社団法人 情報処理学会 研究報告
IPSJ SIG Technical Reports

2008－SLP－74（8）
2008 / 12 / 9

制約付き非負行列因子分解を用いた音声特徴抽出の検討

朴　玄信† 滝口 哲也† 有木 康雄†

† 神戸大学大学院工学研究科 〒657-8501 兵庫県神戸市灘区六甲台町 1-1
E-mail: †silentbattle@me.cs.scitec.kobe-u.ac.jp, ††{takigu,ariki}@kobe-u.ac.jp

あらまし 本稿では，非負行列因子分解を用いた，音声認識のための特徴量抽出手法と，非負行列因子分解における，新しい初期化手法の提案を行う．非負行列因子分解は，非負の制約を用いて，ローカルな特徴抽出を得意とする．テキストや画像，音響データに対する応用がなされているが，本稿では音声特徴抽出に用いる．また，最近非負行列因子分解の初期化手法として特異値分解やクラスタリング法などを用いた手法が提案されているが，本稿では相関情報を用いた新しい初期化手法についても述べる．非負行列分解の付加制約としては，基底ベクトルのスパースネスを考慮した更新アルゴリズムを用いた．非負行列因子分解の初期化手法の性能比較実験では，提案手法が推定誤差と単語音声認識率で有効性を示した．MFCC または，主成分分析，独立成分分析などの特徴量との単語認識比較実験においても，非負行列因子分解を用いた特徴抽出法の有効性が確認された．

キーワード 非負行列因子分解，相関伝播初期化，音声特徴抽出，孤立単語音声認識

Speech Feature Extraction Using Constrained Nonnegative Matrix Factorization

Hyunsin PARK†, Tetsuya TAKIGUCHI†, and Yasuo ARIKI†

† Graduate School of Engineering, Kobe University
Rokkodaicho 1-1, Nada-ku,Kobe,Hyogo, 657-8501 Japan
E-mail: †silentbattle@me.cs.scitec.kobe-u.ac.jp, ††{takigu,ariki}@kobe-u.ac.jp

Abstract In this paper, we propose a speech feature extraction approach using nonnegative matrix factorization (NMF), and a new initialization approach for NMF. NMF is excellent at extracting local feature using nonnegative constraint. This method has been applied to data of text, image, or sound. In this paper, we apply this method to speech feature extraction. Recently, initialization methods based on SVD or clustering methods for NMF have been proposed. In this paper, a new initialization method using correlation information for NMF is described. A sparseness of base vector is additionally constrained to NMF. In comparison experiment among initialization methods, our proposed method shows effectiveness at estimate error and word speech recognition. The new feature extraction method also shows effectiveness at word speech recognition compared to MFCC and features by principal component analysis or independent component analysis.

Key words Nonnegative matrix factorization, Correlation propagation initialization, Speech feature extraction, Isolated word speech recognition

1. はじめに

近年の音声認識システムで一般的に使われる MFCC は事前学習無しの特徴量抽出法である．この特徴量は，正規化手法などと組み合わせることで，高い認識率を示している．しかし，観測される音声信号には音素時系列情報だけではなく，発話者の身体や感情の情報，環境情報など様々な情報が混在する．このような情報の中から，音声認識システムが主に必要とする情報は音素時系列情報である．MFCC などの事前情報を用いない特徴量抽出法は，音声認識に不必要な情報の対処に，特徴量の後処理や音響モデルまたは言語モデルに頼らざるを得ない．この問題を解決するため，主成分分析 [1]，判別分析 [2]，独立成分分析 [3] などの統計的手法をベースにした特徴抽出法が提案されていて，効果が確認されている．

近年，テキストや画像などの非負行列に対して，非負制約を用いた因子分解手法が提案されている．この非負行列因子分

解手法については，[4] で効率的な更新アルゴリズムがはじめて紹介されており，[5] でスパースネス制約を導入した手法が提案されている．[6] は，非負行列因子分解の様々なアルゴリズムと応用分野がまとめられている．最近では，非負行列因子分解における初期化問題が注目されており，[7] [8] [9] などで，いくつかの初期化手法が提案されている．また，音響データに対しても [10] などでも，非負行列因子分解が用いられている．非負行列因子分解は非負の制約により，従来のデータ解析手法に比べ，よりローカルな特徴 (他の文献ではパーツやブロックと表現している) 抽出が可能であり，音声の新しい特徴抽出にも期待できる．

本稿では，2つの新しい提案を行う．1つ目は，相関情報を用いた非負行列因子分解の初期化手法である．2つ目は，非負行列因子分解を用いた音声認識のための特徴抽出法である．評価実験としては，各初期化手法の推定誤差と単語音声認識率の比較を行い，また，主成分分析や独立成分分析などによる音声特徴量との比較も行う．

以降の2章では非負行列因子分解を概説し，3章では1つ目の相関情報を用いた初期化手法，4章では2つ目の非負行列因子分解を用いた音声特徴抽出法について述べる．5章で，評価実験の条件と結果を報告し，最後の6章で，結論と今後の課題について述べる．

2. 非負行列因子分解

非負行列因子分解（以下 NMF と呼ぶことにする.）とは非負行列 $\mathbf{X} \in \Re^{m \times n}$ を二つの非負行列因子 $\mathbf{W} \in \Re^{m \times r}$ と $\mathbf{H} \in \Re^{r \times n}$ に分解する手法であり，次式のように表す．

$$\mathbf{X} \approx \tilde{\mathbf{X}} \equiv \mathbf{WH}, \tag{1}$$

ここで，$\tilde{\mathbf{X}}$ は $\mathbf{X}$ の近似解である．

NMF は目的関数，初期化手法，更新ルール，付加制約により様々なバリエーションがある．この章では，評価実験で用いた手法について主に述べる．

2.1 目的関数

NMF の目的関数としては，$\mathbf{X}$ と $\tilde{\mathbf{X}}$ 間のユークリッド距離または，カルバック・ライブラー情報量が用いられている．ユークリッド距離 D_E はフロベニウス・ノルム ($\| \mathbf{A} \|_F \equiv \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2}$) を用いて，次式のように定義される．

$$D_E(\mathbf{X}, \tilde{\mathbf{X}}) = \| \mathbf{X} - \mathbf{WH} \|_F^2 . \tag{2}$$

また，カルバック・ライブラー情報量 D_K は次式のようになる．

$$D_K(\mathbf{X}, \tilde{\mathbf{X}}) = \sum_{i=1}^{m} \sum_{j=1}^{n} (x_{ij} \log \frac{x_{ij}}{[\mathbf{WH}]_{ij}} - x_{ij} + [\mathbf{WH}]_{ij}). \tag{3}$$

ここで，$[\mathbf{A}]_{i,j}$ は $\mathbf{A}$ の i 行 j 列成分を表す．本研究では，ユークリッド距離の目的関数を用いる．

2.2 初期化

NMF の初期化手法としては，ユークリッド距離 [7]，またはカルバック・ライブラー情報量 [8] を目的関数とした，Spherical K-Means などのクラスタリング手法による初期化と，特異値分解を用いた初期化 [9] が最近提案された．

表 1　SKM (Spherical K-Means) for NMF Initialization

Step 1. Normalize each column vector $\mathbf{X}_i$ to unit L_1-norm.
Step 2. Initialize k centroids $\mathbf{c}_j$ for $j = 1, \cdots, k$.
Step 3. Iterate until convergence.
Compute $d_{ij} = \mathbf{x}_i^T \mathbf{c}_j$ for $i = 1, \cdots, n$ and $j = 1, \cdots, k$.
Update partitioning $\Pi_j = \{X_i \mid \mathrm{argmax}_i \, d_{ij}\}$ for $j = 1, \cdots, k$.
Update each centroids $\mathbf{c}_j = \sum_i \mathbf{x}_i / \| \sum_i \mathbf{x}_i \|$.
Step 4. $\mathbf{W} = \mathbf{C}$ and randomize positive $\mathbf{H}$.

表 1 は SKM (Spherical K-Means) 法を用いた初期化アルゴリズムである．データサンプルのノルムの正規化を行った後，各クラスタの中心ベクトル $\mathbf{c}_j$ をランダム初期化する．サンプルとクラスタ間の距離を計算した後，クラスタの集合を更新し，中心ベクトルの更新も行う．収束するまで更新を行った後，クラスタの中心ベクトルの集合を，NMF の $\mathbf{W}$ とし，$\mathbf{H}$ はランダム初期化を行う．

表 2　NNDSVD (Non-Negative Double Singular Value Decomposition) for NMF initialization

Step 1. Compute SVD: $\mathbf{X} = \mathbf{U\Sigma V}$
Step 2. Initialize $\mathbf{w}_1 = \mathbf{u}_1 \times \sqrt{\sigma_{1,1}}$ and $\mathbf{h}_1 = \mathbf{v}_1 \times \sqrt{\sigma_{1,1}}$.
Step 3. for j from 2 until k
$x = \mathbf{u}_j, y = \mathbf{v}_j$,
$px = pos(x), py = pos(y), nx = neg(x), ny = neg(y)$,
$pn = norm(xp) \times norm(yp)$, $nn = norm(nx) \times norm(ny)$,
if $pn > nn$, $u = px/norm(px)$, $v = py/norm(py)$, $sigma = pn$.
else $u = nx/norm(nx)$, $v = ny/norm(ny)$, $sigma = nn$.
$\mathbf{w}_j = u \times \sqrt{\sigma_{j,j}}$ and $\mathbf{h}_j = v \times \sqrt{\sigma_{j,j}}$.
Step 4. Set all zeros of $\mathbf{W}$ and $\mathbf{H}$ equal to the average of all elements of $\mathbf{X}$.

$pos(x)/neg(x)$ returns vector with only positive/negative elements and 0. $norm(x)$ returns L_1-norm.

表 2 は [9] での特異値分解を用いた初期化手法，NNDSVD (Non-Negative Double Singular Value Decomposition) である．まず，データ行列に対して，特異値分解を行う．第1特異値と対応する特異ベクトルの積を，$\mathbf{w}_1$(列ベクトル) と $\mathbf{h}_1$(行ベクトル) に代入する．第1特異ベクトルの成分は非負であるが，第2特異値からは，特異ベクトルの負の成分を取り除き，$\mathbf{W}$ と $\mathbf{H}$ を構成していく．最後に 0 成分を $\mathbf{X}$ の全成分の平均値に置き換える．

2.3 更新ルール

ユークリッド距離の目的関数に対する乗算更新ルールは次式のようになる．

$$w_{il} \leftarrow w_{il} \frac{[\mathbf{XH}^T]_{il}}{[\mathbf{WHH}^T]_{il}}, \tag{4}$$

$$h_{lj} \leftarrow h_{lj} \frac{[\mathbf{W}^T\mathbf{X}]_{lj}}{[\mathbf{W}^T\mathbf{WH}]_{lj}}. \tag{5}$$

また，カルバック・ライブラー情報量の目的関数に対する乗

算更新ルールは次式のようになる．

$$w_{il} \leftarrow w_{il} \sum_j \frac{x_{ij}}{[\mathbf{WH}]_{ij}} h_{lj}, w_{il} \leftarrow \frac{w_{il}}{\sum_j w_{jl}}, \quad (6)$$

$$h_{lj} \leftarrow h_{lj} \sum_i \frac{x_{ij}}{[\mathbf{WH}]_{ij}} w_{il}. \quad (7)$$

2.4 付加制約

本研究では，$\mathbf{W}$ の各列ベクトル $\mathbf{w}$ に次式のスパースネス尺度を満たすように制約を与えている．

$$sparseness(\mathbf{w}) = \frac{\sqrt{m} - (\sum |w_i|)/\sqrt{\sum w_i^2}}{\sqrt{m} - 1} \quad (8)$$

この式は，L_1 ノルムと L_2 ノルム間の関係に基づくスパースネス尺度になっている．

3. 相関情報を用いた非負行列因子分解初期化

この章では，相関情報を用いた非負行列因子分解初期化手法を提案する．相関情報を伝播していくようになっているため，本稿では相関伝播 (Correlation Propagation: CP) 初期化と呼ぶことにする．

表 3 CP (Correlation Propagation) for NMF initialization

Step 1. Set $\mathbf{W}^0 = \mathbf{E}$ and $\mathbf{H}^0 = \mathbf{X}$.
Step 2. For t from 1 until $m-k$
  Compute correlation matrix $\mathbf{R}^{t-1}$ of $\mathbf{H}^{t-1}$.
  $d = argmax_i(\sum_j r_{i,j}^{t-1})$,
  $c_l^{t-1} = r_{d,l}^{t-1} / \sum_j r_{d,j}^{t-1}$,
  $w_{i,l}^t = w_{i,l}^{t-1} + w_{i,d}^{t-1} \times c_l$,
  $h_{l,j}^t = h_{l,j}^{t-1} + h_{d,j}^{t-1} \times c_l$.
Step 3. $\mathbf{W} = \mathbf{W}^{m-k}$ and $\mathbf{H} = \mathbf{H}^{m-k}$.

提案手法のアルゴリズムを表 3 に示す．ステップ 1 で，$\mathbf{W}^0$ を単位行列 $\mathbf{E}$，$\mathbf{H}^0$ をデータ行列 $\mathbf{X}$ に初期化する．$\mathbf{X} = \mathbf{EX} = \mathbf{W}^0\mathbf{H}^0 = \tilde{\mathbf{X}}^0$ であり，誤差は 0 である．ステップ 2 では，m 次元を k 次元に削減する場合，$m-k$ 回の更新を行う．まず，前回の $\mathbf{H}^{t-1}$ の行ベクトル間の相関行列 $\mathbf{R}^{t-1}$ を求め，他次元との相関和が一番高い次元を削除する次元 d として選択する．削除した次元 d の情報を他の次元 $l(l = 1, \cdots, m-k)$ へ相関比に基づいて伝播させる．そのときの相関比は c_l^{t-1} である．t 回目の $\mathbf{W}$ と $\mathbf{H}$ の各要素は，$t-1$ 回目の要素に削除される要素と相関比の積が加算される．t 回目の $\tilde{\mathbf{X}}$ の i 行 j 列要素 $[\mathbf{WH}]_{i,j}^t$ は次式のように展開される．

$$\begin{aligned}
&[\mathbf{WH}]_{i,j}^t \\
&= \sum_{l=1}^{m-t} w_{i,l}^t h_{l,j}^t \\
&= \sum_{l=1}^{m-t} (w_{i,l}^{t-1} + w_{i,d}^{t-1} c_l)(h_{l,j}^{t-1} + h_{d,j}^{t-1} c_l) \quad (9) \\
&= \sum_{l=1}^{m-t} (w_{i,l}^{t-1} h_{l,j}^{t-1} + w_{i,l}^{t-1} h_{d,j}^{t-1} c_l + w_{i,d}^{t-1} h_{l,j}^{t-1} c_l + w_{i,d}^{t-1} h_{d,j}^{t-1} c_l^2) \\
&= [\mathbf{WH}]_{i,j}^{t-1} + w_{i,d}^{t-1} h_{d,j}^{t-1} \{ \sum_{l=1}^{m-t} (c_l^2 + (\frac{w_{i,l}^{t-1}}{w_{i,d}^{t-1}} + \frac{h_{l,j}^{t-1}}{h_{d,j}^{t-1}}) c_l) - 1 \}
\end{aligned}$$

最後の行の第 2 項は，$\tilde{\mathbf{X}}$ の要素ごとに加算される相関情報であり，$\tilde{\mathbf{X}}^0 = \mathbf{X}$ であるため，各反復での相関伝播による誤差項でもある．第 2 項の累積が，提案手法による誤差値であり，削減する次元数 (反復数) とのトレードオフになる．

4. 非負行列因子分解を用いた音声特徴抽出

NMF は主にテキストや画像データに対して応用されてきた．最近では，音声データに対しても応用されているが，その分野は音源分離などと限られている．本研究では NMF を音声特徴抽出に用いる．音声認識の分野で広く使われている特徴量は MFCC である．この特徴量を抽出する際，音声データに対する事前知識（言語や収録環境など）は用いられていない．そのため，事前知識を用いたデータ依存型特徴抽出法は，音声認識において，MFCC より高い性能を示す傾向がある．計算量においても，線形変換を用いたデータ依存型特徴抽出法は，MFCC と比較しても劣らない．従来の線形変換によるデータ依存型特徴抽出には，主成分分析，線形判別分析，独立成分分析などの統計的手法を用いたものがある．

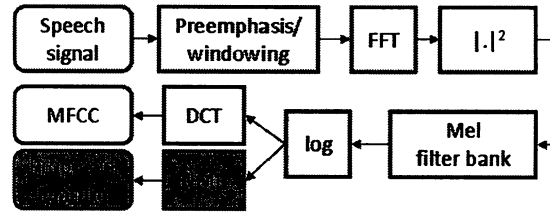

図 1 Feature extraction flow

NMF を用いた特徴抽出のフローを図 1 に示す．音声信号に対し，窓処理と高速フーリエ変換を行い，パワースペクトルを求める．その後，メル周波数上でのフィルタバンク演算の出力に対数変換を行って得られるベクトル $\mathbf{x}$ を NMF の入力データとして用いる．対数をとることで，負になる可能性はあるが，そのときは 0 に置き換える．実験では負になるケースはなかった．データ行列 $\mathbf{X}$ は音声データから音素バランスを考慮しながらランダム抽出して構成する．$\mathbf{X}$ に対して NMF を行い，$\mathbf{W}$ と $\mathbf{H}$ が得られる．スパースネス制約は $\mathbf{W}$ だけに与える．$\mathbf{W}$ の列ベクトルは基底ベクトル，$\mathbf{H}$ の列ベクトルはこのベクトル空間上での 1 つの点にあたる．ひとつのサンプルで表現すると

$$\mathbf{x} = \mathbf{Wh} + \epsilon \quad (10)$$

のようになる．ϵ は NMF による誤差項である．ここで，$\mathbf{h}$ を特徴量としたいので，一般化逆行列を用いて上の式を書き直すと

$$\mathbf{h} = (\mathbf{W}^T\mathbf{W})^{-1}\mathbf{W}^T\mathbf{x} - \epsilon' \quad (11)$$

である．本来，ϵ' は時変誤差項であるが，推定が困難であるため，本研究では定常であるとみなし，$\mathbf{h}$ の後処理として正規化を行い削除する．従来の特徴抽出で用いられている Δ 係数もそのまま用いることにする．

5. 評価実験

5.1 実験条件

評価実験のデータとして，ATR の日本語データベース A-set から男女 5 名ずつ計 10 名話者のデータを用いた．この音声データはクリーン環境で収録された孤立単語発話データであり，音素系列のラベルには各音素の境界時刻情報も含まれている．

音声信号は 16 bits，12 kHz でサンプリングされ時間の長さが 32 ms のハミング窓を 8 ms シフトさせフレーム分析を行った．各フレームに離散フーリエ変換をした後，24 チャネルのメル周波数フィルタバンク分析を行った．本研究では，この 24 次元対数エネルギーを持つ空間をベースのベクトル空間とする．

学習データは 26,200(10×2,620) 単語，評価データは 10,000(10×1,000) 単語である．学習データと評価データの発話内容は異なる．特徴抽出のための学習データは上記の学習データから，音素バランスを保ってランダムに抽出する．データ行列 $\mathbf{X}$ のサイズは 24×5,075 である．

表 4　Initialization methods for NMF

RANDOM	N(0,1) に従うランダム絶対値による初期化
SKM	Spherical K-Means を用いた初期化
NNDSVD	NNDSVD を用いた初期化
CP	相関伝播を用いた初期化 (提案手法)

評価実験で用いた NMF 初期化手法を表 4 にまとめる．NMF の反復学習回数は 20,000 回とした．

表 5　Speech features

MFCC	離散コサイン変換
PCA	$\mathbf{X}$ に対し，PCA
ICA	$\mathbf{X}$ に対し，FastICA [11]
NMF	$\mathbf{X}$ に対し，NMF with CP

また，孤立単語音声認識実験で用いた特徴量を表 5 にまとめる．各特徴量は 12(=k) 次元のベクトルである．これらの特徴量には Δ 係数が加わり，各要素の平均を 0 にする正規化を行った．

音響モデルとしては，54 個のコンテキスト独立のモノフォン HMM を用いた．各 HMM は自己ループの 3 状態で各状態には 20 個のガウス分布が混合されている．HMM の学習と認識にも HTK toolkits [12] を用いた．

5.2 初期化手法による非負行列因子分解の性能比較

表 6 は NMF に対する 4 つの初期化手法による初期誤差を示している．提案手法 CP の初期誤差が他の 3 つの初期化手法に比べ，小さいことが確認できる．

表 6　CP for NMF initialization

Initialization	RANDOM	SKM	NNDSVD	CP
Error	3.5E+06	1.5E+06	2.0E+05	1.7E+04

図 2 は各初期化手法により得られた，$\mathbf{W}$ である．縦軸は

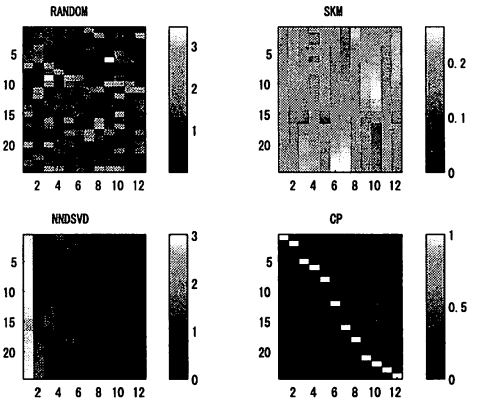

図 2　W by NMF initialization: RANDOM, SKM, NNDSVD, and CP.

m(= 24)，横軸は k(= 12) に対応する．各列が基底ベクトルであり，下が低周波数領域，上が高周波数領域を表す．各行列 $\mathbf{W}$ の最大値が白，最小値の 0 が黒になるように，色のスケーリングがされている．左上は RANDOM による結果である．右上は SKM による結果であり，各基底ベクトルはクラスタの平均値であるため，他の手法よりフラットな分布になっている．左下は NNDSVD による結果である．一番左の第 1 特異ベクトルに高い値が集まっており，他の領域は全体的に低い値になっている．右下は CP による結果である．対角線上に高い値が分布しており，途切れている周波数領域が，削除された基底ベクトルに対応する．右から 4 つの列ベクトルは，途切れることなく低周波数領域を強調しており，特にこの領域が他の領域と相関が低いことを表す．

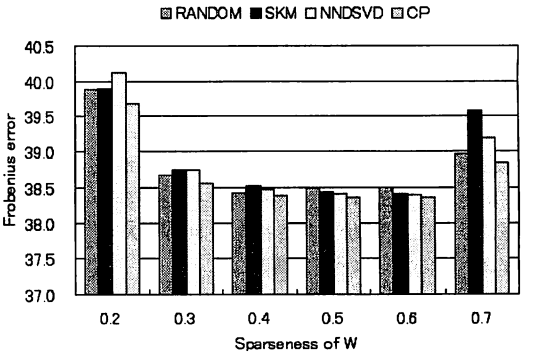

図 3　Frobenius errors after 20,000 iteration

図 3 は 2 万回反復後の誤差値である．すべてのスパースネス制約条件において，提案手法の CP 初期化が他の手法より，高い推定精度を示した．0.5 付近のスパースネス制約が，高い推定精度を示している．

図 4 に 2 万回反復後の $\mathbf{W}$ を示す．CP 初期化による $\mathbf{W}$ は，

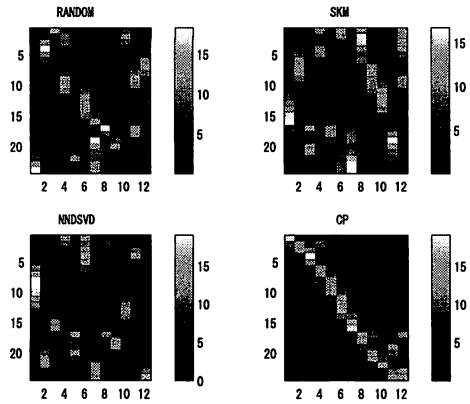

図 4　W after 20,000 iteration: RANDOM, SKM, NNDSVD, and CP.

初期の構造を保った状態で，対角線上に白の領域が繋がっている．特に，音声の低周波数領域を強調する複雑な構造のベクトルが右に並んでいる．他の３つの手法は，２万回の更新後に，初期の構造がほとんど見えなくなった．

5.3　孤立単語音声認識

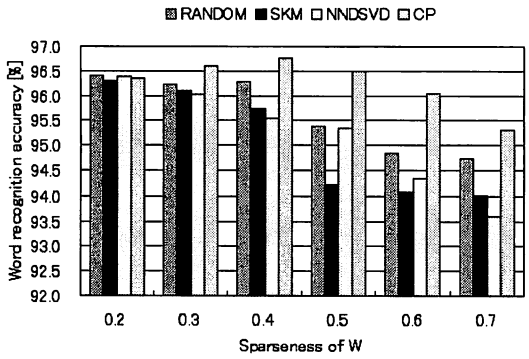

図 5　Comparison of recognition accuracy with RANDOM, SKM, NNDSVD and CP initialization algorithms.

図 5 は単語音声認識実験の認識率を表す．提案手法 CP が，ほとんどのスパースネス制約条件下で他手法より高い認識結果を示した．特に，0.4 のスパースネス制約条件で，一番高い認識率 96.8% が得られた．CP は，図 3 の誤差値が低いほど認識率が高くなるが，他の初期化手法は，誤差値に関係なく，全体的にスパースネス制約が強くなるにつれ，認識率が低下する傾向を見せた．

図 6 の左は主成分分析，右は独立成分分析により得られる $\mathbf{W}$ である．主成分分析によるものは，一番左が第 1 固有ベクトルであり，低周波数領域を強調している．高次になるにつれ，周波数強調領域がぼやけていく．独立成分分析によるものは，基底ベクトルの並び替えは行っていない．９番目の基底ベクトル

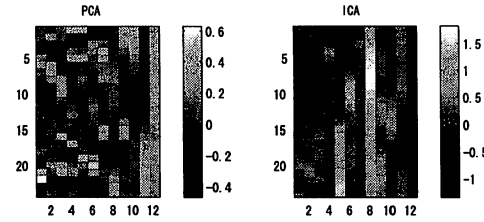

図 6　W by PCA (left) or ICA(right).

が低周波数，７番目の基底ベクトルが高周波数を強調しているが，局所的ではない．

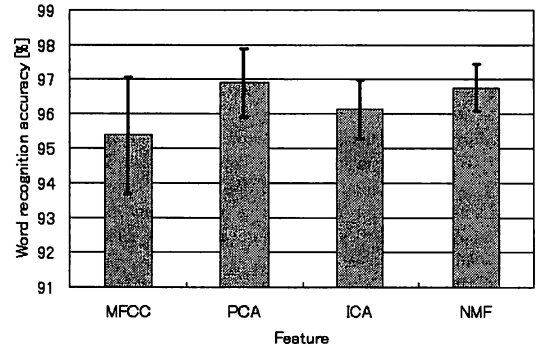

図 7　Comparison of MFCC, PCA, ICA and NMF speech features.

図 7 は，音声特徴抽出法の比較実験の結果を表す．話者 10 人の認識率を標本として，NMF の有意差検定 (t-検定，0.05 水準) を行った結果，MFCC や ICA との比較では，有意差があり，PCA との比較では，有意差がないことが確認された．

ここで，音声認識率における有意差はなかったが，NMF により得られる $\mathbf{W}$ と PCA により得られる $\mathbf{W}$ とは本質が異なる．PCA によるものは正規直交基底であるが，NMF によるものは直交性は保障されない．また，次元数 (k) を減らしていくと，PCA の場合，基底ベクトルが減っていくだけだが，NMF の基底は全体の構造が変わっていく．PCA は各基底ベクトルがグローバルな特徴表現能力を持つ反面，NMF には非負の制約が働いているため，ローカルな特徴表現能力を有していることが，図 4 と図 6 から分かる．

6. おわりに

NMF を用いた音声認識のための特徴抽出法を提案した．また，既存の NMF における初期化手法に代わる相関情報を用いた初期化手法も提案した．初期化手法間の性能比較実験では，推定誤差と単語認識率において，相関情報を用いる提案手法の有効性が確認された．また，特徴抽出法の比較実験でも，NMF を用いて提案手法の有効性が示された．

今後は，我々が提案してきた統合音素部分空間法 [13] との組み合わせによる特徴抽出法と，雑音や残響が混じっている音声データに対する有効性について，検討していく予定である．

文　献

[1] 滝口哲也，有木康雄，“Kernel PCA を用いた残響下におけるロバスト特徴量抽出の検討，” 情報処理学会論文誌，Vol.47，No.6，pp. 1767-1773，2006.

[2] 坂井誠，北岡教英，服部佑哉，中川聖一，武田一哉，“判別分析に基づく音響特徴と識別学習の組み合わせによる単語音声認識，” 日本音響学会 2008 年春季研究発表会論文集，pp. 193-194. 2008

[3] Jong-Hwan Lee, Ho-Young Jung, Te-Won Lee, Soo-Young Lee, “Speech feature extraction using independent component analysis,” Proc. ICASSP2000, Vol. 3, pp. 1631-1634. 2000.

[4] D.D. Lee, H.S. Seung, “Algorithms for Non-negative Matrix Factorization,” Neural Inf. Process. Syst. 13, pp. 556-562, 2001.

[5] P. Hoyer, “Non-negative Matrix Factorization with Sparseness Constraints,” J. Mach. Learning Res. 5, pp. 1457-1469, 2004.

[6] M.W. Berry, M. Browne, A.N. Langville, V.P. Pauca, R.J. Plemmons, “Algorithms and applications for approximate nonnegative matrix factorization,” Comput. Stat. Data Anal., Vol. 52, pp. 155-173, 2007.

[7] S. Wild, J. Curry, A. Dougherty, “Improving non-negative matrix factorizations through structured initialization,” Pattern Recognition, Vol. 37, pp. 2217-2232, 2004.

[8] Y. Xue, C.S. Tong, Y. Chen, W-S. Chen, “Clustering-based initialization for non-negative matrix factorization,” Applied Mathematics and Computation, Vol. 205, pp. 525-536, 2008.

[9] C. Boutsidis, E. Gallopoulos, “SVD based initialization: A head start for nonnegative matrix factorization,” Pattern Recognition, Vol. 41, pp. 1350-1362, 2008.

[10] Y-C. Cho, S. Choi, “Nonnegative features of spectro-temporal sounds for classification,” Pattern Recognition letters, Vol. 26, pp. 1327-1336, 2005.

[11] A. Hyvärinen and E. Oja, “Independent Component Analysis: Algorithms and Applications,” Neural Networks, vol. 13(4-5), pp. 411-430, 2000.

[12] S. Young et. al., “The HTK Book,” Entropic Labs and Cambridge University, 1995-2002.

[13] 朴玄信，滝口哲也，有木康雄，“MDL 基準と ICA を用いた統合音素部分空間による音声特徴量抽出の検討，” 日本音響学会 2008 年秋季研究発表会論文集，pp. 133-134，2008.