

# マハラノビス距離を用いたガラス検査用画像判別の実現

## Scratches Detection for Testing Glasses using Mahalanobis Distance

松崎 裕樹† 山崎 勝弘† 平岡 邦廣‡ 西田 洋隆‡ 三宅 淳司‡

Hiroki Matsuzaki Katsuhiko Yamazaki Hiraoka Kunihiro Nishida Hiroataka Miyake Atsushi

### 1. はじめに

近年、液晶・PDPに代表されるディスプレイ関連の技術革新は著しく、より大型化、より高精細化が急速に進んできている。それに伴い、材料としての基板ガラスや工学フィルムの品質も、より一層、高品位なものが要求されるようになってきている。そのため、ガラス等の製造ラインの検査工程において検出されるべき欠陥のサイズは、より小さなものを検出する必要性が出てきており、また後工程で問題となる欠陥だけ選別するという欠点識別機能も、より高度な機能が要求されるようになってきている[1]。本研究では、ガラスの製造工程において生じる気泡(傷)と、撮影時に生じる気流を正確に判別するために、マハラノビス距離を用いた多変量解析に基づく画像判別を行う。実際のガラス検査で用いるためには、95%以上の判定率を達成することが必要であり、これが本研究の目標である。マハラノビス距離とは、判別分析に用いられる手法であり、判別対象の母集団の分散を考慮して距離の算出を行うのが特徴である。本研究では、まずラベリング対象候補の選定と特徴量の対応付けを行い、次に判定率向上のために、スライスレベルの可変化と気泡・気流の細分化の実験を行って判定率を評価する。

### 2. マハラノビス距離による画像判別

#### 2.1 画像判別の手順

本研究で行った画像判別の手順を図1に示す。対象画像に対し、ラベリングと呼ばれる画像内のオブジェクトを識別するための処理を行う。その後、識別されたオブジェクトから特徴量を抽出し、マハラノビス距離を用いた統計処理を行い、気泡と気流の判定を行う。

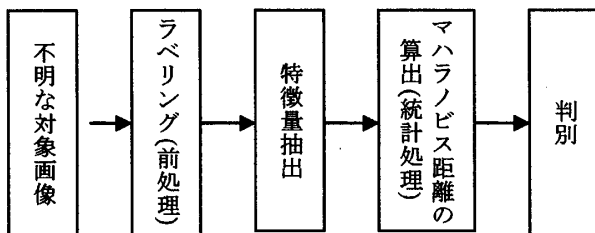


図1: 画像判別手順

対象とするサンプルの気泡画像を図2(a)、気流画像を図2(b)に示す。図2(a)の中央に十字に見えるのが気泡である。また、図2(b)の画面全体にまだらに見えるものが気流である。気泡はガラスの生成工程で発生するものであり、欠陥である。気流はカメラによる撮影時に気流によって生じる屈折で写りこんでしまうものである。気泡と気流を判別し、欠陥である気泡を抽出することが本研究の目的となる。

図2のサンプル画像は気泡・気流が比較的わかりやすい場合を示している。一般的に、気泡画像では気泡が非常に小さい場合や横長になっている場合に、気流画像では気流が中央に収束している場合や縦長な場合に判別が難しくなる。

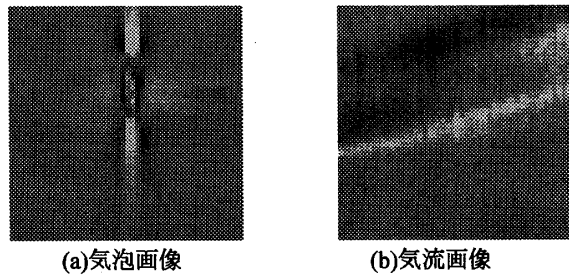


図2: サンプル画像

#### 2.2 ラベリング

ラベリングとは、画素中の同一成分に一つの番号(ラベル)を割付け、別の連結成分には別の番号を割り付ける処理である。ラベリングを行うことで各オブジェクトの識別が可能となり、次節で説明する特徴量抽出が可能となる[2]。今回用いたラベリングでは連結成分を8近傍で行っている。例としてラベリング前の画像(a)とラベリング後の画像(b)を図3に示す。

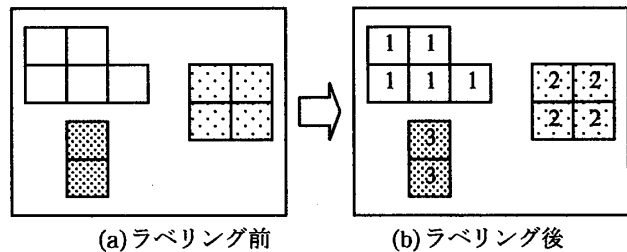


図3: ラベリング

#### 2.3 特徴量抽出

特徴量抽出とは、画像から色や形などの特徴を数値化した情報として抽出する処理のことである。ラベリングにより識別できるようになったオブジェクトに対し、種々の特徴量を抽出する。本研究では特徴量として、面積、体積(輝度の合計)、周囲長、縦の長さ、横の長さ、及び面積比率の6つを用いている。周囲長は各ピクセル間の中心点を結んだ長さで計算している。また、面積比率は縦の長さとの比率である。オブジェクトから特徴量を抽出した例を図4に示す。

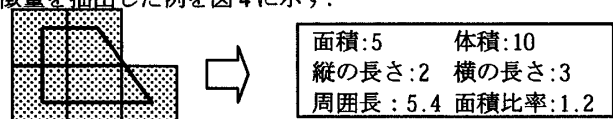


図4: 特徴量抽出

† 立命館大学大学院理工学研究科, Graduate school of Science and Engineering, Ritsumeikan University  
‡ 株式会社ナノスコープ, Nanoscope Co., Ltd

2. 4 マハラノビス距離

マハラノビス距離とは、判別分析に用いられる手法であり、判別対象の母集団の分散を考慮して距離の算出を行うのが特徴である。分散共分散行列とマハラノビスの距離の公式を式(1)、式(2)に示す。

分散共分散行列：

$$S_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x}) \times (y_i - \bar{y})}{n-1} \dots\dots\dots(1)$$

$\bar{x}, \bar{y}$  は母集団の平均

マハラノビス距離：

$$D(x, y) = \sum_{j=1}^T \sum_{k=1}^T S^{JK} (x_j - y_j)(x_k - y_k) \dots\dots\dots(2)$$

$S^{jk}$  : 分散共分散行列 ( $S_{jk}$ ) の逆行列 ( $j, k$ ) の要素

ユークリッド距離などの多変量解析では各変量の重み付けが解析精度に大きな影響を与えるため、重み付けの調整が非常に難しくなるが、マハラノビス距離では母集団の分散共分散の計算と、対象画像との行列計算によって各変量が正規化されることにより、重み付けを意識しなくて済む。また、母集団から共分散とその逆行列を計算してしまえば、マハラノビス距離の計算は対象データとの行列計算を行うだけで済むため、非常に高速に行うことができる。

3. ラベリング対象候補の選定と特徴量の対応

3. 1 実験手順と実験条件

実験手順を図5に示す。実験の手順として、大きく分けて「実験のパラメータ調整」と「判別精度の向上(チューニング)」の2つの段階を踏んでいる。実験に先立ち、まずラベリング対象候補の選定と、画像に対する特徴量の対応付けを行う。次に、さらなる判定率の向上のために、スライスレベルの可変性と気泡・気流の細分化の実験により、プログラムをチューニングする。

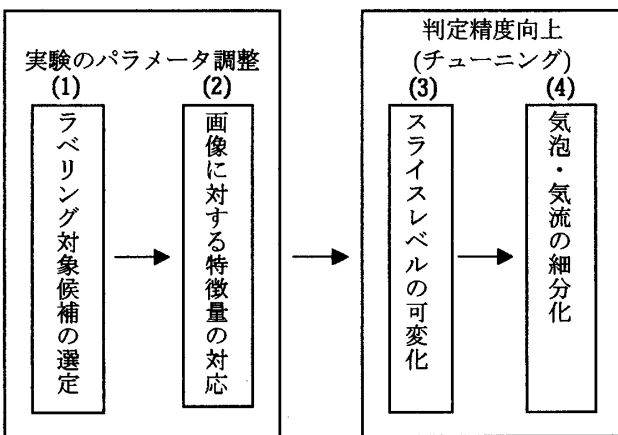


図5：実験手順

実験方法を図6に示す。実験はナノスコープより提供された気泡と気流の画像各283枚(計566枚)を用いて行っている。画像は50×50ピクセル、輝度0~255の濃淡画像である。まず、あらかじめ視認によって気泡のグループと

気流のグループを分別する。次に、設定したパラメータで分散共分散を抽出する。分散共分散を抽出したグループから判別対象とする画像を取り出し、気泡と気流の共分散を用いて各々マハラノビス距離を計算する。算出された距離を比べ、小さい方の距離を判別結果とし、視認による結果と合致しているかどうかで判定率を計算している。

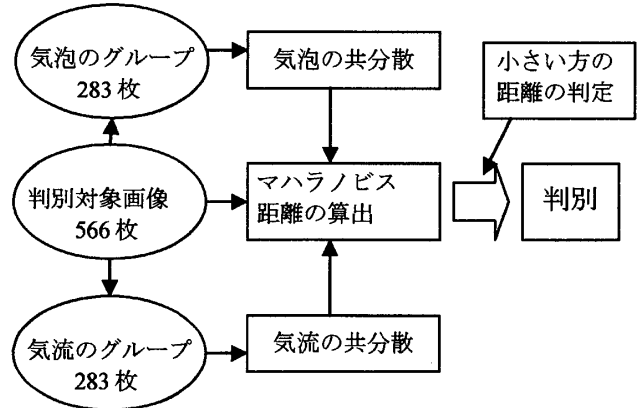


図6：実験方法

3. 2 ラベリングの対象候補の選定

画像に対して、気泡・気流を適切にラベリングできなければ抽出される特徴量も役に立たないものとなる。そこで、パラメータ調整として、ラベリング対象のスライス値の候補を選出した。スライスの閾値について、図7を用いて説明する。ラベル対象となる気泡・気流の輝度値は黒もしくは白に近い色となる。そこで、まず中央値と背景領域幅を変化させて、画像に対してラベリングを行う。ラベリングが正しく行われたかを目視によって判定し、ラベリングのスライス値の候補を決定した。中央値は100~150、背景領域幅は40~90までを10刻みで実験を行った。ラベリングの正当性レベルとして、精度の高いものから順に目視により○、△、×で評価している。実験結果を表1に示す。

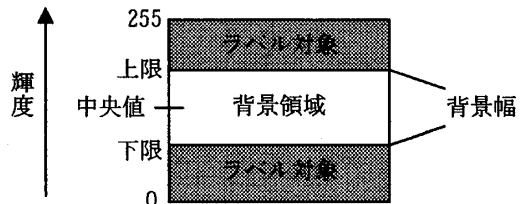


図7：スライスの閾値

表1：スライスレベル(横軸：中央値, 縦軸：背景幅)

	100	110	120	130	140	150
40	×	×	△	△	△	×
50	×	△	○	○	△	×
60	×	△	○	○	△	×
70	×	△	○	△	×	×
80	×	○	○	△	×	×
90	△	○	×	×	×	×

中央値 120 前後、背景幅 50~80 程度できれいにラベリングできることがわかった。表 1 より、○がついている項目を採用し、スライスの上限-下限の候補を 95-145, 90-150, 100-160, 85-155, 70-150, 80-160, 65-155 の 8 通りに設定し、次節の実験を行った。

### 3. 3 画像に対する特徴量の対応

図 8 に示すように一つの画像からラベリングにより、複数のオブジェクトが見つかり、各々特徴量が抽出される。各画像によって、ラベリングされるオブジェクトの数は一定ではないため、解析を行うためには、抽出された特徴量の数を統一する必要がある。そこで、複数のオブジェクトから抽出された特徴量から、その画像へ一つの特徴量に対応させることで、統一を図った。本実験ではラベリングされたデータから、面積が最大の場合と平均の場合で実験を行った。面積が最大の場合というのは、小さい面積の物は画像のゴミであることが多いため、一番大きくラベリングできたオブジェクトをその画像の特徴量としようとするものである。平均の場合は各ラベリングされたデータから特徴量の平均を算出し、特徴量とするものである。面積最大の場合の判定率を表 2 に、平均の場合の判定率を表 3 に示す。

表 2, 表 3 の結果から平均の場合の方が高い判定率であることがわかる。よって、画像と特徴量の対応は平均の場合を採用し、次章以降の実験を行った。

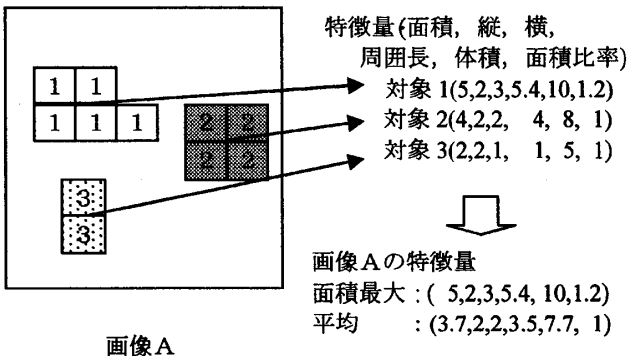


図 8: 特徴量の対応

表 2: 面積最大の場合の判定率

スライス値 (下限-上限)	95-145	105-155	70-150	80-160	85-155	90-150	100-160	65-155
気泡の判定率(%)	17	94	100	35	100	27	80	100
気流の判定率(%)	42	63	6	60	26	51	69	18

表 3: 平均の場合の判定率

スライス値 (下限-上限)	95-145	105-155	70-150	80-160	85-155	90-150	100-160	65-155
気泡の判定率(%)	100	100	99	99	51	99	92	81
気流の判定率(%)	78	68	85	75	98	88	95	89

## 4. スライスレベルの可変化と気泡・気流の細分化

### 4. 1 スライスレベルの可変化

同じラインセンサーカメラから取り込んだ画像でも全体的に薄い画像と、全体的に濃い画像がある。そのため、薄い画像はラベリング対象も薄くなり、濃い画像はラベリング対象も濃くなる。そこで、背景領域の中央値を固定値にするのではなく、まず画像に対してラスタスキャンを行い、輝度の平均値を算出し、その平均値を背景領域幅の中央値とする実験を行った。これは、ラベリングの精度を向上させることにより、特徴量の精度を向上させるためである。実験は背景領域幅を 30~100 まで 10 刻みで変化させて行った。表 4 に実験結果を示す。

表 4: スライスレベルの可変化

背景領域幅	30	40	50	60	70	80	90	100
気泡の判定率(%)	100	100	100	100	100	100	100	100
気流の判定率(%)	92	92	61	78	88	88	89	81

背景領域幅が 30, 40, 70, 80, 90, 100 で非常に高い判定率が得られることがわかった。次節では上記背景領域幅で実験を行った。

### 4. 2 気泡・気流の細分化

気泡の画像でも大きい気泡、小さい気泡など、様々な大きさがある。そこで、大きさ別に気泡で 3 通り、気流で 2 通りに分け、各々分散共分散行列を算出し、マハラノビス距離を計算して判別を行った。マハラノビス距離の計算対象を増やすことにより、抽出される特徴量の分散を小さくすることで、より精度の高い判別を行おうとするものである。気泡の場合のマハラノビス距離の計算を図 9 に示す。大きい気泡、中くらいの気泡、小さい気泡の共分散から、マハラノビス距離を計算し、最も小さい距離を気泡のマハラノビス距離として計算している。実験結果を表 5 に示す。

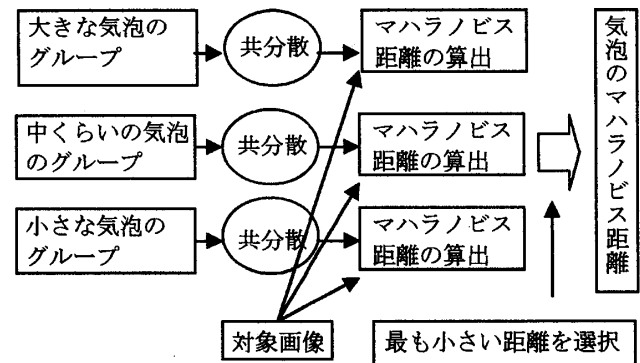


図 9: 気泡の細分化によるマハラノビス距離

表 5: 気泡・気流の細分化

背景領域幅	30	40	70	80	90
気泡の判定率(%)	100	100	100	100	100
気流の判定率(%)	90	79	90	91	90

4. 1の結果と比べて、背景領域 70,80,90 で判定率が向上した。30 と 40 では判定率は低下しているが、全体として判定率が安定してきている。

## 5. 考察

### (1)対象候補の選定と特徴量の対応

画像に対し、各オブジェクト情報の平均を対応させることで、判定率が高くなることがわかった。面積が最大のオブジェクトだけではなく、その他のラベリングされたオブジェクトも判定には有益であるということである。これは、ラベリング候補の選出により、オブジェクトの識別が正確に行われているためと考えられる。

### (2)スライスレベルの可変性

可変にすることで判定率が大きく向上している。ラベリングの正確さ、そしてそこから抽出される特徴量の精度が判別に直結していることがわかる。実験では、背景領域幅が 30,40,70,80,90 で高い判定率が得られている。背景領域幅 50,60 で判定率がそれほど向上しなかったのは、気流の画像の場合、背景領域と気泡との中間の部分において、可変化によってラベリングのゴミができやすくなったためと思われる。

背景領域の中央値を可変化し、背景領域幅を固定値にして実験を行ったが、背景領域幅も可変化することで、さらにラベリングの精度をあげることができるのではないかと考えられる。具体的には、画像成分の輝度の分散、または母集団と画像の輝度の平均との差分で可変化することが考えられる。

### (3)気泡・気流の細分化

背景領域幅が 70,80,90 の場合で判定率が向上した。ラベリングの閾値をシビアに取った場合の方が、細分化には向いていると思われる。判定率が減少した背景幅については、気泡・気流の細分化は特徴量の分散を減らし、精度を向上させようとしたものであるが、マハラノビス距離の場合は分散していること自体が情報として有意に働いたため、減少したものと思われる。本研究では細分化されたマハラノビス距離の最も短いものを用いたが、逆に最も長いものを比べる方法も考えられる。また、細分化を大きさ別に行うのではなく、体積が大きいもの、縦長のものなどのように他の方法で細分化を行うことも考えられる。

### (4)マハラノビス距離

マハラノビス距離が気泡・気流の判別に非常に有効であることがわかった。多変量解析では、各変量(本研究では特徴量)の重み付けが非常に重要であるが、マハラノビス距離では各変量の重み付けが共分散によって行われるため、重み付けを意識する必要がなくなり、順調に実験を進めることができた。本研究では 6 つの変量(特徴量)を用いているが、変量を増やすことでさらに精度をあげる方法も考えられる。変量を増やすと分散の次元も上がるため、逆に有意な変量の情報が小さくなる可能性があり、精度と変量とのバランスが重要であると思われる。

### (5)気泡・気流の判別

気泡・気流の判定率の平均は最高で 96%であった。判別においては、いかに正確に特徴量を抽出するか、そして、判別に有意となる特徴量を選出できるかが重要である。更なる判定率の向上として、用いた特徴量の一部の変更、判定対象の更なる細分化などが考えられる。

## 6. 未知画像に対する実験

既知画像に対する実験結果から、最も判定率が良かったスライスレベルの可変性を用いて、現在、未知画像に対する実験を行い、有効性の検証を進めている。実験方法を図 10 に示す。気泡、気流それぞれ 233 枚に対して、共分散を計算した上で、未知画像 100 枚に対して、マハラノビス距離を算出し、気泡か気流かを判別する。

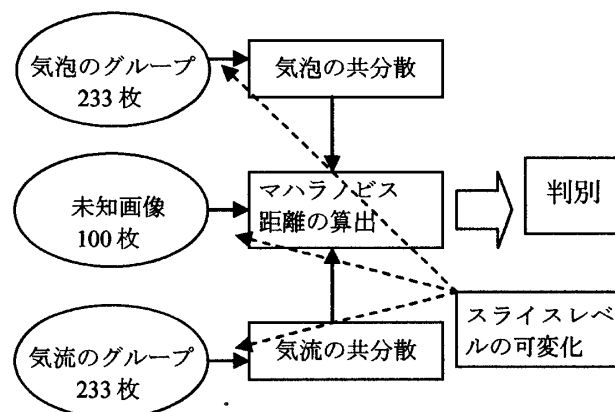


図 10：未知画像に対する実験方法

## 7. おわりに

本論文では、ガラス検査における気泡と気流を正確に判別するために、マハラノビス距離による画像判別の実験結果について述べた。画像判別においては、判定率向上のためにパラメータ調整とプログラムのチューニングを行い、最高で平均 96%の判定率を得ることができ、当初目標としていた 95%以上の判定率を達成した。マハラノビス距離を用いると、各変量の重み付けが共分散によって行われるため、重み付けを意識する必要がなくなり、順調に実験を進めることができた。実験結果については、ラベル対象候補の選定とスライスレベルの可変性が判定率の向上に大きく影響していた。したがって、いかにして高い判定率が得られる特徴量を抽出できるかが重要である。現在、未知画像に対して実験を行い、有効性の検証を進めている。今後、FPGA ボードを用いたハードウェア化による画像判別の高速化を行いたい。

## 参考文献

- [1] 株式会社ナノスコープ：平成 17 年度中小企業創業・経営革新等支援補助金交付申請書, 2005.
- [2] 長尾真：画像認識論, コロナ社, 1990.