

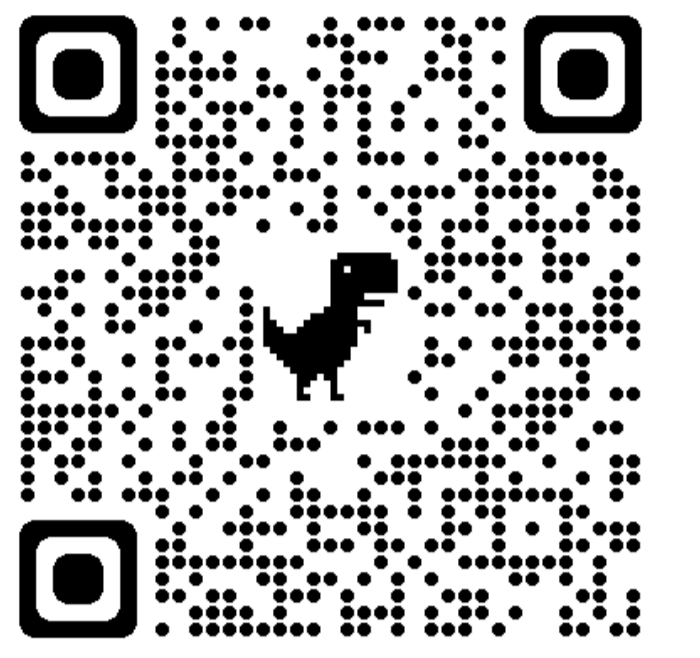
# [2-P-19] 音声モーフィングにおける自動対応付けの提案と品質評価

☆堀部貴紀, 森勢将雅 (明治大学)

## 研究背景 :

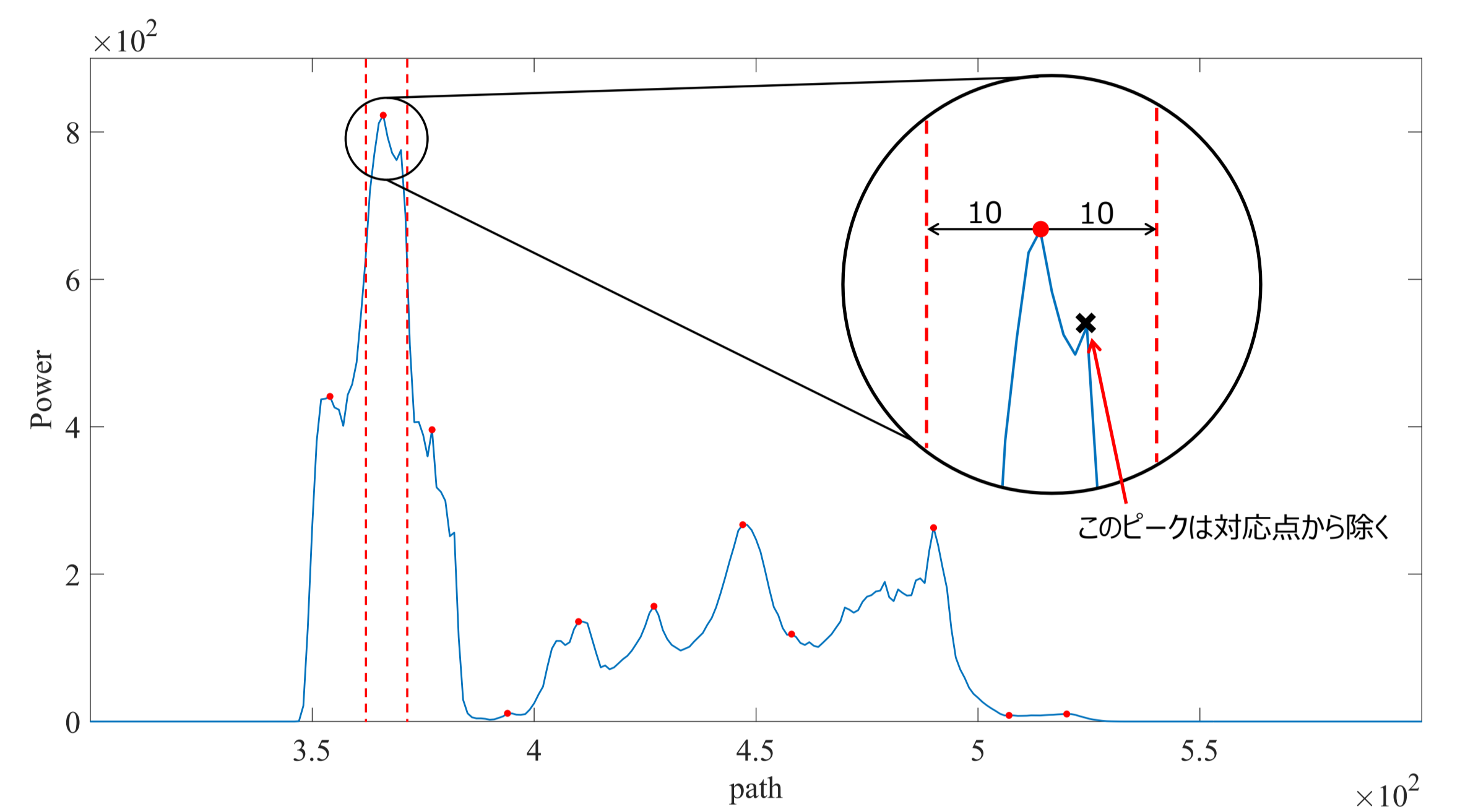
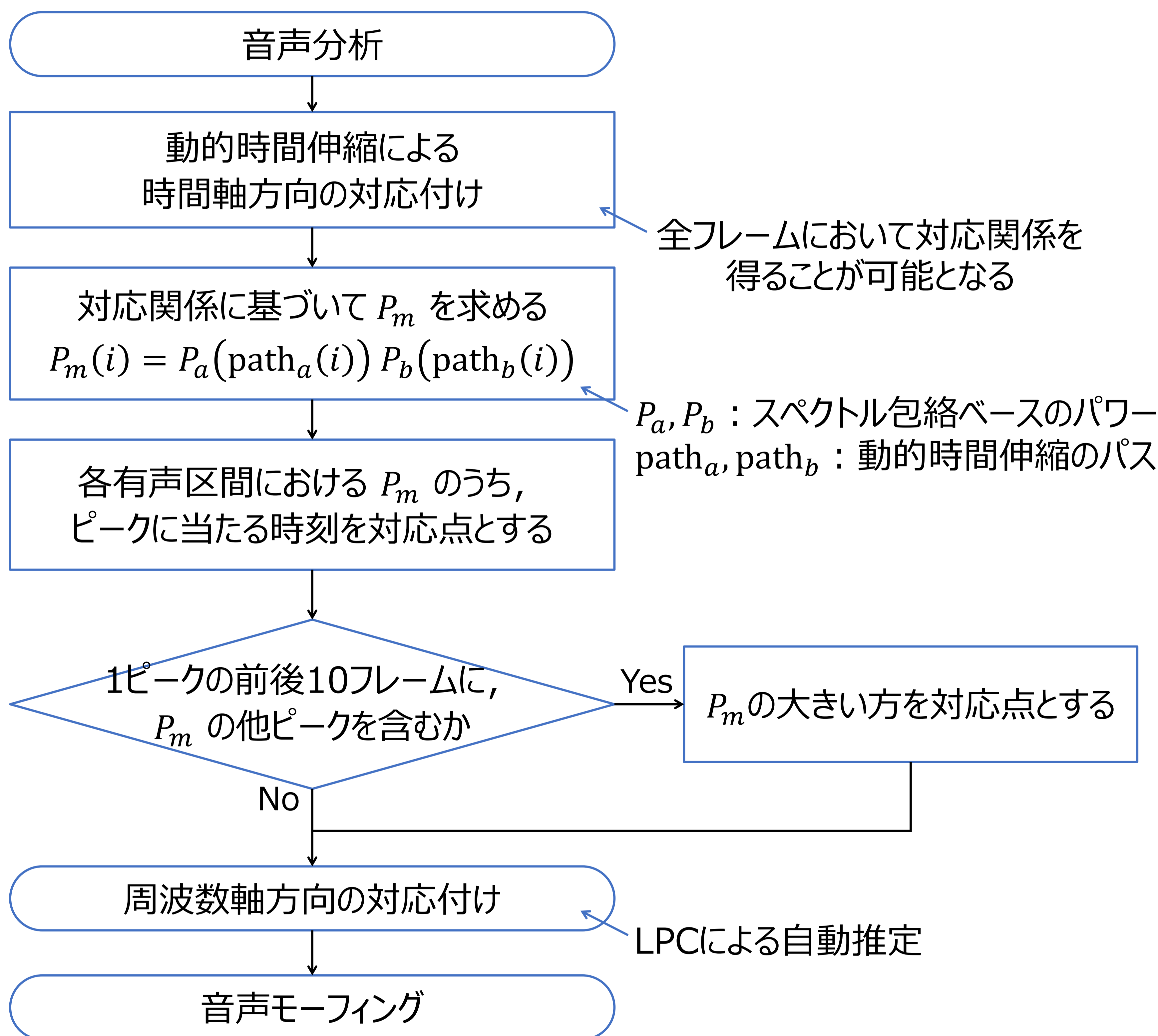
- 音声モーフィング : 同じテキストの2つの音声から, 中間的な印象を与える音声を合成する技術  
→ その過程で, 2つの音声を時間軸・周波数軸に対して対応付けを行う必要がある.
- 対応付け
  - 手作業 : モーフィング音声の品質が対応付けの精度  $\propto$  作業者の経験に依存する.
  - 従来手法 (音素境界) : 音素系列情報が必要である.

モーフィング音声は  
こちらより聴くことができます。

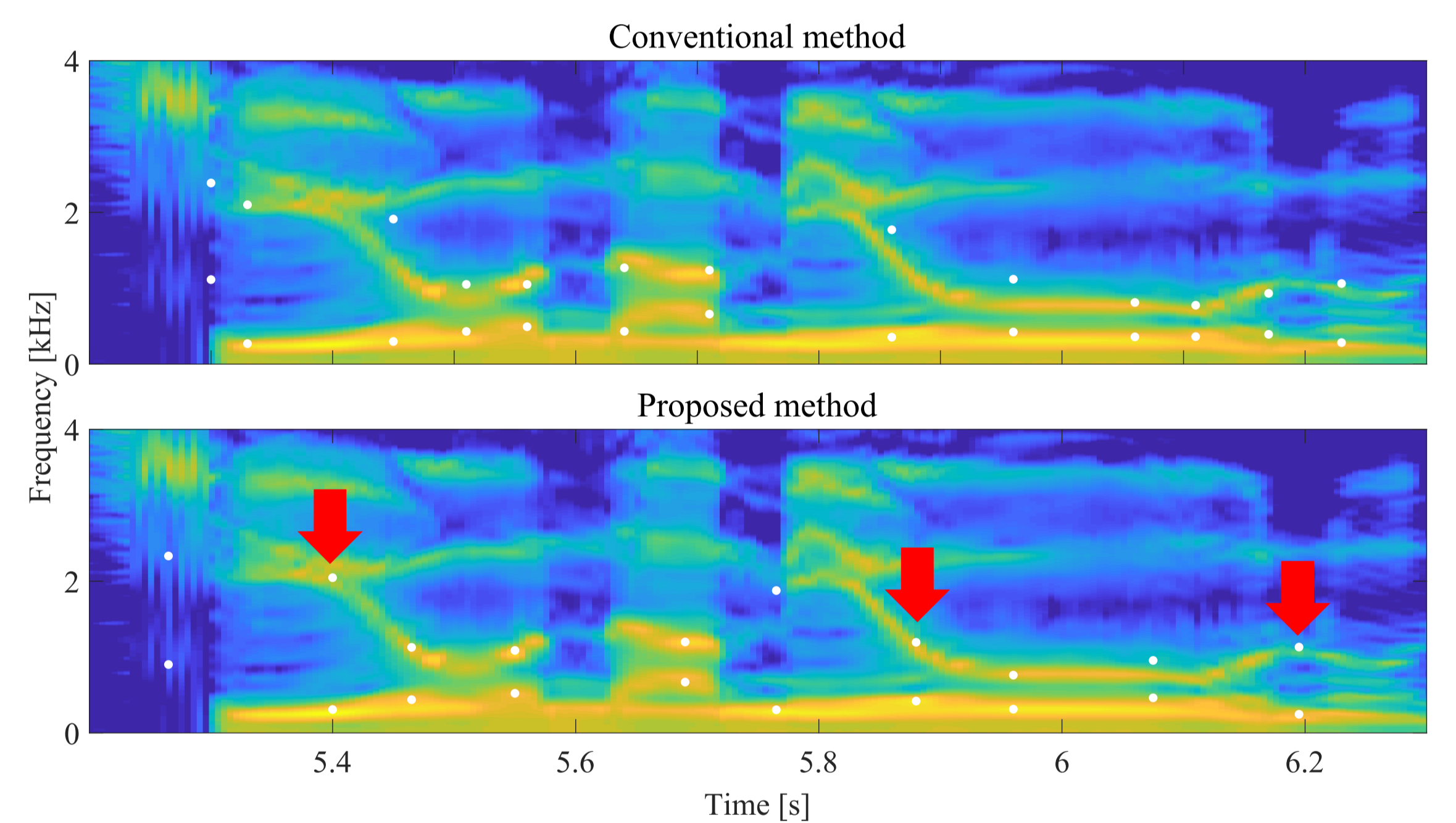


時間軸方向の対応付けについて, 音素境界に依存しない新たな手法を検討する。

## 提案手法 : 動的時間伸縮による時間軸方向の対応付け × パワーによる対応点の選定



有声区間における対応点の選定

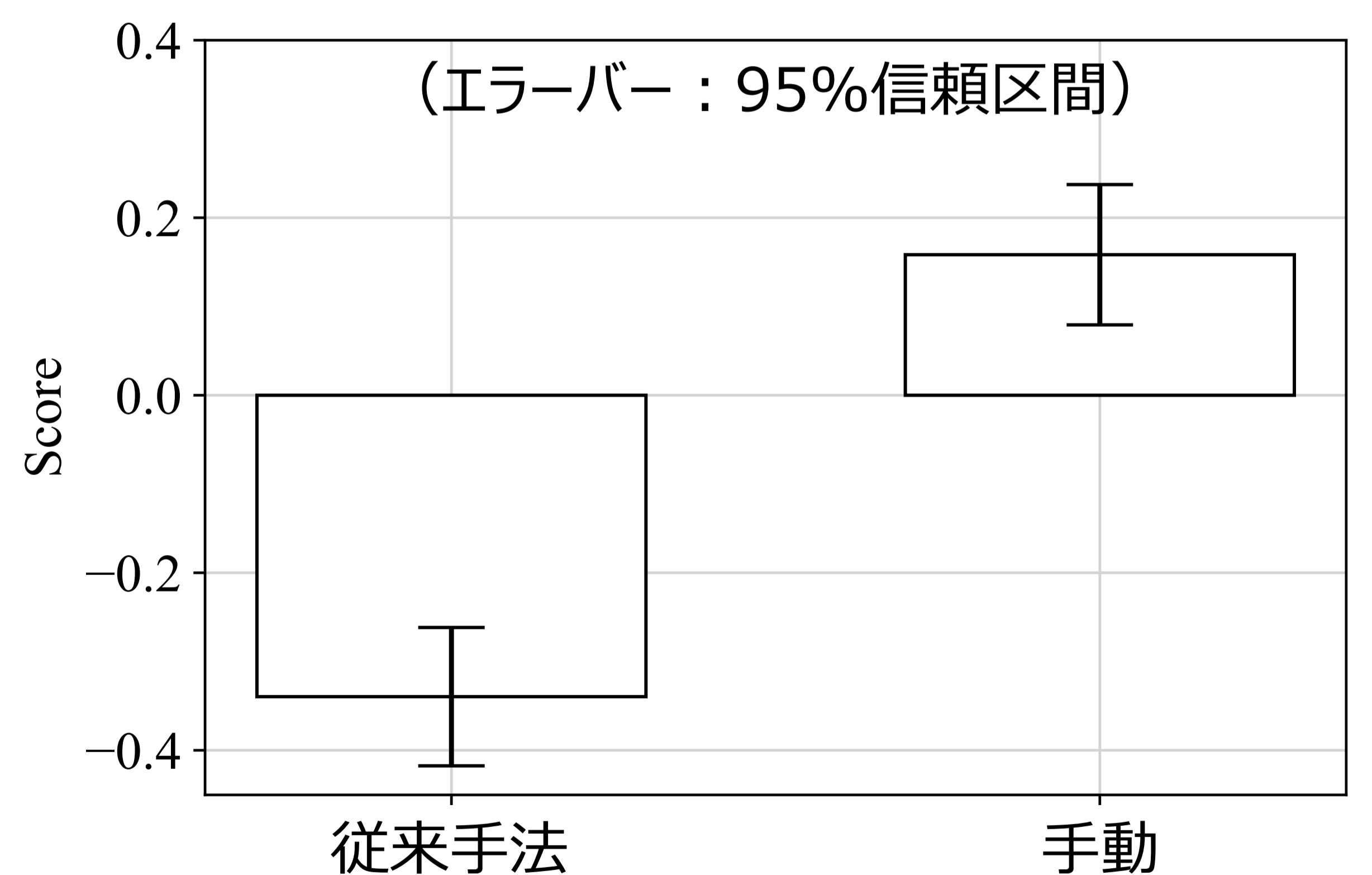


それぞれの手法による対応点をプロットしたスペクトログラム

## 実験 :

評価手法	CMOS (自然性)
実験環境	防音室
被験者数	正常な聴覚を有する20名
音声サンプル	JVS/parallel100コーパス
全試行数	48試行
モーフィング率	50%

- 自然性 : 手作業 > 提案手法 > 従来手法
- **提案手法と従来手法**に有意差あり → 提案手法は時間軸方向の自動対応付け手法として有効である.
- 従来手法 : 音素系列情報が制約となっていた.  
→ 提案手法はこの制約はなく, **同じ発話内容が保証された2つの音声のみ**があれば良い.



## まとめ :

- 動的時間伸縮を用いた時間軸方向の自動対応付け手法の提案
- 今後の展望 : 周波数軸方向の対応付け

