

判別分析の幾何的解釈と楽器特徴抽出への適用

井原 瑞希^{†1} 池田 和司^{†1} 前田 新一^{†2}

簡潔かつ判別精度の高い楽器特徴の抽出は音楽情報検索において重要な課題である。本論文では、判別分析がそのような楽器特徴を抽出可能である理由について考察する。単旋律楽曲においては様々な特徴抽出法が提案されている。パワースペクトルを楽器特徴として使用してもよいが、特徴次元が高すぎて冗長となる可能性が高い。様々な応用を考えると、重要な情報のみを用いた高い楽器判別が好ましい。このような特徴をパワースペクトルから抽出する方法として、前処理として主成分分析 (PCA) によってパワースペクトルの次元削減を行い、その特徴に対して線形判別分析 (LDA)、もしくは LDA に局所性を導入した局所フィッシャー判別分析 (LFDA) による次元削減法をとる。また、これらの方法が適している理由に関して分析する。

Geometrical View of Linear Discriminant Analysis and Its Applications to Instrument Feature Extraction

MIZUKI IHARA,^{†1} KAZUSHI IKEDA^{†1} and SHIN-ICHI MAEDA^{†2}

Extracting only the essential sound attributes from sounds is one of the fundamental issues of music information retrieval. This paper analyzes the discriminant analysis methods and demonstrates their high applicability to the instrument identification. For monophonic music instrument identification, various feature extraction and selection methods have been proposed. Although raw power spectra have enough information for accurate instrument identification, their dimensionality is too high and redundant. It is important to find non-redundant instrument specific characteristics that maintain information essential for high-quality instrument identification to apply them to various instrumental music analyses. As such a dimensionality reduction method, combinations of linear projection methods is introduced: principal component analysis (PCA) followed by linear discriminant analysis (LDA) or local Fisher discriminant analysis (LFDA). Additionally, the reason why linear discriminant analysis algorithms are suitable for instrument identification is explained by the geometrical analysis of algorithms.

1. はじめに

少数の特徴によって音源同定ができれば、楽曲検索、採譜、音合成のような様々な分野に応用が可能であり、このような特徴抽出は音楽情報検索の重要課題の一つである。本研究では、判別分析に基づいた楽器特徴抽出法の提案を単旋律楽曲に適用し、判別分析が楽器特徴抽出に適している理由について考察する。

単旋律楽曲における楽器判別研究の中で頻繁に使われる、スペクトル情報からの楽器特徴抽出手法として、線形予測符合化 (LPC) 係数やメル周波数ケプストラム係数 (MFCC) などが挙げられる¹⁾⁻³⁾。Essid らの実験では、5 種類の木管楽器からそれぞれ 10 個の MFCCs を抽出し、楽器特徴として使用されている⁴⁾。さらに、Marques は MFCCs、線形予測係数 (LPCs) とケプストラム係数による楽器判別率をそれぞれ単独で比較し、MFCCs が最も高い認識率を示したと報告している¹⁾。最近では、Chétry が線スペクトル周波数 (LSFs) のみで楽器同定を行い、6 種類の楽器において LPCs や MFCCs よりも高い、86% の判別率をあげている³⁾。以上の実験は、すべて単旋律の楽曲を用いて実験が行われており、単音や多旋律の結果は含まれない。

本論文では、主成分分析 (PCA) によって前処理されたスペクトルから、二種類の判別分析、線形判別分析 (LDA) もしくは局所フィッシャー判別分析 (LFDA) に基づく方法により特徴抽出する方法を提案する。前者の特徴を *PCA-LDA*、後者を *PCA-LFDA* とよぶ。LDA や LFDA の数学的性質より、これらの手法が理論的に楽器特徴を抽出する方法として適していることを示す。また、実際に SVM によって楽器判別を行い複数の従来手法と比較することで、提案手法が楽器特徴の抽出において優れていることを示す。

2. 次元圧縮アルゴリズム

2.1 線形判別分析 (LDA)

線形判別分析 (*linear discriminant analysis*; LDA) は、識別対象から判別に有効な特徴を抽出する教師あり次元圧縮手法のひとつである。LDA では、 N 次元の入力ベクトル x から M

^{†1}

Graduate School of Information Science, Nara Institute of Science and Technology (NAIST)

^{†2}

Graduate School of Informatics, Kyoto University

次元の出力ベクトル \mathbf{y} ($N > M$) への特徴空間の線形変換

$$\mathbf{y} = \mathbf{W}^T \mathbf{x} \quad (1)$$

を行うことで判別に適した部分特徴空間を決定する．この式でのサイズ $M \times N$ の変換行列 \mathbf{W} の求め方は様々であり，LDA ではクラス内分散とクラス間分散比最大化基準

$$\mathbf{W}_{\text{lda}} = \underset{\mathbf{W}}{\operatorname{argmax}} \operatorname{tr} \left((\mathbf{W}^T \boldsymbol{\Sigma}_W \mathbf{W})^{-1} (\mathbf{W}^T \boldsymbol{\Sigma}_B \mathbf{W}) \right), \quad (2)$$

に基づいている．ここで $\boldsymbol{\Sigma}_B$ はクラス間共分散行列， $\boldsymbol{\Sigma}_W$ はクラス内共分散行列である．

C_i ， $\boldsymbol{\mu}_i$ ， n_i をそれぞれクラス i のサンプルセット，平均，サンプル数，また， C を総クラス数， $\boldsymbol{\mu}$ をすべてのサンプルの平均ベクトルとする． N 次元の入力ベクトル \mathbf{x} が与えられたとき，クラス間分散は各クラスの共分散で表現される．クラス間共分散 $\boldsymbol{\Sigma}_B$ とクラス内共分散 $\boldsymbol{\Sigma}_W$ はそれぞれ

$$\boldsymbol{\Sigma}_B = \sum_{i=1}^{|C|} n_i (\boldsymbol{\mu}_i - \boldsymbol{\mu})(\boldsymbol{\mu}_i - \boldsymbol{\mu})^T, \quad (3)$$

$$\boldsymbol{\Sigma}_W = \sum_{i=1}^{|C|} \boldsymbol{\Sigma}_w, \quad (4)$$

where

$$\boldsymbol{\Sigma}_w = \sum_{\mathbf{x} \in C_i} (\mathbf{x} - \boldsymbol{\mu}_i)(\mathbf{x} - \boldsymbol{\mu}_i)^T \quad (5)$$

で与えられる． \mathbf{W}_{lda} は解析的に求めることができる⁵⁾．

LDA 以外の代表的な次元圧縮手法として，特徴空間の線形変換に分散最大化基準を採用する主成分分析 (PCA) が挙げられる．PCA ではラベル情報は考慮せずにサンプル全体の特徴をうまく表現する空間を求める一方で，LDA は各クラスのサンプル集合を分離する空間を求めている．

本研究で取り扱う楽器判別問題に関しては，楽器の種類をラベル情報とすると教師あり学習が適している．しかし，入力ベクトルの次元数が大きすぎる場合や入力データが多峰性を持つ場合，学習データセットのサンプル数が少なすぎる場合において，LDA によって変換された特徴をクラス判別に用いても思わしくない分類結果となることがある^{5),6)}．

2.2 局所フィッシャー判別分析 (LFDA)

上記のような LDA の弱点より，LDA に局所保存射影 (locality preserving projection; LPP) を導入した教師あり次元圧縮法である局所フィッシャー判別分析 (local Fisher discriminant analysis; LFDA) が考案された．

LPP の射影行列 \mathbf{W}_{lpp} は，制約

$$\mathbf{A}^T \mathbf{X} \mathbf{D} \mathbf{X}^T \mathbf{A} = 1 \quad (6)$$

の元で重み付き自乗誤差

$$\begin{aligned} \mathbf{W}_{\text{lpp}} &= \underset{\mathbf{W}}{\operatorname{argmin}} \left(\frac{1}{2} \sum_{j,k}^N \|\mathbf{W}^T \mathbf{x}_j - \mathbf{W}^T \mathbf{x}_k\|^2 A_{jk} \right) \\ &= \underset{\mathbf{W}}{\operatorname{argmin}} (\mathbf{A}^T \mathbf{X} \mathbf{L} \mathbf{X}^T \mathbf{A}) \end{aligned} \quad (7)$$

を最小化する．ここで， \mathbf{X} はすべての N 個のサンプル行列であり， \mathbf{x}_j と \mathbf{x}_k はそれぞれ X の j 番目， k 番目の縦ベクトルである． \mathbf{L} はラプラス行列 $\mathbf{L} = \mathbf{D} - \mathbf{A}$ であり， \mathbf{D} は N 次元対角行列 $D_{jj} = \sum_{k=1}^N A_{jk}$ ($A_{jk} \in [0, 1]$) である．この \mathbf{D} を決定する類似度行列 A_{jk} を決める方法は様々であり⁷⁾ でも紹介されているが，本論文では以下のようなデータの局所スケーリング法を用いる．

$$A_{jk} = \exp \left(- \frac{\|\mathbf{x}_j - \mathbf{x}_k\|^2}{\sigma_j \sigma_k} \right). \quad (8)$$

ここで， $\sigma_{\{j,k\}} = \|\mathbf{x}_{\{j,k\}} - \mathbf{x}_{\{j,k\}}^{(m)}\|$ であり，この式中の $\mathbf{x}_{\{j,k\}}^{(m)}$ は $\mathbf{x}_{\{j,k\}}$ に対する m 番目の近傍である．射影前のベクトル \mathbf{x}_j と \mathbf{x}_k が近くに射影された場合，類似度行列 A_{jk} は大きな値をとる．LPP は元の空間で近い場所にあるサンプル同士を，出力空間でも近くの場所になるように射影する．つまり，局所性を保持しながら特徴空間の射影を行う方法である⁸⁾．

LFDA の変換行列は LDA の変換行列 (式 (2)) のクラス間共分散行列 $\boldsymbol{\Sigma}_B$ とクラス内共分散行列 $\boldsymbol{\Sigma}_W$

$$\tilde{\boldsymbol{\Sigma}}_B = \frac{1}{2} \sum_{j,k=1}^N \tilde{Q}_{jk}^{(B)} (\mathbf{x}_j - \mathbf{x}_k)(\mathbf{x}_j - \mathbf{x}_k)^T, \quad (9)$$

$$\tilde{\boldsymbol{\Sigma}}_W = \frac{1}{2} \sum_{j,k=1}^N \tilde{Q}_{jk}^{(W)} (\mathbf{x}_j - \mathbf{x}_k)(\mathbf{x}_j - \mathbf{x}_k)^T, \quad (10)$$

に，以下の式 (11) と式 (12) のような局所性 Q_{jk} が導入されている．

$$\tilde{Q}_{jk}^{(B)} = \begin{cases} A_{jk} \left(\frac{1}{N} - \frac{1}{n_c} \right) & (x_j = x_k = c) \\ \frac{1}{N} & (x_j \neq x_k) \end{cases}, \quad (11)$$

$$\tilde{Q}_{jk}^{(W)} = \begin{cases} \frac{A_{jk}}{n_c} & (x_j = x_k = c) \\ 0 & (x_j \neq x_k) \end{cases} \quad (12)$$

LFDA では，このように LPP の局所性を導入することで，同じクラスに属していてもサンプルの集中している箇所が複数ある，つまり多峰性のあるデータセットに対して，誤分類

を減らしている^{7),9)}。

3. 線形判別分析の幾何的解釈

楽器音の生成モデルはしばしばソースフィルタモデルで近似される^{10),11)}。ソースフィルタモデルでは、周波数 ω における音の周波数スペクトル $S(\omega)$ がソース $\alpha(\omega)$ とフィルタ $R(\omega)$ の積

$$S(\omega) = \alpha(\omega)R(\omega) \quad (13)$$

により表現される。ここで、 α はスペクトルの微細構造で表現される音源であり、基本周波数に依存する。また、 R はスペクトル包絡で表現される楽器固有の共鳴特性（調音）である。

すべての音源 α が基本周波数の倍数周波数上に倍音を持つという仮定より、ある音源（ここでは特定の楽器）の異なる高さの音 $\alpha'(\omega)$ は、その音源の基底 $\alpha(\omega)$ の周波数に関する積、つまり κ をスカラの係数としたとき $\alpha'(\omega) = \alpha(\kappa\omega)$ で表現できる¹²⁾。例えば、 $\alpha'(\omega)$ の基本周波数が $\alpha(\omega)$ の二倍であるとき、 $\kappa = 2$ となる。横軸を対数周波数、縦軸をスペクトル強度とした場合、対数周波数軸における離散スペクトルに関して、音高を変えたある音源の音 α' は、 α を要素ごとのシフトとして表現できる。図 1.A から図 1.B への対数周波数軸上におけるシフトはこのようなシフトの一例である。

$\alpha(\omega)$ と $R(\omega)$ それぞれを対数周波数軸に関して等間隔に離散化してベクトルで表現したものを α 、対角行列で表現したものを R とする。 R の対角行列の要素は対数周波数軸のシフト不変性に対応したフィルタ特性（図 1.C）である。以下、ある楽器 i の R は R_i とし、楽器 i の基底スペクトルを $R_i\alpha$ （図 1.D）

$N \times N$ の巡回シフト行列

$$S = \begin{pmatrix} 0 & 1 & 0 & \dots & \dots \\ 0 & 0 & 1 & \ddots & \dots \\ \vdots & \vdots & \ddots & \ddots & \ddots \\ 0 & 0 & \dots & \ddots & 1 \\ 1 & 0 & \dots & \dots & 0 \end{pmatrix} \quad (14)$$

が与えられたときの対数周波数軸に関して m 要素分のシフトを S^m と表現する。つまり、音高をシフトしたある音源の音は $\alpha' = S^m\alpha$ のように書ける。以上のことから、ある特定の楽器、かつ異なる音高をもつ音のスペクトルは、離散化が十分細かい場合に $R_iS^m\alpha$ （図 1.E）と表現できる。

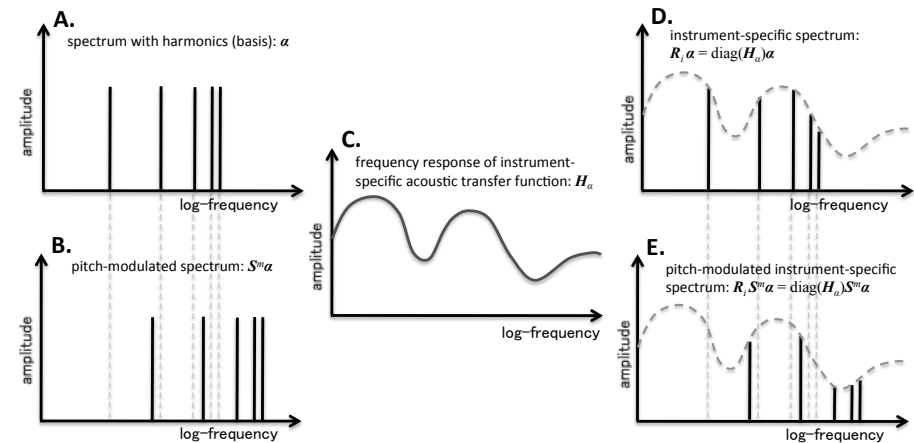


図 1 Basis spectrum (A) and its pitch-modulated spectrum (B) of one instrument through the instrument-specific transfer function (C) resulted in the instrument-specific spectrum (D) and the pitch-modulated instrument-specific spectrum (E) in the log-frequency scale

ここで R_i の逆行列 R_i^{-1} を楽器音スペクトルに掛けることで、 $R_i^{-1}RS^m\alpha = S^m\alpha$ のように音源情報のみ得られることになる。これは音の変化 m に対しては依存せず成立するために、それぞれの単位ベクトル 1 との内積をとっても値が一致、つまり単位ベクトルへの射影が一致し $1^T R_i S^m \alpha = 1^T R_i \alpha$ が成立する。図 2.A はこれを幾何的に描いた図である。

同様に、楽器 i とは異なる楽器 j の楽器音スペクトルを $R_j\beta$ と表現する。ここで β は α と同様に対数周波数に対して離散化した音源のスペクトルである。スペクトルベクトルが α である楽器 i のサンプルを $R_i\alpha^{-1}1$ に、 β である楽器 j のサンプルを $R_j\beta^{-1}1$ に理想的に射影されたとき、小さなクラス内分散が実現できる。これにより、楽器ごとの境界の決定が容易となる。

一方、LDA における射影は、大きなクラス間共分散を保持したままクラス内共分散を最小化する基準に基づいている。LDA の射影は幾何的には図 2.B のようになり、音の高さに依存しない楽器特徴を抽出していると考えられる。

また、可算的な観測ノイズのスペクトル n が楽器 i と j のサンプル集合 $R_iS^m\alpha$ と $R_jS^m\beta$ に対して加えられても、幾何的には図 2.B から図 2.C のように平行移動するだけである。LFDA

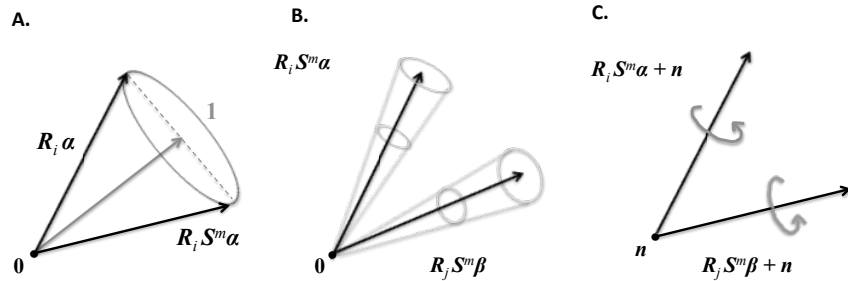


図2 楽器 i の基底スペクトル $R_i \alpha$ と同じ楽器であるが音高を変えたスペクトル $R_i S^m \alpha$ の、単位ベクトル 1 への射影の幾何的解釈

のアルゴリズムは LDA のアルゴリズムに対して局所性の導入以外の変化はない．そのため、バイアスが同クラスのサンプル x_j と x_k に対して加えられたとしてもクラス内共分散 (式 (10)) は変化せず、LDA がもつシフト不変性を保つ．

4. 判別手法

サポートベクタマシン (*support vector machines*; SVM) は機械学習の分野でしばしば用いられるパターン分類器である．分類結果の良さから、単音、単旋律、多旋律に関わらず、これまでの楽器識別研究では頻繁に使われている分類手法である．Marques らは、SVM と混合ガウス分布モデル (*Gaussian Mixture Model*; GMM) による単音分類実験を行い、GMM よりも SVM が単音分類に適していることを示した¹⁾．同様に Agostini も、複数の分類方法を比較し、SVM の良さを提示した¹³⁾．これらの結果に基づき、本研究では SVM を採用する．

SVM はマージン最大化とカーネルトリックという二つの重要な概念に基づいている．サンプルが線形分離可能である場合、学習データを正確に分類する超平面は数多く存在する．未知のテストデータも正確に分類するために、それぞれのクラスで境界に一番近いサンプルと境界の距離を最大化する．これをマージン最大化とよぶ．また、サンプルが線形分離不可能である場合、サンプルを高次元の空間に射影することで線形分離可能な空間を作ることができる．しかし、高次元空間への写像は計算量がかかるため、SVM ではカーネル関数

$$k(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i) \cdot \phi(\mathbf{x}_j) \quad (15)$$

を採用することでこの問題を解決している¹⁴⁾．これをカーネルトリックという．ここで \mathbf{x}_i と \mathbf{x}_j は i 番目と j 番目の学習データ点ベクトル (テストデータ点ベクトルでも良い) $k(\cdot)$

はカーネル関数、 $\phi(\cdot)$ は写像した特徴空間のベクトルを表わしている．

以下の実験では、LIBSVM の matlab インターフェース¹⁵⁾ を使用した．SVM の種類は C-SVC であり、線形カーネルを用いた．SVM はもともと二値分類器であるが、複数組み合わせることで多値分類問題に対応している．対応法は複数あり、主に “one-against-one” が “one-against-all” とよばれる方法が使用される．LIBSVM は学習にかかる時間的な計算量を考慮した結果から one-against-one を採用している．

5. 評価実験

LDA や LFDA によって変換された特徴が楽器判別に適していることを示すため、単旋律楽曲から抽出した特徴を用いて楽器判別を行った．単旋律楽曲の十分なデータベースが無いために単純に従来手法と比較することは不可能であるが、特定の学習データに過学習しないように複数の CD や RWC データベースを用い、スタンダードな楽器特徴抽出法との比較を行った．留意する点として、前述の幾何的解釈では、対数周波数軸にそって等間隔に離散化したスペクトルベクトルに関して LDA ベースの変換をかければ、同楽器で異なる音高をもつサンプル集合のばらつきを小さくすることができるということを主張している．今回の評価実験では、対数音圧-周波数軸のスペクトルに関して LDA ベースの変換をかけたものを楽器特徴として用いている．

5.1 実験に使用したデータ

実験には様々なジャンルの RWC データベース¹⁶⁾ や商用 CD から得たサンプルを使用した．使用楽器は 8 種類 (バイオリン、チェロ、ギター、ピアノ、フルート、オーボエ、ホルン、トランペット) であり、3 種類の楽器カテゴリ、弦楽器、木管楽器、金管楽器からそれぞれ少なくとも 2 種類の楽器を含むように選んだ．サンプリング周波数 44.1kHz の 30 枚の単旋律 CD から 47 種類の楽曲を選び^{*1}、重複しないように総数 38507 (vn: 6612, vc: 5005, gt: 6524, pf: 5366, fl: 5783, ob: 3008, hr: 2498, tp: 3711) の長さ 0.046 秒のサンプルを作成した．無音や音量の小さなサンプルはあらかじめ取り除いてあるが、音高や奏法に関する制限はない．

実験に使用する学習データセットとテストデータセットのサンプルの分け方に二種類の方法を用いた．一つめは単純に各楽器の総サンプルの半分を学習データ、残り半分をテストデータとする方法、二つ目は 47 種類の楽曲のうち一つをテストデータ、残りを学習データ

*1 同じ CD でも異なる楽器で演奏された楽曲が含まれている場合がある．

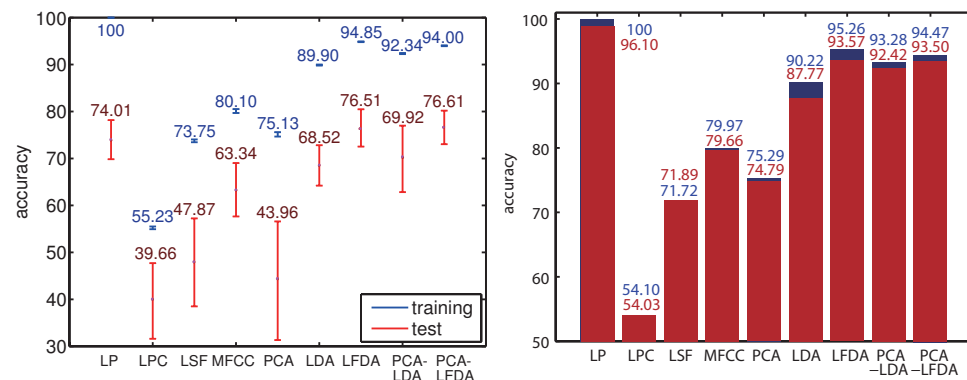


図 3 8 種類の楽器判別を学習データ (赤), テストデータ (青) それぞれに対して 10 トライアル行った平均結果 (左は leave-1CD-out, 右は mixed-CD)

とする方法である．本論文では前者を *mixed-CD*, 後者を *leave-1CD-out* とよぶことにする．

特徴ベクトルの作成方法は次の様に求める．まず，音信号にハミング窓をかけてフーリエ変換した後に対数を取り，1,024 次元の対数パワースペクトルを求める．LDA ベースの変換を用いた場合，学習データに対して過学習する可能性があるため，前処理として PCA をかける．以下，それぞれの手法によって抽出された特徴を PCA-LDA, PCA-LFDA とよぶ．パワースペクトルの次元を PCA によって 310 次元 (PCA-LDA), 563 次元 (PCA-LFDA) に次元削減したのち，LDA あるいは LFDA によって 10 次元に次元削減したものを楽器特徴として用いる．

5.2 楽器判別実験

提案手法，PCA-LDA, PCA-LFDA, LDA, LFDA の他に，元の 1,024 次元対数パワースペクトル (LP), 線形予測分析係数 (LPC)¹⁾, 線スペクトル対係数 (LSF)³⁾, メル周波数ケプストラム係数 (MFCC)³⁾, PCA を比較手法として用いた．対数パワースペクトル以外は，それぞれ提案手法と同様に 10 次元としている．*Leave-1CD-out* のサンプル振り分け方法での判別結果は表 3 左, *mixed-CD* での判別結果は表 3 右に示す．赤が学習データ，青がテストデータの判別結果である．

leave-1CD-out と *mixed-CD* の結果を比較すると，*mixed-CD* の結果は全ての特徴において *leave-1CD-out* の結果よりも高くなっている．この結果は，従来研究でも示唆されているように *mixed-CD* において，楽器特徴ではなくある CD に依存した特徴を抽出している可能性

があると言える^{1),17)}．いくつかの従来手法では学習データとテストデータに同じ CD を用いていることから判別結果が高くなっている可能性がある^{2),18)-20)}．*Leave-1CD-out* におけるそれぞれの実験結果を比較したところ，提案手法の楽器判別精度が高いことが分かる．これらのことから，提案手法によって，判別に適した楽器特徴抽出が可能であると言える．

参考までに，従来研究の判別結果とデータについて表 1 にまとめておく．この表の“source

著者	楽器数	特徴数	判別精度 (%)	信頼区間	source mixed	サンプルの長さ (秒)
<i>leave-1CD-out</i>	8	10	76.61	73.77-83.39	n	0.046
<i>mixed-CD</i>	8	10	93.42	89.92-94.41	y	0.046
Marques ¹⁾	8	16	70	n/a	n	0.2
Livshin ¹⁷⁾	7	62	88	81-94	n	1.0
Chétry ³⁾	6	16	86	72-98	n	300
Eggink ¹⁹⁾	6	120	66	56-85	n,y	2-10
Essid ²¹⁾	10	19	87	66-100	y	0.5
Jinachitra ²⁰⁾	6	28	66	n/a	y	0.5
Ventura ²⁾	5	12	99	97-100	y	10
Brown ¹⁸⁾	4	10	n/a	79-84	y	2.0-7.8

表 1 Summary of classification results of existing studies (y: 学習とテストで同じ CD から得たサンプルを使用, n: 異なる CD から得たサンプルを使用)

mixed”中の“y”は学習データとテストデータに同じ CD から得られたサンプルを用いており，“n”は異なる CD から得られたサンプルを用いている．両方ある場合は両方の実験設定で実験されている．

6. おわりに

本論文では，判別分析に基づく次元圧縮手法の幾何的解釈を示し，これらが楽器判別に適していることを主張した．判別分析に基づく手法は理論的に同楽器サンプルのばらつきを小さくするような射影を行うため，音高に依存しない特徴を求めることができる．この判別分析に基づく手法によって抽出された特徴と頻繁に使用される従来手法を用い SVM により楽器判別した結果，提案手法が従来手法よりも高い判別精度となった．また，LFDA, PCA-LFDA によって抽出された 10 次元の特徴に関しては，元の 1,024 次元の対数スペクトルの特徴よりも高い判別結果を示した．これらのことから，判別分析に基づく次元圧縮手法は楽器特徴抽出に有効であると考えられる．

参 考 文 献

- 1) J.Marques and P.J. Moreno, "A study of musical instrument classification using Gaussian mixture models and support vector machines," Tech. Rep., Compaq Computer Corporation, June 1999.
- 2) R.Ventura-Miravet, F.Murtagh, and J.Ming, "Pattern recognition of musical instruments using hidden markov models," in *Stockholm Music Acoustics Conference (SMAC)*, Aug. 2003, pp. 667–670.
- 3) N. Chétry and M. Sandler, "Linear predictive models for musical instrument identification," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2006, vol.5, pp. 5083–5086.
- 4) S.Essid, G.Richard, and B.David, "Musical instrument recognition based on class pairwise feature selection," in *Proc. of International Conference on Music Information Retrieval (ISMIR)*, Oct. 2004.
- 5) C.M. Bishop, *Pattern Recognition and Machine learning*, Springer Science+Business Media, LLC, New York, NY, Feb. 2006.
- 6) K.Fukunaga, Ed., *Introduction to Statistical Pattern Recognition*, Academic Press. Inc., Boston, 2nd edition, 1990.
- 7) M.Sugiyama, "Dimensionality reduction of multimodal labeled data by local Fisher discriminant analysis," Tech. Rep., Department of Computer Science, Tokyo Institute of Technology, Japan, 2006.
- 8) X.He and P.Niyogi, "Locality preserving projections," *Advances in Neural Information Processing Systems*, vol. 16, 2004.
- 9) A.M. Martínez and A.C. Kak, "PCA versus LDA," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 2, Feb. 2001.
- 10) A.S. Bregman, Ed., *Auditory Scene Analysis: the perceptual organization of sound*, The MIT Press, Hoboken, New Jersey, Sept. 1994.
- 11) G.Fant, *Acoustical Theory of Speech Production: With Calculations based on X-Ray Studies of Russian Articulations*, The Hague, Mouton, 1970.
- 12) S.Sagyama, K.Takahashi, H.Kameoka, and T.Nishimoto, "Specmurt anasyllis: A piano-roll-visualization of polyphonic music signal by deconvolution of log-frequency spectrum," in *Workshop on Statistical and Perceptual Audio Processing*, Oct. 2004, p. 128.
- 13) G. Agostini, M.Longari, and E.Pollastri, "Musical instrument timbres classification with spectral features," in *Proc. of European Conference on Signal Processing (EUSIPCO)*, 2003, vol.1, pp. 5–14.
- 14) V.Vapnik, *The Nature of Statistical Learning Theory*, Springer, New York, 1995.
- 15) C.C. Chang and C.J. Lin, *LIBSVM: a library for support vector machines*, 2001, <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- 16) "RWC music database," <http://staff.aist.go.jp/m.goto/RWC-MDB>.
- 17) A.Livshin and X.Rodet, "Musical instrument identification in continuous recordings," in *Proc. of International Conference on Digital Audio Effects (DAFx)*, Oct. 2004.
- 18) J.C. Brown, O.Houix, and S.McAdams, "Feature dependence in the automatic identification of musical woodwind instrument," *Journal of Acoustical Society of America*, vol. 109, no. 3, pp. 1064–1072, Mar. 2001.
- 19) J.Eggink and G.J. Brown, "Application of missing feature theory to the recognition of musical instruments in polyphonic audio," in *Proc. of International Conference on Music Information Retrieval (ISMIR)*, Oct. 2003, pp. V–553–556.
- 20) P.Jinachitra, "Polyphonic instrument identification using independent subspace analysis," in *Proc. of International Conference on Multimedia and Expo (ICME)*, June 2004, IEEE Computer Society.
- 21) S.Essid, G.Richard, and B.David, "Musical instrument recognition by pairwise classification strategies," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 14, no. 4, pp. 1401–1412, July 2006.