

10

# 番組制作のための音声認識

## Speech Recognition for Smart Production

### 番組から取材映像までさまざまな音声を認識

生放送の情報番組への字幕付与や、取材映像中のコメントを書き起こすための音声認識の研究を進めています。明瞭に発声されない音声が多い情報番組をリアルタイムに認識する技術と、取材映像の書き起こしを効率良く制作する技術を展示しています。

#### 情報番組をリアルタイムに認識

**音声認識** 雑音環境下や不明瞭な発話でも精度よくリアルタイムに認識



情報番組

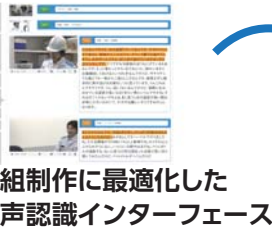


字幕付与



#### 迅速な取材映像の書き起こし制作

**音声認識** 取材映像中のコメントを精度よく認識



番組制作に利用



番組



#### 情報番組への字幕付与と取材映像の書き起こし制作

#### ●情報番組をリアルタイムに認識

多様な話題で構成され、それぞれの話題を解説するゲスト等が出演する情報番組では、解説する音声に背景雑音が含まれることや、明瞭度の低い音声が多いことが音声認識の課題でした。深層学習モデル\*等により、このような情報番組の音声でもリアルタイムに精度よく認識して字幕を付与します。

#### ●迅速な取材映像の書き起こし制作

取材した映像素材からニュース番組等を制作するためには、取材映像に含まれるコメントを素早く書き起こすことが不可欠です。音声認識技術を用いて、取材映像の書き起こしを迅速に効率良く制作するシステムを開発しています。

#### 今後の予定

取材環境や話者の発話状況によらず、精度よく認識できる音声認識技術を目指すとともに、番組制作者が使いやすい音声認識ソフトウェアのインターフェースの開発を進めます。

\* 深層学習モデル: 脳の神経回路網を多層構造でモデル化して学習する手法