

# 音す：AR 公共空間に向けたリアルタイムで言葉を空間に落とすシステム

Wang Hanlin\* 王 雲傑\* 姜 鵬\* 村上 健樹\*  
王 楚\* 劉 宇馨\*

**概要.** 人と公共空間がより深くインタラクションする取り組みとして、リアルタイムで話した言葉を AR 空間内に落とすシステム、「音す」を提案する。このシステムは、音声認識を用いて人々が話した言葉をテキスト化し、音声の特性に基づきユニークな文字に変換し、AR 空間に投影する。この取り組みを具体化するため、学内の学生製作展示にてパイロットスタディを実施した。展示では、来場者が「音」として発話した言葉を AR 空間に「落とす」という体験を通じて、AR 空間との相互作用の可能性を探求し、言葉が溢れている AR 公共空間を作ることを実現した。得られたフィードバックから、参加者は音、文字、空間とのインタラクションや公共空間との相互作用を体感し、多岐にわたる応用が考えられるとの意見が寄せられた。今後の課題として実装に関する問題点を解決する。公共空間と AR 技術の結びつきをさらに深化させるための一つの試みとして、本システムは人々、空間、そしてコミュニティとの関わり方の可能性を生み出す。

## 1 はじめに

人々は日常的に公共空間を通り抜けるが、その空間を編集や参加する意識は薄い。AR の進化により、我々の身の回りの空間とのインタラクションの可能性が増えている。実際、美術館などの公共空間での AR 活用事例が増加している。とはいえ、多くの場合で人々は AR 空間をまだ通過点として扱っており、空間への参加は促進されていない。この課題を解決するために、リアルタイムで話した言葉を AR 空間内に落とすシステム、「音す」を提案する。このシステムは AR と音声入力技術を組み合わせ、話された言葉を音声の特性に基づいてフォント、サイズ、カラーが決められ、ユニークな文字に変換し、AR 空間に投影する。このように、音声情報を視覚情報に分解・再構成することで、人々が話すときの感情の擬似的再現をし、AR 空間に残す。この累積的な AR 公共空間を通じて、人々の空間への関与や空間内の要素とのインタラクションを試み、探索することができる。

## 2 関連研究

### 2.1 AR 空間で音声認識を使う研究

AI 技術の発展と AR グラスの普及とともに、音声認識と AR 技術を統合し、AR 空間にリアルタイム字幕を表示するシステム [3] と翻訳するシステム [5] が開発された。このようなシステムは音声情報を理解することで会話を補助するために使われ、文字情報が話によって更新される。「音す」は、音声情



図 1. ユーザが体験する「音す」の操作画面  
学生展示の AR 空間に言葉が蓄積していく様子

報を理解するより、空間に残すことを目的とし、話した場所に文字を固定するシステムを開発した。音声と空間の関係を作り、言葉が累積する AR 空間を現実空間上に作り出す。

### 2.2 展示や公共空間に AR の応用

AR 技術を応用して展示・公共空間で演出する事例が近年増えている。AR アート作品の一部になる展示 [6]、空間の誘導・説明を機能する応用 [4]、AR オノマトペで集客する [1] 事例がある。しかし、このような応用事例では来客が AR 空間を見ることだけができ、AR 空間を編集する機能は含まれていない。また、AR をコメントツールとして使われる事例は筆者の知見にはない。「音す」は、来客が話した言葉を AR 空間に残すことで更新される AR 公共空間を作る。また、コメントを AR 空間に累積することにより、音と文字を通して人と公共空間の相互作用を作る。

Copyright is held by the author(s). This paper is non-refereed and non-archival. Hence it may later appear in any journals, conferences, symposia, etc.

\* 東京大学

### 3 システム構成

#### 3.1 アプリの全体像と主な機能

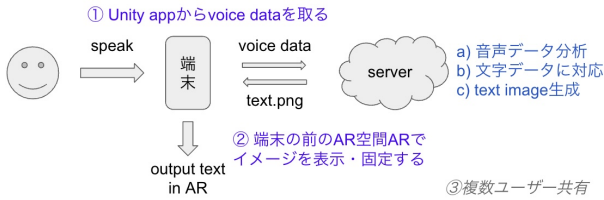


図 2. システムアーキテクチャ

「音す」は、音声認識された言葉をリアルタイムで AR 空間に置くシステムである (図 2)。文字の視覚表現は音声の特徴から決められ、生成された文字はユーザーが話した場所に固定される。実装言語は C# を用い、AR 空間は Unity を用いて作られている。音声認識のために Voice Recognition の API を使って話した内容をテキスト化する。また、周波数分析のために Python Librosa の API を用いて音声データから声の特徴を取得する。認識された言葉と特徴に基づいて、フォント、サイズ、カラーが単語に割り当てられ、ユニークなテキスト画像が生成される。生成された言葉の画像はデバイスに送られ、話した場所の前 0.5m の AR 空間にデバイスの向きに置かれる。

ユーザ体験は以下となる。ユーザが UI のボタンを押すと、録音が始まり、AR 空間内のデバイスの座標が記録される。ボタンを離れた瞬間に録音は停止し、録音されたデータはサーバーに送られる。録音が終わった 3 秒後、サーバーから言葉の画像をもらい、デバイス内の AR 空間の記録された場所に落とされる。言葉はデバイスが移動しても空中に固定されるため、立体感のある演出ができる。

#### 3.2 音声情報の視覚提示

音声入力を使用する際に、話す内容が文字に変換され記録されるが、話のスピード、大きさ、高さなどの「感情」を構成する情報がこのプロセスの中で見落とされる。自然言語処理または韻律と組み合わせで感情を検出し、フォント、カラーなどの視覚情報にかわってユーザーに提示する手法がある [2]。私たちは、以下のような簡単な対応関係で、話すときの音声情報を分解し、視覚情報に提示する。

音の大きさ → 文字のサイズ

音の速さ → 文字のフォント

音の高さ → 文字のカラー

このプロセスでは感情分析が行われず、音声情報の生データを視覚情報に投影することで、生成されたテキスト画像から感情を擬似的に再現する。この手法では、再現された感情に人工処理の影響が少なく、直感的であると期待される。

### 4 展示におけるパイロットスタディ

東京大学制作展で、本システムを用いてパイロットスタディを実践した。パイロットスタディの主旨は、会場内の作品の周りに言葉を落とすことで、来場者、作品、会場とのインタラクションを観察することである。展示は、2023 年 7 月 7 日から 10 日、東京大学本郷キャンパス情報学環オープンスタジオで行われ、本プロジェクトを含め、18 個の作品が展示された。展示で、本システムがインストールされた iPad と iPhone を準備し、iPad 内のアプリの表面をプロジェクターで投影するセットアップをした (図 1)。音声データの特徴分布が事前把握できないため、展示で使われているプロットタイプは音声の特徴量による変化をまだ導入せず、テキストのサイズ、フォント、カラーはランダム生成とした。展示の来場者に端末を渡し、会場内を周りながら自由に発言し、言葉を空間に「落とす」という体験を提供した。5 日間の展示で、約 1400 回の体験を提供した。得られたフィードバックは以下のようであった：

- ・自分が言った言葉が会場空間に残ることに楽しさを感じる
- ・会場と作品の新しい相互作用が観察される
- ・音と空間のかかりができてよかった
- ・周りの作品の作者たちに好かれ、自分の作品にたくさん使われたい
- ・商品の周りにコメントを集めたい。アンケートより柔軟な形なので、顧客にとって話しやすくなる
- ・自分と自分の友達しか見えなコミュニティのような使い方もしたい
- ・ライブやミーティングで AR 技術を使うと、体験が豊かになり、参加者のモチベーションが上がると期待できる。

### 5 議論

今度の実装における技術的制約として、AR 空間の表示には約 3 秒の遅延と座標システムの精度に問題がある。今後の研究と展示の展望としては、多デバイスの座標共有システムの導入、文字の表現が音声の特徴量によって変わる精度の向上、多言語対応の実現、言葉に近づいたら音声の演出、複数の重ねる情報空間の設計を目指して進めている。

### 6 おわりに

本研究で提案された「音す」というシステムを通して、空間を通過する人々の感想や感情の共有や共感が促され、エンゲージングなコミュニティを生み出すと期待される。また、観客参加型の AR 展示形式を提案し、展示やイベントでの観客の参加とエンゲージメントを高め、言葉が溢れている空間を作り出す可能性がある。

## 参考文献

- [1] ARGO. コクヨが「THE CAMPUS SHOP」で AR の実証実験を開始！「ゴゴゴ…」「ぷにぷに」など擬音が飛び出す売り場も. <https://argo.jp/media/news/ar-kokuyo-thecampusshop>, 2021.
- [2] DNP. 感情表現フォントシステム. <https://shueitai.dnp.co.jp/fontsystem/>, 2018.
- [3] I. Suzuki, K. Yamamoto, A. Shitara, R. Hyakuta, R. Iijima, and Y. Ochiai. See-Through Captions in a Museum Guided Tour: Exploring Museum Guided Tour for Deaf and Hard-of-Hearing People with Real-Time Captioning on Transparent Display. In K. Miesenberger, G. Kouroupetroglou, K. Mavrou, R. Manduchi, M. Covarrubias Rodriguez, and P. Penáz eds., *Computers Helping People with Special Needs*, pp. 542–552, Cham, 2022. Springer International Publishing.
- [4] 柴崎 真里, 渡邊 慎二. AR を用いた新しい本との出会い方の研究. 日本デザイン学会研究発表大会概要集, 69:442, 2022.
- [5] 西田 直人, 堀部 咲歩, 暦本純一. Traverco AR: AR グラス上の字幕翻訳を用いた会話補助システム. p. Poster. WISS, 2000.
- [6] 蔡国強. Exhibition Highlight: “Ramble in the Cosmos—From ‘Primeval Fireball’ Onward.