形状情報では識別が困難であると考えられる。この問題を解決するためには、テクスチャを考慮するなど、形状以外の特徴量を野菜から抽出し、識別を行う必要がある。また、本研究では予め手動で前景と背景を分離した画像を用いたが、実際に食品管理システムとして提案手法を適用するためには、自動的に前景背景の分離を行う必要がある。その影響を考慮する必要がある。前景背景の分離手法には、背景差分法やグラフカットなどが挙げられる。しかしこれらの手法は、陰の影響や画像のブレの影響を受けるため、手動で前景背景を分離した場合と比べると、野菜の輪郭をきれいに抽出できない問題がある。以上の問題に対応し実際に食品管理システムを実現することが今後の検討課題である。

謝辞

本研究の機会及び素晴らしい実験環境を与えて下さり，貴重な時間を割いて研究の方向性を御指導頂きました渡辺 裕教授に心から感謝致します。

学部時代にお世話になり，本研究科に入るきっかけを与えて下さいました，東京理科大学の柳田 昌宏准教授に心から感謝いたします。

貴重な御意見，様々な御提案を頂いた AVS ゼミの皆様に御礼申し上げます。特に，数々の有意義な御意見，御助言を下された石井 大祐氏に心から感謝致します。

最後に，私をここまで育てて下さった家族に深く感謝します。

平成 25 年 3 月 15 日

参考文献

- (1) “iPhone,” <http://www.apple.com/jp/iphone/>
- (2) “Android,” <http://www.android.com/>
- (3) “OpenCV,” <http://opencv.jp/>
- (4) “はてなモノリス,” <http://mono.hatena.ne.jp/>
- (5) “Google Goggles,” <http://www.google.com/mobile/goggles/>
- (6) 藤吉弘亘, 山下隆義, 岡田和典, 前田英作, ノジク・ヴァンソン, 石川 尋代, ドゥソルビエ・フランソワ, コンピュータビジョン 最先端ガイド 2, アドコム・メディア株式会社, 2011.
- (7) Y.Boykov, V.Kolmogorov, “An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision,” *IEEE Transactions Pattern Analysis and Machine Intelligence*, vol.26, pp.1124-1137, 2004.
- (8) D.G.Lowe, “Distinctive Image Features from Scale-Invariant Keypoints,” *International Journal of Computer Vision*, Vol.60, No.2, pp.91-110, Jan.2004.
- (9) H.Bay, T.Tuytelaars, L.V.Gool, “SURF: Speeded Up Robust. Features,” *Proc. of European Conference on Computer Vision*, pp.404-415, 2006.
- (10) Y.Ke, R.Sukthankar, “PCA-SIFT: A more distinctive representation for local image descriptors,” *CVPR2004*, Vol.2, pp.506-503, 2004.
- (11) C.Cortes, V.Vapnik, “Support-Vector Networks,” *Machine Learning*, Vol.20, No.3, pp.273-297, 1995.
- (12) 津田宏治, “サポートベクターマシンとは何か,” *電子情報通信学会誌*, Vol.83, No.6, pp.460-466, 2000.
- (13) C.M. Bishop, “パターン認識と機械学習 上 - ベイズ理論による統計的予測,” Springer, 2007.
- (14) C.M. Bishop, “パターン認識と機械学習 下 - ベイズ理論による統計的予測,” Springer, 2008.
- (15) 栗田哲平, 近山隆, “多クラス Support Vector Machine を用いた一般物体認識での複数候補提示下における分類性能の傾向情報処理学会研究報告,” Vol.2008, No.115, pp.251-258, 2008.
- (16) 矢部博, “工学基礎 最適化とその応用,” 数理工学社, 2006.

- (17) Nello Cristianini, John Shawe-Taylor, “サポートベクターマシン入門,” 共立出版, 2008.
- (18) Y.Freund, R.E.Schapire, “A decisiontheoretic generalization of on-line learning and an application to boosting,” *Journal of Computer and System Sciences*, Vol.55, No.1, pp.119-139, 1997.
- (19) Thorsten Joachims, “Text categorization with support vector machines: Learning with many relevant features,” In Claire Nedellec and Celine Rouveirol editors *Machine Learning*, Vol.1398, pp.137-142, 1998.
- (20) E.Osuna, R.Freund, F.Girosit, “Training support vector machines: an application to face detection,” In *Computer Vision and Pattern Recognition*, 1997.
- (21) G.Csurka, C.Bray, C.Dance, L.Fan, “Visual categorization with bags of keypoints,” *Proc. of ECCV Workshop on Statistical Learning in Computer Vision*, pp.59-74, 2004.
- (22) C.D.Manning, and H.Schfutzte, “Foundation of statistical Natural Language Processing,” The MIT Press, 1999.
- (23) S.S.Keerthi, S.K.Shevade, C.Bhattacharyya, K.R.K.Murthy, “Improvements to Platt’s SMO Algorithm for SVM Classifier Design,” *Neural Computation*, Vol.13, No.13, pp.637-649, 2001.
- (24) 野田一雄, 宮岡悦良, “入門・演習 数理統計,” 共立出版株式会社, 2004.
- (25) 上坂吉則, “開曲線にも適用できる新しいフーリエ記述子,” *電子情報通信学会論文誌 A*, Vol.J67-A, No.3, pp.166-173, 1984.
- (26) 望月宏史, 柳井啓司, “物体認識技術を用いたモバイル物品管理システム,” *データ工学と情報マネジメントに関するフォーラム*, 2012.
- (27) 森元亜美, 向井信彦, 小杉信, “葉形状特徴を用いた植物の識別,” *映像情報メディア学会技術報告*, Vol.34, No.15, pp.69-72, 2010.
- (28) 原田幹子, 川村裕介, 島村和典, “RFID タグによる冷蔵庫内食品把握に関する一提案,” *電気関係学会四国支部連合大会講演論文集*, Vol.2009, No.ROMBUNNO.12-16, pp.166-173, 2009.
- (29) 原田幹子, “RFID タグによる冷蔵庫管理制御に関する研究,” *高知工科大学情報システム工学科 学士学位論文*, 2010.
- (30) 山崎太一, 石井大祐, 渡辺裕, “フーリエ記述子と色情報を用いた野菜識別手法,” *情報処理学会 AVM 研究会研究報告*, Vol.112, No.335, pp.145-148, 2012.

図一覧

2-1 冷蔵庫管理システム全体図 (文献 ⁽²⁹⁾ より引用)	4
2-2 望月らの提案システムの動作イメージ (文献 ⁽²⁶⁾ より引用)	5
3-1 画像から求められる SIFT 特徴量の例	7
3-2 LoG オペレータ例 (文献 ⁽⁶⁾ より引用)	8
3-3 DoG 画像を用いた 26 近傍による極値の検出	9
3-4 キーポイント周辺の各ブロックの勾配方向 (文献 ⁽⁶⁾ より引用)	12
3-5 学習データが線形分離可能な場合	12
3-6 学習データが線形分離不可能な場合	15
3-7 BoK 法の処理フロー	18
4-1 曲線の折れ線近似	20
4-2 P 型フーリエ記述子から得られる再生曲線の例 (文献 ⁽³⁰⁾ より引用)	21
4-3 提案手法の処理フロー (文献 ⁽³⁰⁾ より引用)	23
5-1 LOOCV 法	25
5-2 野菜のリスト	26
5-3 P 型フーリエ記述子の項数 K を変化させた際の平均識別率	27

5-4 形状情報のみの場合の識別率	27
5-5 青梗菜のサンプル	28
5-6 形状情報と色情報を用いた場合の識別率	28
5-7 形状情報のみの場合と形状情報と色情報を用いた場合の識別率の比較	29
5-8 似た形状となってしまうニンジンとサツマイモ	29

表一覧

4-1 色クラスと対応する野菜	21
---------------------------	----

研究業績

	題 名	発表年月	発表掲載誌	連名者
(1)	Multi Size Eye Detection on Digitized Comic Image	2012 年 11 月	IIEEJ 3rd Image Electronics and Visual Computing Workshop (IEVC) 2012	D.Ishii T.Yamazaki H.Watanabe
(2)	フーリエ記述子と色情報を用いた野菜識別手法	2012 年 12 月	情報処理学会 AVM 研究会	山崎 太一 石井 大祐 渡辺 裕
(3)	マンガ固有の特徴を利用したマンガ登場人物識別に関する一検討	2013 年 2 月	情報処理学会 AVM 研究会	石井 大祐 山崎 太一 渡辺 裕