

音声入力と感情推定情報を活用する対話インタラクション方式 ：感情適応提示技術の提案と検証

原 梓織[†] 深澤 伸一[†] 赤津 裕子[†]

概要：音声入力と感情推定情報を活用する、公共機器の新たな対話インタラクション方式と、それを実現させるための感情適応提示技術を創案し、有効性をユーザー評価実験により検証した。感情適応提示技術とは、機器の情報提示に対するユーザーの表情・しぐさ等の感情の表出をセンシングし、提示した情報がユーザーの意図する内容と近かったかどうかを判定して、遠かった場合は異なる内容の情報を提示する技術である。効果として、音声入力の補助や、機器の操作性・親しみ感等の向上が期待される。評価実験として、券売機での自力購入へのこだわりのなかった 12 名の被験者に、感情適応提示機能の有／無と GUI/VUI（音声入出力）の条件を組合せた、従来 GUI、感情適応 GUI、感情適応 VUI の 3 条件の試作システムを操作させ、行動および心理指標により比較した。その結果、感情適応 GUI および VUI 条件は従来 GUI 条件よりも操作時間や誤操作回数が有意に少なかった。また、感情適応 VUI 条件では機器の操作負荷感、機器に対する親しみ感、機器の使用満足感等が他の GUI 2 条件よりも有意に高かった。すなわち、感情適応提示技術は被験者の機器操作の支援に有効であるが、それがユーザーに体感されたのは VUI 条件だけであり、効果がユーザーインタフェースの方式（使用モダリティの違い）と相互作用を有することが示唆された。

1. はじめに

1.1 背景

近年、音声を用いて機器を操作する Voice User Interface（以下“VUI”）を備えたシステムが増加している。増加の背景としては、VUI が人間の特性に適合したシステム操作方法であることが挙げられる。葛岡[1]は、VUI を利用する利点として「特別な訓練なく利用ができ、理解も容易である」点を挙げており、VUI は人間にとって自然に扱うことのできるユーザインタフェース（以下“UI”）であると考えられる。さらに Google I/O 2017 にて Google Speech API の音声認識時の誤認識率が 4.9% に低下したことが報告され、人間の誤認識率である約 4%[2] に近付き、音声認識技術が実用フェーズに入ったことを受け、この数年の間に Google Home や Amazon Echo などの AI スピーカーが登場してきた。このような流れの中、今後 VUI を搭載するシステムはさらに増加していくことが予想される。

一方で、Graphical User Interface（以下“GUI”）を備えた従来の対話型システムに VUI を追加導入する場合には、GUI と VUI を複合させた UI の設計法に関する体系的な指針は既存にほとんどなく、研究知見の拡充が求められている。

1.2 関連研究

VUI を備えた音声対話システムの目指す姿として、河原[3]はホテルのコンシェルジュやデパート店員の例を挙げ、その実現には、ユーザーの情報要求が明確でない場合でもユーザーの意図や思考を対話を通じて明確化できること、さらに目標を明確に有していないユーザーに対しシステムが積極的に情報を提示することが必要であると述べている。

また、中嶋ら[4]は、システム設計時に目標達成の効率という観点だけではなくユーザーのフィーリングや満足度も

考慮するべきであるとし、社会的応答として情緒的情報をやり取りするエージェントシステムの効果に関する評価実験を行った。その結果、エージェントなしよりもエージェントが存在する条件が、さらにそれらよりエージェントが感情表出を行う条件の方がユーザーの学習定着度の効果が高かった。そして、主観評価より、システムに対するユーザーの使いやすさや満足度の印象は、システムのエージェントに対する「親しみ」感の度合いに影響されていた。

1.3 研究課題

1.3.1 VUI による券売機のユーザテストによる課題

このような中、著者らは、社内で VUI を新たに搭載した次世代の駅指定席券売機プロトタイプシステムのユーザビリティ評価実験に観察者として参加する機会を得た。その際の調査経験をまとめ、駅指定席券売機への VUI の導入にあたり次のような課題を導出した。

課題 1：音声操作の最適な操作フローはタッチ操作とは異なる

音声操作の券売機では手続き型の操作フローを用いるのではなく、みどりの窓口の対面対話で発話をするように、一度に入力が出来るような対話設計を行うべきである。

課題 2：音声による入力情報の曖昧さ

音声入力の場合、入力情報に対して曖昧さの許容が必要である。そして、曖昧さの許容によって発生する情報の不足は、システムがユーザーに問いかけたり、センサデータを活用したりすることで補う仕組みが必要である。

課題 3：音声の誤認識や認識失敗時にユーザーへの聞き直し処理が発生し操作が進まない

VUI を搭載した指定席券売機が高雑音環境下で誤認識した場合や認識自体が出来なかった場合にシステムがユーザーにどうふるまうのかを適切に設計する必要がある。

まとめると、システムに VUI を搭載する際には、ユーザ

[†] 沖電気工業株式会社 研究開発センター

一が人と喋るような自然な発話での入力を可能にすることが必要であり、そのためにはいかに「ユーザーの曖昧な入力を許容」しながら「ユーザーの意図する操作内容にたどり着くか」が課題となると言える。また、中嶋ら[4]によれば、対話システムに対して「親しみ」感を持てることがユーザーの使いやすさや満足度の印象につながる。VUIを備えた対話システムにおいて、音声による自然な対話と、曖昧な情報入力の許容による操作支援が実現されることで、システムに対するユーザーの親しみ感が醸成され、システム利用の満足度の向上につながる可能性がある。

2. 「感情適応提示技術」の提案

前章で挙げた課題の解決手段として、対話システムの新たなインタラクションの方式を実現させる「感情適応提示技術」を創案した。本技術は、システムが関連情報を提示し、その情報提示に対するユーザーの表情・しぐさ等の感情表出行動のセンサデータに基づいて、提示した関連情報がユーザーの意図に適合していたかどうかを推定処理して、その結果に応じて異なる情報の提示を行うものである。従来の手続き型のフローとは異なり、検索に最低限必要な条件のみ入力を行えば情報検索ができ、システムがユーザーの意図を読み取り異なる情報提示を行うという補助のふるまいを行うことで、自然な発話入力を可能にし、曖昧な入力を許容することができる。

音声による入力は不充分または曖昧な情報入力になることが多いこと、さらに音声のセンサデータを追加収集できることでユーザーの意図推定の精度を上げられ、その結果ユーザーの曖昧な入力からでもユーザーの意図する操作を実現できる可能性があることから、感情適応提示技術はVUIとの相性が良い技術と考えられる。

2.1 スロットフィリングとスロットの種類

本技術では、ユーザーの自由発話による対話システムへの入力があった際に、スロットフィリング方式[5]を用いて入力情報が埋められていく。スロットの種類には必須スロット[5]と呼ばれるものがある。これは関連情報の検索に最低限必要なスロットであり、例えば駅指定席券売機の場合、「出発駅」、「到着駅」等がこのスロットに該当する。

一方、必須スロットの他に、関連情報の検索に必須ではないがユーザーが関連情報を検索する際に考慮する要素になるスロットが存在する。このスロットを“任意スロット”と定義する。任意スロットは、ユーザーのその時々感情や状況によって入力情報の必要性や重要性が変化するスロットであり、例えば駅指定席券売機の場合、「乗換回数」、「金額」等がこのスロットに該当する。

まとめると、本技術で用いるスロットには以下の2種が存在する。

- 必須スロット
- 任意スロット

2.2 関連情報の検索と情報提示の調整方法

前述の必須スロットの入力内容を基に、システムは関連情報を検索する。また、任意スロットに入力があった場合、システムはユーザーがそのスロットに対応する経路条件を重要視しているとみなし、重要度のパラメータ設定・調整を行う。そして検索された複数の関連情報の中で重要度の数値が最も高いスロットの関連情報について候補となる複数の経路条件間で比較をし、ユーザーにとって高い評価値をとるような関連情報を選択し、提示する。

さらに、提示された情報に対するユーザーの表情、しぐさ、音声情報などをセンサによって取得し、そのセンサデータから、ユーザーの感情状態を推定することによって、ユーザーが提示された情報に対し、どのような感情になったかを把握して、他の関連情報と比較した際に、提示された関連情報がユーザーにとって最も低い評価値をとる関連情報のスロットの重要度が高くなるように、重要度を再設定する。そして再設定された重要度を基に、再度前述の情報提示に係る処理を行う。情報提示と重要度の変更を繰り返すことで、最終的にユーザーの意図に最も適合する関連情報を探し出すことができる。

3. 評価実験

前章で提案した感情適応提示技術の応用先のシステムを公共機器の駅指定席券売機とし、感情適応提示技術やGUI/VUIの有無による使用性比較のための3種類のプロトタイプシステムを開発して、各システムを被験者ユーザーに操作させ、評価した。評価指標は、主観評価アンケートやインタビューによる心理指標と、システムの操作時間や誤操作回数、表情やしぐさ表出等による行動指標とした。

3.1 仮説

本実験で検証する仮説は次のとおりである。

仮説1: VUIはGUIに比べ、ユーザーのシステム操作負荷感が小さくなり、システムに対する親しみ感と満足度が高くなる。(1.3.1参照)

仮説2: 感情適応提示型のUIは従来手続き型のUIと比べ、誤操作回数や困り回数が少なくなり、システムに対する親しみ感と満足度が高くなる。(2.参照)

3.2 実験タスク

券売機プロトタイプシステムを用いて、特定の条件を満たすきっぷを購入させるタスクを被験者に行わせた。後述する3種類の券売機プロトタイプシステムそれぞれに同程度の難易度で入力情報が異なるタスクを用意し、到着駅が関東で新幹線を用いる経路案を2種類用意した。また、経路選択条件として被験者が守るべき「到着時刻」を設定し、さらにできるだけ「乗換回数が少ない」経路を2経路の中から選んで購入を行うように被験者に教示した。本実験では、券売機プロトタイプのユーザー応答にWizard of OZ法[6]（以下“WOZ法”）を用い、感情適応提示技術の挙動お

よび被験者の感情表出の認識を、システムではなく実験者が行った。各タスクの開始および終了時点は、被験者が購入するきっぷの券種を指定し、次の操作画面へ遷移した段階で開始とし、上記2経路の比較を完了し、きっぷ料金の支払い要求画面へ遷移した段階で終了とした。詳しい実験手順は3.6節で述べる。

3.3 実験条件

実験に使用した3条件のプロトタイプシステムは、以下A～Cの通りである。図1にそれぞれの実験条件の入力方式と画面例の比較を示す。

A. 従来（手続き）型 GUI 条件（以下，“従来 GUI”）

基準的な比較条件として設定し、現在駅で利用されている指定席券売機のUIを模した画面構造と遷移を行う条件である。手続き型のフローであり、被験者は駅名や経路情報など全ての項目を段階的に入力しなくてはならず、操作負荷がかかると考えられる。画面数は、条件終了時に表示される紙幣投入画面を合わせ全部で16画面あり、被験者には新幹線到着駅の異なる2経路を検索させた。

B. 感情適応提示型 GUI 条件（以下，“感情適応 GUI”）

2章で説明した感情適応提示技術とタッチパネルによるGUI入出力を組み合わせた条件である。重要条件の入力には手続き型と同様の操作フローでタッチ入力させ、任意条件はソフトキーボードを用いて自由文形式（複数語句による条件入力）でタッチ入力させた。画面数は、条件終了時に表示される紙幣投入画面を合わせて全部で10画面あり、被験者には新幹線到着駅の異なる2経路を比較させた。

C. 感情適応提示型 VUI 条件（以下，“感情適応 VUI”）

この条件は感情適応提示技術と音声によるVUI入出力、さらに券売機画面によるGUI出力を組み合わせた条件である。被験者は、乗車日等経路の検索に必要な情報だけではなく経路情報のような被験者の好みや状態、その時の状況によって変化をする項目に関しても音声発話によって自由に入力することが出来る。システムは被験者に対し、画面出力と音声出力を併用し情報の提示を行う。実験条件の統制のため、出力音声には音声合成ソフトウェア『OKI SMARTTALK 3.0』を用いて作成した合成音声を利用した。

ユーザーの入力内容に応じた画面表示や音声応答の出力には前述のようにWOZ法を用いた。画面は、実験終了時に表示される紙幣投入画面を合わせて全部で10画面あり、被験者には新幹線到着駅の異なる2経路を比較させた。

3.4 実験環境

図2に実験風景の例を示す。被験者は立った状態で駅指定席券売機の形状を模したモックアップ筐体の前に立ち、操作を行った。また、モックアップ筐体正面には被験者の顔動画像を観測するためのWebカメラを設置した。

条件	従来（手続き）型 GUI	感情適応提示型 GUI	感情適応提示型 VUI
駅名入力画面例			
入力方式	タッチパネル／一覧表示 ボタン入力	タッチパネル／一覧表示 ボタン入力	音声入力／一問一答形式入力
経路条件入力画面例			
入力方式	タッチパネル／一覧表示 ボタン入力	タッチパネル／自由形式入力	音声入力／自由形式入力

図1 実験3条件の入力方式と入力画面例



図2 実験風景

3.5 被験者

16名の被験者で実験を実施したが、実験後のインタビューやアンケート結果から、従来GUIを強く好み感情適応提示機能のような機器の支援を好まない心理特性を持った被験者が4名いたことがわかったため、それら4名を除いた12名の被験者のデータを分析した。被験者12名の属性は次のとおりである：性別・女性6名／男性10名、年齢・20～40代、指定席券売機の利用頻度・2～3か月に1回：3名／半年に1回：3名／1年に1回：3名／それ以下：7名。

3.6 実験手続き

実験は、システムの操作および評価を行う被験者と、WOZ法でシステムの挙動を操作する実験者で実施した。被験者は実験室に入室後、実験参加同意書及びユーザー属性を調査するアンケートに回答した。次に、旅行に行くために指定のきっぷを購入するというカバーストーリーを用いてシステム操作のタスクの説明を被験者に行った。タスク説明の例として、従来GUI条件の説明を以下に記載する。

- あなたは、友人と旅行へ行くことになった
- 到着駅の近くのお店を予約している
- 同伴者の友人は足を痛めているため、乗換回数が少な

い経路が望ましい

- システムを使って、予約の時間に間に合うかつ乗換回数が少ない経路のきっぷを購入してほしい

その後、被験者に購入するきっぷの条件を指示し、システムの操作を開始させた。なお、3条件の実験順序は被験者毎にカウンタバランスをとった。実験者は被験者の顔動画像、操作ログおよび操作画面を目視確認しながら WOZ 法の反応を行い、被験者が 3 回以上誤操作した場合には次の操作内容を教える介入を行った。タスク終了後、被験者に対し直前に操作したシステムに関する後述の主観評価を回答させた。

3.7 計測項目

3.7.1 心理指標

(1) 既存質問項目

前東ら[7]による自動化システムのユーザビリティ評価アンケートを用い、本実験の各条件のシステムに対する有効さ、効率、満足度、意図の理解、違和感、モチベーションを評価した(次の項番 10 項目を抜粋使用:A1, A2, B1, B2, C1, C2, D2, D3, E2, E3, F1, F2, 質問文は結果グラフに記載)。

(2) 独自質問項目

本実験の各条件のシステムの操作負荷感、補助感、親しみ感といった印象に関する質問として次の 4 項目を作成し、7 段階のリッカート尺度(全くそう思わない~非常にそう思う)で被験者に回答させた:

- このシステムによる駅名の指定操作はわずらわしかった(駅名指定の操作操作負荷感)
- このシステムによる経路条件の指定操作はわずらわしかった(経路条件指定の操作操作負荷感)
- このシステムは自分が行いたい操作内容を推測して助けてくれようとしていた(システムの補助感)
- 使用中に、このシステムに親しみを感じた(システムに対する親しみ感)

3.7.2 行動指標

(1) 操作時間

各条件のタスク完了までの操作時間を測定した。

(2) 誤操作回数

被験者の誤操作回数を測定した。誤操作の判定は、タスクの正しい手順以外の操作、つまり次の画面に遷移しない操作を行う毎に 1 回の誤操作としてカウントした。

(3) 表情表出回数

仮説 2 の困り回数を検証するために、被験者の表情・しぐさ等による困りの表情変化回数の測定を行った。困りの表情変化の回数は、被験者がシステムを操作している顔動画像を確認し、表情が無表情状態から大きく変化したことを録画映像から実験者が判定した回数をカウントした。

4. 結果

4.1 分析方法

各測定項目に対し、従来 GUI 条件、感情適応 GUI 条件、感情適応 VUI 条件の 3 条件間で多重比較を行った。まず各項目に対し Shapiro-Wilk の正規性検定を行い、パラメトリックであった場合は TukeyHSD 法を、ノンパラメトリックデータであった場合は Steel-Dwass 検定による多重比較を行った。有意水準は 5%、有意傾向は 10%に設定した。

4.2.1 心理指標 1: 既存項目 (有効さ, 効率, 満足度)

既存項目心理指標の前東ら[7]のアンケート 10 項目は、3 条件間の多重比較の結果、有効さ、効率、満足度の 3 要素に属する 3 項目(項番 A1, B2, C2)で感情適応 VUI は従来 GUI より有意に高い評定値であり、さらに効率と満足度においては、感情適応 VUI は感情適応 GUI よりも有意または有意傾向で高い評定値であった(図 3, 4, 5)。他の 7 項目については、条件間に有意な差は認められなかった。

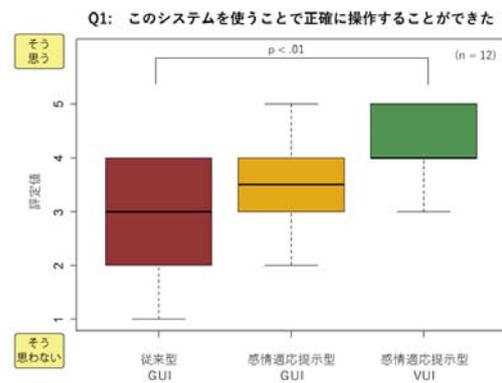


図 3 システムの有効さ

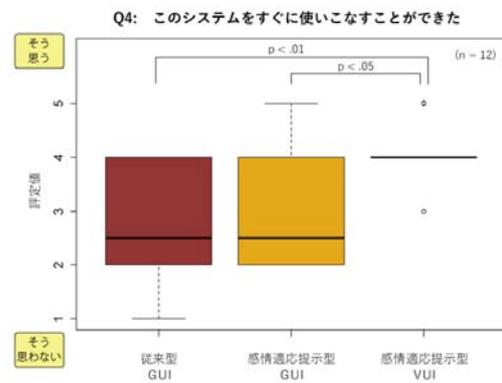


図 4 システムの使用効率

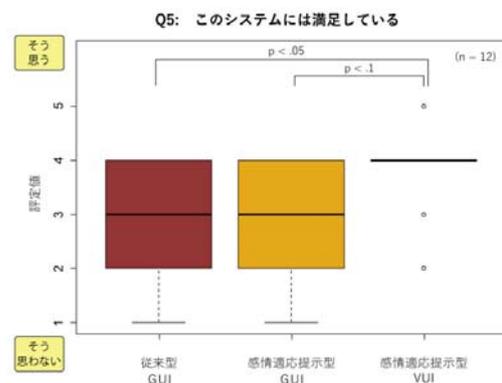


図 5 システムの使用満足度

4.2.2 心理指標 2：独自項目（操作難易感，補助感，親しみ感）

本研究の独自項目心理指標について，3条件間の多重比較の結果，駅名指定の操作負荷感，システムの補助感，システムに対する親しみ感で，感情適応 VUI は従来 GUI および感情適応 GUI よりも有意に高い評価値であった(図 6, 7, 8)．経路条件指定の操作負荷感については，条件間に有意な差は認められなかった．

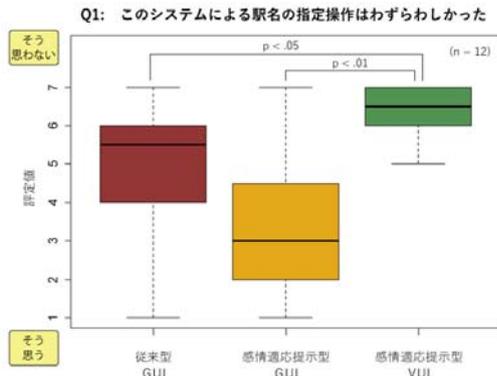


図 6 駅名指定の操作負荷感

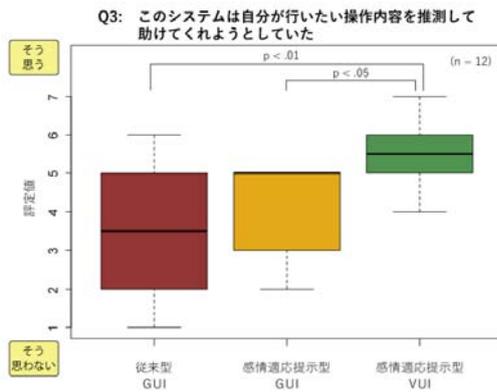


図 7 システムの補助感

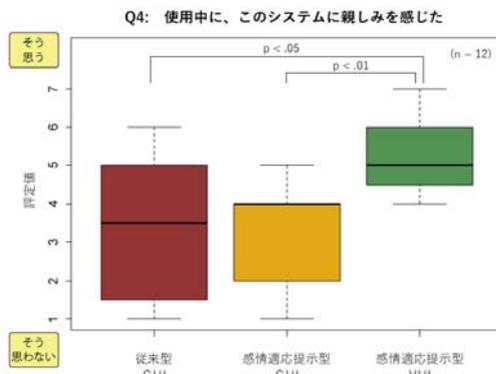


図 8 システムに対する親しみ感

4.2.3 行動指標（操作時間，誤操作回数，感情表出回数）

行動指標について，3条件間の多重比較の結果，システムの操作時間，誤操作回数で，感情適応 VUI は従来 GUI および感情適応 GUI よりも有意に少ない測定値であった(図 9, 10)．また，システム操作中の困り感情表出回数で，感情適応 VUI は従来 GUI よりも有意に少ない表出回数であった(図 11)．

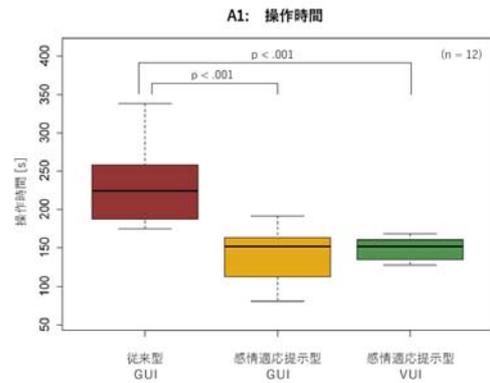


図 9 システムの操作時間

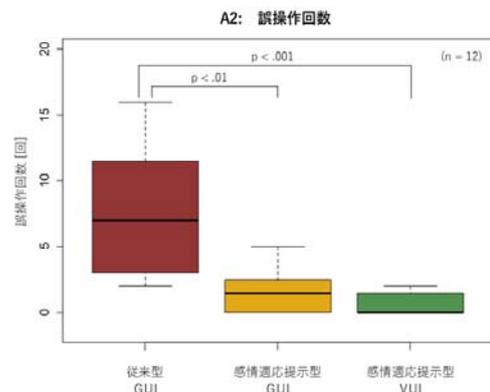


図 10 システムの誤操作回数

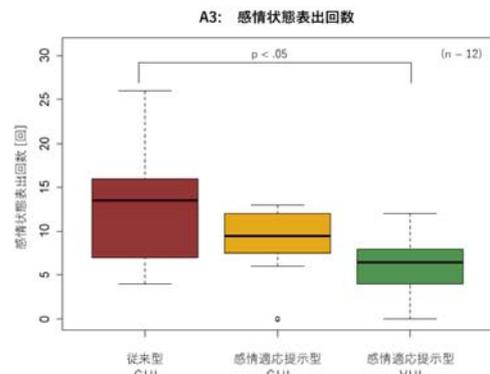


図 11 システム操作中の困り感情表出回数

5. 考察

5.1 仮説の検証

仮説 1:「VUI は GUI に比べ、ユーザーのシステム操作負荷感が小さくなるため、システムに対する親しみと満足度が高くなる。」

実験の結果、仮説 1 は採択された。感情適応 VUI 条件は GUI 2 条件と比べ、駅名指定の操作負荷感が有意に低く(図 6)、補助感(図 7)やシステムへの親しみ感(図 8)、さらに、満足度(図 5)も最も高かった。

仮説 2:「感情適応提示型のシステムは手続き型に比べ、誤操作回数や困り回数が少なくなるため、システムに対する親しみ感と満足度が高くなる。」

実験の結果、仮説 2 は一部採択された。感情適応 GUI および VUI 条件は従来 GUI 条件よりも操作時間(図 9)や誤操作回数(図 10)が有意に少なかった。一方で、困り感情表出回数(図 11)、親しみ感(図 8)、満足度(図 5)は、感情適応 GUI と従来 GUI 条件間で有意な差は認められなかった。

5.2 感情適応 GUI と VUI 条件間の効果の差(仮説 2)

仮説 2 の検証結果は一部採択となり、一部の評価項目について想定していた結果を得られなかった。

まず、操作時間や誤操作回数では有意差が認められていたのにも関わらず、困り感情表出回数で感情適応 GUI と従来 GUI 条件間に有意な差が認められなかった点に関しては、GUI と VUI の操作特性の違いに起因する可能性がある。従来 GUI は実際の駅券売機の GUI を模した UI になっており被験者にもある程度の慣れがある可能性があるが、感情適応 GUI は本実験が初見の UI となることから被験者に困惑の感情が生じたのかもしれない。一方で、感情適応 VUI も初見の UI ではあるが、背景で述べたように音声入力は事前訓練を必要としない特性を持った UI であることから、感情適応 GUI で見られた初見操作による困惑発生の影響を抑えられたのかもしれない。

本実験の結果から、感情適応提示技術は UI の使用モダリティの違いと相互作用を有し、2章で述べたように VUI との相性が良い技術であることが示された。

5.3 総合考察

(1) 感情適応提示技術の有効性と VUI との相性

感情適応提示機能を有する感情適応 GUI および VUI 条件では、従来 GUI 条件よりも、操作時間と誤操作回数が有意に少なく、主観的にも有意に容易感が高かった。一方で、親しみ感と満足度は、感情適応 GUI と従来 GUI 条件間で有意な差は認められなかったことから、感情適応提示機能はシステム操作行動の改善に有効であるが、その効果がユーザーに主観的に認められやすいのは、音声入力を備えた UI (VUI) の場合であることが示唆された。ただし、本実験では、感情適応 GUI のデザインを従来 GUI から大きく

は変えないように設計した。従来の駅券売機の UI の仕様に縛られることなく、大幅な画面構造や遷移の設計を行えば、GUI でも感情適応提示機能の有効性が向上する可能性があり、今後の検討課題である。

(2) VUI は操作のための事前訓練が必要ない

本実験の結果を総合すると、感情適応提示機能を有した VUI は、従来の駅券売機 UI と比較して、操作の事前訓練なく有意に少ない操作時間と誤操作回数で操作でき、さらに、親しみ感や満足度が有意に高くなった。感情適応提示技術のある VUI は、さまざまな特性や状況(初回使用を含む)のユーザーが利用する公共端末と相性が良いインタラクション方式といえるだろう。

6. まとめ

VUI を備える対話システムの新たなインタラクションの方式を実現させる「感情適応提示技術」を創案し、同方式を体験できる UI を備えた公共機器のプロトタイプシステムを開発して、ユーザー評価実験により効果検証を行った。実験の結果、感情適応提示機能を有する VUI の UI 条件において、システムの操作負荷感は GUI よりも有意に低く、親しみ感、満足度は有意に最も高かった。VUI を備える対話システムのインタラクションの方式として、感情適応提示技術の方法が有効であることが示唆された。

参考文献

- [1] 葛岡英明他. ヒューマンコンピュータインタラクション 改訂 2 版. オーム社, 2002, pp.82-83.
- [2] Lippmann, R. P. Speech recognition by machines and humans. *Speech Communication*, 1997, vol. 22, no.1, pp.1-15.
- [3] 河原達也. 音声対話システムの進化と淘汰: 歴史と最近の技術動向. *人工知能学会誌*, 2013, vol. 28, no.1, pp.45-51.
- [4] 中嶋宏, 森島泰則, 山田亮太他. 人間-機械協調システムにおける社会的知性. *人工知能学会論文誌*, 2004, vol. 19, no. 3, pp.184-196.
- [5] 関口真理子, 荒金陽助, 阿部伸浩. 1bit 状態遷移表シナリオ記述方式のスロット状態拡張に関する検討. *情報処理学会 研究報告 高度交通システム (ITS)*, 2004, vol. 114, pp.71-78.
- [6] Fraser, N. M. and Gilbert, G. N. Simulating speech systems, *Computer Speech and Language*, 1991, vol. 5, no. 1, pp.81-99.
- [7] 前東晃礼, 三輪和久, 小島一晃, 寺井仁, 鈴木達也. 自動化システムのユーザビリティ評価アンケートの開発. *電子情報通信学会論文誌 A*, 2017, vol. 100, no.6, pp.240-250.