

外観画像分析によるクリ果実の品種判定

佐野健人¹⁾

(茨城県農業総合センター園芸研究所)

Determination of Chestnut Cultivars by Appearance Image Analysis

Taketo SANO¹

要約

クリ 8 品種を対象に側果を外側面側から撮影した外観画像分析による品種判定モデルの構築を試みた。外観として果皮色と輪郭を分析した結果、果皮色は貯蔵により変化したが、輪郭は貯蔵による変化が認められなかったことから、輪郭が品種の判定項目として適当と考えられた。輪郭をフーリエ記述子に変換して主成分分析を行い、マハラノビス距離を算出し、最小のマハラノビス距離を与える群を該当品種と判定する品種判定モデルを構築できた。1190 果を判定モデルで検証した結果、個体判定としての正答率は 54% であった。一方、生産現場では樹（品種）ごとにコンテナ収穫することが多いため、コンテナ単位など集団の品種判定モデルとしてボルダ方式による集計を採用すると、8 品種とも正しく判定できた。

キーワード：クリ，画像分析，品種判定，輪郭，果皮色

1 はじめに

本県は生産量・出荷量ともにクリの全国一の産地である。クリ果実は、‘利平’等の品質が良く高単価が期待される品種や‘ぼろたん’のように性質（剥皮性）の異なる品種以外では、品種を区別して取り扱われることは一般的ではないが、本県の一部の産地・団体・業者ではクリ果実を品種別に集出荷している。出荷までの貯蔵中に品種を混同、判定困難になる懸念から、著者は既報（佐野ら，2012）で、クリ‘ぼろたん’と‘国見’の果実の近赤外分光法による非破壊品種判別に取り組み、近赤外線 2 次微分スペクトルを PLS 回帰法による検量モデルに適合することで 9 割程度の正答率で判別可能と報告した。しかし、産地で取り扱われている品種はより多岐にわたる。また、果皮が厚いクリ果実への近赤外分光法の測定は、測定時間が一果実あたり数十秒～数分かかることから、より簡便な手法が期待されていた。品種判定の手法としては DNA 解析によるものも試みられているが（矢野，2004 門田，2017）、簡便な手法とは言い難い。そのような中、近年では画像処理技術が発達し、外観を撮影した写真をもとに、形状や色を定量的に取り扱うことができるようになった。そこで、本県で栽培されるクリの 8 品種について、果実の形状および果皮色を定量化し、品種の分類・判定を試みたので、その結果について報告する。なお、一部の産地・団体・業者では、低温貯蔵後に出荷をしているため、貯蔵前と貯蔵後の外観変化についても検討した。

2 材料および方法

2. 1 外観定量化のための画像データの収集

クリ果実は目視で選別した側果（外グリ）を用い、外側面側から外観を撮影した。撮影には、zoom を 3、focus を 40 に設定したビジュアルアナライザー VA300（A 株式会社）を使用した。

1) 現 茨城県霞ヶ浦環境科学センター

1 Address : Ibaraki Agricultural Center Horticultural Research Institute, 3165-1 Ago, Kasama, Ibaraki 319-0292, Japan

外観に、貯蔵による変化が見られるか検証するため、貯蔵前と貯蔵後に撮影を行い、それぞれの果皮色および輪郭を評価した。2013年は2時期（貯蔵前・9～10月と貯蔵後・11月）に、2014年は3時期（貯蔵前・9～10月と貯蔵後1・10月、貯蔵後2・12月）に撮影を行った。

2013年の貯蔵前の写真撮影は‘丹沢’が9月11日、‘ぼろたん’が9月18日、‘筑波’が10月2日、‘石鎚’が10月11日に、収穫当日の果実を用いて行った（表1）。なお、いずれも中庸な大きさの果実を目視で選び撮影に用いた。

2013年の貯蔵後の写真撮影は、収穫後0℃で貯蔵していた果実を用い、11月11日に行った。果実は、品種ごとに複数の収穫日の果実を混ぜて貯蔵したものから無作為に選別した8品種（‘丹沢’、‘ぼろたん’、‘国見’、‘利平’、‘銀寄’、‘筑波’、‘石鎚’、‘岸根’）の果実を用い、‘ぼろたん’は3L規格果、その他の品種は目視で中庸な大きさの果実を選び撮影対象とした。なお、2013年の貯蔵前と貯蔵後の写真撮影は別の果実を用いた。

表1 クリの収穫及び撮影日等

年	撮影時期等	撮影の有無	品種							
			丹沢	ぼろたん	国見	利平	銀寄	筑波	石鎚	岸根
2013	収穫日（貯蔵前）	有	9/11	9/18	—	—	—	10/2	10/11	—
	貯蔵後	有	11/11							
2014	収穫日	無	9/8	9/17	9/17	9/26	9/26	9/26	10/7	10/7
	収穫～撮影までの一時保管条件	無	9/17～9/18, 18℃			9/26～9/29, 18℃				
			2℃	9/18～9/19, 0℃		9/29～9/30, 2℃			2℃	
	貯蔵前	有	9/11	9/19	9/19	9/30	9/30	9/30	10/9	10/9
	貯蔵後1	有	10/17							
貯蔵後2	有	12/24								

※ “—”：実施せず。

2014年は1品種あたり200果程度を供試し、貯蔵前と貯蔵後に撮影したが、個体の区別はつけず、また、完全に一致していない。つまり、‘丹沢’では9月8日収穫の2L果258果のうち100果を9月11日に貯蔵前の撮影を行い、258果を貯蔵して10月17日と12月24日に個体を区別せずに100果をそれぞれ選び撮影した。‘ぼろたん’は9月17日収穫の2L果を9月19日に貯蔵前の撮影を行った。‘国見’は9月17日収穫の果実のうち側果のみを選び、9月19日に貯蔵前の撮影を行った。‘利平’、‘銀寄’、‘筑波’は9月26日に収穫して中庸な大きさの側果のみを選び出し、9月30日に貯蔵前の撮影を行った。‘石鎚’は2L果、‘岸根’は3L果を10月7日に収穫、10月9日に貯蔵前の撮影を行った。貯蔵後の撮影は各品種とも‘丹沢’と同日の2回行った。

2013、2014年に撮影した果数を表2に示した。

なお、果実は水洗し、表面についた水をペーパータオルでふき取ってから撮影した。撮影後に果実を低密度ポリエチレン袋に入れ、袋の口を軽く折った状態で、次の撮影まで0℃で貯蔵した。

表2 撮影したクリの果数

年	撮影時期	品種								計
		丹沢	ぼろたん	国見	利平	銀寄	筑波	石鎚	岸根	
2013	貯蔵前	30	30	—	—	—	30	30	—	120
	貯蔵後	52	43	11	49	64	106	67	92	484
<hr/>										
	貯蔵前	100	100	53	100	68	100	100	85	706
2014	貯蔵後 1	100	100	52	100	64	100	100	84	700
	貯蔵後 2	100	100	45	100	63	100	100	81	689
品種累計		382	373	161	349	259	436	397	342	2,699

※ “—”：データなし。

2. 2 果皮色による品種判定

得られた画像は、ビジュアルアナライザーVA300 (A 株式会社) 付属の「アルファソフト ver12.42」により表3の設定で各画像を処理し、それぞれの果実の画像を構成する画素を色別に集計した。

表3 アルファソフト ver12.42 による果皮色の解析時の設定

項目	設定値
閾値	Lab
データタイプ	L : in 0-200, a : in 0-255, b : in 0-255
プリミティブパラメータ	オブジェクト, スペクトラム, プリミティブ
マルチオブジェクト	カラー距離 : 50, 最小サイズ : 10
スペクトラム	最小サイズ : 50
	値 最小 : 1, 平均値 最小 : なし

※ 各項目及び設定値の表記は、アルファソフト ver12.42 上の表記による。

なお、「閾値」とは、撮影画像から目的領域を抽出する際の基準値であり、ここでは表色形式として Lab を選択し、L 値 0~200 かつ a 値 0~255 かつ b 値 0~255 の色の領域のみを選択するよう設定している (L 値 201~255 の明るすぎる色領域は、背景として扱い、目的領域としないように設定した)。

2. 3 輪郭による品種判定

2. 1 で得た累計 2699 果分の各画像を Iwata ら (2002) によるソフトウェアパッケージ「SHAPE Ver.1.3」により、各果実の輪郭を標準化楕円フーリエ記述子 (調和数 20) に変換した (表4 参照)。この処理により、標準化楕円フーリエ記述子のフーリエ係数は、x 座標分 2 ケ、y 座標分 2 ケ、調和数 (表4 の Max Harmonic

表4 SHAPE Ver.1.3 による果形特徴の解析時の設定

項目	内容
【プログラム】 ChainCoder	
Gray Scale	G (Green)
Ero Dil Filter	1
Dil Ero Filter	0
【プログラム】 Chc2Nef	
Max Harmonic No.	20
Normalization Method	based on the Longest radius

No.での設定項目) の 20 ケ分、合わせて $2 \times 2 \times 20 = 80$ ケが各果実当たり得られる。なお、標準化楕円フーリエ記述子へ変換時の角度調整は目視で行い、果頂部を最上部に配置した。

得られた標準化楕円フーリエ記述子の係数を、年、貯蔵前後および品種ごとに平均し、平均したフーリエ記述子の係数から年、貯蔵前後および品種ごとに平均果形（輪郭）をフリーソフトウェア「R」により作図した。

次に、2013年は貯蔵後の8品種484果、2014年の貯蔵前706果、合わせて1190果を対象に、Iwataら(2002)の方法を参考に、楕円フーリエ記述子のフーリエ係数を説明変数とした主成分分析を行い、主成分得点をもとに品種の判定モデルを作成した。

判定モデル作成の手順については以下の通りである（図1参照）。

まず、品種（目的変数）とフーリエ係数（説明変数）のデータを用意する。試料件数を n 件、説明変数の数を m ケとすると、 n 件の試料それぞれに、1ケの品種情報と m ケの説明変数データがあるため、 n 件 \times $(1+m)$ ケのデータセット（ n 行 \times $(1+m)$ 列）となる。

手順（1）～（12）により、判定モデルを作成した。

- （1） n 件 \times m 点の説明変数に対して主成分分析を行い、 n 件 \times m 点（ n 行 \times m 列）の主成分得点を得る。
- （2）各列の主成分得点に対し、統計的に有意な品種差があるか分散分析を行う。
- （3）有意差のない k 列を除く。 n 行 \times $(m-k)$ 列の主成分得点データセットが得られる。

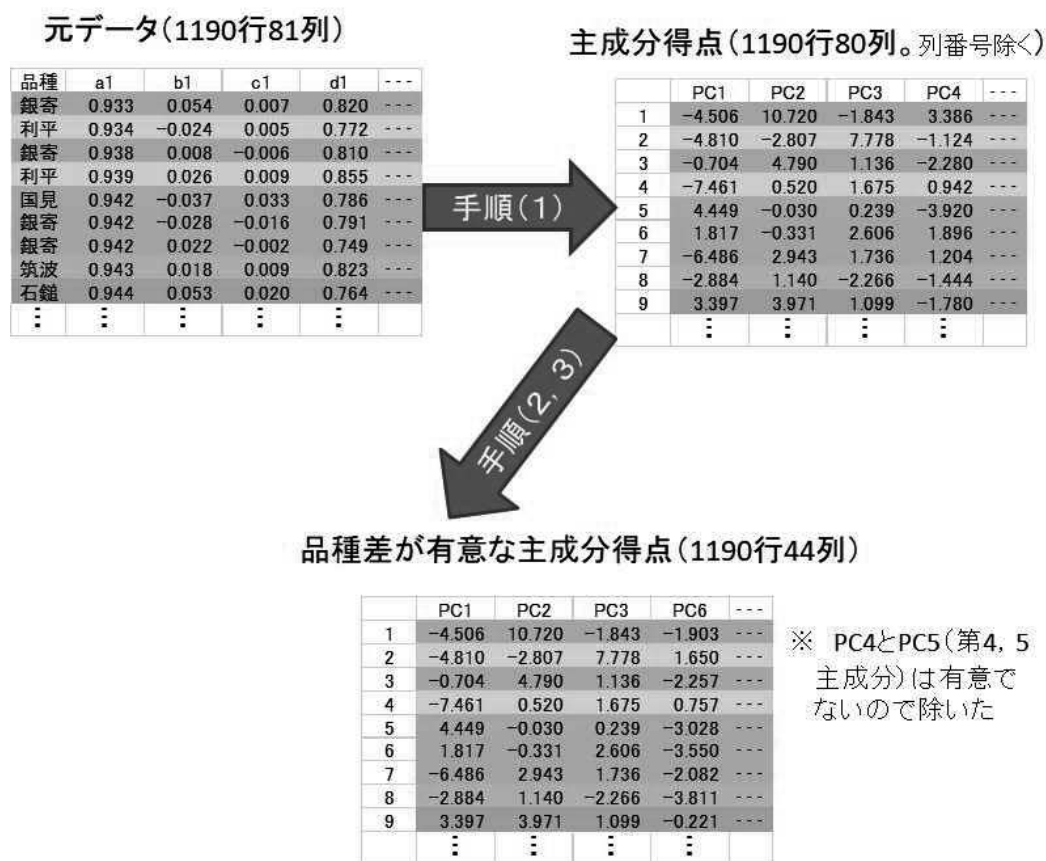


図1 (1/5) クリの輪郭をもとにした品種判定モデルの作成手順（本文（1）～（3））

- （4）1列目のみを抜き出して、 n 行 \times 1 列の主成分得点データセットを得る。
- （5）任意の1件の試料データ x_1 を除く（ x_1 は1行1列）。残りの $(n-1)$ 行 \times 1 列のデータセットを A とすると、 A を品種別に分ける（品種別に $A_1 \sim A_8$ とする）。
- （6） x_1 , $A_1 \sim A_8$ の1列目の主成分得点を用い、 x_1 と A_i ($i=1 \sim 8$) それぞれのマハラノビス距離を求める。
- （7） x_1 は、最小のマハラノビス距離を与える A_i 群の品種だと判定する。
- （8）（7）の「 x_1 が品種 i である」という判定が、正か誤か評価する。

- (9) (5) ~ (8) の手順を別の試料データ x_h ($h=1\sim n$) についても繰り返す。
 (10) x_1 から x_n までの品種判定の正・誤件数をまとめ、正答率を求める。

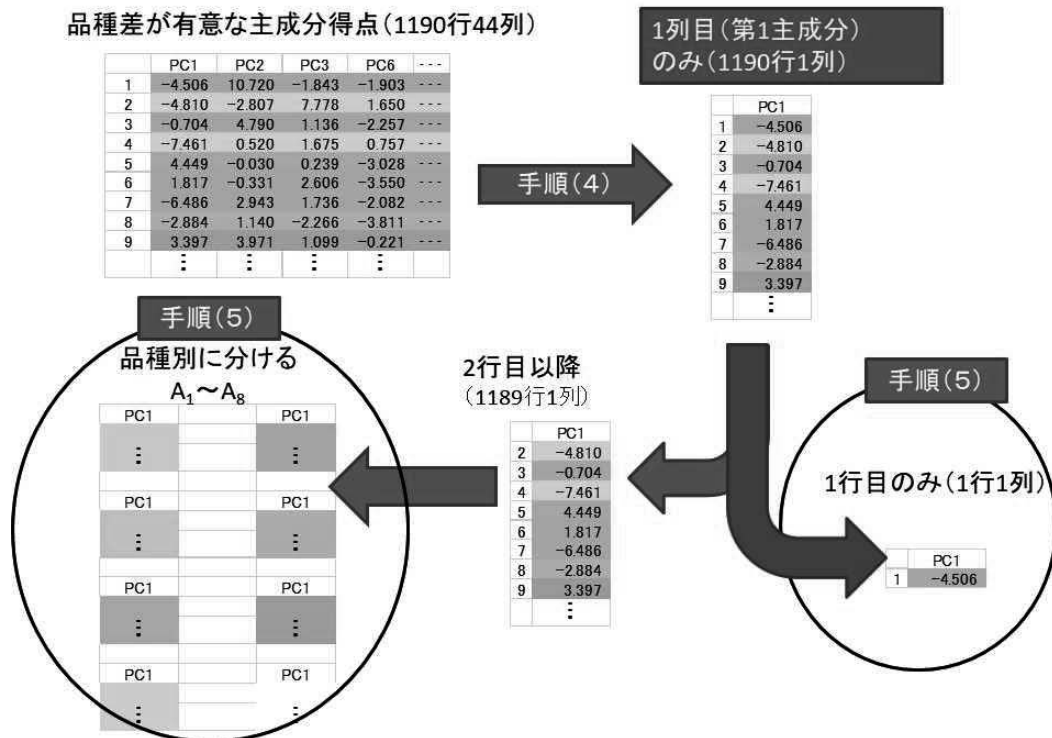


図1 (2/5) クリの輪郭をもとにした品種判定モデルの作成手順 (本文 (4) ~ (5))

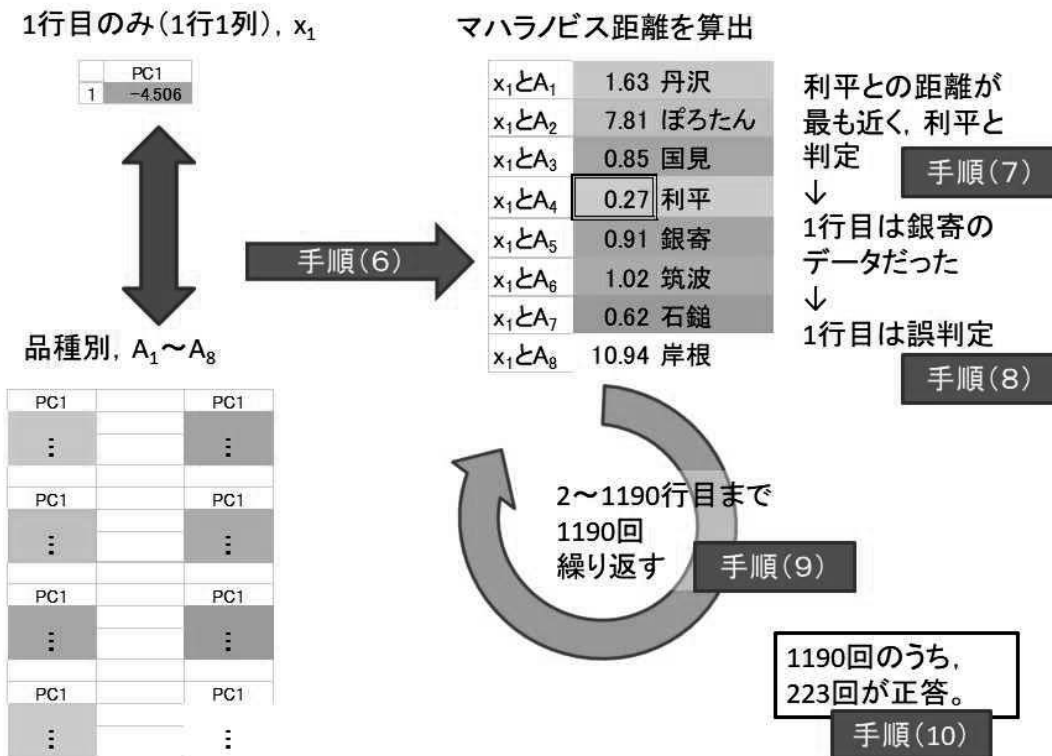


図1 (3/5) クリの輪郭をもとにした品種判定モデルの作成手順 (本文 (6) ~ (10))

(11) (3) で得た n 行 \times $(m-k)$ 列のデータセットから、 $1 \sim g$ 列目 ($g \leq m-k$) を抜き出した n 行 g 列のデータセットに対し、(5) ~ (10) を行う。

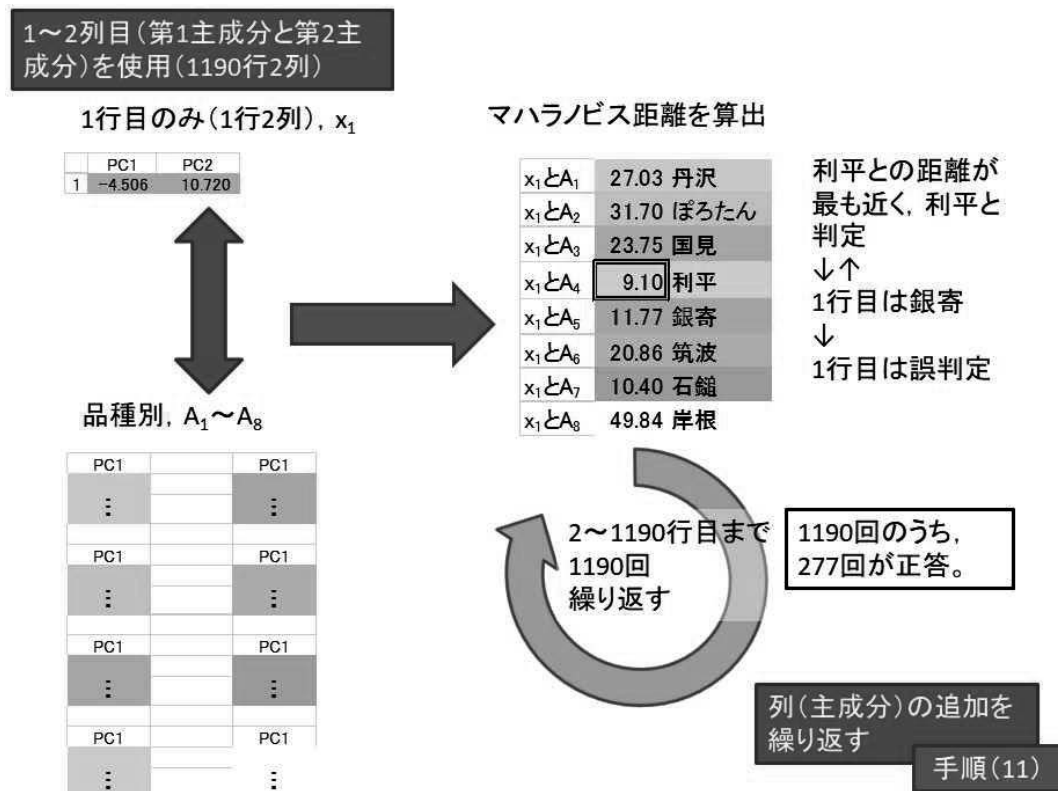


図1 (4/5) クリの輪郭をもとにした品種判定モデルの作成手順 (本文 (11))

(12) 1列目のみを抜き出した場合の正答率と、 $1 \sim 2$ 列目を抜き出した場合の正答率、 $1 \sim g$ 列を抜き出した場合の正答率を比較し、正答率が減少に転じる列数を f ($f \leq g$) とすると、 $f-1$ 列目までの主成分得点を利用する。

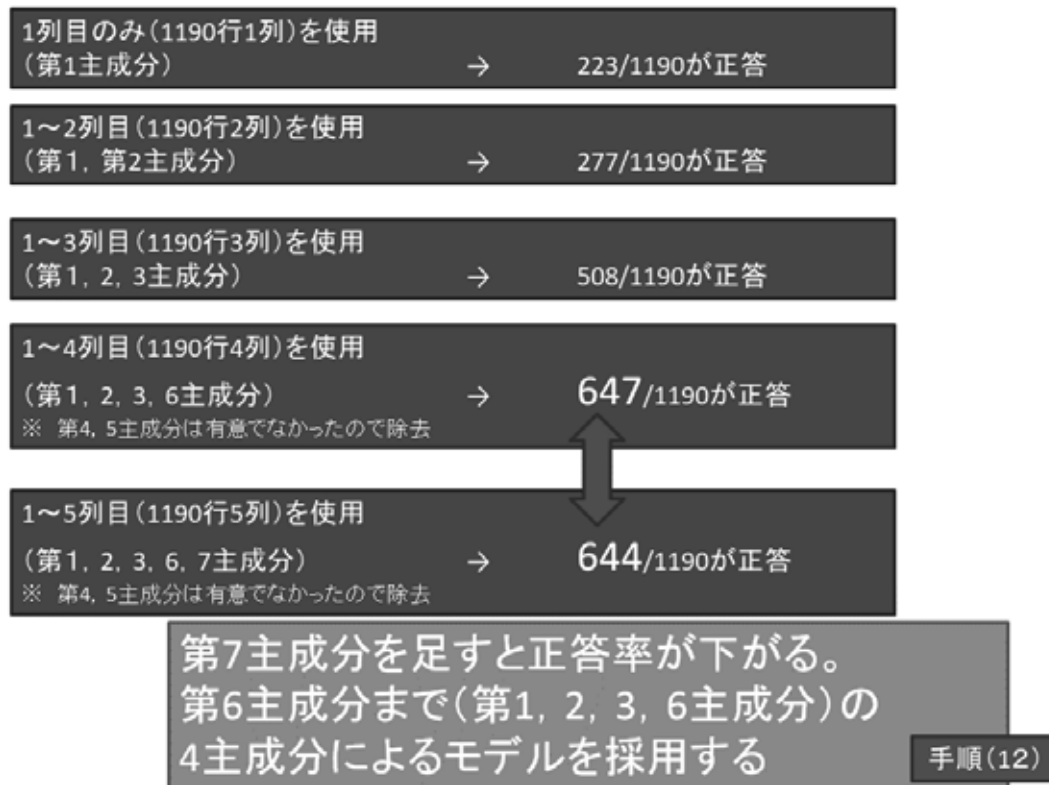


図1 (5/5) クリの輪郭をもとにした品種判定モデルの作成手順 (本文 (12))

ところで、クリの収穫時期は品種ごとに概ね一定しており、品種ごとに収穫し、コンテナ等に入れて貯蔵される。そのため、果実個体での判定ではなく、コンテナ毎にどの品種か判定するような、果実を集団として品種判定する場面が想定される。「果実3ヶが品種Aにもっとも良く似ていて、品種Cにやや似ており、品種Bには似ていないが、4ヶは品種Bにもっとも良く似ていて、品種Aにやや似ており、品種Cには似ていない。」といったような状況で、どの品種か判定するための方法として、ボルダ方式による集計を試みた。単純集計とボルダ方式による品種判定の概念図を図2に示した。ボルダ方式は、 n 個の順位のうち、最高位に n 点、次点に $n-1$ 点、次次点に $n-2$ 点…最下位に1点と各順位に得点を傾斜配分した上で集計する方式で、品種判定の場合では「もっとも似ているわけではないが、かなり似ている」というケースも判定に反映させることができる。すなわち、最小のマハラノビス距離を与える品種候補にボルダ得点8,次に小さい距離の品種候補にボルダ得点7と、マハラノビス距離の小さい順に8から1へと点数を付し、ボルダ得点の合計が最も大きい品種候補と判定した。

品種がわからない 10 果の測定結果

番号	マハラノビス距離の小さい順		
	品種A	品種B	品種C
1	1位	3位	2位
2	2位	1位	3位
3	1位	3位	2位
4	2位	1位	3位
5	3位	2位	1位
6	2位	1位	3位
7	1位	3位	2位
8	2位	3位	1位
9	2位	3位	1位
10	2位	1位	3位

正答回数による集計

	品種A	品種B	品種C
1位	3	4	3
2位	6	1	3
3位	1	5	4

最もマハラノビス距離が短い(1位である)回数の多い、**品種 B だと判定**する。

しかし、**品種 B だとマハラノビス距離が遠い(3位の)回数も多い。**

ボルダ方式による集計

番号	ボルダ得点		
	品種A	品種B	品種C
1	3点	1点	2点
2	2点	3点	1点
3	3点	1点	2点
4	2点	3点	1点
5	1点	2点	3点
6	2点	3点	1点
7	3点	1点	2点
8	2点	1点	3点
9	2点	1点	3点
10	2点	3点	1点
合計	22点	19点	19点

ボルダ方式では、1位に3点、2位に2点、3位に3点で集計する。

↓

最高得点の**品種 A だと判定**する。

図2 マハラノビス距離の単純集計とボルダ方式による集計の比較概念図

3 結果および考察

3.1 果皮色による品種判定

果皮色は貯蔵に伴い変化した(図3)。貯蔵に伴い、明度の低い色(各図の横軸左方向)が減少し、明度の高い色(横軸右方向)が増加した。また、さらに貯蔵により特定少数の色に集中する傾向が見られた。貯蔵期間の長短でも果皮色占有率が変化した。(図3(2/3)および図3(3/3)の貯蔵後1と貯蔵後2を品種ごとに比較)。

以上のように、果皮色は貯蔵により変化した。そのため、貯蔵後の任意の時期で品種判定を行うことに、果皮色の情報は不相当と考えられる。

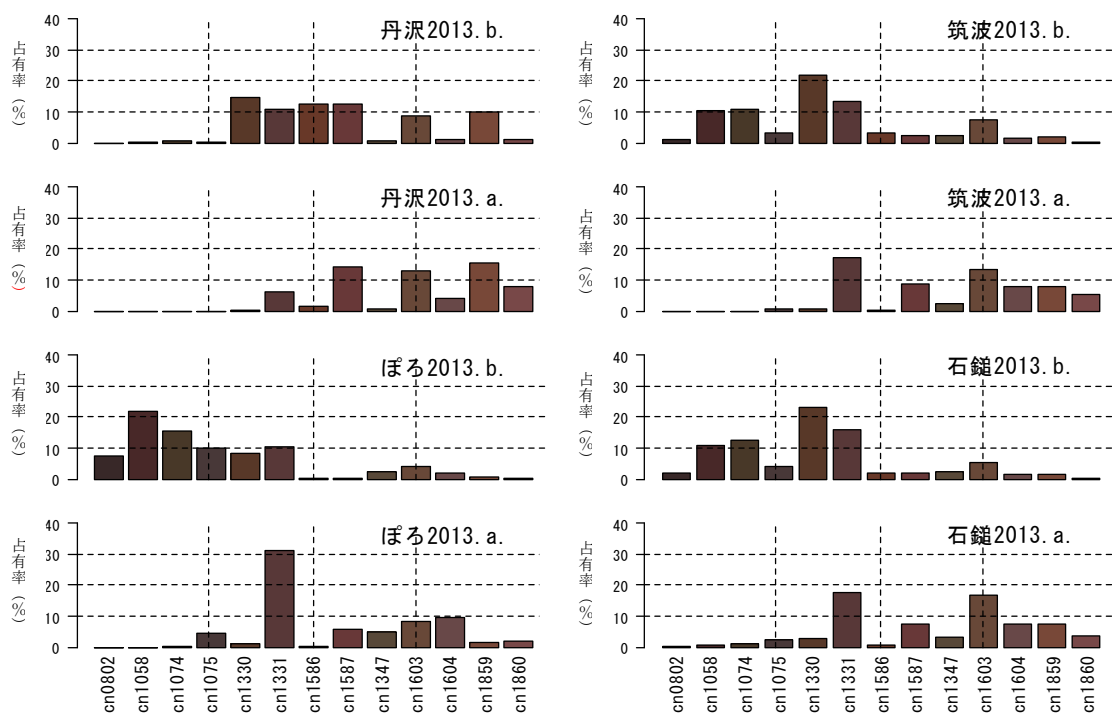


図 3(1/3) 貯蔵前後の品種ごと平均果皮色 (2013 年, ‘ぼろたん’ は「ぼろ」と略記)
 ※ 左に明度の低い, 右に明度の高い色を配置。貯蔵前果実 826 果を対象に, 各色の占有率の標準偏差が 3%を超えた 13 色のみを掲載した。横軸の“cn”を付した番号はビジュアルアナライザーの色番号
 b: 貯蔵前 (before), a: 貯蔵後 (after, 2013 年 11 月 11 日)

図 3 付表 ビジュアルアナライザーの色番号と RGB 値および色名*

色番号	R	G	B	色名
802	56	40	40	dark reddish brown
1058	72	40	40	dark reddish brown
1074	72	56	40	dark grayish yellowish brown
1075	72	56	56	dark grayish reddish brown
1330	88	56	40	dark brown
1331	88	56	56	dark reddish brown
1347	88	72	56	dark grayish yellowish brown
1586	104	56	40	deep brown
1587	104	56	56	deep reddish brown
1603	104	72	56	moderate brown
1604	104	72	72	dark grayish red
1859	120	72	56	moderate reddish brown
1860	120	72	72	dark red

※ いずれもビジュアルアナライザー独自の表記

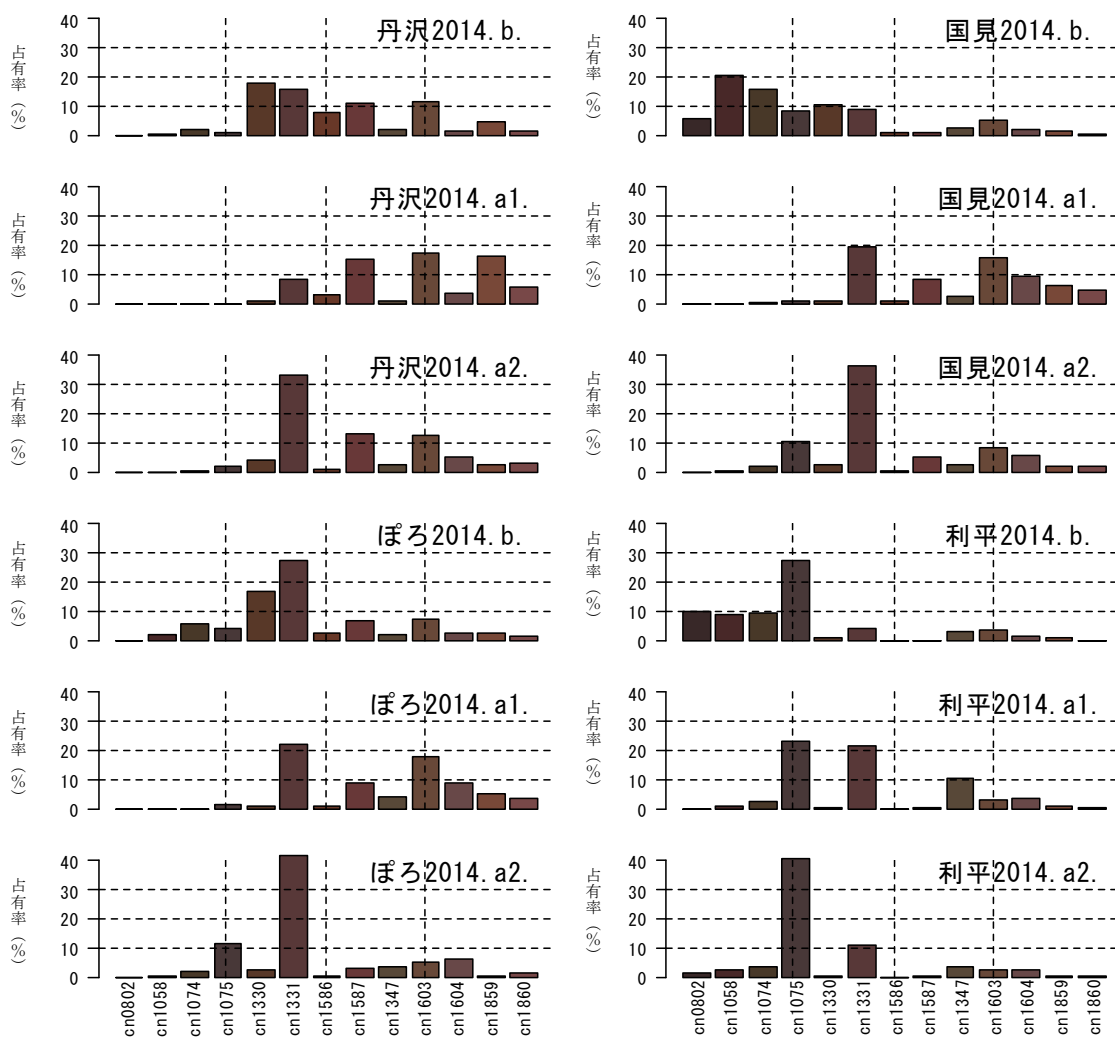


図 3(2/3) 貯蔵前後の品種ごと平均果皮色 (2014 年, 'ぼろたん' は「ぼろ」と略記)

※ b : 貯蔵前, a1 : 貯蔵後 1 (2014 年 10 月 17 日), a2 : 貯蔵後 2 (2014 年 12 月 24 日)。表 2 参照

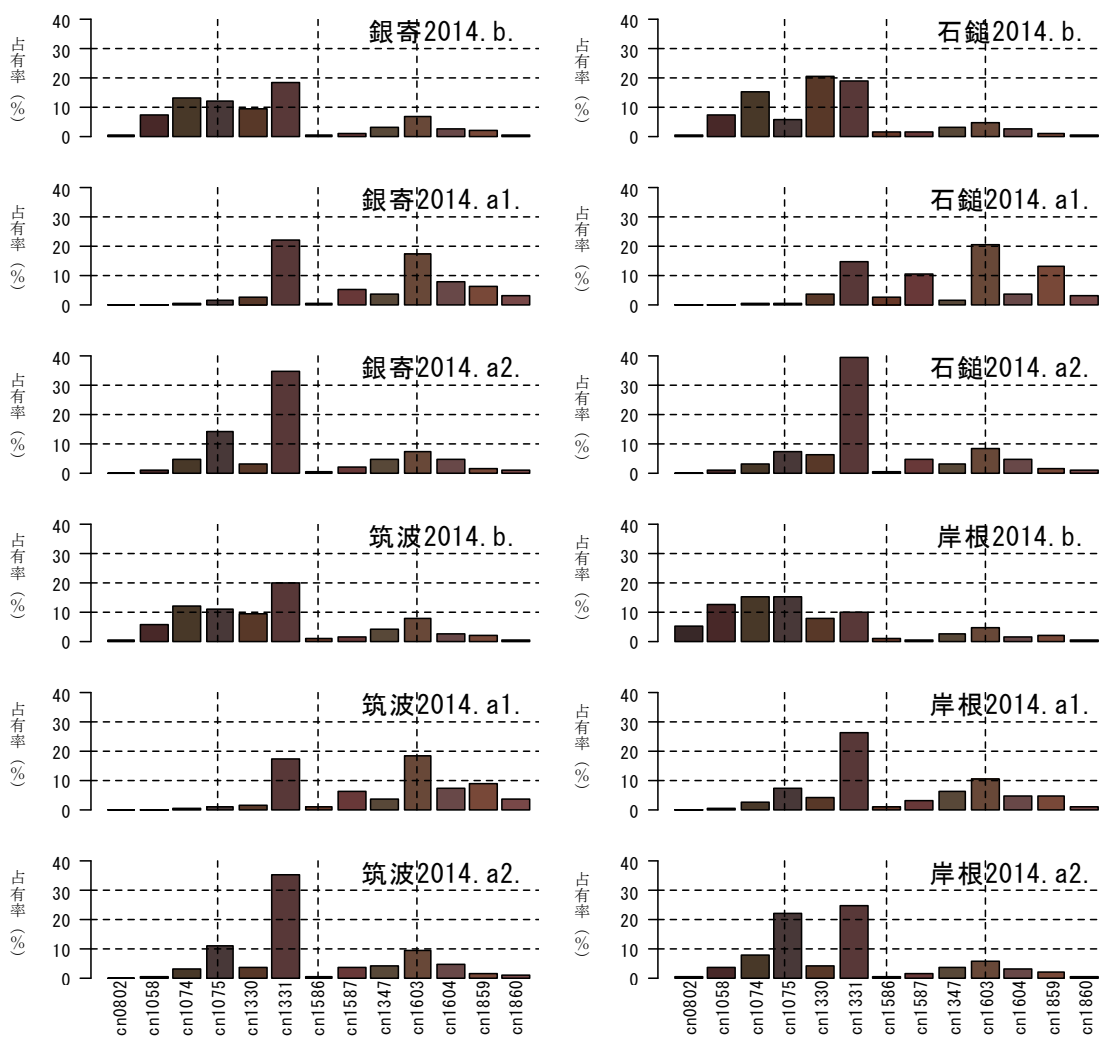


図 3(3/3) 貯蔵前後の品種ごと平均果皮色 (2014 年つき)

※ b: 貯蔵前, a1: 貯蔵後 1 (2014 年 10 月 17 日), a2: 貯蔵後 2 (2014 年 12 月 24 日)。表 2 参照

3. 2 輪郭による品種判定

2. 3により得られた標準化楕円フーリエ記述子の係数を, 年, 貯蔵前後および品種ごとに平均し, 平均したフーリエ記述子の係数から年, 貯蔵前後および品種ごとに平均果形 (輪郭) をフリーソフトウェア「R」により作図したが, 輪郭には貯蔵による変化は認められなかった (図 4)。

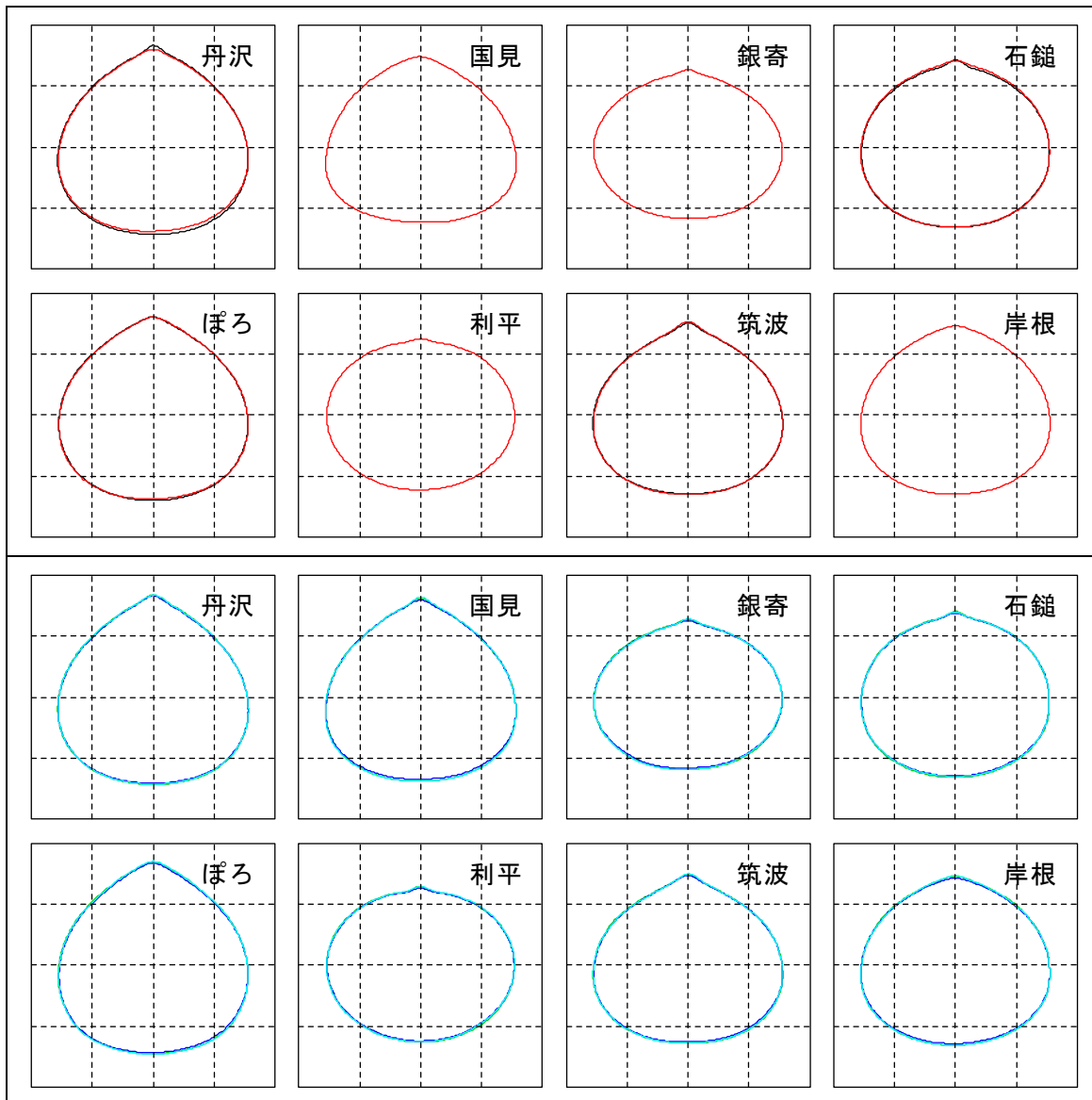


図4 貯蔵前後の品種別平均果形（輪郭）

※ 上段は2013年で、貯蔵前が黒線、貯蔵後が赤線

下段は2014年で、貯蔵前が緑線、貯蔵後1が青線、貯蔵後2が水色線

輪郭には貯蔵による変化は認められなかったため、以後2013年は貯蔵後の8品種484果、2014年の貯蔵前706果、合わせて1190果を対象に検討を行った。

2.3により、各果実の標準化楕円フーリエ記述子の係数の主成分分析を行ったところ、品種差が優位な44主成分を得た(図1(1/5))。44主成分のうち、品種の正答率が最も高くなる第1, 2, 3, 6主成分の4主成分による品種判定モデルを採用した(図1(5/5))。その結果、合計1190果中647果が正しい品種に判定され、全果実数ベースでの正答率が54%の判定モデルが作成できた(表5)。

なお、正答率が低い品種は‘丹沢’、‘国見’、‘岸根’だった。‘丹沢’と‘国見’では、互いの誤判定が多かったが‘国見’は‘丹沢’の子品種であり、また、‘岸根’は‘筑波’や‘石鎚’との誤判定が多かったが‘筑波’‘石鎚’ともに‘岸根’の子品種であり、これらの遺伝的関係が形質(果形)に影響したものと推測されるが、詳細については遺伝的解析も合わせた検討が必要である。

表5 果形（輪郭）からの品種判定結果

	判定結果(ヶ)								果数 小計 (ヶ)	正答 数 ¹⁾ (ヶ)	正答 率 ¹⁾ (%)	
	丹沢	ぼろ たん	国見	利平	銀寄	筑波	石鎚	岸根				
元の 品種 ²⁾	丹沢	53 (35)	24 (16)	39 (26)	0 (0)	0 (0)	21 (14)	1 (1)	14 (9)	152	53	35
	ぼろ たん	21 (15)	86 (60)	19 (13)	0 (0)	0 (0)	6 (4)	1 (1)	10 (7)	143	86	60
	国見	14 (22)	9 (14)	25 (39)	0 (0)	0 (0)	7 (11)	3 (5)	6 (9)	64	25	39
	利平	0 (0)	0 (0)	0 (0)	101 (68)	36 (24)	0 (0)	12 (8)	0 (0)	149	101	68
	銀寄	0 (0)	0 (0)	0 (0)	29 (22)	83 (63)	5 (4)	15 (11)	0 (0)	132	83	63
	筑波	13 (6)	2 (1)	19 (9)	0 (0)	8 (4)	129 (63)	20 (10)	15 (7)	206	129	63
	石鎚	0 (0)	0 (0)	7 (4)	13 (8)	15 (9)	13 (8)	107 (64)	12 (7)	167	107	64
	岸根	7 (4)	10 (6)	21 (12)	1 (1)	6 (3)	32 (18)	37 (21)	63 (36)	177	63	36
									1190	647	54	

1) 元の品種がどのくらい正確に判定されているか評価

例：‘丹沢’ 果実を判定すると、35%が‘丹沢’と正しく判定される。

2) 上段は果数（ヶ），下段のカッコ付き数字は果数小計を母数とした割合（%）

ボルダ方式による集計結果を図5に示した。8品種共に本来の品種が最も高いボルダ得点を獲得しており、正しく判定できたと認められる。

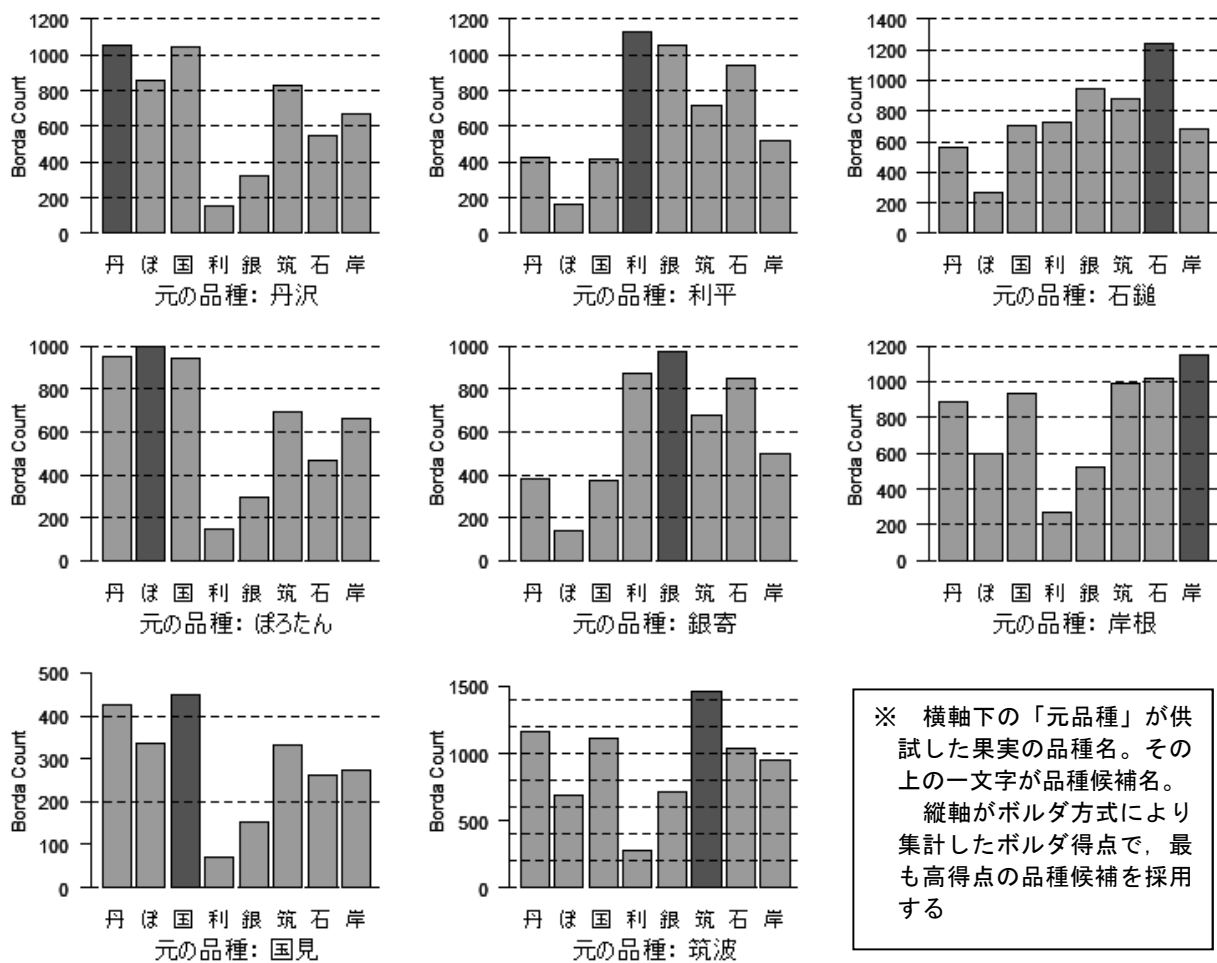


図5 ボルダ方式による集団の品種判定結果

輪郭からの品種判定では、一果実を対象に判定する場合（最小のマハラノビス距離となる品種と判定する場合）では54%程度の正答率に留まる。一方、果実集団を対象とする場合は、マハラノビス距離順の集計にボルダ方式を採用することで、品種判定が可能と考えられた。

今回の試験では、品種毎に中庸な大きさの果実を使用したが、大きさの違いが輪郭に差異を生じるかについては未検討である。

また、近年クリでは‘美玖里’や‘ぼろすけ’といった新品種も出てきており、今後栽培面積も増えることが予想されるので、今後はそれらの品種も併せた上で、判定可能か検証する必要がある。

摘要

クリ8品種を対象に、側果外側面側の外観を定量化した。クリ果実の果皮色は貯蔵により変化することが分かった。貯蔵により変化しない果形の画像情報から、クリ8品種の判定モデルを構築した。一果実の判定では正答率は54%程度にとどまるが、果実集団を対象とする場合はボルダ方式で集計することで品種を正確に判定できた。

引用文献

Iwata H. and Ukai Y (2002) SHAPE : A computer program package for quantitative evaluation of biological shapes based on elliptic Fourier descriptors. Journal of Heredity 93:384-385

門田有希 (2017) 農作物・食品の品種判定検査技術の開発日本の大切な品種を守るために。化学と生物。

55(12) : 817-824

佐野健人・森田陽一・鹿島恭子 (2012) 近赤外分光法によるクリ ‘ぼろたん’ と ‘国見’ の非破壊判別.

茨城農総セ園芸研究所研究報告. 19 : 47-52

矢野 博 (2004) DNA 多型分析による品種識別の可能性. 農業および園芸. 79(1) : 131-136

Summary

An attempt was made to construct a variety determination model for eight chestnut varieties by analyzing the appearance image of the lateral side. As a result of analyzing the skin color and contour as the appearance, the skin color changed with storage, but the contour did not change with storage, so the contour was considered to be appropriate as a judgment item of the variety. A contour determination model was converted into a Fourier descriptor, a principal component analysis was performed, a Mahalanobis distance was calculated, and a model for determining a group that gave the smallest Mahalanobis distance as a corresponding variety could be constructed. As a result of verifying 1190 fruits with a judgment model, the correct answer rate as an individual judgment was 54%. On the other hand, containers are often harvested for each tree (variety) at the production site, so if the boulder method is used as a model for determining the variety of groups such as containers, it was possible to correctly determine all 8 types.

Keywords: chestnut, image analysis, cultivars determination, contour, skin color