

博士学位論文

動的計画法による
音と画像の高精度認識に関する研究

2012 年

服部 公央亮

目次

第1章	はじめに	1
第2章	自己組織化マップを用いた顔方向の認識	3
2.1	まえがき	3
2.2	顔姿勢の自動生成	4
2.2.1	自己組織化マップ (SOM)	5
2.2.2	SOM を用いた中間姿勢画像の補間生成	7
2.3	補間画像を用いた姿勢推定	9
2.3.1	推定する姿勢領域のクラスタリング	9
2.3.2	中間姿勢を追加したクラスタリング	9
2.4	実験データベース	10
2.5	姿勢推定実験	11
2.5.1	5度の分解能に対する姿勢推定実験	12
2.5.2	2.5度の分解能に対する姿勢推定実験	12
2.5.3	テンプレートマッチングによる推定	13
2.6	姿勢推定に対する考察	14
2.7	ジェスチャ認識に対する検討	16
2.7.1	2自由度の姿勢推定への拡張	16
2.7.2	首振り, 頷きの判定	18
2.8	まとめ	22
第3章	機械駆動異音の自動検出	23
3.1	まえがき	23
3.2	各種法の概要	25
3.2.1	CHLAC による異音検査	25
3.2.2	連続 DP マッチングによる異音検査	26
3.3	音声分析基礎理論	26
3.3.1	フーリエ変換	26
3.3.2	離散フーリエ変換(DFT)	28
3.3.3	高速フーリエ変換(FFT)	28
3.3.4	線形予測分析	30
3.3.5	線形予測法によるスペクトル分析	33
3.4	連続 DP マッチングによる検査手法	36
3.4.1	連続 DP マッチング	37
3.5	CHLAC による検査手法	38
3.5.1	立体高次局所自己相関特徴 (CHLAC)	38
3.5.2	検査区間の選択	39
3.5.3	CHLAC 特徴の抽出	40
3.5.4	良品空間の学習	41
3.5.5	良否判定	42
3.6	連続 DP マッチングによる検査手法	44
3.6.1	良品リファレンスの登録	44
3.6.2	良否判定	45
3.7	実験データベース	46
3.7.1	データグループ 1	46

3.7.2	データグループ2	47
3.7.3	データグループ3	48
3.7.4	データグループ4	49
3.8	識別実験結果	50
3.8.1	CHLACによる良否判定結果	51
3.8.2	連続DPマッチングによる良否判定結果	52
3.8.3	パワー値による不良判定	53
3.8.4	各周波数帯のパワー値を用いた不良判定	53
3.9	ライン検査用ソフトウェア開発	54
3.10	まとめ	58
第4章	非接触指紋・指静脈認証	60
4.1	まえがき	60
4.2	非接触指紋認証システム	61
4.2.1	非接触指紋撮影システム	61
4.2.2	接触型指紋センサとの画像の特徴比較	64
4.2.3	認証アルゴリズム	65
4.2.4	特徴抽出	65
4.2.5	端点フリーDPマッチング	66
4.2.6	LEDの色と電圧の検討	69
4.2.7	切り出し処理	71
4.2.8	オフセットとシェーディングの除去	72
4.2.9	エッジ抽出	73
4.2.10	考察	75
4.3	指静脈認証システム	76
4.3.1	非接触指紋撮影システム	76
4.3.2	光源の波長選定	81
4.3.3	認証アルゴリズム	84
4.3.4	実験用データベース	89
4.3.5	静脈強調フィルタの検討	91
4.3.6	マッチング方向の検討	91
4.3.7	静脈強調フィルタの適用方向の検討	93
4.3.8	切り出しサイズの検討	94
4.3.9	考察	94
4.3.10	LED自動輝度調整システム	95
4.3.11	静脈認証システムの試作	105
4.4	まとめ	107
第5章	むすび	108
	謝辞	110
	参考文献	111
	研究業績一覧	117

第1章はじめに

センサや計算機の低価格化・高性能化に伴い、音や画像を用いたセンシングが盛んに研究・開発されている^{[1]-[8]}。特に、生産設備用の不良検査システムやセキュリティシステムについては多数実用化されている。しかしながら、近年、そのシステム規模、精度、価格、速度、自動化などの面で、要求されるスペックがさらに飛躍しようとしている。しかし、センサやコンピュータの性能が著しく向上しているのに比べて、処理アルゴリズムの発達は遅れており、センシングシステム全体としては未だに多くの課題を残している。

画像を用いたセンシングの中で、顔のセンシングは利用時の抵抗感の少なさ、インターフェイスの簡便さ等の特徴を有しており、様々な分野での応用が期待されている^{[8]-[10]}。特にロボットビジョンなどでの応用を考えた場合、人の視線方向に対応する顔の向きという情報は、自然なインタラクションを返す際に非常に有用な情報として使用できる。

同様に、画像センシングの進歩に伴い、近年の生産現場では多くの画像センサが導入されている。特に検査員による目視検査は、画像センサを用いた自動不良検査システムに置き換えられ、省力化・検査の高精度化が図られている。しかし、音の検査については画像に比べて利用例が少なく、多くの現場では今なお検査員の聴覚による官能検査が主流である。目視検査とは異なり、聴覚による不良検査は検査指標を定量化しにくく、さらに聴覚の個人差が大きいことから安定した検査が難しい。こうした背景のもと、自動検査システムによる高精度かつ安定した検査が望まれている。

さらに、バイオメトリクスシステムでは、指紋センサの場合、接触型センサから非接触型センサへの需要が高まりつつあり、一方、静脈センサの場合、小型化、低価格化、および高精度化が期待されている。

本研究では、こうした問題に対して動的計画法を適用することで、高精度な認識、識別を行うシステムを開発した。動的計画法は、1962年にリチャード・E・ベルマンにより発表されたアルゴリズム^[11]である。一般に2パターンをマッチングする際には、ズレや歪みの影響によりマッチング精度が低下する。これをマッチングするために、しばしばパターンをゴムのように非線形に伸縮さ

せマッチングさせることがある。このような技術は、音声やジェスチャなどの 1 次元パターンその他、2 次元パターンである顔・文字などのマッチングにも必要不可欠である。動的計画法を用いたマッチング (DP マッチング) は、弾性マッチングの手法として 1970 年代に開発されたアルゴリズムであり、アルゴリズム実装の容易さや非常に少ない計算量などの優れた性質を持つ。そのため、マッチングのほか、領域のセグメンテーションや、リファレンスとテストパターンの位置合わせや対応点探索など、音声・画像認識の分野で広く使用されている^{[12]- [15]}。

本論文では、動的計画法を応用した音と画像の高精度認識システムについて述べる。第 2 章では自己組織化マップを用いた顔方向のセンシング手法について述べる。隣接画像間を繋ぐユニットを動的計画法により選択することで、リファレンスに使用する画像を選択した。次に、第 3 章では動的計画法と CHLAC 特徴による音の不良検査手法について述べる。DP マッチングを機械駆動音に対して適用することで、良品パターンからの距離算出と生産ラインにおいて検査に使用するリファレンスパターンの自動切り出しシステムを開発した。最後に指紋・静脈によるバイオメトリクスについて第 4 章で述べる。ここでは、撮影システムと認証アルゴリズムについてそれぞれ検討した。指紋や静脈を照合する際には、位置ずれや歪みが発生する。そのため、DP マッチングを取り入れた認証アルゴリズムを検討した。

第2章 自己組織化マップを用いた 顔方向の認識

2.1 まえがき

顔は人間のコミュニケーションにおいて重要な役割を演じていることは、多くの心理学的研究を通じて確認されている^[16]。顔画像においても人物の個人性を示す情報のほか、感情、意図、関心などの心理状態や行動に関する情報が含まれると考えられ、盛んに研究されてきた^[7]。このような情報をコンピュータにより読み取ることができるようになれば、医療や福祉、セキュリティ、マンマシンインターフェイスなどのさまざまな分野において応用が可能であると考えられている。また、顔画像はデータ取得に非接触なビデオカメラを用いるため、利用者に与える不快感が少なく実生活において自然に利用できるという点からも、高い期待を受けている。中でも顔方向は人の視線に対応した情報であり、顔画像を用いて姿勢を推定する事で、人が注視している方向の推定が可能となる。そのため、顔画像を用いた姿勢推定は人物の注意や興味を知るための手法として注目され、研究が行われている。実用化されれば、例えば脇見運転の検知などが可能となると期待される。ドライバの運転支援を目的としたシステムは、交通事故件数増加の抑止を目的に研究・開発が盛んに進められている^[17]。

姿勢推定の手法としては、2次元画像、もしくは3次元情報を用いた手法が主に行われている。前者には顔の特徴点、すなわち目、鼻、口などの位置を検出して推定を行うもの、濃淡画像そのものを用いて推定を行うものなどがある^{[18]- [20]}。

一方、後者の3次元情報を用いた手法としては、3次元情報から特徴点を検出する方法や、距離画像を用いる方法が挙げられる^{[21]- [23]}。3次元情報を用いた手法は高精度だが、データの取得にコストがかかるという問題がある。一方、2次元画像のみによる手法では、3次元データ取得のための処理や装置が必要

無く、システムが簡易なもので済むため、実用面において優位性が高い。しかし、2次元画像から高精度に特徴点を抽出して、高精度な姿勢推定をすることは難しい。また、濃淡画像をそのまま用いたシステムは、3次元情報を用いたものに比べて推定精度が低い。その理由の一つとして、人間の姿勢は連続的なものであり、全ての姿勢の2次元画像を撮影して登録することが困難であるということが考えられる。

姿勢変動のような変化が連続的であることを利用して、2次元画像を低次元のベクトルに圧縮した上で、その固有空間上で照合や姿勢推定を行う手法として、パラメトリック固有空間法が提案されている。しかし、主成分分析などにより得られた空間は、画像集合の次元圧縮という意味では最適であるが、パラメータ推定を行う上で最適な空間とは言えないことが知られており、パラメータ推定のための様々な空間生成法が検討されている^{[24][25]}。

本章では、入力画像に対する次元圧縮を行わずに、連続的な動きをニューラルネットワークの一種である自己組織化マップ(SOM)により補間することで、少ない学習データから高精度に姿勢推定を行う手法を提案する。これにより、2次元画像を用いながらも従来研究に比べて分解能の高い推定を行う。

2.2 顔姿勢の自動生成

自己組織化マップ(Self-Organizing Maps, SOM)は、コホーネン型ニューラルネットワークとも呼ばれる人口神経回路網の一種である。多次元のデータを2次元に写像することが可能であり、高次元空間の可視化に用いることが可能である。各分野で取り扱われるデータは多次元であることが多く、また、データごとの特徴などの予備知識がない中で、その分類や特徴を抽出する際や、音声などのパターン認識を行う際に利用される手法の1つとしてSOMは知られている^{[26]-[31]}。しかしながら、提案法ではSOMを画像の補間を行うニューラルネットワークとして扱う^{[32][33]}。これにより、少ない学習画像から、その画像間の遷移状態を生成することができる。SOMは、マップ上の情報を可視化することに秀でているため、特徴ベクトルとして視認しやすい情報を入力することで、その特性を把握しやすい。そのため、本研究では実験の簡略化とマップの視認性の確保を目的に、輝度情報を特徴ベクトルに使用する。照明変動や不特定者へ対応するためには、前処理としてPCAなどの検討が必要である。

2.2.1 自己組織化マップ (SOM)

SOM は、ヘルシンキ大学情報科学教授のコホーネンによって開発された大脳皮質の神経機能をモデル化したニューラルネットワークである^{[29] [30]}。また、その学習アルゴリズムから競合学習型ニューラルネットワークとして知られている。ここで、競合学習型とは、ニューロンの学習段階で、任意の入力に対して最大の応答を示したニューロンについてのみ、重みの修正(学習)を行うものである。この学習により、データの統計的分布や相互の類似関係を反映したテンプレートが形成されるという特徴がある。SOM は多次元のベクトルにより表されたデータを、その特徴を残し他のデータとの相互関係を保ったまま、2次元のマップに写像することが出来る。これにより、多次元のデータの関係が2次元平面上の距離として表され、視覚的に理解しやすいという特徴を持つ。この性質を利用して、テクスチャ画像の濃淡値や距離画像の値をそのまま特徴ベクトルとして入力することにより、入力画像の中間に位置する画像をマップ上で補間生成することができる^[33]。また、SOM により補間された画像を利用することで、表情を高精度に認証することができるという先行研究も行われている^[34]。提案法では、SOM の補間画像作成を利用して隣接姿勢間の画像を補間することにより、姿勢推定を行う際に使用するリファレンスデータを作成する。本稿で扱う SOM の構成を図 2-1 に示す。

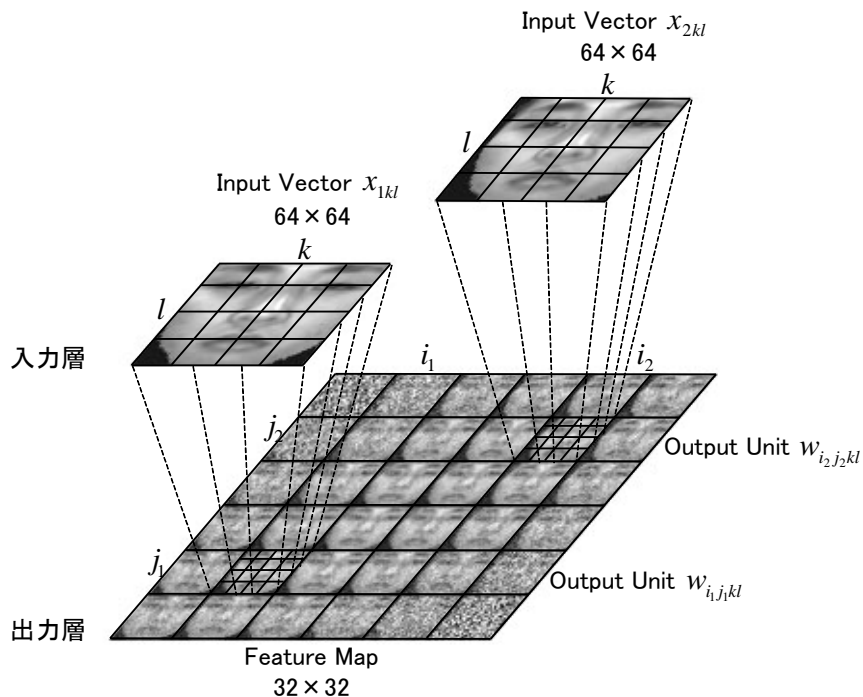


図 2-1 自己組織化マップの構成

ここで、入力ベクトル(Input Vector)を 64×64 [pixel]の顔姿勢画像とし、出力層 32×32 個の出力ユニット(Output Unit)で構成される。入力顔画像には9枚の顔姿勢画像(64×64 [pixel])を用いる。

学習過程について説明する。まず、各出力ユニット w_{ij} を乱数で初期化する。次に、各入力ベクトル x と出力ユニット w_{ij} との類似度 u_{ij} を(2-1)式により算出する。

$$u_{ij} = \frac{\sum x_{kl} w_{ijkl}}{\sqrt{\sum x_{kl}^2 \sum w_{ijkl}^2}} \quad (2-1)$$

類似度が最大になる出力ユニットを中心に、半径 d 以内の出力ユニットを(2-2)式により、入力ベクトルに近づけるように更新する。

$$w'_{ijkl} = w_{ijkl} + \alpha(t)(x_{kl} - w_{ijkl}) \quad (2-2)$$

ここで、 $\alpha(t)$ は学習係数であり、(2-3)式により、学習が進むにつれて減少する単調減少関数で更新される。

$$\alpha(t+1) = \alpha(t) \left(1 - \frac{t}{T}\right) \quad (2-3)$$

T は学習回数、 t は現在の学習回数を示す。また、半径 d を時間と共に狭くするほうがパターンの分類能力は向上する。半径 d は(2-4)式により更新される。

$$d(t+1) = d(t) \left(1 - \frac{t}{T}\right) \quad \{d(0) = d\} \quad (2-4)$$

通常の SOM では、学習が進むにつれて学習半径を小さくするが、提案法では重み更新時に(2-5)式を用いる。

$$f(X) = \cos\left(\frac{\pi X}{2d_s}\right) \quad (2-5)$$

ここで、 d_s は学習時に設定される半径、 X はマップ上における類似度最大ユニットから更新されるユニットまでの距離である。類似度最大ユニットからの距離が離れるに従い重み更新量を減少させることで、なだらかな変化が生じることが先行研究により確認されている^[34]。重みの更新は(2-6)式を用いて行う。

$$w'_{ijkl} = w_{ijkl} + f(d_i)f(d_j)\alpha(t)(x_{kl} - w_{ijkl}) \quad (2-6)$$

ここで、 d_i と d_j は最適整合ユニットから更新する重みまでの i 方向と j 方向の距離($d_s > d_i$, $d_s > d_j$)である。これにより、出力ユニットの更新量は類似度最大のユニットからの距離が離れるにつれて小さくなる。以上の操作を各学習画像に対して行い、設定した学習回数に至るまで繰り返す。学習時のマップの状態を示す。図 2-2 に学習における出力ユニットの変化の様子を示す。

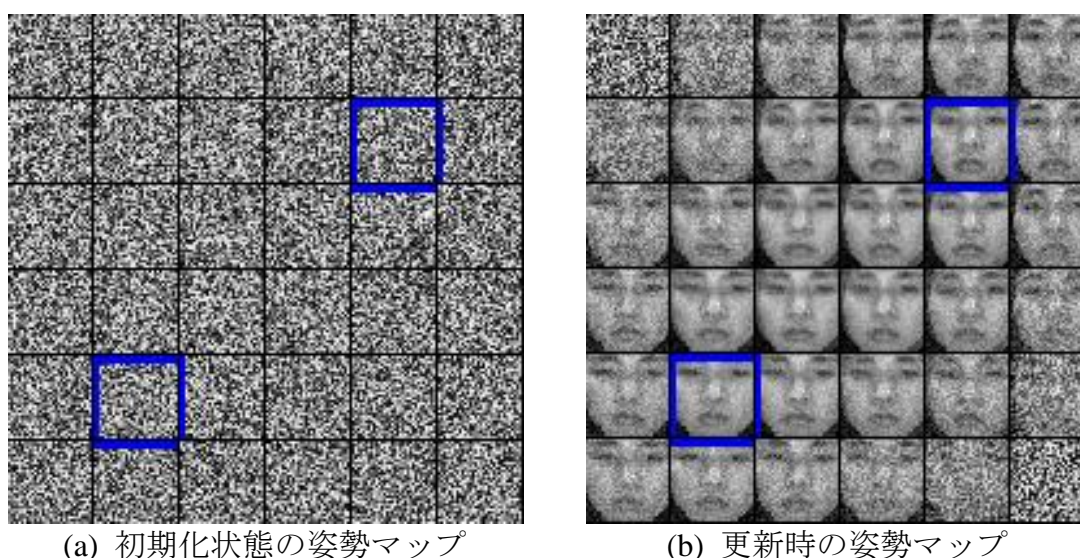


図 2-2 自己組織化マップ学習の様子

2.2.2 SOM を用いた中間姿勢画像の補間生成

SOM を用いて、学習データから姿勢推移の過渡状態の画像を補間する。各隣接姿勢間の画像を補間することで、姿勢が微小に変化した画像を補間する。補間された画像を用いて姿勢推定を行うことで、高精度な姿勢推定の実現を図る。

まず、SOM に 5 度刻みの仮想生成画像を入力してマップの作成を行う。学習画像には 5 度刻みの仮想生成画像 1 セットを用いる。マップサイズに対して学

学習半径 を小さくした場合，学習時の重み更新の幅が狭くなりマップ全体を使い切らず，マップ上の白色雑音（初期化信号）の領域が更新されずに学習が終了してしまうことがある．逆に大きくしすぎると，各入力ベクトルのセントロイド位置に他のクラスが大きな影響をおよぼすため，滑らかにクラス間が補間されない．本実験では，先行研究^[34]で有効性が確認されているパラメータを使用して学習を行う．具体的には，入力ベクトルのサイズは 64×64 ，出力ユニット数は 32×32 ，学習回数を 300 回，学習半径 を 10，学習係数 を 0.05 とした．予備実験の結果，本パラメータによりマップ全体を使い切り，滑らかに中間画像が補間されていることを確認した．

次に，学習により作成したマップ上において，隣接する姿勢 A と姿勢 B の学習画像（例えば 0 度と 5 度）に対する類似度最大位置をそれぞれ始端・終端として，その間の顔画像の変化量が最小となる経路（姿勢推移画像列）を求める．このとき類似度にはユークリッド距離を用いた．経路は図 2-3 に示すように DP マッチングにより求め，使用する DP パスは(2-7)式に示す対称型を用いる． -15 度 ~ 15 度の 5 度刻みの離散的な角度のデータに適用し，各姿勢間の推移が補間された画像列を取得する．この DP マッチングによる補間画像生成は，先行研究^[33]において 3 次元顔表情の滑らかな変化を実現しており，本タスクにおいても有効性が期待できると考えた．

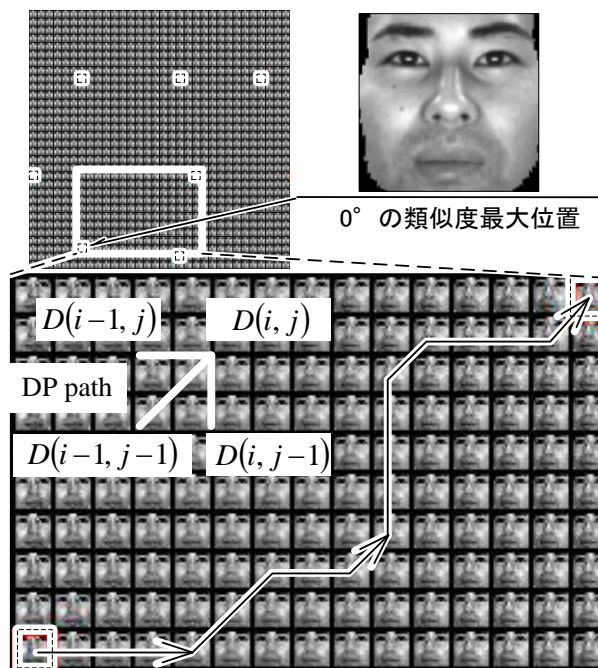


図 2-3 SOM 上での DP マッチング

$$D(i, j) = \max \begin{cases} D(i-1, j) + d_1 \\ D(i-1, j-1) + d_2 \\ D(i, j-1) + d_3 \end{cases} \quad (2-7)$$

d_1 : (i, j) と $(i-1, j)$ の画像の相関値 c

d_2 : (i, j) と $(i-1, j-1)$ の画像の相関値 c

d_3 : (i, j) と $(i, j-1)$ の画像の相関値 c

2.3 補間画像を用いた姿勢推定

作成した姿勢推移画像列を角度（離散値）でクラスタリングして、姿勢推定に用いる。2.3.1 項では、入力画像と姿勢推移画像列中の画像との類似度を求め、最も高い類似度を示した画像の属する角度を姿勢の結果として出力する。2.3.2 項では、学習した姿勢画像間の中間姿勢の領域をクラスタリングにより作成することで、未学習の中間姿勢を推定する。

2.3.1 推定する姿勢領域のクラスタリング

姿勢を学習画像と同じ5度刻みの分解能で推定するために、姿勢推移画像列の始端と終端をセントロイドベクトルとしてクラスタリングを行い、5度ずつの領域に分割する。クラスタリングの一例を図 2-4 に示す。

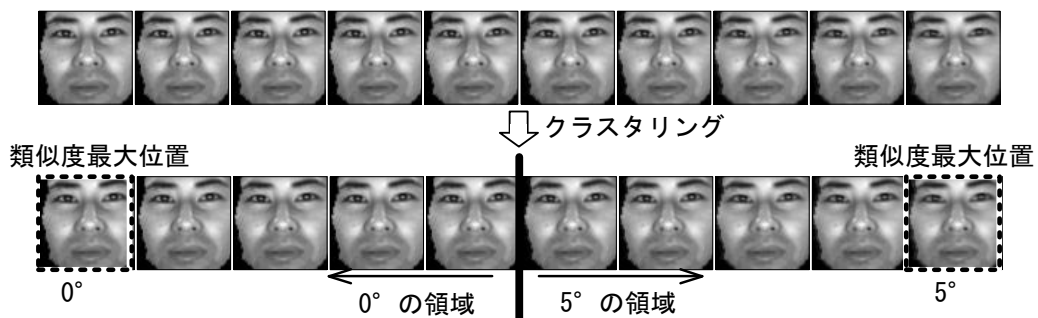


図 2-4 領域分割方法

2.3.2 中間姿勢を追加したクラスタリング

さらに、学習した姿勢画像間の中間姿勢の領域をクラスタリングにより作成することで、未学習の中間姿勢を推定する手法を検討した。中間姿勢のクラス

を生成するため、マップ上に中間姿勢のセントロイドベクトルを設定した。中間姿勢のセントロイドベクトル決定法として、以下の2つの手法を比較した。

手法1 : 始端と終端の中間にある画像を中間姿勢のセントロイドとする (図 2-5)。

手法2 : 始端と終端の画像に対して類似度の最大値が最小となる画像を中間姿勢のセントロイドとする (図 2-6)。

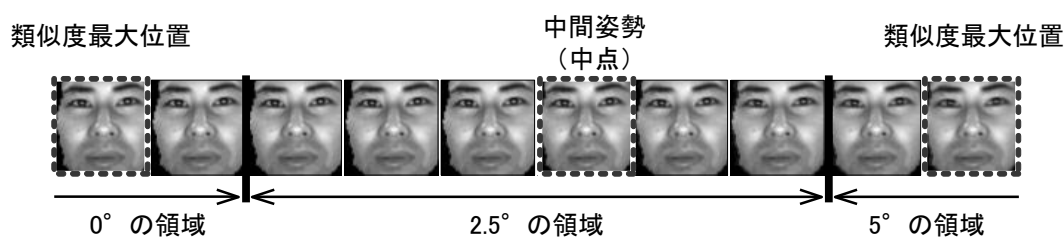


図 2-5 中間姿勢を含む領域分割方法 (手法 1)

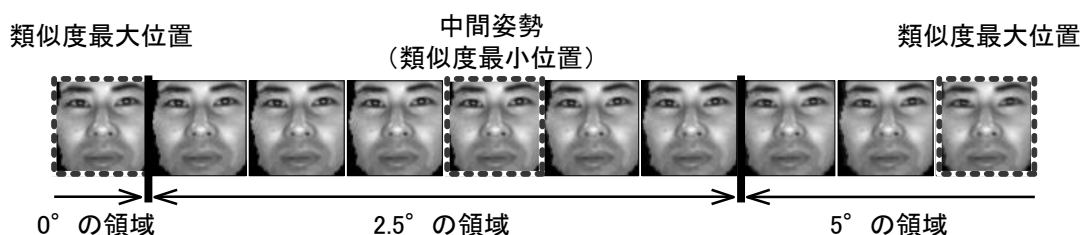


図 2-6 中間姿勢を含む領域分割方法 (手法 2)

2.4 実験データベース

姿勢の角度精度が保証されたデータを用いて実験を行うため、正面から計測された顔の3次元形状データを使用する。3次元形状は、株式会社スペースビジョン社製のレーザレンジファインダ Cartesia を使用して取得した。3次元形状と同時に、顔前面より拡散照明を照射してテクスチャを撮影する。

実験に使用する2次元画像は、3次元形状データを回転させ、テクスチャをマッピングすることにより生成する。回転中心には、3次元形状データより抽出が容易な鼻の頂点を使用した。本章の実験では、正面画像を0度の姿勢として、顔の垂直軸まわりに-15度から15度まで回転させ2次元に投影した顔画像

を扱う。実験で使用する画像（以降、仮想生成画像とする）は、生成した画像から鼻の頂点を中心に一定領域を切り出した後に 64×64 [pixel] のサイズに縮小したものである。このデータを 10 名（男性 8 名，女性 2 名）に対して 3 回ずつ撮影・生成して実験に使用した。このとき 1 回分の撮影・生成で得られる仮想生成画像を 1 セットとして実験に用いる。3 次元データから生成した姿勢画像の一例として、2 名分、各 1 セットを図 2-7 に示す。提案法の評価実験は、5 度刻みの仮想生成画像を教師データとして学習に使用して、2.5 度刻みの仮想生成画像をテストデータとして行う。

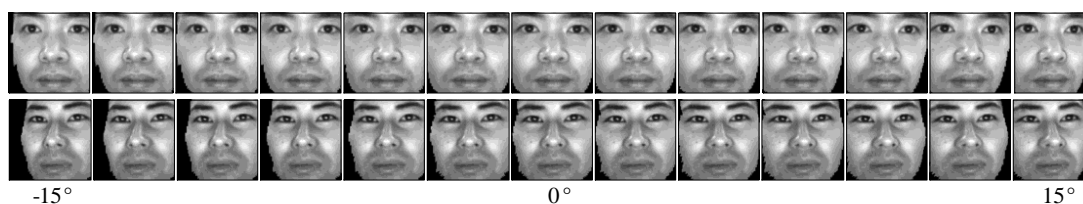


図 2-7 3 次元画像から生成した画像例

2.5 姿勢推定実験

まず SOM により補間された画像を用いて、5 度の分解能における性能について実験した。次に、5 度刻みの仮想生成画像から提案法により中間姿勢のリファレンス画像を補間生成することで、2.5 度の分解能での姿勢推定を試みる。提案法に対する比較として、2.5 度刻みの仮想生成画像をテンプレートとしたテンプレートマッチングによる姿勢推定実験を行い、提案法の有効性について検証する。テンプレートマッチングによる姿勢推定では、SOM と同様に相互相関係数(2-1)式によりテンプレートからの距離を算出して、最も類似度の高い姿勢の角度を推定結果とした。

実験を行う際には、1 名につき 3 セットのデータを使用する。交差確認法により精度評価を行うこととした。まず、1 セット分の仮想生成画像を学習もしくはテンプレートに使用して、残りの仮想生成画像の姿勢を推定する。その後、使用したデータセットを入れ替えて実験を繰り返す。これを被験者 10 名に対して、それぞれ行う。

2.5.1 5度の分解能に対する姿勢推定実験

2.3.1 項のクラスタリングを行い、5度刻みの仮想生成画像に対する姿勢推定実験をした。1セットを学習に、2セットを姿勢推定対象として使用して、交差確認法によりテストデータを入れ替えて実験をした。その結果、平均推定精度 87 [%]を得た（表 2-1）。このときの平均二乗平方根誤差（RMSE）として、1.78 度を得た。

表 2-1 提案法による推定結果

出力 入力	-15.0°	-10.0°	-5.0°	0.0°	5.0°	10.0°	15.0°	精度 [%]
-15.0°	59	1						98
-10.0°	1	54	5					90
-5.0°		5	49	6				82
0.0°			4	48	8			80
5.0°				6	49	5		82
10.0°					5	53	2	88
15.0°						5	55	92

2.5.2 2.5度の分解能に対する姿勢推定実験

2.3.2 項のクラスタリングを行い、2.5度刻みの仮想生成画像を対象として分解能 2.5度で姿勢推定を試みた。手法 1 の結果を表 2-2 に、手法 2 の結果を表 2-3 に示す。手法 1 の平均推定精度は 57 [%]となり、学習画像と同角度の姿勢についての平均推定精度 70 [%]、学習画像に無い角度についての平均推定精度は 43 [%]という結果が得られた。一方、手法 2 を用いた場合、平均推定精度が 58 [%]となり、学習画像と同姿勢の角度では平均推定精度 69 [%]、学習データにない角度については平均推定精度 46 [%]であった。両手法共の RMSE はそれぞれ 1.91 度、1.87 度であった。手法 1、手法 2 を比較すると、特に補間した中間姿勢について手法 2 が手法 1 に比べて高い認識結果となることが確認された。また、手法 2 では 95 [%]以上のデータにおいて推定誤りが±2.5 度以内で推定できていることが確認された。

表 2-2 提案法による推定結果（手法 1）

出力 入力		-							+							精度 [%]
		15.0°	12.5°	10.0°	7.5°	5.0°	2.5°	0.0°	2.5°	5.0°	7.5°	10.0°	12.5°	15.0°		
-	15.0°	53	7												88	
	12.5°	8	35	17											58	
	10.0°		4	47	8	1									78	
	7.5°			17	28	13	1	1							47	
	5.0°			3	7	40	7	3							67	
	2.5°				2	13	29	14	2						48	
	0.0°					3	5	39	10	3					65	
+	2.5°						1	21	19	17	2				32	
	5.0°							3	13	29	12	3			48	
	7.5°								2	19	19	19	1		32	
	10.0°									3	10	40	5	2	67	
	12.5°										1	22	24	13	40	
	15.0°											2	12	46	77	

表 2-3 提案法による推定結果（手法 2）

出力 入力		-							+							精度 [%]
		15.0°	12.5°	10.0°	7.5°	5.0°	2.5°	0.0°	2.5°	5.0°	7.5°	10.0°	12.5°	15.0°		
-	15.0°	52	8												87	
	12.5°	9	33	18											55	
	10.0°		4	45	10	1									75	
	7.5°			16	28	14	1	1							47	
	5.0°			2	8	38	9	3							63	
	2.5°				2	10	32	14	2						53	
	0.0°					3	5	36	14	2					60	
+	2.5°						1	19	23	15	2				38	
	5.0°							2	13	29	13	3			48	
	7.5°								2	17	21	19	1		35	
	10.0°									2	11	39	6	2	65	
	12.5°										1	19	27	13	45	
	15.0°											1	10	49	82	

2.5.3 テンプレートマッチングによる推定

比較実験としてテンプレートマッチングによる姿勢推定をした(表 2-4). 平均推定精度は 58 [%], RMSE は 1.88 度であった. この実験により, 提案法により未学習の中間姿勢を含めて推定した場合においても, 予め 2.5 度のテンプレートを持つ場合と同程度の推定精度であることが確認された.

表 2-4 テンプレートマッチングによる推定結果

出力 入力		-							+							精度 [%]
		15.0°	12.5°	10.0°	7.5°	5.0°	2.5°	0.0°	2.5°	5.0°	7.5°	10.0°	12.5°	15.0°		
-	15.0°	51	9												85	
	12.5°	6	44	10											73	
	10.0°		9	41	9	1									68	
	7.5°			11	37	10	1	1							62	
	5.0°			1	12	34	10	3							57	
	2.5°				3	12	31	11	3						52	
	0.0°					3	10	30	15	2					50	
+	2.5°						2	14	25	17	2				42	
	5.0°							2	17	23	15	3			38	
	7.5°								2	17	24	15	2		40	
	10.0°									3	14	30	12	1	50	
	12.5°										1	12	36	11	60	
	15.0°												15	45	75	

2.6 姿勢推定に対する考察

提案法を用いて中間姿勢画像を補間して姿勢推定を行うことで、角度分解能向上を実現できることが確認された。手法 1, 2 で推定を誤るデータについて確認すると、多くの場合は比較対象としたテンプレートマッチング法と同じデータの推定に対して推定誤りが生じることが確認された。さらに、提案法における推定誤りは、クラスタリングの境界付近の画像との相関が高い場合に多く発生していることが確認された。どの手法においてもテストデータの姿勢角度が-7.5 度の際に推定角度 0 度という大きな推定誤りをするデータが存在するが、このデータは全て同一のデータであることが確認された。このときのテスト及びリファレンスデータの画像（圧縮前）を図 2-8 に示す。

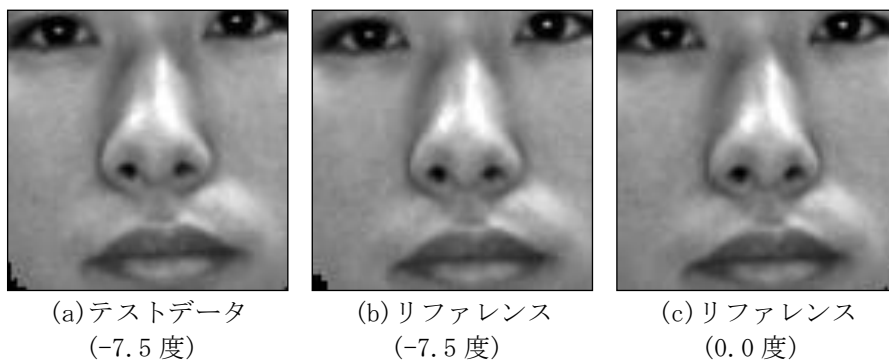


図 2-8 推定誤り例

このデータは、リファレンスの顔全体が若干右にずれているため、正対した場合との類似度が高くなり 0 度と推定された。これは、位置合わせに使用している鼻の頂点抽出のずれにより発生しており、抽出精度を向上することで改善されると考えられる。

中間姿勢の領域決定方法として使用した 2 つのクラスタリング方法を比較すると、手法 2 で高い推定精度を示した。単純に中央を取るのに比べ、類似度を使用した方が高い精度であることから、姿勢推移画像列中において座標と姿勢角度の推移は必ずしも線形の関係ではないと考えられる。そこで関係を確認するため、0.5 度刻みの仮想生成画像をテストデータ、同じ 3 次元データから生成した姿勢推移画像列をリファレンスとしてマッチングした。その結果の例（男女各 1 名分）を図 2-9 に示す。縦軸が 0.5 度刻みの仮想生成画像の姿勢角度、横軸が姿勢推移画像列中の座標である。この結果から、姿勢角度と画像列の座標との間には、線形ではないものの一定の相関があることが確認された。このことより、今後クラスタリングやマッチング方法を改良することで、学習姿勢の中央に位置する姿勢以外もクラスとして、より分解能の高い姿勢推定が可能であると考えられる。

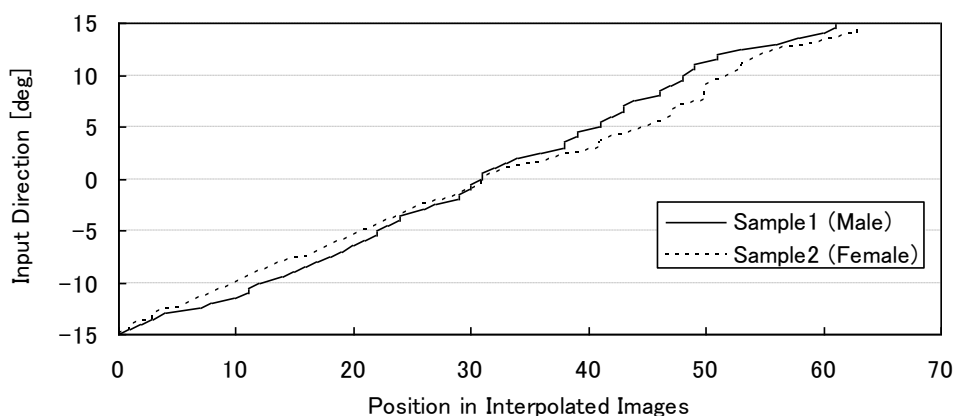


図 2-9 姿勢角度と姿勢推移画像列座標との関係

推定精度でなく RMSE で比較した場合にも、手法 2 の誤差が小さくなることが確認された。手法 2 と比較法における RMSE をテストデータの姿勢角度別に算出した結果を図 2-10 に示す。提案法の結果は学習に使用した姿勢を白丸、未学習の中間姿勢を黒丸としてプロットする。バツ印で示した 2.5 度刻みの姿勢をテンプレートとして持つ場合と比較すると、誤差の範囲及び傾向はほぼ同

じであることが確認できる。このとき、提案法の未学習姿勢と学習姿勢に対する RMSE の平均はそれぞれ 2.04 度と 1.70 度であり、全体の平均は 1.87 度である。一方、全ての姿勢をテンプレートに含むテンプレートマッチング法の RMSE は 1.88 度であることから、提案法の未学習姿勢を含めた姿勢推定は、補間を行わずに全姿勢をテンプレートとして使用した場合と比較しても遜色ない精度であるといえる。これにより、提案法による中間姿勢推定の有効性が確認された。今後、補間された画像について詳細に検証して、精度向上を目指したい。

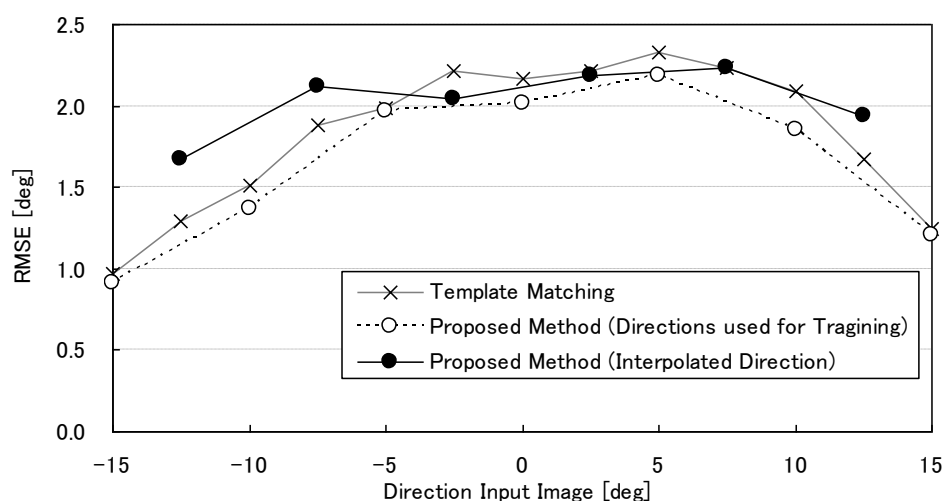


図 2-10 提案法と比較法における RMSE 比較

2.7 ジェスチャ認識に対する検討

さらに、顔の姿勢を連続的に捉えることで、頷きや首ふりなどのジェスチャの認識が可能ではないかと考えた。このようなユーザの意思を伝える情報をコンピュータにより読み取ることができるようになれば、ロボットなどの対話システムで利用できる。

2.7.1 2 自由度の姿勢推定への拡張

首振り・頷きを認識するために、SOM に 2 自由度の姿勢変動を学習した。2 自由度の姿勢変動について検討を行うため、前述した 1 自由度のときと同様に、3 次元データから 2 自由度の姿勢変動を含む画像を生成した (図 2-11)。

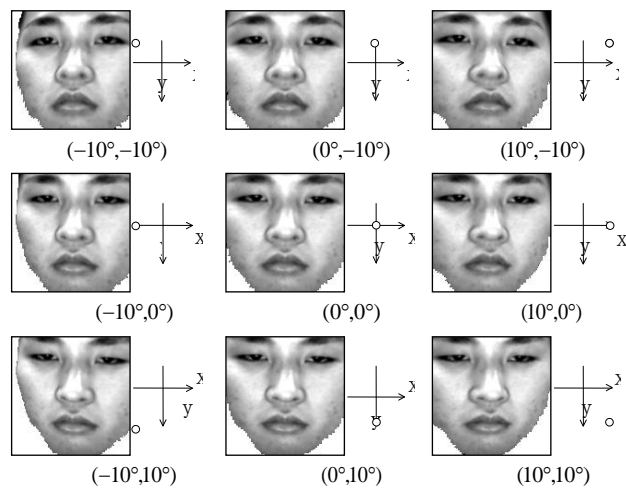


図 2-11 学習姿勢画像の例

SOM に画像を学習させる際には、マップの縦軸を頷き方向の動きに、横軸を首振り方向の動きにそれぞれ対応させるため、マップ上のセントロイド位置を固定して学習した。中間姿勢についても、学習姿勢の中央位置に中間姿勢のセントロイド配置することで、推定可能であることを確認している (図 2-12)。この 2 自由度の姿勢変動を学習した SOM に対して、入力画像の類似度最大位置がどう動くかを判定することで、首振り・頷きなどのジェスチャ認識が可能ではないかと考えた。

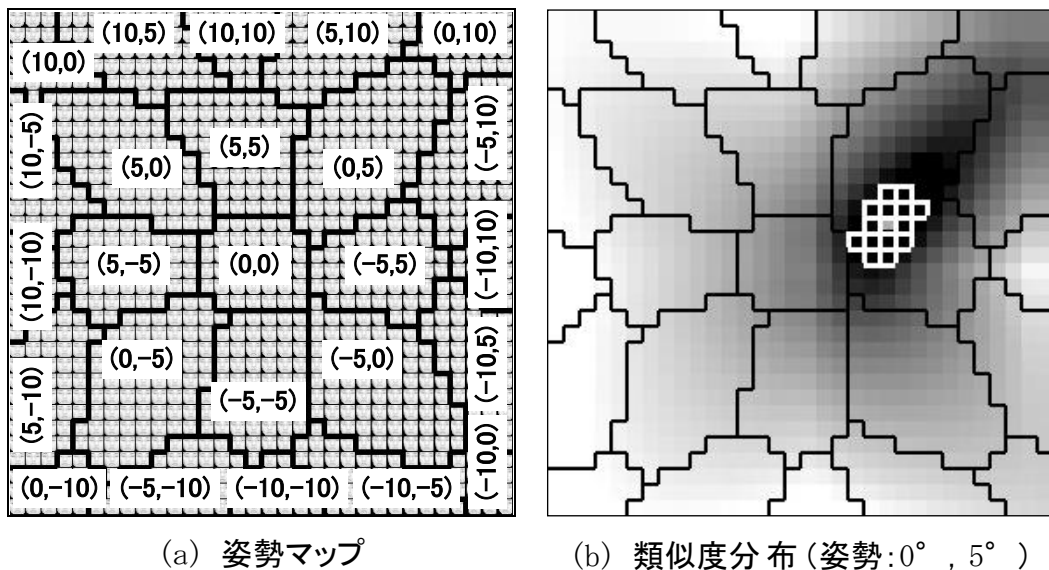


図 2-12 2 自由度の姿勢認識の例

(黒 : 類似度高. 白 : 類似度低)

2.7.2 首振り， 頷きの判定

2 自由度の姿勢変動を学習した SOM を用いて，首振り， 頷きを認識する。

まず， 予め SOM 上に 2 自由度の姿勢変動を学習しておく． 認識時には， 学習済みの SOM 上の各画像と入力画像とを比較することで， 類似度最大位置を算出する． ジェスチャ認識を行う際には， 類似度最大位置の座標変動を時系列で処理する． 図 2-13 に頷きと首振りの各動作時の座標変動例を示す． 各ジェスチャの動作に合わせて類似度最大位置が移動していることが分かる． この変移の縦横それぞれの大きさを比較することで， 動作が「首振り」「頷き」の判定を行うことができる．

変移の大きさの比較は， 座標変動の分散値を基準に行うこととした． まず， 単位時間当たりの座標変動の分散値をマップの縦横それぞれに対して算出する． 次に求めた分散値の一定時間内での平均値を求める． これにより， 時間方向に平滑化を行う． 最後に算出した平均分散値を比較して閾値処理を行うことで， 縦横どちらの座標変動が大きいかを判定して， 首振りか頷きかを認識する．

今回は， 過去 10 フレーム分の座標から分散値を求め， 30 フレーム分の平均を識別に使用した． この値は， カメラのフレームレートや， 使用するコンピュータの処理速度に依存する． 今回は， 実験には Core2Duo 2.26 [GHz] のコンピュータを使用して， 15 [fps] の Web カメラから入力された動画像に対して処理を適用した．

図 2-14 と図 2-15 では， i 方向の分散値が j 方向よりも高い値を示しており， 頷きと判定可能である． これより， 微小な提案法により微小な頷きの認識が可能なが確認された． 同様に， 図 2-16 と図 2-17 は首振りを入力した結果であるが， 双方ともに j 方向の分散が大きな値を示しており， 首振りと判定された． これらの結果より， SOM 上の座標変動を使用した首振り・頷きの認識が可能なが確認された． 大きな頷きを入力した際の結果である． 図 2-15 では， 細かな頷きを入力した図 2-14 に比べて i 方向の分散値がさらに大きな値を示していることが確認できる． 同様に図 2-16 と図 2-17 の分散値のグラフを比較すると， 微小な首振りに比べて大きな首振りを入力した際に大きな値を示している． このことから， 提案法により算出された分散値の大きさから， 首振り・頷きの大きさの評価が可能であると考えられる． 今後， 被験者数を増やして定量的に評価を行いたい．

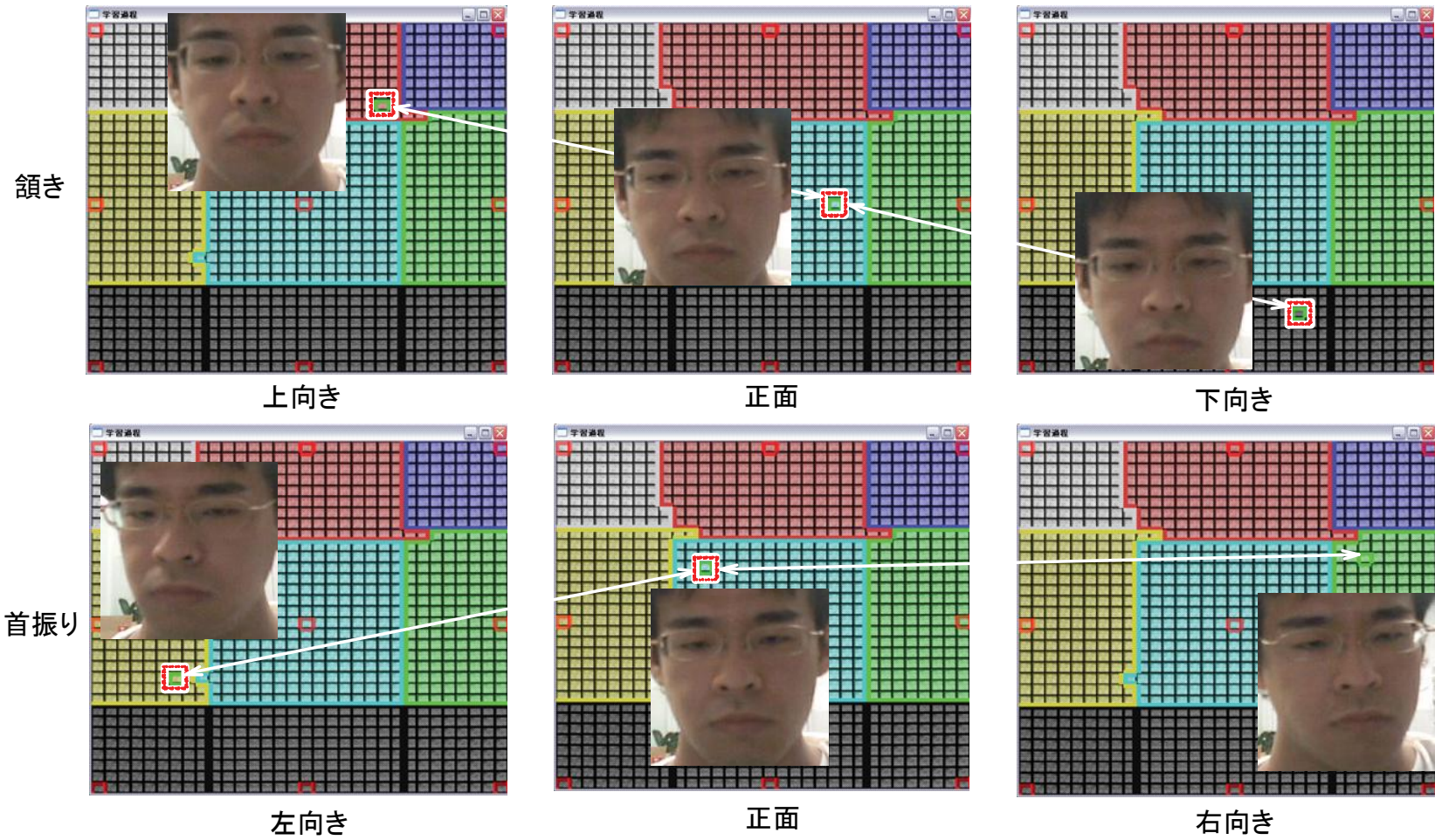


図 2-13 類似度最大位置の移動例

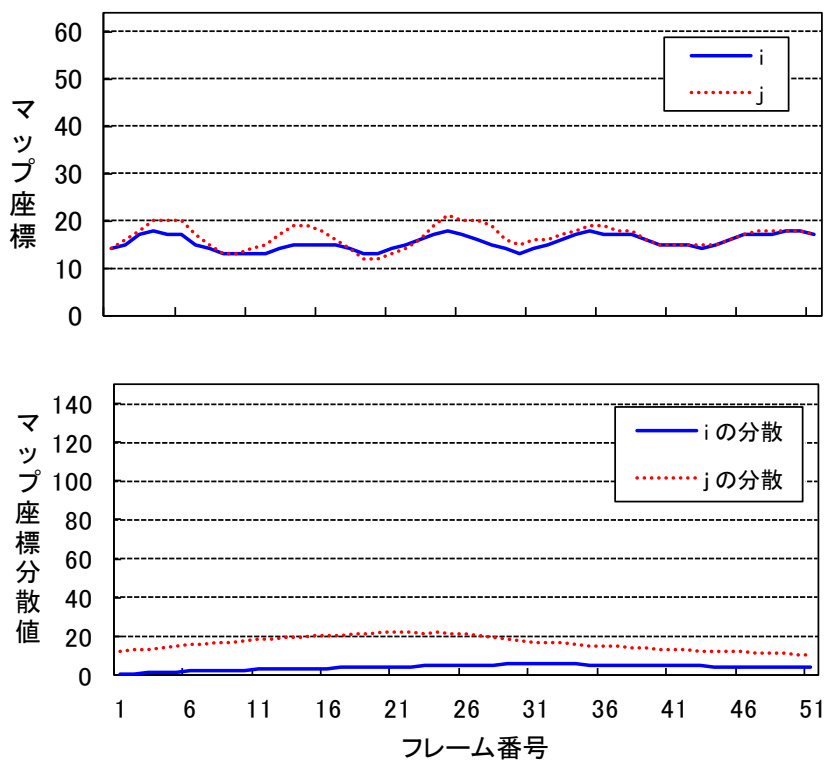


図 2-14 微小揺き時の変化

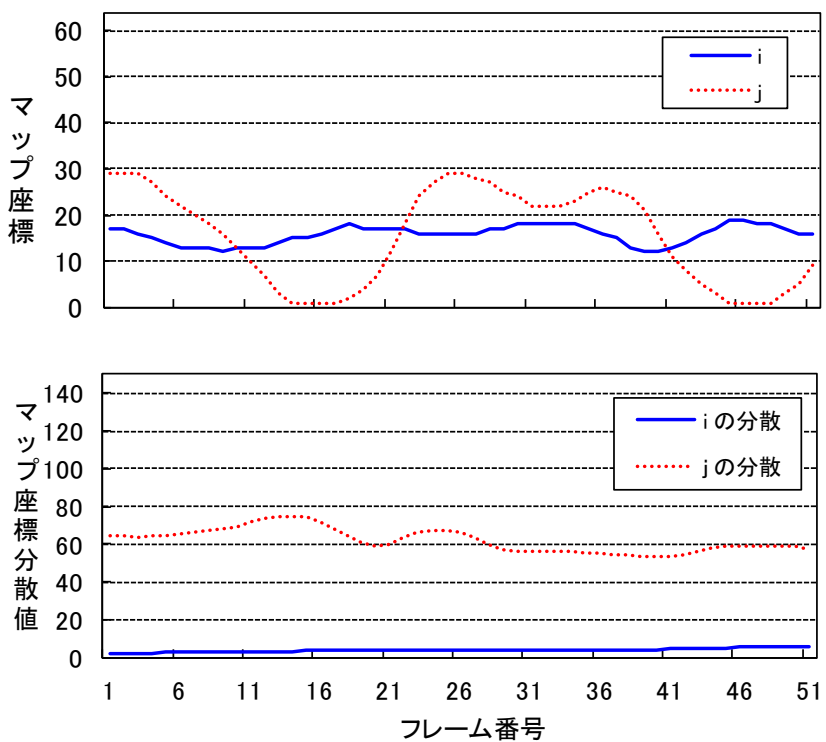


図 2-15 大きな揺き時の変化

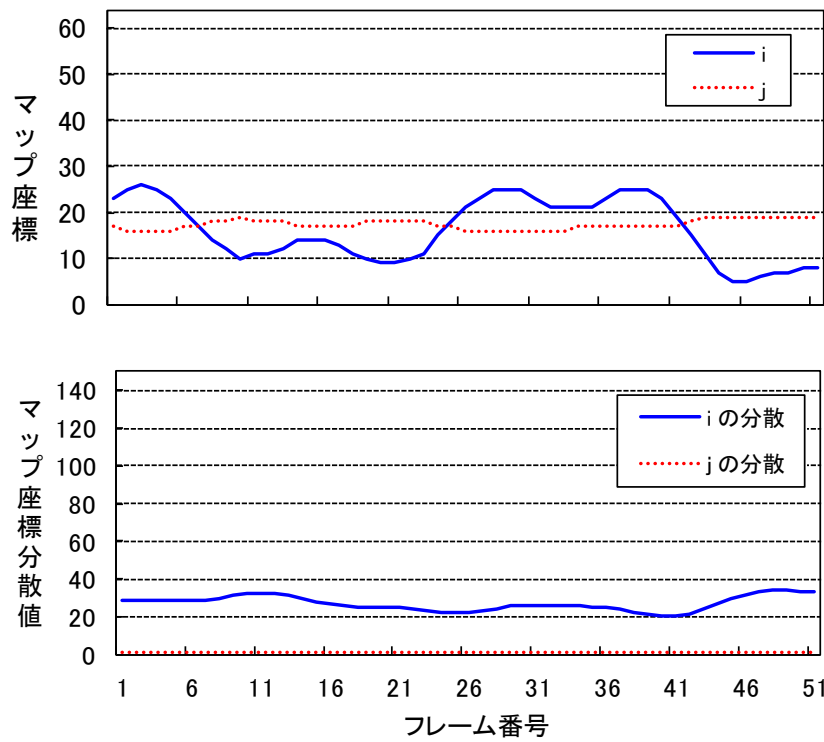


図 2-16 微小首振り時の変化

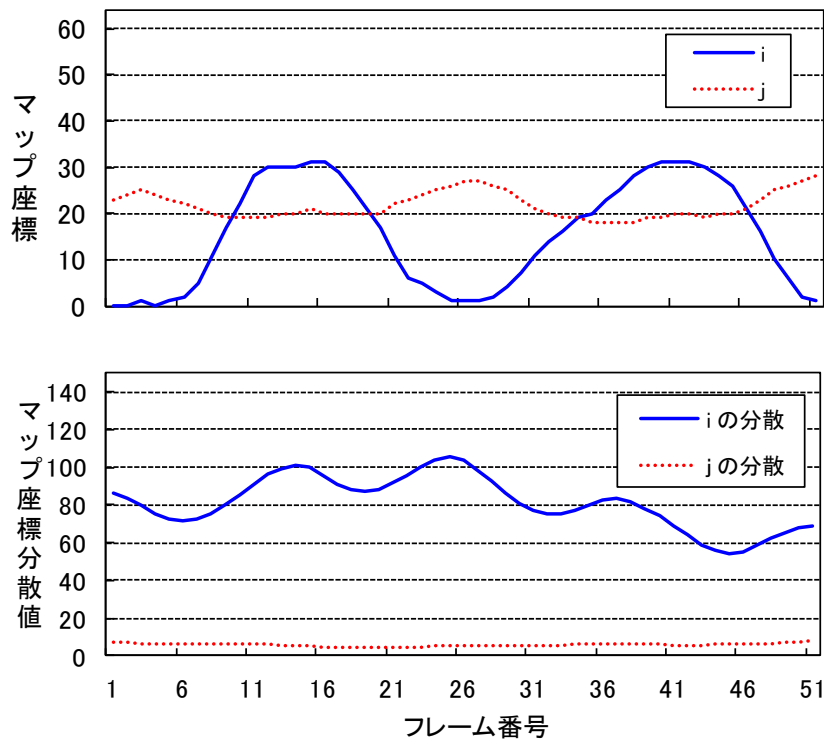


図 2-17 大きな首振り時の変化

2.8 まとめ

本章では、コンピュータとの豊かなインターフェイスの実現への応用が期待できる人間の顔方向を認識する手法について述べた。自己組織化マップにより顔画像を補間生成することで、局所的な領域の検出を必要とせず、少数の画像からあらゆる顔方向画像を生成し、精度よく人の顔方向を推定できることを確認した。

実験では、まず1自由度の姿勢変動に対して検討した。3次元データから顔方向の異なる画像を生成することで、精度保証された実験用データベースを実験に使用した。まず5度刻みのデータをSOMに学習することで、離散的な姿勢変動の顔画像から連続的な姿勢変動を学習したマップを生成する。次に動的計画法により隣接する姿勢画像の間を補間する画像を選択することで、姿勢推定に使用する姿勢推移画像列を選択する。最後に補間生成された姿勢推移画像列を使用してマッチングを行うことで、姿勢推定を行う。実験では、5度刻みの仮想生成画像を用いて補間生成した姿勢推移画像列を用いて、未学習の2.5度刻みの中間姿勢を含めた姿勢推定を行い、2.5度刻みの仮想生成画像をテンプレートとして与えた場合と同等の精度が得られることを確認した。このときのRMSEは1.87度という高い精度であることが確認された。これにより、姿勢推定を行う際に必要となるリファレンス画像の枚数削減が可能となると考えられる。

次に2自由度の姿勢変動について検討した。3次元データから生成された2自由度の姿勢変動を伴う顔画像をSOMに学習して、マップ上の類似度最大位置が2自由度の姿勢変動に対応することを確認した。さらに、この姿勢推定を動画に対して適用することで、首振りと頷きというジェスチャの認識が可能なことを確認した。

姿勢認識処理により、ドライバの姿勢推定や、人間の視線推定など、ユーザインターフェイスへの応用が期待できる。さらに、ジェスチャの認識を併用することで、より自然な対話システムなどへの応用も期待できる。

今後、顔検出システムと連動して姿勢推定やジェスチャ認識を動作させることで、実用的なシステムを構築したい。さらに、現在は顔画像をそのままマッピングすることで特定ユーザに対して認識をしているが、主成分分析などにより抽出した変数をマッピングすることで、不特定多数のユーザへの対応を検討したい。

第3章機械駆動異音の自動検出

3.1 まえがき

工業製品の異常検出・不良検査は、現在でも検査員による官能検査が広く用いられている^[35]。このような検査においては、誰がどこで何度検査しても同じ結果が得られ、同じ評価が得られなければならない。しかしながら、判断基準には個人差があり、体調や精神状態にも左右されるため一定の品質を保つことは困難である。また、検査員による官能検査は、経験則に従った検査であるため、不良判定を正しく行うには検査員の十分な経験が必要であり、検査員の育成には時間が掛かるという問題もある。そのため、生産量が増加した場合などに検査員を増員しても、増員がすぐに検査力アップとはならない。それに加え、国内においては、小子化に起因する新卒就労人口の減少と高年齢化という避けられない事実があり、検査員の採用・確保の問題が将来的に一層深刻化する恐れが高く、良質な検査員を確保し、増員するのは、ますます難しくなることが予想される。もし仮に増員できた場合においても、生産量に見合う増員が必要であるため、人件費等のコスト増加も伴うこととなる。このため、製造現場では自動検査システム導入によるコスト削減並びに検査品質の向上が注目されている^{[36] [37]}。

本研究では官能検査の中でも機械駆動部に関するものを対象として、録音された機械駆動音から、駆動部の良否判定（異音検査）を自動で行うことを目的とする。聴覚による官能検査は個人差が大きく、安定した品質の検査が難しい。このような検査に対する自動検査の簡単な手法としては、騒音計を使用して一定値以上のパワーが検出された場合に異常と判定する方法が用いられる。騒音計による手法は、良品サンプルの取り得る値を事前に確認することで、駆動音に混ざる異音に対する検査を実現できる。しかし、周期的な音など複雑な非連続音に含まれる駆動音は、時間とともにパワー値が変動することから、この手法は不向きであると考えられる。

そこで、音源波形のスペクトルから抽出した特徴量を用いてニューラルネッ

トワークや SVM を用いて判別を行う方法^[38]や、スペクトルの音圧帯域圧縮信号を用いてニューラルネットワークにより判別を行う手法^[39]が提案されている。さらに、良不良サンプルを学習した SOM マップと比較することで、良否を判定する手法も提案されており、製品化されている^[40]。これらの手法は、事前に良品サンプルと異音を含む不良品の駆動音サンプルの各クラス空間を学習することで、正常音と異音の特徴を分離する手法である。そのため、事前に検知すべき異音を想定して、異音サンプルを入手する必要がある。現場で発生する異音には様々な種類があり、それらに対応した高精度な識別には全ての異音パターンを網羅した異音サンプルセットが必要である。さらに、発生率の低い不良サンプルの収集には多大な労力と時間が必要となるため、これらの手法を適用すること自体が難しいことも多い。また、適切に検査を行うためには、入力された音の中から検査対象となる駆動音をセグメンテーションして、対象音の特徴を正しく抽出する必要がある。

本章では、画像・動画処理と音声認識に基づく検査手法をそれぞれ提案する。提案法は良品からの距離に基づき良否判定を行うため、学習用に多数の不良サンプルを必要としない。類似の研究として、溝口が自動車用のエンジン駆動音を対象として、各周波数帯のパワー値を用いた判別分析法を提案している^[41]。本稿では、エンジン駆動音のような連続的な音ではなく、非連続的な音を含む機械駆動音も対象として、CHLAC と連続 DP マッチングによる検査手法を提案する。

CHLAC は、テクスチャ分類や文字認識、顔認識などに用いられる高次局所自己相関 (HLAC)^{[42] [43]}を動画像用に拡張したものである。この特徴量を用いた研究として、人物動作動画からのジェスチャ認識や異常動作検出、不審動作検出などが行われている^{[44]-[46]}。提案法においては、駆動音から算出される時間一周波数平面の画像を時系列に並べ動画化し、その動画から CHLAC を抽出する。異常・不審動作の検出と駆動音の良否判定は、時系列データを分析し通常動作と異常動作に分類する点は同じであり、CHLAC の有効性が期待できる。特に CHLAC を用いることで、信号の形状（短時間の周波数成分）とその時間変化を同時に抽出することができるため、駆動音の不良らしさを時系列に追跡するような細かな解析が必要なくなる。本手法では、前処理としてセグメンテーションが必要であるため、後述する連続 DP マッチングにより処理対象区間を選択する。

次に、ケプストラム特徴量を用いて連続 DP マッチングにより識別する手法

を提案する。本手法は、学習やモデル化を必要とせず、音の周波数的特徴を良品リファレンスと比較することにより良否判定を行う。このとき連続 DP マッチングを用いて全領域に対してマッチングを行うため、セグメンテーションにより検査駆動音を切り出す必要もない。このため、新機種などに対して容易に適用することができる。さらに、連続 DP マッチングでマッチングされた領域を新たに 1 周期分のデータとしてリファレンスデータに使用することができる。従来は、新たにリファレンスデータを生成する際に全てのデータに対して手動で切り出す必要があったが、手動で 1 リファレンスを登録しておいて提案法によるマッチングを適用することで、半自動で新規リファレンスの生成が可能となる。

連続 DP マッチングとケプストラム特徴量による不良検査手法は、良品からの単純な距離値に基づく検査手法である。そのため、サンプル数が少ない場合にも適用することが可能であり、生産現場などにおいて実用性が高い。しかしながら、良品の駆動音に多くのバリエーションが存在する場合には、良品の特徴空間の拡がりに合わせてリファレンスを増やす必要があり、マッチングに必要な計算量が増すという欠点がある。一方、CHLAC と部分空間法による不良検査手法は、良品の部分空間を予め定義しているため、適用には一定量の良品サンプルが必要であるが、バリエーションが増えた場合においても検査時の計算コストは一定であるという利点を有している。

3.2 各種法の概要

本稿では、画像・動画処理と音声認識に基づく検査手法をそれぞれ提案する。それぞれの手法の概要について述べる。

3.2.1 CHLAC による異音検査

CHLAC による検査手法は良品データ学習部と良否判定部の 2 つに分けられる。学習および判定はグループ毎に行う（各入力データのグループは既知とする）。良品データ学習部では、事前に収録された良品データから検査区間を半自動で切り出す。切り出したデータを周波数解析して、スペクトログラム動画を作成する。その動画から CHLAC を抽出する。複数の良品データ（実験では 20 サン

ブルとした) から抽出された CHLAC から, 良品の部分空間を構成する. 良否判定部では入力データからスペクトログラム動画を作成し CHLAC を抽出する. 抽出した特徴と良品の部分空間との距離を求め, 良否を判定する.

3.2.2 連続 DP マッチングによる異音検査

連続 DP マッチングによる検査手法は, リファレンスの登録部と良否判定部の 2 つに分けられる. リファレンスの登録部では, 事前に用意した 1 つの良品リファレンスを基準として, 良品データ群からリファレンス区間を自動的に切り出して登録する. このとき, 機械駆動の 1 周期が入るようリファレンス範囲を設定した. 良否判定では, 登録された複数の良品リファレンス群と入力データとの距離を求めて一定の閾値以上の場合, 異音と判定する.

3.3 音声分析基礎理論

音声信号を 1 次元信号として観測した場合, 音声信号の時間的な変化をそのまま利用する方法と, 音声信号に含まれている周波数成分を利用する方法がある. 前者の方法は, 外部環境の変化に対し変動しやすいが, 後者の方法は変動が比較的小さく, その音声を特徴づける情報が容易に得られることから, 現在の音声分析の主流になっている. この周波数成分を用いて音声信号の分析を行う方法を一般にスペクトル分析 (Spectrum analysis) と呼んでいる. 本章で扱う音声分析はスペクトル分析が中心となる. 以下に, 音声信号を時間領域から周波数領域に変換するための理論を述べる.

3.3.1 フーリエ変換

音声のようなある繰り返し信号波形に対して, どのような周波数成分が含まれているかを求めるために用いられるのがフーリエ変換である.

ある区間 T における波形が $t = -\infty$ から $+\infty$ まで繰り返している信号 $s(t)$ に対し, その周波数成分を求める場合, 少なくとも T を一周期とする正弦波 $\sin \omega t$ または余弦波 $\cos \omega t$ の成分を含むことになる. ただし, $\omega = 2\pi/T$ である. この繰り返しの基本区間 T をピッチ周期といい, この周期 T の周波数を基本周波数と

いう。

次に、信号 $s(t)$ に基本周波数の正弦波と余弦波成分がどの程度含まれているかを求めるには、 $s(t)$ と $\sin \omega t$, $\cos \omega t$ との相関を求めればよい。すなわち、

$$\begin{aligned} \text{基本余弦成分} &= \int_{-\infty}^{+\infty} s(t) \cos \omega t dt \\ \text{基本正弦成分} &= \int_{-\infty}^{+\infty} s(t) \sin \omega t dt \end{aligned} \quad (3-1)$$

となる。しかし、この式は正規分区間が $-\infty$ から $+\infty$ までとなっているので、実際の計算はこのままでは不可能である。そこで、 $s(t)$ も $\sin \omega t$, $\cos \omega t$ も周期 T の波形となっているので、その成分は一周分だけの波形だけを切り出して、次式のように相関を計算することにより求めることができる。一周期に対する基本余弦波成分を a_1 , b_1 とすると、

$$\begin{aligned} a_1 &= \int_0^T s(t) \cos \omega t dt \\ b_1 &= \int_0^T s(t) \sin \omega t dt \end{aligned} \quad (3-2)$$

同様にして、基本周波数の2倍、3倍、…の周波数に対する成分 $a_2, b_2, a_3, b_3, \dots$ が求まる。したがって、一般には基本周波数の n 倍の周波数に対する成分を a_n , b_n とすると、次式のようなになる。

$$\begin{aligned} a_n &= \int_0^T s(t) \cos(n\omega t) dt \\ b_n &= \int_0^T s(t) \sin(n\omega t) dt \end{aligned} \quad (3-3)$$

また、 a_0 として、一周期における信号 $s(t)$ の平均値を次のように求める。

$$a_0 = \int_0^T s(t) dt \quad (3-4)$$

(3-3), (3-4)式を一般にフーリエ変換といい、その係数 a_0, a_n, b_n をフーリエ係数という。また、余弦波成分を実数、正弦波成分を虚数として、複素数で表すと次式のように定義される。

$$S(nw) = \int_0^T s(t)e^{-jn\omega t} dt \quad (3-5)$$

3.3.2 離散フーリエ変換(DFT)

フーリエ変換は、アナログ信号のスペクトル分析に非常に有効な方法である。それに対し、サンプリングによりデジタル化された信号を分析する場合に用いるのが離散的フーリエ変換(DFT : Discrete Fourier Transform)である。

音声波形 $s(t)$ をサンプリング周期 T [s] でデジタル化した時の n 個の標本値時系列を $\{s(n)\}$, $s(n) = s(nT)$, $n = 0, 1, \dots, N-1$ とする。アナログ波形に対する(3-5)式に対応して、

$$S(k\Omega) = \sum_{n=0}^{N-1} s_n e^{-j\Omega nk} \quad \Omega = \frac{2\pi}{NT} = \frac{w}{N} \quad (3-6)$$

$$k = 0, 1, \dots, N-1$$

を $\{s(n)\}$ の DFT と定義する。

3.3.3 高速フーリエ変換(FFT)

実際にデジタル化された信号系列 $\{s_n\}$ からその離散的スペクトル $\{S_k\}$ を計算するのに(3-7)式からもわかるように、特定の k について F_k の値の計算は N 回の複素乗算と N 回の複素加算が必要である。したがって、 $\{F_k\}$, $k = 0, 1, 2, \dots, N-1$ の全体を求めるには、 N^2 の複素乗算と加算が必要となる。 $N = 128$ とすれば、 $N^2 = 16384$, $N = 256$ とすれば、 $N^2 = 65536$ となり、高速の計算機を使用してもその演算時間は無視できず、実時間処理は不可能に近くなる。

この有用な演算を高速に実行するために、 $e^{j2\pi mk/NT}$ の周期性を利用して考察されたのが高速フーリエ変換(FFT : Fast Fourier Transform)である。この手法によれば、必要な複素乗算と加算の回数は $N \log_2 N$ 回ですむことになる。したがって、DFT に比べて、 $N \log_2 N / N^2 = \log_2 N / N$ の比率で少なくてすむ。例えば、 $N = 128$ の時、比率は $\log_2 N / N = 7/128$, $N = 256$ の時、

$\log_2 N/N = 8/256 = 1/32$, となり, N が大きい程, この高速化の効果は顕著になる.

次に FFT の原理を示す. 処理する信号の標本値 N を 2 のべき乗 ($N = 2^m$) であるとする. まず, 信号 $s(t)$ の標本値の系列 $\{s(n)\}$, $n = 0, 1, \dots, N-1$ を偶数番目の標本値の系列 $\{g_l\}$ と奇数番目の標本値の系列 $\{h_l\}$ に分ける.

$$\begin{aligned} \{g_l\}: s_0, s_2, \dots, s_{N-2} \{g_l\} &= \{s_{2l}\} \\ l &= 0, 1, \dots, \frac{N}{2} - 1 \\ \{h_l\}: s_1, s_3, \dots, s_{N-1} \{h_l\} &= \{s_{2l+1}\} \\ l &= 0, 1, \dots, \frac{N}{2} - 1 \end{aligned} \quad (3-7)$$

この 2 つの系列は, 共に項数が $N/2$ で, 周期が $NT/2$ の系列とみなすことができる.

したがって, その DFT は,

$$\begin{aligned} G(k\Omega) = G_k &= \sum_{l=0}^{\frac{N}{2}-1} g_l e^{-j2\pi \frac{2}{NT} kl} = \sum_{l=0}^{\frac{N}{2}-1} g_l (W^2)^{kl} \\ H(k\Omega) = H_k &= \sum_{l=0}^{\frac{N}{2}-1} h_l e^{-j2\pi \frac{2}{NT} kl} = \sum_{l=0}^{\frac{N}{2}-1} h_l (W^2)^{kl} \end{aligned} \quad (3-8)$$

ただし, $W = e^{-jw2\pi \frac{1}{NT}}$

一方, $\{s_n\}$ の DFT は,

$$\begin{aligned} F(k\Omega) = F_k &= \sum_{i=0}^{N-1} s_i e^{-jw2\pi \frac{1}{NT} ki} \\ &= \sum_{l=0}^{\frac{N}{2}-1} \{s_{2l} W^{2lk} + s_{2l+1} W^{(2l+1)k}\} \\ &= \sum_{l=0}^{\frac{N}{2}-1} \{g_l W^{2lk} + h_l W^{(2l+1)k}\} \\ &= G_k + W^k H_k \end{aligned} \quad (3-9)$$

$$k = 0, 1, \dots, N-1$$

ここで、 $\{G_k\}, \{H_k\}$ は周期が $NT/2$ であることを使えば、

$$F_k = \begin{cases} G_k + W^k H_k & k = 0, 1, \dots, \frac{N}{2} - 1 \\ G_k - \frac{N}{2} + W^k H_k - \frac{N}{2} & k = \frac{N}{2}, \frac{N}{2} + 1, \dots, N - 1 \end{cases} \quad (3-10)$$

上式は、 N 点のDFTの計算が、 $N/2$ 点のDFTの計算2回に分解できることを示している。例えば $N=8$ の時にこの処理を繰り返していけば、 $2N \log_2 N$ 回の計算で N 点のDFTが求まることになる。しかも、各点での積和の計算のうち、一方は必ず乗算の係数が1であり単なる加算の演算になり計算が大幅に最適化される。

3.3.4 線形予測分析

音声波形は単に一つ前の信号波形だけではなく、過去の数サンプルの値とも関係している。

そこで、標本値間に相関に基づく高い予測性があると仮定し、予測形式として線形予測として線形予測(線形1次結合)、すなわち、

$$s_n = a_1 s_{n-1} + a_2 s_{n-2} + \dots + a_p s_{n-p} \quad (3-11)$$

の形で推定する。(3-11)式は、音声波形の任意の標本値 s_n がそれより過去の p 個の標本値のおのおのに、ある係数 a_i をかけて加え合わせたもの、つまり過去の p 個の値の加重平均で近似できるということを意味している。

(3-11)式を仮定し、その線形予測誤差の自乗平均を最小にするという条件で、線形予測係 $\{a_i\}, i=1, 2, \dots, p$ を求めることを、線形予測分析(Linear Prediction Coding: LPC)という。

(3-11)式の関係より線形予測係数 $\{a_i\}$ の求め方を導いてみる。

\hat{s}_n を s_n の予測値、

$$\hat{s}_n = a_1 s_{n-1} + a_2 s_{n-2} + \dots + a_p s_{n-p} \quad (3-12)$$

s_n をその実現値としたとき、

$$\varepsilon_n = s_n - \hat{s}_n = s_n - \sum_{i=1}^p a_i s_{n-i} \quad (3-13)$$

を予測誤差と定義する． $-a_i$ を a_i と置き直して考えれば，予測誤差は，

$$s_n + \sum_{i=1}^p a_i s_{n-i} = \sum_{i=0}^p a_i s_{n-i}, \quad a_0 = 1 \quad (3-14)$$

となる．したがって，予測誤差の自乗平均は，

$$\overline{\varepsilon_n^2} = \overline{\left(\sum_{i=0}^p a_i s_{n-i} \right)^2} = \overline{(s_n + a_1 s_{n-1} + a_2 s_{n-2} + \cdots + a_p s_{n-p})^2} \quad (3-15)$$

これは正もしくは0の量であり，極値が1つしかなければ，それが最小値である．よって，予測誤差の自乗誤差を最小にするような係数 $\{a_i\}$ は，上式の ε_n^2 の各 a_i についての変微分係数を0とおく連立 p 元1次方程式の解として求められる．

$$\frac{\partial \overline{\varepsilon_n^2}}{\partial a_i} = 0 \quad i = 1, 2, \dots, p \quad (3-16)$$

(3-15)式の右辺を展開し，項別に平均値をとり，(3-16)式の条件を用いることにより，次の連立 p 元1次方程式が得られる．

$$\begin{pmatrix} r_0 & r_1 & r_2 & \cdots & r_{p-1} \\ r_1 & r_0 & r_1 & \cdots & r_{p-2} \\ r_2 & r_1 & r_2 & \cdots & r_{p-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{p-1} & r_{p-2} & r_{p-3} & \cdots & r_0 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_p \end{pmatrix} = - \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ \vdots \\ r_p \end{pmatrix} \quad (3-17)$$

ただし， $r_0 = \overline{s_i s_i}$ $r_j = \overline{s_i s_{i+j}} = \overline{s_{i-j} s_i}$

ここで， r_j は波形 $\{s_i\}$ の相関係数であり， $r_j = r_{j-1}$ と s_i の定常性を仮定している．上式から，左辺の相関係数 r_j が正常値であれば，その逆行列が存在し，上式

から $\{a_i\}$ を求めることができる. 実際には有限個 (N 個) の実測データ $\{s_n\}$ から, 窓をかけたデータ $s'_n = s_n w_n$ を作り, 次式によって相関係数を求める.

$$r_j = \frac{1}{N} \sum_{n=0}^{N-j-1} s'_n s'_{n+1} = \frac{1}{N} \sum_{n=0}^{N-j-1} s_n s_{n+1} = \frac{1}{N} \sum_{n=0}^{N-j-1} s_n s_{n+j} w_n w_{n+j} \quad (3-18)$$

これは実測データ $\{s_n\}$ として, $n=0,1,\dots,N-1$ の N 個のデータがあり, その外側では $\{s_n\}$ は 0 であると仮定していることを意味している. この方法を相関法といい, 窓掛けと(3-18)式の相関係数の定義から係数行列の正常値性が保障され, 線形予測係数 $\{a_i\}$ として安定な解が得られる.

一方, 無限に続いているデータの一部分による演算という考え方から, r_j の演算を,

$$r_{ij} = \frac{1}{N} \sum_{n=0}^{N-1} s_{i+n} s_{j+n} \quad (3-19)$$

と定義することもできる. このとき, 窓掛けは意味がなく, $r_{ii} = r_{jj}$ であり, 係数行列の正常値性は必ずしも保障されない. このような係数行列による解き方を共分散法という.

波形 $\{s_n\}$ の系列が長く, 定常であるときは, 相関法と共分散法はほとんど同じ $\{a_i\}$ を与えるが, 系列が短く, $\{s_n\}$ が時間的な変化を伴う場合には異なった $\{a_i\}$ の値を与える. $\{s_n\}$ の定義からいって, 共分散の方が時間的な変化に対する応答性が良いが, 解が必ずしも安定な $\{a_i\}$ を与えないという欠点がある.

線形システム同定理論によれば, 系の入力 $\{x_n\}$ と出力 $\{s_n\}$ の間に,

$$s_n + \sum_{i=1}^p a_i s_{n-1} = \varepsilon_n \quad (3-20)$$

の関係があるときこれを自己回帰 (AR : auto-regressive) 過程といい, そのシステム関数は,

$$H(z) = \frac{1}{1 + \sum_{i=1}^p a_i z^{-i}} \quad (3-21)$$

であり、極のみを有することから全極モデル(all-pole)モデルともいう。また、

$$s_n + \sum_{i=1}^p a_i s_{n-i} = x_n + \sum_{j=1}^q \beta_j x_{n-j} + \varepsilon_n \quad (3-22)$$

の関係があるとき、これを自己回帰・移動平均(AR・MA: auto-regressive・mean-average)過程といい、そのシステム関数は、

$$H(z) = \frac{1 + \sum_{i=1}^q \beta_i z^{-i}}{1 + \sum_{i=1}^p a_i z^{-i}} \quad (3-23)$$

となるから極零(pole-zero)モデルという。

波形のサンプル値間の線形予測分析過程は、制御理論的に言えば音声波形をAR過程とみなし、その全極形システム関数を規定する係数 $\{a_i\}$ を求める(2.7)式が、制御理論の正規方程式、またはYule-Walkerの方程式と言われるものに他ならないのである。

3.3.5 線形予測法によるスペクトル分析

線形予測法を用いてスペクトル分析を行うことにより、きわめて正確に、音声のホルマント周波数や帯域幅が求まる。時刻 n における音声信号の予測値 \hat{s}_n とすると、 \hat{s}_n は次式のように、過去の p サンプル点の値から線形結合によって求められる。

$$\hat{s}_n = \sum_{i=1}^p a_i s_{n-i} \quad (3-24)$$

これより、実際の音声 s_n と予測値 \hat{s}_n の誤差信号 ε_n は次式で表せる。

$$\begin{aligned}
\varepsilon_n &= s_n - \widehat{s}_n \\
&= s_n + \sum_{i=1}^p a_i s_{n-i}
\end{aligned} \tag{3-25}$$

両者を z 変換すると,

$$E(z) = \left(1 + \sum_{i=1}^p a_i z^{-i} \right) S(z) \tag{3-26}$$

となり, 音声信号のパルス伝達関数 $S(z)$ は, (3-26)式より,

$$S(z) = \frac{E(z)}{1 + \sum_{i=1}^p a_i z^{-i}} \tag{3-27}$$

で求められる. 我々が聞いている音声は, 音源波形が声道の伝達関数 $H(z)$ を通過した出力波形である. そこで, 音声を一つの入出力システムと考え, 次式のような線形の離散値モデルで表現する.

$$s_n = -\sum_{k=1}^p a_k s_{n-k} + G \sum_{l=1}^q \beta_l u_{n-l} \tag{3-28}$$

時間遅れを表す z^{-1} を用いて, (3-28)式を書き換えると次式のようになる.

$$\left(1 + \sum_{k=1}^p a_k z^{-k} \right) S(z) = G \sum_{l=1}^q \beta_l z^{-l} U(z) \tag{3-29}$$

ただし, $S(z)$, $U(z)$ はそれぞれ s_n , u_n の z 変換である. これより, 声道の離散値モデルの伝達関数 $H(z)$ は次式で求まる.

$$H(z) = \frac{S(z)}{U(z)} = \frac{G \sum_{l=0}^q \beta_l z^{-l}}{1 + \sum_{k=1}^p a_k z^{-k}} \tag{3-30}$$

一般的には、次式のような全極モデルが用いられる。

$$H(z) = \frac{G}{1 + \sum_{k=1}^p a_k z^{-k}} \quad (3-31)$$

このような全極モデルの係数 a_k は、線形予測法により求めた係数となる。
(3.30) 式から、声道のパワー伝達関数は次式によって表せる。

$$|H(z)|^2 = \frac{G^2}{\left| 1 + \sum_{k=1}^p a_k z^{-k} \right|^2} \quad (3-32)$$

このパワー伝達関数から、声道のパワースペクトルを求めることができる。
 z^{-1} は信号の 1 サンプリグ時間 T の遅れ要素で、複素周波数 s を用いて、 $z^{-1} = e^{-sT}$ で表される。したがって、周波数スペクトルを求めるために、 $s = j\omega$ とすると、 $z^{-1} = e^{-j\omega T}$ で与えられ、 z^{-1} は、角周波数 ω の関数となり、 ω が与えられれば z^{-1} の値が求まる。これを周波数 f で表すと、ナイキスト周波数 $f_{\max} = 1/2T$ 、 $\omega = 2\pi f$ となっているので、

$$z^{-1} = e^{-j\pi \frac{f}{f_{\max}}} \quad (3-33)$$

となり、(3-32)式に代入すると、任意の周波数 f に対するそのパワースペクトル値は次式(3-34)式で与えられる。

$$|H(z)|^2 = \frac{G^2}{\left| 1 + \sum_{k=1}^p a_k e^{-j\pi \frac{f}{f_{\max}} k} \right|^2} \quad (3-34)$$

(3-34)式の予測係数で表されているパワー伝達関数から、高速にパワースペクトルを求めるアルゴリズムとして FFT を用いる方法について示す。まず、(3-34)式より、次式が求まる。

$$10\log_{10}|H(z)|^2 = 20\log_{10}G - 10\log_{10}\left|1 + \sum_{k=1}^p a_k e^{-j\pi\frac{f}{f_{\max}}k}\right|^2 \quad (3-35)$$

この式より，パワースペクトルを求めるには，第 2 項の $\left| \right|^2$ の中の式の計算を行えばよい．このスペクトルを，周波数間隔 f を細かい刻みで求めるには， f を少しずつ変化させて，繰り返し計算しなければならない．この計算は， f を一定間隔で求める場合，FFT を利用することができる．

$$e^{-j\pi\frac{f}{f_{\max}}} = W \quad (3-36)$$

とすると，そのパワースペクトルは，次のように与えられる．

$$\begin{aligned} \text{直流分} &: 1 + a_1W^0 + a_2W^0 + \dots + a_pW^0 \\ f \text{ 成分} &: 1 + a_1W^1 + a_2W^1 + \dots + a_pW^1 \\ 2f \text{ 成分} &: 1 + a_1W^2 + a_2W^2 + \dots + a_pW^2 \\ &\vdots \\ nf \text{ 成分} &: 1 + a_1W^n + a_2W^n + \dots + a_pW^n \end{aligned} \quad (3-37)$$

この演算は，求めるスペクトルの個数を 2^L とすると，データ d として次式，

$$d = [1, a_1, a_2, \dots, a_p, 0, \dots, 0] \quad (3-38)$$

を与えた FFT で順次求められる実数部，虚数部から計算できることになる．

3.4 連続 DP マッチングによる検査手法

CHLAC による検査手法では，学習する良品データを選択する際と，入力データ中から良否判定を行う区間を設定する際に，連続 DP マッチングと一次の自己相関演算による検査区間の抽出を行う．一方，連続 DP マッチングによる検査手法では，リファレンスの自動生成と良否の判定時に連続 DP マッチングを使用する．本節では，どちらの手法においても共通して使用する連続 DP マッチングについて述べる．

3.4.1 連続 DP マッチング

連続 DP マッチングは、DP マッチングを拡張した手法である。リファレンスデータに対して入力データの全点を始点として DP マッチングを行い、距離を計算する。このとき、得られた正規化累積距離が最小の値を距離として選択することで、入力データ中のどの区間に検査区間がある場合でも距離を計算可能である（図 3-1）。さらに、正規化累積距離最小の位置からバックトレースを行うことで、リファレンスに最も近い駆動音を 1 周期分選択することができる。DP パスには、音声認識などでも広く用いられる傾斜制限 $1/2 \sim 2$ の対称型 DP パスを用いる。

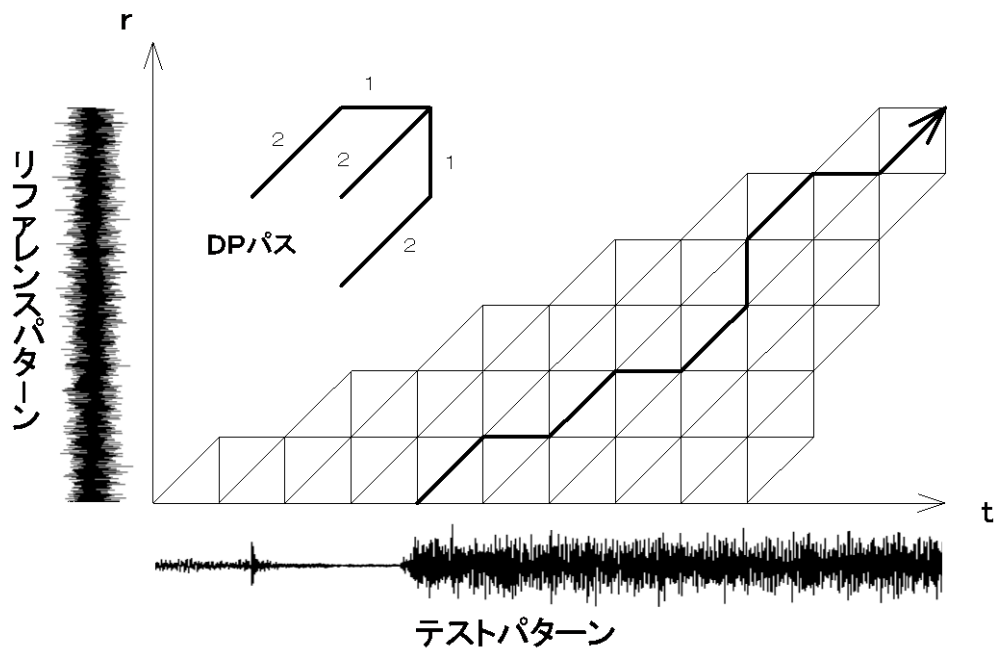


図 3-1 連続 DP マッチング

フレーム間距離はケプストラム特徴量のユークリッド距離より求める。この場合の正規化累積距離は、(3-39)式に示す最小累積距離 P と (3-40)式に示す最適積分路の長さ c より、(3-41)式のように求められる。

$$P(t, r) = \min \left\{ \begin{array}{l} P(t-1, r-2) + 2Q(t, r-1) \\ P(t-1, r-1) + Q(t, r) \\ P(t-2, r-1) + 2Q(t-1, r) \end{array} \right\} + Q(t, r) \quad (3-39)$$

$$C(t, r) = \begin{cases} C(t-2, r-1) + 3 \\ C(t-1, r-1) + 2 \\ C(t-1, r-2) + 3 \end{cases} \quad (3-40)$$

$$P(t, T) = \frac{P(t, T)}{C(t, T)} \quad (3-41)$$

3.5 CHLAC による検査手法

本節では、CHLAC による検査手法について詳細を述べる。

3.5.1 立体高次局所自己相関特徴 (CHLAC)

CHLAC は従来テクスチャ分類や文字認識等に用いられていた HLAC (高次局所自己相関特徴) を、時間軸情報も扱えるように拡張したものである (図 3-2)。N 次の HLAC (H_N) は(3-42)式のように表される。

$$H_N(a_1, a_2, \dots, a_N) = \int f(r)f(r+a_1)\dots f(r+a_N) dr \quad (3-42)$$

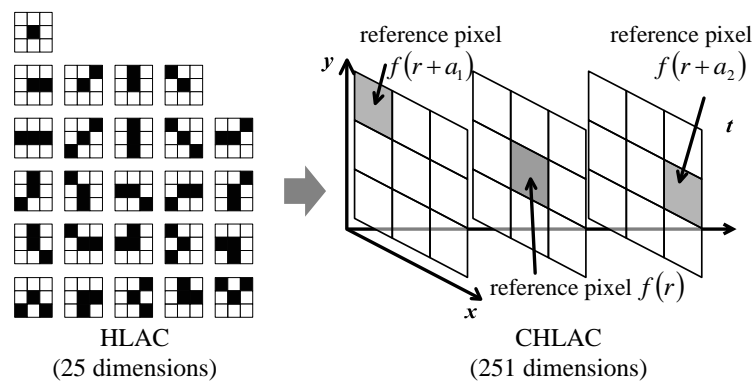


図 3-2 立体高次局所自己相関特徴 (CHLAC)

ここで $r \in R^3$ は位置ベクトル (x, y, t) , $f(r)$ は着目点でのデータ値 (例えば輝度値), $a_1, a_2, \dots, a_N \in R^3$ は変位ベクトルをあらわす. CHLAC は、さらに時間方向の積分範囲を持たせることで、時間方向の相関関係も考慮した特徴量である. 本稿では、N は 2 まで、変位ベクトルの範囲は r を中心とした $3 \times 3 \times 3$ の

領域までとする．このとき，CHLAC 特徴は，0 次が 1 個，1 次が 13 個，2 次が 237 個の計 251 次元のベクトルとなる．具体的な演算方法を次に示す．

- ① 着目点と参照点画素の輝度値積を算出
- ② 着目点をシフトし，画像全体に対して①を実行
- ③ ①，②の操作を A フレーム先まで繰り返して積分
- ④ ①～③の操作を 251 種の組み合わせ全てで実行

3.5.2 検査区間の選択

まず，連続 DP マッチングにより抽出した 1 周期分のデータを使用して自己相関を算出する．自己相関算出結果のピーク点から周期性を判断して，検査区間を決定する．自己相関演算式 R_τ を(3-43)式に示す．ただし， $1 \leq \tau \leq t_{all} - N$ であり， N は駆動音 1 周期におけるサンプリング点数， t_s は連続 DP により選択した駆動音の開始時刻， x は原波形振幅値， t_{all} は駆動音データ全体のサンプリング点数を示す．

$$R_\tau = \frac{1}{N} \sum_{t=1}^N x(t+\tau)x(t+t_s) \quad (3-43)$$

自己相関値の高い区間を検査区間として選択することで，周期的な駆動音の動作部を切り出すことができる．**図 3-3**，**図 3-4** に本手法により抽出した検査区間の例を示す．サンプル全体の 90 [%]程度は**図 3-3** のように正確な抽出に成功したが，**図 3-4** のように一部が欠落する場合も確認された．

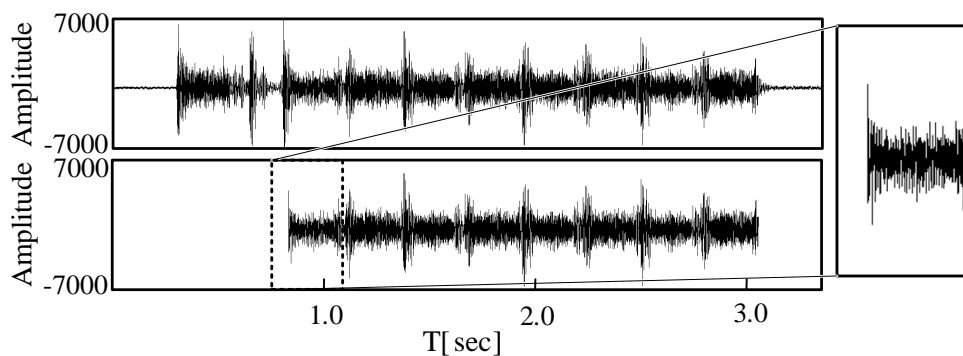


図 3-3 セグメンテーションの成功例
(上段：切り出し前， 下段：切り出し結果)

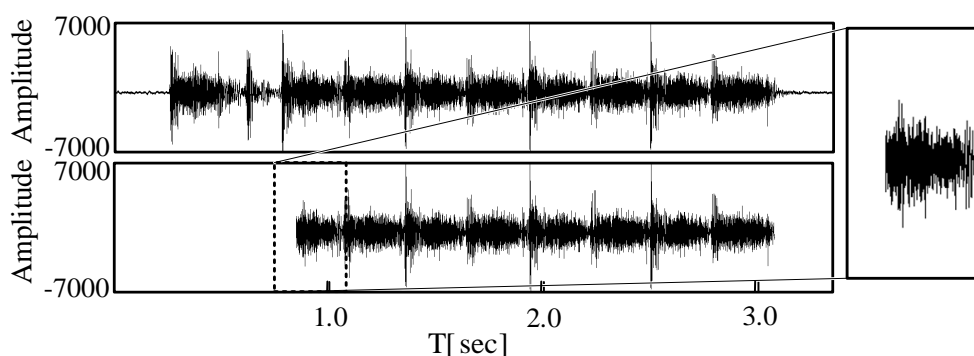


図 3-4 セグメンテーションの失敗例
(上段：切り出し前, 下段：切り出し結果)

3.5.3 CHLAC 特徴の抽出

選択された区間の駆動音データから CHLAC 特徴量を抽出する．はじめに，駆動音からサウンドスペクトログラム動画を生成する．動画を生成するために，駆動音を FFT および LPC（線形予測分析）により周波数－時間平面に変換する（図 3-5）．駆動音原波形から直接 CHLAC を算出するのではなく，FFT や LPC により周波数－時間平面に変換するのは，音声特徴の幾何学的構造はフォルマントやその時間遷移に強く表わされている^[47]ためである．

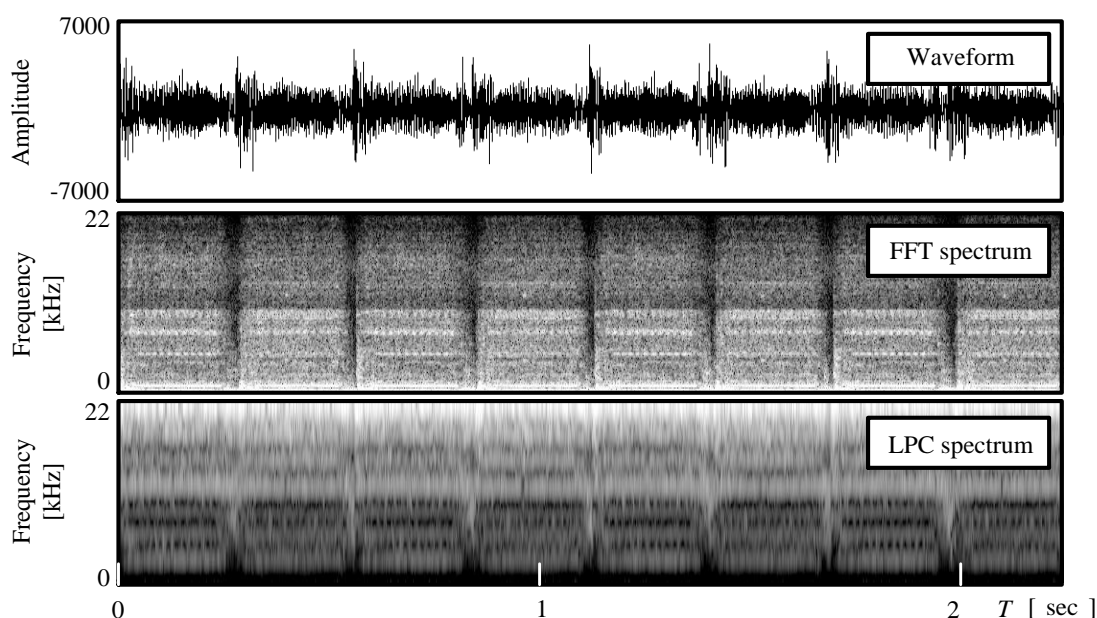


図 3-5 サウンドスペクトログラムの一例

次に、スペクトルの時間遷移に関する特徴量を得るために、周波数-時間平面を重ね合わせた動画を生成して、CHLAC を抽出する。CHLAC は、スペクトルの時間遷移に伴う特徴であり、位置不変の特徴量であるため、重ね合わせの始点位置には依存しない特徴抽出を行うことができる。この際、LPC はフォルマントが極として現われるため、時間遷移特徴が顕著に現れると考え、FFT と共に検討した。次に、一定の切り出し幅 [pixel]、シフト幅 [pixel] で画像を切り出し、切り出し画像を時系列に重ねていくことでスペクトログラム動画を生成する (図 3-6)。動画の横 1 [pixel] がスペクトルの 1 分析フレームを表す。

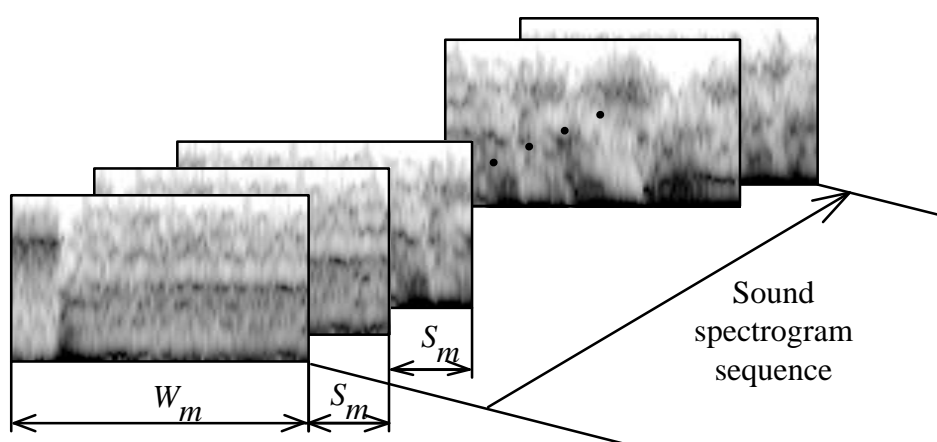


図 3-6 サウンドスペクトログラム動画の生成方法

作成したスペクトログラム動画より良否判定に使用する CHLAC を抽出する。なお、実際の抽出には、スペクトル変化をより捉え、より長時間の相関を考慮した特徴量を算出するために、スペクトログラム動画のフレーム間差分動画を用いた。

3.5.4 良品空間の学習

予め用意した良品サンプルから抽出される CHLAC を用いて良品データの部分空間を構成する。部分空間法を用いて次元圧縮を行うことで、識別に不要な特徴を取り除く。

部分空間は、主成分分析で求めた主成分ベクトル $U = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_M]$ より求めることができる。主成分ベクトルを固有値の大きいほうから順に d 個選ぶことで d 次の部分空間が張られる。本稿では固有値より累積寄与率 η を求め、 η がある閾値となる次元 d までの固有ベクトル (主成分ベクトル) により

張られる d 次の空間を部分空間として構成する。

テストパターンとなる駆動音の異常度（不良らしさ）は，異常行動，不審者検出の先行研究^{[44][45]}を参考とし，学習された部分空間からの距離 d_{\perp} として定義する． d_{\perp} は駆動音のスペクトログラム動画から抽出した CHLAC 量ベクトルと，学習データから構成した部分空間との直交補空間への射影距離である．学習データとして良品データのみを用いることで良品空間が構成されるため，入力特徴ベクトルと良品空間との距離 d_{\perp} は異常度合いと捉えることができる．この手法は不良品の性質を事前に知る必要がないという利点がある．距離 d_{\perp} の導出を(3-44)式に示す．

$$\begin{aligned}
 d_{\perp}^2 &= \|P_{\perp} \mathbf{x}\|^2 \\
 &= \|(I_M - U_K U_K^T) \mathbf{x}\|^2 \\
 &= \mathbf{x}^T (I_M - U_K U_K^T)^T (I_M - U_K U_K^T) \mathbf{x} \\
 &= \mathbf{x}^T (I_M - U_K U_K^T) \mathbf{x}
 \end{aligned} \tag{3-44}$$

ここで \mathbf{x} はテストデータ動画中の任意フレームから得られる特徴ベクトル（入力ベクトル）， P は入力ベクトルの部分空間への射影子 ($P = U_K U_K^T$)， P_{\perp} は入力ベクトルの部分空間の直交補空間への射影子 ($P_{\perp} = I_M - P$)， U_K は主成分直交基底， I_M は M 次の単位行列 (M は入力ベクトルの次元) である．サウンドスペクトログラム動画の各フレームにおいて良品空間との距離 d_{\perp} を算出して，次節の手法で良否を判定する．

3.5.5 良否判定

テストデータ動画からフレーム毎に得られる特徴ベクトルより，学習空間との時系列距離値を得る．良否判定はこのデータを FFT して，そのパワースペクトル \mathbf{p} の N 次までの低次成分平均 $g(\mathbf{p})$ より行う ((3-45)，(3-46)式)． b は閾値である．予備実験として N を 1 次から順に増やして識別率を確認した結果， $N=3$ で良好な判別結果が得られたため，以降 $N=3$ として実験を行う．

$$g(\mathbf{p}) = \text{Ave}\{p_i | 1 \leq i \leq N\} \tag{3-45}$$

$$\begin{cases}
 g(\mathbf{p}) < b & : \text{ 駆動音を良品と判定} \\
 \text{otherwise} & : \text{ 駆動音を不良品と判定}
 \end{cases} \tag{3-46}$$

時系列距離値をFFTすることで距離の時間変動を考慮し、概形成分を得る。図 3-7、図 3-8 はグループ 2 の実験結果例である。図 3-7 のような場合は不良品サンプルの良品学習空間からの距離が常に大きいため、時系列距離値の平均を用いても良否閾値を決定することができる。しかし、図 3-8 に示した良品サンプルの先頭部分のように検査対象区間の抽出ミス等により異常値が発生し、平均値を用いただけでは誤判定が発生する可能性がある。このような場合においてもFFTの概形を得ることで異常値の影響を抑えた識別を実現する。

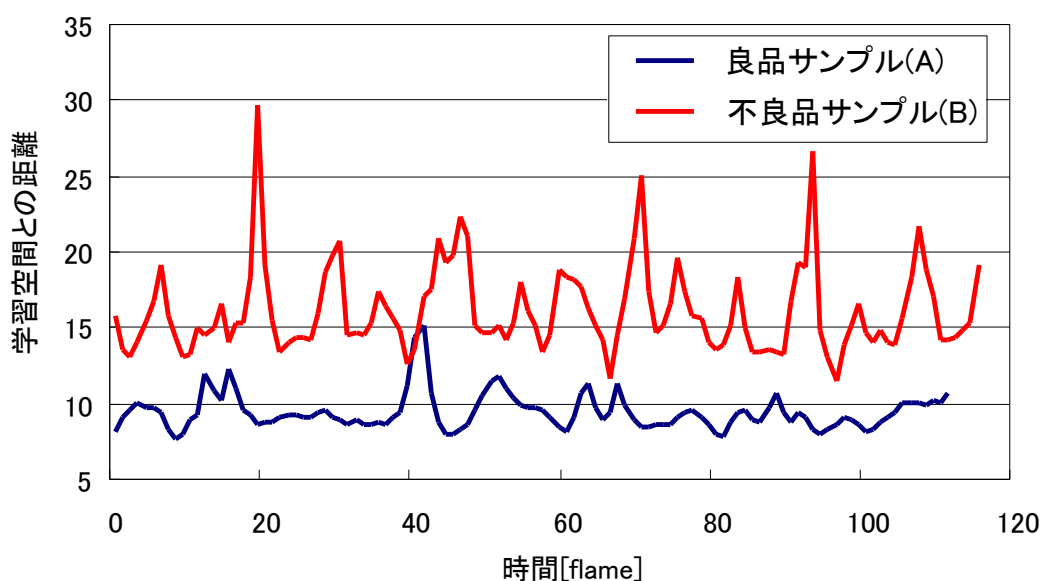


図 3-7 距離値の算出例 1

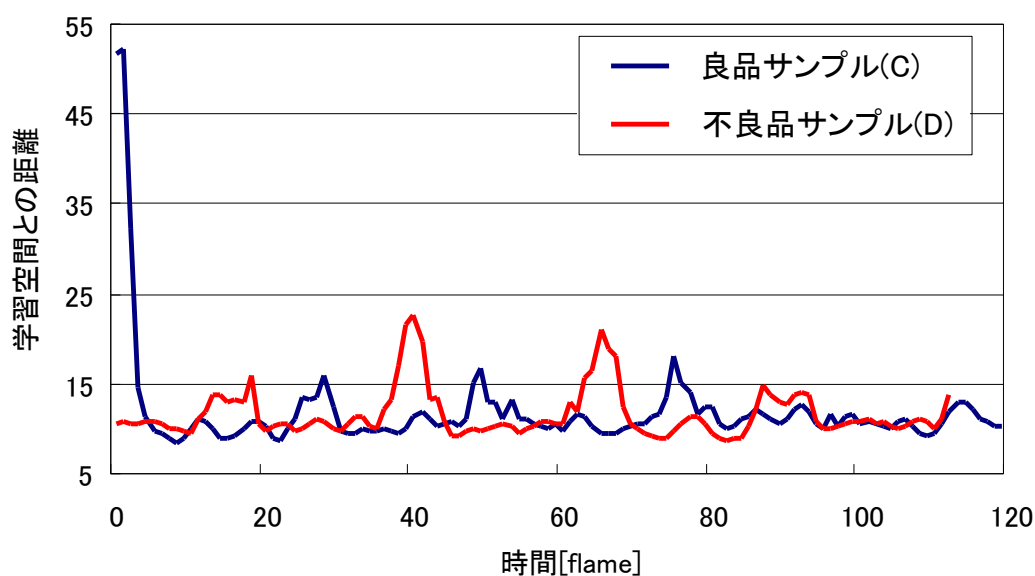


図 3-8 距離値の算出例 2

3.6 連続 DP マッチングによる検査手法

本節では、連続 DP マッチングによる検査手法について詳細を述べる。連続 DP マッチングによる不良判定処理は、良品リファレンスの登録と良否判定の 2 ブロックに大きく分かれる。

3.6.1 良品リファレンスの登録

まず、良品データ中から 1 周期分のデータを手動で選択する。このとき、周期性のあるデータについては自己相関 $C(d)$ を算出して、最も相関の高い d を選択範囲の長さとして採用することで 1 周期分を正確に選択する ((3-47), (3-48) 式)。ただし、 t_s は DP もしくは手動で選択した駆動音の開始時刻、 x は原波形振幅値を示す。また、 d は初期選択範囲の長さに対して 75~150 [%] の範囲で検索するものとした。

$$C(d) = \sum_{t=0}^{d-1} \frac{(x(t_s+t) \times x(t_s+d+t))}{P(t_s, d)} \quad (3-47)$$

$$P(t_s, d) = \frac{\sum_{t=0}^{d-1} (x(t_s+t) \times x(t_s+t))}{d} \quad (3-48)$$

上記の選択により、1 周期分のデータを「良品リファレンス作成に使用するリファレンス（以降、基準リファレンスと表記）」として登録する。次に、基準リファレンスをリファレンスとして良品データに対して連続 DP マッチングを行い、検査に使用する 1 周期分の良品リファレンスデータを自動的に生成する。さらに、周期的な駆動音データについては、前述した自己相関による周期検出を用いて、長さを補正する。これにより、作業員のスキルに依存しないリファレンス作成を実現すると共に、作業員の負担を軽減する。自己相関は、グループ 1 に対しては原波形に対して算出する。一方、高周波成分の多いグループ 2 に対しては原波形での自己相関では周期の算出が難しいため、(3-49) 式により正規化した信号に対して自己相関を求めて、相関値のピーク位置を基準に 1 周期の長さを求めるか、パワー値に基づき一定区間を切り出す。図 3-9 に自動切り出しの動作例を示す。1 周期分のサンプリング点数が正確に算出される。

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^n (\bar{x} - x_i)^2$$

$$x_i' = \frac{x_i - \bar{x}}{\sigma}$$
(3-49)

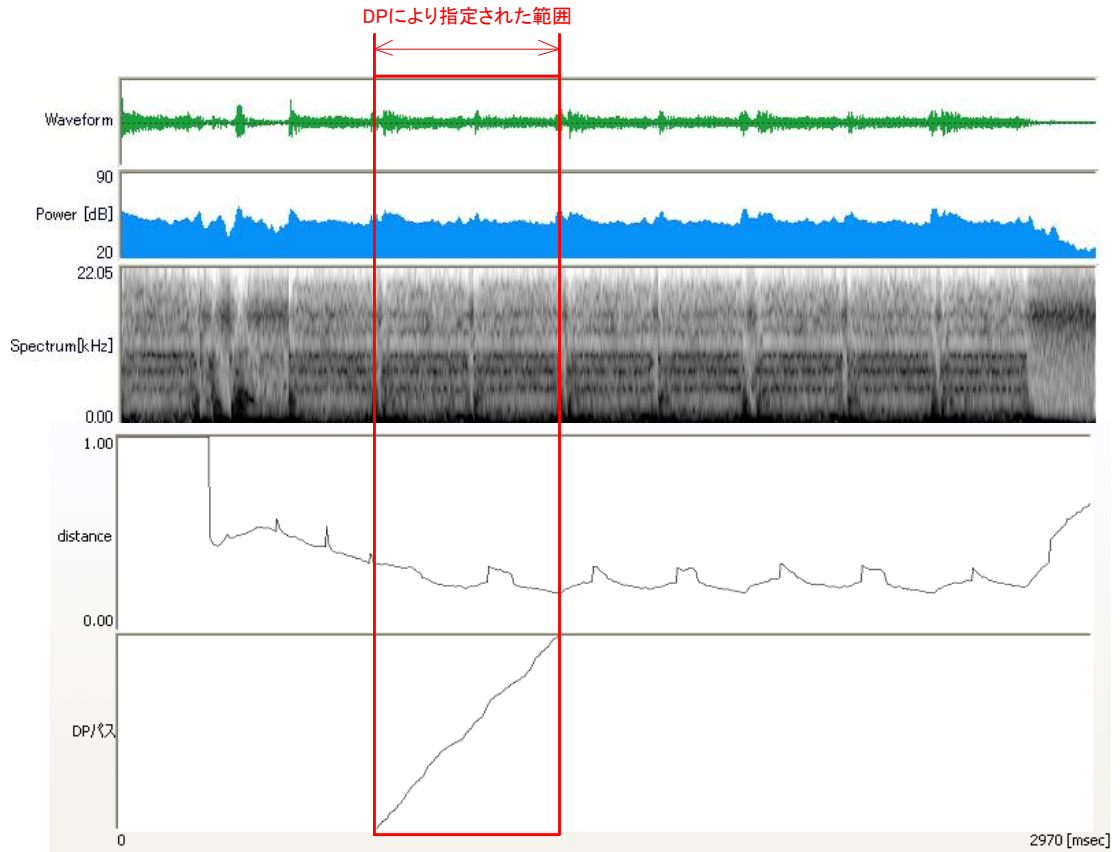


図 3-9 連続 DP マッチングによるリファレンス切り出し

3.6.2 良否判定

入力波形に対して連続 DP マッチングを行い、(3-41)式により求められる正規化累積距離のうち最小のものを最小正規化累積距離として選択する。これを登録済みの q 個のリファレンスに対して繰り返し、 q 個の距離を求める。次に、 k -Nearest Neighbor 法により k 個の距離を選択して平均値を算出することで、リファレンスからの距離を定義する。この距離に対して、閾値処理をして良否判定を行う (図 3-10)。このとき、検査に使用する k の最適値は識別実験により実験的に求めた。

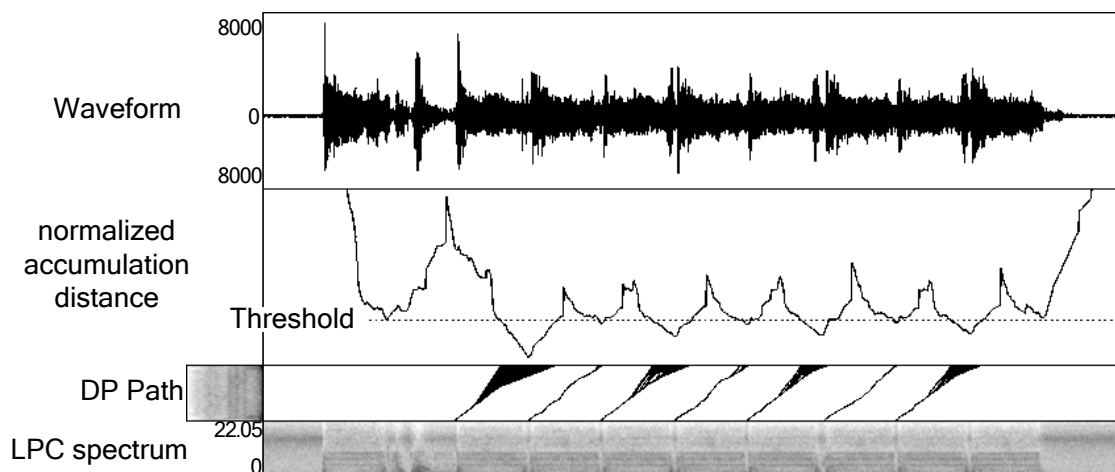


図 3-10 連続 DP マッチングの適用例

3.7 実験データベース

本稿では 4 種類の機械駆動音を対象として、その駆動音データから駆動部の良否を判定する。駆動音データは、サンプリング周波数 44.1 [kHz]、量子化ビット数は 16 [bit] で事前に録音した。これらのデータの良・不良は、予め複数の官能検査員により厳格に確認されている。これらのデータは、生産ライン内に設置された簡易防音環境にて約 1 週間にわたり採取されたものである。それぞれをグループ 1~4 と定義して実験に使用することとした。各グループの特徴について以下に述べる。

3.7.1 データグループ 1

良否判定実験の対象となる 4 種の駆動音データのなかで、図 3-11、図 3-12 に示す種の駆動音データを以下データグループ 1 と定義する。図 3-11、図 3-12 はそれぞれデータグループ 1 の良品データ、不良品データの一例である。データグループ 1 の駆動音は周期約 0.28 [sec] で振幅が時系列に変化する駆動音であり、その周期内で振幅は時系列変化を持つ。実験に用いるデータベースの内訳を表 3-1 に示す。

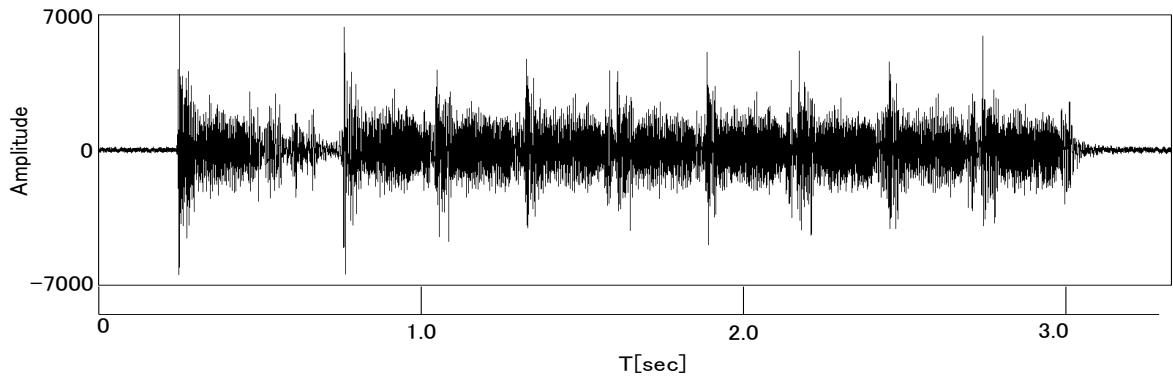


図 3-11 グループ1 良品データ例

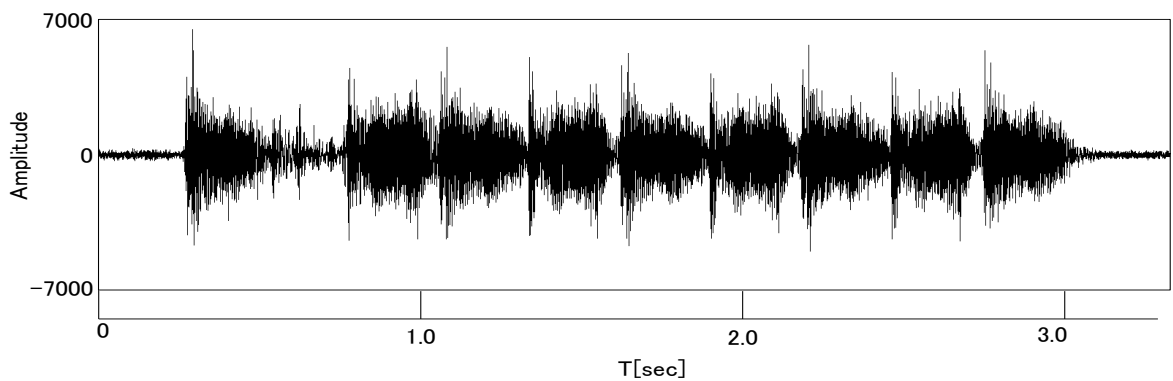


図 3-12 グループ1 不良品データ例

表 3-1 グループ1 データベース内訳

	良品	不良品	合計
サンプル数	41	18	59

3.7.2 データグループ2

良否判定実験の対象となる4種の駆動音データのなかで、図 3-13、図 3-14 に示す種の駆動音データを以下データグループ2と定義する。データグループ2の駆動音は周期約0.55[sec]を持ち、その一周期内において、白色雑音性が強い振幅を繰り返すが、振幅最大値はほぼ一定である。図 3-13、図 3-14 はそれぞれデータグループ2の良品データ、不良品データの一例である。実験に用いるデータベースの内訳を表 3-2 に示す。

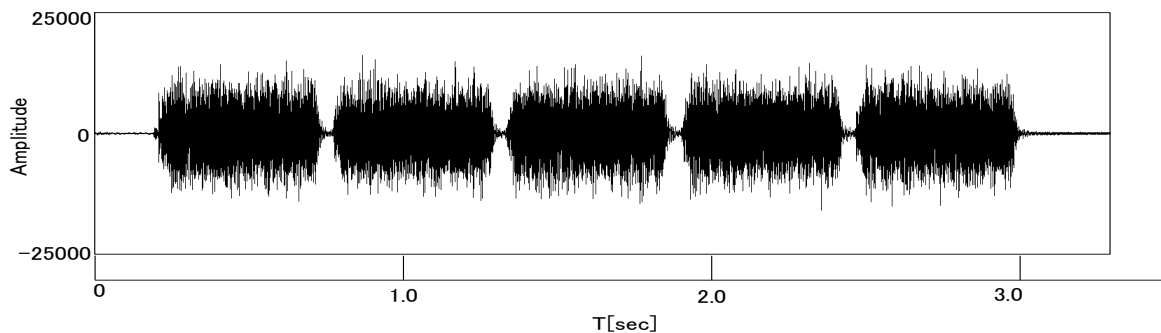


図 3-13 グループ 2 良品データ例

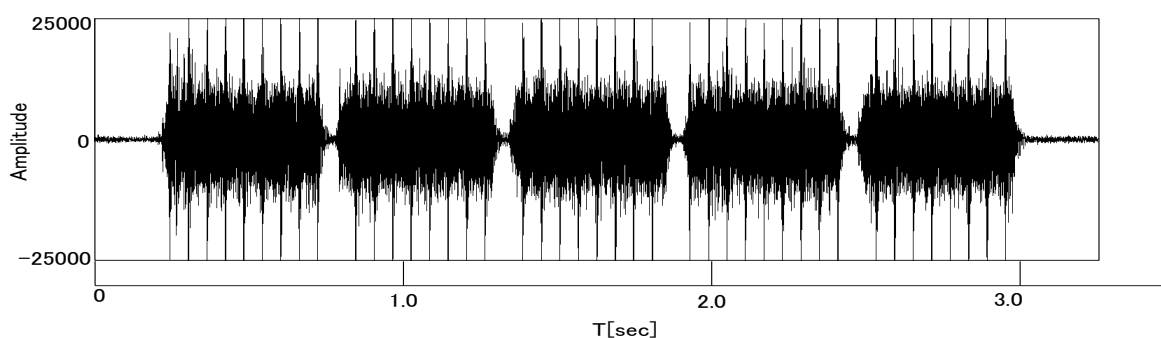


図 3-14 グループ 2 不良品データ例

表 3-2 グループ 2 データベース内訳

	良品	不良品	合計
サンプル数	40	4	44

3.7.3 データグループ 3

良否判定実験の対象となる 4 種の駆動音データのなかで、図 3-15、図 3-16 に示す種の駆動音データを以下データグループ 3 と定義する。データグループ 3 の駆動音は連続駆動音であり、グループ 2 同様に白色雑音性が強い振幅を繰り返すが、振幅の最大値はほぼ一定である。図 3-15、図 3-16 はそれぞれデータグループ 3 の良品データ、不良品データの一例である。実験に用いるデータベースの内訳を表 3-3 に示す。

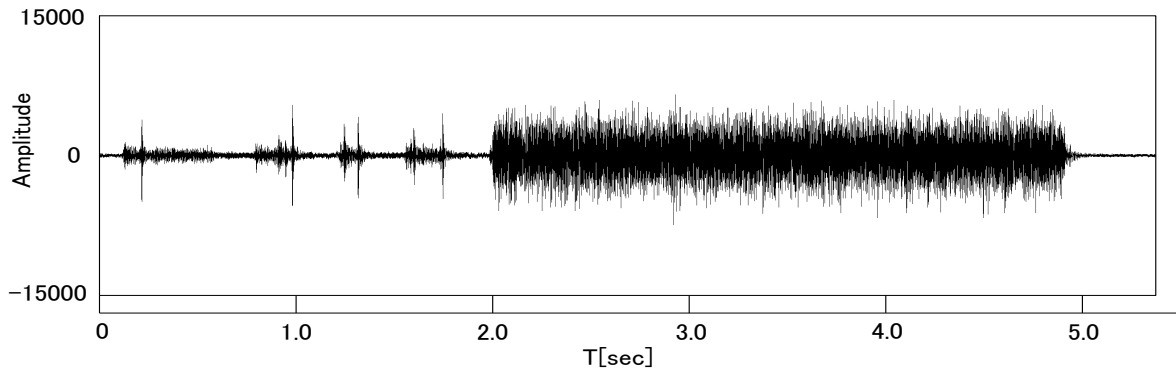


図 3-15 グループ 3 良品データ例

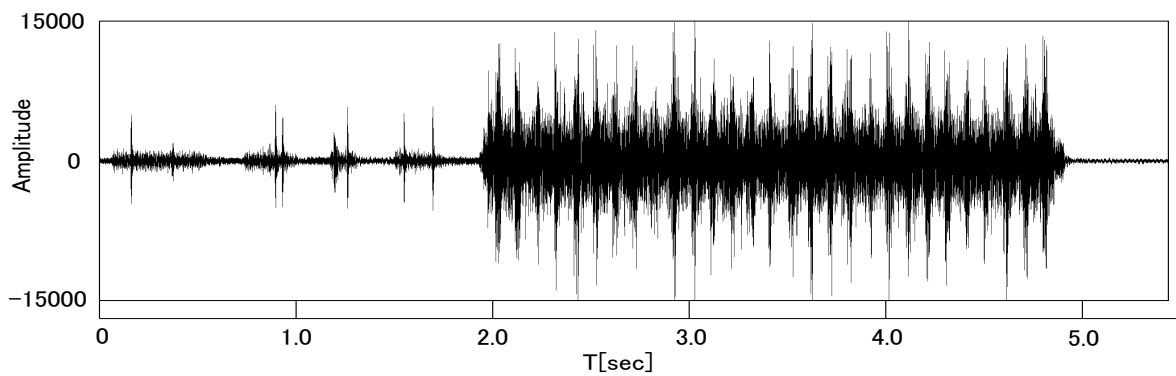


図 3-16 グループ 3 不良品データ例

表 3-3 グループ 3 データベース内訳

	良品	不良品	合計
サンプル数	37	18	55

3.7.4 データグループ 4

良否判定実験の対象となる 4 種の駆動音データのなかで、図 3-17、図 3-18 に示す種の駆動音データを以下データグループ 4 と定義する。データグループ 4 の駆動音は一定の周期を継続し続けるものではないため、駆動音開始から終了までを 1 つの周期ととらえる。図 3-17、図 3-18 はそれぞれデータグループ 4 の良品データ、不良品データの一部である。実験に用いるデータベースの内訳を表 3-4 に示す。

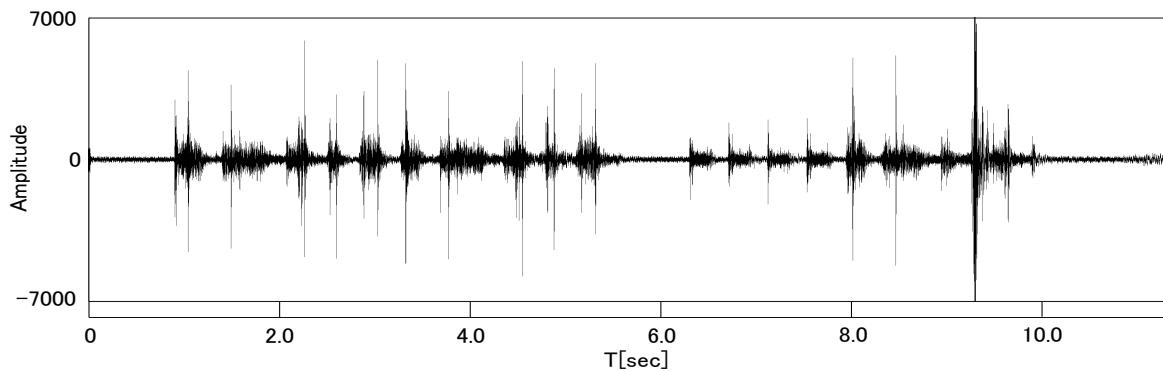


図 3-17 グループ 4 良品データ例

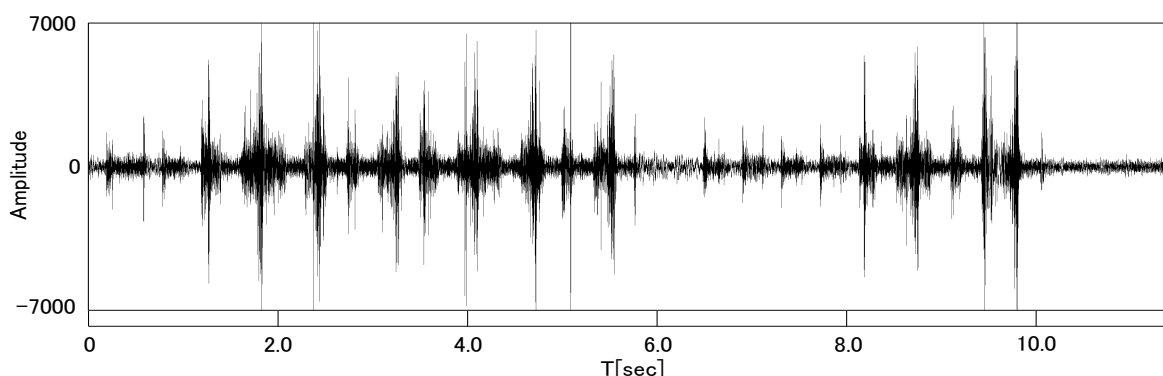


図 3-18 グループ 4 不良品データ例

表 3-4 グループ 4 データベース内訳

	良品	不良品	合計
サンプル数	38	15	53

3.8 識別実験結果

それぞれの手法について、良否判定の識別率を確認した。実験に使用したデータの内訳を表 3-5 に示す。グループ 2 の音については、不良発生頻度が低いため、極端に不良サンプル数が少ない。しかし、今回は様々な音に対する検査を検討することを目的として使用することとした。評価指標としてはすべてのデータに対して同一の指標を用いることとして、データを全て使用した場合

の平均誤り率でシステムを評価した。

表 3-5 実験に使用するサンプル数の内訳

カテゴリ 駆動部	リファレンス (良品)	テストデータ		合計
		良品	不良品	
グループ 1	20	21	18	59
グループ 2	20	20	4	44
グループ 3	20	17	18	55
グループ 4	20	18	15	53

3.8.1 CHLAC による良否判定結果

まず、CHLAC による良否判定の結果を示す。実験条件として、部分空間の寄与率は 0.999 のところまでの次元数を使用することとした。さらに、FFT および LPC の分析フレーム長 5.8[msec]、分析シフト幅 2.3[msec]として、動画シフト幅 s_m は駆動音原波形の 1 周期の半分（半周期）より短い値について検討した。これは s_m を半周期以上とし CHLAC の相関幅（前後フレームの変化時間）を大きくした場合、周期内での変化が計算できなくなり、短時間変動情報を失うからである（サンプリング定理による）。 $w_m > s_m$ とすることでオーバーラップして CHLAC を抽出することもできるが、スペクトル解析時にもオーバーラップをしていることから、 $w_m = s_m$ の条件のもと CHLAC の時間方向積分幅 A を変化させることで同様の効果が得られると考えられる。そこで、今回は $w_m = s_m$ として固定することとして、 w_m と A の 2 変量についての最大認識率を実験的に求めた。同様に、実験の簡易化のため、LPC の線形予測次数についても 14 に限定して実験することとした。上記の実験条件において、良否判別実験を試みた。

各グループに対して、それぞれ FFT、LPC スペクトルから生成した動画を使用して認識実験を行う。グループ 1 に対して FFT を用いた識別実験時の w_m を検討した結果を一例として図 3-19 に示す。各グループの最大識別率を表 3-6 に示す。FFT 動画から抽出した CHLAC による実験では、全グループで 95[%] 以上の識別率を得た。

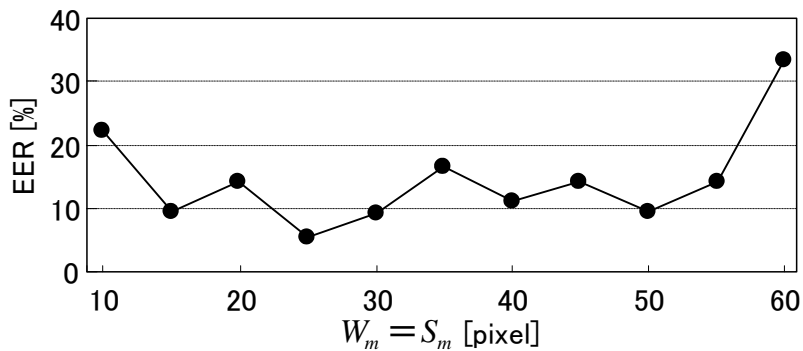


図 3-19 平均誤り率の算出例 (グループ 1 : FFT スペクトルを用いた場合)

表 3-6 識別結果 (CHLAG)

特徴量 駆動部	FFT [%]	LPC [%]
グループ 1	95	95
グループ 2	100	50
グループ 3	100	95
グループ 4	100	80

3.8.2 連続 DP マッチングによる良否判定結果

同様に、連続 DP マッチングによる良否判定の結果を示す。FFT ケプストラム、LPC ケプストラムの両特徴量による良否判定を試みた。判別時の LPC 次数、ケプストラム次数を変動させて実験をした。各特徴量・グループにおいて、最大となる識別率を表 3-7 に示す。全てのグループに対して 97[%]以上の識別率を得た。グループ 1, 4 では、FFT ケプストラムに比べてスペクトルの極を強調することのできる LPC ケプストラムが良好な結果を示した。グループ 2 については、LPC/FFT のどちらをした場合にも同等の結果を示したが、不良品サンプル数が少ないため、今後サンプル数を増やした上で検証を行う必要がある。グループ 3 のデータは白色雑音性の高い駆動音であるため、LPC による分析が適さないことが認識率低下の原因であると考えられる。尚、良否判定処理は機械駆動の動作時間以内に終了しているため、リアルタイムな検査が可能である。

表 3-7 識別結果（連続 DP マッチング）

特徴量 駆動部	FFT [%]	LPC [%]
グループ 1	95	97
グループ 2	100	100
グループ 3	97	94
グループ 4	97	100

3.8.3 パワー値による不良判定

提案法の有効性を確認するため、比較実験としてパワー値を用いた不良判定を試みた。これは騒音計などを用いた異音検出法に該当する。まず、入力音のパワー値を算出して、その最大・平均を特徴量として抽出する。この特徴量を用いて、良品リファレンスデータからのマハラノビス距離を計算して、その値が一定以上であれば不良として判別する。このとき、パワー値の最大・平均は相関が高いパラメータであると考えられるため、逆行列が正しく求められるように修正グラム・シュミット法による直交化を適用した。使用するデータは、提案法と同様である。その結果を表 3-8 に示す。その結果 70~80 [%] 程度の識別率しか得られないことが確認された。このことより、対象データは単なるパワー値では識別できないことが確認された。

表 3-8 識別結果（パワー値の最大・平均）

特徴量 駆動部	FFT [%]
グループ 1	67
グループ 2	79
グループ 3	78
グループ 4	86

3.8.4 各周波数帯のパワー値を用いた不良判定

この手法^[41]では、はじめに検査対象音を一定の分析長の窓で切り出し、切り出し窓をシフトしながら FFT を適用することで、時系列に変化するパワースペクトルを求める。次に、算出したパワースペクトルの時系列方向の分散と平均を求めて、特徴量とする。最後に、特徴量を用いて良品リファレンスセットか

らマハラノビス距離を算出する。このとき、良品リファレンスセットの特徴量に対して乱数で微小なランダマイズを適用してデータセットを仮想的に生成して、良品空間の作成に使用する。

比較実験として、連続 DP マッチングによる実験と同じ分析長 (256 フレーム)、シフト幅 (64 フレーム) で特徴量を算出した場合の識別率を確認した。このときの特徴次元数は平均 128, 分散 128 の計 256 次元とする。良品リファレンスセットの特徴量に対して ± 1.5 [%] のランダマイズを適用して、良品リファレンス 20 個を 40 倍の計 800 個にした。これはマハラノビス距離を求める際に特徴次元数の 3 倍程度が必要であると言われているためである。ランダマイズ幅は、各特徴量の良品リファレンスの標準偏差と平均の比から決定した。なお、CHLAC を用いた提案法と同様に、連続 DP マッチングによるセグメンテーションを前処理として適用した。その結果を表 3-9 に示す。この結果は、CHLAC と同様の特性を示しているが、グループ 1 において CHLAC の方が良好な識別率であることから提案法の有効性が確認された。グループ 1 の識別率が低い理由としては、グループ 1 の駆動音が一定の音ではなく、不良の特徴が微妙な波形の違いとして表れるため、平均や分散という特徴では不良の特徴が表現できないためであると考えられる。

表 3-9 識別結果 (各周波数帯パワー値の平均と分散)

特徴量 駆動部	FFT [%]
グループ 1	87
グループ 2	100
グループ 3	100
グループ 4	100

3.9 ライン検査用ソフトウェア開発

作成した検査システムを生産現場で使用するために、ソフトウェアを作成した。その概要を以下に記す。

アプリケーションには、前述した連続 DP マッチングによる不良判定システムを実装した。ソフトウェアは、主に 2 つのウィンドウより構成されている。

起動時に立ち上がる図 3-20 のウィンドウをメインウィンドウ，図 3-21 の表示メニューより呼び出すことの出来るウィンドウを操作メニューと定義する。生産ライン上で操作しなければならない基本的な部分をメインウィンドウに，その他の現場での微調整など細々とした設定項目を操作メニューにまとめた。尚，現在はオフライン用に設計されているが，音源取り込みを組み込むことでオンライン検査に対応可能である。

メインウィンドウには，生産ラインでの検査に必要な最低限の機能がまとめられており，基本的にはこちらを操作することで検査が可能ないように設計されている。ウィンドウ内の各項目の概要を表 3-10 に示す。

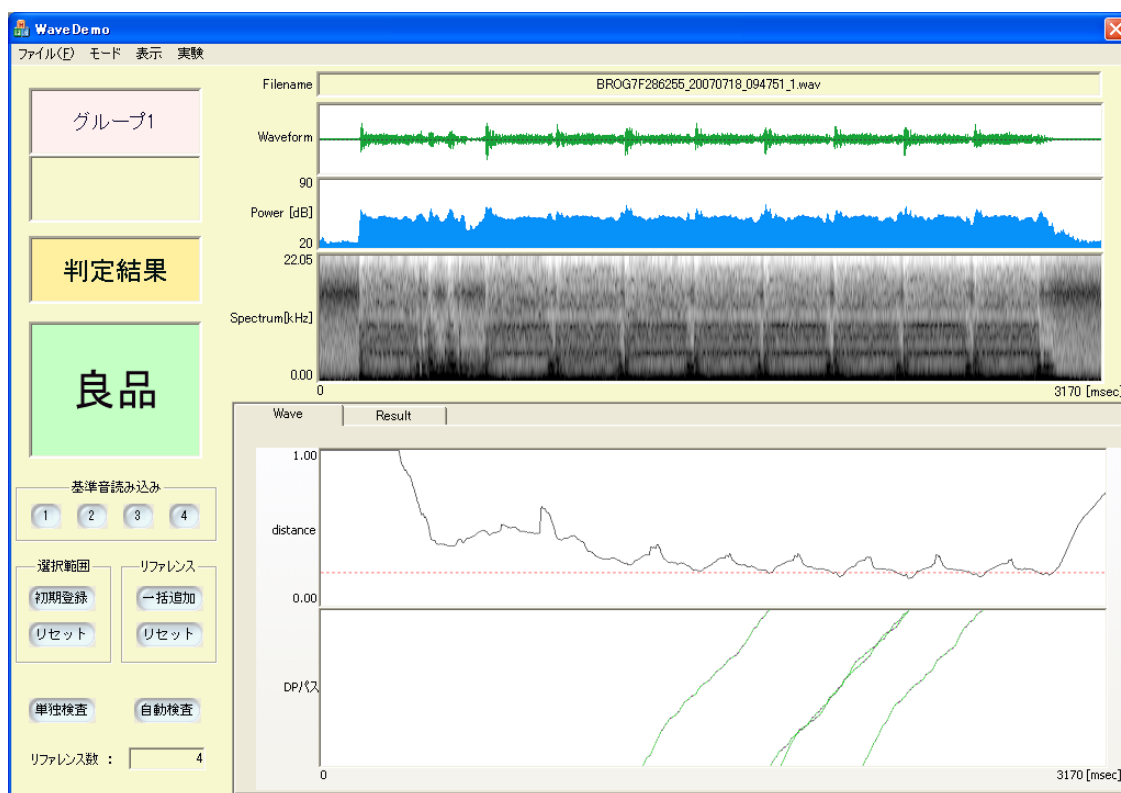


図 3-20 メインウィンドウ

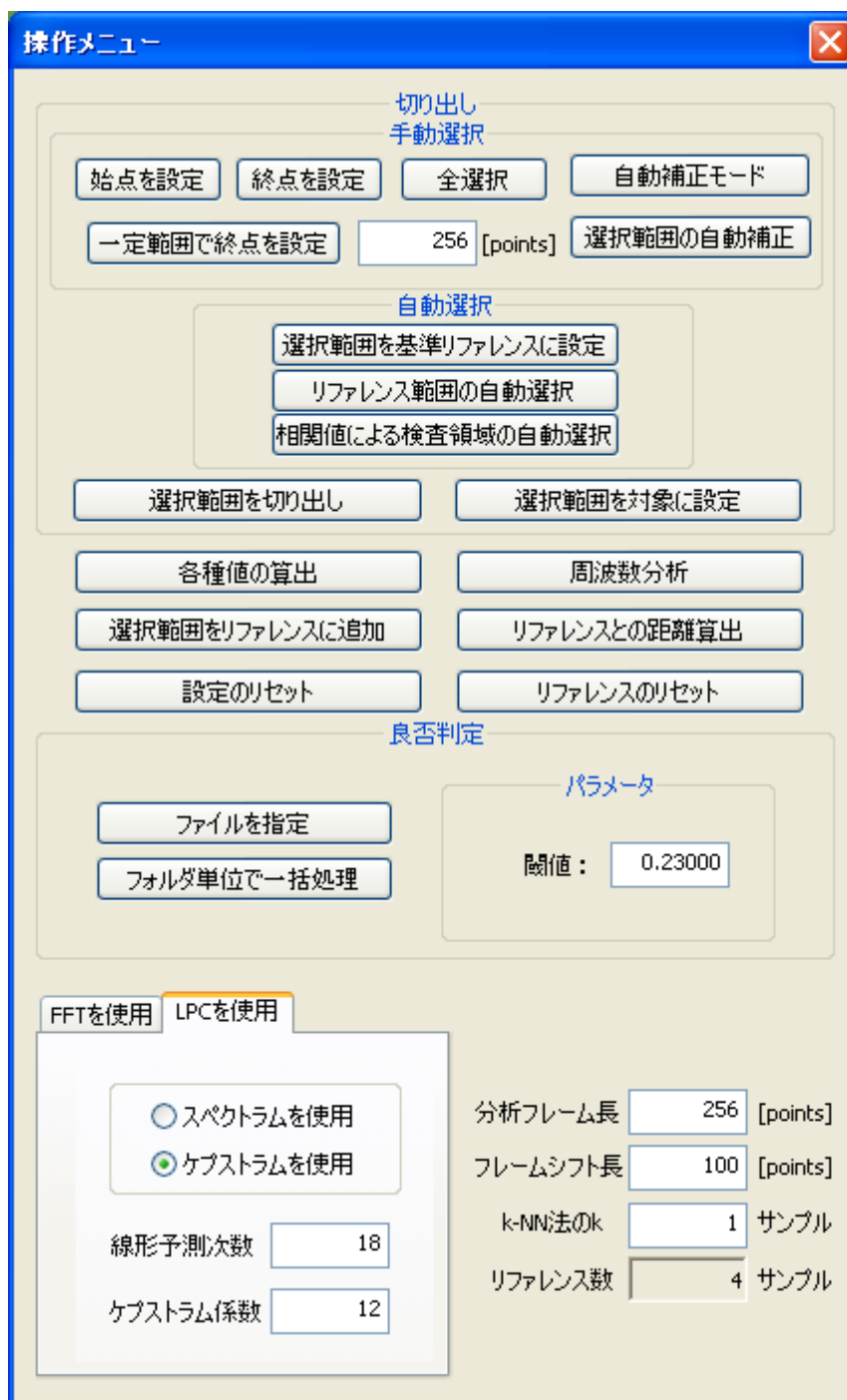


図 3-21 操作メニュー

表 3-10 メインウィンドウ機能一覧

番号	機能
1	検査部位 (モードの表示)
2	良否の正解データ (自動検査デモ時に使用)
3	現在の処理内容
4	良否判定結果
5	基準音読み込み
6	選択範囲を基準リファレンスに登録
7	選択範囲をリセット
8	リファレンスの一括登録 (モード1~4を一括で自動登録)
9	リファレンスのリセット (選択中のモードをリセット)
10	ファイルを読み込んで1つ検査 (選択中のモードで検査)
11	DetaList.ini ファイルのデータを自動検査
12	選択中のモードに登録されているリファレンス数
13	開いているファイル名
14	音源波形
15	音源波形のパワー値
16	LPC スペクトラム
17	マッチング結果表示用タブ
18	計算結果表示用タブ (現在は、距離値のみ表示)
19	連続 DP マッチングで算出した距離値のグラフ
20	DP 経路

基本的な検査は、以下の流れで行うことができる。リファレンスの生成などを出来る限り自動化して、現場の作業員のスキルによらず検査が行えるよう配慮をした。

- Step1 (5) 基準音読み込みで読み込んだ各グループの基準音源を読み込む
- Step2 検査対象とする区間を(14)のウィンドウをクリックして選択
この際、周期性のある音については、自動的に長さが調整される
- Step3 (6) 初期登録で選択範囲を基準リファレンスに設定
- Step4 (8) 一括登録で、基準リファレンスを元に検査用リファレンスを自動生成
- Step5 (11) 自動検査を押すことで、入力音が自動で検査されてゆく

3.10まとめ

本稿では、機械駆動異音の検出手法を提案した。手法として、画像・動画処理と音声認識に基づく手法をそれぞれ提案した。また、4種の駆動音データを対象として実験を行い、官能検査員と同等かそれ以上の性能を確認することにより、提案法の有効性を確認した。

はじめに、CHLACを用いた異音検出法を提案した。連続DPマッチングにより切り出された駆動音のサウンドスペクトログラムを動画像化して、そこから抽出したCHLACで良品の部分空間を作成することにより良否判定を実現した。実験で使用したFFTとLPCを比較すると、LPCに比べてFFTスペクトルを動画像化した際により良い良否判定精度を得た。これは、CHLACを $3 \times 3 \times 3$ という局所領域で求めているため、LPCスペクトルのような局所領域での変化が少ないデータから有効な特徴量が抽出できていないという可能性が考えられる。今後、局所領域のサイズや、画像内での積分範囲に関する実験を行うことで、提案法の有効性をより詳細に検討したい。また、CHLACを用いた判別においてグループ2はLPCでFFTに比べて極端に悪い識別率を示している。これは全4サンプルの不良データのうち半分を誤識別しているためである。グループ2はグループ3と同様に白色雑音性のような駆動音であるため、FFTで良好な結果を示したと考えられる。この手法は、入力音のセグメンテーションを必要とせず、実時間処理が可能であることを確認した。

次に、CHLACでセグメンテーションに使用した連続DPマッチングで算出される距離値に基づく異音検出手法を提案した。その結果、グループ3を除けばLPCケプストラムを用いることで高い良否判定精度を得た。白色雑音性の高いグループ3についても、FFTケプストラムを用いることで高精度に判定可能であることを確認した。

今回は約1週間にわたり採取されたデータを対象に実験をして安定した識別を確認した。今後、長期にわたり採取されたデータについて検討を行い、判定に使用する良品リファレンスの選定法を検討することで、より高精度かつ利便性の高いシステム構築を目指したいと考えている。さらに、今後、駆動音も含む様々なデータに対する有効性を検証することで、より汎用的なシステムを目指したい。本稿で対象としたデータの良品クラスは、比較的類似した特徴を有していたため、連続DPマッチングのみでも高い精度で識別可能なことが確認された。しかし、駆動音には様々な良品パターンが存在する場合も考えられる。

連続 DP マッチングは，様々な良品パターンが存在する場合でもリファレンスを増やすことで対応が可能であるが，リファレンス数の増加に伴い処理時間が増すという欠点がある．そのため良品のパターンが非常に多い場合には，CHLAC で良品の部分空間を構成しておき，連続 DP マッチングを 1 度適用してセグメンテーションした後に，CHLAC 法などで部分空間との比較を行うことで，複数回の連続 DP マッチングによる識別に比べて高速な処理が可能である．

第4章非接触指紋・指静脈認証

4.1 まえがき

近年、情報化社会の発達に伴い、個人認証の需要が増加している。2009年度で214億円であり、さらに年15パーセントの成長が見込まれている。現在、最も普及している個人認証手法として、印鑑、IDカード、パスワード等がある。しかし、これらは常に、盗難、紛失、忘却等の危険性が伴う。そのため、より強固なセキュリティを保持しかつ利便性の高い個人認証システムが求められている。その中で最も注目されている技術として本人の生体的特徴を利用したバイオメトリクス認証が挙げられる。

バイオメトリクス認証には、顔認証、虹彩認証、網膜認証、指紋認証等様々な手法がある^{[48]-[53]}。中でも指紋による個人認証は、照合率の高さやユーザにかかる負担が少ないなどヒューマンインターフェイスの点で特に有効であるため、既に多数の指紋照合法や接触型の指紋入力装置が研究開発されており、システムとして実用化され、広く利用されている。当研究室も接触型の指紋認証システムを開発し、携帯電話、パソコンなど様々な場所で使用されている^{[54]-[62]}。しかしながら、現在利用されている指紋入力装置は、指紋の入力時にセンサデバイスとの接触を要する接触型である。接触型指紋入力装置では、指の乾燥等の影響で、隆線の途切れ、かすれ、つぶれ等、指紋画像の劣化を生じる欠点がある^{[54]-[58][63]}。また、不特定多数の人が装置を使用するため不衛生であり、装置の使用に抵抗を感じる人も少なくない。そのため、これらの問題点を解決する非接触型指紋入力装置が期待されている。

一方、指静脈は指紋同様に、ユーザへの負担が少ない。さらに、指紋と指静脈は共に指から画像取得を行うため、将来的には一度に両者の画像を取得することも可能であると考えられる。指紋・指静脈のハイブリッド認証システムを実現すれば、2つの異なる個人特徴を用いることでより認証率の大幅な向上が見込める。さらに、指紋、静脈ともになりすましの問題が存在するが、それらを同時になりすますのは非常に困難であることから、より安全なシステムを提

供できる．そこで，我々はカメラと近赤外 LED を用いた指静脈撮影装置と認証アルゴリズムを独自に開発した．

本章では，非接触指紋認証システムと指静脈認証システムの撮影システムと認証アルゴリズムについてそれぞれ述べる．

4.2 非接触指紋認証システム

本節では，非接触指紋撮影システムとその認証アルゴリズムについて述べる．

4.2.1 非接触指紋撮影システム

中部大学葛谷研究室と共同で開発した非接触型指紋撮影システムの概要を図 4-1 に，構築した撮影環境の外観を図 4-2 にそれぞれ示す．このシステムは，穴の空いたガイドに指を置き，両側から 3 波長(赤，青，緑)LED を当てることによって生じる指紋の拡散反射光を CCD カメラで読み取ることで指紋画像を撮影するものである．使用した 3 波長 LED と CCD カメラの外観を図 4-3 と及び図 4-4 使用に，外観を

表 4-1 と表 4-2 にそれぞれ示す．取得される指紋画像は，グレースケール(8bit 階調)の画像である．

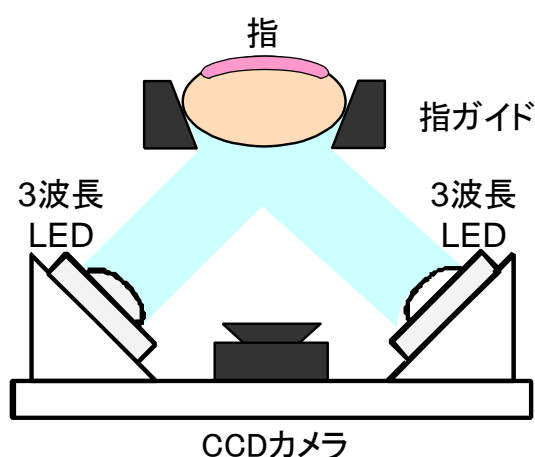


図 4-1 非接触型指紋入力装置の概要

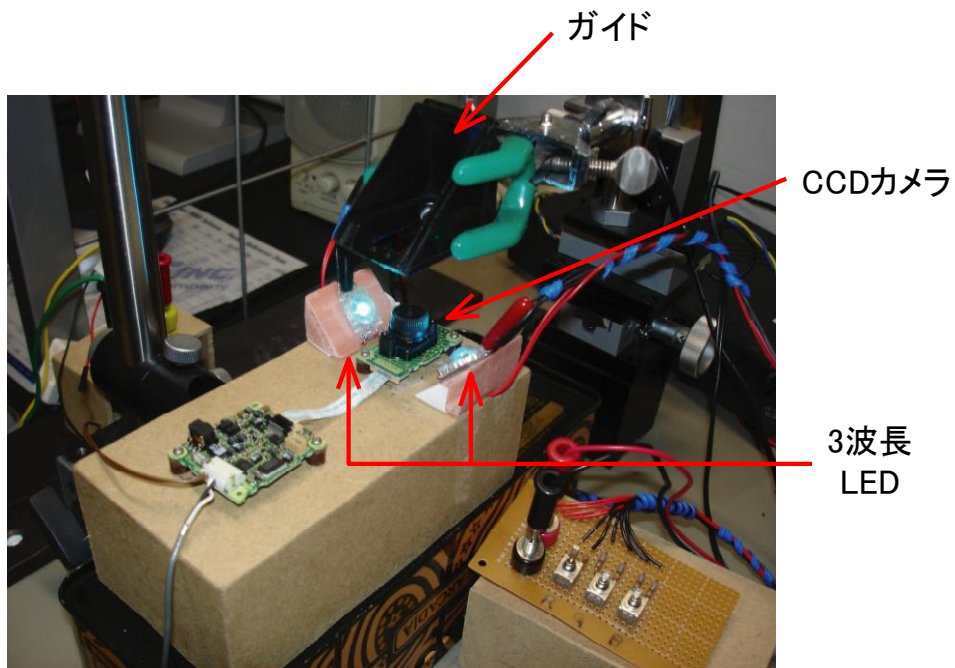


図 4-2 非接触型指紋入力装置の概観



図 4-3 3波長LEDの概観



図 4-4 カメラの概観

表 4-1 3 波長 LED の仕様

型番	EP204K-35RGB
波長	約 425nm～約 670nm
使用温度範囲	-40℃～100℃
保管温度範囲	-40℃～100℃
標準カンデラ	2.6cd～3.7cd
電圧	赤 2V, 緑 3.5V, 青 3.5V
電流	1mA～20mA

表 4-2 CCD カメラの仕様

使用温度範囲	-20℃～+60℃
動作温度範囲	-20℃～+60℃
撮影方式	色差線順次方式 フィールド蓄積
走査方式	2 : 1 インターレース 525 本 60 フィールド
同期方式	内部同期
画像出力	EIA 準拠 (白黒)
水平周波数	fH=15.734kHz
垂直周波数	fV=fH×2/525 (59.94Hz)
画像素子	補色フィルター内臓 1/4 インチ 510HCCD 固定撮影素子出力
電子シャッター	固定
電源電圧	5.0±1.0V
焦点距離	f=5.0mm
F 値	1 : 2.85
画角	水平 38° 垂直 29° 対角 48°
撮影範囲	H=21.5mm、W=17mm

4.2.2 接触型指紋センサとの画像の特徴比較

接触型指紋入力装置と非接触型指紋入力装置の特徴を比較する。接触型の指紋入力装置は様々な種類があるが、今回使用する非接触型指紋入力装置が光学式であるため、光学式接触型指紋入力装置との比較を行う。接触型と非接触型の特徴を表 4-3 に示す^{[64]-[66]}。さらに、指紋画像の撮影例を図 4-5 に示す。

表 4-3 接触型と非接触型の特徴

	接触型	非接触型
照明条件	均一	変化
歪み	あり	なし
かすれ	あり	なし
残留指紋	あり	なし
背景	なし	あり
血流の影響	なし	あり



接触型指紋センサ



非接触型指紋センサ

図 4-5 指紋の撮影例

接触型の指紋入力装置は、指を装置に押し付けるため平らになり、明るさがほぼ均一である。非接触型の指紋入力装置では、指の形状が平らでないため、部位により明るさが変化する。また、接触型は接触させる際に指を押し付けることで指紋画像が歪む、指の乾燥により指紋画像が不鮮明になる(かすれ)欠点がある。しかし、非接触型の指紋入力装置ではこれらの問題は無い。

非接触型の指紋入力装置では、指紋と共に背景画像が取得され、接触型の指紋入力装置では、指紋と残留指紋(センサ部に残留した以前の指紋画像)が取得される。また、非接触型指紋入力装置の指紋画像は、指内部の血流の影響で微小領域ごとに輝度値が変化する。

このように、接触型と非接触型では特徴が異なるため、接触型で用いてきた従来の指紋画像前処理とは異なる手法が必要である。

4.2.3 認証アルゴリズム

接触型指紋認証においても実績のある周波数解析法に基づく認証を検討した。フローチャートを図 4-6 に示す。特徴抽出後、端点フリーの DP マッチングにより照合を行う。

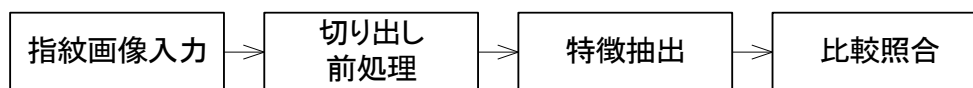


図 4-6 認証フロー

4.2.4 特徴抽出

切り出しを施した後の画像から、横方向 1 ライン分の輝度値情報を取り出す。取り出した情報を音声波形の信号とみなし、周波数領域の情報である、LPC スペクトルに変換し、指紋の周期的スペクトル情報を取得する。更に、LPC スペクトルから LPC ケプストラムに変換することで、スペクトルピークを平滑化する^{[59][60]}。特徴抽出の流れを図 4-7 に示す。

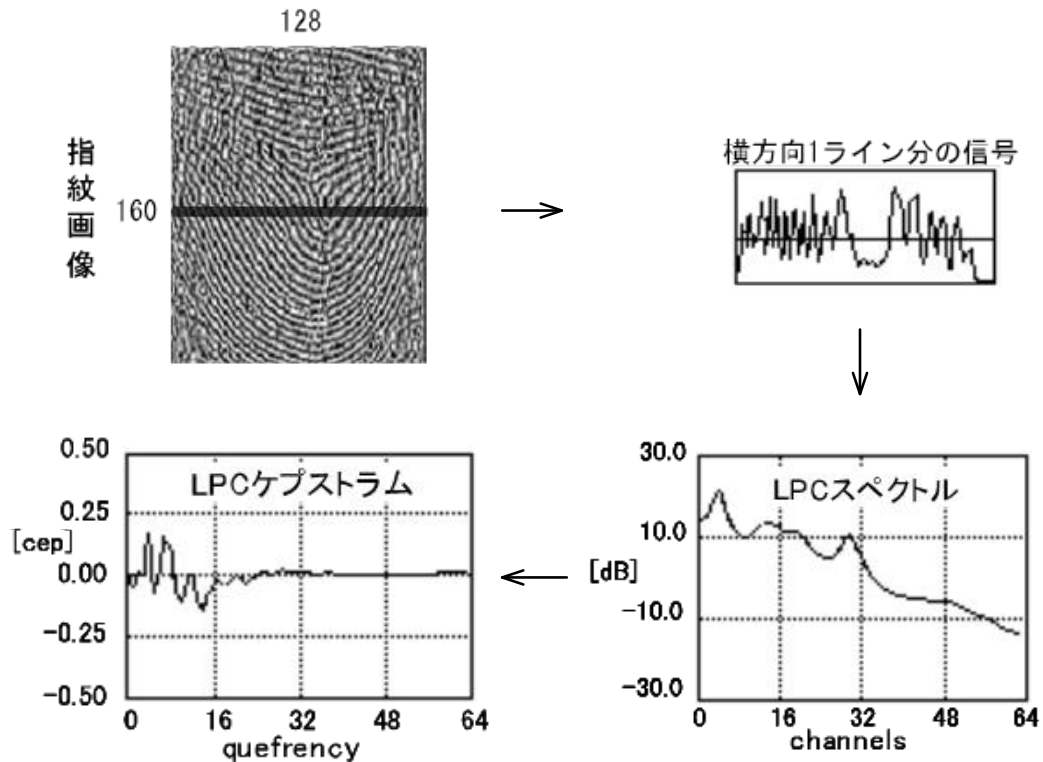


図 4-7 特徴抽出の流れ

4.2.5 端点フリーDP マッチング

DP マッチングは音声認識の分野で広く用いられているアルゴリズムであり、特に端点フリーDP マッチングはパターンの伸縮および位置ずれを考慮した識別を行うことができるという特徴を持つ。

DP マッチングで用いるパスは、図 4-8 に示す対称型 DP パスを用いる。

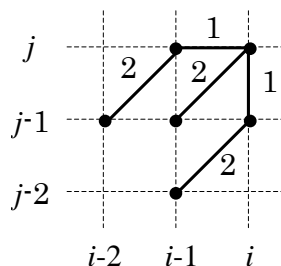


図 4-8 対称型パス

DP マッチングにより識別処理を行うとき、あらかじめ登録した登録パターンと識別対象の入力パターンとのマッチングをとる。このとき、パターン同士の

フレーム間距離を求める．まず，図 4-9 に，入力パターンに対して端点フリー DP マッチングを行う模式図を示す．

このとき，フレーム間距離 $d(i, j)$ は(4-1)式のように定義する．

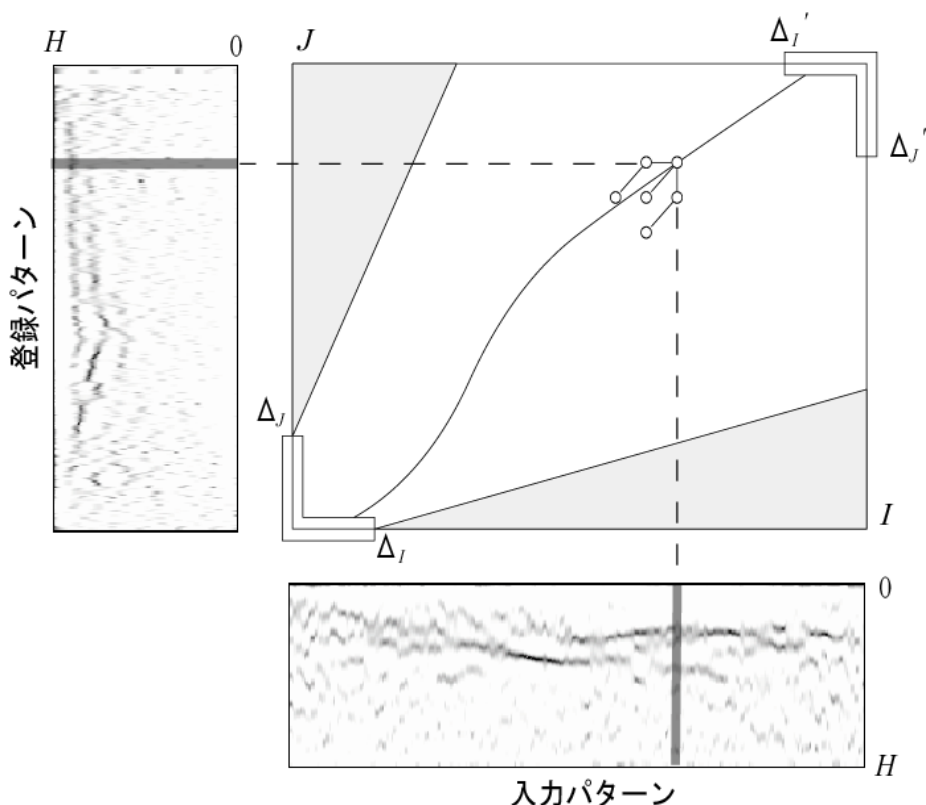


図 4-9 端点フリー DP マッチング

$$d(i, j) = \frac{1}{H} \sum_{k=0}^{H-1} (x_{ref}(j, k) - x_{tes}(i, k))^2$$

$(0 \leq i < W_{tes}, 0 \leq j < W_{ref})$

(4-1)

x_{ref}, x_{tes} : 登録，入力パターンのピクセル濃淡値

W_{ref}, W_{tes} : 登録，入力パターンの幅

H : パターンの高さ

図 4-8 に示す対称型パスを用いた DP マッチングにおいては，(4-2)式による最小累積距離 $g(i, j)$ を計算する．

$$g(i, j) = \min \begin{cases} g(i-1, j-2) + 2d(i, j-1) \\ g(i-1, j-1) + d(i, j) \\ g(i-2, j-1) + 2d(i-1, j) \end{cases} + d(i, j) \quad (4-2)$$

始端点フリー区間においては、(4-3)式に示すような $(j+1)$ および $(i+1)$ で重み付けされたパスを用いる。

$$\begin{cases} g(0,0) = 0 \\ g(1, j) = g(0,0) + (j+1)d(1, j) \\ g(i,1) = g(0,0) + (i+1)d(i,1) \end{cases} \quad (4-3)$$

$$(1 \leq i \leq \Delta_I, 1 \leq j \leq \Delta_J)$$

終端点フリー区間においては、(4-4)式および(4-5)式で示される正規化累積距離 $G(i, j)$ の最小値を正規化最小累積距離 D として、識別の距離尺度に用いる。

$$D = \min \begin{cases} G(i, J), I - \Delta_I' \leq i \leq I \\ G(I, j), J - \Delta_J' \leq j \leq J \end{cases} \quad (4-4)$$

$$G(i, j) = \frac{g(i, j)}{i + j} \quad (4-5)$$

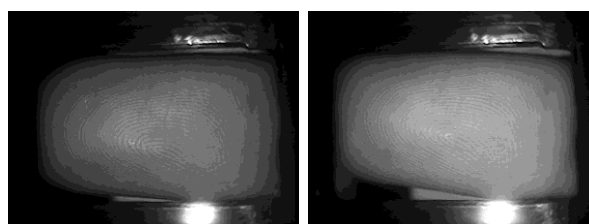
(4-2)式により定義される対称型パスを使用する場合、識別処理可能な動体パターンに制約が生じる。この対称型パスは、図 4-9 のマップ中に示される灰色領域に最小累積距離のパス軌跡を描かない。すなわち、入力データは登録データのフレーム長の 1/2 から 2 倍の範囲の長さであるときに限り、パターン間の距離計算が可能となる。

4.2.6 LEDの色と電圧の検討

撮影装置に使用した3波長LEDは、赤、青、緑の3色を組み合わせて照射することが可能である。すなわち、光源色として、赤(r)、青(g)、緑(b)、赤と青(br)、赤と緑(rg)、青と緑(bg)、赤と青と緑(rgb)の7種類を使用できる。それぞれの組み合わせで取得したデータを用いて照合実験を行うことで、認証に適する光源色を選択した。このとき、LED回路に印加する電圧を3.5V、4.0V、4.5Vの3通りとして、明るさに対する違いも検討した。両手の人差し指と中指の計4指を各10枚ずつ、27人分のデータを取得した。取得したデータベースの内訳を表4-4に示す。さらに、印加する電圧を変えて撮影した画像を図4-10、LEDの組み合わせを変更して撮影した画像を図4-11にそれぞれ示す。

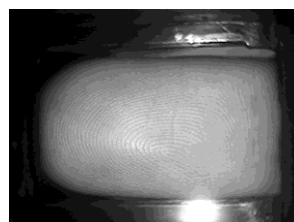
表 4-4 画像データベース

光源の組み合わせ	7種類 [r, g, b, rg, br, gb, rgb]
LEDの電圧	3種類 [3.5V, 4.0V, 4.5V]
指の本数	27人×4指=108指
各指の枚数	10枚



(a) 3.5 [V]

(b) 4.0 [V]



(c) 4.5 [V]

図 4-10 電圧による画像の違い

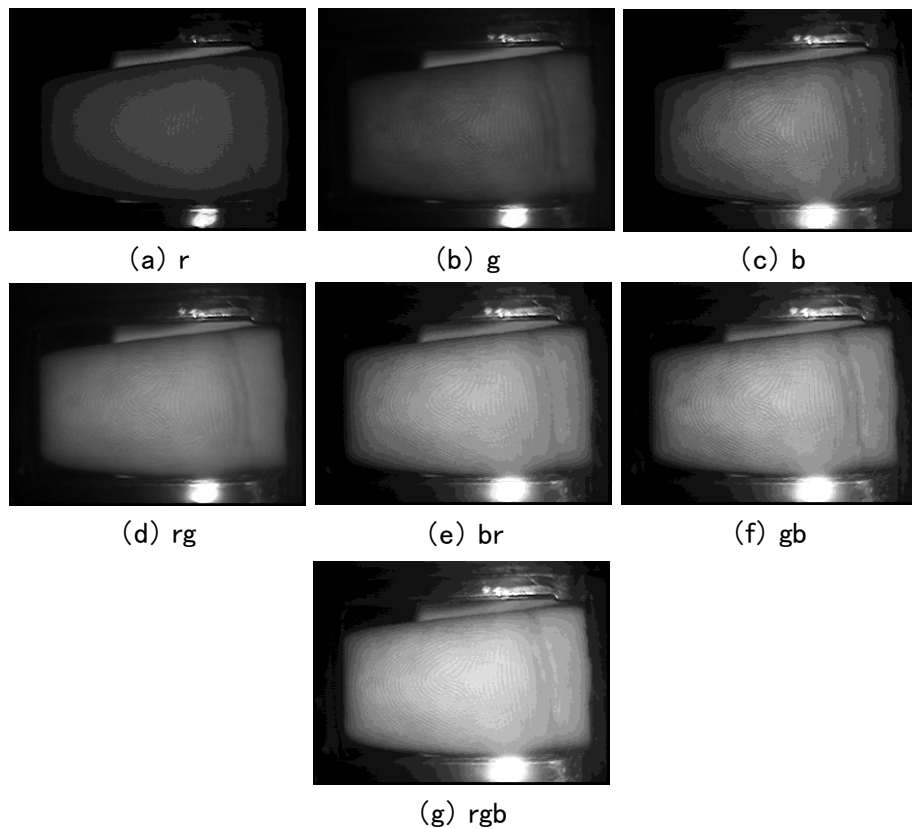


図 4-11 LED の組み合わせによる画像の違い

照合の前処理には、ヒストグラムの伸長化を適用した。その結果を表 4 に示す。すべての組み合わせについて、電圧 4.0 [V] が最良の照合率を示した。また、LED の組み合わせは gb で最も EER の際の照合率が高いことが確認された。これは、人の指の補色に相当する色を使用することで、指紋のコントラストが高い画像が取得できたためであると思われる。以降、波長 gb、印加電圧 4.0 [V] で撮影された画像を使用して前処理を検討することとする。

表 4-5 認証結果

色 \ 電圧	3.5 [V]	4.0 [V]	4.5 [V]
r	75.0%	77.7%	75.7%
g	80.6%	83.8%	82.3%
b	84.8%	88.9%	84.5%
rg	82.5%	87.8%	83.5%
br	84.8%	89.5%	84.9%
gb	86.2%	90.7%	86.1%
rgb	86.7%	90.1%	85.7%

4.2.7 切り出し処理

まず、位置合わせのための切り出し方法について検討した。切り出す際の基準位置として指紋の中心位置を検出することで、画像中から指紋領域を切り出す。指紋の中心付近は複雑な形をしており、輝度値の変化回数が多くなることが予想される。そこで、輝度値の振幅回数により指紋の中心を検出する。輝度値の振幅回数として、移動平均による包絡線との交差回数を用いる。まず、画像から横1ラインの輝度値の情報を取得する。1ラインの輝度値を見ると、指の端部分と中央部分では照明条件が異なるため、山形のオフセットが存在している。そのため、ある直線との交差回数を算出しても指紋の複雑さが反映されない。そこで、取得した1ラインの情報に移動平均を取ることで包絡線を求め、求めた包絡線と取得データの交差回数を算出する。図4-12に算出した交差回数の分布を示す。交差回数の分布を見ると、指紋の中心付近では交差回数の回数が多いことが分かる。よって、交差回数の最大値を求めることで、指紋の中心点の検出を行う。検出した指紋の中心位置を基準に、一定サイズを切り出し照合に使用する。このときの、切り出しサイズについて検討した。このとき、縦横のサイズをそれぞれ変動させて、照合率を比較した。その結果、切り出しサイズ 128×160 [pixel] のときに最も照合率が高くなることが確認された。

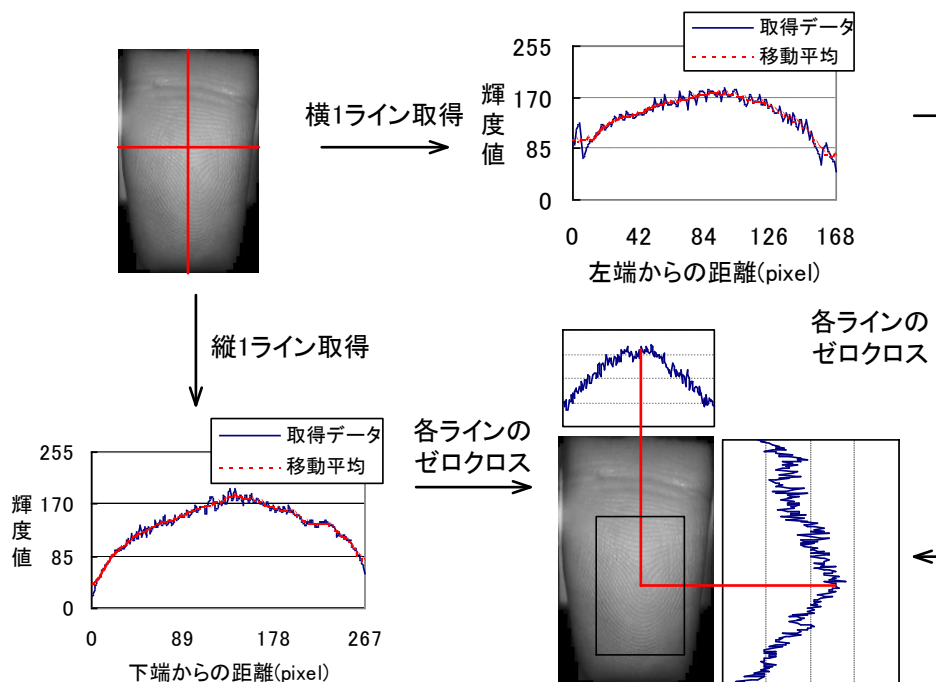


図 4-12 中心点の検出と切り出し

4.2.8 オフセットとシェーディングの除去

非接触型指紋入力装置から取得される指紋画像に生じるシェーディングは、指の置き方などにより変動するため、称号率に悪影響を及ぼすと考えられる。そのため、照合前にこのような影響については画像処理により除去する必要がある。今回は、指紋画像から横1ラインの情報を取得して包絡線を算出して、データから包絡線成分を減ずることでシェーディングを取り除く手法を検討した。包絡線の算出には、移動平均を用いる方法とフーリエ変換を用いる方法をそれぞれ試みた。フーリエ変換を用いる方法では、空間周波数領域において*i*次成分以上の高周波成分を取り除き、逆フーリエ変換することで包絡線を算出する。(4-6)式の高周波領域を単純に切断するフィルタ、(4-7)式の線形に切断するフィルタの2種類のフィルタによる切断法を検討する(図4-13)。照合実験により、各種法を評価した。その結果、EERの際の照合率を比較した場合、フーリエ変換を用いて11次成分までを線形に切断して求めた包絡線が最も高くなることが判明した。このときの包絡線の例を図4-14に示す。この処理により、照合率はオフセットとシェーディング除去前の92.4%から96.4%に向上した。

$$f(x) = \begin{cases} 1 & (0 \leq x \leq i) \\ 0 & (i < x \leq \pi) \end{cases} \quad (4-6)$$

$$f(x) = \begin{cases} -\frac{x}{i} + 1 & (0 \leq x \leq i) \\ 0 & (i < x \leq \pi) \end{cases} \quad (4-7)$$

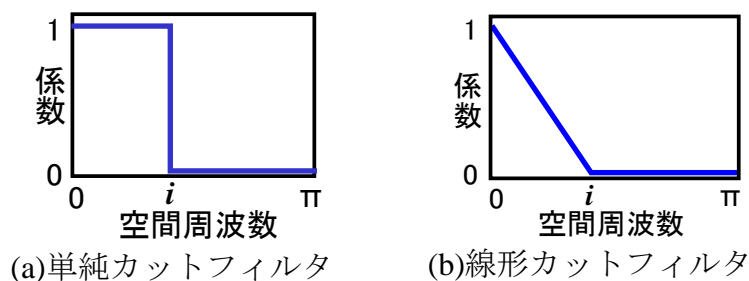


図 4-13 ローパスフィルタ

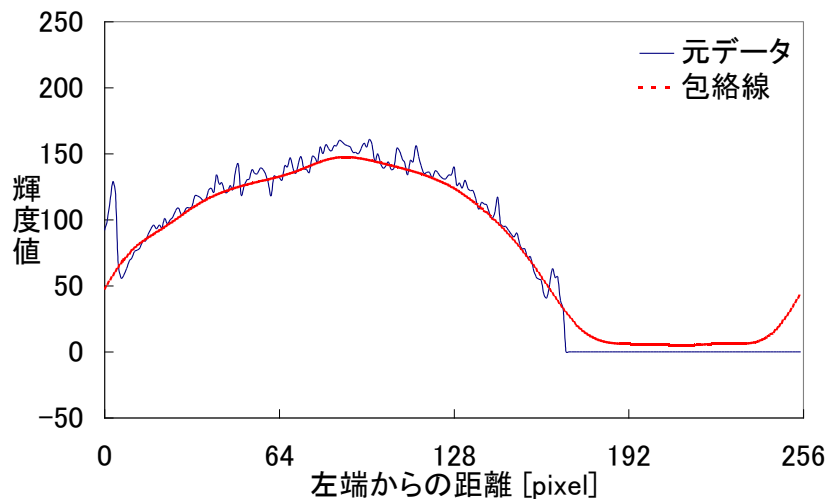


図 4-14 11 次成分までで線形にカットした結果

オフセット除去処理の例を図 4-15 に示す。

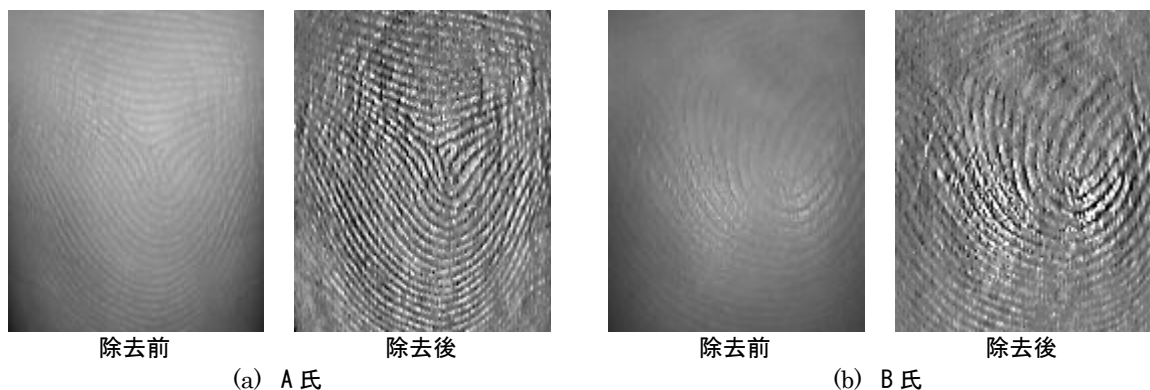


図 4-15 オフセットを除去した例

4.2.9 エッジ抽出

非接触型指紋入力装置から得られる指紋画像は、血流の影響により、部位ごとに皮膚の輝度値が変化する(図 4-16)。血流による輝度変化は、指紋の情報ではないため、除去を行うことで照合率が向上すると思われる。そこで、画像にラプラシアンフィルタを掛けることで指紋の隆線を抽出した。処理後の画像と、その 1 ラインの輝度分布を図 4-17 に示す。また、2 名分の原画像とエッジ抽出画像を図 4-18 に示す。この処理ではノイズ成分も強調されるため、距離計算を行う前に平滑化処理を施した。このとき、ガウシアンフィルタのサイ

ズが 9×9 , $\sigma=1.2$ で最良の結果となり, 照合率は 96.4% から 98.1% に改善することが確認された.



図 4-16 血流の影響例

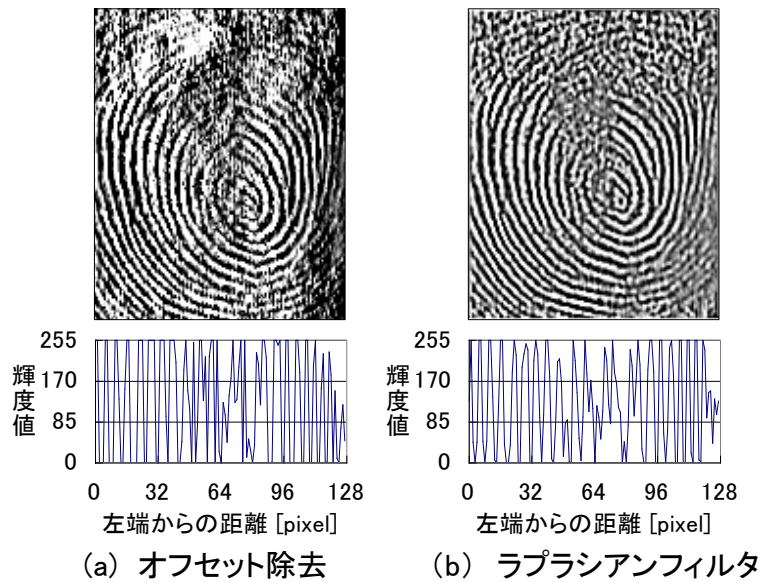


図 4-17 エッジ抽出結果

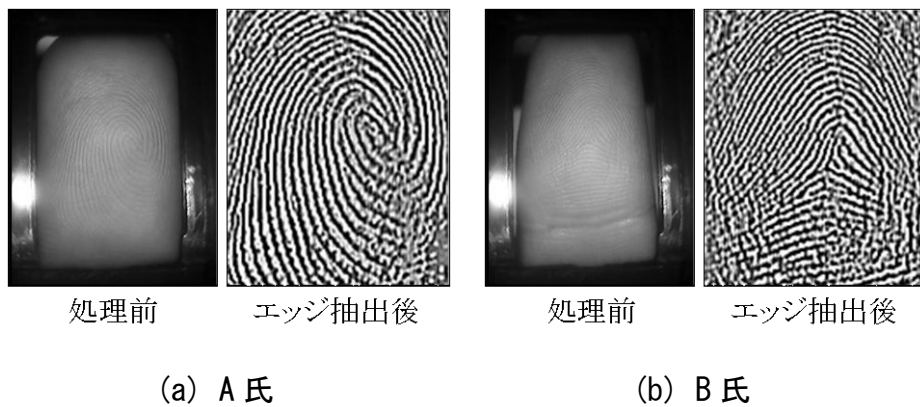
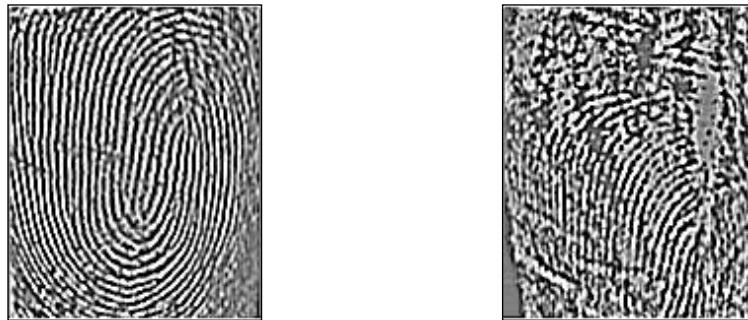


図 4-18 隆線抽出例

4.2.10 考察

同一人物の指紋の中で、誤識別されたデータの一部に図 4-19 の位置ずれが発生していることが確認された。このデータの位置ずれについて、手動で位置補正を行い、照合した。しかし、位置補正後のデータも他人として棄却された。これは、右のデータの指紋領域が潰れたことが原因だと考えられる。その他の誤識別例としては、図 4-20 のような隆線の潰れも認められた。今後、リファレンス登録時の画質判定処理などが必要だと考えられる。



縦横方向の位置ずれ : (x = 2, y = 41 [pixel])

図 4-19 本人同士のデータで距離値の大きな例

	切り出し後	オフセット除去後	エッジ抽出後
データ A			
データ B			

図 4-20 隆線がつぶれる例

4.3 指静脈認証システム

本節では、指静脈撮影システムとその認証アルゴリズムについて述べる。

4.3.1 非接触指紋撮影システム

実験に使用する静脈撮影画像を取得するため、撮影装置を構築した。静脈撮影は近赤外光を用いることにより可能である。近赤外光には生体組織を透過しやすく、血中ヘモグロビンで吸光しやすい特性がある。従って、指に近赤外光を照射することで静脈が暗く、それ以外が明るい画像を取得できる。一般的に、撮影方式には図 4-21 のような指を透過した光である透過光を撮影して画像を取得する透過型装置と反射光を撮影する反射型撮影装置の二つの方式が存在する。それぞれの特徴を表 4-6 に示す。本研究では実験用として鮮明な静脈画像が得やすい透過型の撮影装置を構築した。構築した撮影装置と撮影の様子を図 4-22 に示す。

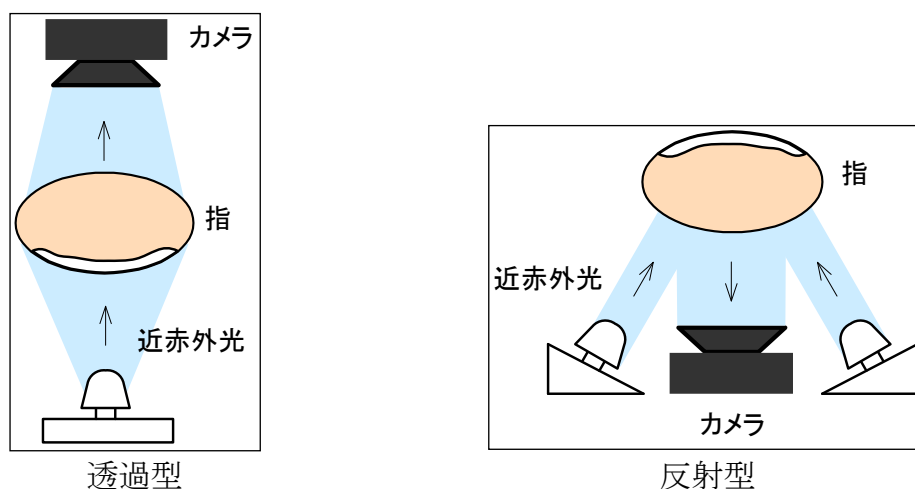


図 4-21 静脈撮影の方式

表 4-6 各静脈撮影方式の特徴

透過型	反射型
反射型に比べ、鮮明な静脈が撮影しやすい	光源とカメラの位置が同方向であることから装置の小型化が可能
皮膚表面状態の影響を受けづらい	光の透過率の影響を受けづらい

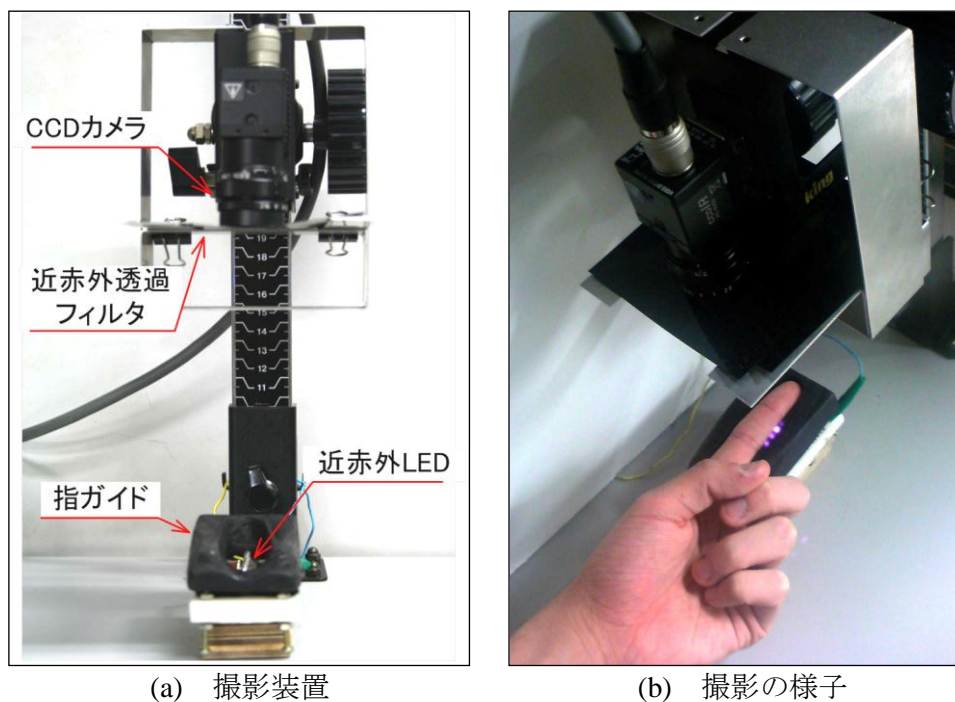


図 4-22 撮影装置図と撮影図

本撮影装置では CCD カメラを用いて撮影を行う。カメラの仕様を表 4-7、レンズの仕様を表 4-8 に示す。図 4-23 の特性のように、近赤外領域にも感度を持つ為、近赤外 LED を用いた指静脈の撮影に適していると考えた。

表 4-7 カメラ仕様

画像	
型番	XC-EI50
撮影素子	1/2 IT 方式 CCD
有効画像素数[pixels]	768×494
外形寸法(mm)	29(W)×29(H)×32(D)
電源電圧(V)	DC 12

表 4-8 レンズ仕様

<p>画像</p>	
<p>型番</p>	<p>VS-LD10</p>
<p>焦点距離[fmm]</p>	<p>10.27</p>
<p>倍率可変範囲</p>	<p>0.02x~0.1x</p>
<p>最小撮影距離[mm]</p>	<p>89.3</p>
<p>適合カメラ/マウント</p>	<p>1/2 以下, C マウント</p>

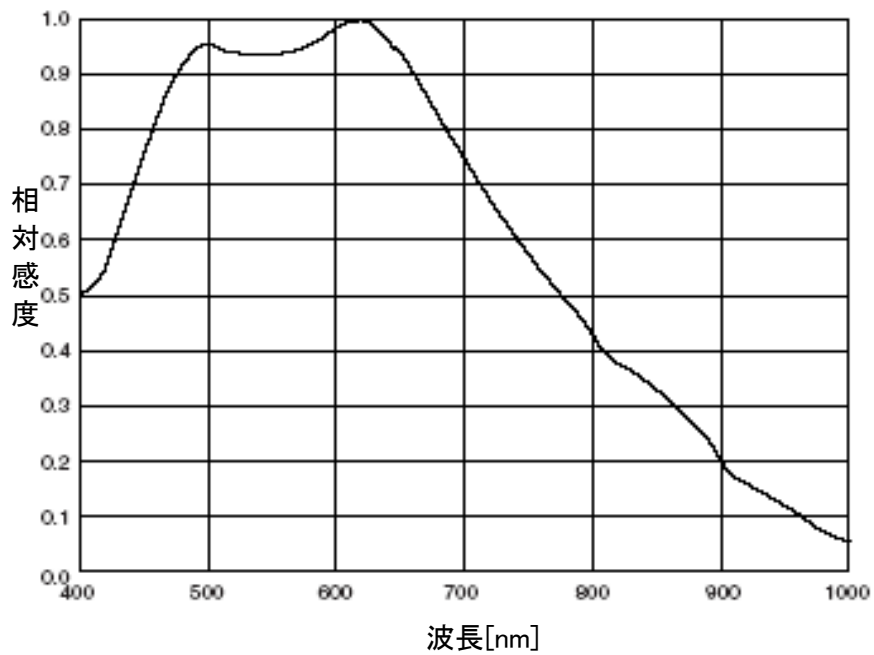


図 4-23 カメラの分光特性

静脈撮影には近赤外光を指に照射する必要がある。そこで、光源には近赤外 LED を使用する。LED を光源とすることで安価な装置の構築が可能である。使用する光源波長により、静脈の鮮明さは変化するため、静脈撮影に適した光源波長を選定する必要がある。本研究では波長 9 種類 (700, 760, 810, 850, 870, 890, 910, 940, 970) の中から静脈撮影に適する光源波長を選定した。選定に関しての詳細は次章において述べる。ここでは使用する LED の仕様を示す。波長 9 種類の LED の仕様を表 4-9 に示す。

表 4-9 LED 仕様

型番	ピーク波長 [nm]	順電圧 [V]	半値幅 [nm]	放射力 [mW]	半減角 [°]	放射強度 [mW/sr]
L700-33AU	700	2.0	30	13.0	±18	30
L760-33AU	760	1.8	30	19.0	±18	40
L810-33AU	810	1.7	35	22.0	±18	40
L850-33UP	850	1.5	40	26.0	±18	50
L870-33UP	870	1.5	40	26.0	±15	50
L890-33AU	890	1.4	40	15.0	±18	20
L910-33	910	1.4	45	12.0	±18	25
L940-33V	940	1.3	50	18.0	±18	20
L970-33	970	1.3	55	5.5	±18	20

使用する LED の個数により光源の明るさ、撮影可能な範囲は変化する。LED の個数が少ないと明るさが足りないことや光が当たる面積が小さいことから鮮明な静脈パターンを得ることはできない。逆に、大きくすると照射面積を大きくすることができ、さらに光量がふえることで明るい画像を撮影することができる。実際に、各 LED の個数で静脈を撮影して、それらの違いを確認した。図 4-24、図 4-25 に各 LED の個数での静脈撮影画像を示す。

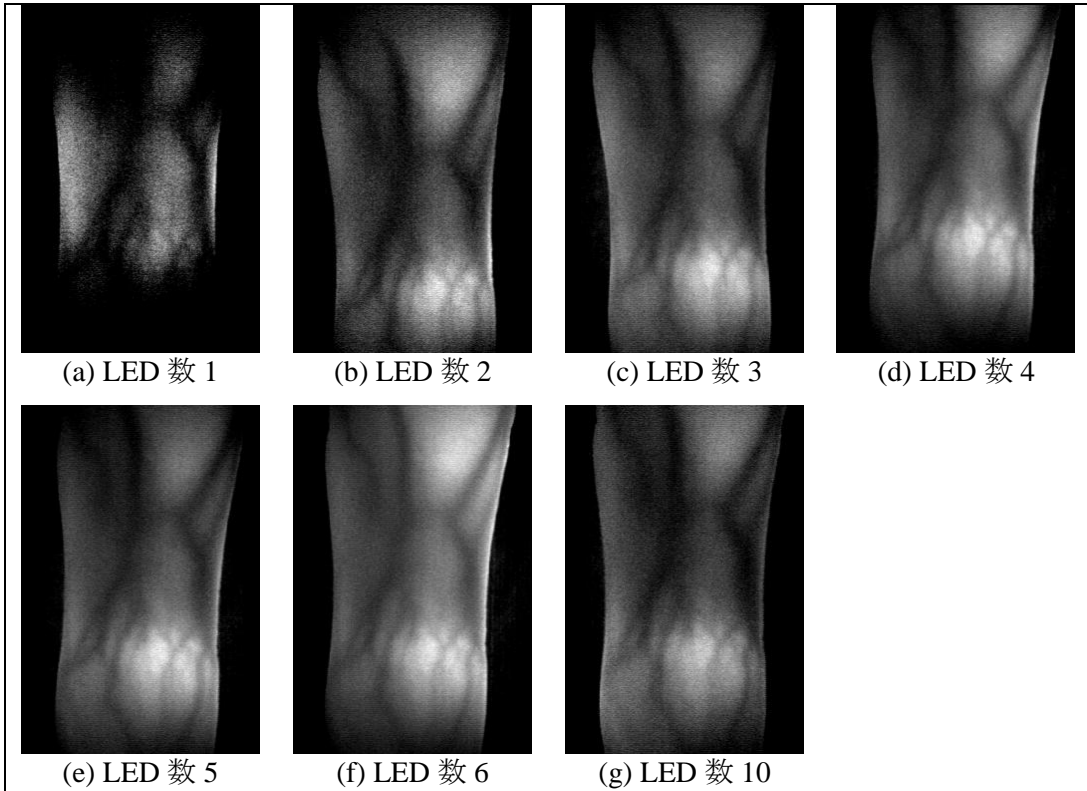


図 4-24 各 LED 数による撮影画像 1

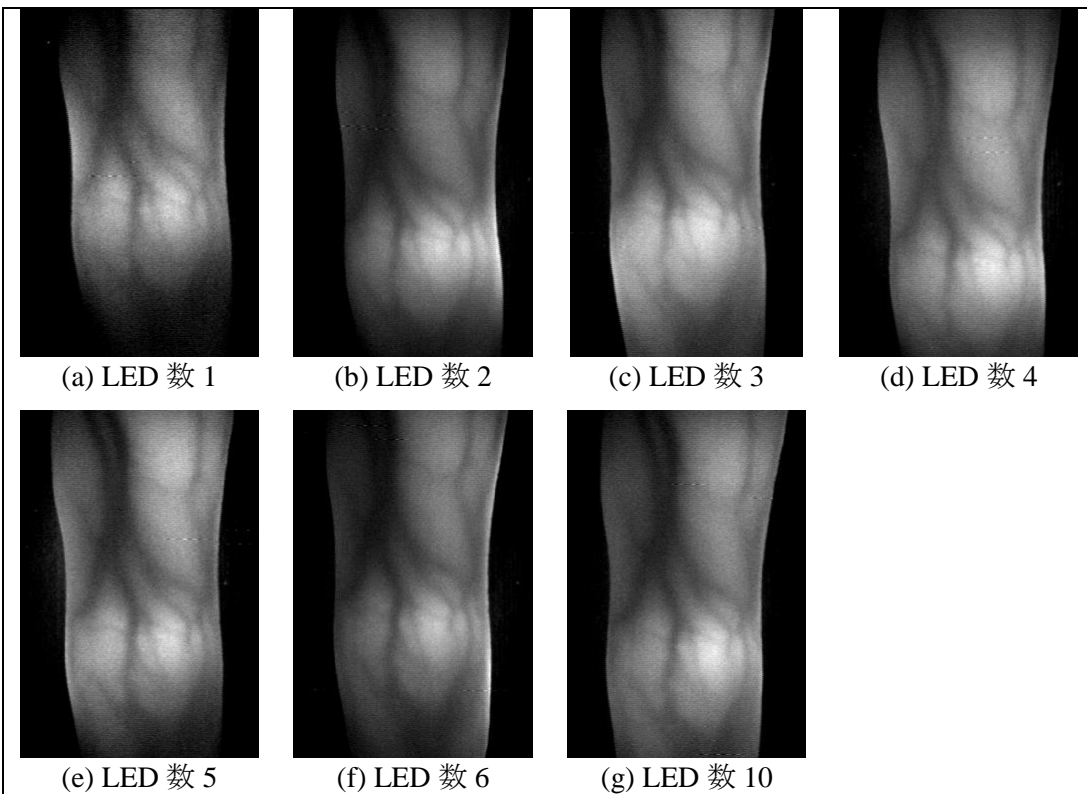


図 4-25 各 LED 数による撮影画像 2

以上の結果より 4~5 個以上の LED を使用することで、視野をカバーする照射範囲となることが確認された。実験では、6 ないし 7 個の LED を使用することとした。

撮影では近赤外領域での生体組織と血中ヘモグロビンの吸光特性の違いを利用していることから、可視光などそれ以外の光は不要である。そこで、近赤外透過フィルタを用いて可視光の外乱を遮断することとした。使用したフィルタは富士フィルムの SC68 である。フィルタ特性を図 4-26 に示す。フィルタにより静脈撮影に有効な近赤外光による撮影が可能となる。

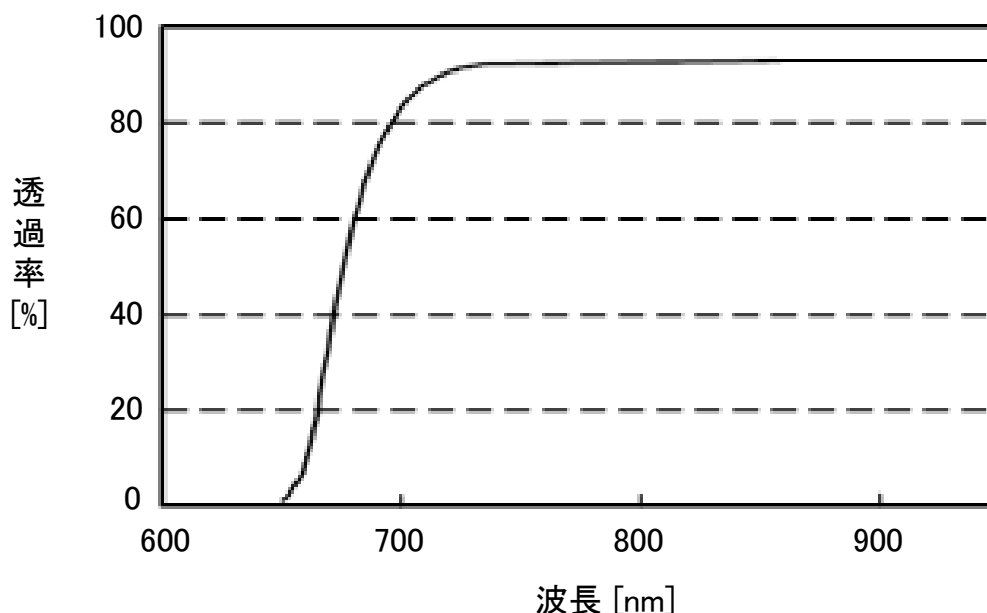


図 4-26 フィルタ特性

4.3.2 光源の波長選定

静脈の鮮明さは指に照射する近赤外光の波長により変化する。そこで、撮影に使用する光源波長の選定を行う。候補光源には波長 700~970[nm] の近赤外領域から選んだ 9 種 (710, 760, 810, 850, 870, 890, 910, 940, 970[nm]) を使用する。波長毎に LED の輝度、カメラの感度が異なるため、電流値を調整して波長毎に LED の輝度を変更する。電流値は 3 人の人物が各々最適と判断した値を使用する。6 人の左人差し指を 3 回ずつ、波長を変えて繰り返し撮影して、486 枚の静脈画像を取得した。各波長での取得画像例を図 4-27 に示す。

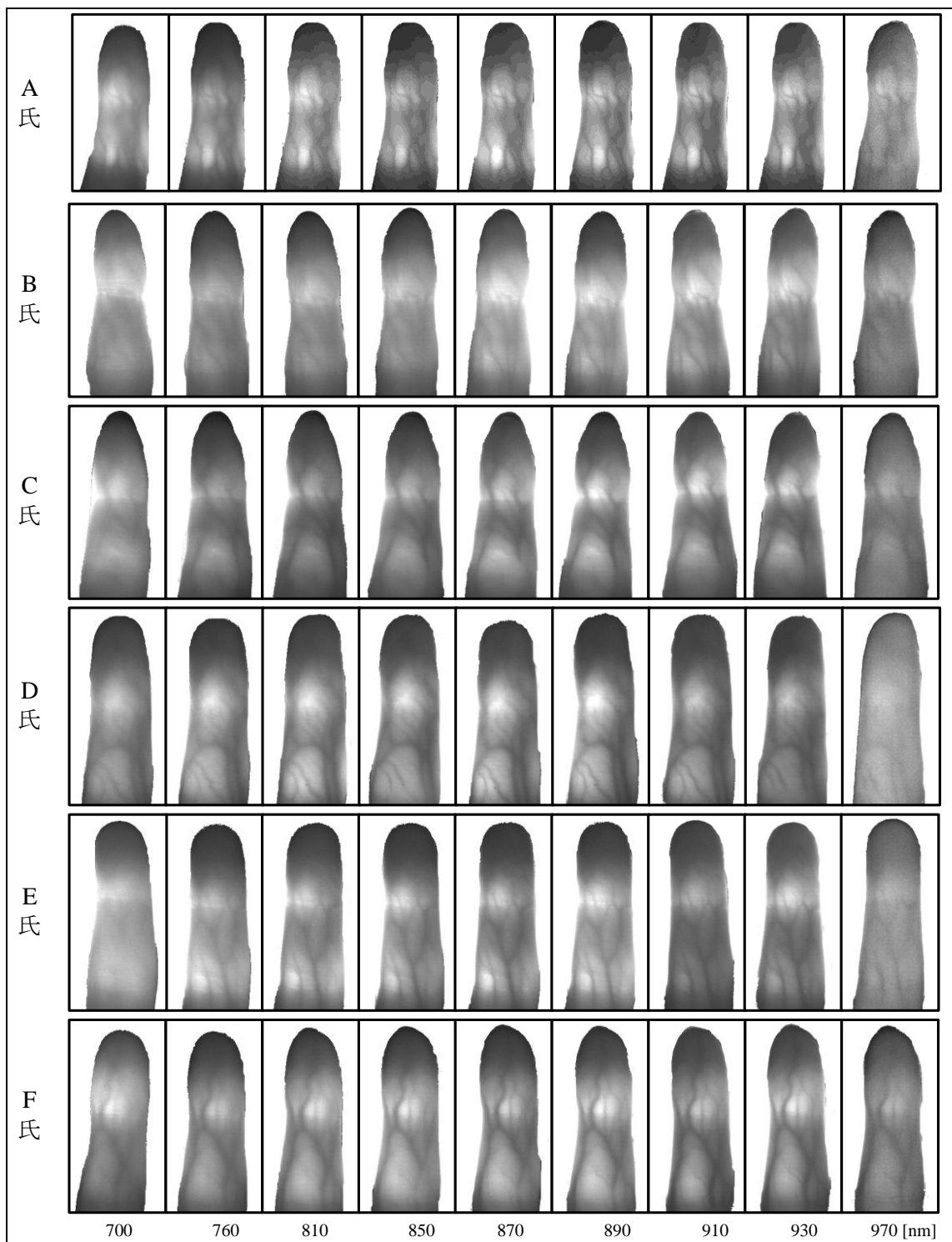


図 4-27 光源波長選定実験用画像

選定は対比較実験により行う。実際の実験方法を以下に示す。

- ① 6人分の画像から1名分をランダムに選択
- ② 提示する2波長をランダムに選択
- ③ 各波長の画像(各9枚)から提示画像をランダムに選択
- ④ 2枚の提示画像のうち、どちらの静脈が鮮明であるかを被験者が選択
- ⑤ 各波長の組み合わせを2回ずつ(計432通り)提示

評価はサーストンの一対比較法により行う。サーストンの一対比較法により算出されるスコアを静脈の鮮明度を示す尺度とする。

対比較実験により得られた2データ間の優劣を表す投票行列を表4-10に示す。また、投票行列より算出したサーストンの一対比較法のスコアを図4-28に示す。スコアは小さいほど優位であることを示す。同図には静脈の鮮明さの基準とされる血液中の還元ヘモグロビンの吸光特性も示す。左軸にスコアを右軸に還元ヘモグロビンの吸光特性を示す^{[67]-[69]}。

表 4-10 投票行列

		不鮮明な光源波長[nm]								
		700	760	810	850	870	890	910	940	970
鮮明な光源波長 [nm]	700		14	6	5	0	1	4	3	61
	760	106		20	24	10	2	8	7	108
	810	114	100		44	18	24	24	23	107
	850	115	96	76		31	26	35	30	110
	870	120	110	102	89		51	63	56	117
	890	119	118	96	94	69		51	64	120
	910	116	112	96	85	57	69		71	118
	940	117	113	97	90	64	56	49		120
	970	59	12	13	10	3	0	2	0	

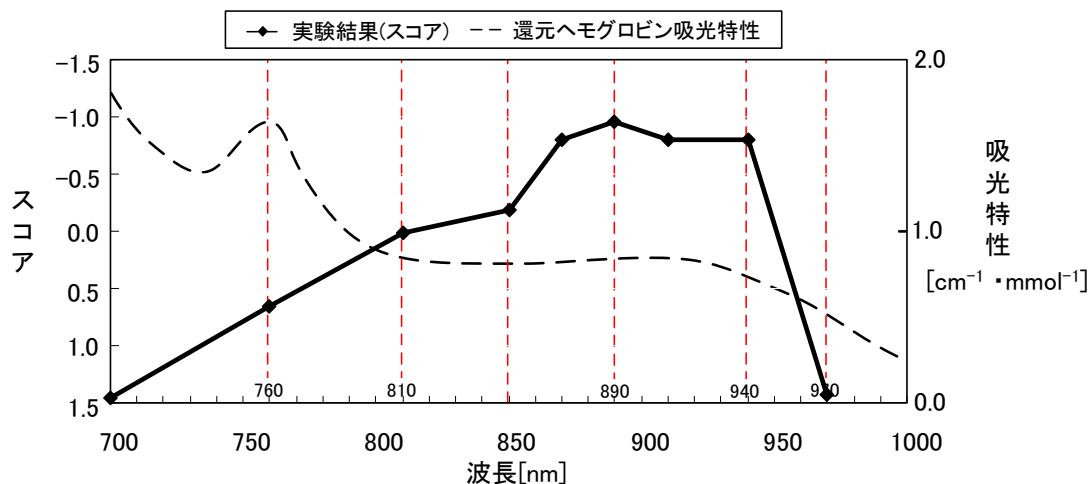


図 4-28 光源波長選定実験結果

図 4-28 より，対比較実験により波長 890 [nm] を光源波長とした際に最も鮮明な静脈画像を取得できる結果が得られた．従来鮮明な静脈画像を取得する際の指標とされてきた還元ヘモグロビンの吸光特性に従うと 760 [nm] を光源波長とした際に鮮明な静脈画像が得られるはずであり，本実験結果と反する．このような異なる結果が得られたのは皮膚などの生体組織の影響が考えられる．生体組織は 900 [nm] 付近の波長の光を透過しやすい特性を持つ^[70]．鮮明な静脈画像撮影には静脈で多く吸光し，それ以外の生体組織では多く透光することがよい．これらの条件が最もよいのが 890 [nm] であると考えられる．この結果より，撮影システムには波長 890 [nm] の近赤外 LED を光源として採用した．

4.3.3 認証アルゴリズム

以下に，認証アルゴリズムの詳細を述べる．図 4-29 に個人認証の流れを示す．認証に使用する領域を切り出し，切りだされた画像に対して周波数領域でフィルタリング処理を行うことで静脈を強調した画像を生成する．最後に端点フリーの DP マッチングにより登録済みパターンとの照合を行う．

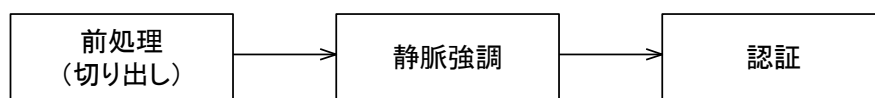


図 4-29 認証の流れ

1. 画像の前処理

撮影装置には指ガイドが設置してあり，被験者は指先がガイドに触れるように指を置く．この際，指先を中心とした傾きのずれと位置ずれが生じる．そのため，まず始めに前処理によりこれらのずれを取り除く．

指の傾きは，指の中心線を検出し，中心線が垂直になるように回転することで補正する．中心線は指輪郭を抽出して，高さ 140～320 までの領域の指輪郭 L_n ， R_n の座標から算出する(図 4-30)．輪郭抽出はソベルフィルタによるエッジ検出を行い，閾値を超える部分を指輪郭として検出することで行う．

次に傾き補正を行った画像から認証に使用する領域を切り出す(図 4-31)．縦方向は指先端から画像下方向に 90[pixel] の位置から 280[pixel]，横方向は指中心線の左右 50[pixel] を切り出し，100×280 の画像を作成した．指の先端座標，指中心線を基準として切り出すことで撮影装置に指を置く際に発生する位置ずれを取り除く．

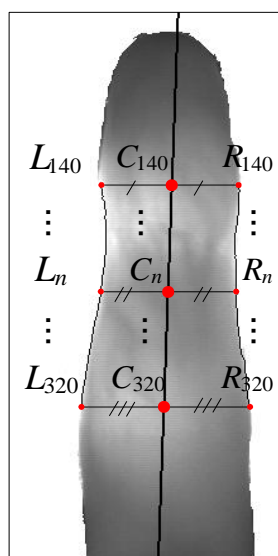


図 4-30 中心線検出

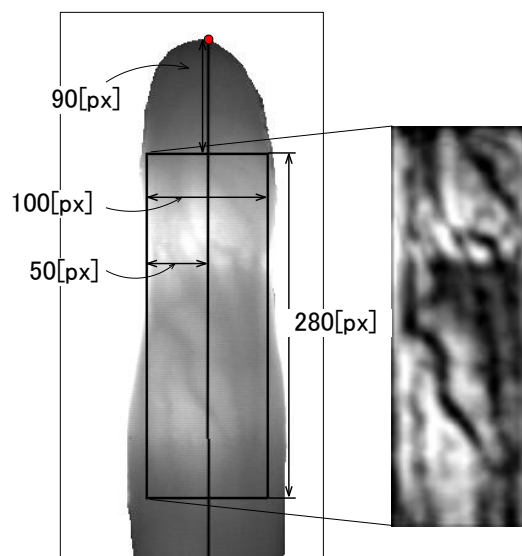
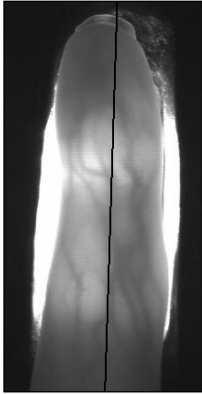
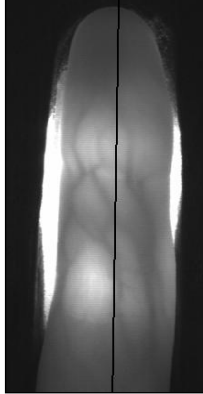
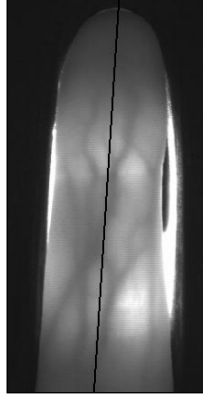
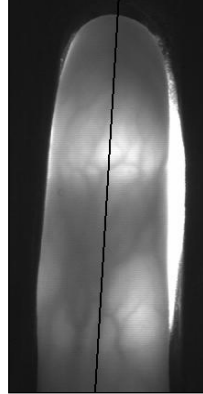






図 4-31 画像切り出し

中心線検出と切り出し処理の例を表 4-11 に示す．中心線検出により指の回転が補正され，静脈領域が伐り出されていることが確認できる．

表 4-11 中心線検出と切り出し結果の例

	A 氏	B 氏	C 氏	D 氏
中心線検出結果				
切り出し結果				

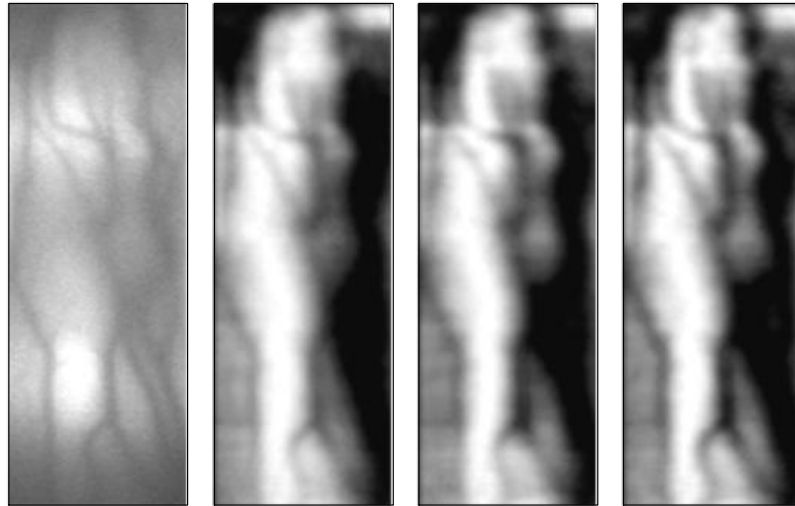
2. 静脈強調

静脈画像から横 1 ラインのデータを取得し，空間周波数領域における m 次成分以上の高周波成分を除去するフィルタを掛け，静脈を強調する．

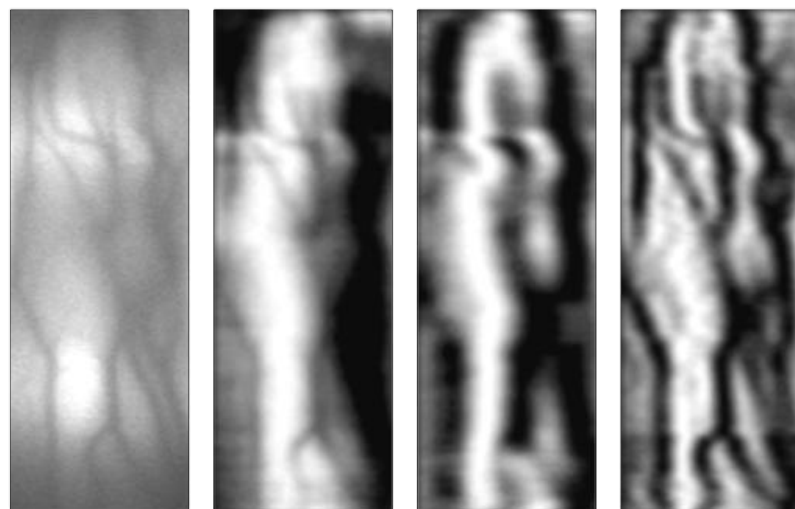
認証に用いる特徴量算出のため，切り出し画像に対して静脈強調を行い，画像中の個人特徴を強調する．まず，画像の横 1 ラインの輝度に対し，フーリエ変換を行う．そして，フィルタ処理により静脈を強調する．本稿では 2 種類のフィルタを検討する．(4-8)式にフィルタ 1 を(4-9)式にフィルタ 2 を示す．なお， m はカットオフ周波数である．強調例を図 4-32 に示す．

$$H(i) = \begin{cases} 1 & (i \leq m) \\ 0 & (i > m) \end{cases} \quad (4-8)$$

$$H(i) = \begin{cases} i & (i \leq m) \\ 0 & (i > m) \end{cases} \quad (4-9)$$



(a) フィルタ 1



(b) フィルタ 2

図 4-32 静脈強調画像例

3. 比較照合

各画像間の類似度を用いて個人認証を行う。類似度算出には端点フリーDP マッチングを用いる。端点フリーDP マッチングでは図 4-33 のように縦方向のパ

ターンの伸縮および位置ずれを考慮した類似度算出を行うことができる。

さらに、横方向の位置ずれを考慮する為に、ずらし幅 a ($-p < a < p$, p は最大ずらし幅) を導入することで、横方向の位置ずれを吸収する。画像間の正規化最小累積距離 $G(I, J)$ が最小となるとき、その値を画像間の類似度とした。

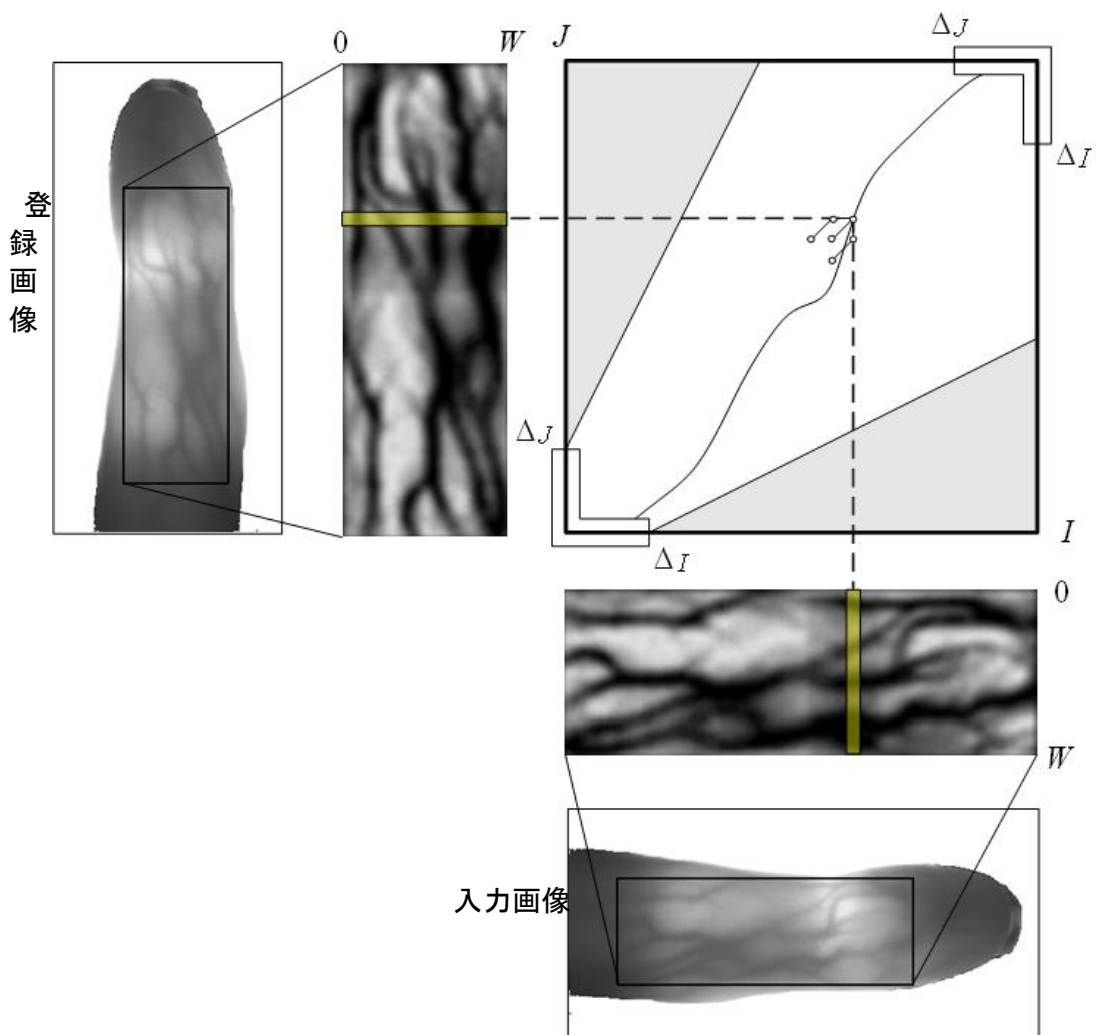


図 4-33 DP マッチングによる認証

正規化最小累積距離 $G(I, J)$, 最小累積距離 $g(i, j, a)$ は式(8), 式(9)のように定義する。また、登録パターンと入力パターンのフレーム間距離 $d(i, j, a)$ は式(10)のように定義する。なお、 x_{reg} , x_{in} は登録, 入力パターンの輝度値, w は画像の幅, I, J は入力, 登録画像での探索長である。

$$G(I, J) = \min \left(\frac{g(I, J, a)}{I + J} \right) \quad (4-1 \quad 0)$$

$$g(i, j, a) = \min \left\{ \begin{array}{l} g(i-1, j-2, a) + 2d(i, j-1, a) \\ g(i-1, j-1, a) + d(i, j, a) \\ g(i-2, j-1, a) + 2d(i-1, j, a) \end{array} \right\} + d(i, j, a) \quad (4-1 \quad 1)$$

$$d(i, j, a) = \frac{1}{W - |a|} \sum_{k=0}^{W-|a|-1} (x_{ref}(k+a, j) - x_{test}(k, i))^2 \quad (4-1 \quad 2)$$

4.3.4 実験用データベース

選定した波長 890nm の近赤外 LED を組み込んだ撮影装置を使用して指静脈を撮影することで、実験用データベースを構築した。なお、光の透過率が個人ごとに異なり、一定の照度では静脈を取得することが出来ないため、光源の明るさは電流値を個人毎に調整して撮影した。撮影対象は人差し指とした。20 人の両手人差し指をそれぞれ 10 回ずつ撮影した。撮影したデータベースの内訳は、表 4-12 に示す通りである。

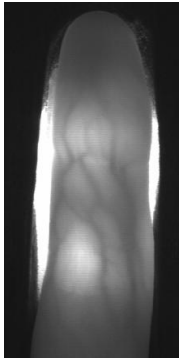
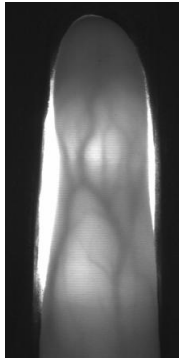
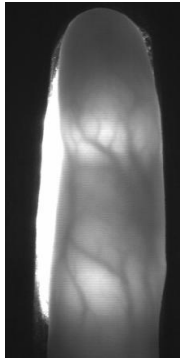
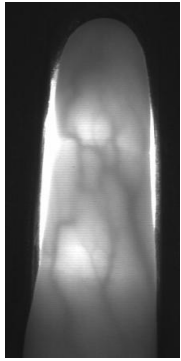
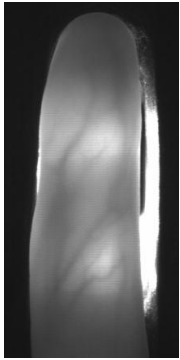
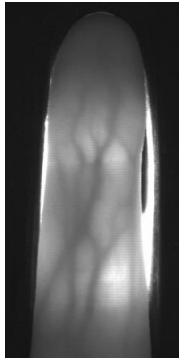
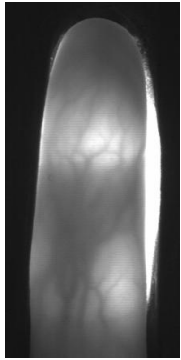
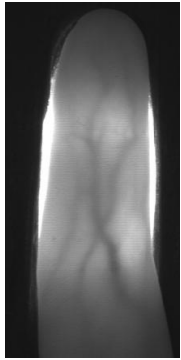
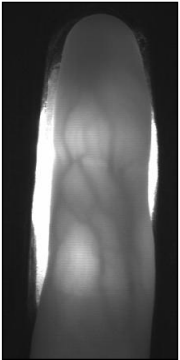
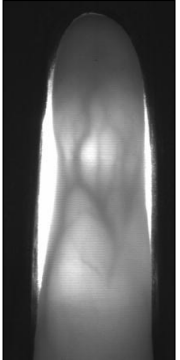
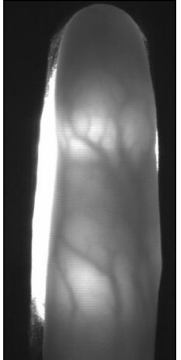
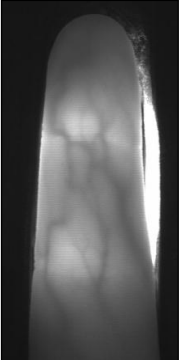
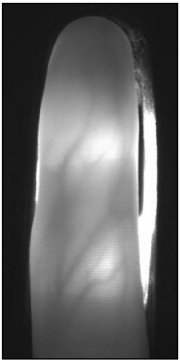
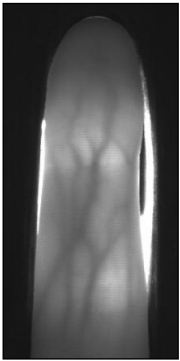
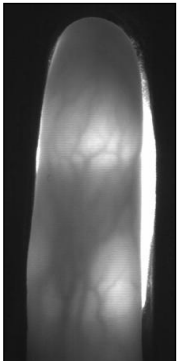
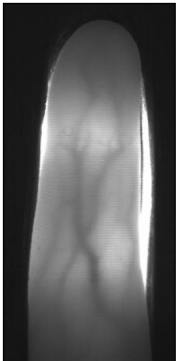
表 4-12 実験用データベース画像枚数

撮影人数	一人当たりの指数	一指あたりの撮影回数	総データ数
20	2	10	400

認証実験では登録画像 1 枚、入力画像 1 枚で認証実験を行うことから、本人比較の組み合わせが 3600 通り、他人比較の組み合わせが 156000 通りであり、総組み合わせ数は 159600 通りである。

撮影画像の一例を示す。これらをデータベースとして、認証実験を行うことで、認証アルゴリズムの有効性を確認する。

表 4-13 実験用データベース

		Person1	Person2	Person3	Person4
Data1	右				
	左				
Data2	右				
	左				

4.3.5 静脈強調フィルタの検討

静脈強調を行う為のフィルタとして2種類のフィルタから検討する。検討を行うフィルタは(4-8), (4-9)式に示すフィルタである。認証実験結果を図 4-34 に示す。

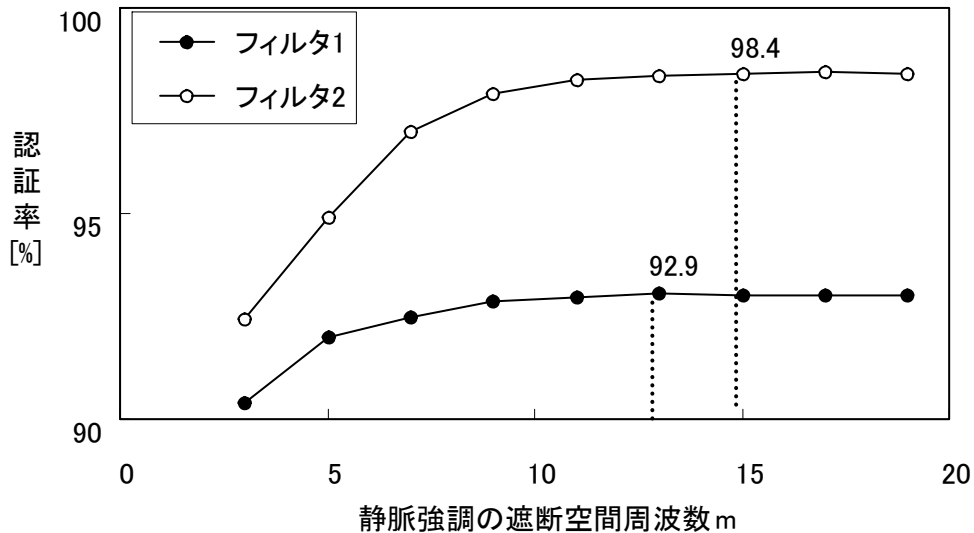


図 4-34 2種類のフィルタでの認証実験

図 4-34 より、ローパスフィルタを用いて静脈強調をした場合での最大認証率は $m=13$ の際、92.9[%] (11290/159600)、微分とローパス両方を持つフィルタを用いた場合の最大認証率は $m=15$ の際、98.4[%] (誤認証数 2611/159600) であり、フィルタ 2 を用いた場合の方が高い認証率が得られた。フィルタ 1 のようなローパスフィルタを用いた静脈強調では低周波成分が残るため m を変化させても細部の静脈が強調されない。それに対し、フィルタ 2 はバンドパスフィルタに似た特性を持ち、 m を大きくすることにより低周波成分が減衰されて、細部まで静脈を強調できると考えられる。

4.3.6 マッチング方向の検討

本認証手法では DP マッチングの特性よりマッチング方向の画像の伸縮が吸収できる。そこで、縦横どちらの方向の伸縮を吸収するのがよいかを検討する。各認証方向での認証実験結果を図 4-35 示す。実験より横方向に DP マッチングを適用した認証で 98.8[%] の認証率であり、縦方向の認証に比べ高い認証率

が得られた(図 4-36)。横方向の認証では指の中心を軸とした回転に前処理よるずれを吸収できると考えられ、その回転ずれの吸収が認証率を向上させることを確認した。

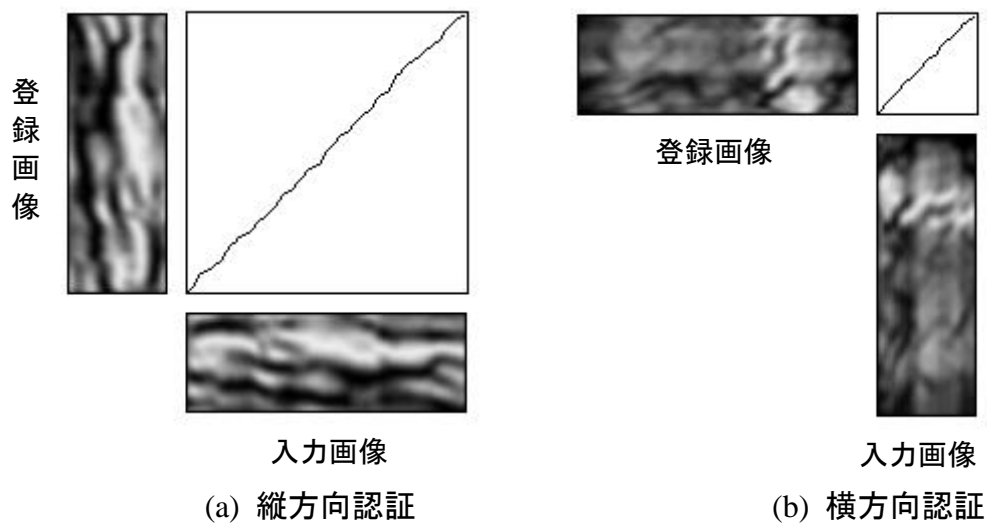


図 4-35 各方向での認証概念図

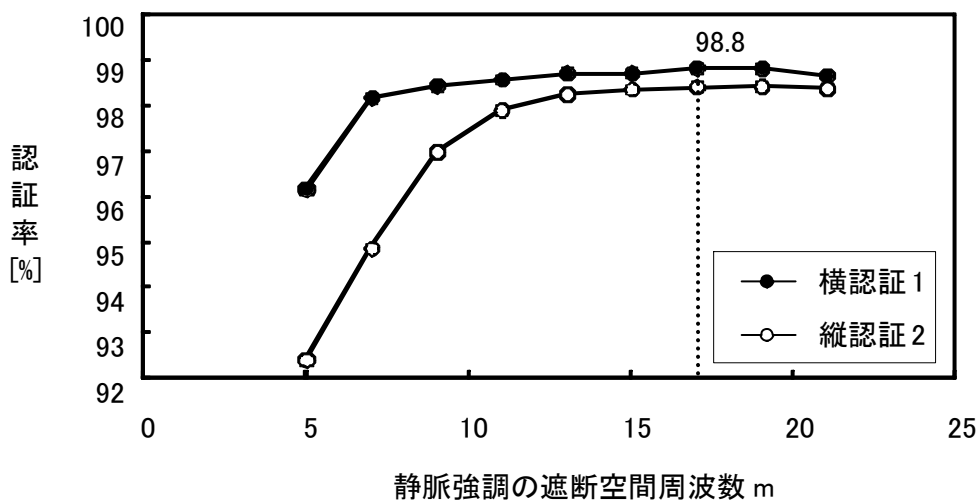


図 4-36 各方向での認証実験結果

4.3.7 静脈強調フィルタの適用方向の検討

静脈強調において1ラインを取り出す方向により静脈の強調のされ方が異なる。そこで、認証実験により認証に適した分析方向を検討する。検討するのは横方向分析画像、縦方向分析画像、さらに各方向での分析画像を足し合わせた画像である。各分析方向での静脈強調画像を表 4-14 に示す。横方向分析では縦方向に伸びる静脈が強調されやすいが縦方向分析では強調されにくいことが確認できる。認証実験結果を表 4-15 に示す。実験より縦方向分析の方が横方向分析より高い認証率が得られた。身体的特性上、縦方向に伸びる静脈は全ての人々が持つ特徴である。従って、それのみでは他人との区別が難しくなる場合がある。一方、横方向に伸びる静脈は共通の特徴ではないため、誤認証が生じにくい。以上のことから、静脈の個人性は横方向に伸びる静脈が多く持つと考えられる。また、それらの足し合わせでは、両方で強調された静脈パターンを併せ持つパターンとなり、高い認証率が得られたと考えられる。

表 4-14 各分析方向での静脈強調画像







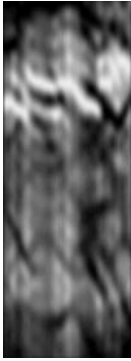

元画像	横方向分析	縦方向分析	足し合わせ画像
			
			

表 4-15 各分析方向での認証率[%]

	横方向	縦方向	両方向
最大認証率	98.8	99.5	99.7

4.3.8 切り出しサイズの検討

認証実験により認証に適した縦、横の画像サイズを検討した。画像サイズ(横)に関する実験結果を図 4-37 に示す。画像サイズ(横)が 80, 90[pixel]の際に最大認証率 99.9% が得られた。画像サイズ(縦)に関しても同様に実験を行い、240[pixel]の際に最大認証率 99.9% が得られた。

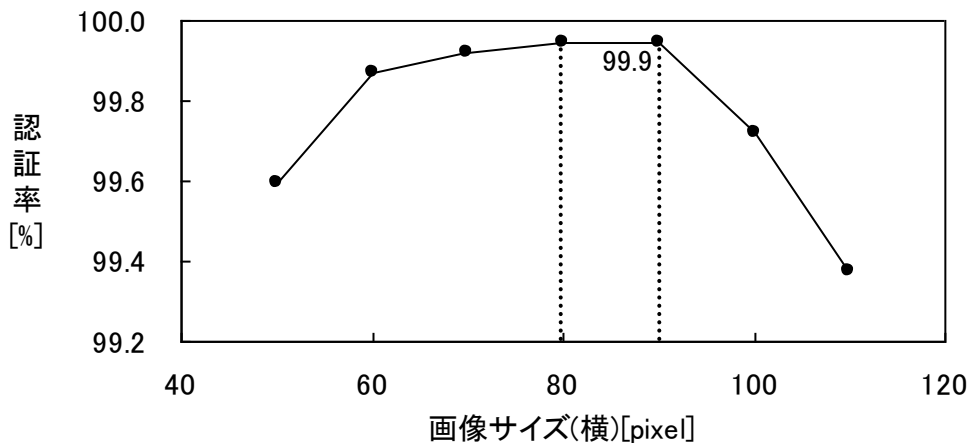


図 4-37 切り出しサイズの違いによる認証率の変化

4.3.9 考察

本稿では登録画像 1 枚の実験で 99.9% の精度が得られた。また、アルゴリズムによる性能差が出やすいよう登録画像を 1 枚としたが、個人認証では登録画像を複数枚とすることでさらに認証率が向上する。従って、本手法は非常に高い認証結果が得られたと言える。

また、官能評価実験と個人認証実験の結果から、共に最も良い結果が得られた 890[nm] を静脈撮影に用いる単波長光源として選定した。しかし、全ての人において 890[nm] が最も良い結果が得られるわけではなく、主観評価では個人ごとの結果を見ると 870[nm] や 910[nm] で最もよいスコアとなるものも存在す

る。すなわち、指の個人性により最適な波長が異なると考えられる。今回、単波長光源での実験のみであることから、複数波長を組み合わせた光源を用いた際、認証率が向上する可能性があり、今後検討していく必要がある。

さらに、認証システムとして構築するには、LED 輝度調整を自動化する必要がある。

4.3.10 LED 自動輝度調整システム

一定の輝度の赤外光を照射して静脈を撮影した場合、人によりサチュレーションや黒つぶれが生じてしまうため静脈が明瞭に映らないという問題が発生する。そのため、実用的なシステムを構築するためには輝度を自動的に調整する機構が必要となる。輝度の調整法としては、光源の照射輝度・ゲイン・アイリス・露光時間など様々な方法が考えられる。同じ光量の照明を用いた場合でも、人により完全に黒潰れあるいは白飛びしてしまうことがあるため、ゲインの調整は難しいと考えた。さらに、アイリスの調整にはアイリスを変動させる機構が必要であること、露光時間の調整は撮影時間の変化を伴うことから、今回は単純に照射輝度の調整するシステムを検討した。

LED の輝度の調整は、電圧制御により行うこととした。COM 通信により、0～255 の 256 段階で輝度を調整する基板を作成した（図 4-38、表 4-16）。3 つの LED を直列として、それらを並列に接続することで、USB バスパワーで駆動するよう設計した。輝度調整のため情報取得には、特にセンサの追加などは行わず、画像情報からフィードバック制御を掛けることで行う。

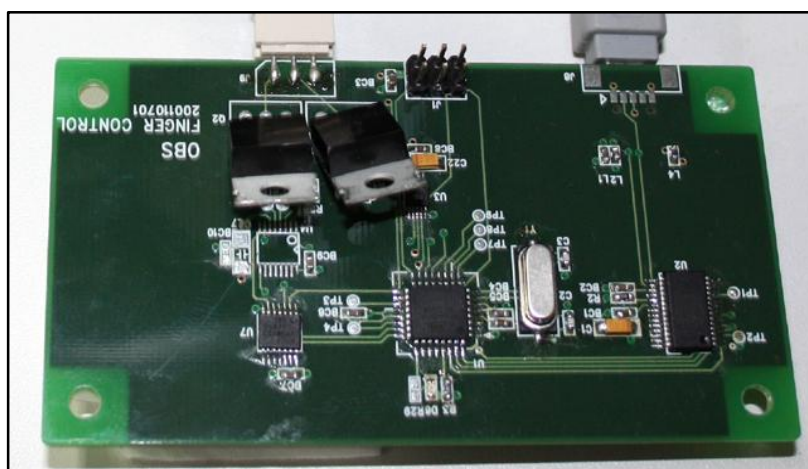


図 4-38 基板の外観

表 4-16 LED 調整基盤の仕様

インターフェイス	USB
制御信号	デジタル 8 ビット
最大出力	5 [V]

男女各 4 人について、輝度を変えた場合の撮影画像を図 4-39、図 4-40 に示す。同一の照射輝度でも人により明るさが大きく異なり、人によりサチュレーションや黒潰れが発生することが確認できる。このことから輝度調整の必要性が裏付けられた。撮影結果から、特に女性データについてはサチュレーションが生じやすいことが示唆されている。このとき、静脈画像が鮮明に見えるのは、サチュレーションや黒潰れが少なく、指領域の平均輝度値が 90 程度であるため、その輝度を目標輝度として LED を調光することとした。

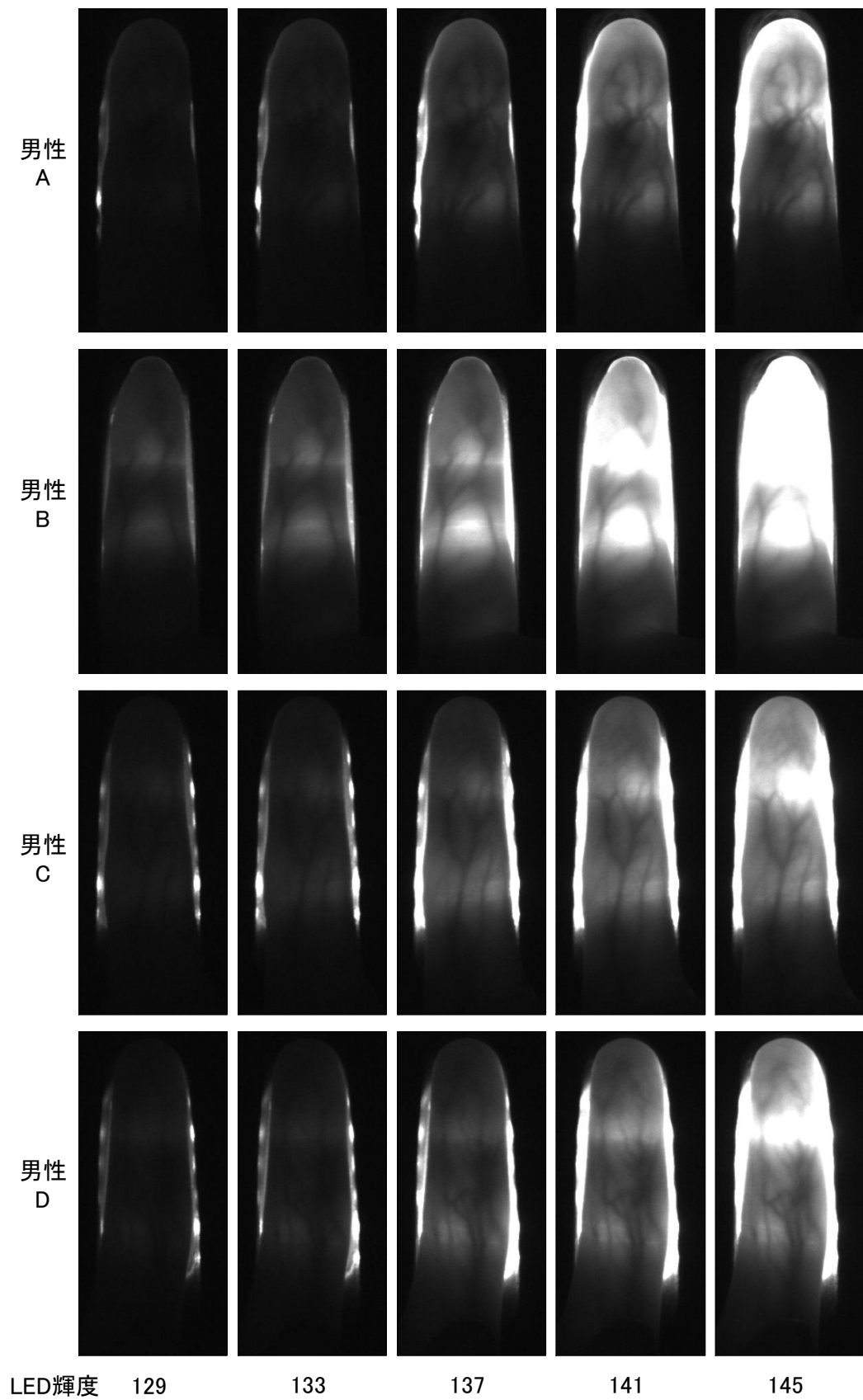


図 4-39 男女各 4 人分の静脈撮影例

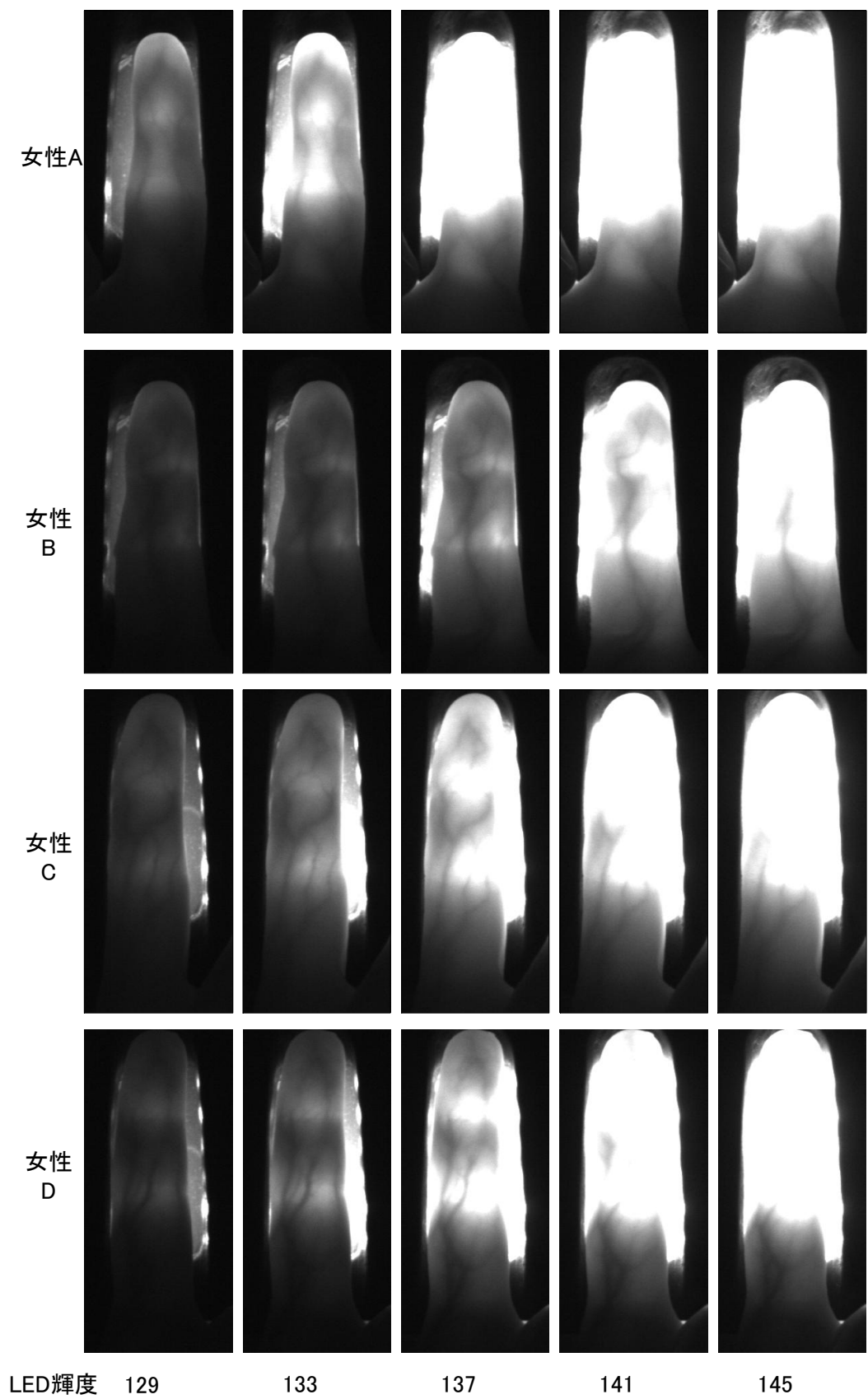


図 4-40 男女各 4 人分の静脈撮影例

撮影実験をもとに，LED 輝度調整のアルゴリズムを図 4-41 のように設計した．サチュレーション及び黒潰れのピクセルが一定以上の場合に大きく輝度値を調節することで，最適な輝度に設定されるまでの時間を短縮している．

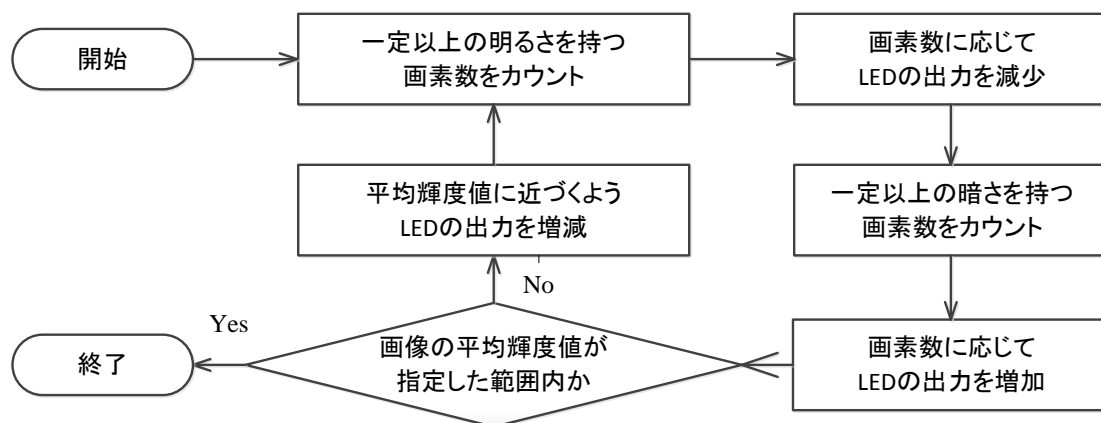


図 4-41 輝度調整フロー

輝度調整の過程を図 4-42，図 4-43 に示す．7～8 フレーム程度の間には輝度が調整されていることが分かる．図 4-43 の結果を見ても分かる通り，輝度調整により調整前に比べて鮮明な静脈の画像を捉えることになることが分かる．

図 4-39，図 4-40 で撮影した人についても，輝度調整後の静脈強調画像を確認した．その結果を図 4-46 に示す．どの被験者の指に対しても静脈の抽出に成功していることが確認された．

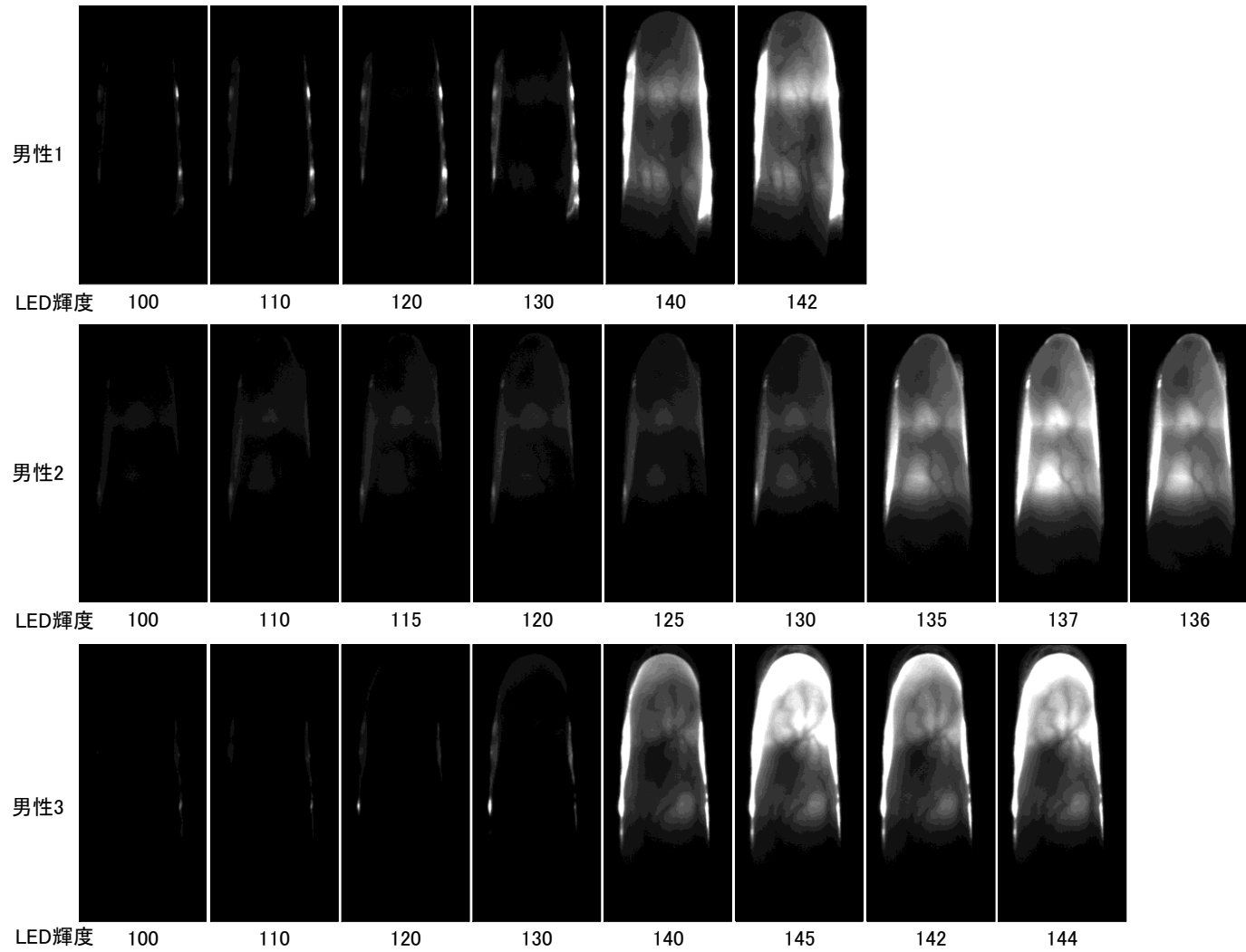


図 4-42 男性 3 名分の LED 輝度調整過程

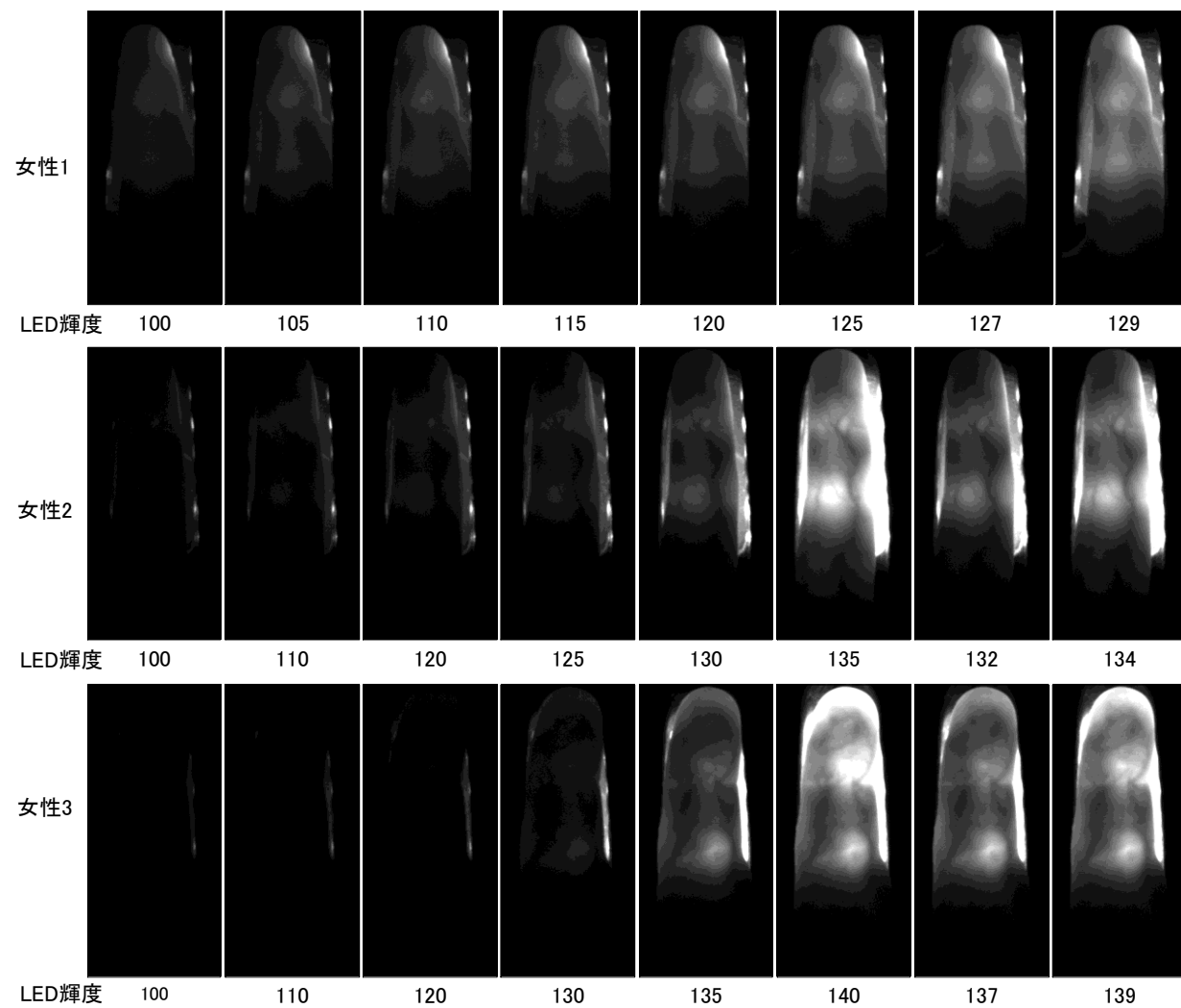


図 4-43 女性 3 名分の LED 輝度調整過程

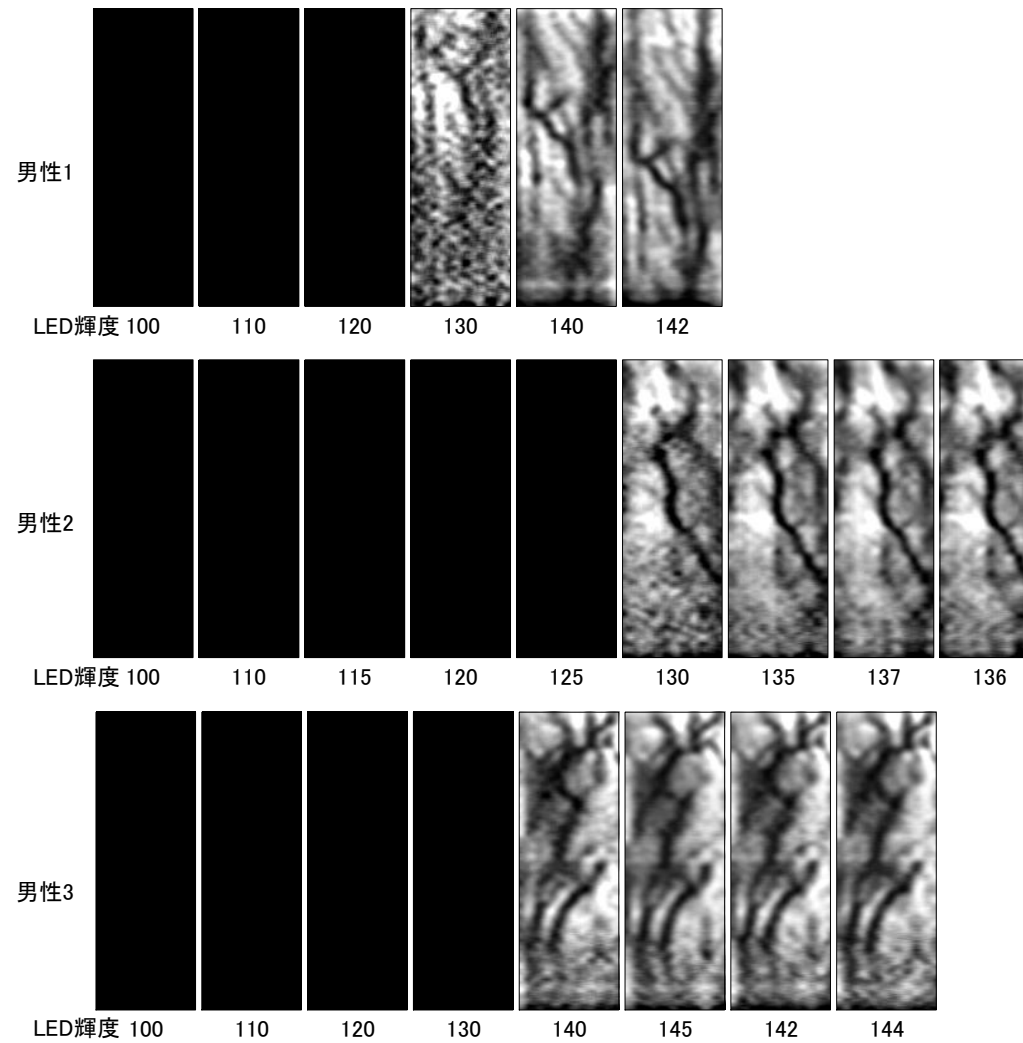


図 4-44 男性 3 名分の LED 輝度調整過程（静脈強調処理後）

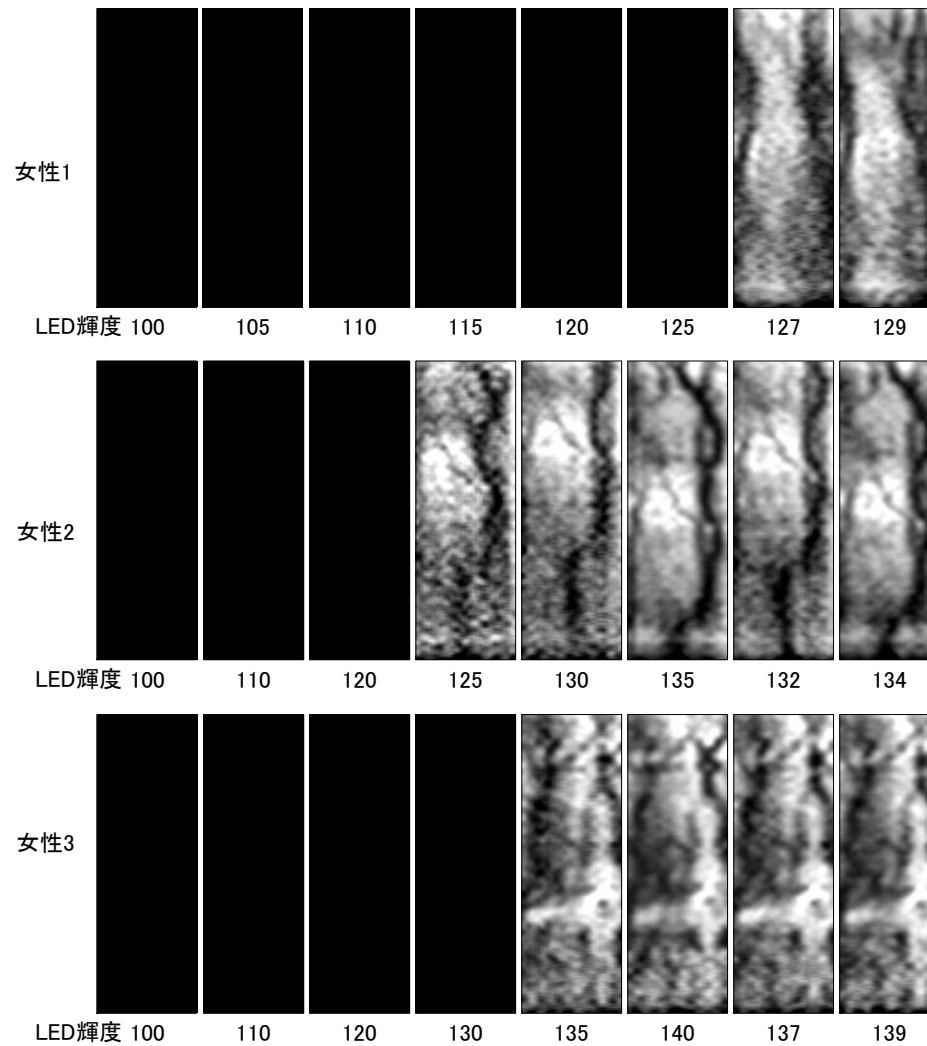


図 4-45 女性 3 名分の LED 輝度調整過程 (静脈強調処理後)

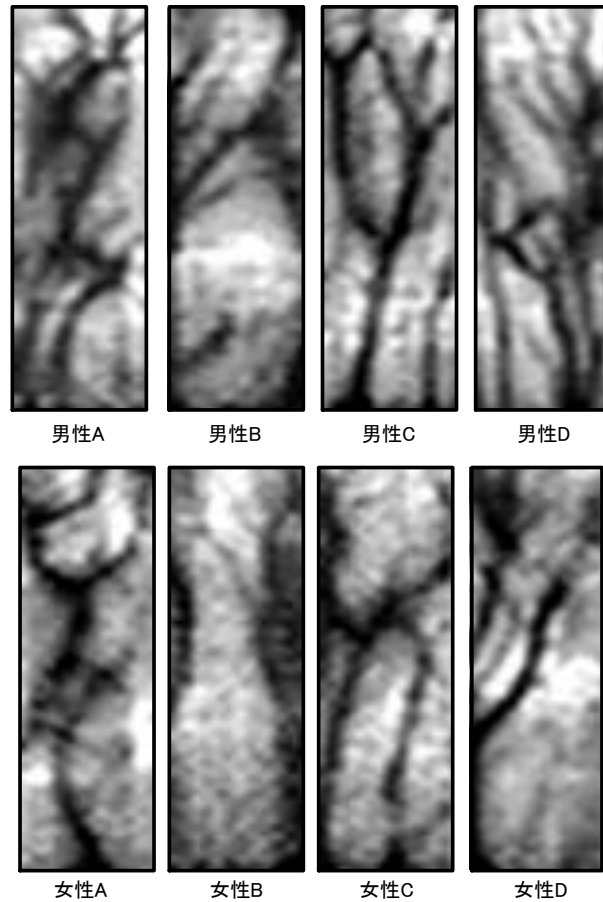


図 4-46 男女各 4 人分の静脈撮影例（静脈強調処理後）

LED 輝度調整アルゴリズムを使用して、男女各 10 人の左右の人差し指（計 40 指）の最適な輝度レベルを調査した。その結果を図 4-47 に示す。

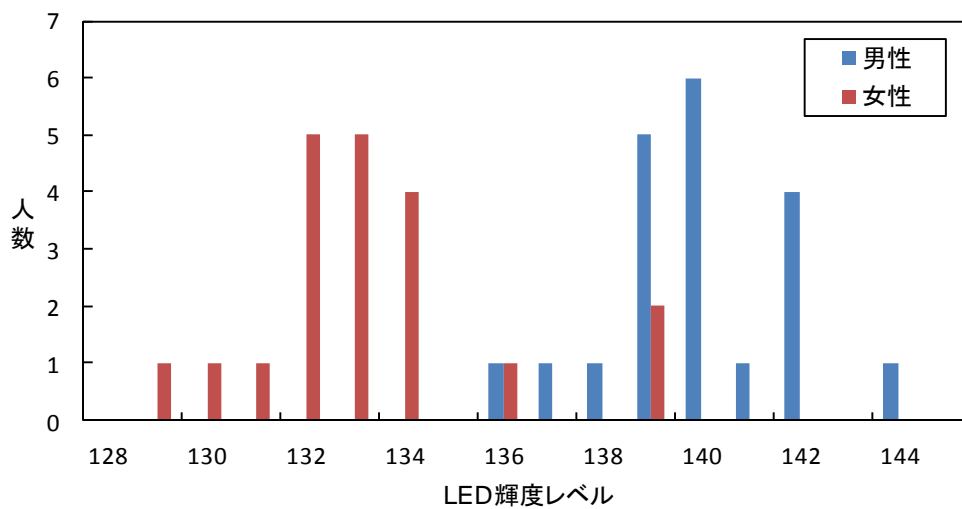


図 4-47 最適 LED 輝度レベル分布

撮影実験で示唆されている通り，男性に比べて女性の方がより弱い照明輝度レベルで静脈画像の撮影が可能なが確認できた。

現状，画像処理と認証アルゴリズムについては，30fps 程度で動作しているため，リアルタイムに処理できていると言える．しかし，輝度調整に2秒程度を要しており認証全体のボトルネックである．全体にかかる時間を短縮するためには，より短時間で輝度調整を行う必要がある．今後，二分探索や線形予測分析などを適用することで，探索回数の削減や，最適な輝度レベルを推定する枠組みを導入して，より高速に輝度の調整を行う手法を検討したいと考えている．さらに，透過特性にも個人性が表れていると考えられるため，複数の輝度レベルで撮影された画像を利用した精度向上手法について検討したい。

4.3.11 静脈認証システムの試作

指静脈認証を行うためのソフトウェアを試作した．作成したソフトウェアのインターフェイスを図 4-48 に，ハードウェアを含む全体の構成を図 4-49 にそれぞれ示す．このソフトウェアにより，LED 輝度補正と認証処理を 15 [fps] に程度で行うことが可能なことを確認した。

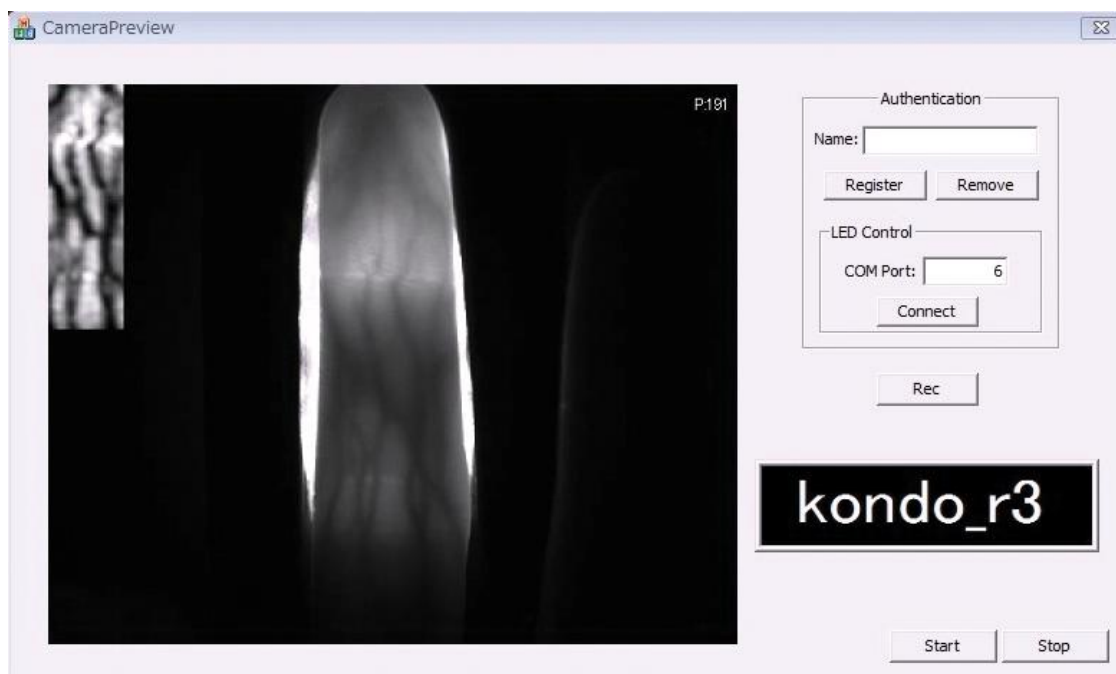


図 4-48 指静脈認証ソフトウェア

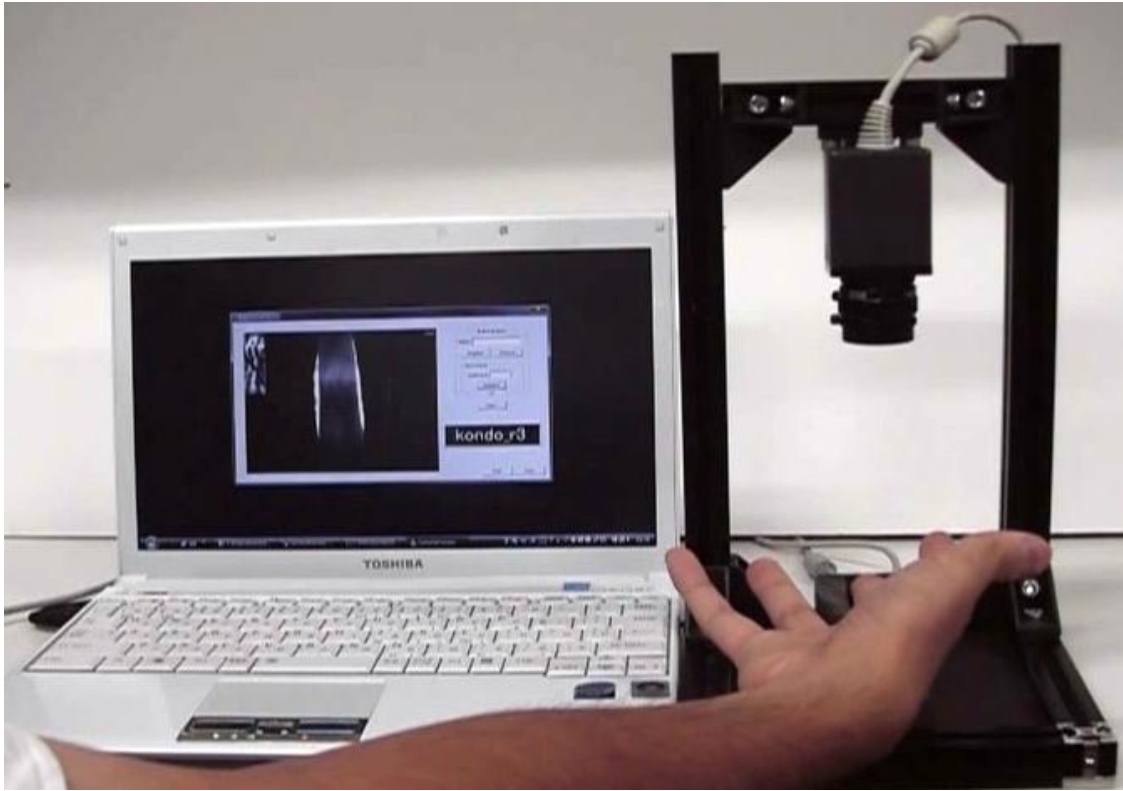


図 4-49 指静脈認証システム試作機

さらに、試作機をもとに認証ユニットの小型化が可能なのも確認されている（図 4-50）。今後、カメラをより小型なものにすることにより、コンパクトな認証システムが構築可能だと考えられる。



図 4-50 指静脈認証システム試作機

4.4 まとめ

本章では，次世代のバイオメトリクスシステムの構築を目指して，非接触型の指紋認証システムと指静脈認証システムを開発した．

非接触型指紋入力システムでは，カメラとLEDを用いて指紋を撮影して，それにより得られた画像に対する前処理法について検討した．その結果，照合実験にて98.1%の照合率を得た．しかし，接触型センサによる照合率と比較すると，精度が劣る．今後，登録画像枚数を複数にすることや，中心点検出誤りの改善，可視光以外の波長の使用により，さらなる照合率向上が必要となる．

指静脈認証システムでは，光源波長として従来の800[nm]付近とは異なり，波長890[nm]で最も鮮明な静脈画像が得られることが判明した．また，指中心線を基準とした傾き補正と切り出しを行い，静脈強調フィルタを用いることで，最大認証率99.9[%]が得られた．さらに，指静脈認証のための輝度調整システムの開発をした．その結果，撮影画像をから得られた輝度情報に基づいてLEDに印加する電圧を調整することで，男女20人（計40指）で静脈を安定して撮影することが出来ることを確認した．さらに，作成した輝度調整システムを静脈認証システムに組み込むことで，リアルタイムに認証を行うシステムが実現できることを確認した．今後，輝度調整が認証率に及ぼす影響の調査が必要である．さらに，より高精度な認証を行うために，複数レベルの輝度で撮影された静脈画像を用いた認証を検討した．また，高速な輝度調整手法についても検討を行い，認証システムの完成度を高める必要がある．また，静脈撮影システムをよりコンパクトにするためには，反射型静脈撮影システムの開発が必須であると考えられる．透過型と同様に，照射角度や輝度，波長についての検討を行いたい．

第5章 むすび

本論文では動的計画法を応用した音と画像の高精度認識システムについて検討した。音声画像を用いたセンシング手法として、自己組織化マップを用いた顔方向のセンシング、動的計画法と CHLAC 特徴による音の不良検査手法、指紋・静脈によるバイオメトリクスという事例について述べた。

まず、ロボットなどのマンマシンインターフェイスへの応用が期待できる人間の顔方向を認識する手法について述べた。自己組織化マップにより顔画像を補間生成することで、精度よく人の顔方向を推定できることを確認した。実験では、3次元計測データから生成された精度保証された2次元画像データを用いて RMSE は 1.87 度という高い精度で推定可能なことを確認した。さらに、提案法は2自由度にも拡張可能であり、それを用いた首振りと頷きの認識が可能なことを確認した。

次に機械駆動音による不良検査システムについて述べた。対象とした駆動音は非連続的な動きに伴う駆動音であり、従来の不良判定法では難しいものを対象としている。検査手法としては、連続 DP マッチングを用いた手法と、立体高次局所自己相関特徴を用いた良否識別手法をそれぞれ提案した。どちらの手法についても、従来法に比べて高精度に不良判定を行うことが出来ることを確認した。さらに、ライン用のシステムに必要なリファレンスの自動生成手法についても検討を行い、実際に使用可能なソフトウェアも実現した。本システムの実用化により、検査コスト削減、検査員の負担軽減、検査の高品質化が見込まれる。しかし、背景音等の外乱音の影響を受けることが確認されており、それらに対する検討が必要である。また、現場で安定した検査を行うためには、良否判定を行う際に使用するリファレンスの自動選定機能があることが望ましいと考えている。これらの手法は、他の駆動音や加速度センサから入力された振動情報に対しても適用可能なことを確認している。今後、どのような対象まで応用可能なのかを検討したい。

最後に、非接触指紋と指静脈認証システムについて述べた。非接触指紋は、カメラと LED を用いて指紋画像を取得して、前処理を適用することで接触型指紋と同様の画像を生成することに成功した。これにより、非接触で衛生的な指紋認証システムを安価に構築可能だと考えられる。指静脈認証システムでは、近赤外 LED を光源とする静脈画像撮影装置を構築して、静脈の鮮明さに影響を及ぼす光源波長として 890[nm]を選定した。認証アルゴリズムについても検討を行い高精度な認証を実現した。さらに、認証システムの構築に必要な輝度調整機構も開発した。今後、反射型の小型静脈認証システムや、指紋・静脈のハイブリッド認証システムを実現したい。

近年、コンピュータやセンサが低価格化・高性能になり、音声や画像を用いた様々なセンシングが実用化されている。本論文では、動的計画法を応用した高精度な認識システムについて述べた。今後、これらの技術についてさらに検討を行い、より高精度かつ使用しやすい実用的なシステムを実現していきたい。

謝辞

本研究を行うにあたり，終始懇切丁寧なるご指導を頂きました名古屋工業大学 梅崎太造教授，岩田彰教授，北村正教授，中部大学 藤吉弘亘教授に謹んで感謝いたします。

また，本研究を行うにあたりご協力頂いた，田口亮助教をはじめとする梅崎研究室関係者各氏に深謝いたします。研究に対して様々なご協力を頂いた中部大学 葛谷幹夫教授，保黒政大准教授に感謝いたします。

研究のためのデータ提供や様々なご助言，惜しみないご協力を頂きました共同研究先の企業の皆様に対しても深く感謝申し上げます。

なお，静脈認証の研究の一部は，公益財団法人堀科学芸術振興財団の研究助成を受けて実施しております。関係者各位に深謝いたします。

参考文献

- [1] 島田 竜也, 柳下 達也, 河口 尚広, 熊谷 拓哉, 内藤 恵介, 山田 博三, 森 晃徳, “屋内侵入者検知システムの開発”, 電子情報通信学会技術研究報告. PRMU, Vol.102, No.531, pp.87-92, 2002.
- [2] 羽下 哲司, 八木 康史, “時空間動き特徴に着目した屋外侵入者監視技術に関する研究”, 情報処理学会研究報告. CVIM, Vol.2006, No.51, pp.259-274, 2006.
- [3] 井手徳隆, 石川眞澄, “視覚情報に基づく異種物体の実時間検出”, 電気情報通信学会論文誌, Vol.104, No.758, pp. 161, 2005.
- [4] 西田健陽, 永菅久嗣, 永菅久嗣, “表面検査の自動化・インライン化の動向 : レーザーによる目視検査”, 溶接学会誌, Vol.67, No.6, pp. 502, 1998.
- [5] 岡本紳二, 吉村一成, “外観検査への新しい手法の導入”, 溶接学会誌, Vol.58, No.4, pp. 300, 1990.
- [6] 横井茂樹, “実装回路基板はんだ面の自動外観検査”, 精密工学会誌, Vol.56, No.8, pp. 1407, 1989.
- [7] 藤吉弘亙, 梅崎太造, 竹内英世, 佐分利伸, 鈴木宣夫, “スペクトルの遷移確率を利用した指紋照合法”, 電子情報通信学会論文誌, Vol.J80-D-2, No.5, pp.1169-1177, 1997.
- [8] 吉嶺達樹, “指紋/顔パターン認識アルゴリズム”, インターフェース, CQ 出版, No3, pp.62-68, 2005.
- [9] 赤松茂, “コンピュータによる顔の認識—サーベイ—”, 電子情報通信学会論文誌, Vol.J80-D-2, No.8, pp.2031-2046, 1997.
- [10] 角所考, 李立群, 美濃導彦, “擬人化エージェントを介したコミュニケーションのための表情マッピングの獲得”, 情報処理学会研究報告, ICS2004 - 105, pp.2031-2046, 2004.
- [11] Richard E Bellman, Stuart E Dreyfus, “Applied Dynamic Programming”, Princeton University Press, 1962.
- [12] 迫江博昭, 千葉成美, “動的計画法を利用した音声の時間正規化に基づく連続単語認識”, 日本音響学会誌, Vol.27, No.9, pp.483-490, 1978.

- [13] 山田圭, 内田誠一, 迫江博昭, “フレーム間単調連続周波数ワープによる不特定話者単語音声認識の検討”, 電子情報通信学会技術研究報告, PRMU97-205, pp. 73-78, 1997.
- [14] Forstmn Sven, 加納裕, 大谷 淳, Thuring Sven, Schmitt Alfred, “動的計画法による実時間ステレオ”, 電子情報通信学会技術研究報告, PRMU103-453, pp. 103-108, 2003.
- [15] 迫光裕章, “DP マッチングとの 30 年”, 電子情報通信学会技術研究報告, SLP2002-121, pp.207-212, 2002.
- [16] V.Bruce, “Recognizing faces, Lawrence Erlbaum Assoc”, London, 1988.
- [17] 高橋宏, 下村倫子, “運転支援のための画像処理”, 映像情報メディア学会誌, vol.51, No.6, pp.746-750, 1997.
- [18] 青山晃治, 山本毅, 大西昇, 杉江昇, “一台のカメラによる顔と視線方向の推定”, 電子情報通信学会技術研究報告, PRU95-233, pp.131-136, 1996.
- [19] Andrew Gee, Roberto Cipolla, “Determining the gaze of faces in images” Image and Vision Computing Vol.12, No.10, pp.639-647, 1994.
- [20] 渡辺彰裕, 斎藤英雄, “固有空間法を用いた濃淡画像からの顔の向き of 推定法”, 電子情報通信学会技術研究報告, PRMU97-71, pp.111-116, 1997.
- [21] 吉川元基, 呉海元, 塩山忠義, “ステレオによる顔の 3 次元特徴モデルの作成と姿勢推定”, 電子情報通信学会技術研究報告, PRMU2000-160, pp. 1-6, 2001.
- [22] 高橋大介, 平岡正徳, 鳥居恵多, 岡本教佳, “2D-3D システムによる実時間为目标とした顔姿勢推定”, 電子情報通信学会技術研究報告, PRMU2003-267, pp. 67-72, 2004.
- [23] 松井俊樹, 菅沼直樹, 藤原直史, 景山一郎, 栗谷川幸代, “ステレオビジョンによる顔表面 3D モデリングと頭部姿勢計測”, 日本機械学会論文誌 C, vol.72-724, pp.118-123, 2006.
- [24] 山上晃弘, 石井雅博, 小池康晴, 佐藤誠, “画像認識のための新しいパラメトリック識別空間の提案”, 電子情報通信学会技術研究報告, HIP99-78, pp. 15-22, 2000.
- [25] 中野雄介, “判別空間での姿勢多様体を用いた顔画像の姿勢推定”, 映像情報メディア学会技術報告, 26(7), pp. 7-12, 2002.

- [26] 井上紀文, 鹿島雅之, 佐藤公則, 高橋行俊, “自己組織化特徴写像による表情認識”, 電子情報通信学会技術研究報告, PRMU2001-154, pp.7-12, 2001.
- [27] 石井雅樹, 佐藤和人, 間所洋和, 門脇さくら, 西田眞, “顔の動的位相変化に着目した表情空間モデルの最適化に関する検討”, 電子情報通信学会技術研究報告, TL2004-82, PRMU2004-250, pp.109-114, 2004.
- [28] 石井雅樹, 佐藤和人, 間所洋和, 門脇さくら, 西田眞, “顔パターンの階層的位相変化に着づく表情空間マップの形成”, 画像の認識・理解シンポジウム(MIRU2007), pp.935-940, 2004.
- [29] Kohonen.T, “Self-organization and associative memory”, Series in Information Sciences, Vol.8 , 1984.
- [30] Kohonen.T, “The Self-organization map”, EEE, ASSP-78-9, pp.1464-1480, 1990.
- [31] 田中利幸, 斉藤正男, “Kohonen のモデルにおける学習の収束性について”, 電子情報通信学会技術研究報告, NC91-154, pp. 169-176, 1992.
- [32] 安藤哲也, 藤吉弘亘, 松岡智美, 梅崎太造, “コホーネン型ニューラルネットワークによる顔表情の自動生成”, 平成9年度電気関係学会東海支部連合大会講演論文集, 849, 1997.
- [33] 梅崎太造, 平野崇, 佐藤幸男, “3次元顔表情の自動生成”, 電気学会論文誌C, Vol.121-C, No.2, pp.417-422, 2001.
- [34] 加藤弘貴, 梅崎太造, 段野幹夫, 宮治正廣, “コホーネン型ニューラルネットワークによる顔表情の認識”, SSII07, IN2-05, 2007.
- [35] 佐藤信: “新版 官能検査ハンドブック”, 日科技連出版社, 1973.
- [36] 漢野救泰, 船田哲男: “有声音検出に基づくロンバート音声認識と工場における検反システムへの応用”, 電気情報通信学会論文誌, Vol.J85-D-II, No.5, pp.851-862, 2002.
- [37] 大瀧淳司, 小菅太, 三譯勉, 植原孝之, 藤江龍一, “異音を定量的に検出する手法の開発”, 富士通テン技報, Vol.46, pp.37 -46, 2005. “異音を定量的に検出する手法の開発”, 富士通テン技報, Vol.46, pp.37-46, 2005.
- [38] 堀内匡, 別府俊幸, 藤岡美博, 原元司, “音響信号に基づく不良シジミ貝判別に対するパターン認識手法の適用と評価”, 日本知能情報ファジィ学会誌, Vol.20, No.5, pp.817-822, 2008.

- [39] 岡田賢, 上條哲平, 石川稜威男, “ニューラルネットワークを用いた打音の特徴抽出”, 電子情報通信学会技術報告, EA, Vol.102, No.398, pp.1-6, 2002.
- [40] 池田和隆, 橋本良仁, 石黒雄大, “音・振動情報による官能検査の自動化法”, パナソニック電工技報, Vol.54, No.2, pp.42-48, 2006.
- [41] 溝口健二, “未知の異音を検出する異音判定技術 エンジン異音の市場流出防止を目指して”, オムロンテクニクス, Vol.45, No.1, pp.32-36, 2005.
- [42] 石原辰也, 大津展之, “高次局所自己相関特徴の自己回帰係数によるジェスチャー認識”, 電子情報通信学会技術報告, PRMU, Vol.103, No.737, pp.61-66, 2004.
- [43] 豊田崇弘, 長谷川 修, “テクスチャ識別のためのマスクパターンによる特徴抽出法”, 電子情報通信学会技術報告, PRMU, Vol.104, No.290, pp.77-84, 2004.
- [44] 白木孝義, 石黒勝彦, 深野亮, 鴨志田良和, 白井達也, 斉藤秀雄, 田浦健次朗, 大武美保子, 佐藤知正, 大津展之, “CHLAC 特徴と Grid コンピューティングを併用したリアルタイム動作認識”, 電子情報通信学会技術報告, PRMU, Vol.105, No.615, pp.97-102, 2005.
- [45] 南里卓也, 大津展之: “複数人動画像からの異常動作検出”, 電子情報通信学会技術報告, Vol.104, No.291, pp.9-16, 2004.
- [46] 岩田健司, 佐藤雄隆, 依田育士, 坂上勝彦, 大津展之, “超高速 CHLAC によるリアルタイム異常動作検出”, ViEW2006 講演論文集, pp.43-48, 2006.
- [47] 加藤俊祐, 滝口哲也, 有木康雄, “局所特徴量によるフィッシャー重みマップに基づく音素認識”, 電子情報通信学会技術報告, SP, Vol.106, No.444, pp.19-24, 2006.
- [48] 瀬戸洋一, “ユビキタス時代のバイオメトリクスセキュリティ”, 日本工業出版, 2003.
- [49] 瀬戸洋一, “総論ーバイオメトリック認証技術の最新動向ー”, OPTRONICS, No2, pp.88-93, 2006.
- [50] 瀬戸洋一, “サイバーセキュリティにおける生体認証技術”, 共立出版, 2002.
- [51] A. Jain, R. Bolle, and S. Pankanti eds, “BIOMETRICS Personal Identification in Network Society”, KluwerAcademic Publishers, 1999.
- [52] D. Maltoni, D. Maio, A. K. Jain, and S. Prabhakar, “Handbook of Fingerprint Recognition”, Springer, 2003.

- [53] 小松尚久, 内田薫, 坂野鋭, 和田誓一, 池野修一, “バイオメトリクスの精度評価”, 計測と制御, Vol.43, No.7, pp.539-543, 2004.
- [54] 木村聡仁, 梅崎太造, “季節変化に影響されないロバストな指紋照合法”, 情報科学リサーチジャーナル, Vol.6, pp.65-72, 1993.
- [55] 木村聡仁, 梅崎太造, 松本憲幸, “かすれ指紋を考慮した照合法”, 電子通信学会技術研究報告, PRMU98-209, No.1, pp.127-134, 1999.
- [56] 竹内英世, 梅崎太造, 松本憲幸, 平林勝巳, “かすれ指紋画像の判定・復元機能をもつ指紋照合法”, 電子情報通信学会論文誌, Vol.J87-D-II, pp.574-585, 2004.
- [57] 梅崎太造, 佐藤省三, 木村聡仁, 松本憲幸, “画像品質の変動を考慮した指紋照合法”, 電気学会論文誌 C, Vol.122, No.7, pp.1127-1136, 2002.
- [58] 竹内英世, 松本憲幸, 桑山清明, 梅崎太造, “光学式指紋スキャナにより採取された指紋画像の評価方法”, 電気学会論文誌 C, Vol.124, No.3, pp.761-768, 2004.
- [59] 松本憲幸, 藤吉弘亘, 梅崎太造, 浜田敏雄, “FFT 及び LPC 分析に基づく指紋照合法の評価”, 電子通信学会技術研究報告, PRU92-4, No4, pp.25-31, 1992.
- [60] 松本憲幸, 佐藤省三, 藤吉弘亘, 梅崎太造, “LPC 分析に基づく指紋照合法の評価”, 電気学会論文誌 C, Vol.122, No5, pp.799-807, 2002.
- [61] 松本憲幸, 梅崎太造: “2次元フーリエ変換に基づく指紋照合法の評価”, 電子通信学会技術研究報告, Vol.PRU93-39, No.134, pp.37-44, 1993
- [62] 藤吉弘亘, 梅崎太造, 竹内英世, 佐分利伸, 鈴木宣夫: “スペクトルの遷移確率を利用した指紋照合法”, 電子情報通信学会論文誌, D-II, J80-D-2(5), 1169-1177, 1997
- [63] 佐野恵美子, 松下雅仁, 笹川耕一, 春名正光, 前田卓志, 鹿井正博, 近江雅人, “指内部の光学特性を利用した指紋センサ”, 電気学会論文誌 C, Vol.127, No.4, pp.472-479, 2007.
- [64] 鳥居恵多, 岡本教佳, “非接触型指紋認証を目的とした構造線抽出法”, 映像情報メディア学会誌, Vol.59, No7, pp.1001-1010, 2005.
- [65] 鳥居恵多, 岡本教佳, “非接触型指紋認証のための Wavelet 変換による画像補正”, 電子情報通信学会総合大会講演論文集, 2006年 情報・システム, No2, pp.227, 2006.

- [66] 鳥居恵多, 岡本教佳, “非接触型指紋認証のための構造線抽出法の改良”, 電気学会論文誌 C, Vol.126, No8, pp.972-979, 2006.
- [67] “生体情報の可視化技術”, コロナ社
- [68] 小澤利行, 浅野薫, 沼田成弘, 蓮井康嗣, 高地泰浩, 石原謙: “近赤外分光画像計測法による血中ヘモグロビン濃度の無侵襲測定”, 生体医工学会, Vol. 43, No.1, pp.93-102, 2005.
- [69] 高谷淳, 北間正崇, 加藤祐次, 清水孝一, 山本克之, 金子守: “後方散乱光による体内機能情報取得のための基礎的検討—低コヒーレント光干渉応用の可能性—”, 電子情報通信学会技術報告, MBE95-177, 1996.
- [70] Yeon-Sin Kim, Mee-Eum Kim, Ki-Suk Kim: “Healing Effect on Infected Wound by Two Different Wavelengths and Output Energies of GaAs Semiconductor Diode Laser”, A Comparative Study, Korean Journal Medicine Vol.28, No.4, 2003

研究業績一覧

学術論文（査読有）

- [1] **服部公央亮**, 田口亮, 梅崎太造, “自己組織化マップを用いた顔方向の認識”, 日本顔学会誌, Vol.10, No.1, p.29-36, 2010.
- [2] **服部公央亮**, 近江泰志, 田口亮, 梅崎太造, “CHLAC と連続 DP マッチングを用いた機械駆動異音の自動検出”, 電気学会論文誌, Vol.131, No.2, Sec.C, pp.367-374, 2011.
- [3] 加藤嗣, **服部公央亮**, 野村卓也, 田口亮, 保黒政大, 梅崎太造, “毛髪を含む頭部形状の 3 次元計測”, 電気学会論文誌, Vol.131, No.12, Sec.C (採録決定)
- [4] 蒲谷嘉代子, 竹村景史, 渡邊暢浩, 江崎伸一, 佐藤雄二, 田中史子, 中山明峰, 羽柴基之, **服部公央亮**, 梅崎太造, 村上信五, “Quantitative evaluation of acceleration in the vestibular organ”, Nagoya Medical Journal (採録決定)

国際会議（査読有）

- [5] **Koosuke Hattori**, Ryo Taguchi, Taizo Umezaki, Masashi Hayashi, Tomoyasu Saigo, Susumu Shibata, Masahiro Hoguro, Mikio Kuzuya, “Development of Smart Security System”, Proc. of 2010 IEEE Nanotechnology Materials and Devices Conference Joint 1st IEEE International Symposium on Energy, Environment, Safety and Security, pp.183-188, 2010.
- [6] Ryo Taguchi, Yuji Yamada, **Koosuke Hattoki**, Taizo Umezaki, Masahiro Hoguro, Naoto Iwahashi, Kotaro funakoshi, Mikio Nakano, “Learning Place-Names from Spoken Utterances and Localization Results by Mobile Robot”, Proc. of INTERSPEECH2011, pp.1325-1328, 2011.

研究会・学会（査読有）

- [7] **服部公央亮**, 加藤弘貴, 田口亮, 梅崎太造, “コホーネン型ニューラルネットワークによる顔方向の認識”, ViEW 2008 (Vision Engineering Workshop), B-7H(I-17), pp. 70-75, 2008.
- [8] 山本琢麿, **服部公央亮**, 田口亮, 梅崎太造, “ニューラルネットワークを用いた3次元顔画像の自動生成”, SSII 2009 (Symposium on Sensing via Image Information), IS4-10, 2009.
- [9] **服部公央亮**, 山口みなみ, 田口亮, 梅崎太造, 保黒政大, 葛谷幹夫, “非接触型指紋入力装置による個人認証システム”, ViEW 2009 (Vision Engineering Workshop), A-7H(I-5), 2009.
- [10] **服部公央亮**, 石川諒治, 田口亮, 梅崎太造, 保黒政大, “近赤外 LED を用いた指静脈認証システムの開発”, SSII 2010 (Symposium on Sensing via Image Information), IS4-21, 2010.
- [11] **服部公央亮**, 石川諒治, 田口亮, 梅崎太造, 保黒政大, “静脈画像を用いた個人認証システムの開発”, MIRU2010 (Meeting on Image Recognition and Understanding), IS3-61, 2010.
- [12] 蒲谷嘉代子, 渡邊暢浩, 宝来慶, 田中史子, 佐藤雄二, 竹村景史, 中山明峰, 羽柴基之, 村上信五, **服部公央亮**, 梅崎太造, “6軸センサを用いた平衡器官への負荷に関する研究 その4”, 第69回日本めまい平衡医学会総会・学術講演会, P46, 2010.
- [13] **服部公央亮**, 田口亮, 梅崎太造, 保黒政大, “頷き, 首振りなどのジェスチャ認識システムの開発”, DIA2011 (Dynamic Image processing for real Application), I2-26, 2011.
- [14] **服部公央亮**, 田口亮, 梅崎太造, 保黒政大, “自己組織化マップを用いた頭部ジェスチャ認識システム”, SSII 2011 (Symposium on Sensing via Image Information), IS1-21, 2011.
- [15] 廣田隼, **服部公央亮**, 田口亮, 保黒政大, 梅崎太造, “GPGPUを用いたグラウンド利用者計数システムの開発”, SSII 2011 (Symposium on Sensing via Image Information), IS3-20, 2011.
- [16] 村瀬雅之, **服部公央亮**, 田口亮, 保黒政大, 梅崎太造, “筆跡に含まれる個人性を学習する書道ロボットの開発”, SSII 2011 (Symposium on Sensing via Image Information), IS4-22, 2011.
- [17] 本野明大, 兼松佳弘, 石川坂幸, **服部公央亮**, 梅崎太造, “X線による金属部品内部欠陥検査システムの開発”, SSII 2011 (Symposium on Sensing via Image Information), IS4-28, 2011.
- [18] 竹村景史, 蒲谷嘉代子, 宝来慶, 田中史子, 佐藤雄二, 中山明峰, 渡邊暢浩, 羽柴基之, 村上信五, **服部公央亮**, 梅崎太造, “6軸センサを用いた平衡器官への負荷に関する研究 その5”, 第70回日本めまい医学会総会・学術講演会, P5-11, 2011.

- [19] **服部公央亮**, 近藤真司, 田口亮, 梅崎太造, 保黒政大, “透過型指静脈認証システムにおける照射輝度自動補正の検討”, ViEW 2011 (Vision Engineering Workshop), 2011.
- [20] 廣田隼, **服部公央亮**, 田口亮, 保黒政大, 梅崎太造, “頭・肩検出によるグラウンド利用者数計測システムの開発”, ViEW 2011 (Vision Engineering Workshop), 2011.
- [21] ウキエツ, 三谷大志, **服部公央亮**, 田口亮, 梅崎太造, 保黒政大, 本野明大, “生産情報自動読み取りシステムの開発”, ViEW 2011 (Vision Engineering Workshop), 2011.
- [22] 平野俊幸, 山田貴哉, **服部公央亮**, 田口亮, 柴田進, 保黒政大, 堀米秀嘉, 梅崎太造, “非侵襲内部計測のための共焦点レーザー顕微鏡の開発”, ViEW 2011 (Vision Engineering Workshop), 2011.

学会（査読なし）

- [23] **服部公央亮**, 梅崎太造, 竹内良徳, “アーク溶接部の自動不良判別”, 平成 19 年度電気関係学会 東海支部連合大会, 0-289, 2007.
- [24] **服部公央亮**, 加藤弘貴, 田口亮, 梅崎太造, “コホーネン型ニューラルネットワークによる顔方向の認識”, 平成 20 年度電気関係学会 東海支部連合大会, 0-140, 2008.
- [25] **服部公央亮**, 近江泰志, 田口亮, 梅崎太造, “連続 DP マッチングと CHLAC を用いた機械駆動異音の検出”, 平成 21 年度電気関係学会 東海支部連合大会, 0-201, 2009.
- [26] **服部公央亮**, 近江泰志, 田口亮, 梅崎太造, “サウンドスペクトログラムの立体高次局所自己相関による機械駆動異音の検出”, 平成 21 年 電気学会 電子・情報・システム部門大会, MC10-6, 2009.
- [27] 丸山哲裕, **服部公央亮**, 田口亮, 梅崎太造, 保黒政大, “低コスト住宅向け画像監視システム”, 生活生命支援医療福祉工学系学会連合大会 WWLS (Welfare, Wellbeing, Life Support) 2010, 2D3-3, 2010.
- [28] **服部公央亮**, 田口亮, 梅崎太造, 中村隆, 鈴木光久, 林真司, “低価格 3 次元計測器とレーザー焼結機を用いた簡易義足の造形”, 生活生命支援医療福祉工学系学会連合大会 WWLS (Welfare, Wellbeing, Life Support) 2010, 2D4-3, 2010.
- [29] **服部公央亮**, 田口亮, 梅崎太造, 保黒政大, 蒲谷嘉代子, 渡邊暢浩, 村上信五, “頭部加速度・角速度センサからの内耳負荷推定”, 平成 22 年度電気関係学会, 東海支部連合大会, 講論集, E3-3, 2010.
- [30] 梅崎太造, **服部公央亮**, 田口亮, 保黒政大, 葛谷幹夫, 林真司, 西郷知泰, 柴田進, “スマートセキュリティシステム”, 平成 22 年 電気学会 電子・情報・システム部門大会 TC14-5, 2010.

- [31] 山田雄治, **服部公央亮**, 田口亮, 梅崎太造, 保黒政大, 岩橋直人, 船越孝太郎, 中野幹生, “連続音声から場所名を学習する自律移動ロボット”, 情報処理学会 第73回全国大会, 1ZB-3, 2011.
- [32] 梅崎太造, **服部公央亮**, 田口亮, 保黒政大, 葛谷幹夫, “統合型スマートセキュリティシステム”, 平成23年電気学会電子・情報・システム部門大会 TC15-2, 2011.
- [33] **服部公央亮**, 田口亮, 梅崎太造, 保黒政大, “SOMを用いた首振り・頷きのリアルタイム認識システム”, 平成23年度電気関係学会東海支部連合大会, C4-5, 2011.
- [34] 川井将人, **服部公央亮**, 田口亮, 梅崎太造, 本野明大, 兼松佳弘, “X線画像を用いた不良品検査に関する研究”, 平成23年度電気関係学会東海支部連合大会, D4-1, 2011.
- [35] 于慕悦, **服部公央亮**, 田口亮, 梅崎太造, 保黒政大, 近藤悠太, “インライン型カシメ形状検査装置の開発”, 平成23年度電気関係学会東海支部連合大会, D4-3, 2011.
- [36] 平野俊幸, 山田貴哉, **服部公央亮**, 田口亮, 梅崎太造, 保黒政大, “非侵襲内部計測のための共焦点レーザ顕微鏡の開発”, 平成23年度電気関係学会東海支部連合大会, P1-6, 2011.
- [37] 廣田隼, **服部公央亮**, 田口亮, 梅崎太造, 保黒政大, “GPGPUを用いたクラウド利用者の検出”, 平成23年度電気関係学会東海支部連合大会, Po2-38, 2011.

表彰

- [1] 社団法人電気学会 平成21年電子・情報・システム部門大会 奨励賞
- [2] 名古屋工業大学 平成22年度名古屋工業大学学生表彰 (学術活動部門) 学長表彰

採択奨励研究費

- [1] 堀科学芸術振興財団 2010年度 研究助成 一般の部 75万円