

## 第3章 ラン実生無菌培養苗の 画像処理による形状認識

---

### 第1節 本章の目的

第2章では、苗の選別を、上から見た葉の投影面積によって行った。しかしながら、苗の葉は、それぞれの個体で傾きが異なるため、投影面積が苗の大きさを的確に表しているとはいえない。そこで、本章では、培養容器から取り出された苗の横から見た形状を用いて選別を行う。また、大きさだけではなく、苗の良否を示す形状の特徴パラメータについての検討を行う。形状特徴の抽出方法として、これまでも研究例のある、図形の重心から境界線までの距離を等角度間隔にサンプリングする方法を用いる。また、得られたサンプル値をフーリエ変換し、苗の形状を表すことが可能か検討を行う。

### 第2節 実験材料および装置

#### 第1項 ラン実生苗

実験試料として、無菌培養されたラン(*Paph. (Spotglen x Mahasuka) x Paph. Arapaho*)を用いた。これは、堂ヶ島洋らんセンターから入手した播種後約12ヶ月の実生苗である。この苗は播種後3回の移植を経ており、出荷前の最後の移植をされた直後のものである(図1)。マシンビジョンによる判定結果と比較するため、熟練者によって大、中、小、不良に分けられている苗を用いた。この品種の苗は、ランの中では比較的単純な形状をしており、左右に展開した葉を持っている。今回使用した苗では、良苗と判断できる苗は、左右にほぼ均等に展開している2~3枚の葉と1本の根を備えて

いる(図 2(a), (b)). 一方, 不良苗は葉が 1 枚だけのものや(図 3(a)), 片方の葉が極端に短いもの, 正常な生長をしておらず複雑な形状をしているもの(図 3(b))がある.

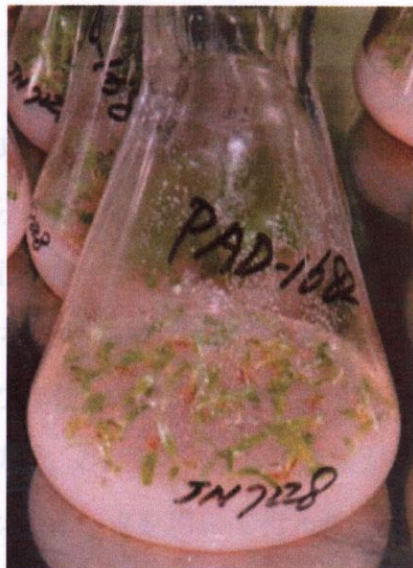


図 1 最終移植直後のフラスコ苗

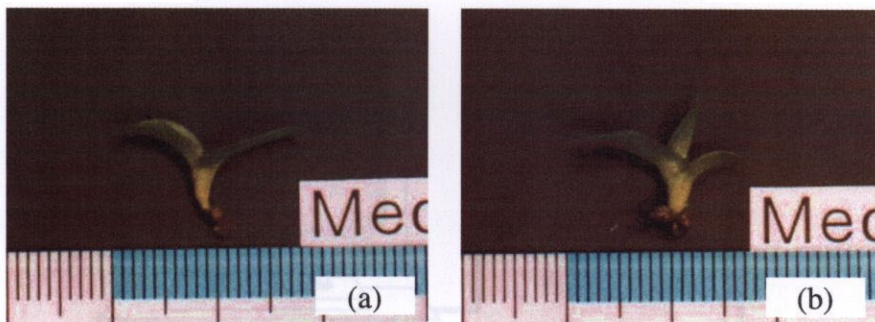


図 2 良苗の例

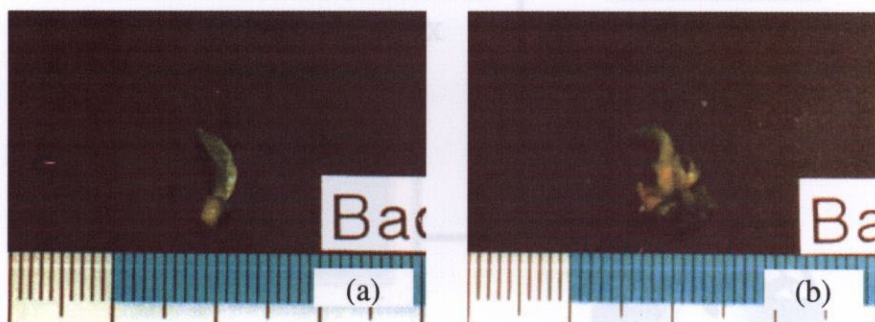


図 3 不良苗の例

## 第2項 画像処理装置

画像入力装置として、モノクロ MOS 型カメラ (EPSON, GT-20), パーソナルコンピュータ (NEC, PC-9801VX), カメラから入力される画像のモニターとして, AV モニター (Victor) を用いた. 画像処理アルゴリズムの開発には N88BASIC を用いた. 画像処理の処理に時間を要する部分は, アセンブラ言語 MASM を使用し, マシン語で実行した (木谷, 1992). カメラの画素数は  $244 \times 244$  ピクセル. 輝度の解像度は 4 ビット 16 階調であった. モノクロカメラを三脚に取り付け, 垂直下方を向けて撮影を行った. 照明には蛍光灯を用いた. 苗の背景として, つや消しの黒い塩ビの板を使用した. これによって蛍光灯の乱反射を防ぎ, ノイズを防ぐ上で効果があった. 統計処理や, 選別結果の出力はパーソナルコンピュータを用いて行った. 統計処理のソフトとして, Math Works 社の Matlab 6.5 (The MathWorks, 2000) を用いた. このソフトは, 行列計算に優れている.

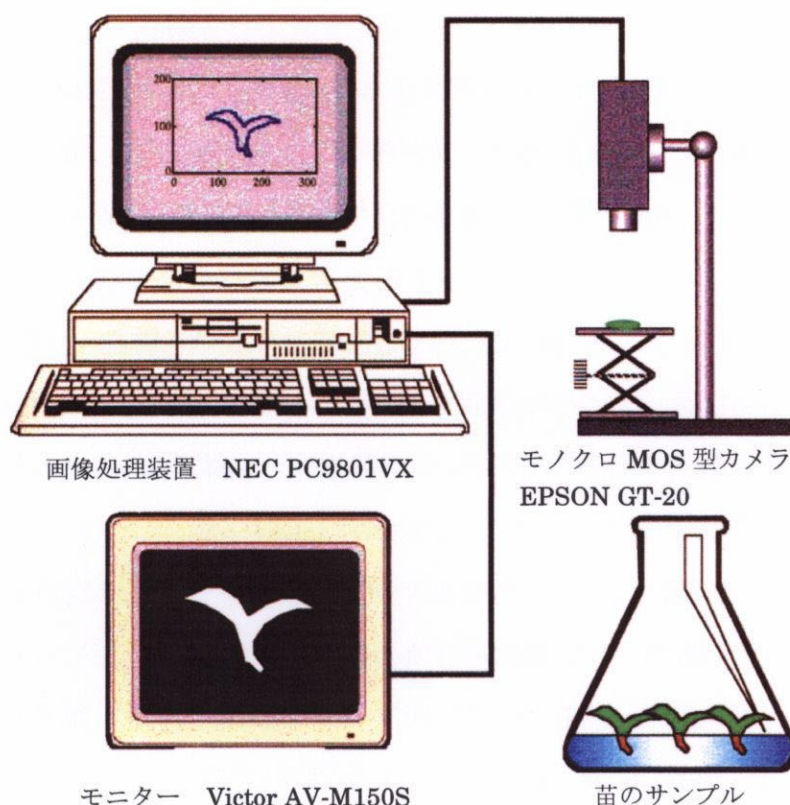


図 4 実験装置の構成図

## 第3節 画像認識アルゴリズム

### 第1項 境界線の表現方法

任意の二次元の閉曲線は図心(重心)から曲線までの距離と角度によって近似することが可能である(図 5) (Dubois and Glanz, 1986). 図中,  $t$  はサンプル番号,  $R(t)$  は  $t$  番目の重心から曲線までの距離,  $\Delta\theta$  は式(1)で表されるサンプリングの間隔である.

$$\Delta\theta = 360 / N \quad (1)$$

ここで  $N$  はサンプル数である.  $N$  の数を増やしたり, 減らしたりすることで, 近似の細かさを細かくしたり, 粗くしたりする事が可能である.

図 6に横軸を  $t$ : サンプル番号, 縦軸を  $R(t)$ : 図心から境界線までの距離で表したグラフを示す. 図 6の(a), (b)はそれぞれ, 図 5の(a), (b)に対応している. それぞれの境界線の形状によって, 異なるパターンが得られる事がわかる. 本実験ではこのパターンを用いて, パターン認識を行うこととする.

この境界線の近似方法は, 図心からある角度に引いた直線が, 境界線上の 1 点とだけ交わる場合は, 正しく近似を行う事が可能である. しかし一方, 境界線に凹な部分があり, 直線が  $3+2n(n=0,1,2,\dots)$  点で交わる場合は(図 7(a)), 一番外側の交点のみを採用するため, 凹な部分の情報が欠損するという欠点がある. また, 図心が図形の外側にあるときは, 直線との交点が存在しない. この場合, その角度における距離を 0 とする. その為, 元の図形とは異なった形状の近似がなされる(図 7(b)).

図心を中心として, 境界線までの距離をサンプリングする手法はこれまでにいくつかの研究例がある. Dubois and Glanz(1986)は, アルファベットや工業製品を対象とし, サンプル値を自己回帰モデルによってモデル化することで, 選別を行った. 図形の凹んだ部分もサンプリングするために, 重心から等角度で引いた線と境界線の交わった点を全て, 境界線に沿った順番でサンプリングした. また, 一部が隠れた図形についても選別を行ったが, 良い結果は得られなかった. Tao et al.(1995)は, 等角度に線を引くのではなく, 境界線上の点と重心の距離を順番にサンプルして, フーリエ係数を求めた.

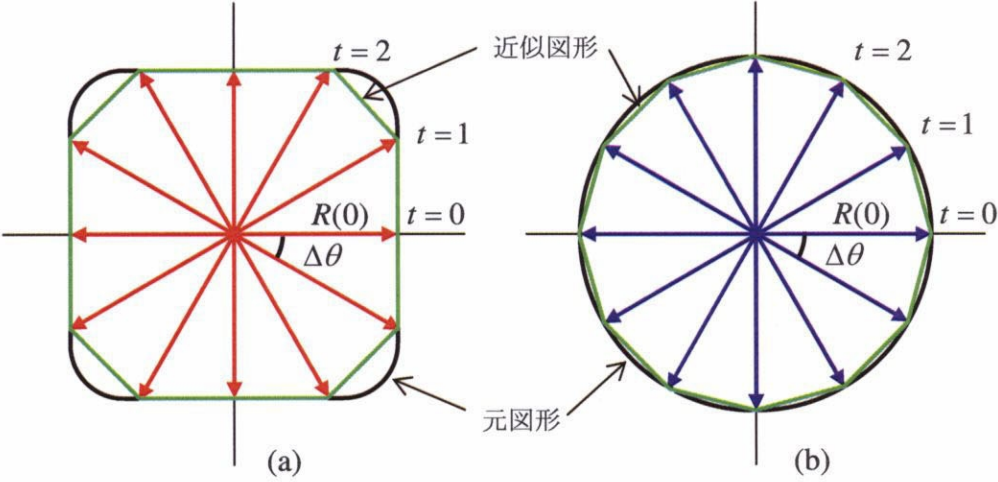


図 5 閉曲線の近似方法

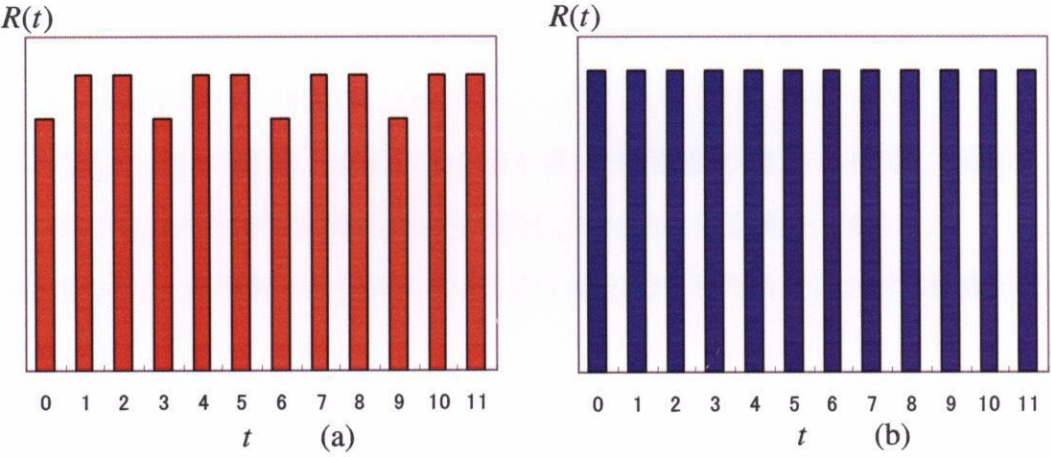


図 6 図心からの距離

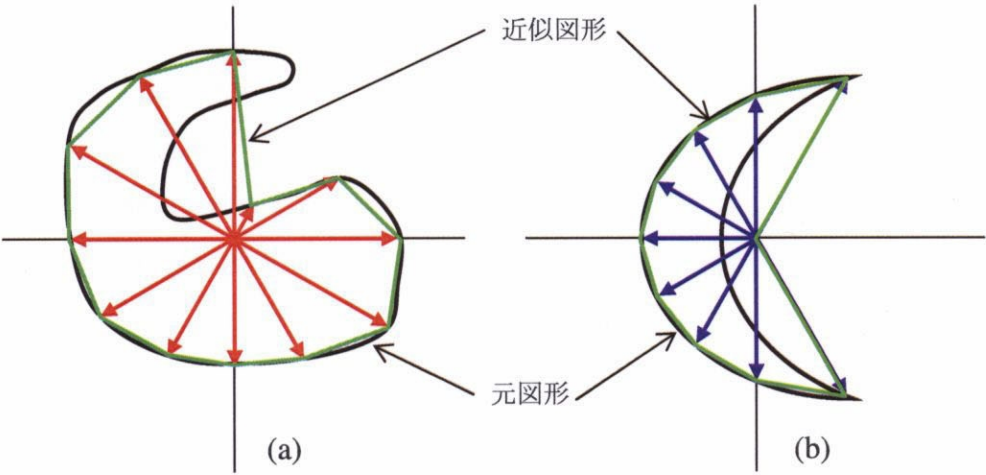


図 7 凹な部分を持つ閉曲線の場合

## 第2項 サンプルングの方法

以下に具体的なサンプルングの方法を示す.

1. 4ビット16階調の濃淡画像を適当な閾値で二値化する. その際, 閾値は, 照明条件が変わらない限り固定値とする.
2. ラベリング処理を行い, 最も大きな領域のみを抽出する.
3. 抽出した領域の図心 $(xc, yc)$ を式(2)より求める(井上ら, 1999).

$$(xc, yc) = \left( \frac{1}{n} \sum_{i=0}^{n-1} x_i, \frac{1}{n} \sum_{i=0}^{n-1} y_i \right) \quad (2)$$

$n$  は, 抽出領域のピクセル数,  $x_i$  と  $y_i$  は, 抽出領域に含まれるピクセルの座標である.

4. 図心から直線を仮想的に等角度で引く. 本実験においては, 一周 $360^\circ$ を64等分した. ゆえに $\Delta\theta$ は,  $5.625^\circ$ となった.
5. 各角度において直線上を図心から最も遠い所(画像の端)から探索し, 最初に境界線に出会った場所を境界線の位置とし, 図心からの距離を記録した.
6. 前述の通り, 境界線との交点が存在しない場合は, その角度における距離は0とした.

## 第3項 離散的フーリエ級数展開

本実験は, 境界線の形状による, 苗のグループ分けを目的としている. 一般的に, 複数の特徴量を用いて, 未知の対象物がどのグループに属するかを決定する手法をパターン認識と呼ぶ(Duda and Hart, 2001). パターン認識を行う場合, まず前段階として, グループ間に顕著な差がある特徴量を見つける必要がある. 本節で求めた, 図心からの距離をそのまま特徴量として用いることも可能だが, 図心からの距離は, 角度との対応付けがされているため, 同じ図形であってもその置かれた角度によって, 特徴量が変わってしまうという欠点がある. そこで, 求めた図心からの距離が周期関数とみなす事ができるという利点を生かし, 離散的フーリエ級数展開を行い, その各次のフーリエ係数を特徴量として用いる事とする. サンプルングの間隔が, 実用上十分な細かさであれば, フーリエ係数は, 図形の回転に影響されないと言う事ができる. 以下に離

散的フーリエ級数展開の式を示す(安居院ら, 1994).  $R(t)$  が以下のように級数展開されるとする.

$$R(t) = a(0) + a(1) \cdot \cos 2\pi / N + a(2) \cdot \cos 4\pi / N + a(1) \cdot \cos 6\pi / N + \dots \\ + b(1) \cdot \sin 2\pi / N + b(2) \cdot \sin 4\pi / N + b(3) \cdot \sin 6\pi / N + \dots \quad (3)$$

この時,  $a(n), b(n) (0 \leq n < N)$  は  $n$  次のフーリエ係数である.  $t$  はサンプル番号,  $N$  はサンプル数である.  $a(n), b(n)$  はそれぞれ以下の式によって求める事ができる.

$$a(0) = \frac{1}{N} \sum_{t=0}^{N-1} R(t) \quad (4)$$

$$a(n) = \frac{1}{N} \sum_{t=0}^{N-1} R(t) \cos(nt \frac{2\pi}{N}) \quad (5)$$

$$b(n) = \frac{1}{N} \sum_{t=0}^{N-1} R(t) \sin(nt \frac{2\pi}{N}) \quad (6)$$

$n$  次の正弦波(周期= $N/n$ )の位相は次式で表される.

$$\phi(n) = \tan^{-1} \{b(n)/a(n)\} \quad (7)$$

また,  $n$  次の正弦波(周期= $N/n$ )の振幅は次式で表される.

$$f(n) = \sqrt{a(n)^2 + b(n)^2} \quad (7)$$

以降  $f(n)$  を  $n$  次のフーリエ係数と呼ぶ.  $f(n)$  を  $n=11$  まで求め, 特徴量とした.  $n$  は実際には  $2/N$  まで求める事が可能だが, 本実験においては  $n=12$  以上の高次のフーリエ係数は必要ないと考え, 計算しなかった.

この, フーリエ係数を求めることでその閉曲線の特徴を得ることが可能となる. あらかじめ予測しうる何種類かの図形のフーリエ係数の特徴パターンがわかっているならば, それらと未知の図形のフーリエ係数の特徴パターンを比較することで, 未知の図形がどの図形に近いかを推測することができると考えられる.

## 第4項 その他の特徴量

前項で求めたフーリエ係数の  $f(0)$  は, 図心から境界線までの距離の平均値を表している. そのため, この値は苗のおよその大きさを表しているといえる. しかし, 前述の通り, 苗に窪んだ部分があると, 正しい近似を行う事ができない. その為,  $f(0)$  は常に実際の面積よりも若干大きな値を示すと考えられる. そこで苗の大きさについての特徴

量として、抽出された領域の面積を用いることとした。また、苗の大きさは、その葉の大きさによっても表す事ができる。そこで、図心から境界線までの距離の最大値も、特徴量として用いることとした。

## 第4節 パターン認識アルゴリズム

### 第1項 マハラノビス距離

パターン認識のアルゴリズムには、様々な種類が存在する(Duda and Hart, 2001)。その中でも最もシンプルな方法であるパターン間最小距離を用いる(安居院と長尾, 1996a)。この方法では、入力パターンと、あらかじめ記憶した各クラスタ(グループ)の標準パターンとの間の距離を計算し、最小距離を与える標準パターンを入力パターンに対する識別結果とする方法である。距離を表す方法は、ユークリッド距離や市街区距離、マハラノビスの距離などがある。本章ではマハラノビスの距離を用いてグループ分けを行う。

マハラノビスの距離  $D_j$  は次のように定義される。

ある、クラスタ  $C_j$  に属する  $n_j$  個のパターン  $\mathbf{P}_k (\in C_j, k \in N_j \equiv \{1, 2, \dots, n_j\})$  の平均のパターン  $\boldsymbol{\mu}_j$  と、共分散行列  $\boldsymbol{\Sigma}_j$  を次式から求める。

$$\boldsymbol{\mu}_j = \frac{1}{n_j} \sum_{m \in N_j} \mathbf{P}_m \quad (8)$$

$$\boldsymbol{\Sigma}_j = \frac{1}{n_j} \sum_{m \in N_j} (\mathbf{P}_m - \boldsymbol{\mu}_j)(\mathbf{P}_m - \boldsymbol{\mu}_j)^T \quad (9)$$

このとき、未知のパターン  $\mathbf{P}_i$  と、クラスタ  $C_j$  の間のマハラノビス距離  $D_j$  は次の式で表される。

$$D_j^2 = (\mathbf{P}_i - \boldsymbol{\mu}_j)^T \boldsymbol{\Sigma}_j^{-1} (\mathbf{P}_i - \boldsymbol{\mu}_j) \quad (10)$$

未知のパターン  $\mathbf{P}_i$  について  $D_j^2$  を求め、 $D_j^2$  を最も小さくする  $j$  のクラスタ  $C_j$  に属すると判断する。

## 第5節 苗の測定結果

### 第1項 目視による形状の分類

大, 中, 小の良苗をそれぞれ 40 個ずつ, 不良苗を 51 個, 計 171 個の苗をサンプルとして用いた. まず, 目視により, 苗の形状を記録した. 今回使用した苗の形状の特徴として最も適切なものは, 葉の枚数である. 苗は葉の枚数によって成長の度合いが異なる. また, 葉の枚数に応じて, フーリエ係数のパターンが変化すると予想される. 表 1 に等級と葉の枚数の関係を示す. 表中で「不明」とあるものは, 苗が正常な生長をしておらず, 葉の枚数が把握できないものを指す. このような苗を以後未発達な苗と呼ぶ. 2 枚(1 短)は, 葉が 2 枚観察できるが, 1 枚がもう一方に比べて, 極端に短いもの, 2 枚(均等)は, 葉が 2 枚均等に展開しているものを指す. また同じように 3 枚(1 短)は, 葉が 3 枚あるが 1 枚が極端に短いもの, 3 枚(均等)は, 3 枚の葉が均等に展開しているものである. この分類結果より以下の事が推測される.

表 1 等級と葉の枚数の関係

|        | 不明 | 1 枚 | 2 枚(1 短) | 2 枚(均等) | 3 枚(1 短) | 3 枚(均等) | 計   |
|--------|----|-----|----------|---------|----------|---------|-----|
| 良苗(小)  | 0  | 1   | 10       | 22      | 4        | 3       | 40  |
| 良苗(中)  | 0  | 0   | 0        | 25      | 4        | 11      | 40  |
| 良苗(大)  | 0  | 0   | 1        | 28      | 6        | 5       | 40  |
| 良苗(合計) | 0  | 1   | 11       | 75      | 14       | 19      | 120 |
| 不良苗    | 28 | 6   | 12       | 4       | 1        | 0       | 51  |
| 全合計    | 28 | 7   | 23       | 79      | 15       | 19      | 171 |

1. 良苗は, その 90%が, 2 枚の均等に展開した葉か, 3 枚以上の葉を持っている.
2. 不良苗は正常な生長をしていない萎縮した苗が 50.9%と最も多く, 次いで葉が 2 枚確認できるが, 1 枚が極端に短い苗が 23.5%と多く見られた. また, 葉が 1 枚しか確認できないものは 11.8%だった. 2 枚均等に展開しているものも, その葉の長さは短かった,

3. 良苗(小)は, 葉が 2 枚均等に展開している苗が最も多かった. 次いで, 不良苗のように, 葉が 2 枚確認できるが, 1 枚が短い苗が存在した. 葉が 3 枚のものは少なかった.
4. 良苗(中)は, 葉が 2 枚均等に展開しているものが最も多く, 次いで 3 枚均等に展開しているものも多く見られた.
5. 良苗(大)も同じく, 2 枚均等に展開している苗が最も多かった.
6. 良苗と不良苗を分離するには, 1) 未発達な苗, 2) 葉が 1 枚しかない苗, 3) 葉が 2 枚だがその内の 1 枚が極端に短い苗を形状より抽出すれば良い事がわかる. 葉が 2 枚だがその内の 1 枚が短い苗については, 良苗(小)である場合もあるので, 苗の大きさにも着目して分離を行う必要がある.

## 第2項 苗の形状とフーリエ係数の分布

良苗, 不良苗に関係なく, 苗をその形状で分類したときの, 形状とフーリエ係数のパターンの関係を以下に述べる. 以下, 図心から境界線までの距離をサンプル値と呼ぶ.

### 1. 苗が未発達で, 葉の枚数が把握できない場合

苗が, 図 8(a)のように未発達な場合図 8(b)のような境界線画像と, 図 8(c)のような近似画像を得る事ができる. この場合, はっきりとした葉が見られないので, 図心から境界線までの距離(以後サンプル値)の変動が少ない(図 8(d)). そのため, 低次のフーリエ係数には顕著なピークは見られない(図 8(e)).

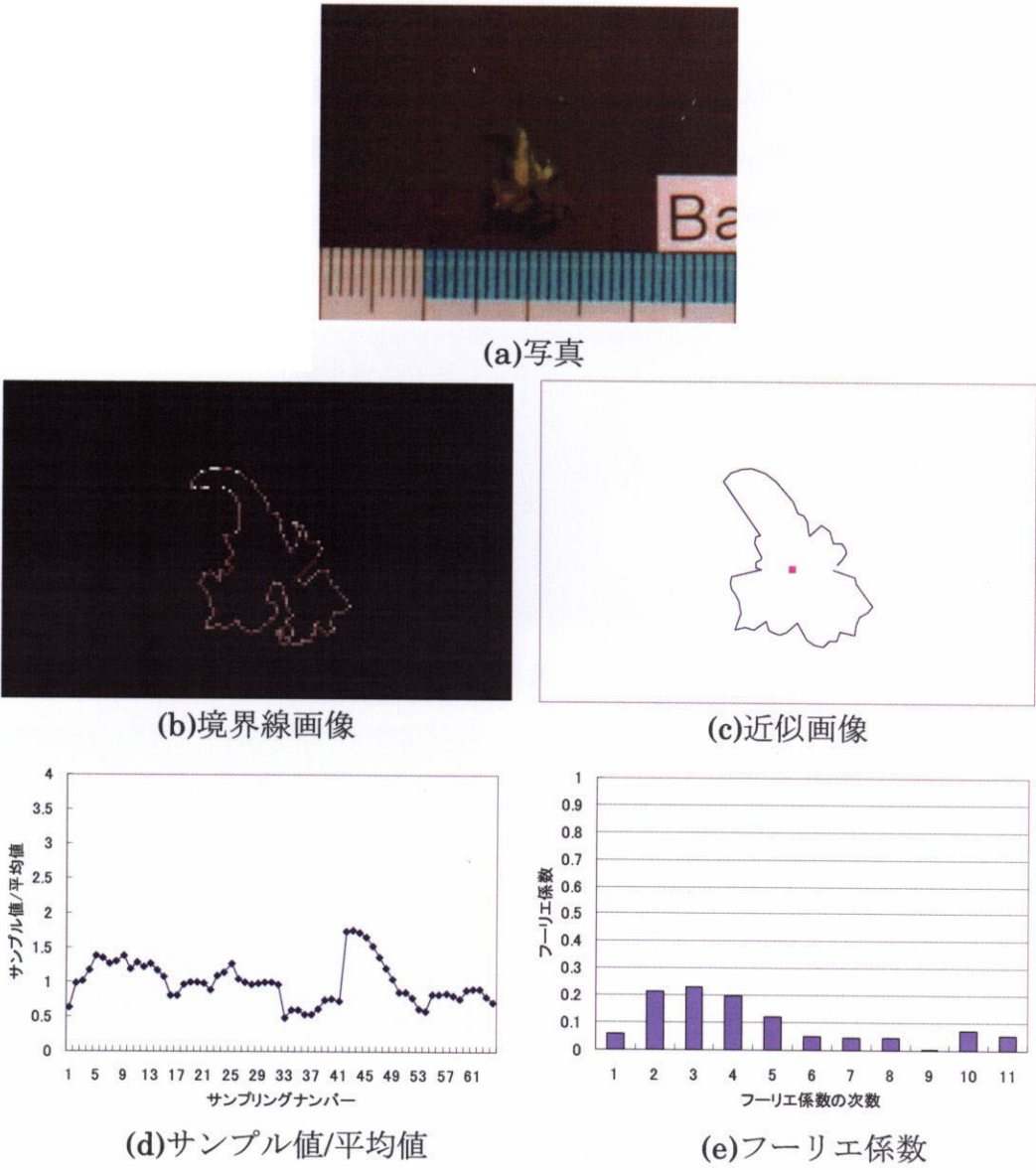
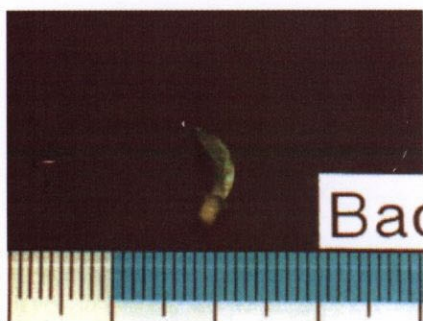


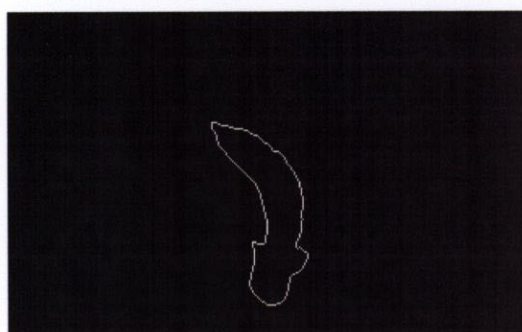
図 8 苗が未発達な場合

## 2. 葉が 1 枚の場合

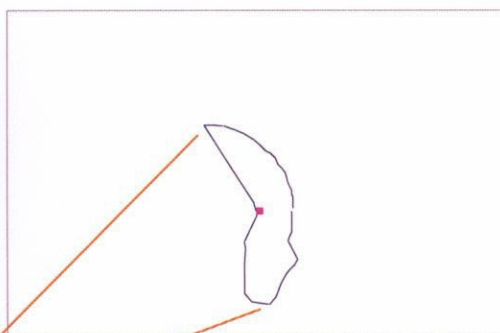
図 9(a)の様に葉が 1 枚の場合, 境界線は, 図心を中心にほぼ  $180^\circ$  に展開する. そのため, サンプル値には2つの大きなピークが生じる. その結果, 2 次のフーリエ係数が他の次数のものに比べて大きくなる. 図中の矢印は対応するピークを表している.



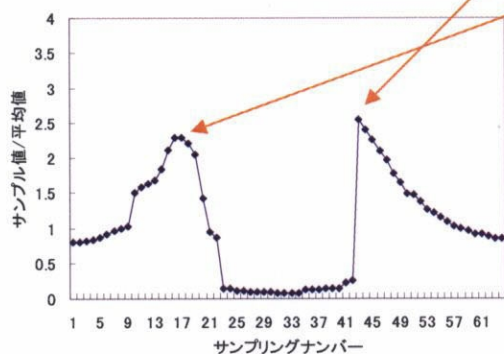
(a)写真



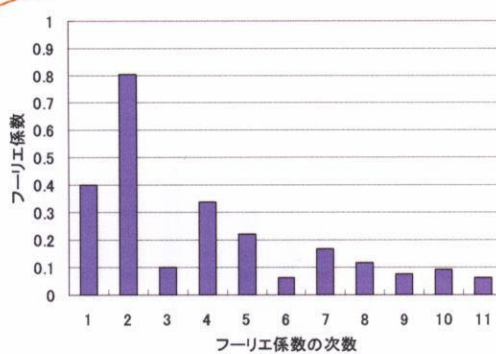
(b)境界線画像



(c)近似画像



(d)サンプル値/平均値

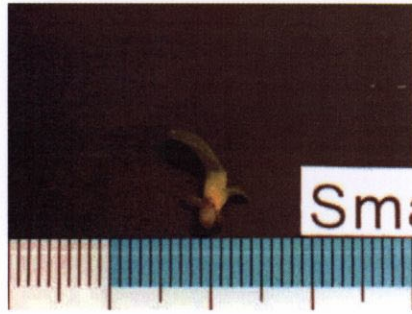


(e)フーリエ係数

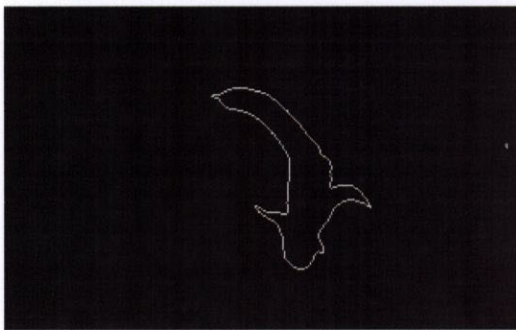
図 9 葉が 1 枚の場合

### 3. 葉が2枚だが1枚が短い場合

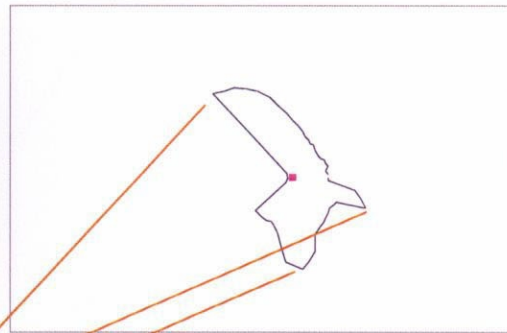
図 10(a)の様に葉が 2 枚確認できるが、1 枚がもう片方と比較して短い苗の場合、境界線は、葉が 1 枚の場合と似たものとなる。しかし、短い葉があるため、サンプル値には2つの大きなピークに加えてもう1つピークが生じる。その結果、2 次のフーリエ係数以外に 3 次も大きくなる。



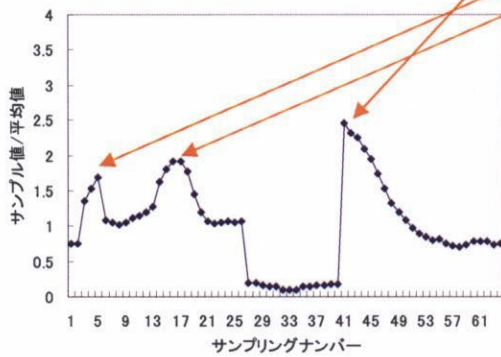
(a)写真



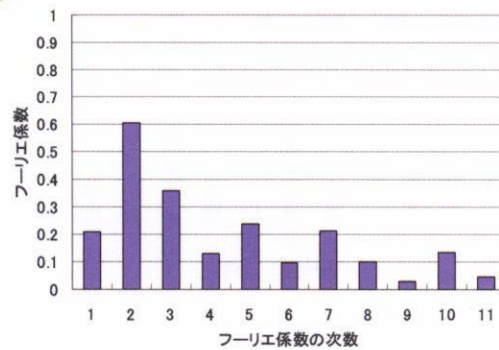
(b)境界線画像



(c)近似画像



(d)サンプル値/平均値



(e)フーリエ係数

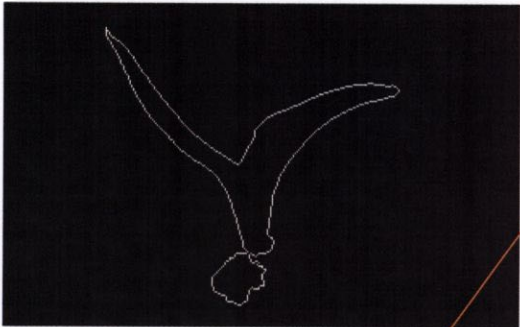
図 10 葉が2枚だが1枚が短い場合

4. 葉が2枚均等に展開している場合

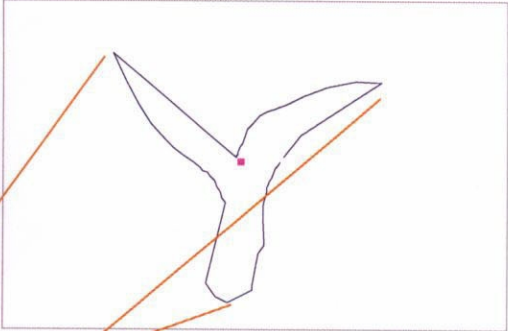
図 11(a)の様に葉が2枚均等に展開している苗の場合、境界線は、図心を中心として3方向に均等に展開する。そのためサンプル値には、同じ間隔で3つの大きなピークが生じる。その結果、3次のフーリエ係数が他の次数に比べて大きくなる。



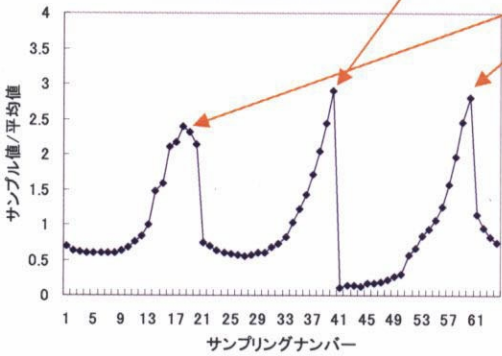
(a)写真



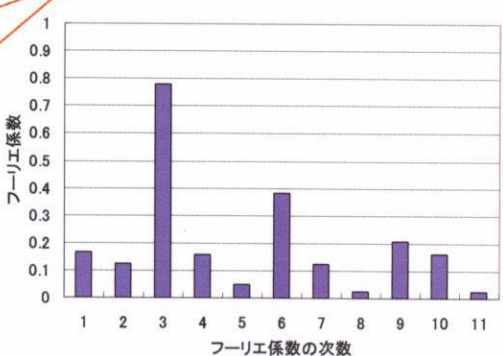
(b)境界線画像



(c)近似画像



(d)サンプル値/平均値

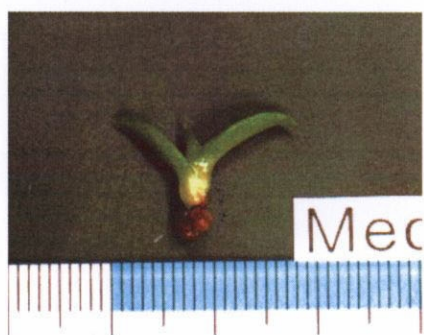


(e)フーリエ係数

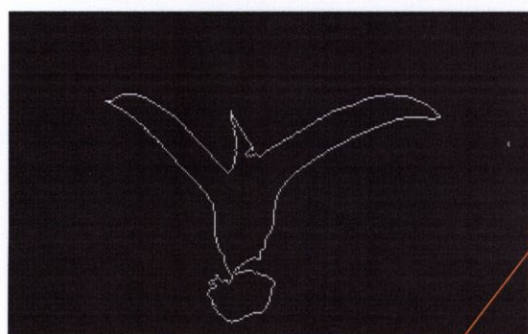
図 11 葉が2枚均等に展開している場合

## 5. 葉が3枚だが1枚が短い場合

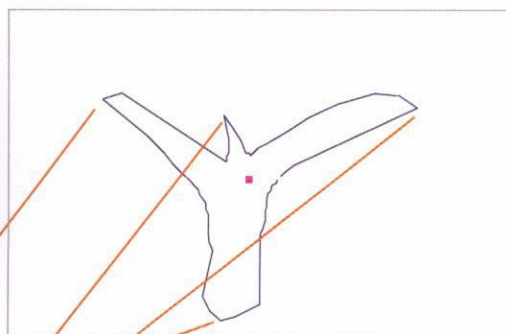
図 12(a)の様に葉が3枚確認できるが、1枚が完全に伸長していない場合、境界線の形状は、葉が2枚均等に展開しているものに類似している。3枚目の短い葉のところで、小さなピークが生じるものの、それ以外の3つのピークの影響が大きく、フーリエ係数は3次が最も大きくなる。



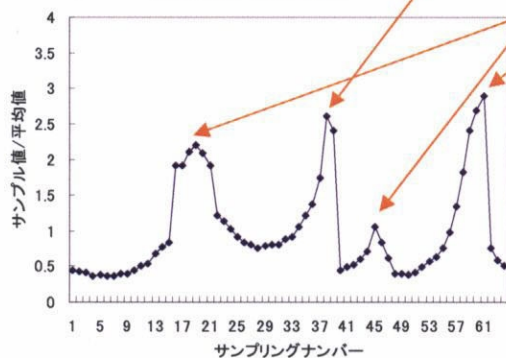
(a)写真



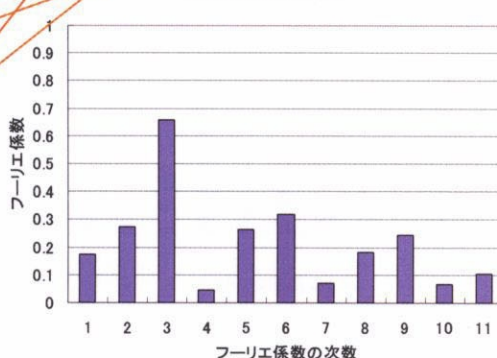
(b)境界線画像



(c)近似画像



(d)サンプル値/平均値

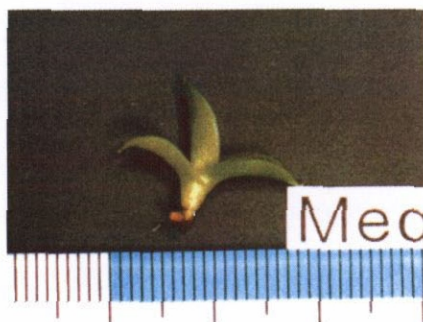


(e)フーリエ係数

図 12 葉が3枚だが1枚が短い場合

## 6. 葉が3枚均等に展開している場合

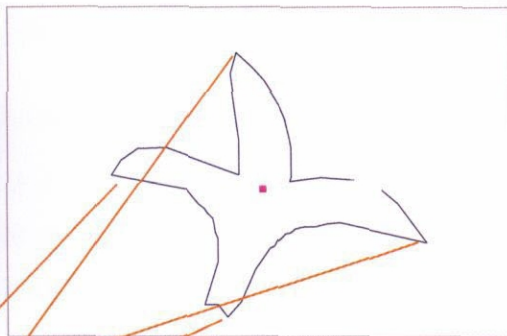
図 13(a)の様に葉が3枚均等に展開している場合、境界線の形状は、図心を中心として4方向に展開する. サンプル値においても、同じ間隔で大きなピークが4つ生じる. その結果、4次のフーリエ係数が最も大きくなる.



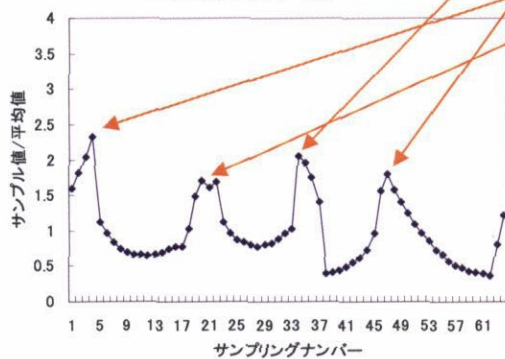
(a)写真



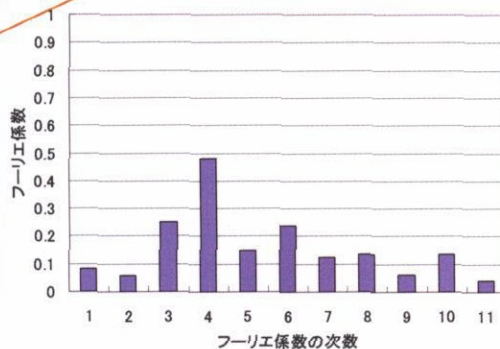
(b)境界線画像



(c)近似画像



(d)サンプル値/平均値



(e)フーリエ係数

図 13 葉が3枚均等に展開している場合

第3項 面積の分布

表 2に、各等級における投影面積の主要統計量を示す。また、図 14にヒストグラムを示す。良苗の各等級の面積の分布は比較的、等級と合致しているが、不良苗と、良苗(小)は、その分布が重なっている。この原因としては、以下の2点が挙げられる。

- 1. 不良苗の中で、未発達なものは、面積が広い傾向がある事。これは、未発達な苗が、短い萎縮した複数の葉を持っている場合に見られる。
- 2. 不良苗と判断されている中にも、良苗(小)と判断できるものが存在する事。苗の大きさは、作業者が目視で選別しているため、このような分布の重なりは十分に起こりうる。

表 2 各等級における投影面積の主要統計量

|       | 平均値(pixel) | 標準偏差(pixel) | 最小値(pixel) | 最大値(pixel) |
|-------|------------|-------------|------------|------------|
| 不良    | 2956.4     | 1010.9      | 1369       | 6258       |
| 良苗(小) | 4039.3     | 1010.7      | 2022       | 6959       |
| 良苗(中) | 5852.9     | 1331.3      | 3359       | 9093       |
| 良苗(大) | 9168.2     | 1855.5      | 6352       | 13683      |

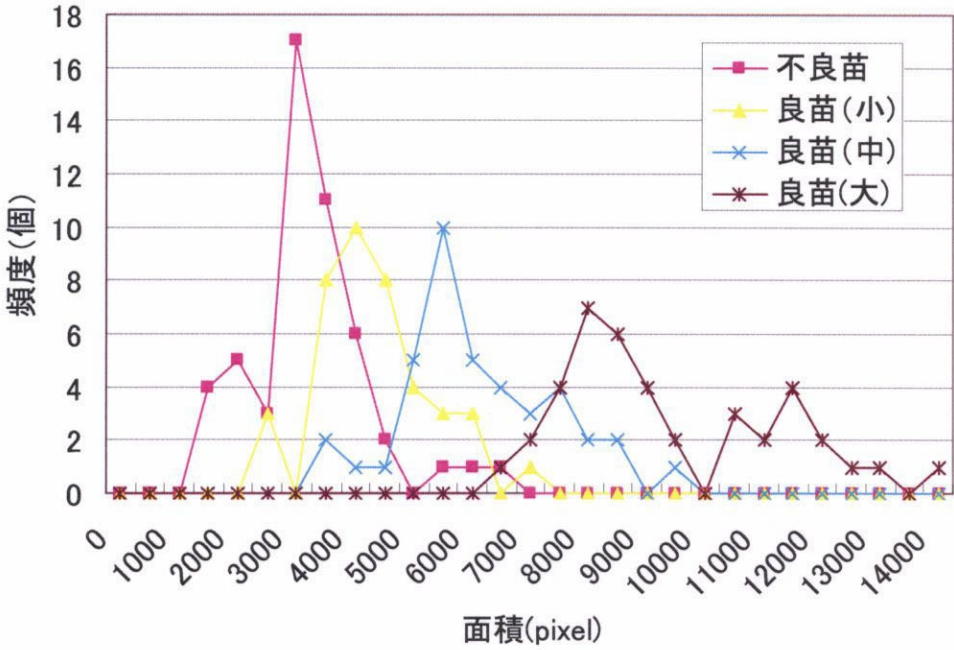


図 14 投影面積のヒストグラム

第4項 最大サンプル値の分布

苗の大きさを示す指標として、最大サンプル値(重心から境界線に至る直線の距離の最大値)を計測した. 表 3にそれぞれの等級における最大サンプル値の主要統計量を示す. また, 図 15に最大サンプル値のヒストグラムを示す. この指標でも, 各等級には重なりが見られる事がわかる.

表 3 各等級における最大サンプル値の主要統計量

|       | 平均値(pixel) | 標準偏差(pixel) | 最小値(pixel) | 最大値(pixel) |
|-------|------------|-------------|------------|------------|
| 不良    | 59.8       | 12.6        | 34.1       | 90.8       |
| 良苗(小) | 84.5       | 13.9        | 47.3       | 113.2      |
| 良苗(中) | 111.6      | 12.8        | 88.9       | 153.8      |
| 良苗(大) | 143.7      | 19.7        | 115.6      | 201.1      |

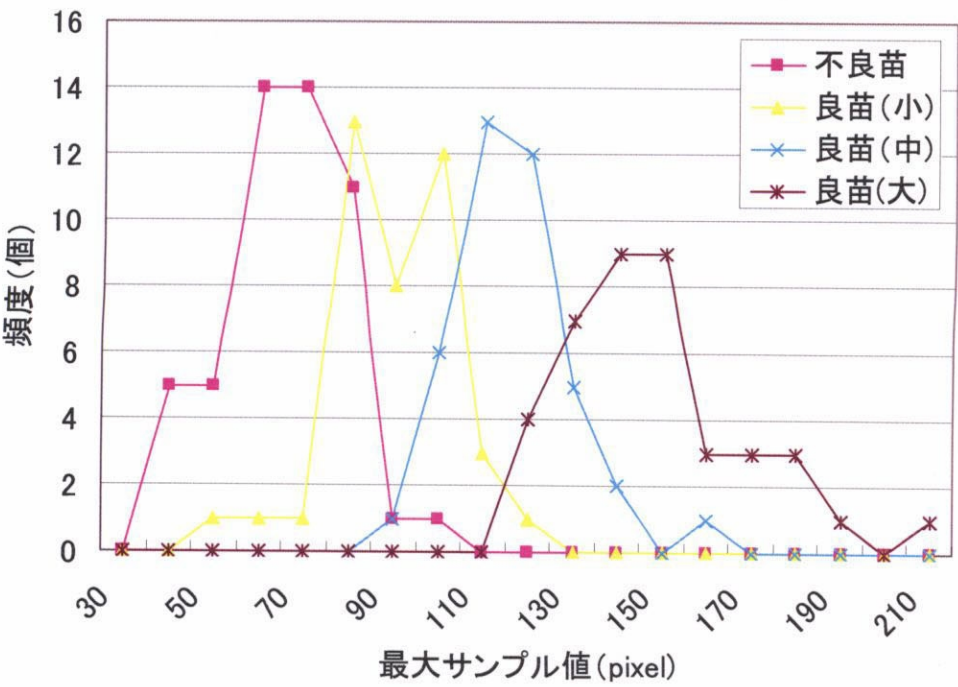


図 15 最大サンプル値のヒストグラム

第5項 平均サンプル値の分布

表 4に各等級における平均サンプル値(図心から境界線までの距離の平均)の主要統計量を示す. また, 図 16にヒストグラムを示す. 苗の大きさは, 主に葉の長さで表現されるが, 葉の部分は細く, たとえ葉が長くてもサンプリングされる個数が少ないため, このような結果となったと考えられる. 各等級の分布が大きく重なっていることから, この指標は, 大きさを示す指標としては使えない事がわかる.

表 4 各等級における平均サンプル値の主要統計量

|       | 平均値(pixel) | 標準偏差(pixel) | 最小値(pixel) | 最大値(pixel) |
|-------|------------|-------------|------------|------------|
| 不良    | 28.1       | 5.6         | 18.5       | 45.6       |
| 良苗(小) | 32.6       | 5.4         | 21.4       | 48.6       |
| 良苗(中) | 38.5       | 5.4         | 27.2       | 48.4       |
| 良苗(大) | 47.5       | 6.9         | 31.2       | 59.5       |

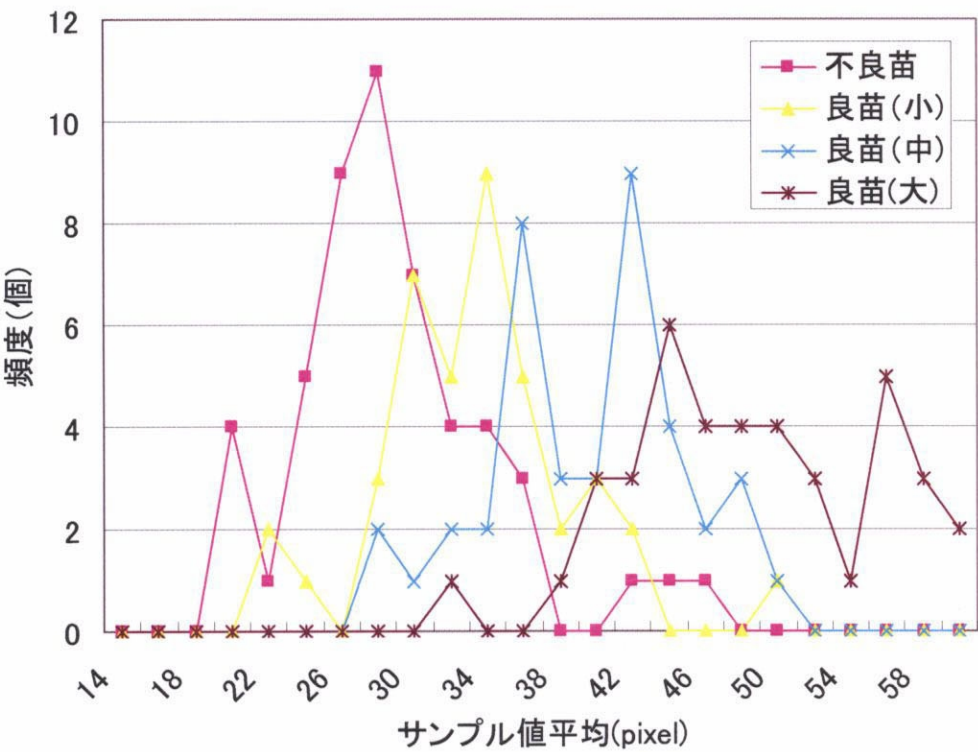


図 16 平均サンプル値のヒストグラム

第6項 面積×最大サンプル値の分布

図 17に面積と最大サンプル値を掛け合わせたもののヒストグラムを示す。各等級の分布の重なりが、面積だけや、最大サンプル値だけの時に比べて、減っている事がわかる。また、図 18に各等級における面積と最大サンプル値の散布図を示す。この散布図より、各等級が、適切な境界線によって、分離できる事が予想される。

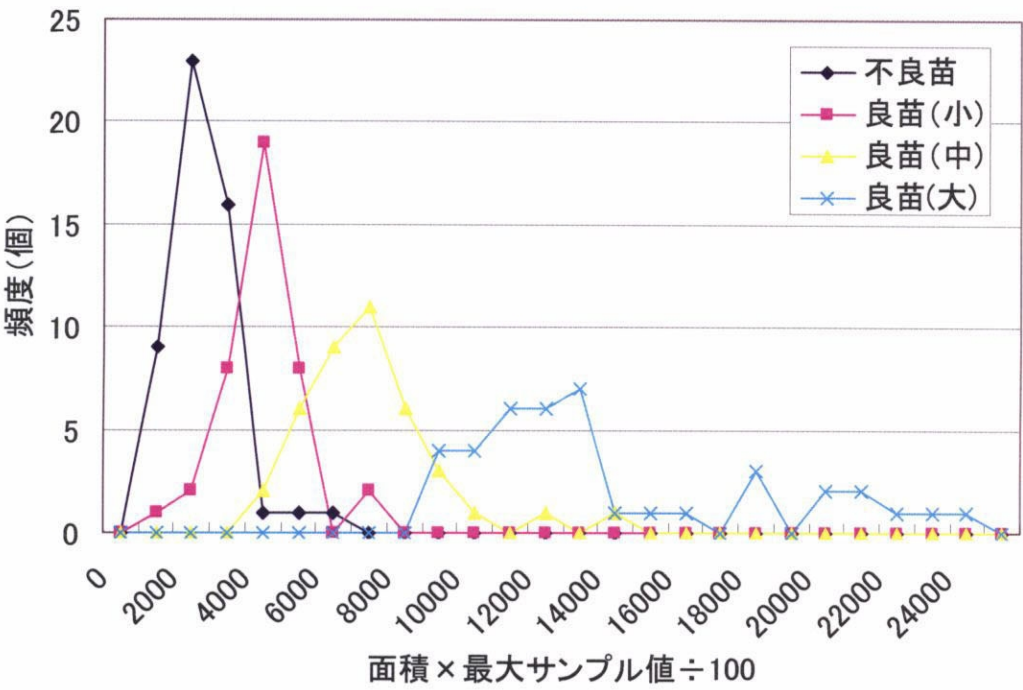


図 17 面積×最大サンプル値のヒストグラム

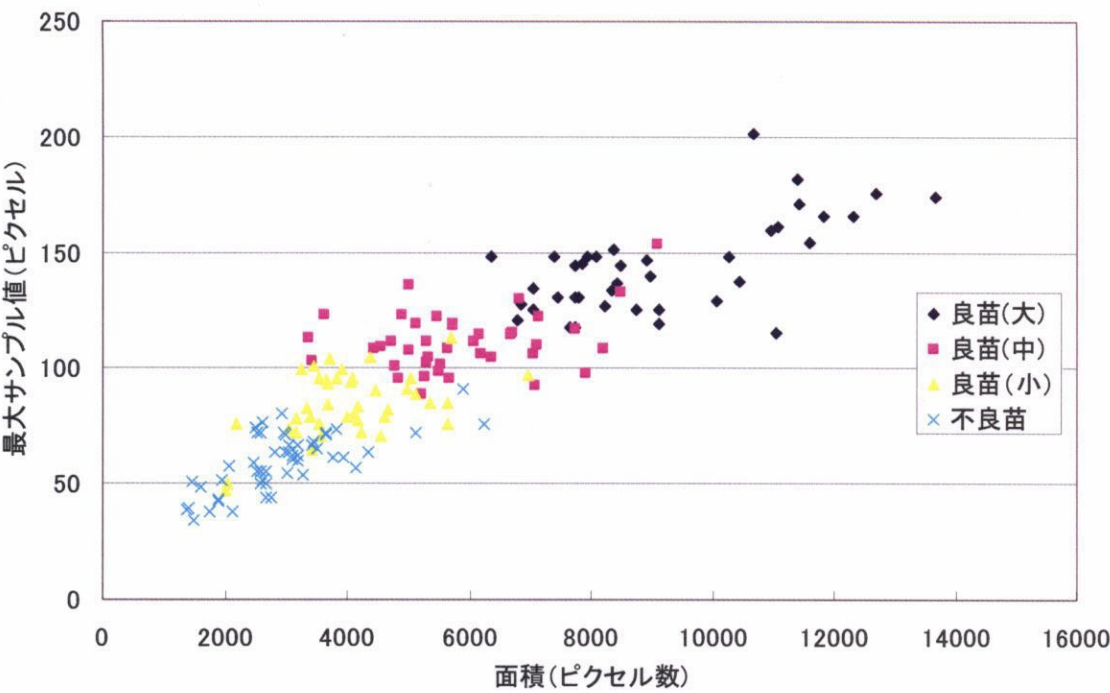


図 18 各等級の面積と最大サンプル値の分布

## 第6節 等級選別

### 第1項 等級選別の戦略

等級選別を以下の2つの戦略に基づいて行い、その結果を比較する。

1. 形状によって、明らかに不良苗と判断できるもの(未発達なもの)を除く。これによって、形状から言えば今後、正常に成長する可能性があるが、未熟である苗の混入を認める。その後、大きさによって、不良苗、良苗(小)、良苗(中)そして良苗(大)に選別する。
2. 大きさのパラメータのみによって、等級選別を行う。

本来であれば、無作為に取り出したサンプルから、特徴量を抽出し、選別アルゴリズムを作成し、それを未知のサンプルに対して適用して、その選別アルゴリズムの妥当性を評価すべきである。しかし、今回の実験の大きな目的は、選別に有効な特徴量の選定と、選別アルゴリズムの妥当性の評価である。その為、全てのサンプルを用いて、平均と分散共分散行列を計算し、判別した後の誤判定率を算出した。

選別アルゴリズムとしては、前述のマハラノビスの距離を用いた最短距離法を用いる。

第2項 形状による未発達な苗の除去

1. 単一のフーリエ係数をパラメータとする未発達な苗と正常な苗のヒストグラム

前節第 2 項で述べたように、未発達な不良苗のフーリエ係数は全体的に低く、その平均値が小さくなる。良苗と不良苗含めて正常に生長している苗のフーリエ係数は、2 次や 3 次、4 次の特徴的なピークを持っている。以下に、未発達な苗と、正常な苗のそれぞれのフーリエ係数のヒストグラムを示す。最もはっきりとヒストグラムが分かれているものはフーリエ係数の平均値だった(図 19)。1 次や 2 次のフーリエ係数の分布はかなり重なりが大きかった(図 20, 図 21)。一方で、3 次、4 次のフーリエ係数はその分布に差異が認められた(図 22, 図 23)。しかしながら、その分布には重なりが見られる。より高い精度で選別を行うためには、複数のパラメータを組み合わせる必要があると考えられる。

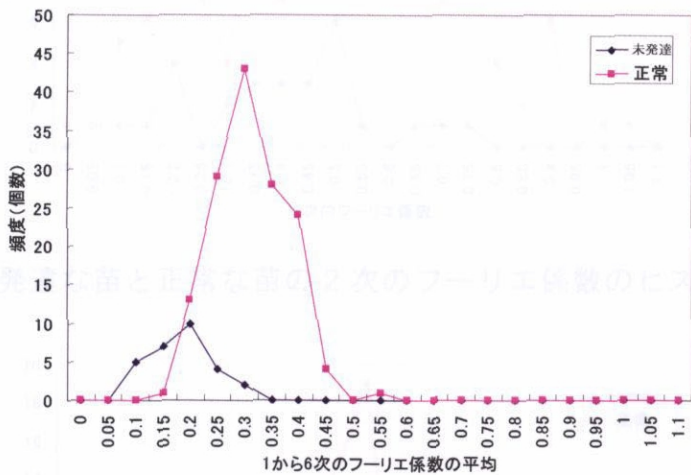


図 19 未発達な苗と正常な苗の 1 次から 6 次までのフーリエ係数の平均

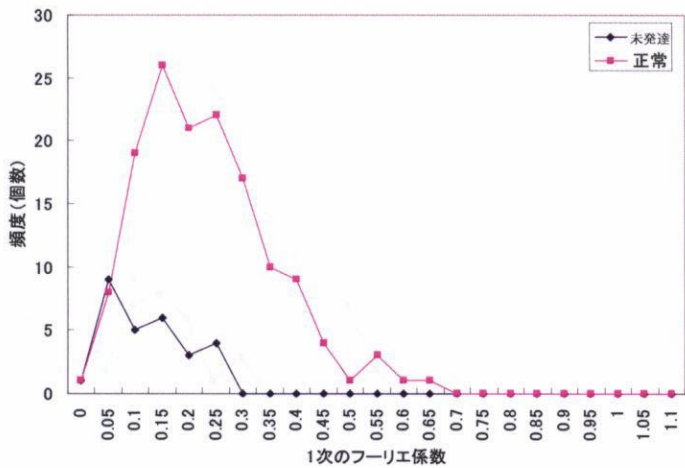


図 20 未発達な苗と正常な苗の 1 次のフーリエ係数のヒストグラム

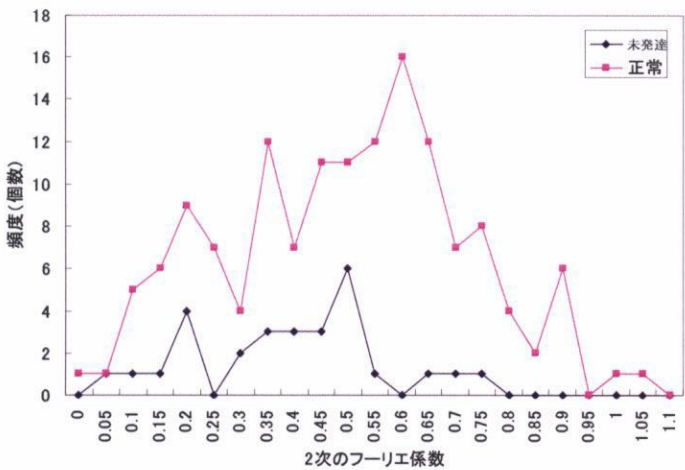


図 21 未発達な苗と正常な苗の 2 次のフーリエ係数のヒストグラム

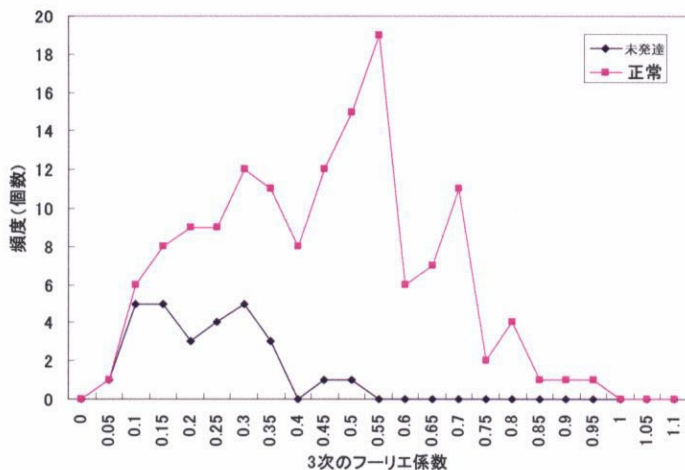


図 22 未発達な苗と正常な苗の 3 次のフーリエ係数のヒストグラム

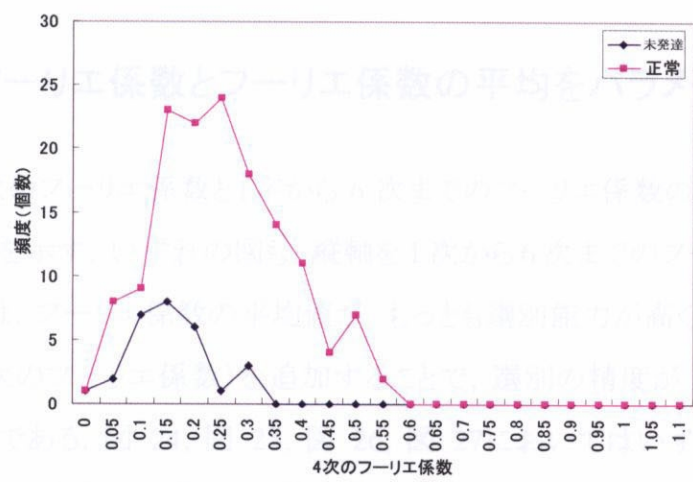


図 23 未発達な苗と正常な苗の 4 次フーリエ係数のヒストグラム

2. 各次のフーリエ係数とフーリエ係数の平均をパラメータとする散布図

以下に、各次のフーリエ係数と1次から6次までのフーリエ係数の平均値をパラメータとする散布図を示す. いずれの図も、縦軸を1次から6次までのフーリエ係数の平均値とした. これは、フーリエ係数の平均値が、もっとも選別能力が高く、これにその他のパラメータ(各次のフーリエ係数)を追加することで、選別の精度が上がるのではないかと考えたためである. 図 24, 図 25, 図 26, 図 27においてはいずれも、単独のパラメータに比較して、未発達な苗と、正常な苗の分布の重なりが少なくなっている.

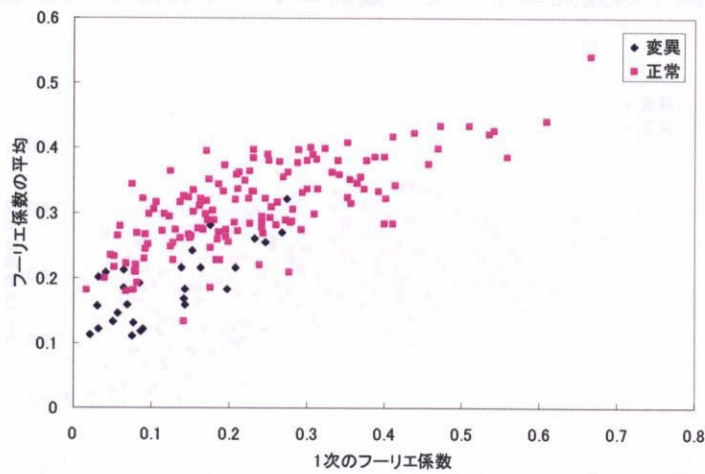


図 24 1 次のフーリエ係数－フーリエ係数の平均

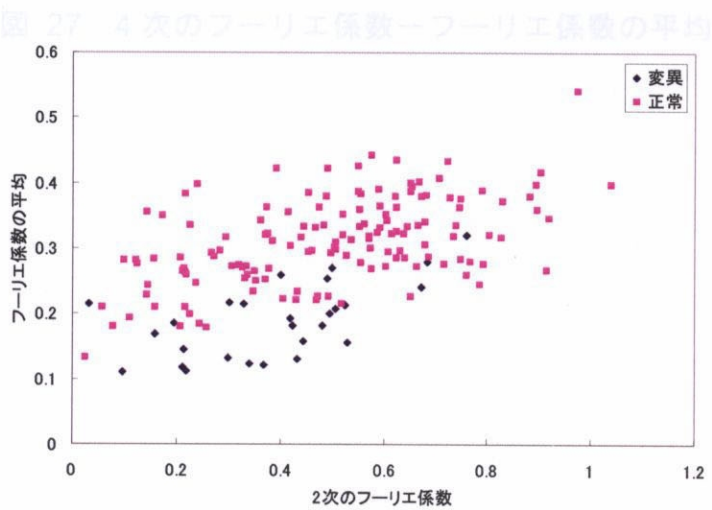


図 25 2 次のフーリエ係数－フーリエ係数の平均

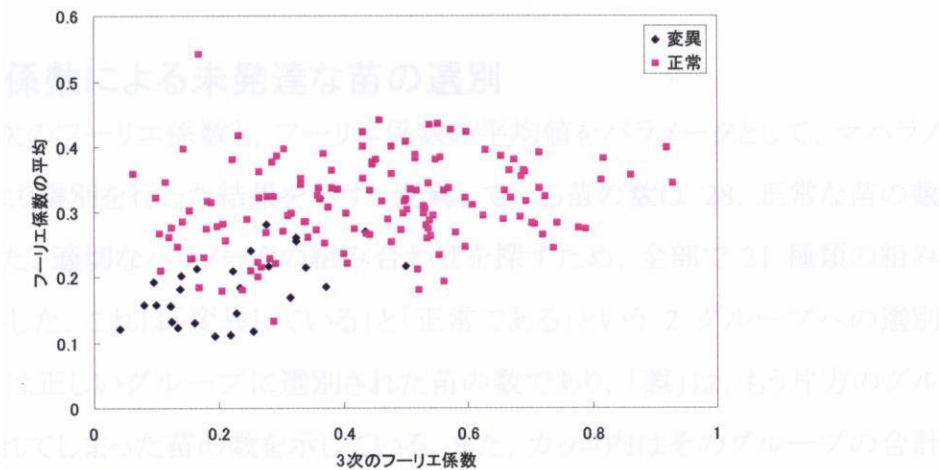


図 26 3 次のフーリエ係数－フーリエ係数の平均

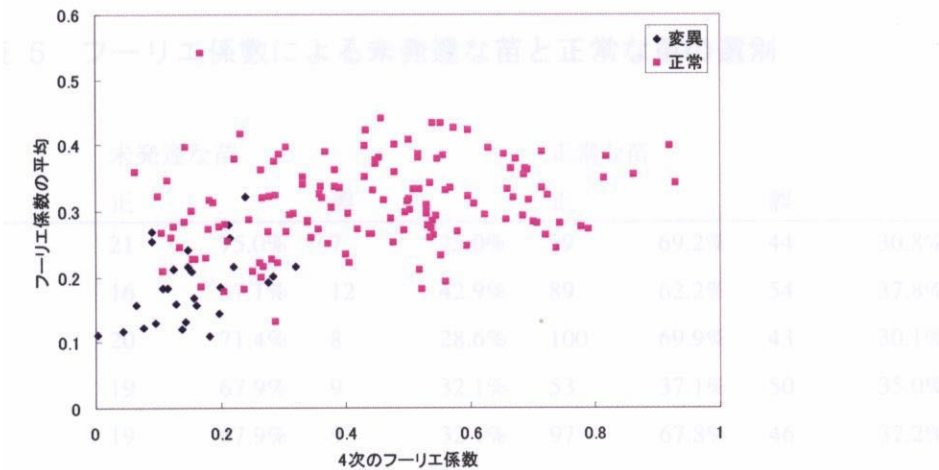


図 27 4 次のフーリエ係数－フーリエ係数の平均

### 3. フーリエ係数による未発達な苗の選別

以下に、各次のフーリエ係数と、フーリエ係数の平均値をパラメータとして、マハラノビスの距離により選別を行った結果を示す。変異している苗の数は 28、正常な苗の数は 143 であった。適切なパラメータの組み合わせを探すため、全部で 21 種類の組み合わせをテストした。これは「変異している」と「正常である」という 2 グループへの選別である。「正」とは正しいグループに選別された苗の数であり、「誤」は、もう片方のグループに選別されてしまった苗の数を示している。また、カッコ内はそのグループの合計に対する割合をパーセント表示したものである。F1, F2, ... は、各次のフーリエ係数、平均は、1次から 6 次までのフーリエ係数の平均を表している。

表 5 フーリエ係数による未発達な苗と正常な苗の選別

| パラメータ          | 未発達な苗 |       |    |       | 正常な苗 |       |    |       |
|----------------|-------|-------|----|-------|------|-------|----|-------|
|                | 正     |       | 誤  |       | 正    |       | 誤  |       |
| F1             | 21    | 75.0% | 7  | 25.0% | 99   | 69.2% | 44 | 30.8% |
| F2             | 16    | 57.1% | 12 | 42.9% | 89   | 62.2% | 54 | 37.8% |
| F3             | 20    | 71.4% | 8  | 28.6% | 100  | 69.9% | 43 | 30.1% |
| F4             | 19    | 67.9% | 9  | 32.1% | 53   | 37.1% | 50 | 35.0% |
| F5             | 19    | 67.9% | 9  | 32.1% | 97   | 67.8% | 46 | 32.2% |
| F6             | 18    | 64.3% | 10 | 35.7% | 94   | 65.7% | 49 | 34.3% |
| 平均             | 23    | 82.1% | 5  | 17.9% | 122  | 85.3% | 21 | 14.7% |
| F1+平均          | 20    | 71.4% | 8  | 28.6% | 129  | 90.2% | 14 | 9.8%  |
| F2+平均          | 22    | 78.6% | 6  | 21.4% | 127  | 88.8% | 16 | 11.2% |
| F3+平均          | 22    | 78.6% | 6  | 21.4% | 129  | 90.2% | 14 | 9.8%  |
| F4+平均          | 21    | 75.0% | 7  | 25.0% | 129  | 90.2% | 14 | 9.8%  |
| F5+平均          | 21    | 75.0% | 7  | 25.0% | 127  | 88.8% | 16 | 11.2% |
| F6+平均          | 22    | 78.6% | 6  | 21.4% | 124  | 86.7% | 19 | 13.3% |
| F2+F3+F4       | 21    | 75.0% | 7  | 25.0% | 133  | 93.0% | 10 | 7.0%  |
| F2+F3+平均       | 21    | 75.0% | 7  | 25.0% | 134  | 93.7% | 9  | 6.3%  |
| F3+F4+平均       | 20    | 71.4% | 8  | 28.6% | 135  | 94.4% | 8  | 5.6%  |
| F1+F2+F3+平均    | 18    | 64.3% | 10 | 35.7% | 137  | 95.8% | 6  | 4.2%  |
| F2+F3+F4+平均    | 19    | 67.9% | 9  | 32.1% | 139  | 97.2% | 4  | 2.8%  |
| F1+F2+F3+F4+平均 | 17    | 60.7% | 11 | 39.3% | 139  | 97.2% | 4  | 2.8%  |
| F2+F3+F4+F5+平均 | 17    | 60.7% | 11 | 39.3% | 138  | 96.5% | 5  | 3.5%  |

この表より、未発達な苗の誤判定率は、「平均」をパラメータとして用いたときに最も低く、また、正常な苗の誤判定率は「F2+F3+F4+平均」をパラメータとして用いたときに最も低くなることがわかる。「F2+F3+F4+平均」をパラメータとして用いたときに正常な苗の中で誤判定された4個の苗の内、3個はもともと、不良苗に分類されていたもので、1個は良苗(小)に分類されていたものであった。

4. 面積と最大長さをを用いた等級選別

前項で「F2+F3+F4+平均」をパラメータとして用いたときに、「正常」と判断された苗を、面積とサンプル値の最大値を用いて、等級選別した結果を以下に示す。選別方法としては、「正常」と判断された苗を、面積、サンプル値の最大値または、面積×サンプル値の最大値の大きな順に並べ替え、上から 40 個ずつ、「良苗(大)」、「良苗(中)」、「良苗(小)」、「不良苗」とした。

面積のみを用いた選別(表 6)と、最大サンプル値による選別(表 7)とでは、正しく判定された苗の割合に大きな違いはない。しかしながら、面積×最大サンプル値を用いた選別では、より高い正判定率を得る事ができた。実際に苗を生産、販売する場合問題となるのは、明らかに不良な苗を、良苗として販売してしまうことである。これらの選別では、苗の大きな順番に各等級に分類している。不良苗でありながら、良苗(小)に分類された苗は、いずれも、面積×最大サンプル値が、良苗(小)の中でも低いものばかりであった。適切な閾値さえ設定すれば、不良苗の混入を防ぐことは可能だと考えられる。

表 6 面積による選別

| 面積による選別 |       |       |       |     |    |       |
|---------|-------|-------|-------|-----|----|-------|
|         | 良苗(大) | 良苗(中) | 良苗(小) | 不良苗 | 合計 | 正判定率  |
| 良苗(大)   | 35    | 5     | 0     | 0   | 40 | 87.5% |
| 良苗(中)   | 5     | 28    | 7     | 0   | 40 | 70.0% |
| 良苗(小)   | 0     | 7     | 26    | 7   | 40 | 65.0% |
| 不良苗     | 0     | 0     | 6     | 22  | 28 | 78.6% |

表 7 最大サンプル値による選別

| 最大サンプル値による選別 |       |       |       |     |    |       |
|--------------|-------|-------|-------|-----|----|-------|
|              | 良苗(大) | 良苗(中) | 良苗(小) | 不良苗 | 合計 | 正判定率  |
| 良苗(大)        | 35    | 5     | 0     | 0   | 40 | 87.5% |
| 良苗(中)        | 5     | 29    | 6     | 0   | 40 | 72.5% |
| 良苗(小)        | 0     | 6     | 26    | 8   | 40 | 65.0% |
| 不良苗          | 0     | 0     | 7     | 21  | 28 | 75.0% |

表 8 面積×最大サンプル値による選別

| 面積×最大サンプル値による選別 |       |       |       |     |    |       |
|-----------------|-------|-------|-------|-----|----|-------|
|                 | 良苗(大) | 良苗(中) | 良苗(小) | 不良苗 | 合計 | 正判定率  |
| 良苗(大)           | 37    | 3     | 0     | 0   | 40 | 92.5% |
| 良苗(中)           | 3     | 34    | 3     | 0   | 40 | 85.0% |
| 良苗(小)           | 0     | 3     | 30    | 7   | 40 | 75.0% |
| 不良苗             | 0     | 0     | 5     | 23  | 28 | 82.1% |

第3項      大きさのパラメータによる等級選別

前項では、あらかじめ、未発達な苗と考えられる苗を除いた後、苗が大きい順に同数ずつ選別している。苗のサンプルを見ると、不良苗は、未発達なものと正常なものいずれも、良苗に比べて小さい。大きさを示すパラメータは多々あるが、ここでは、前項でも用いた、面積×最大サンプル値を用いて、選別を行う。具体的には、全てのサンプルを面積×最大サンプル値によって並び替え、上から 40 個ずつ、良苗(大)、良苗(中)、良苗(小)として、残りを不良苗とした。その結果を表 9に示す。

表 9    大きさのパラメータのみによる選別

| 面積×最大サンプル値による選別 |       |       |       |     |    |       |
|-----------------|-------|-------|-------|-----|----|-------|
|                 | 良苗(大) | 良苗(中) | 良苗(小) | 不良苗 | 合計 | 正判定率  |
| 良苗(大)           | 37    | 3     | 0     | 0   | 40 | 92.5% |
| 良苗(中)           | 3     | 33    | 4     | 0   | 40 | 82.5% |
| 良苗(小)           | 0     | 5     | 29    | 6   | 40 | 72.5% |
| 不良苗             | 0     | 0     | 7     | 44  | 51 | 86.3% |

良苗(大)は、同じ正判定率、良苗(中)および(小)に関しては、若干の低下が見られた。不良苗では、前項では形状によって取り除かれた苗が合計に含まれているため、見た目の正判定率が向上する結果となった。

## 第7節 まとめ

ランの実生苗をモノクロカメラで撮影し、その形状を用いて、等級選別を行った。以下に本章の結果を述べる。

1. ランの実生苗の形状特徴量を得る方法として、境界線の座標を、図形の図心からの距離と角度で近似した。その結果、図形の画面上での位置に関係のない特徴量を抽出した。この方法は、今回対象としたランの幼苗のように放射状かつ等角度間隔に展開した形状を持つ物体の特徴の表現に優れていることが明らかになった。しかしながら、苗が大きくなるにつれ葉が長くなり、凹な部分が増える。今回採用したサンプリングの方法では、凹な部分はサンプリングできず、凹な部分の前後ではサンプル値が不連続になり、正しい形状の近似が困難であった。より長い葉を持った苗については、第5章で述べるような境界線の曲率を用いた方法が有効であると考えられる。
2. 図心からの距離を周期 $2\pi$ の周期関数と見なし、離散フーリエ級数展開することで、図形の回転および、大きさに左右されない形状特徴量を得た。その結果苗の形状によって、図心からの距離のパターンが変化し、また同時にフーリエ係数のパターンも変化する事がわかった。そして、葉の枚数と、フーリエ係数との関係を明らかにした。
3. 大きさのパラメータを用いなくても、2次、3次、4次のフーリエ係数と、1次から6次までのフーリエ係数の平均値をパラメータとして用いてマハラノビス距離により、未発達な苗と正常な苗の選別を行えることを示した。これは、苗の葉の枚数が同じならば、苗全体の大きさに関わらず、フーリエ係数のパターンが類似するであろうという予想に合致していた。
4. 面積と最大サンプル値を掛け合わせた値が、苗の大きさを表す指標としては最も適していることが判明した。