

# 顔から声を予測する 深層学習に基づく音声合成

名古屋工業大学 工学専攻  
情報工学系プログラム  
准教授 橋本 佳

本研究の特徴を一言で言うと、、、

# 顔画像からその人の声を再現することができる音声合成技術

を、研究・開発している。

# 社会背景と技術的課題

- **音声合成技術の普及**

- 音声合成：テキストを音声へと変換する技術
- 音声案内システム、音声対話システム、スクリーンリーダー、など

- **特定の人で話す音声合成システムの要望**

- 有名人の声
- 自分の声

- **音声合成システムを作るためには大量の音声データが必要**

- 収録コストが大きい
- そもそも音声データが残っていない場合も

# 従来技術

## ● 特定話者音声合成システム

- ある特定の人物の音声データのみから音声合成システムを構築
- その人物の声質を再現
- その人物の大量の音声データが必要

## ● 多数話者音声合成システム

- 多数の人物の音声データから音声合成システムを構築
- 目標の人物の声質を真似ることが可能（話者適応技術）
- 目標の人物の音声データは少量で良い

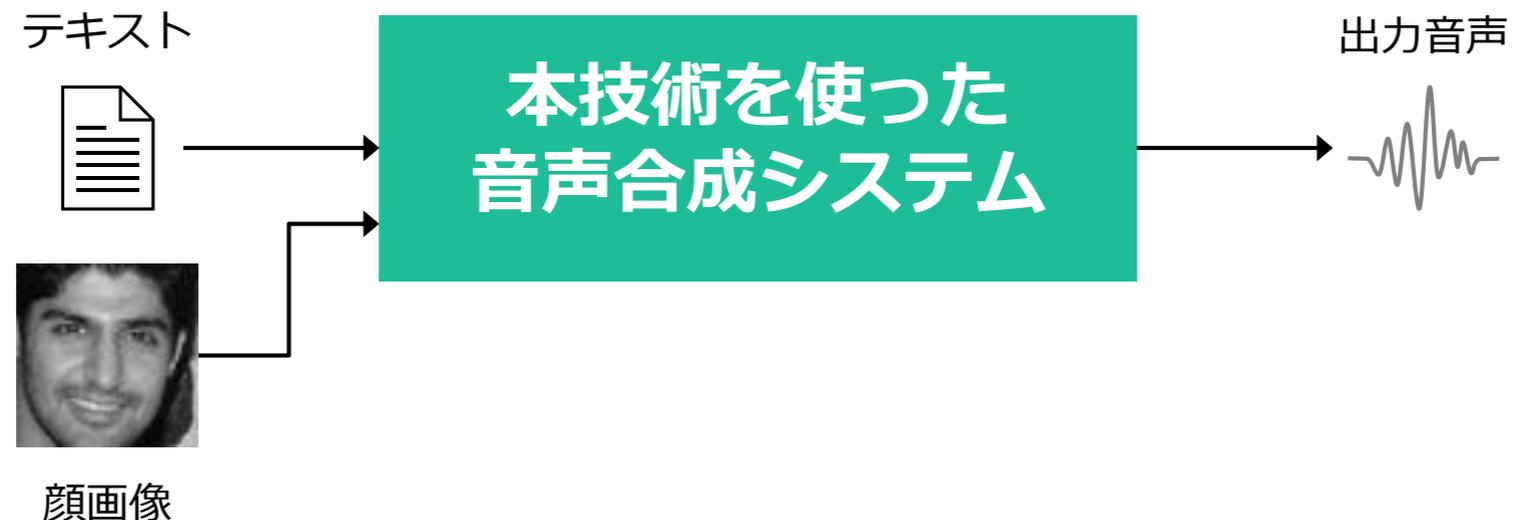
# 本技術の特徴

- 画像認識技術と音声合成技術で顔と声の関係をモデル化



# 本技術の特徴

- **画像認識技術と音声合成技術で顔と声の関係をモデル化**
  - 多数の人物の顔画像データと音声データを利用
  - 顔画像を入力することで声質を予測して音声を合成
  - 顔画像データがあれば目標の人物の音声データは不要
    - 特定の人物の音声合成システムの構築コストを大きく削減
    - 音声データがない場合でも音声合成システムを構築可能

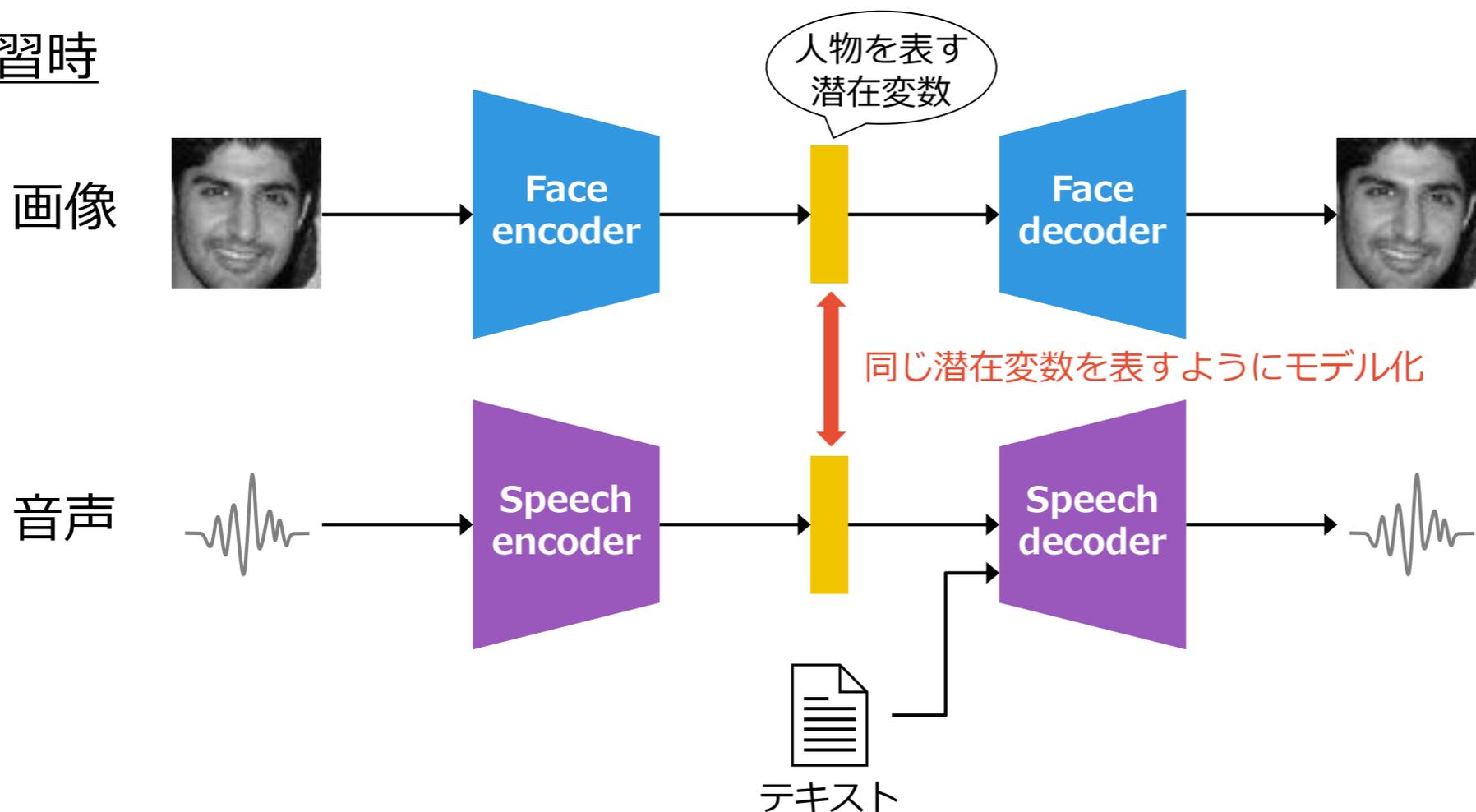


# 具体的な取り組み

## ● 深層学習を用いたモデル化

– Variational Autoencoder (VAE) を利用

学習時

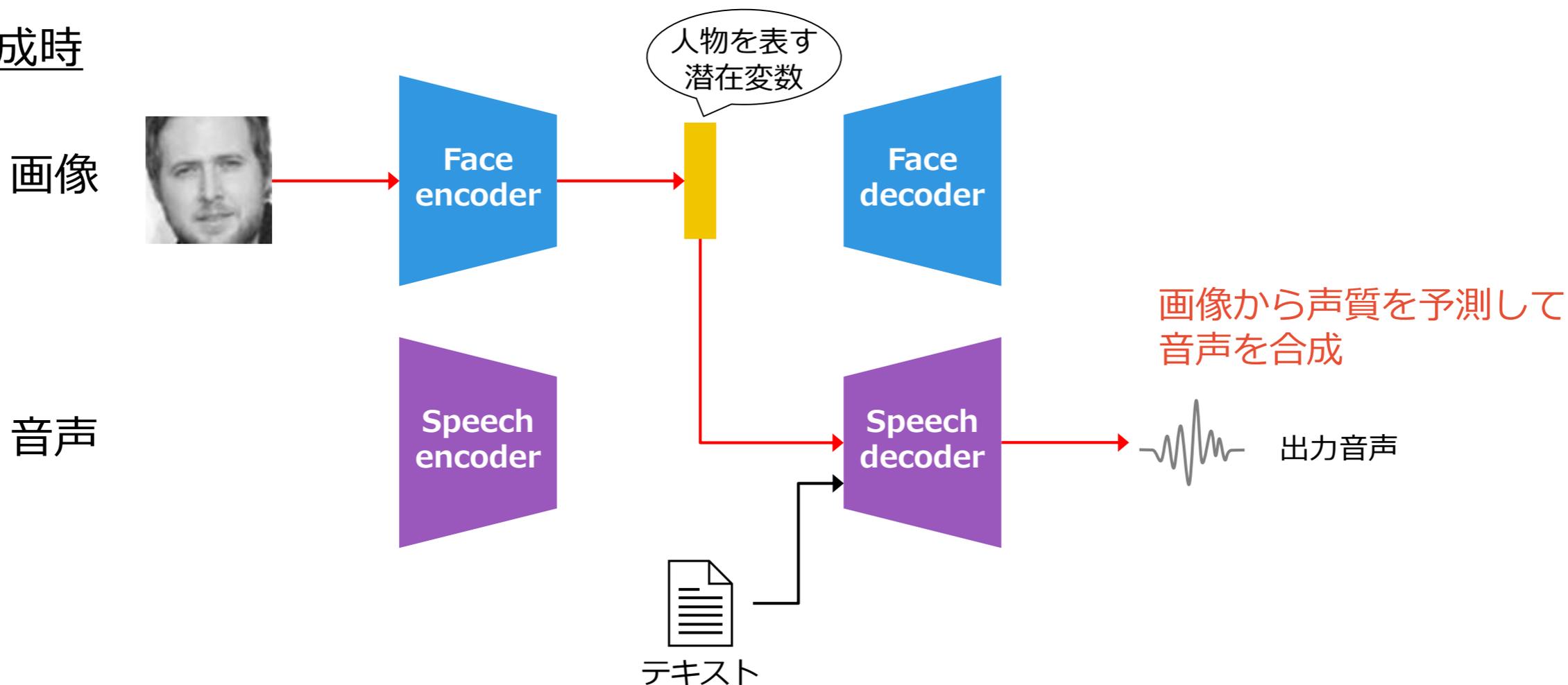


# 具体的な取り組み

- 深層学習を用いたモデル化

- Variational Autoencoder (VAE) を利用

合成時



# 求める提携先・メッセージ

- **顔から声を予測する音声合成の研究をしています**
  - 顔から声質だけでなく表情から声質などの応用も考えられます
  - 顔画像と音声以外にも○○と音声の組み合わせも考えられます
  - 関連技術の開発に興味があればご相談ください
  
- **ほかにも様々な音声情報処理に関連する研究をしています**
  - 音声合成、歌声合成、感情音声合成、声質変換など
  - 音声情報処理の研究に興味があればお問い合わせください

# 本技術に関する情報

## 試作品の状況

**未定**

※提供の際は諸手続が必要となるため、下記問合せ先までご連絡願います。

## 研究フェーズ



# 【お問合せ】

名古屋工業大学 産学官金連携機構

〒466-8555 名古屋市昭和区御器所町字木市29番

TEL:052-735-5627

E-mail: [nitfair@adm.nitech.ac.jp](mailto:nitfair@adm.nitech.ac.jp)

URL: <https://technofair.web.nitech.ac.jp/>