テキスト依存型話者照合のためのDP経路の利用*

○鎌田敏明[†‡], 峯松信明[‡], 長内　隆[†], 蒔苗久則[†], 谷本益巳[†] ([†]科警研, [‡]東大･新領域)

1　はじめに

我々が行っている話者照合[1,2]は，研究結果の求められる環境となる法科学の観点から，電話を通した音声を対象としている。これらの対象となる音声は，一般的に行われている研究と比較すると非常に悪い状況が多く，スペクトル包絡情報だけを利用した照合手法で高い照合率を得るのは難しい。

一方, 人間が聴覚で話者を認識する場合は，周波数構造だけでなく，ピッチやイントネーション，アクセントなどの韻律的特徴や，発話速度，方言などといった“話し方”についても判断の基準としている。低次のケプストラムによる高い認識率が得られない場合は，これらの“話し方”を第二の判断材料とすることにより，認識率を向上させることができるのではないかと考えられる。

そこで本研究では，“話し方”の特徴の１つがテキスト依存型話者照合における DP マッチングの経路に現れるという観点から，DP 経路を利用した話者照合実験を行い，話者照合において DP 経路が利用できる可能性について調べた。

2　DP マッチング

2.1 DP 経路

/KURUMA/の単語について，同一人及び別人の DP マッチングにおける経路の例を図 1 に示す。同一人と別人の DP 経路を比較すると，同一人の場合は経路と対角線の距離は小さく，別人の場合は大きい傾向がある。

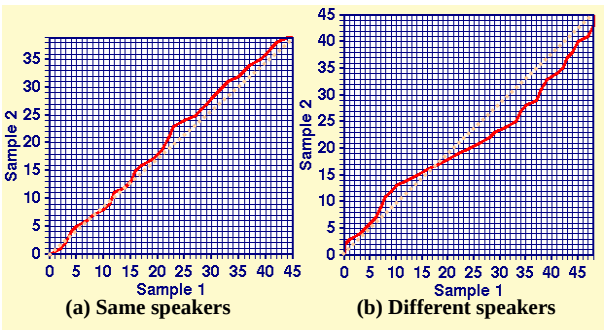

図1 同一人及び別人の DP 経路

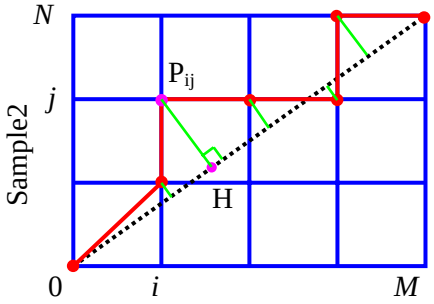

図2 DP 経路と対角線間の距離

2.2 DP 経路による話者間距離

図 2に DP 経路の模式図を示す。2 つの資料の時系列数を M+1, N+1 とし，DP 経路の点 P_{ij} が 2 資料の時系列の i, j 番目とすると，対角線との距離 $\overline{P_{ij}H}$ は式 1 のようになる。

$$\overline{P_{ij}H} = \frac{|Ni - Mj|}{\sqrt{M^2 + N^2}} \qquad (1)$$

DP 経路の全ての点に対する対角線との距離の総和を求め，対角線の長さ $\sqrt{M^2+N^2}$ で割って標準化を行い，これを DP 経路から得られた 2 資料間の距離とした。

3　DP 経路による話者照合

3.1 音声データ

照合実験に使用した音声データを表 1に示す。電話音声であることから，11.025 kHz から 8 kHz にダウンサンプリングを行った。

表1 音声資料

話者数	300 人
発声時期差，時期	3～4 ヶ月，3 時期
1 時期の発声回数	3 回
発声内容	5 単語（はい,車,連絡,爆弾,銀行），2 短文（爆弾を仕掛けた,逆探するな）
サンプリング周波数	8 kHz
量子化精度	16 bit

*Use of Dynamic Time Warping Function for Text-Dependent Speaker Verification, by KAMADA T.[†‡], MINEMATSU N.[‡], OSANAI T.[†], MAKINAE H.[†] and TANIMOTO M.[†] ([†]National Research Institute of Police Science, [‡]Graduate School of Frontier Sciences, The University of Tokyo).

表2 分析条件

分析窓	ハミング窓
フレーム長	32 ms
フレーム周期	16 ms
高域強調	1 次の適応型
分析方法	LPC 分析
分析次数	16 次
特徴量	LPC ケプストラム係数

3.2 話者照合実験

表 2 に示す分析を行い，テキスト依存型の話者照合実験を行った。DP マッチングを行い，DP 経路から得られた距離(trace distance)による同一人及び別人の距離分布を発話内容ごとに求めた。発話ごとの距離分布から，等誤り率となるしきい値を求め，このしきい値から照合率を得た。DP マッチングで利用する局所的な DP パスは，1/2～2 の傾斜制限を行うと最大 10% 以上の照合率の低下が見られたため，最も単純なパス(斜めの経路の重みは 2)を利用した。

3.3 話者照合結果

5 単語，2 短文による話者照合結果を図 2に示す。比較の為に，DP マッチングによる LPC ケプストラム係数のユークリッド距離 (DP distance) から得られた照合実験も行った。w01～w05 は 5 単語，s01, s02 は 2 短文である。

この結果から，trace distance のみによる話者照合は DP distance には及ばないものの，DP 経路から得られた距離(話し方)だけを用いることによって，ある程度の話者照合が可能であることがわかった。

4 DP 経路を利用した話者照合

4.1 DP 経路と DP 距離の統合

重み w により trace distance と DP distance の線形結合を行った。それを式 2 に示す。

$$dist = w\frac{dist_{trace}}{\mu_{dist_{trace}}} + (1-w)\frac{dist_{DP}}{\mu_{dist_{DP}}} \qquad (2)$$

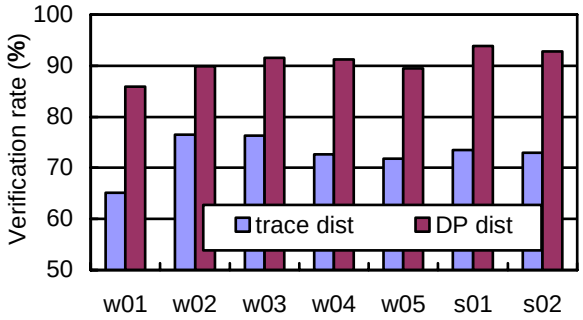

図3 DP 経路による話者照合結果

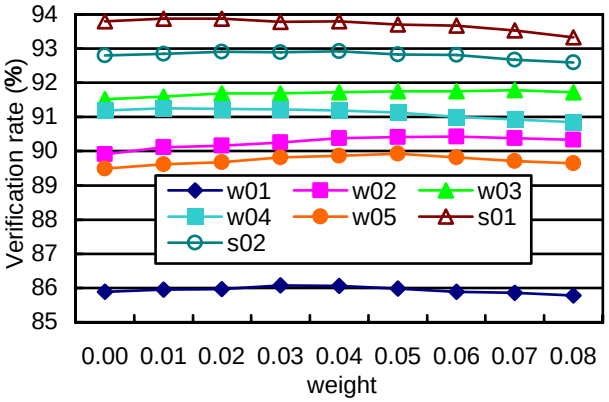

図4 DP 経路による話者照合結果

4.2 話者照合結果

式 2 から得られた距離により w を変化させた話者照合実験を行い，その結果を図 4 に示す。この結果から，5 単語及び 2 短文のいずれの場合においても，$w = 0.05$ 付近で照合率は最大となり，0.1～0.5 % 程度照合率が向上することがわかった。

5 まとめ

本研究では，“話し方”の特徴の 1 つが DP 経路に現れるという観点から，話者照合において DP 経路が利用できる可能性について調べた。その結果，ケプストラムの特徴量による DP 距離には及ばないものの，DP 経路のみを用いた話者照合がある程度可能であることがわかった。一方従来の DP 距離による話者照合に対して DP 経路を利用した場合は，わずかな照合率の向上は見られたものの，その効果は非常に低いことがわかった。

今後は DP 経路に対する発声時間の解釈や，DP 経路の話者間距離を数値化する手法について，経路と対角線によって得られた面積を使うなどの実験を行い，更に効果的な手法について検討する予定である。また従来の手法に対する効果的な DP 経路の利用方法についても調べる必要がある。

謝辞

本研究の一部は，日本学術振興会科学研究費補助金(課題番号:17500133)の助成を得た。

参考文献

[1] 長内, “テキスト依存型話者照合における判断基準の単語，話者依存性”，日鑑科技会，pp.117，2003.

[2] 鎌田，長内，蒔苗，谷本，“劣悪な雑音環境下における話者照合”，音講論(秋)，pp.130−131，2005.