

# 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

湯浅 美野<sup>1</sup> 吉野 孝<sup>1</sup> 青木 加苗<sup>2</sup>

概要：美術館において個人が行う鑑賞は、解説等の作品情報を読み取ることが主となり、作品を鑑賞者自身で解釈できないという問題点がある。それに対して、対話型鑑賞法とは、作品に対するそれぞれの考えを深めるため、ファシリテーターが適切な質問を鑑賞者にし、作品についての新たな解釈を生み出すことを目的とした鑑賞法である。一般にこの鑑賞法は、ファシリテーターやともに鑑賞する他の鑑賞者を必要とするため、鑑賞者が1人で行うことは困難である。そこで、本研究では、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 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1人でも対話型鑑賞を行うことを可能にする。システムを用いた実験では、本システムを利用することで、対話型鑑賞を実現できることを確認した。

## Proposal of Spoken Dialogue System for Aesthetic Development on Interactive Appreciation

YOSHINO YUASA<sup>1</sup> TAKASHI YOSHINO<sup>1</sup> KANAE AOKI<sup>2</sup>

### 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、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。そして、美術館における学芸員やボランティアの数は限られているため、これらのイベントを常に行うことも難しいと考えられる。本研究では、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、ユーザが一人でも対話型鑑賞を行うことを可能にする。本稿では、KANAI の概要と、WOZ 法を用いた事前実験およびシステムを用いた比較実験と、その結果に対する考察について述べる。

### 2. 関連研究

#### 2.1 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 [3]。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

<sup>1</sup> 和歌山大学

<sup>2</sup> 和歌山県立近代美術館

<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。

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 [4]。これらのシステムは、美術館で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。

## 2.2 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 [5]。アートコードとは、ユーザーが作り出した画像やパターンを、機械可読マーカー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 [6]。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。

## 2.3 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発した [7]。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 [8]。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の考えを深め、作品に関する意味や価値をユーザ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。

## 3. WOZ 法を用いた事前実験

### 3.1 概要

システム作成前に、合成音声を用いたファシリテーションによる対話型鑑賞の有用性を検証するため、WOZ 法を用いた事前実験を行った。人との対話型鑑賞を行ったグループと、WOZ 法を用いた合成音声との対話型鑑賞を行ったグループの比較実験である。

### 3.2 ファシリテーターによる質問内容

本実験におけるファシリテーターの質問内容は、小泉による論文 [1] をもとに行った。小泉は「対話を介した鑑賞における作品内容に関する発問」[1] で 7 個の基本的な発問を述べている。さらに、著者の経験と他の実践者から学んだ発問として 9 個の発問を挙げている [1]。この中で絵画

内の服の材料に関する質問など、適応する絵画に限られる質問を省き、5 個の質問を選んだ。7 個の基本的な質問と合わせて、12 個の質問を行った。

## 3.3 結果

事前実験の結果を以下にまとめる。

### ● システムについて

- (1) 合成音声を用いたファシリテーションによって対話型鑑賞を実現できる。
- (2) 合成音声を用いたファシリテーションによる対話型鑑賞は有用である。

### ● システム機能について

- (1) システムとの対話においても、相槌や肯定をその都度行う必要がある可能性がある。
- (2) ユーザの返答内容に応じた質問文の作成を行う必要がある。

### ● 鑑賞者について

- (1) 対話型鑑賞を行う際には、対話型鑑賞の目的を事前に詳しく説明し、質問に対して正解があるわけではないということを鑑賞者に理解してもらう必要がある。
- (2) 作品に関する考えを言葉で表せないとき、作品に関する知識を求める傾向がある。

## 4. KANAI

### 4.1 概要

KANAI は、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>\*3</sup>と KNP<sup>\*4</sup>と Web Speech API<sup>\*5</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、ユーザと作品の間に作り出すことを目指す。

### 4.2 事前調査

システムを作成するにあたり、まず、一つ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ものを目指した。今回は石垣栄太郎の《街》<sup>\*6\*7</sup>を題材とした。そこで、この絵画を題材として、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 2 度行ってもらった。1 度目は、ファシリテーターを学芸員が、鑑賞者を著者が行い、2 人で対話型鑑賞を行った。2 度目は、ファシリテーターを 1

<sup>\*3</sup> <http://nlp.ist.i.kyoto-u.ac.jp/index.php?JUMAN>

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

<sup>\*5</sup> [https://developer.mozilla.org/ja/docs/Web/API/Web\\_Speech\\_API](https://developer.mozilla.org/ja/docs/Web/API/Web_Speech_API)

<sup>\*6</sup> <http://www.momaw.jp/press/2013-4/ishigaki-pr.php>

<sup>\*7</sup> 著作権は切れていないが、使用可能な絵画

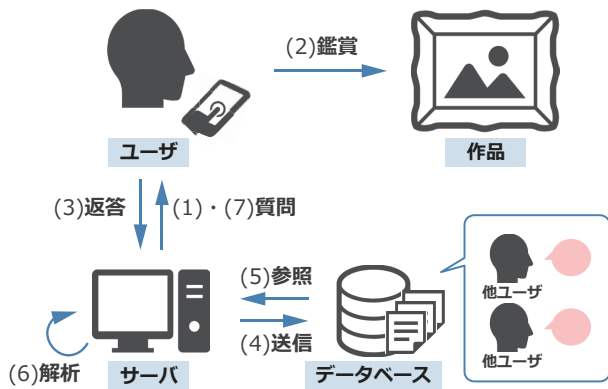


図 1 システム構成図

度目とは異なる学芸員が、和歌山大学美術館部の男女3人と著者の4人を鑑賞者とし、複数人での対話型鑑賞を行った。その際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3点を以下に示す。

- (1) 「この絵の中に何が見える？」に対する返答が、どの鑑賞者も「人」に関するものであった点  
返答例：「怪我をした男の人」
- (2) 鑑賞者が注目した事物の特徴に関する質問を行う点  
質問例：「足はどうなってる？」
- (3) 鑑賞者がまだ気づいてない内容に関して気付かせるための質問を行う点  
質問例：「帽子を持ってるのはどうしてかな？」

#### 4.3 設計方針

共同研究者である和歌山県立近代美術館の学芸員に、事前実験(3章)のアンケート結果と実験の様子を撮影した映像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3点である。

- (1) 鑑賞者の注目位置の確認
- (2) 鑑賞者の返答をまとめ肯定
- (3) 回答内容に応じた質問文の作成

事前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、これらの内容を追加する形で、システムの質問内容を作成した。

#### 4.4 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に対して最初の質問を行う。図1(2)では、ユーザが作品を鑑賞する。図1(3)では、その質問に対して、ユーザが返答を行う。図1(4)では、その返答をデータベースに送

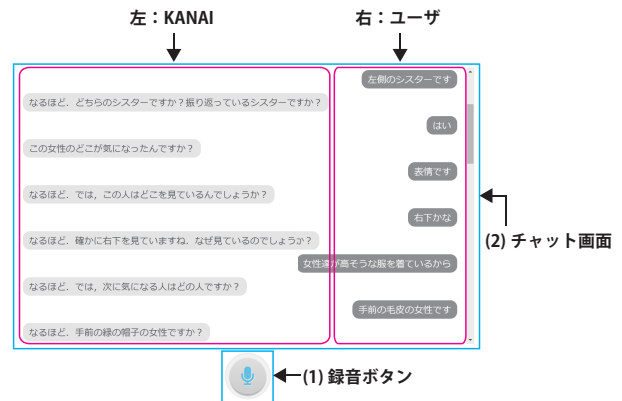


図 2 システム画面

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、ユーザと対話を行う。

#### 4.5 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。システム画面を図2に示す。

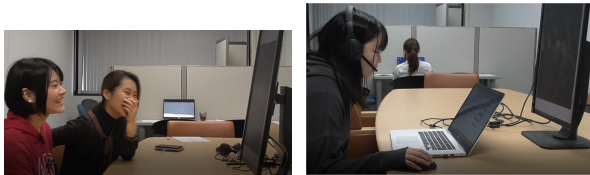
##### 4.5.1 音声認識機能

本機能は、図2(1)の録音ボタンを押すと開始する。音声認識の終了のタイミングは、ユーザの言葉が長く途切れた部分をWeb Speech APIが検知し、自動で終了する。

##### 4.5.2 質問内容作成機能

本機能は以下の3つ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が自分の解釈を示した際、次の質問に移る前にユーザが示した解釈を肯定する機能である。本システムは、ユーザの解釈を事実として認め、それによってKANAIのデータを蓄積する。この機能は、ユーザ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) 位置確認  
本機能は、ユーザが示した絵画内の事物の位置を確認する機能である。確認は以下の流れにより行う。
  - (a) ユーザの返答を分析  
ユーザの返答を分析し、返答に含まれる単語を抽出する。
  - (b) ユーザが示した絵画内の事物の推定  
抽出した単語から、ユーザの示す事物を推定する。
  - (c) ユーザに推定した事物の正誤を確認  
質問により、推定した事物がユーザの示すものか確認する。
  - (d) ユーザが示す絵画内の事物の特定



(a) 人との対話型鑑賞を行うグループの実験の様子 (b) システムとの対話型鑑賞を行うグループの実験の様子

図 3 実験の様子

ユーザの返答によって、ユーザ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。
- ユーザの返答を解析  
ユーザの返答内容を 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。
  - ユーザの返答によって質問を変更  
ユーザが返答に詰まった際に、異なる観点の質問を投げかける。

#### 4.5.3 音声合成機能

本機能は、音声認識機能によりユーザの返答を認識後、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の返答に応じた質問内容が作成された時に行われる。

#### 4.5.4 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。

### 5. システムを用いた比較実験

#### 5.1 概要

本実験では実際に KANAI を用いて、事前実験と同様に、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。

以下の 2 つの仮説をもとに実験を行った。

仮説 (1) KANAI によって対話型鑑賞を実現できる

仮説 (2) KANAI による対話型鑑賞は有用である

#### 5.2 実験内容

本実験は 2020 年 1 月 22 日、1 月 27 日、1 月 28 日の 3 日間、和歌山大学の学生 20 名を対象に行った。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

表 1 対話型鑑賞の評価項目

	評価項目
(1)	作品をじっくりと観察
(2)	自己の判断で言葉を選び表現
(3)	作品についての知識が不必要
(4)	美術作品への親近性の獲得
(5)	作品の意味を創る楽しさの獲得

との対話型鑑賞を行ったグループ、それぞれ 10 名の学生に対して実験を行った。1 人の被験者が片方の実験のみに参加する被験者間実験である。実験協力者には、今までに一度は美術館に行った経験がある人を選んだ。本システムは、美術館で利用することを想定したものであるからである。人との対話を行ったグループは、女性が 5 名、男性が 5 名の計 10 名で行った。システムとの対話を行ったグループは、女性が 4 名、男性が 6 名の計 10 名で行った。実験後、実験協力者にアンケート調査を行った。アンケートには、5 段階のリッカートスケール（以下、5 段階評価と表記する）と自由記述の設問を用いた。5 段階評価の項目は、「1：強く同意しない」「2：同意しない」「3：どちらともいえない」「4：同意する」「5：強く同意する」である。

図 3(a) に、人との対話型鑑賞を行ったグループの実験の様子を示す。右側の女性が著者であり、この対話型鑑賞におけるファシリテーターである。左側の女性が実験協力者である。2 人が見ている画面に題材とした絵画の画像が映し出されており、それを見ながら対話を行った。図 3(b) に、システムとの対話型鑑賞を行ったグループの実験の様子を示す。手前のヘッドセットを着けた女性が実験協力者である。奥の後ろ姿が見える女性は、著者である。実験協力者の前にあるノート PC には KANAI が映し出されている。そして、実験協力者の前にあるディスプレイには、題材とした絵画の画像が映し出されている。

#### 5.3 対話型鑑賞の評価項目

本実験における対話型鑑賞の評価項目は、小泉による論文 [1] をもとに考えた。「作品をじっくりと観察し、対話を介して多様な価値を受容する経験と、考えながら自己の判断で言葉を選び表現する知をもつことを、ギャラリー・トークは求めている」と小泉は述べている。それにより、表 1(1), (2) を設定した。「ギャラリー・トークは、鑑賞者が作品についての知識を持たなくても可能であるということに意味がある」と小泉は述べている。それにより、表 1(3) を設定した。「ギャラリー・トークの可能性と課題」において、「自己の感性（発話）が受容されることによる美術作品への新たな親近性の獲得」が必要であると小泉は述べている。それにより、表 1(4) を設定した。「対話を介した鑑賞の価値は、作品の意味を創りあげていく楽しさである」と小泉は述べている。それにより、表 1(5) を設定した。これらを表 1 に示す。

表 2 成功した対話例 1

	話者	対話内容
(1)	システム	なるほど、手前のオレンジの帽子の女性が気になったんですね？
(2)	ユーザ	はいそうです
(3)	システム	この女性のどこが気になったんですか？
(4)	ユーザ	どこを見ているか
(5)	システム	なるほど、では、この人はどこを見ているんでしょうか？
(6)	ユーザ	右下の緑の女性
(7)	システム	なるほど、確かに右下を見えていますね、なぜ見ているのでしょうか？

## 6. 実験結果および考察

### 6.1 KANAI による対話型鑑賞の実現に関する考察

対話型鑑賞の実現に関するアンケート結果を表 4 に示す。この表の質問項目 (1)～(5) は、上述した表 1 の評価項目 (1)～(5) に対応している。グループ名は、人との対話型鑑賞を行ったグループを「人」、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を「システム」とした。評価項目ごとに分析したアンケート結果と考察を以下に示す。

実験協力者に「美術作品を丁寧にみることができた」(表 4(1)) という質問を行ったところ、人のグループでは 5 段階評価で中央値 4、最頻値 4 と 5 という評価を得られた。それに対して、システムのグループでも、5 段階評価で中央値 4、最頻値 4 という高い評価を得られた。有意確率は 0.32 となり、人とシステムのグループで差があるとは言えないという結果になった。システムのグループで、「強く同意する」と回答した実験協力者とシステムの対話を表 2 に示す。表 2(1)～(2) において、ユーザの示した事物の位置特定ができている。さらに、表 2(6)～(7) において、ユーザの発言から「右下」を抽出し、ユーザの対話内容を次の質問に反映することができており、自然な対話を行っている。システムのグループで、「同意しない」と回答した実験協力者とシステムの対話を表 3 に示す。表 3(3) の質問後の注目点の特定が十分に出来ていなかったことが分かった。原因としては、音声認識の誤変換を想定できていなかったことが考えられる。表 3 の (4) の「家」が「いえ」の誤変換である。ユーザが、絵画内の複数人について同時に注目する返答を行うことを、著者が想定していず、システムが対応できなかったことも原因として考えられる。表 3 の (4) の「2 人ともにです」がそれにあたる。以上のことから、システムによるファシリテーションでも、人によるファシリテーションと同様に、「美術作品をじっくりと観察する」(表 1(1)) ができたと考えられる。しかし、今後は、著者が想定した絵画内の事物の位置情報だけで、ユーザの示した事物の特定を行うのではなく、上手くいかなかった場

表 3 失敗した対話例 1

	話者	対話内容
(1)	システム	なるほど、では、どの人に一番注目しましたか？手前、真ん中、奥のどの人ですか？
(2)	ユーザ	手前の人に注目しました
(3)	システム	なるほど、手前の緑の帽子の女性ですか？
(4)	ユーザ	家 2 人ともにです
(5)	システム	すいません。合っていますか？

合に他の方法で特定することができるようにシステムを作り直す必要があると考えられる。

実験協力者に「作品に関する自分の考えを言葉で表現できた」(表 4(2)) という質問を行ったところ、人のグループでは 5 段階評価で中央値 4、最頻値 4 という評価を得られた。それに対して、システムのグループでも、5 段階評価で中央値 4、最頻値 4 という高い評価を得られた。有意確率は 0.083 となり、人とシステムのグループで差があるとは言えないという結果になった。以上のことから、システムによるファシリテーションでも、人によるファシリテーションと同様に、「自己の判断で言葉を選び表現」(表 1(2)) ができたと考えられる。

実験協力者に「対話の中で、絵画に関する知識を必要とした」(表 4(3)) という質問を行ったところ、人のグループでは 5 段階評価で中央値 2、最頻値 2 という評価を得られた。それに対して、システムのグループでも、5 段階評価で中央値 1.5、最頻値 1 という高い評価を得られた。有意確率は 0.046 となり、人とシステムのグループで差があるという結果になった。評価の分布から、システムの方が高い評価を得たということが分かった。以上のことから、システムによるファシリテーションでも、人によるファシリテーションと同様、またはそれ以上に、「作品についての知識が不必要」(表 1(3)) であったと考えられる。

実験協力者に「美術作品に対する親近感が増した」(表 4(4)) という質問を行ったところ、人のグループでは 5 段階評価で中央値 3.5、最頻値 4 という評価を得られた。それに対して、システムのグループでも、5 段階評価で中央値 4、最頻値 3 という高い評価を得られた。有意確率は 0.025 となり、人とシステムのグループで差があるという結果になった。評価の分布から、システムの方が高い評価を得たということが分かった。以上のことから、システムによるファシリテーションでも、人によるファシリテーションと同様、またはそれ以上に、「美術作品への親近性の獲得」(表 1(4)) ができたと考えられる。

実験協力者に「鑑賞を楽しいと感じた」(表 4(5)) という質問を行ったところ、人のグループでは 5 段階評価で中央値 4、最頻値 4 という評価を得られた。それに対して、システムのグループでも、5 段階評価で中央値 4.5、最頻値 5 という高い評価を得られた。有意確率は 0.41 となり、人とシステムのグループで差があるとは言えないという結果

表 4 対話型鑑賞の実現に関するアンケート結果（5段階評価）

	質問項目	グループ名	評価の分布					中央値	最頻値	有意確率
			1	2	3	4	5			
(1)	美術作品を丁寧に鑑賞することができた	人	0	2	0	4	4	4	4 と 5	0.32
		システム	0	1	0	5	4	4	4	
(2)	作品に関する自分の考えを言葉で表現できた	人	0	0	2	5	3	4	4	0.083
		システム	0	0	3	6	1	4	4	
(3)	対話の中で、絵画に関する知識を必要とした	人	3	4	2	0	1	2	2	0.046*
		システム	5	3	1	1	0	1.5	1	
(4)	美術作品に対する親近感が増した	人	0	2	3	4	1	3.5	4	0.025*
		システム	0	0	4	3	3	4	3	
(5)	鑑賞を楽しみ感じた	人	0	0	0	6	4	4	4	0.41
		システム	0	1	1	3	5	4.5	5	

・評価の分布はそれぞれ「1: 強く同意しない」「2: 同意しない」「3: どちらともいえない」「4: 同意する」「5: 強く同意する」である。

・人がファシリテーターを行い、対話型鑑賞を行ったグループを「人」、合成音声によるファシリテーションを行ったグループを「システム」とする。

・\*: 有意差あり  $p < 0.05$

・有意確率には、Wilcoxon の符号付き順位検定を使用。

表 5 対話型鑑賞の有用性に関するアンケート結果（5段階評価）

	質問項目	グループ名	評価の分布					中央値	最頻値	有意確率
			1	2	3	4	5			
(1)	この鑑賞方法に精神的負担を感じた	人	5	2	3	0	0	1.5	1	0.083
		システム	6	3	1	0	0	1	1	
(2)	今後もこの鑑賞方法を用いたいと思った	人	0	0	1	4	5	4.5	5	0.56
		システム	0	0	0	7	3	4	4	

・評価の分布はそれぞれ「1: 強く同意しない」「2: 同意しない」「3: どちらともいえない」「4: 同意する」「5: 強く同意する」である。

・人がファシリテーターを行い、対話型鑑賞を行ったグループを「人」、合成音声によるファシリテーションを行ったグループを「システム」とする。

・\*: 有意差あり  $p < 0.05$

・有意確率には、Wilcoxon の符号付き順位検定を使用。

になった。以上のことから、合成音声によるファシリテーションでも、人によるファシリテーションと同様に、「作品の意味を作る楽しさの獲得」(表 1(5)) ができたと考えられる。

## 6.2 KANAI による対話型鑑賞の有用性に関する考察

KANAI による対話型鑑賞の有用性に関するアンケート結果を表 5 に示す。実験協力者に「この鑑賞方法に精神的負担を感じた」(表 5(1)) という質問を行ったところ、人のグループでは 5 段階評価で中央値 1.5、最頻値 1 という評価を得られた。それに対して、システムのグループでも、5 段階評価で中央値 1、最頻値 1 という高い評価を得られた。有意確率は 0.083 となり、人とシステムのグループで差があるとは言えないという結果になった。

実験協力者に「今後もこの鑑賞方法を用いたいと思った」(表 5(2)) という質問を行ったところ、人のグループでは 5 段階評価で中央値 4.5、最頻値 5 という評価を得られた。それに対して、システムのグループでも、5 段階評価で中央値 4、最頻値 4 という高い評価を得られた。有意確率は 0.56 となり、人とシステムのグループで差があるとは言

ないという結果になった。

## 6.3 実験協力者の意見と考察

### (1) 音声認識の終了判定の自動化

自由記述欄において「自分の話した内容が間違っていたり、途中までしか音声入力されなかったりしたので、修正する機能があればいいと思った」というコメントが得られた。今回、返答の録音時間は、自動でシステムがユーザが話し終えたと判断すると、録音を終える設定になっていた。しかし、対話型鑑賞において鑑賞者は、返答をしながら絵画についての考えをまとめ、話し始めた時より多くのことに途中で気付いていく傾向がある。そこで今後は、録音時間をユーザ自身で決めることができる機能が必要であると考えられる。

### (2) ユーザによる返答の修正

自由記述欄において「音声入力が間に合わないことがあったので、やり直し機能があるといいと思った」というコメントが得られた。今後は、返答を修正することができる機能が必要であると考えられる。

## 7. おわりに

本研究では、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 を開発した。本システムを利用することで、ユーザが一人でも対話型鑑賞を行うことができるシステムを目指す。本稿では、KANAI の概要と、WOZ 法を用いた事前実験およびシステムを用いた比較実験と、その結果に対する考察について述べた。

- (1) 音声認識の終了判定を自動ではなくユーザ自身で、決定できるようにする必要がある。
- (2) 返答の言い直しをする機能の追加が必要である。
- (3) 作品内の事物の位置特定方法について検討し、変更・追加する必要がある。

今後は上記の改善点を解決するため、システムを改良していく。また、実際に美術館で本システムを利用した評価実験を行う。

## 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] 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)。
- [4] Amartya Banerjee, Rovik Robert, Michael S.Horn: FieldGuide: Smartwatches in a Multidisplay Museum Environment, CHI'18 Extended Abstracts, ACM, pp.1–6 (2018)。
- [5] 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)。
- [6] Kari Salo, Merja Bouters, Tommi Mikkonen: User Generated Soundscapes Activating Museum Visitors, SAC 2017, ACM, pp.220–227 (2017)。
- [7] 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)。
- [8] 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)。