

# 対話型美術鑑賞音声対話システムのオンライン上での利用の提案

## Proposal of Online Use of a Spoken Dialogue System for Interactive Art Appreciation

湯浅 美野<sup>†</sup> 吉野 孝<sup>††</sup> 青木 加苗<sup>†††</sup>  
Yoshino Yuasa Takashi Yoshino Kanae Aoki

### 1. はじめに

美術館において鑑賞者個人が行う鑑賞は、解説等の作品情報とともに作品を鑑賞するものが一般的である。しかし、この鑑賞法は、文字で記された作品情報を読み取ることが主となりがちで、作品を鑑賞者自身で解釈できないという問題点がある。一方で、近年、学校教育、美術館教育のみならず、社会人の自己研鑽において、「対話型」と呼ばれる鑑賞法が注目されている [1]。これは、参加者が鑑賞対象についてファシリテーター<sup>1</sup>や他の鑑賞者とやり取りを行い、興味や理解を深め、作品に関する意味や価値を鑑賞者の中で作り出していく、鑑賞の方法である [2]。

多くの美術館では、ギャラリートークやワークショップといった形で、学芸員やボランティアがファシリテーターの役割を担い、対話型鑑賞を行っている。しかし、それらに参加せず、鑑賞者が対話型鑑賞を行うことは困難であると考えられる。鑑賞経験の少ない鑑賞者が対話型鑑賞を行うためには、的確なファシリテーションを行うファシリテーターと、意見を交換し合うことのできる他の鑑賞者が必要なためである。

現在、新型コロナウイルスの世界的蔓延により、感染防止の観点から多くの美術館が休館している [3]。そこで注目されているのが、オンラインでの美術鑑賞である。多くの美術館が、実際に美術館を訪れずとも作品を鑑賞出来るように、オンライン上で作品画像の公開や動画での作品紹介を行った。それによって鑑賞者は、美術館以外の場所でも作品の鑑賞を行うことが出来るようになった。しかし、上述の通り、鑑賞経験の少ない鑑賞者が対話型鑑賞を行うためには、的確なファシリテーションを行うファシリテーターと、意見を交換し合うことのできる他の鑑賞者が必要なため、オンライン上での対話型鑑賞はあまり行われていない。そこで我々は、オンライン上で美術鑑賞における対話型鑑賞を支援する音声対話システム KANAI<sup>2</sup>を開発した。KANAIはファシリテーターの役割をシステムが担い、複数人のユーザが利用することで、ユーザが一人でもオンライン上で対話型鑑賞を行うことを可能にする。本稿では、オンラインでの対話型鑑賞を実現するため開発したシステムの概要について述べる。

### 2. 対話型鑑賞

対話型鑑賞は、1980年代からニューヨーク近代美術館で始まったVTC (Visual Thinking Curriculum) やその発展系として開発されたVTS (Visual Thinking Strategies) がもとになっている [4]。VTSの特徴は、鑑賞中に作品名や作者名、解説文など背景となる情報を第一に用いない点である [1]。作品に関する情報を用いないことは、作品についての知識を否定することではない。作品に関する情報を用いない理由は、幼児期や小学生、また美的経験が少ない青少年や大人達の最初のかかわり方として、最初に知識がなくても鑑賞が可能であることを保障するものである。ファシリテーターが鑑賞者を導き、作品について他の鑑賞者と意見を交換することで、作品に関する意味や価値を鑑賞者の中で見つけ出していく鑑賞法である。

### 3. 関連研究

#### 3.1 作品情報の提供を支援するシステム

Toyamaらは、ヘッドマウント式モバイルアイトラッカーを用いて、Museum Guide 2.0を開発した [5]。これは、ユーザが特定の作品を見ていることを検出すると、イヤホンを通じてその作品に関する音声ガイドを自動で提供し、それによってユーザは作品知識の獲得をしながら、作品を鑑賞することができるガイドアプリケーションである。Banerjeeらは、美術館内のマルチディスプレイとスマートウォッチを連携させた作品情報提示システムを開発した [6]。これらのシステムは、美術館の作品を鑑賞する際に利用するという点が本システムと類似しているが、本システムは、作品に関する情報の理解を重視した鑑賞方法ではなく、対話によって作品に対するそれぞれの考えを深める対話型鑑賞を、システムによって実現することを目的としている。

#### 3.2 能動的鑑賞を支援するシステム

Aliらは、アートコードを用いた音声ガイドの提供と、アートコードを用いた鑑賞者の作品に対する意見の共有を提案した [7]。アートコードとは、ユーザが作り出した画像やパターンを、機械可読マーカとして使用できる画像認識技術である。Saloらは、展示物に対するサウンドスケープの作成により、鑑賞者に能動的鑑賞を促すシステムを開発した [8]。サウンドスケープとは、作曲家M. シューパーが提唱する概念で、「音の風景」を意味する造語である。これらのシステムは、ユーザに能動的鑑賞を促すという点が本システムと類似しているが、本システムは、対話型鑑賞をシステムによって実現することを目的としている。

#### 3.3 対話システム

Yoshinoらは、ユーザの発話の焦点を発見し、発見されたユーザの興味に応じた情報を提示する対話システムを開

<sup>†</sup> 和歌山大学大学院システム工学研究科, Graduate School of Systems Engineering, Wakayama University

<sup>††</sup> 和歌山大学システム工学部, Faculty of Systems Engineering, Wakayama University

<sup>†††</sup> 和歌山県立近代美術館, The Museum of Modern Art, Wakayama

<sup>1</sup> 対話型鑑賞における司会進行役

<sup>2</sup> ユーザに対して、「~かな (“KANA”)」と考えさせるという点と、ユーザからみて本システムは、「AI」との対話を感じさせるものであるという点から「KANAI」と名付けた。

発した [9]. Bessho らは、リアルタイムのクラウドソーシングとウェブから自動的に収集される大規模なデータベースにもとづく対話システムを提案した [10]. これらのシステムは、音声対話を行っているという点で本システムと類似しているが、本システムは、絵画鑑賞においてユーザーの考えを深め、作品に関する意味や価値をユーザーの中で深めるための、適切な質問を形成する方法について考える。

#### 4. 事前調査

システムを作成するにあたり、特定の作品に対してファシリテーションを行うことができるシステムの構築を目指した。今回は石垣栄太郎の《人民戦線の人々》<sup>3</sup>を題材とした。まず、この絵画を題材として、遠隔会議システム ZOOM を用いて、オンライン上で共同研究者である和歌山県立近代美術館の学芸員に対話型鑑賞を 3 度行ってもらった。その際の鑑賞者は以下の通りである。

- (1) 鑑賞者：著者
- (2) 鑑賞者：和歌山大学の女子学生
- (3) 鑑賞者：和歌山大学の男子学生 2 人と女子学生 1 人

これらの対話内容をもとに、システムを作成した。特に注目し、システムに取り入れたファシリテーション方法とファシリテーターの質問内容を以下に示す。

- ファシリテーション方法
  - (1) 鑑賞者が示した解釈を肯定した後に、鑑賞者の返答内容をより具体的なものに言い換えることで、鑑賞者に自身の解釈の根拠を明確に理解させる方法
  - (2) 人数を数える動作によって、鑑賞者に作品に描かれた全ての人物に対して自然と注意を向けさせる方法
  - (3) 作品の歴史的背景と他の作品とのつながりについて最後に述べ、鑑賞者の作品に対する理解をより深める方法
- ファシリテーターの質問内容
  - (1) 時代についての質問
  - (2) 場所についての質問
  - (3) 構図についての質問
  - (4) 人物の持ち物についての質問
  - (5) 人物の表情についての質問
  - (6) 人物の立場についての質問
  - (7) 作者の立場についての質問

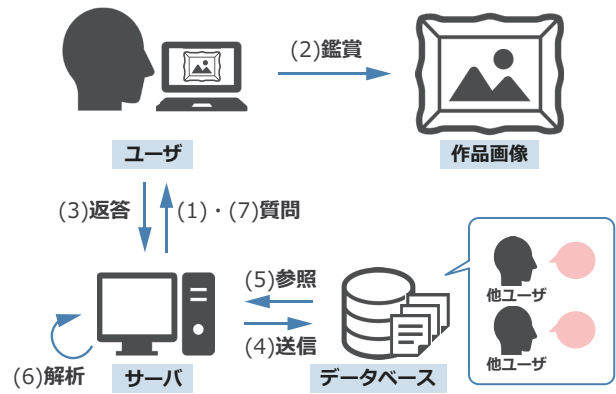


図 1: システム構成図

#### 5. KANAI2

##### 5.1 概要

KANAI2 は、オンライン上で利用可能な、美術鑑賞における対話型鑑賞を支援する音声対話システムである。対話型鑑賞におけるファシリテーターの役割をシステムが担い、ユーザーが一人でも対話型鑑賞を行うことを可能にする。開発は、JUMAN<sup>4</sup>と KNP<sup>5</sup>と Web Speech API<sup>6</sup>を用いて行った。本システムは、作品に対する鑑賞者の考えを深めるために適切な質問をすることで、作品についての新たな解釈を生み出し、さらに、他の鑑賞者とその解釈を共有することで、作品との新たな関係をユーザーと作品の間に作り出すことを目指す。

##### 5.2 システム構成

図 1 にシステムの構成図を示す。本システム使用時の流れについて説明する。まず、図 1(1)では、本システムからユーザーに対して最初の質問を行う。図 1(2)では、ユーザーが作品を鑑賞する。図 1(3)では、その質問に対して、ユーザーが返答を行う。図 1(4)では、その返答をデータベースに送信し、蓄積する。図 1(5)では、データベース内の情報を参照し、図 1(6)では、返答の解析結果と参照した情報をもとに、次の質問を作成する。そして、図 1(7)で再び、本システムからユーザーに対して質問を行う。本システムでは、この一連の流れを繰り返すことで、ユーザーと対話を行う。

##### 5.3 システム機能

本節では、本システムの音声認識機能、質問内容作成機能、音声合成機能、対話内容表示機能について述べる。本システムのシステム画面を図 2 に示す。

###### 5.3.1 音声認識機能

本機能は、図 2(1)の録音ボタンを押すと開始し、もう一度録音ボタンを押すと音声認識を終了する。音声認識中は録音ボタンの色が赤色に変化する。

<sup>3</sup> 石垣栄太郎《人民戦線の人々》1936 - 37 (昭和 11 - 12) 頃、油彩、キャンバス 148.5 × 207.0 cm 和歌山県立近代美術館蔵

<sup>4</sup> <http://nlp.ist.i.kyoto-u.ac.jp/index.php?JUMAN>

<sup>5</sup> <http://nlp.ist.i.kyoto-u.ac.jp/index.php?KNP>

<sup>6</sup> [https://developer.mozilla.org/ja/docs/Web/API/Web\\_Speech\\_API](https://developer.mozilla.org/ja/docs/Web/API/Web_Speech_API)

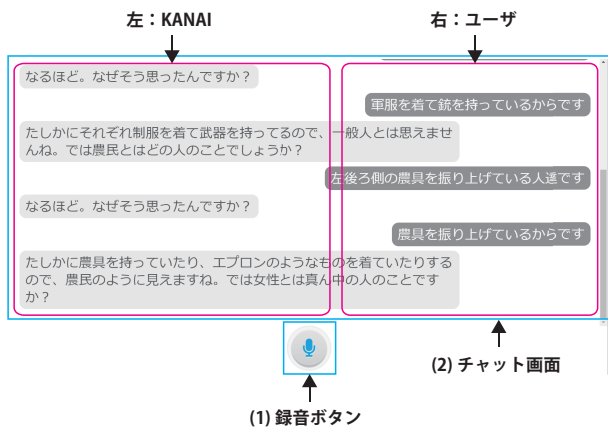


図 2: システム画面

### 5.3.2 質問内容作成機能

我々は先行研究として、美術館での対話型鑑賞を支援する音声対話システム KANAI を開発した [11]。KANAI を用いて、人との対話型鑑賞を行ったグループと、システムを用いて合成音声との対話型鑑賞を行ったグループの比較実験を行った。共同研究者である和歌山県立近代美術館の学芸員に、比較実験のアンケート結果を確認してもらった。その際に学芸員から指摘された、ファシリテーションにおける不足点が以下の 4 点である。

- (1) 絵画内における、描かれたそれぞれの事物が持つ役割に関する質問
- (2) 絵画内の事物間の関係性に関する質問
- (3) 絵画全体に関する質問
- (4) 絵画の解釈を容易にするための情報の整理と誘導

比較実験で用いた小泉による論文 [1] にもとづいた質問内容に、これらの内容と事前調査で得られた知見を追加する形で、KANAI2 の質問内容を作成した。

本機能は以下の機能を組み合わせたものである。

#### (1) 肯定

本機能は、システムからの質問に対してユーザーが自分の解釈を示した際、次の質問に移る前にユーザーが示した解釈を肯定する機能である。本システムは、ユーザーの解釈を事実として認め、それによって KANAI2 のデータを蓄積する。この機能は、ユーザーにとっても発言への抵抗感を減らすことにつながると考えられる。

#### (2) 注目事物の特定

本機能は、ユーザーが注目した事物の特定を行う機能である。特定は以下の流れにより行う。

- (a) ユーザーの返答を分析  
ユーザーの返答を分析し、返答に含まれる単語を抽出する。
- (b) ユーザーが示した絵画内の事物の推定  
抽出した単語から、ユーザーの示す事物を推定する。

#### (c) ユーザーが示す絵画内の事物の特定

ユーザーの返答によって、ユーザーが示す絵画内の事物を特定する。

#### (3) 返答文の内容を次の質問文に反映

本機能は、ユーザーの返答内容を解析し、次の質問をそれに応じた内容にする機能である。ユーザーの返答内容を JUMAN と KNP によって解析し、抽出した返答の一部を次の質問内容に加える。

#### (4) ユーザーの返答の言い換え・具体化

本機能は、ユーザーの返答内容をより具体的なものに言い換えることで、ユーザーに自身の解釈の根拠を明確化させる機能である。

#### (5) 作品情報の提示

本機能は、ユーザーがシステムとの対話により十分に自分なりの解釈を深めた後に、作品情報の提示を行うものである。具体的には作品のテーマ、実際の作品の大きさ、描かれた年代、国、同じテーマで描かれた他の作品について述べる。一方的に情報を提示するのではなく、ユーザーとの対話の流れに沿って対話の中で行う。

### 5.3.3 音声合成機能

本機能は、音声認識機能によりユーザーの返答を認識後、質問内容作成機能により質問を作成し、それを音声合成によりユーザーに発話する機能である。発言は、質問内容作成機能によりユーザーの返答に応じた質問内容が作成された時に行われる。

### 5.3.4 対話内容表示機能

本機能は、対話内容を図 2(2) のチャット画面に表示する機能である。右側に音声認識された内容が表示され、左側にシステムの返答が表示される。音声認識されたユーザーの発話内容は、発話終了後に表示される。システムの返答は、音声認識機能によりユーザーの返答を認識後、質問内容作成機能により質問を作成した後に表示される。

## 6. おわりに

本研究では、オンライン上で利用可能な、美術鑑賞における対話型鑑賞を支援する音声対話システム KANAI2 を開発した。本システムを利用することで、ユーザーがオンライン上の作品データを用いて、一人でも対話型鑑賞を行うことを可能にする。本稿では、遠隔会議システム ZOOM を用いて行った事前調査に基づき開発した KANAI2 の概要について述べた。今後は、本システムを用いたオンライン上での評価実験を行い、有用性を検証する。

## 参考文献

- [1] 小泉卓：対話を介した鑑賞教育の方法—専門性と論理的思考を育成する鑑賞教育—, 桜花学園大学学芸学部研究紀要, Vol.10, pp.49–68(2018).
- [2] Housen, A., Yenawine, P.: VTS Basic Manual. Visual Understanding in Education (2000).

- [3] 日経新聞：博物館 13%が再開不可も ユネスコ、支援呼び掛け, <https://www.nikkei.com/article/DGXMZO59287960Z10C20A5CR8000/> (参照 2020 年 7 月 13 日)
- [4] 加藤悦子, Ai Wee Seow, 宇野慶, 柿崎博孝, 佐藤由紀, 高橋愛, 林卓行：美術作品を中心とした視覚媒体を活用した教育の研究—VTS 美術鑑賞教育を日本に適用した教育方法の形成—, 玉川大学学術研究所紀要, Vol.22, pp.37–55(2016).
- [5] Takumi Toyama, Thomas Kieninger, Faisal Shafait, Andreas Dengel: Museum Guide 2.0 – An Eye-Tracking based Personal Assistant for Museums and Exhibits, In Deutsches Forschungszentrum für Künstliche Intelligenz German Research Center for Artificial Intelligence, pp.1–10 (2011).
- [6] Amartya Banerjee, Rovik Robert, Michael S.Horn: Field-Guide: Smartwatches in a Multidisplay Museum Environment, CHI'18 Extended Abstracts, ACM, pp.1–6 (2018).
- [7] Susan Ali, Boriana Koleva, Ben Bedwell, Steve Benford: Deepening Visitor Engagement with Museum Exhibits through Hand-crafted Visual Markers, DIS 2018, ACM, pp.523–534 (2018).
- [8] Kari Salo, Merja Bouters, Tommi Mikkonen: User Generated Soundscapes Activating Museum Visitors, SAC 2017, ACM, pp.220–227 (2017).
- [9] Yoshino, K., Suzuki, Y., Nakamura, S.: Information navigation system with discovering user interests. Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue, Association for Computational Linguistics, pp.356–359 (2017).
- [10] Bessho, F., Harada, T., Kuniyoshi, Y.: Dialog System Using Real-Time Crowdsourcing and Twitter Large-Scale Corpus, Proceedings of the 13th Annual SIGdial Meeting on Discourse and Dialogue, Association for Computational Linguistics, pp.227–231 (2012).
- [11] 湯浅美野, 吉野孝, 青木加苗：対話型鑑賞による美的発達のための音声対話システムの提案, マルチメディア, 分散, 協調とモバイル (DICOMO) シンポジウム論文集, pp.1147–1153(2020).