

アプリケーション

- **雑音環境下**での話者クラスタリング

アプローチ

1. **因子分析モデル**により発話を表現

$$\mathbf{m}(\mathbf{o}_u) = \mathbf{m}_0 + \mathbf{T}\mathbf{w}(\mathbf{o}_u) + \varepsilon$$

- \mathbf{o}_u : u 番目の発話
- $\mathbf{m}_0(\mathbf{o}_u)$: 発話 \mathbf{o}_u で話者非依存GMMをMAP適応し得られたGMMのスーパーベクトル
- \mathbf{m}_0 : 話者非依存 GMM のスーパーベクトル
- \mathbf{T} : Total variability space
- $\mathbf{w}(\mathbf{o}_u)$: 発話 \mathbf{o}_u の 因子ベクトル (i-vector)
- ε : 残渣ベクトル

2. 因子ベクトルに**線形判別分析**を適用
3. 得られたベクトルを cosine 類似度基準で**スペクトラルクラスタリング**

話者クラスタリング実験結果

- 様々な雑音を重畳した音声で評価
- **クリーン音声と同程度の精度**でクラスタリングできることを示した

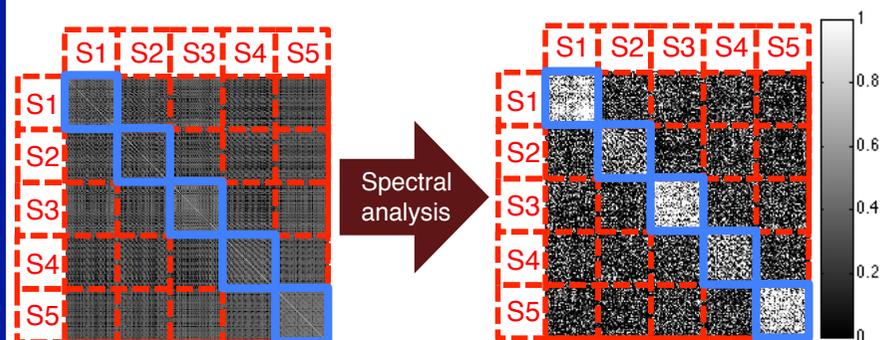


図. 雑音を重畳した音声の類似度行列

図. 低次スペクトラルを用いて得られた類似度行列

