

卒業論文

3D顔画像を用いた 個人識別に関する基礎的研究

東北大学 工学部 電気・情報系
阿曾研究室 4年

荀込卓明

目次

1	序論	1
1.1	研究の目的と背景	1
1.2	色符号化スリット光投影法の概要	1
1.3	本論文の構成	1
2	照合方法	4
2.1	データ	4
2.2	マッチング	4
2.3	パターン間の相違の尺度	4
2.4	評価基準	4
2.4.1	1位認識率 RR (%)	4
2.4.2	本人棄却率 ER (%)	5
2.4.3	分離能 $Separability$	5
3	特徴量抽出	6
3.1	3次元画像からの特徴の抽出	6
3.1.1	2点間の勾配の特徴抽出 <i>surface two-point gradient feature</i>	6
3.1.2	3点間の角度の特徴抽出 <i>surface three-point angular feature</i>	7
3.2	領域における特徴量の抽出	8
3.3	補正	9
4	実験方法	10
4.1	特徴量の抽出法の絞り込み	10
4.1.1	x 、 y 偏微分と領域の並べ方	10
4.1.2	全領域の y 成分の範囲	12
	上限もしくは下限を設定した時の下限もしくは上限の効果	12
	中心を設定した時の全領域の上下の幅の効果	16
4.1.3	領域の重なり	19
4.2	2点間の勾配を用いた特徴量について	21
4.3	3点間の角度を用いた特徴量について	21
4.4	領域における特徴量につて	22
5	結果	23
5.1	2点間の勾配の特徴抽出	23
5.2	3点間の角度の抽出	23
5.3	検討	23

6 特徴量抽出の改良	24
6.1 特徴の分布	24
6.2 量子化の改良	24
6.3 線型拡大	26
6.4 実験の結果	26
6.4.1 2点間の勾配を用いる特徴量	26
6.4.2 3点間の角度を用いた特徴量	27
7 結論	28
8 今後の課題	30
謝辞	31
参考文献	32

図目次

1	3次元形状システムの構成	2
2	3次元形状システムの座標系	2
3	観測画像の例	2
4	顔の3次元データの例	3
5	2点間の勾配	6
6	特徴の抽出	6
7	3点間の角度	7
8	特徴抽出	7
9	領域を半分ずつ重ねる方法	8
10	領域を隣接させる方法	8
11	補正の方法	9
12	垂直方向に領域を並べた場合	10
13	水平方向に領域を並べた場合	11
14	水平方向に領域を並べた状態	12
15	上限を動かした際の領域の状態 (上限が0付近)	14
16	上限を動かした際の領域の状態 (上限が1200付近)	14
17	下限を動かした際の領域の状態 (下限が-1200付近)	14
18	下限を動かした際の領域の状態 (下限が0付近)	14
19	領域を隣接させた場合	15
20	領域を半分ずつ重ねた場合	16
21	y成分の幅の小さい領域の状態 (隣接)	18
22	y成分の幅の大きい領域の状態 (隣接)	18
23	y成分の幅の小さい領域の状態 (重複)	18
24	y成分の幅の大きい領域の状態 (重複)	18
25	中心設定の領域の幅の効果	19
26	重なりのない領域の状態	20
27	3分の1重なった領域の状態	20
28	半分重なった領域の状態	20
29	4分の3重なった領域の状態	20
30	重なりの効果	21
31	3点間の角度の分布	24
32	2点間の勾配の分布	24
33	3点間の角度の分布 (使用した領域)	26
34	2点間の勾配の分布 (使用した領域)	26

1 序論

1.1 研究の目的と背景

人の顔を見てそれが誰であるか認識する能力は、人間にとってごくあたりまえのことである。これがコンピュータでできれば、セキュリティの分野でおおいに有用である。

かつては、正面あるいは真横から撮影した画像を対象として、2次元データから得られる特徴量のみを用い識別を行なおうとするものであった。これらは、顔の向きや顔-カメラ間の距離など被験者とカメラの位置関係が正確に固定されてないと識別が困難である。また、多数の顔を識別するためには、2次元画像より得られる情報量以上の情報が必要であると考えられる。

このため絶対的な3次元形状を計測する必要がでてくる。絶対的な3次元座標を対象とすれば、顔-カメラ間の距離の変動について識別の段階で考慮する必要がない。さらにデータから顔の向きを推定し回転することによってデータ計測時に被検者の向きが変動しても、その影響を受けにくい識別が可能となる。従来は人間の顔の3次元形状を計測すること自体、被検者に特別な姿勢を強いる等の問題があり困難であった。しかし色符号化スリット光投影法により、容易に3次元データを得ることができるようになった。本論文では、この色符号化スリット光投影法により得られた3次元データを用いた。

ただし3次元データでは、顔全体のデータをそのまま用いることは計算コストやデータの蓄積の面で効率的ではない。そこで、顔の特徴領域をできるだけ少ないパラメータで記述した特徴量が必要となる。本論文では、顔表面の勾配、または角度を抽出し量子化した後、顔の中心付近に領域を与え、その個数を数えるという手法を提案する。この手法によれば、特徴量の位置情報を大幅に削減し、次元数を減少させることができる。また、計測誤差の吸収にもなり、良好な識別ができると考えられる。

1.2 色符号化スリット光投影法の概要

この方法は平面コード方と呼ばれる3次元データ計測法の一つで、色で符号化されたスリット光を対象に投影し、それを撮影することによってスリットの投影方向から見て得られる情報と等価な情報を、対象を撮影した観測画像から得られるようにしたものである。図1に3次元形状システムの構成を、図2に3次元形状システムの座標系を示した。図3は撮影された観測画像の例であり、図4は図3から得られた3次元画像を補間し、メッシュ状のデータにして表示させたものである。本論文ではこのメッシュ状のデータを用いている。座標は図4に示したとおりであり、鼻の頂点が原点となっている。

1.3 本論文の構成

第1章は序論であり、本研究の背景と目的について述べる。

第2章では、照合方法について述べる。

第3章では、特徴量の抽出方法について説明する。

第4章では、識別実験方法、及び識別実験結果について述べる。
 第5章で、本研究の結論について述べ、今後の課題について考える。

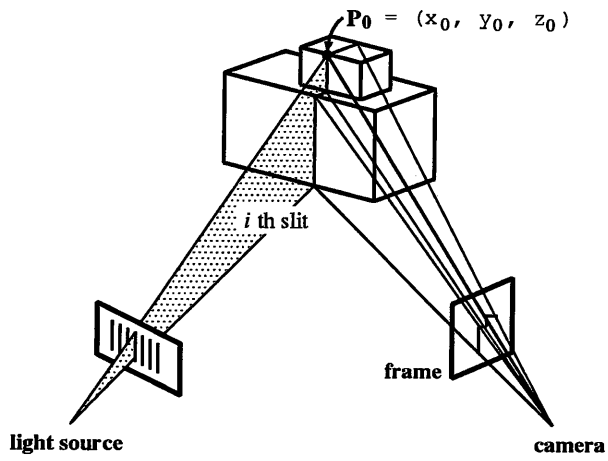


図 1: 3次元形状システムの構成

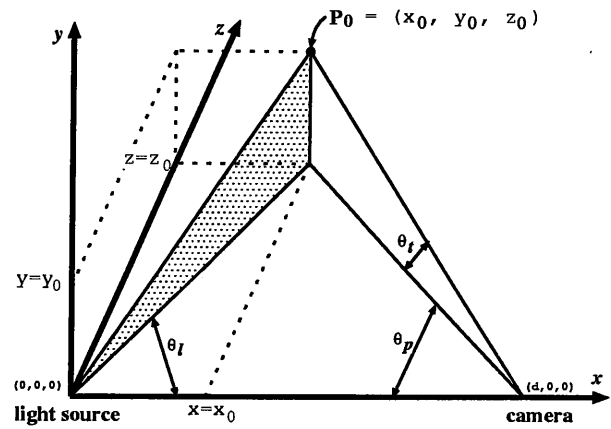


図 2: 3次元形状システムの座標系

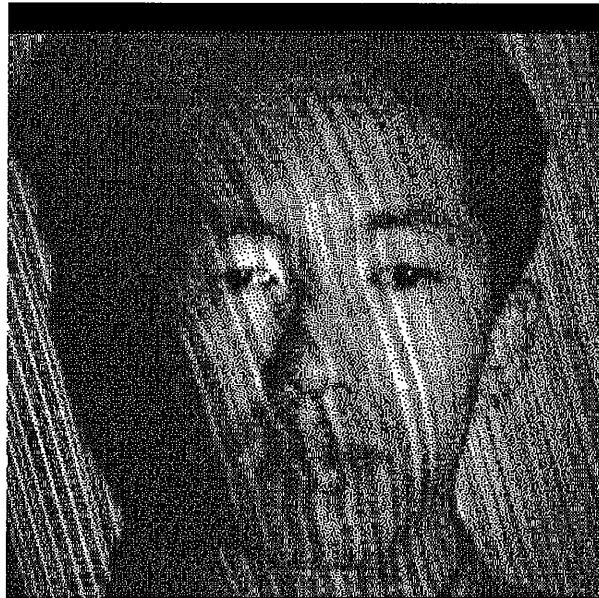


図 3: 観測画像の例

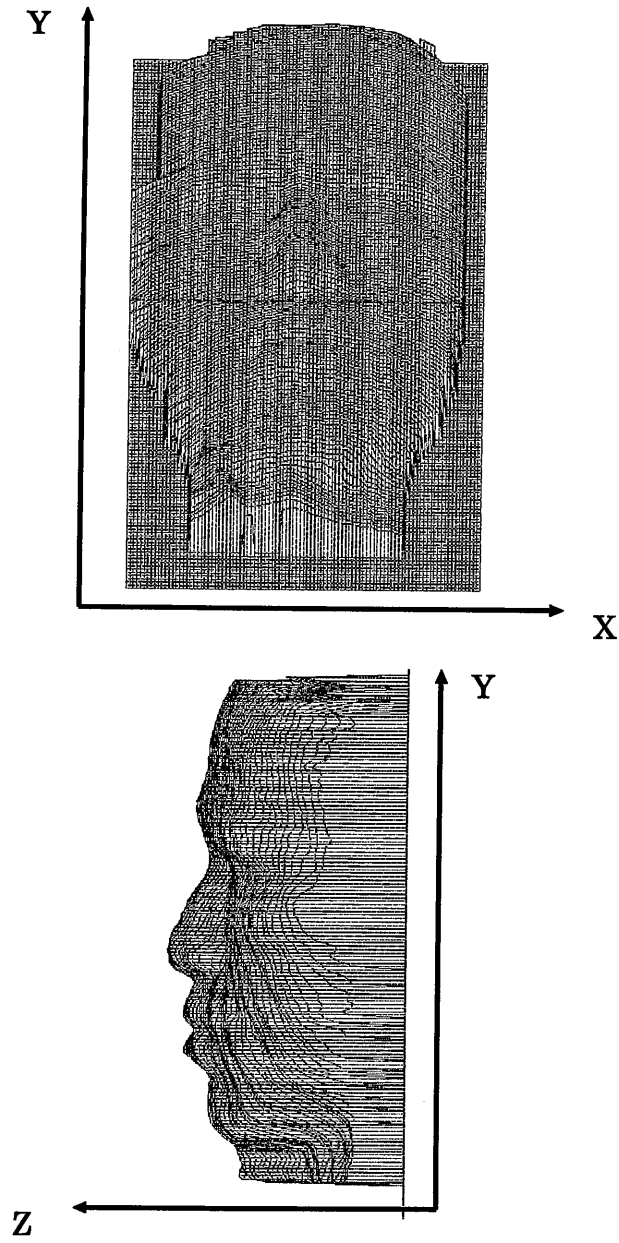


図 4: 顔の 3 次元データの例

2 照合方法

2.1 データ

今回用いたデータは、1人につき5枚で34人分の全部で170枚の3次元顔データである。画素は縦161、横101で、メッシュ状になっているデータであり、鼻頂点が原点となるようにできている。また $y-z$ 平面で回転しているものがあるので、最少二乗法により補正した。

2.2 マッチング

今回の実験では用いた顔画像のサンプル数が少ないので、次の方法を採用している。顔のデータを p_{ij} とする。ただし i は人物を表し1から34の値をとり、 j は同一人物の画像の番号を表し1から5の値をとる。この時ある人物 M の N 番目のデータを本人の入力パターンとすると、登録パターン $\overline{p_{MN}}$ は以下のようになる。

$$\overline{p_{MN}} = \frac{\sum_{j=1}^5 p_{Mj} - p_{MN}}{4}$$

これは、5枚の同一人物の顔画像のうち、1枚を入力パターンとすると、登録パターンは残り4枚の平均とすることである。このような登録パターンは、同一人物について5枚、全部で170枚できる。

1枚の入力パターンに対して、本人の登録パターン1枚と他人の登録パターン165枚の計166枚の登録パターンのマッチングを1セットとして、全部で170セットのマッチングを行なった。

2.3 パターン間の相違の尺度

2つのパターン間の違いは、以下の $L2$ 距離（ユークリッド距離） d で計算した。

$$d(\overline{p_{MN}}, p_{ij}) = \sqrt{\frac{\sum_{k=1}^L (\{\overline{p_{MN}}\}_k - \{p_{ij}\}_k)^2}{L}}$$

ここで $\{\overline{p_{MN}}\}_k$, $\{p_{ij}\}_k$ はそれぞれ登録パターン、入力パターンの特徴ベクトルであり、 L はベクトルの次元である。 $L2$ 距離が大きいほど2つのパターンの相違は大きいと見ることにした。

2.4 評価基準

2.4.1 1位認識率 RR (%)

1枚の入力パターンとの $L2$ 距離が小さい順に166枚の登録パターンを並べる。第1位の登録パターンが、その入力パターンと同一人物である場合をカウントする。その数が全

セット数 170 に対し r とすると、

$$RR = \frac{r}{170} \times 100 (\%)$$

これは正しく本人と判定される率である。

2.4.2 本人棄却率 ER (%)

170 セットのマッチングにおいて、すべての異なる人物同士の $L2$ 距離の最小値が t のとき、同一人物同士のマッチングのうちで $L2$ 距離が t より大きい場合の数を e とすると、

$$ER = \frac{e}{170} \times 100 (\%)$$

これは全ての登録パターンに対し、他人を全て棄却するような一定のしきい値を定めて照合したとき、本人が誤って詐称者と判定される率である。

つまり、本人棄却率が 0% であるならば、認識対象が ID などを持っているとすると、登録パターンとの一度の距離計算で、このしきい値を越えていなければ本人であると識別することができ、計算量を大幅に削減できる。

また、認識対象が ID を持っているならば、このしきい値を各登録パターンごとに設定するという考え方もあるが、未知の入力パターンへの信頼性が少ないことから、上記の本人棄却率を評価対象とした。

2.4.3 分離能 *Separability*

170 セットのマッチングにおいて、正しく本人と識別できた際の識別順位 1 位の距離値の平均値を D_1 、識別順位 2 位の距離値の平均値を D_2 とすると、分離能は以下の式によって求められる。

$$Sep. = \frac{D_2 - D_1}{D_1}$$

これは特徴量が同一人物と他人とを分離する能力の目安となる。

また、170 セットのマッチングのうち k 番目のセットにおいて、正しく認識できた際の識別順位 1 位の距離値を d_{1k} 、識別順位 2 位の距離値を d_{2k} とすると、最低分離能は以下の式によって求められる。

$$min.sep. = \min_{1 \leq k \leq 170} \left\{ \frac{d_{2k} - d_{1k}}{d_{1k}} \right\}$$

これは分離する時の能力の最低の値を表している。

3 特徴量抽出

3.1 3次元画像からの特徴の抽出

3.1.1 2点間の勾配の特徴抽出 *surface two-point gradient feature*

y 座標の等しい、または x 座標の等しい隣り合った点もしくは一定間隔をおいた点において、水平に対する角度を求め、それを数種類の角度に量子化する。ただしこの時、求める角度の範囲は $-90\sim+90$ である。

実際の計算では隣り合った点は常に等しい間隔であるので、2点間の差を特徴として抽出することと同じである。

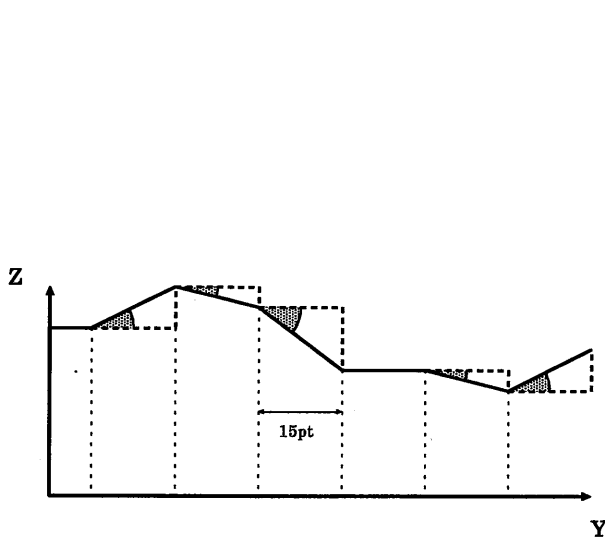


図 5: 2点間の勾配

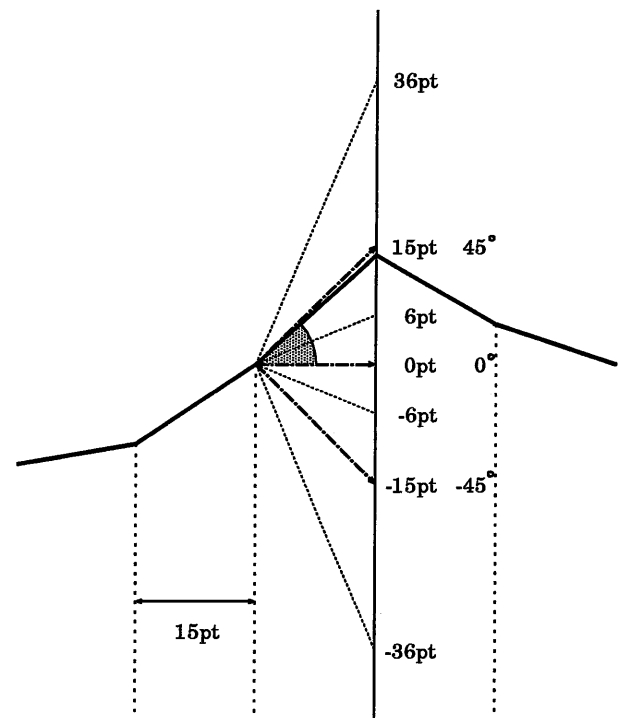


図 6: 特徴の抽出

図 5は顔画像上の2点間の勾配の様子を模式的に示したものである。水平より下側を向いている角は負とし、上側を向いている角は正としている。ただし、ここで用いられている長さの単位 pt は $1pt = 0.1mm$ である。図 6は、例として角度を3種類(-45° 、 0° 、 45°)に量子化する時の様子を示したものである。36pt、15pt、6pt、0pt、-6pt、-15pt、-36pt、はそれぞれ、 67.5° 、 45° 、 22.5° 、 0° 、 -22.5° 、 -45° 、 -67.5° の時の2点間の差を表している。つまり、2点間の差が6pt~36ptの時(斜線の角のような場合)に、 45° とすると決めている。

この手法の特徴の1つは、微少な回転には対応できるということである。 z 座標をそのまま特徴量とすると、データが微少に回転していた場合でも、原点付近では誤差は少なく済むが、原点から離れた点ではかなり大きくずれてしまい、信頼性が小さくなる。この手

法によれば、原点付近であろうと原点より離れていようと、データの誤差はその微小回転分だけとなる。さらに、量子化によりその微小誤差は吸収されることが期待できる。

2つめの特徴は、原点(鼻)付近の特徴を抽出できることである。 z 座標でみると、前述のように微小な回転を考えた時、信頼性の大きい点は原点に近いところであると言える。また原点は、人間の顔の中でも凹凸がはっきりし、特徴がある場所と考えられる鼻となっている。しかし、原点はすべてのデータを合わせる点であり、異なるデータ間でも違いが少ない。この手法によれば、前述の通り回転に対する信頼性はどの点でも同じではあるが、データを原点で合わせてあるということに対しては影響がない。

以上のことから、このような特徴を抽出したのである。

3.1.2 3点間の角度の特徴抽出 *surface three-point angular feature*

y 座標の等しい、または x 座標の等しい隣り合う3点、もしくは一定間隔をおいた3点において、水平に対して下側の角度を求め、それを数種類に量子化する。ただしこの時、角度の範囲は $0\sim+360$ である。

この場合は、実際の計算で、角度を逆三角関数で求めなくてはならないので、2点間の勾配の特徴抽出より計算時間が多くなる。

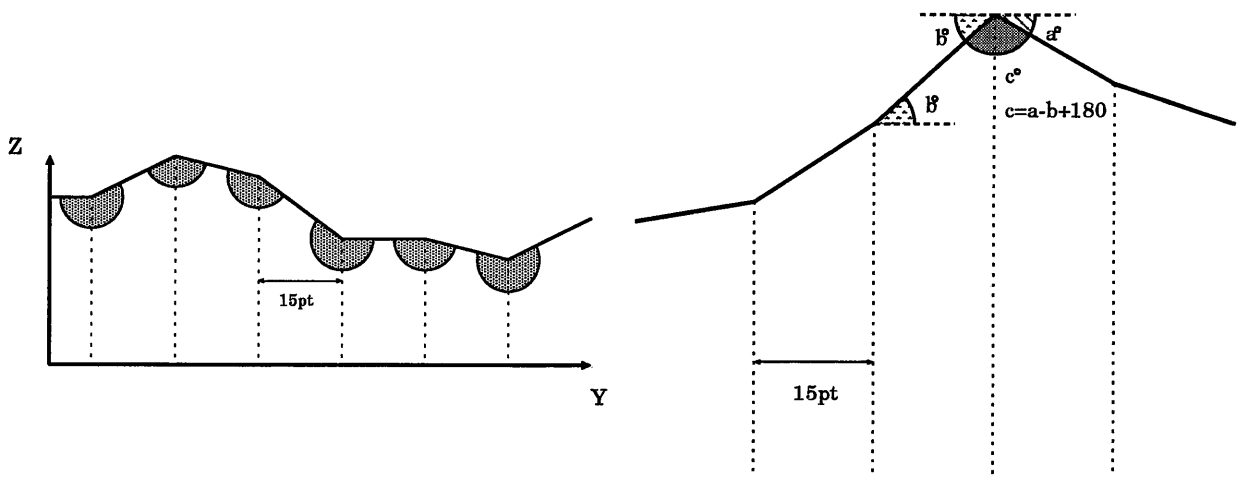


図 7: 3点間の角度

図 8: 特徴抽出

図 7は顔画像上の3点間の角度の様子を模式的に示したものである。図 8は角度の計算法を示したものである。求めるポイントで、前述の2点間の勾配の角度を \arctan で実際に求め、一つ前のポイントにおける2点間の勾配の角度との差をとって 180° を加えることで、角度が得られる。その後量子化においては、例えば、 $0^\circ\sim40^\circ$ 、 $40^\circ\sim80^\circ$ 、 $80^\circ\sim120^\circ$ 、 $120^\circ\sim160^\circ$ 、 $160^\circ\sim200^\circ$ 、 $200^\circ\sim240^\circ$ 、 $240^\circ\sim280^\circ$ 、 $280^\circ\sim320^\circ$ 、 $320^\circ\sim360^\circ$ に分け、それぞれ、 20° 、 60° 、 100° 、 140° 、 180° 、 220° 、 260° 、 300° 、 340° 、で代表させるようにする。

この手法の特徴は、3点間の角度は回転に対して保存されるということである。それ以外は2点間の勾配の特徴抽出と同じである。

3.2 領域における特徴量の抽出

領域は y 軸に平行な線、または x 軸に平行な線を中心とし、その両側に一定の幅を与え、その中心線に垂直に数個、均等に分割して作る。この領域において先の量子化された角度の個数を数え、これを特徴量とする。つまり、次元数は (領域の数) \times (角度の種類) となる。また、分割の方法として次の 2 種類を選ぶ。

- 領域を半分ずつ重ねていく
- 領域を隣接させる

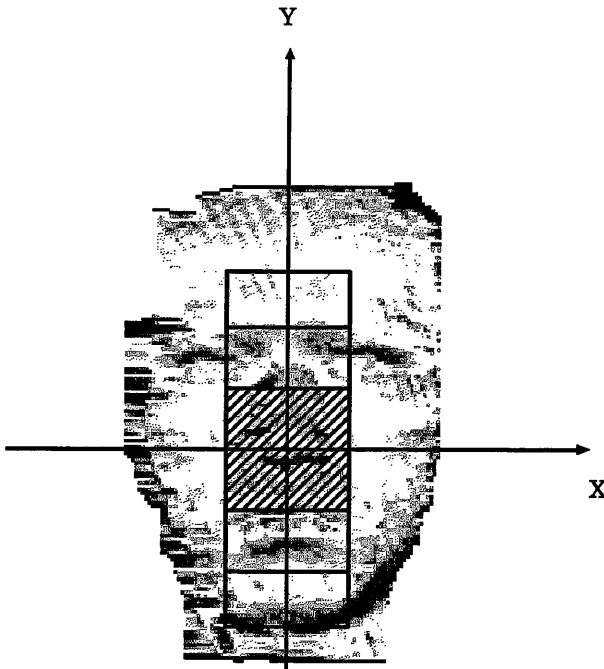


図 9: 領域を半分ずつ重ねる方法

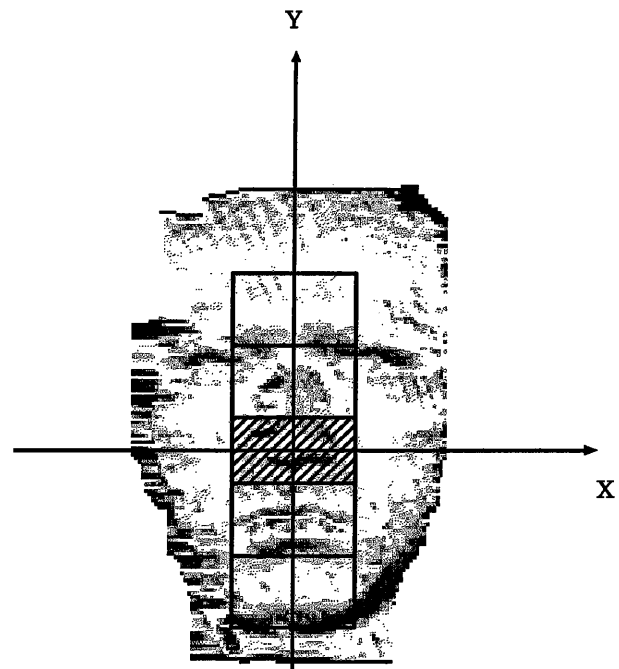


図 10: 領域を隣接させる方法

図のどちらも、5 つ領域を取っている。斜線部分が 1 つの領域である。この 2 つの方法を選んだのは、以下の理由である。領域を隣接させる方法は、全領域を網羅するための方法として一般的であると考えたためである。領域を半分ずつ重ねる方法は、前者の領域を隣接させる方法だと、境界となっている部分の情報が生かされていないと考えたためである。ただし、この方法だと両端の領域の半分だけが 1 度しか数えられず、残りの部分は 2 度数えられることになり、不自然感が残る。そのことについては、両端の領域が計測誤差などで、中心付近に比べ信頼性が少ない、という理由で採用の理由としておく。

この手法の有効性は、位置情報の大幅な削減にある。誤差の吸収に対応し、前段階で量子化することによる、位置情報をもった特徴としての不適格さを吸収することにある。つまり、量子化を行うと、例えば、「 $x = 50$ 、 $y = 25$ の点の特徴は 42.3 である」という情報が、「 $x = 50$ 、 $y = 25$ の点の特徴は 50 である」と変わるわけで、この「 $x = 50$ 、 $y = 25$ の点の」の部分に対しては「特徴」は不適格となる。しかし領域を取って個数を数えること

はこの「 $x = 50$ 、 $y = 25$ の点の」の部分が無意味にすることができ、位置情報を削減することが有効に働くことになるのである。残念ながら、位置情報を削減することが、有効に働くことについて、パラメータとの相関関係までは言及することができない。以下の実験においても、目標が特徴量の次元の削減と、認識率の向上にあるため、この部分の考察については多くを語れないであろう。

3.3 補正

勾配の角度を求める際に量子化を行うわけだが、量子化誤差が大きくなるように補正を行なうことにする。補正の方法は、量子化をする際に次の点の z 座標を量子化した角度に合わせた z 座標にするという方法である。次元数を減らすために、実際の分布に対して隔たりの大きい量子化をするので、有効であると考えられる。図 11はこのことを説明したものである。

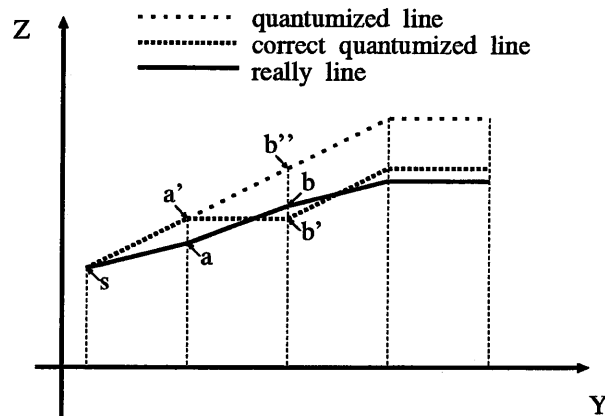


図 11: 補正の方法

図 11において、右方向に特徴の抽出を行っているとする。点 s と点 a との角度は、点 s と点 a' の角度に量子化される。もしも、次に点 a と点 b の角度を量子化すると、それは点 a' と点 b' との角度に量子化される。“量子化された線 (quantumized line)”とは、このように作られた線である。補正した場合はどうなるかという、点 s と点 a の角度を求めた後、次の点 a は点 a' になったとみなして、点 a' と点 b との角度を量子化し、それは点 a' と点 b' との角度となる。そこでまた新たに点 b を点 b' の代わりとする。このようにして作成されたのが“量子化され、補正した線 (correct quantumized line)”である。図を見て明らかのように、補正を行ったほうがよく特徴を捉えている。

ただしこれは 2 点間の勾配の特徴抽出にしか使えないものである。その理由は、3 点間の角度の特徴抽出の場合は図 8 において a' が -90° に近い場合で、量子化で実際の角度より小さくなる時、かなり大きなずれとなってしまうからである。

4 実験方法

4.1 特徴量の抽出法の絞り込み

4.1.1 x 、 y 偏微分と領域の並べ方

x 偏微分 (y 座標の等しい2点での勾配) を用いたもの、 y 偏微分 (x 座標の等しい2点での勾配) を用いたものをそれぞれ垂直方向 (y 方向) に領域を並べたものと、水平方向 (x 方向) に領域を並べたものを行った。領域の中心線の座標がどこあれば最適かを見るとともに、4つの方法の検討し、結果の最良のものを後の実験で用いることにする。

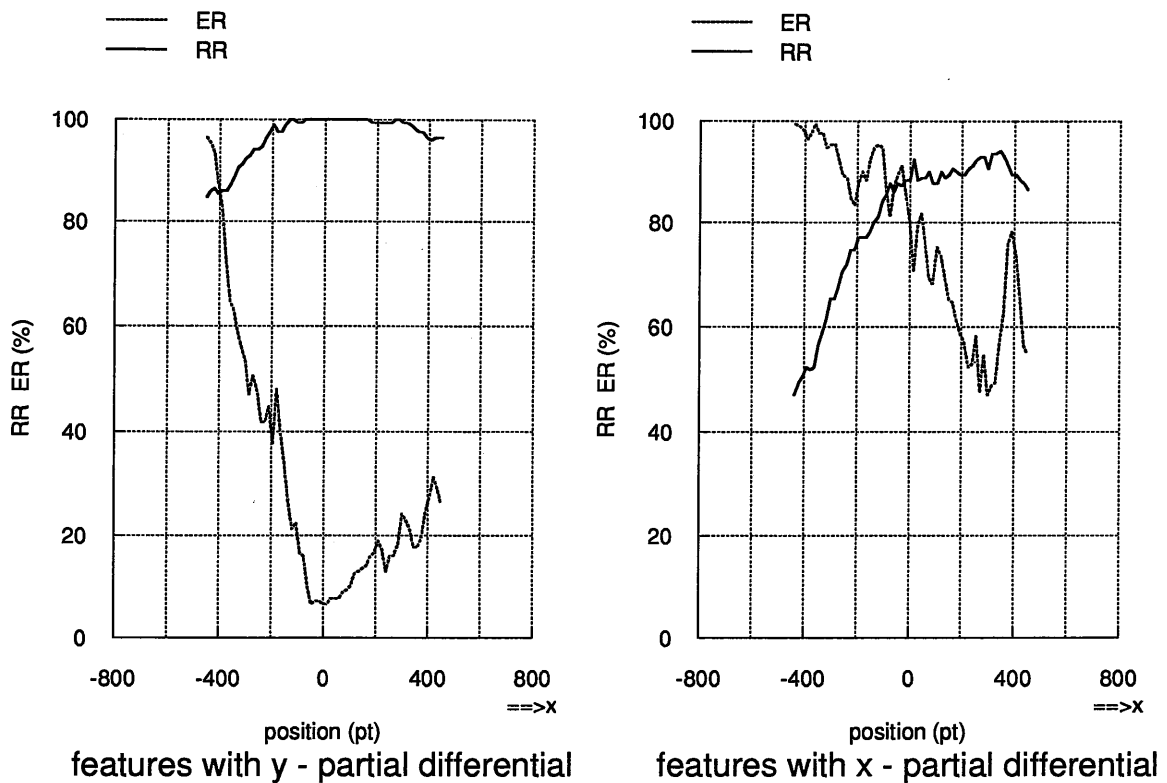


図 12: 垂直方向に領域を並べた場合

図 12は、前章の図 10のように、 y 軸と平行な線を中心線とした領域をとり、 x の小さい方から大きい方へと移動させた時の、1位認識率と本人棄却率を表したものである。ただし、横軸の $position$ は、領域の中心線の x 座標を表している。領域は、 $-900 \leq y \leq 900$ (pt)、 $\Delta x = 600$ (pt) の範囲を y 軸と垂直に、かつ隣接するように5個に分けたものであり、中心線が $x = 0$ の時は、図 10の (位置、比率とも、ほぼ等しい) 状態である。これより、 y 偏微

分を用いた方が、格段に結果が良いことが分かる。

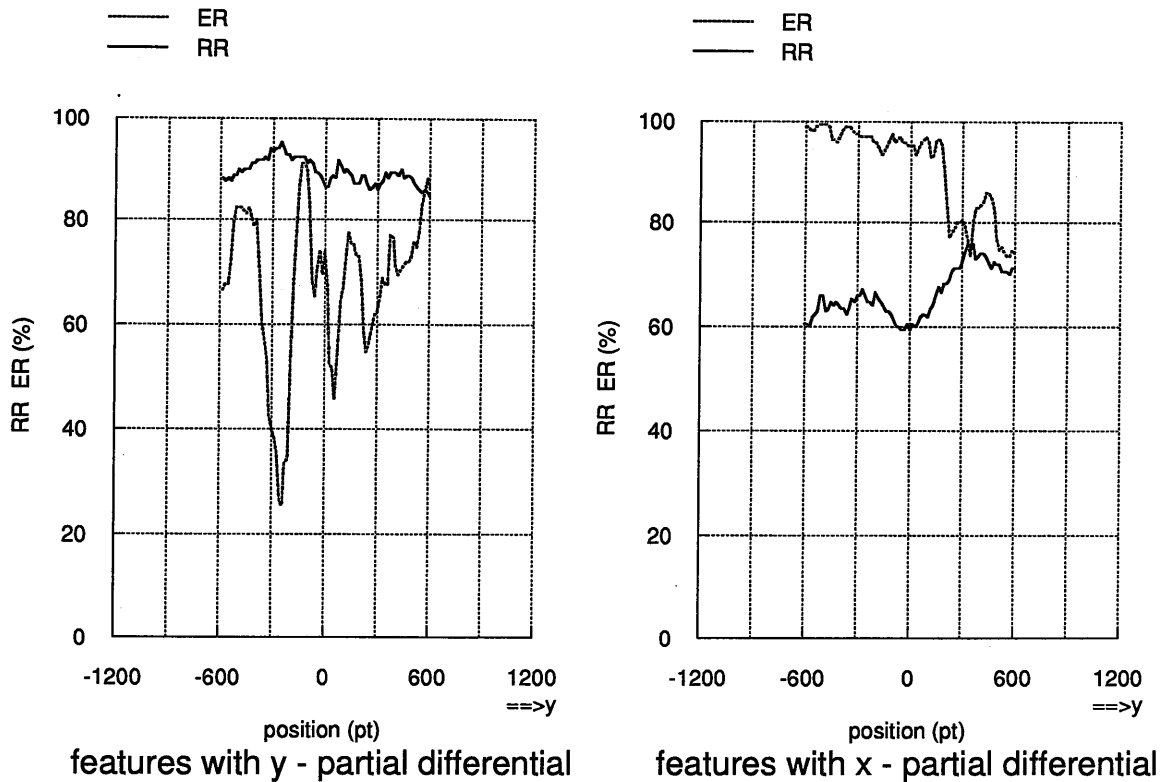


図 13: 水平方向に領域を並べた場合

図 13は、図 14のように領域を x 方向に並べたとき x 軸と平行な線を中心線とした領域をとり、 y の小さいほうから大きい方へと移動させた時の、1 位認識率と本人棄却率を表したものである。ただし、横軸の $position$ は、領域の中心線の y 座標を表している。領域は、 $-525 \leq x \leq 525$ (pt)、 $\Delta y = 1200$ (pt) の範囲を x 軸と垂直に、かつ隣接するように 5 個に分けたものであり、中心線が $y = 0$ の時は、右図の実線の部分となる。このような 5 つの領域を取ることで、 y 方向に領域を並べる方法では最も良い結果を出しているのだが、先に挙げた x 方向に領域を並べる方法と比べると、結果はあまり良くない。したがって、後の実験はすべて y 偏微分を用い、中心線を y 軸と重ね (x 座標が 0)、 y 方向に領域を並べる方法を採用する。

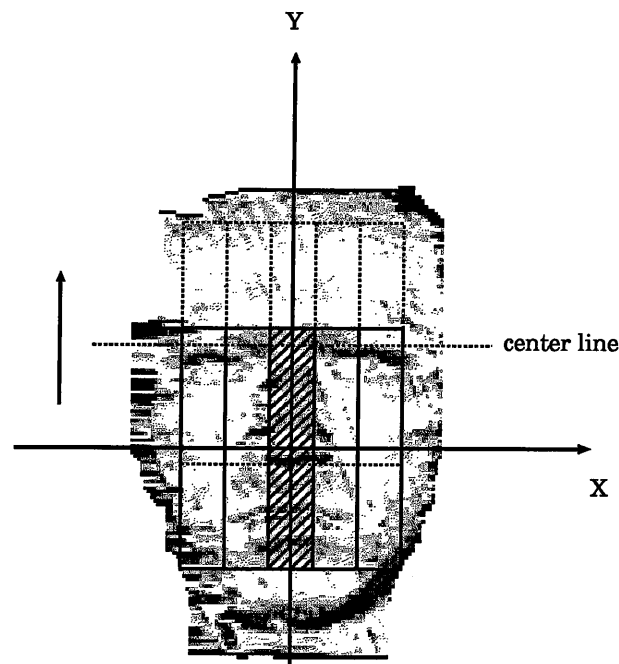


図 14: 水平方向に領域を並べた状態

4.1.2 全領域の y 成分の範囲

今回、決定すべきパラメータが多いので、絞り込む対象として、領域全体の y 成分の範囲を考えた。予備実験等により、経験的に y 成分の範囲は $-900 \leq y \leq 900$ が良好であることを、以下の実験で効果を検証した。

上限もしくは下限を設定した時の下限もしくは上限の効果 まず、 y 成分の下限を -900 に設定し、 y 成分の上限を $0 \sim 1200$ まで動かす。与えられた範囲で均等に 5 つ領域を取り、特徴量を抽出した後、認識実験を行い 1 位認識率と本人棄却率を求めた。図 15 は上限の位置が 0 に近い時、図 16 は上限の位置が 1200 に近い時の領域の状態を表している。この上限を動かす方法では領域を半分ずつ重ねる方法で示してある。

次に、 y 成分の上限を 900 に設定し、 y 成分の下限を $-1200 \sim 0$ まで動かす。下限を設定する方法と同様に、与えられた範囲で均等に 5 つ領域を取り、特徴量を抽出した後、認識実験を行い 1 位認識率と本人棄却率を求めた。図 17 は下限の位置が -1200 に近い時、図 18 は下限の位置が 0 に近い時の領域の状態を表している。この下限を動かす方法では領域を隣接させる方法で示してある。

以上の2つの実験を領域を隣接させる方法と、領域を半分ずつ重ねる方法についてそれぞれ行った。

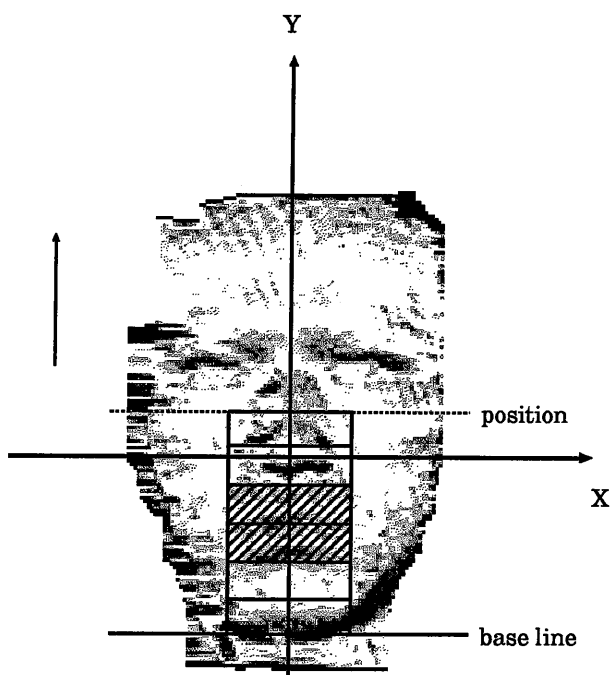


図 15: 上限を動かした際の領域の状態 (上限が 0 付近)

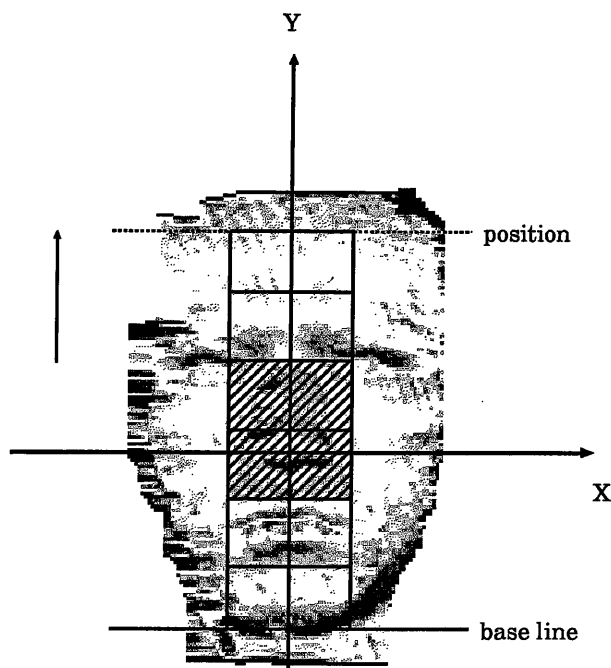


図 16: 上限を動かした際の領域の状態 (上限が 1200 付近)

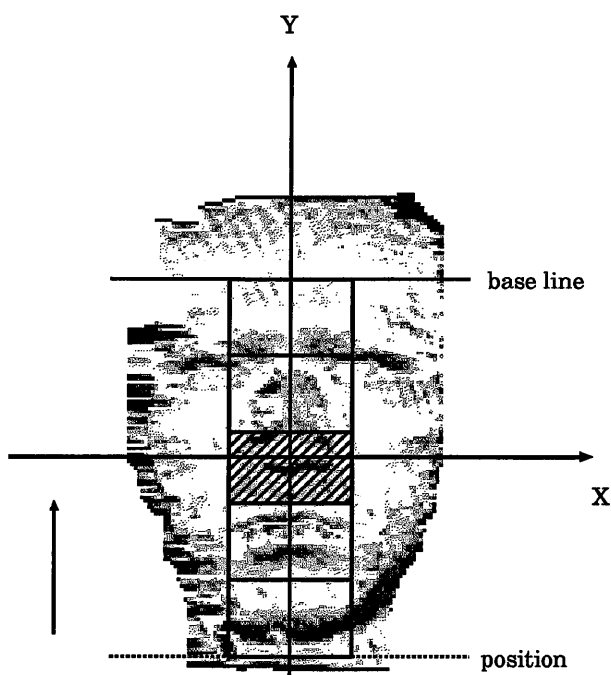


図 17: 下限を動かした際の領域の状態 (下限が -1200 付近)

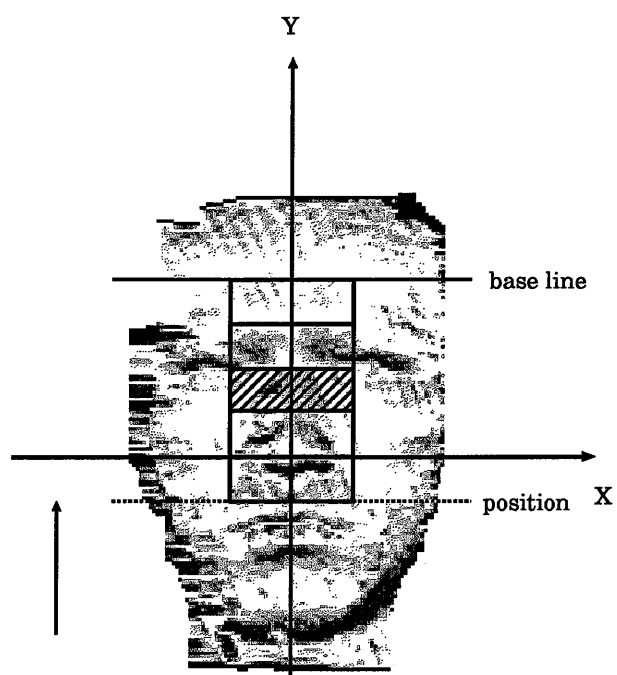


図 18: 下限を動かした際の領域の状態 (下限が 0 付近)

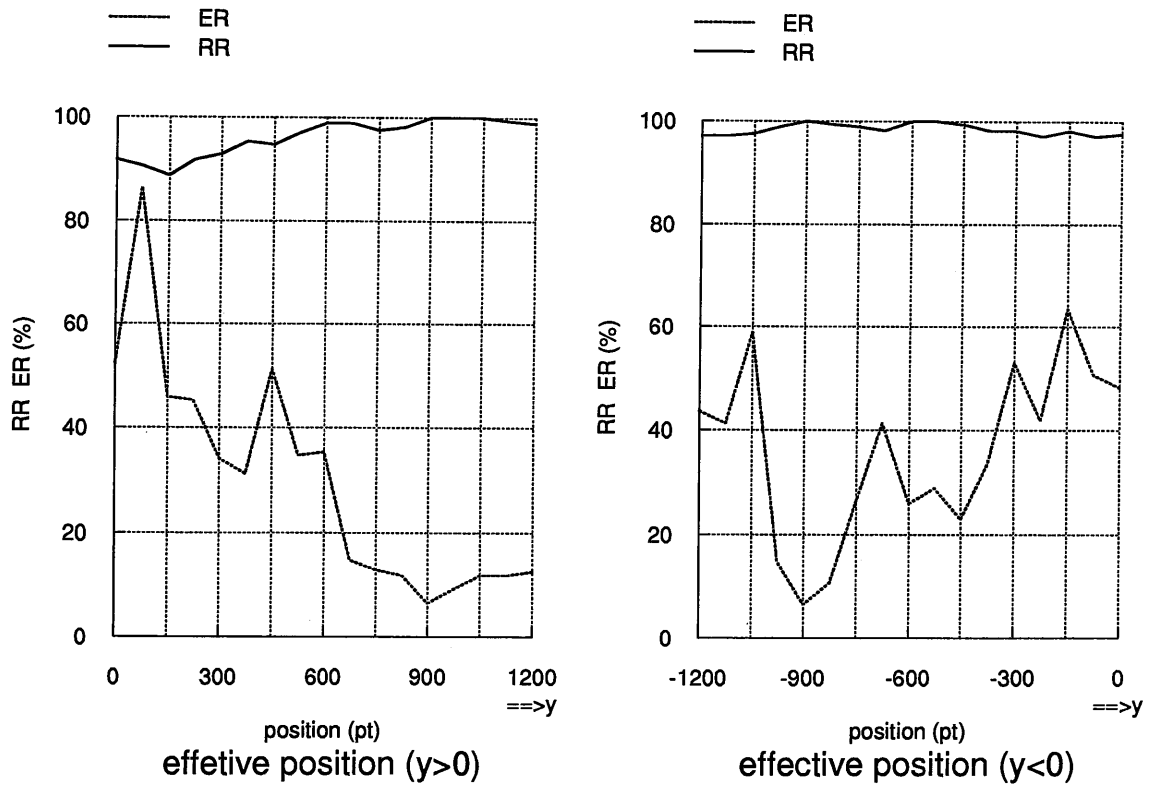


図 19: 領域を隣接させた場合

図 19は領域を隣接させた場合についての結果である。領域の横幅は $40pt$ で行った。これを見ると、1位認識率は100%になっている部分がある所があるため、どの位置に上限、下限を決定すれば効果的かということは判断できない。しかし本人棄却率を見ると、下限を $-900pt$ に設定した時は上限が $900pt$ 、上限を $900pt$ に設定した時は下限が $-900pt$ で、最低となっている。本人棄却率は0%に近い方が結果が良好であると評価できるものであるため、効果的な y 成分の範囲は $-900 \leq y \leq 900$ であると言える。

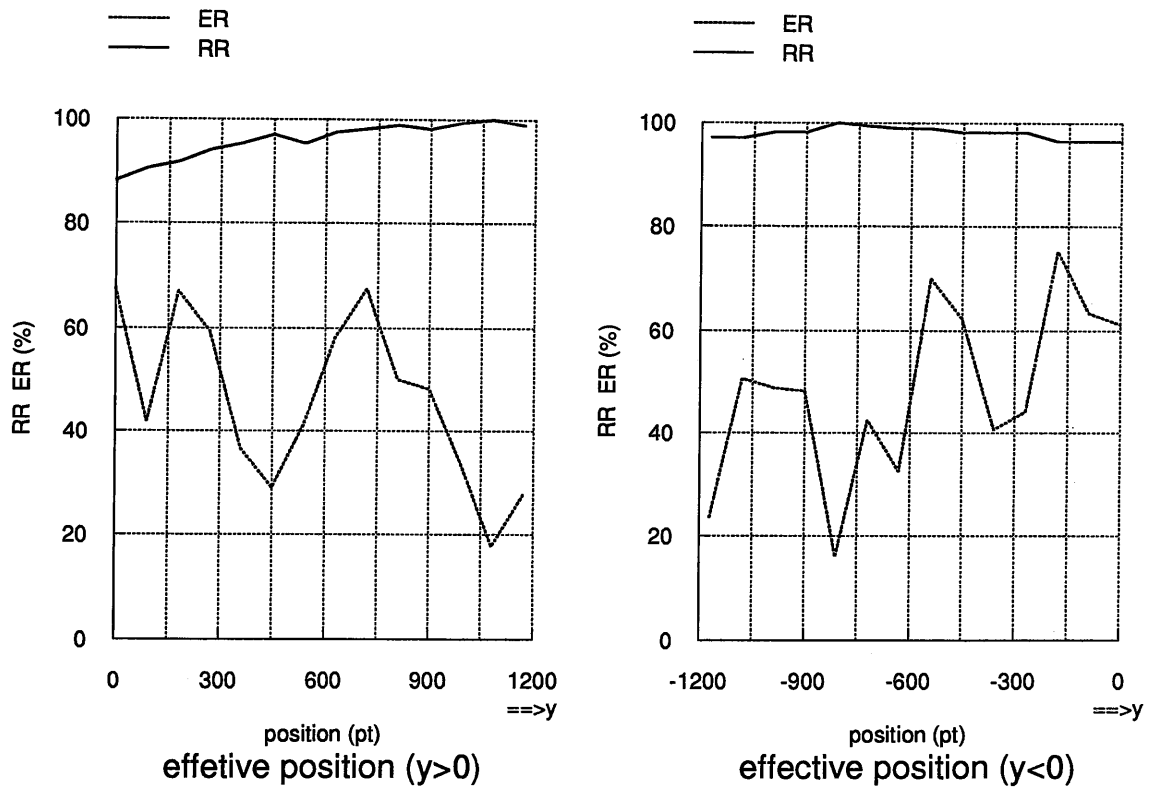


図 20: 領域を半分ずつ重ねた場合

図 20は領域を半分ずつ重ねた場合についての結果である。領域の横幅は $40pt$ で行った。これを見ると、下限を $-900pt$ に設定した時は上限が $1100pt$ 付近、上限を $900pt$ に設定した時は下限が $-800pt$ 付近で結果が良好となっている。この後さらに、上限を $1100pt$ に設定し下限を動かしたり、上限を $-800pt$ に設定し上限を動かしたり、と繰り返せば最も効果的な y 成分の範囲が見つかるかもしれない。しかし最も効果的な y 成分の範囲というのはデータ、もしくは他のパラメータに依存する可能性があるので、この実験からは、 $-900 \leq y \leq 900$ を y 成分とすると効果的であるとして良いと判断することにした。

中心を設定した時の全領域の上下の幅の効果 さらに次のような実験も行った。 $y = 0$ 、つまり x 軸を中心に上下に同じ幅を取り、それを領域全体の y 成分の上限、下限とする。与えられた範囲に均等に 5 つ領域を取り、特徴量を抽出し、識別実験を行い、1 位認識率と本人棄却率を表す。領域全体の y 成分の範囲は $1200 \sim 2400(pt)$ で動かした。

図 21、図 22は領域を隣接させる方法について、この実験方法を説明したものであり、図 23、図 24は領域を半分ずつ重ねる方法について説明したものである。

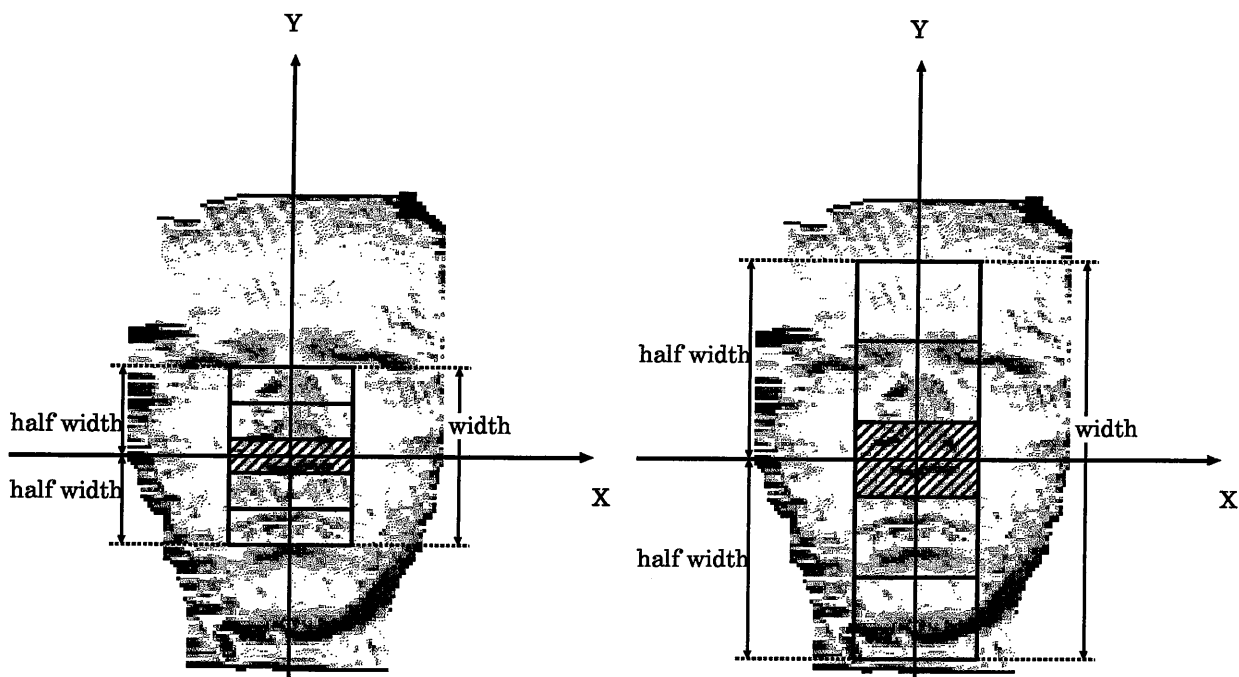


図 21: y 成分の幅の小さい領域の状態 (隣接) 図 22: y 成分の幅の大きい領域の状態 (隣接)

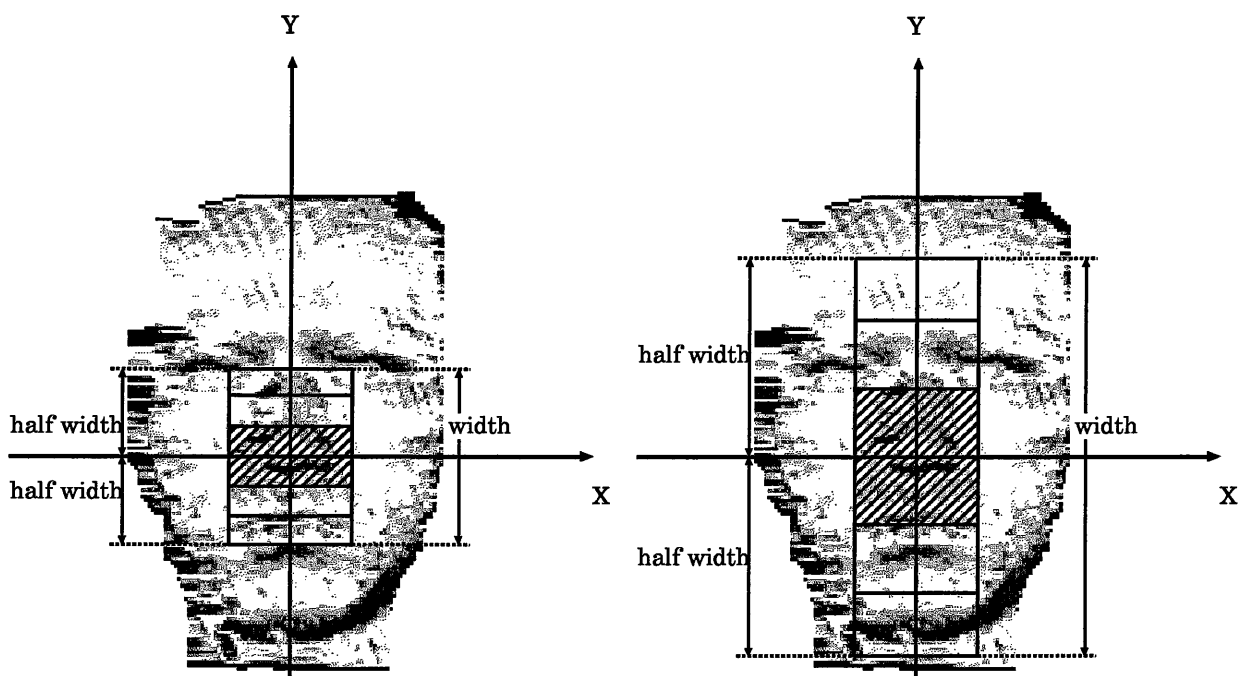


図 23: y 成分の幅の小さい領域の状態 (重複) 図 24: y 成分の幅の大きい領域の状態 (重複)

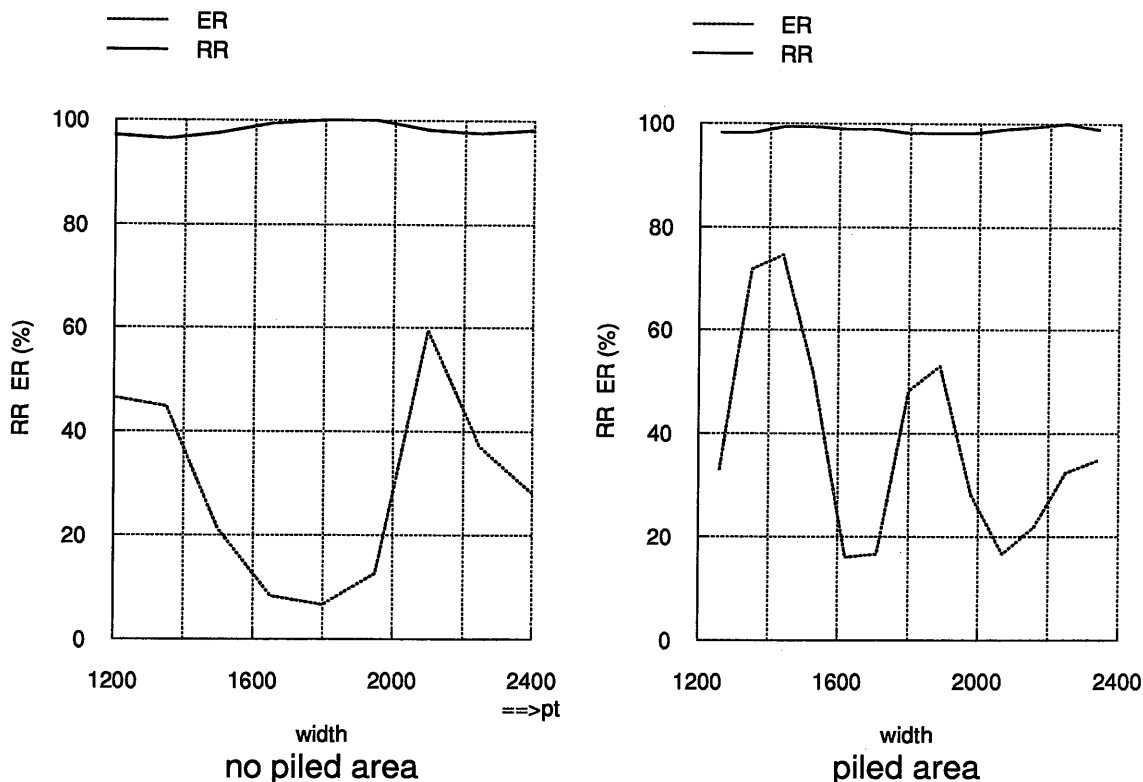


図 25: 中心設定の領域の幅の効果

図 25 は実験結果を示したものである。領域の横幅は $40pt$ にしてある。左側は領域を隣接させる方法についての結果で、右側は領域を半分ずつ重ねる方法についての結果である。これを見ると、領域を隣接させる方法では $1800pt$ 、つまり $-900 \leq y \leq 900$ の範囲を取れば良好な結果を得ることが分かる。また、領域を半分ずつ重ねる方法では $1600pt$ 付近、つまり $-800 \leq y \leq 800$ の範囲を取れば良好な結果を得ることがわかる。

以上のことから、領域全体の y 成分の範囲は $-900 \leq y \leq 900$ とすることに決めた。

4.1.3 領域の重なり

領域の重なりについて、前述のように領域を半分ずつ重ねる方法と領域を隣接させる方法について採用しているが、領域を重ねることで、どの程度影響が出るものか実験してみた。

領域全体の y 成分の範囲を $-900 \leq y \leq 900$ に設定し、重なりが $0pt$ の状態から $50pt$ の状態になるように領域を作成する。その際、特徴量を抽出し識別実験を行い、1 位認識率、

本人棄却率を求めた。図 26 は重なりが $0pt$ の状態である。図 27 は重なりが $10pt$ で、領域の幅が $32pt$ の時で、領域の約 $\frac{3}{10}$ が重なっている状態である。図 28 は重なりが $20pt$ で、領域の幅が $40pt$ の時で、領域の半分が重なっている状態である。図 29 は重なりが $45pt$ で、領域の幅が $60pt$ の時で、領域の $\frac{3}{4}$ が重なっているじょうたいである。また各図に、顔の左側に各領域を重ねずに並べたものを示した。

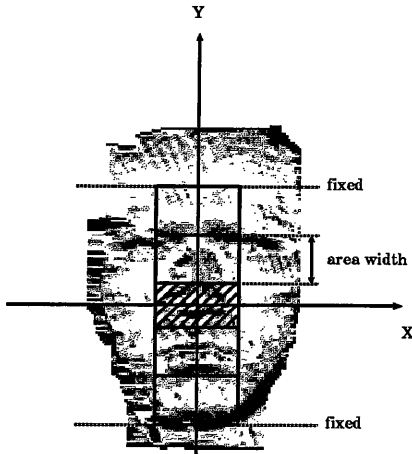


図 26: 重なりのない領域の状態

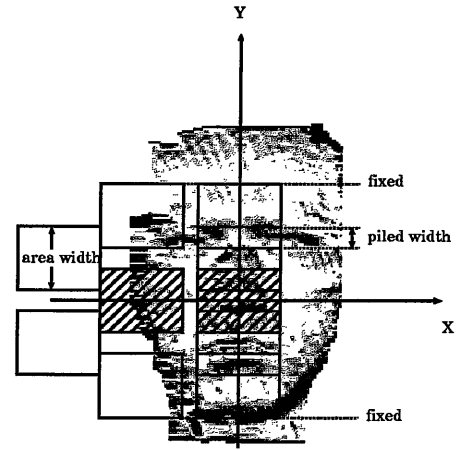
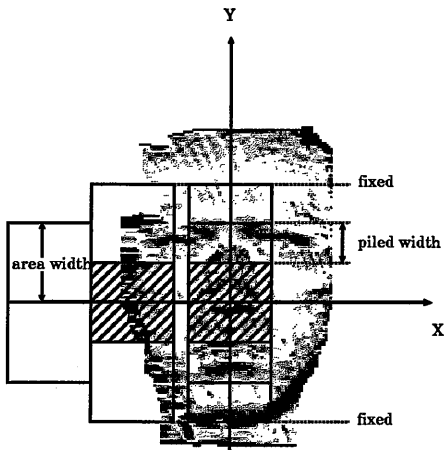
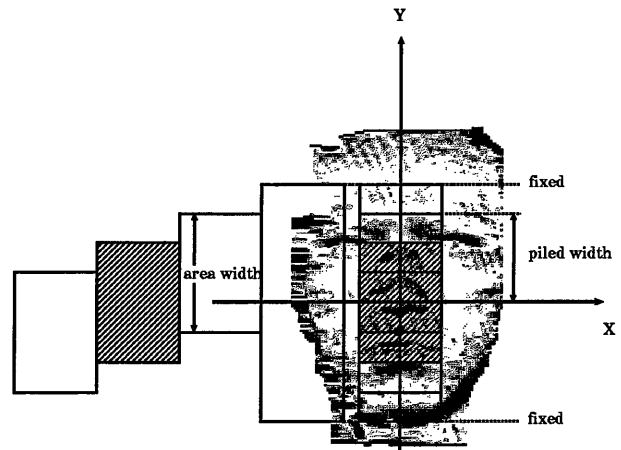
図 27: $\frac{3}{10}$ 重なった領域の状態

図 28: 半分重なった領域の状態

図 29: $\frac{3}{4}$ 重なった領域の状態

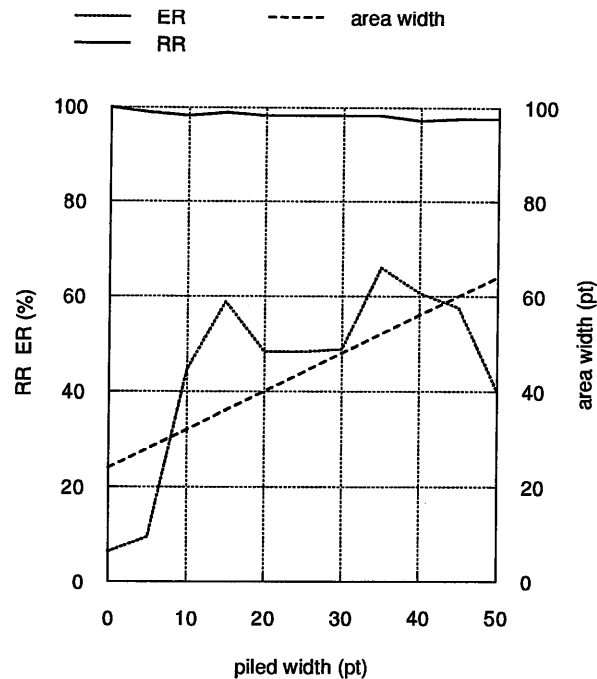


図 30: 重なりの効果

図 30は実験結果を表している。領域の横幅は40ptにしてある。これによると、1位認識率は重なりが大きいほど悪くなっていくのがわかる。本人棄却率も重なりが大きいとほど悪くなっている。領域全体の y 成分の上限もしくは下限を設定し、それぞれ下限、上限を動かす実験からも、領域を半分重ねる方法は $-900 \leq y \leq 900$ が最良ではないことが分っている。しかし、結果はデータもしくは他のパラメータに依存する可能性があるので、ここでも $-900 \leq y \leq 900$ の範囲で良いと判断することにする。

4.2 2点間の勾配を用いた特徴量について

- 2点間の勾配を用いた特徴抽出は3、5、7、8、9、11、17種に量子化したものを行った。
- 角度を求める時の点の間隔を1、5、10ptと変えたものを行った。(7種に量子化したものだけ)
- 量子化の際、補正をしないものも対照として行った。(7種に量子化したものだけ)

4.3 3点間の角度を用いた特徴量について

- 3点間の角度を用いた特徴量は9、17、19種に量子化したものを行った。

- 角度を求める時の点の間隔を1、5、10pt と変えたものを行った。

4.4 領域における特徴量につて

- 領域の個数を3~11 と変えた。ただし、目標が特徴量の次元 $\{(\text{領域の数}) \times (\text{角度の種類})\}$ の削減にあるため、角度の種類によって、変動する。
- 領域の横幅を30~50pt と2pt ずつ変え、角度の種類と領域の数の1つの組み合わせに対し、11回の認識実験を行った。た。

5 結果

5.1 2点間の勾配の特徴抽出

以下に角度の種類別に次元数が最も小さく、1位認識率100%となったものを挙げる。

	3種類	5種類	7種類	8種類	9種類	11種類	17種類
次元数	該当なし	35次元	35次元	32次元	45次元	55次元	51次元
領域の 形態		no piled	no piled	no piled	no piled	no piled	no piled
					piled	piled	

3種類に角度を量子化したものは、量子化する際に67.5°より大きい角度は特徴として抽出しないようにしてしまったので、欠落が生じ、良い結果が出ていないと考えられる。8種類に量子化したものは、0°が量子化する際の境界となってしまうので、量子化時に不自然な切り方になっているが、結果は良い。不自然に切るならば、他にも良い方法が考えられるので、次章で示す。

量子化時に補正をしないものは、角度を7種類に分ける方法のみを行ったが、補正したものと同じく35次元で100%が出た。ただし、補正した方が、領域の横幅のパラメータを11種に振らせたうち、9種で100%になったのに対し、補正しなかった方は1種だけであった。つまり補正を行ったほうが結果は良いといえる。

2点間の距離を5、10ptにしたものは、同じく7種に量子化したものだけを行ったが、5ptの方は35次元で100%が出たものの、補正しなかったものと同様に、横幅のパラメータを変え時に1つしか100%にならず。10ptの方は100%の該当はなかった。これより、角度の特徴を出すときに点の間隔を空けるものは有効ではないと言える。

5.2 3点間の角度の抽出

この方法では1位認識率100%の出たものはなかった。また、点の間隔を5pt、10ptにしたものでも50次元程度で100%となったものはなかった。

5.3 検討

3点間の角度を用いた特徴が良い結果が出ない。これには抽出された特徴の分布に対し、量子化の際の切り方が有効でないと考えられる。次章において、改良した実験を行なう。

2点間の勾配の角度を用いた方は、良い結果が得られた。8種類に量子化したものは、不自然な量子化となっているが32次元で、他にも35次元で100%となったことで、認識率の向上と、次元の削減の両方の課題を達成していると言える。

6 特徴量抽出の改良

6.1 特徴の分布

3点間の角度の特徴抽出では、その特徴の分布が良くないと考えられるため、分布を確かめてみた。170枚全てのデータから特徴のとれる全ての点で量子化をしない角度(正確には整数に量子化しているが)の個数を数え、グラフに表す。

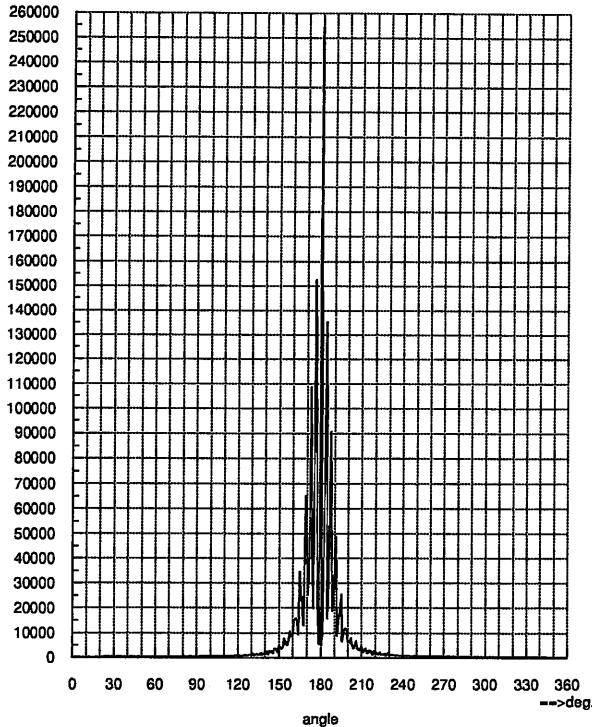


図 31: 3点間の角度の分布

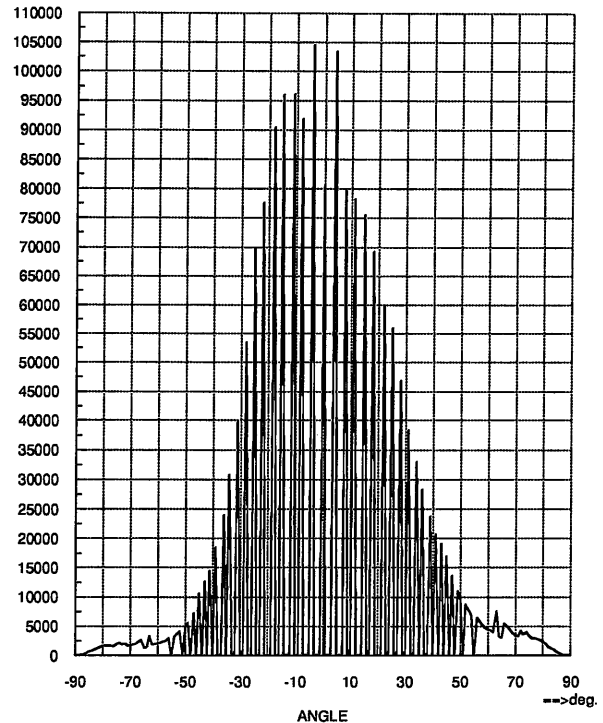


図 32: 2点間の勾配の分布

図 31の3点間の角度の分布と、図 32の2点間の勾配の分布では大きな違いが見られる。3点間の角度の分布は180°とその付近が非常に多くなり、残りはほとんどない状態となっている。2点間の勾配だと角度のちらばりがあるが、隣り合った3点だと、ちらばりがない、つまり特徴が出にくいと言える。そこで次の方法を考えた。

6.2 量子化の改良

今までは量子化の際、2点間の勾配では、8種にわけるとは、0°を必ず入れて、前後90°を均等に数個に分割したものに量子化するように(ただし、-90°、90°はありえないので除く)していた。例えば5種の場合、まず0°、そして、 $90^\circ/3 = 30^\circ$ より、 $\pm 30^\circ$ 、 $\pm 60^\circ$ に量子化する。つまり、それぞれ $-75^\circ \leq \theta < -45^\circ$ を -60° に、 $-45^\circ \leq \theta < -15^\circ$ を -30° に、 $-15^\circ \leq \theta \leq 15^\circ$ を 0° に、 $15^\circ < \theta \leq 45^\circ$ を 30° に、 $45^\circ < \theta \leq 90^\circ$ を 60° に量子化するという

ことである。そして75°より大きい、もしくは-75°より小さいものは取り扱わなかった。

3点間の量子化では、微妙に異なり、360°を数個に均等に分割して作った幅で、その各幅の中心の角度に量子化するようにしていた。実際には行っていないが、例えば5種に量子化しようとする、 $360^\circ/5 = 72^\circ$ より、 $0^\circ < \theta \leq 72^\circ$ を36°に、 $72^\circ < \theta \leq 144^\circ$ を108°に、 $144^\circ < \theta < 216^\circ$ を180°に、 $216^\circ \leq \theta < 288^\circ$ を252°に、 $288^\circ \leq \theta < 360^\circ$ を324°に量子化するというのである。ただし、量子化した数値は用いないので、何でもよい。

今回は、「均等に」というのをやめて、分布に合わせた切り方にする。「分布に合わせた」というのは量子化する際、それぞれの角度の幅に入る個数を均等にするということである。つまり、数の多い、2点間の勾配の方では0°付近、3点間の角度の方では180°付近は角度の幅を小さくし、数の少ない、2点間の勾配の方では±90°付近、3点間の角度の方では0°、360°付近は角度の幅が大きくなるのである。ただし、角度の幅は、0°、もしくは180°に対し対称に取った。

170枚の全てのデータより得られた分布より、角度の幅を決めるのだが認識時に使うデータまで入れてしまうので、クローズ実験になるとも言えるが、170枚中の1枚のデータでは影響がほとんどないとみることにする。

この方法では、2点間の勾配の方は-90°~90°までのすべての範囲を使用し、3点間の角度の方は0°~360°までのすべての範囲を使用するので、違いはない。また、どちらも補正は行わない。3点間の角度の方は、前回の理由もあって行わないのだが、角度の幅の中心を代表にするか、個数で中心になる角度を代表とするか、さらに(個数)×(角度)の平均を代表とするかなど、どれが適切か考えにくいので、今回はやらなかった。

また、領域における特徴量を出す時に使用する全領域で、分布を確かめる方が有効かと考えられるが、領域の大きさも変えるので、その度に計算させるのは大変なので、顔全体から得たもので行なう。領域が決まり、さらに最適なものを求める時にはそのようにした方がよいと考える。

図33と図34に主に使用する、特徴量を出す時の全領域($-900 \leq y \leq 900$ 、 $-300 \leq x \leq 300$)から得られた分布も示しておく。さほど変わらないのが分かる。

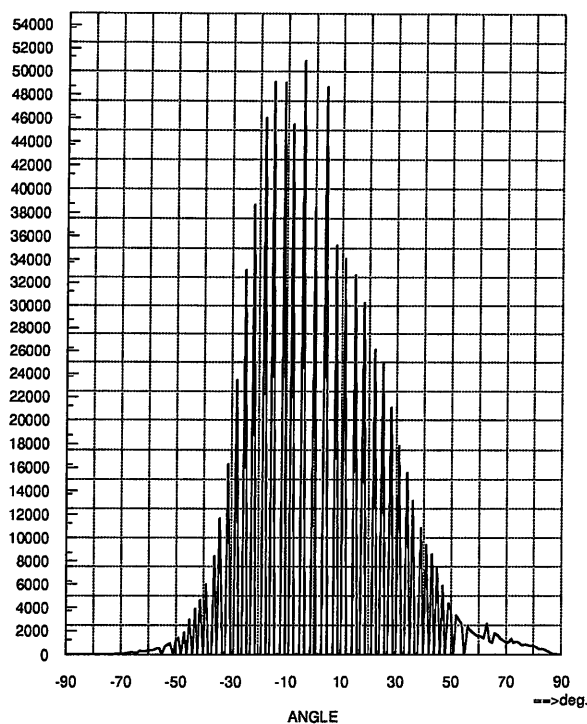
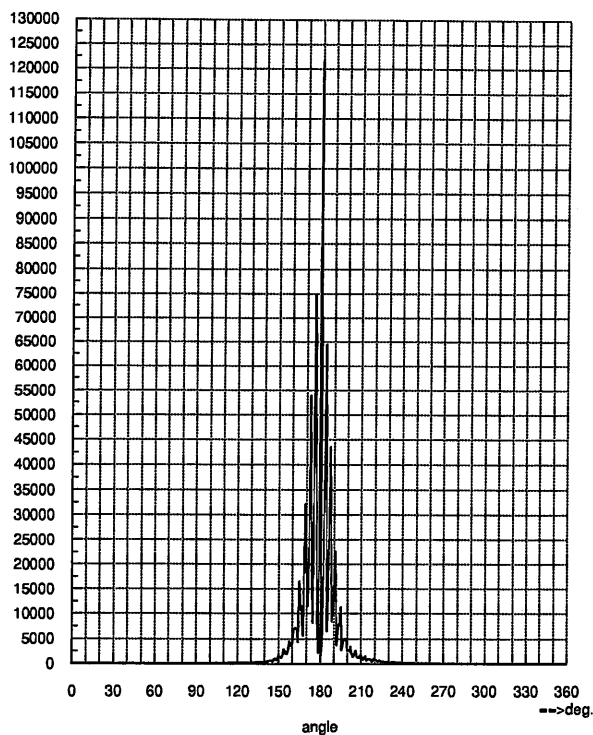


図 33: 3 点間の角度の分布 (使用した領域)

図 34: 2 点間の勾配の分布 (使用した領域)

6.3 線型拡大

さらに、3 点間の角度では 180° ばかりが多くなってしまっているので、それを改善するために、顔データの z 座標だけを数倍することにする。このことにより、 180° 近辺に集まってしまう特徴を分散させることができる。

6.4 実験の結果

6.4.1 2 点間の勾配を用いる特徴量

以上の改良をした実験の結果を示す前述の結果と同様に、100%の1位認識率を得た、最も次元数の少ないものを示す。

	3 種類	5 種類	7 種類	9 種類	11 種類	17 種類
次元数	27 次元	20 次元	35 次元	45 次元	55 次元	85 次元
領域の形態	no piled	no piled	no piled	no piled	no piled	no piled
				piled		

これより、角度の種類が少ないものには効果的であると言える。しかし、角度の種類が多いものについては逆効果であった。

6.4.2 3点間の角度を用いた特徴量

線型拡大を併用して、角度を9種類に分け、領域を2つとった18次元で1位認識率100%を得ることができた。これが3点間の角度を用いた特徴量では最低の次元数である。

ここではその18次元に着目して、線型拡大と、特徴を出す時の点の距離についての効果の結果を示す。

	1倍	5倍	10倍	15倍	20倍	40倍
5pt	80.0,80.4	91.2,90.7	91.7,91.6	91.9,90.7	93.2,91.8	
10pt	87.6,90.0	97.7,96.8	99.5,98.0	99.2,98.5	99.3,98.7	97.0,97.1
15pt	90.2,87.6	95.4,94.2	94.3,91.2	93.6,90.9	93.3,91.0	

この表で、5pt、10pt、15ptというのは、点の間隔である。また、1倍、5倍、10倍、15倍、20倍、40倍というのは、z座標を拡大した率である。また中の2つの数字は、領域の横幅を30ptから50ptまで2ptずつ変えた11回の実験で、半分ずつ重なった2つの領域をとった場合と隣接した2つの領域をとった場合のそれぞれの1位認識率の平均である。このうち10倍、10ptで、領域を半分ずつ重ねた方法で1位認識率100%が得られ、15倍、20倍のそれぞれ10ptで、領域を半分ずつ重ねる方法と、領域を隣接させる方法で本人認識率100%を得た。

この結果によると、線型拡大は15倍~20倍程度が効果が上がり、特徴を求める際の点の距離は10pt付近が良いと考えられる。今回、線型拡大は3点間の角度を用いた特徴量のみ適用したが非常に効率が上がることが分かった。

7 結論

次元数の少ないもので、1位認識率100%を達成したものを示す。ただし、3点間の角度を用いる方法で、線型拡大を用いたものは、10倍で、点の距離が10ptのものだけ使用した。

次元数	角度の種類 ×領域の数	3pt or 2pt	RR	ER	Sep.	min. sep.
18	9 × 2	3*	100.0	13.5	0.96	0.05
20	5 × 4	2*	100.0	31.2	0.85	0.02
25	5 × 5	2*	100.0	14.1	0.90	0.02
27	3 × 9	2*	100.0	14.1	1.18	0.02
27	9 × 3	3*	100.0	8.8	1.06	0.07
32	8 × 4	2	100.0	19.4	1.15	0.12
33	3 × 11	2*	100.0	11.2	1.20	0.14
34	17 × 2	3*	100.0	4.1	0.93	0.05
35	7 × 5	2	100.0	6.5	1.15	0.03
35	5 × 7	2	100.0	12.9	1.11	0.03
35	5 × 7	2*	100.0	18.2	0.96	0.04
35	7 × 5	2*	100.0	15.9	0.90	0.04
38	19 × 2	3*	100.0	3.5	0.89	0.02

*は特徴量の改良をしたもの。

今回、1位認識率を100%のまま次元数を削減することにばかり目を向けていたため、本人棄却率、分離能については、考察できない。ただし、次元数が大きければ、本人棄却率も低いものができる。しかし、0%まで落とせたものはなかった。

2点間の勾配を用いる方法については、

- 2点間の距離が大きくなると特徴が出なくなる。
- 量子化時に補正を行うとわずかに良くなる。
- 次元数が小さい場合、角度の切り方により、非常に効率が良くなる。

ことが言える。

3点間の角度を用いる方法については、

- 線型拡大をした場合、3点間の距離が大きくなると特徴が出てくる。
- 角度の切り方により、非常に効率が良くなる。

ことが言える。

・ *B*スプライン曲面の制御点を特徴量とした認識では、次元数 86 次元、1 位認識率 98.8% であり、顔の各部位の抽出を手法とした認識では、次元数 49 次元、1 位認識率 95.0% となっているため、本手法では、次元の削減、認識率の向上が同時に達成されたと言える。

8 今後の課題

データが増えた場合に本手法が有効であるか確認する必要がある。確かに大幅な次元数の削減には成功したのだが、これが、大量のデータだと、通用しない可能性もある。例えば、170枚のデータに対して、170次元の特徴量を用いるとすると、データが340枚になっても、次元数とデータの比は2倍とあまり変わらない。しかし本手法のように20次元程度とすると、はじめから確かに8.5倍だが、さらに、17倍ともなってしまう。こう見ると、もっと高い次元数でも安定した認識のできる手法を探る必要もあると思われる。

もし本手法を本実験程度のデータで扱うものとするれば、まだ、改良の余地がある。例えば、次元圧縮を行うとか、重み付きユークリッド距離やマハラノビス距離を用いるとかが考えられる。その際本人棄却率や分離能の評価もする必要がある。

また、顔画像認識においては、3次元計測の精度も重要な要素である。この精度が向上した時に本手法通用するかどうかは分からない。本手法のように現在は、誤差の吸収がうまくいっているものが、成績が良いはずだが、誤差がなくなった場合を想定する必要もあるだろう。もちろん、人間は生き物なので、計測の精度が向上しても、毎回全く等しいデータが得られることはない。しかし、人間の顔だからこそという、変動する場所なり、その変動の大きさなりがあるはずなので、それに対応する手法が必要であると考えられる。

謝辞

本研究を行うにあたり、多大な御指導、御助言、を戴いた阿曾弘具教授に心より感謝いたします。また、阿部亨博士には、本研究全般にわたり、親身な御指導を賜りました、ここに深く感謝いたします。さらに、有益な御示唆、御助言をいただいた成富敬氏に感謝の意を表します。最後に、日頃より数多くの御指導、御討論、御協力を賜りました阿曾研究室の皆様に感謝いたします。

参考文献

1. 阿部、阿曾、木村：「3次元データより得られる横顔輪郭線を用いた人間の顔の自動識別」
電子情報通信学会論文誌 D- II Vol.J73-D- II No.9 pp.1468-1476 (1990.9)
2. 阿部：「色符号化スリット光投影法による3次元形状計測」
3. 長峯、植村、増田：「3D顔画像における個人性の分析」
信学技報、PRU91-95 (1991)
4. 久保 賢：「顔画像の識別に関する研究」
東北大学工学研究科情報工学専攻 修士学位論文 (1991)