

# i-vectorに基づく賑わい音の推定方式の検討

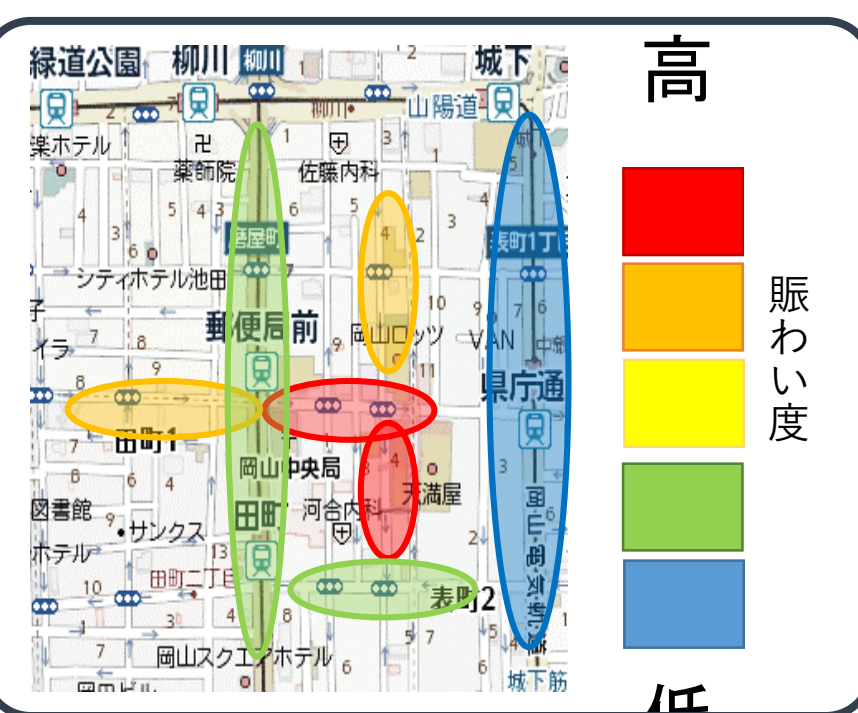
呉 セン陽(岡山大学 阿部研究室)

## 1. 研究背景・目的

**賑わい** 商業施設や観光地の周辺などに多くの人が集まり、活気のあること

### 想定される用途

- ✓ イベントの盛り上がり度合いを数値化・評価
- ✓ 政策立案や店舗の出店場所の策定



- i-vectorに基づく環境音分類で高い識別性能が得られることを目指す
- 最良の識別性能の実験条件を見つける

## 2. 提案方式

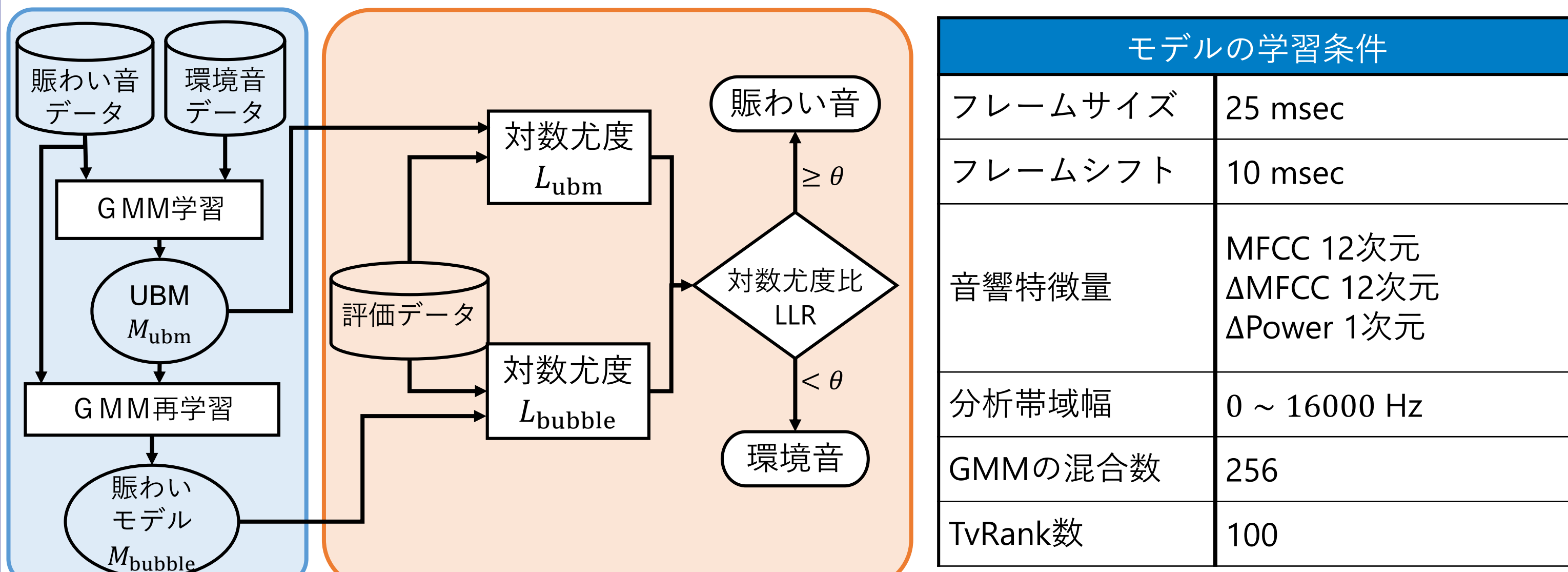
### GMM-UBMに基づく賑わい音識別

- ◆ 全ての学習データを用いて  $M_{ubm}$  と賑わいモデル  $M_{bubble}$  を作成
- ◆ 各モデルを用いて、評価データのLLRを計算

### i-vectorに基づく話者識別手法

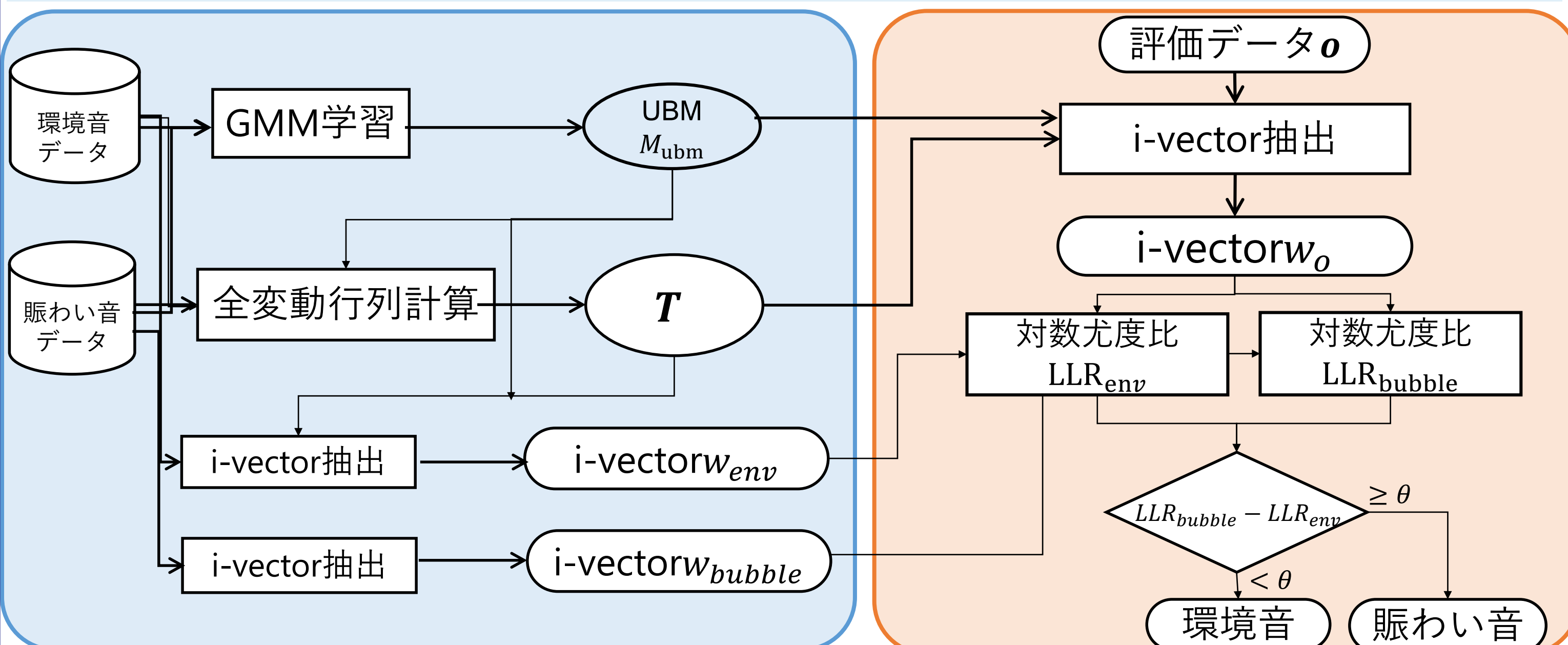
- ◆ 話者ごとに固有の特徴を抽出
- ◆ 得られた特徴を登録話者の特徴と比較

### GMM-UBMによる賑わい音を検出



$$LLR(o) = \log L(o; M_{bubble}) - \log L(o; M_{ubm})$$

### i-vectorによる賑わい音を検出



## 3. 実験条件

環境音データ		モデルの学習条件	
場所・行事	住宅街・駅の近辺・商店街	フレームサイズ	25 msec
データ数	7,499	フレームシフト	10 msec
賑わい音データ		音響特徴量	MFCC38次元, MFCC59次元(メルフィルタバンク24)
場所・行事	成人式・センター試験・入学式・花火大会	分析帯域幅	0 ~ 16000 Hz
データ数	263	GMMの混合数	128, 256, 512, 1024
評価データ		TvRank数	100, 400, 600
場所・行事	催し物のない休日		
データ数	1139		

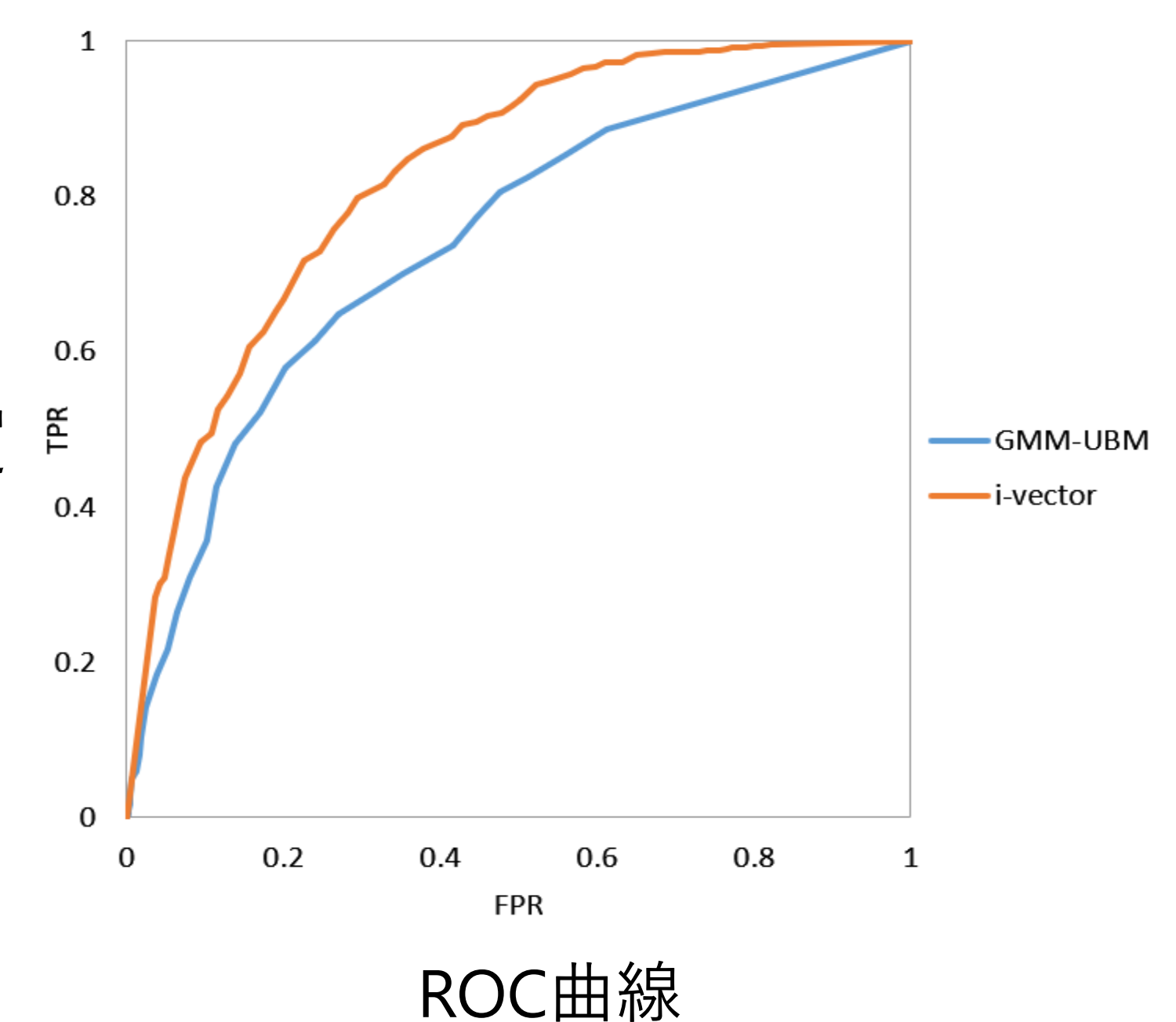
## 4. 性能評価

- ◆ F尺度：適合率と再現率の調和平均
- ◆ ROC曲線：識別器の特性を可視化

## 5. 実験結果

### 比較実験の結果

	F尺度
i-vector	0.74
GMM-UBM	0.64



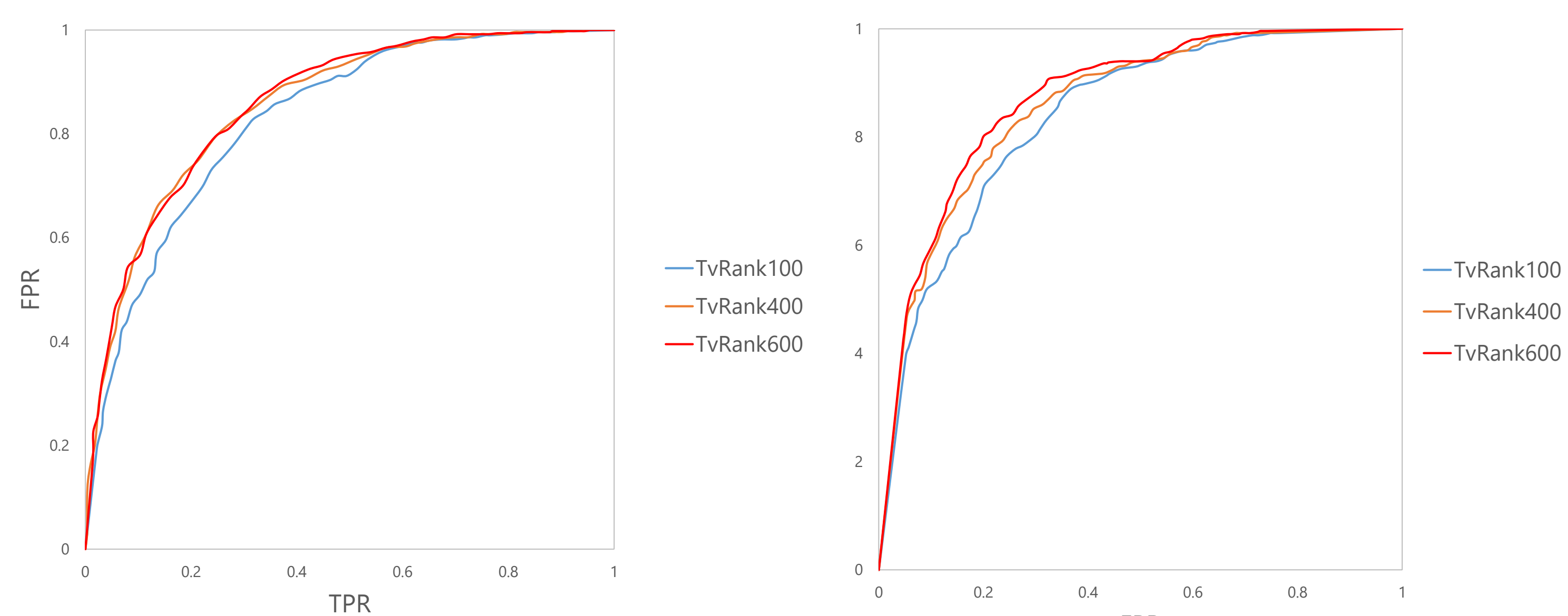
i-vectorに基づく賑わい音の推定方式とGMM-UBMに基づく賑わい音の推定方式の比較

- ◆ i-vectorはGMM-UBMよりもF尺度の最大値が高い
- ◆ 全ての範囲においてi-vectorがGMM-UBMより上側

i-vectorに基づく賑わい音識別器はGMM-UBMによる賑わい音識別器よりも優れた性能を発揮

### 提案方式によるパラメータを変化させた実験結果

MFCC38	TvRank100	TvRank400	TvRank600
128mix	0.7496	0.7632	0.7686
256mix	0.7515	0.7695	0.7833
512mix	0.7524	0.7711	0.7759
1024mix	0.7509	0.7684	0.7801
MFCC59	TvRank100	TvRank400	TvRank600
128mix	0.7631	0.7720	0.7899
256mix	0.7589	0.7772	0.7893
512mix	0.7640	0.7776	0.7863
1024mix	0.7601	0.7772	0.7851



MFCC59次元, 混合数128mix, TvRank600が最良

## 6. まとめ・今後の課題

### まとめ

- ◆ 二つ方式の比較実験を行う
  - ✓ 提案方式の賑わい音の識別性能はより優れている
- ◆ 提案方式によるパラメータを変化させた実験を行う
  - ✓ 混合数を増加させても識別性能が変わらない
  - ✓ 最良の識別性能の実験条件を得られる

### 今後の課題

- ◆ 他の賑わい音の推定方式の検討
- ◆ 地図上に賑わい音を可視化するシステムの構築