

ユーザの位置計測に基づいた 自然なインタラクション遷移を実現する デジタルサイネージシステム

菅井 ゆり佳¹ 杉村 大輔¹

概要：公共の場に大型ディスプレイやプロジェクターなどの映像表示装置を設置し、広告やニュース、株価などの情報を発信するデジタルサイネージが普及してきた。しかしながら、ディスプレイ上で画像を切り替えたり動画を流し続けたりするなど、一方的な情報の提示が主となっている。また、大規模なデジタルサイネージと比べると、小規模なデジタルサイネージは費用対効果が小さい。このような問題に対処するため、ユーザの位置計測に基づいた自然なインタラクション遷移を実現するデジタルサイネージシステムを提案する。提案システムの特徴は2つある。1つ目は、スクリーンとユーザとの距離に応じて情報提示方法を4つのフェーズに分けることにより、ユーザとシステムの暗黙的なインタラクションから明示的なインタラクションへの遷移を実現することである。ユーザとシステムのインタラクションを可能にすることで、一方的な情報提示という従来システムの問題を解決する。2つ目は、特別な装置を必要とせず、iPhone、iPadとプロジェクターを用いるだけでシステムを構成することができることである。これにより、システム全体のコストを抑えることが可能になる。本システムのプロトタイプを試作し、有効性を確認した。

1. はじめに

駅や店舗、施設、オフィスなどの公共の場に大型ディスプレイやプロジェクターなどの映像表示装置を設置して、広告やニュース、株価などの情報を発信するデジタルサイネージが普及してきた。デジタルサイネージは、紙媒体での看板広告や同じ静止画を切り替えるだけのロール・スクリーン看板とは異なり、デジタル情報を表示する電子看板である。そのため、広告などの情報の差し替えや変更が容易になる。また、動画や3Dコンテンツなどを表示することが可能であるため、ユーザの関心を惹きやすい。そして、不特定多数に同じ広告を流すのではなく、通信ネットワークを用いて、設置場所の地域性やターゲット層を考慮した広告表示ができる。このような利点を踏まえ、デジタルサイネージでの広告は年々増加している。

一方で、問題点は2つ挙げられる。1つ目は、ディスプレイ上で画像を切り替えたり動画を流し続けたりするなど、一方的な情報の提示が主となっている点である。このような一方的な情報の提示では、紙媒体での広告と大差なく、広告に関心のないユーザの注目を得にくい。また、これを解決するための方法の一つとしてディスプレイとユーザとのインタラクションを促すタッチパネルがあるが、ユーザ

は画面に接触する必要がある。さらに、大きなディスプレイを用いる場合は、手が届かない欠点がある。これを解決するために、マーカや装置を用いてユーザの体や手の位置を測定し、ディスプレイの表示を変化させる手法 [1][2] があるが、ユーザが特別な装置を装着する必要がある。

2つ目の問題は、小規模な広告になればなるほどユーザの関心を集めることが難しくなっている点である。2021年の広告市場では、プロモーションメディア広告費のうち、「屋外」に分類される大型で目立つデジタルサイネージなどは増加した一方で、「POP」や「イベント・展示・映像ほか」は減少した [3]。以上より、デジタル広告や屋外での大規模なデジタルサイネージと比べると、小規模なデジタルサイネージは費用対効果が小さい。このことから、小規模なデジタルサイネージの費用対効果を高めるために、コストを抑えたシステムが望まれる。

上記の問題点を解決するために、本研究ではユーザの位置計測に基づいた自然なインタラクション遷移を実現するデジタルサイネージシステムを提案する。本手法では、スクリーンとユーザとの距離に応じて情報提示方法を4つのフェーズに分けることにより、ユーザとシステムの暗黙的なインタラクションから明示的なインタラクションへの遷移を自然に行う。ユーザの頭部の位置やハンドポーズを用いた3Dコンテンツの操作により、前述した1つ目の問題

¹ 津田塾大学 学芸学部 情報科学科

である一方的な情報の提示を解決し、ユーザとシステムの自然なインタラクションを可能にする。これにより、情報表示に触れた際に、その場にいるからこそ得られる情報や体験をユーザに提供できるため、更なるユーザの関心を得ることが可能になると考える。また、特別なセンサやカメラではなく、大多数の人が保有している iPhone や iPad などの端末とプロジェクターを用いたシステムを構成する。広告を掲出する側が、3D コンテンツと画像さえあれば表示内容を容易に設定することを可能するために、iPhone や iPad などの端末で動く iOS アプリケーションを開発した。加えて、関連研究で使用されているステレオカメラ [4] や深度センサ [5]、測域センサ [6] を用いることなく、iPhone や iPad に搭載されているカメラを用いているため、システム全体のコストを抑えることが可能になる。

2. 提案システム

2.1 提案システムの概要

本研究では、スクリーンとユーザの距離に応じて 3D コンテンツ表示を変化させるシステムを提案する。また、ハンドジェスチャを用いた 3D コンテンツの操作や、bluetooth やメールを用いたユーザとの情報共有を行う。本システムは、以下のような 3 つの性質を持つ。

暗黙的なインタラクションから明示的なインタラクションへの誘導

ユーザがスクリーンに表示された情報を見る時、ユーザは徐々にスクリーンに近づくと考えられる。この際、本システムでは、スクリーンとの距離に応じて、画面に表示された 3D コンテンツが変化する。この変化によって、ユーザは自分の立ち位置に応じてスクリーンの表示が変化することを認識する。すなわち、ユーザは画面に表示された 3D コンテンツとユーザ自身の行動が関連づいていることに気づくことができる。その後、画面上に表示された情報について興味を持ったユーザは、スクリーンに近い場所まで移動すると考えられる。その際、ハンドポーズ認識により、ユーザは自分の意志で 3D コンテンツを操作することができる。

特別な端末が不要

3D コンテンツの操作は、ユーザのハンドポーズによって行われる。コンテンツに関する情報共有は、bluetooth やメールなどのツールを用いる。そのため、ユーザはスマートフォン 1 つを所持していれば良い。このように、ユーザは 3D コンテンツを操作するための特別な端末を必要としない。

スクリーンとユーザの距離に応じた情報表示

上述したように、本システムではスクリーンとユーザの距

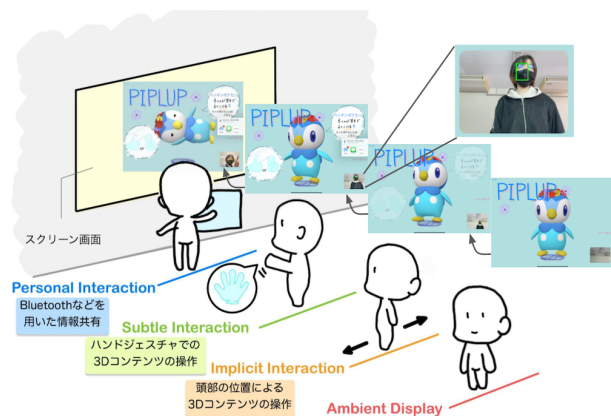


図 1: システム概要図

離に応じて画面に表示される情報が変化する。具体的には、ユーザがスクリーンに近づくほど表示される文字やイラスト情報が増えていく。また、ユーザの位置やハンドポーズより、表示される 3D コンテンツの回転やアニメーションを制御する。このように、スクリーンとユーザの距離に応じて、ユーザは表示されているコンテンツに即した情報を効果的に閲覧・操作することができる。

2.2 システム構成

図 1 に示すように、提案システムは、プロジェクタ 1 台、スクリーン 1 基、iPad Pro (12.9 インチ 第 5 世代) 1 台から構成されている。iPad では、情報をスクリーンに投影する前に 3D コンテンツの設定、文字やイラストの編集などを設定することができる。また、コンテンツ表示の際には、ユーザの顔認識による距離推定を行う。測定した距離に応じて、表示する情報を変化させることができる。また、ユーザは指定されたハンドポーズを行うことで、3D コンテンツの操作・閲覧をすることができる。加えて、bluetooth やメールを用いてユーザとの情報共有を行う。このように、iPad で処理された情報は、プロジェクターを通してスクリーンに投影される。

3. ユーザの行動誘導のためのフレームワーク

従来研究の一つである Vogel らにより提案された Interactive Public Ambient Displays[7] は、ディスプレイとユーザとの距離に応じてインタラクションを切り替える手法である。この手法を用いることで、暗黙的なインタラクションから明示的なインタラクションへの状態遷移を自然に行えるようになった。しかしながら、ディスプレイとユーザとの距離を測定するために、ユーザが赤外線マーカを装着しなければならない。また、ユーザのディスプレイに対する奥行き方向の動きと体の向きで表示内容を変更している。

これに対し本研究では、カメラ映像からの顔検出により得られた顔の大きさを用いて、ディスプレイとユーザとの距離を測定し、情報提示方法を切り替える。そして、ユー



図 2: Ambient Display フェーズでの情報表示の例

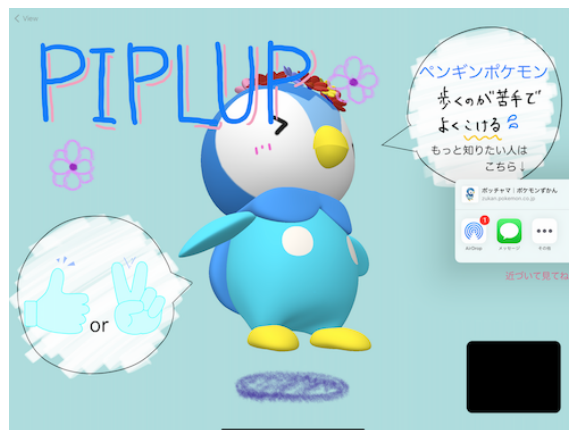


図 5: Subtle Interaction フェーズにおける操作方法の提示の例。スクリーン左側に指定ハンドポーズを表示する。

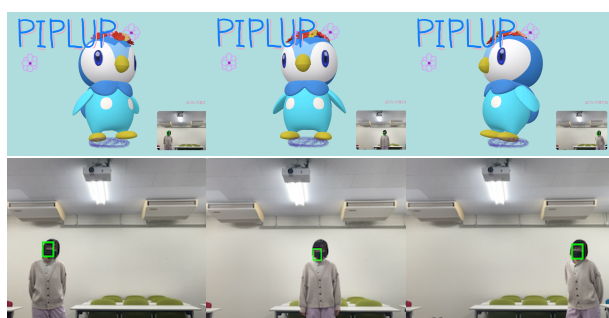


図 3: Implicit Interaction: ユーザの動きに伴う表示の 3D コンテンツの回転

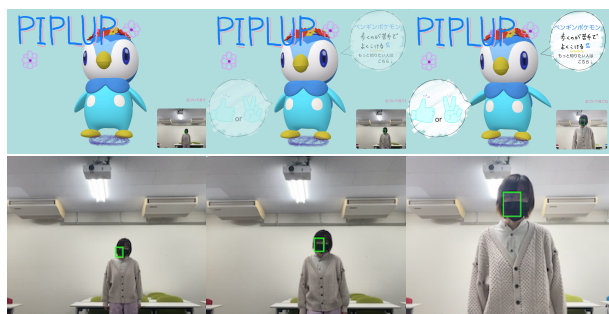


図 4: Implicit Interaction: ユーザの奥行き位置に伴う表示の不透明度の変化

ザのディスプレイに対する奥行き方向の動きだけでなく、ディスプレイに対する横方向の動きで表示内容を操作することができる。

3D コンテンツを用いた情報表示に関して、ユーザの関心を集めるためのフレームワークを構築する。本研究で提案するフレームワークは 4 つのフェーズに分かれている。4 つのフェーズとは、図 1 に示すように、Ambient Display, Implicit Interaction, Subtle Interaction, Personal Interaction である。以下、それぞれのフェーズについて説明する。

3.1 Ambient Interaction

Ambient Display のフェーズでは、ユーザーが一目見ただけで、全体の情報を把握できるような表示を行う。スクリーン上の表示は 3D コンテンツ、文字やイラストで構成されている。図 2 にその例を示す。

3.2 Implicit Interaction

Implicit Interaction のフェーズでは、ユーザがスクリーンの前を通り過ぎた時に、ユーザの頭部の位置に応じて 3D コンテンツの見える角度が変化をさせることで、ユーザとの暗黙的なインタラクションを行う。また、ユーザがスクリーンに近づくにつれて、コンテンツに関する詳細な説明を行うために提示する画像の不透明度を上げる。ユーザがスクリーンに対して横方向に動く際は、3D モデルがユーザの動きに即して回転する(図 3)。また、ユーザがスクリーンに対して奥行き方向に動く際は、コンテンツに関する詳細な情報が、ユーザがスクリーンに近づくにつれて表示される(図 4)。これにより、ユーザーをスクリーンの近くに引き寄せ、次のインタラクションフェーズに誘導することができる。

3.3 Subtle Interaction

Subtle Interaction のフェーズでは、ユーザはスクリーンに表示されている 3D コンテンツの操作やコンテンツに関する詳細な説明を閲覧するなど、システムと明示的なインタラクションを行う。このフェーズの前までは、ユーザは 3D コンテンツを暗黙的に操作していたが、このフェーズでは、ユーザはハンドポーズを用いて 3D コンテンツを明示的に操作することができる。

操作方法の提示は、フェーズが移行した時、ユーザが違うハンドポーズを行なっている時、フェーズの範囲内にいるがハンドポーズをしていない時に行う。また、フェーズが移行したことをユーザに提示するために、詳細な説明を不透明度を変化させることなく表示する。その様子を図 5

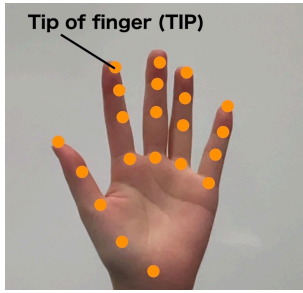


図 6: 手の骨格検出

に示す。

3.3.1 CoreML を用いたハンドポーズの認識

CoreML[8]とは、Appleのデバイス、iPhoneやMac上のアプリケーションで動く機械学習ライブラリである。デバイス上で機械学習を実行でき、作成したモデルを利用することができる。本研究では、このCoreMLを作成するにあたり、CreateMLを用いる。CreateML[9]とは、サンプルデータからパターンを認識し、モデルの訓練を行うシステムである。モデルが十分な精度になれば、CoreMLとしてアプリケーションに組み込むことができる。

本研究では、3Dコンテンツの回転と3Dコンテンツのアニメーションの切り替えをハンドポーズで操作する。3Dコンテンツの回転を操作するために、画像からパー及びそれ以外の手の形を識別する。識別するために用いたトレーニングデータ用の画像はそれぞれ235枚、バリデーションデータ用の画像はそれぞれ30枚、テストデータ用の画像はそれぞれ30枚とした。学習の結果、識別率が98%となった。

また、3Dコンテンツのアニメーションを切り替えるために、画像から、グッドとピースとそれ以外の手の形を識別する。識別するために用いたトレーニングデータ用の画像はそれぞれ250枚、バリデーションデータ用の画像はそれぞれ30枚、テストデータ用の画像はそれぞれ30枚とした。学習の結果、識別率が96%となった。この学習させたモデルを用いて、ハンドポーズの識別を行う。

3.3.2 ハンドポーズによる3Dコンテンツ操作

3Dコンテンツをユーザの意図した方向に回転させる処理を行う。まず、appleが提供するAPI[10]を用いてハンドトラッキングを行う。これにより、片手ずつ21つの骨格を検出できる。ユーザの手がパーと識別した時、図6におけるTip of finger(TIP)を追跡する。

本研究では、画面上のカーソルの位置によって3Dコンテンツの回転を制御している。カーソルは、ユーザがパーのハンドポーズを行なっている時のみ、手の骨格TIPに伴って動く。3Dコンテンツを回転させたい向きにカーソルを移動させると、カーソルの方向に回転し始める。図7

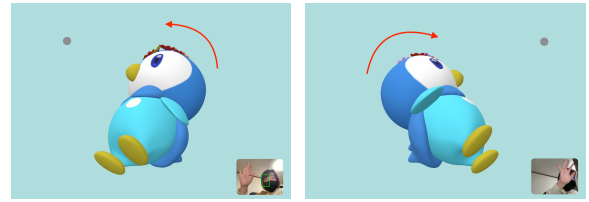


図 7: Subtle Interaction:TIP 追跡による 3D のモデルの左回転操作 (左) と 3D のモデルの右回転操作 (右)

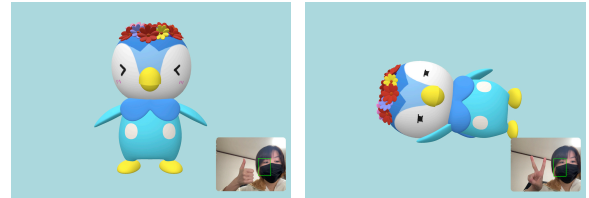


図 8: Subtle Interaction: ハンドポーズ識別によるグッドの時の 3D コンテンツのアニメーション (左) とピースの時の 3D コンテンツのアニメーション (右)

にその様子を示す。

3.3.3 ハンドポーズによる 3D コンテンツアニメーション変更

アニメーションを含む3Dコンテンツがある場合、ユーザのハンドポーズに伴い表示するアニメーションを切り替えることができる。ユーザのポーズや表示画像に即したアニメーションを表示することで、3Dコンテンツの特徴を提示することが可能となる。図8にその様子を示す。

3.4 Personal Interaction

Personal Interactionのフェーズでは、ユーザはbluetoothやメールなどを用いてコンテンツに沿った情報を得ることができる。ユーザが実際にiPadを操作して、ユーザのスマートフォンに欲しい情報を送信することができる。

3.5 フェーズの遷移

各インタラクションフェーズは、ユーザとカメラの距離によって切り替える。カメラとユーザの距離は、iPadのカメラ映像からの顔検出により得られる顔領域の大きさによって算出している。顔領域の面積が大きければユーザとカメラの距離は近く、顔領域の面積が小さければユーザとカメラの距離は遠いと考える。つまり、顔領域の面積によってシステムはフェーズを遷移させ、スクリーンに投影される情報表示を変化させる。

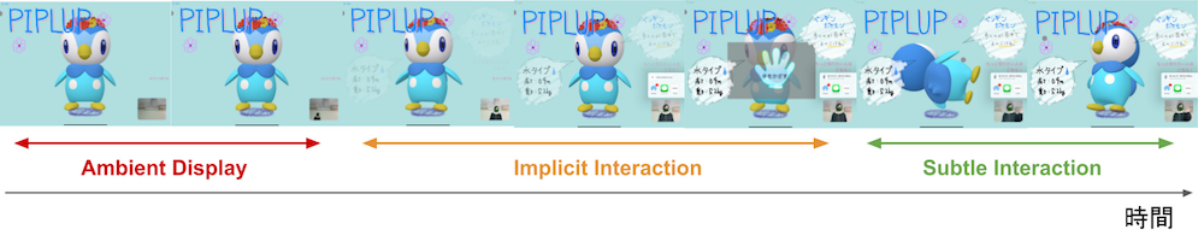
3.6 プロトタイプ

提案したフレームワークを利用して、2つのプロトタイプを開発した。ハンドポーズによる3Dモデルの回転制御(アニメーションなし)のプロトタイプ(図9(1))と、

ユーザ位置



スクリーン画面

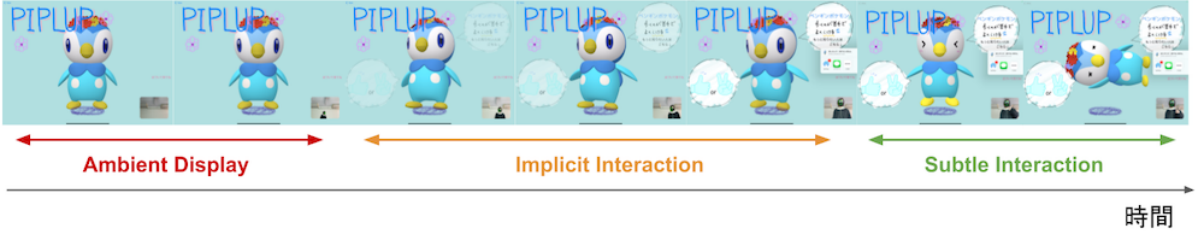


(1) ハンドポーズによる 3D モデルの回転制御 (アニメーションなし)

ユーザ位置



スクリーン画面



(2) ハンドポーズによる 3D モデルのアニメーションの切り替え (アニメーションあり)

図 9: 実験の様子

ハンドポーズによる 3D モデルのアニメーションの切り替え (アニメーションあり) のプロトタイプ (図 9(2)) である。プロトタイプ の 3D モデルは、3DCG ソフト Blender を用いて作成したポケモンを用いる。Ambient Display では、ポケモンの名前が表示がされる。Implicit Interaction では、ユーザがスクリーンに対して横方向に動く際は、3D モデルがユーザの動きに即して回転する。また、ユーザがカメラに対して奥行き方向に動く際は、コンテンツに関する詳細な情報が、ユーザがスクリーンに近づくにつれて表示される。Subtle Interaction は、ハンドポーズによる 3D モデルの回転制御 (アニメーションなし) のプロトタイプでは、ユーザが手をカメラに写さない場合や違う手の形を出している場合、ユーザがパーの手を出すよう誘導する表示を出す。ハンドポーズによる 3D モデルのアニメーションの切り替え (アニメーションあり) のプロトタイプでは、詳細な説明を表示する際に、ユーザがグッドとピースの手を出すよう誘導する表示を出す。Personal Interaction では、ポケモンずかん [11] が、bluetooth またはメールでユーザのスマートフォンに送信される。

4. 実験

ハンドポーズによる 3D モデルの回転制御 (アニメーションなし) のプロトタイプ (図 9(1)) と、ハンドポーズによる 3D モデルのアニメーションの切り替え (アニメーションあり) のプロトタイプ (図 9(2)) について、評価実験を行った。被験者の数は 18 人である。本実験では、被験者が、4 つのフェーズのうち、どのフェーズまで進むのかを観察した。また、2 つのプロトタイプに対して、それぞれ「どのくらい操作しやすかったか」、「どのくらいコンテンツに興味を湧いたか」について、5 段階に分けてアンケートをとった (大きい値の方が高評価)。これに加え、自由記述で感想を求めた。

4 つのフェーズのうちどのフェーズまで進むかを調べた結果を以下に説明する。Personal Interaction のフェーズまで進んだ被験者は 3 人、Subtle Interaction のフェーズまで進んだ被験者は 9 人、Implicit Interaction のフェーズまで進んだ被験者は 1 人、Ambient Display のフェーズのみであった被験者は 3 人となった。被験者のうち半分の 9 人は Subtle Interaction のフェーズまで、Personal Interaction のフェーズには 3 人移動した。広告に興味の薄

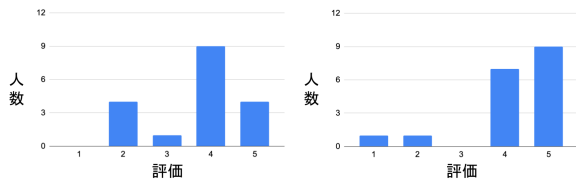


図 10: 「どのくらい操作しやすかったか」に対する評価 (アニメーションなしのプロトタイプ (左) とアニメーションありのプロトタイプ (右))

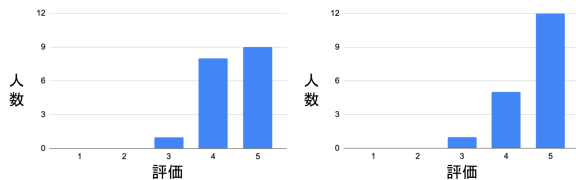


図 11: 「どのくらいコンテンツに興味を湧いたか」に対する評価 (アニメーションなしのプロトタイプ (左) とアニメーションありのプロトタイプ (右))

いユーザは、Ambient Display の状態にとどまると考えられるが、この結果にあるように、被験者の多くがスクリーンに近づいた。すなわち、提案システムによる広告コンテンツ表示に興味を持ったとみなすことができる。

「どのくらい操作しやすかったか」に関する評価値の分布を図 10, 「どのくらいコンテンツに興味を湧いたか」に関するものを図 11 に示す。ハンドポーズによる 3D モデルの回転制御 (アニメーションなし) (図 10 (左)) において、プロトタイプが操作しにくいと回答した被験者が多かった。自由記述において、カメラと手の距離感が掴みにくかったという意見が多かった。実際、右下にカメラ画像が配置されていても、ユーザの手がカメラが撮影できる範囲外に出てしまうことが多かった。しかし、終始、手がカメラが撮影できる範囲内にあるユーザは、手をどのように動かせば 3D コンテンツが回転するのか、説明しなくとも操作できていた。また、ハンドポーズによる 3D モデルの回転制御 (アニメーションなし) よりも、ハンドポーズによる 3D モデルのアニメーションの切り替え (アニメーションあり) のプロトタイプ (図 11 (右)) を体験してコンテンツに興味を湧いた被験者が多かった (評価値が高い傾向にあった)。自由記述の感想において、ポケモンの表情や動きを見て楽しかったという意見が多くあった。このことから、ユーザとシステムのインタラクション機能は、ユーザに表示情報 (コンテンツ) への興味を持たせることができると考える。

5. おわりに

本研究では、ユーザの位置計測に基づいた自然なインタラクション遷移を実現するデジタルサイネージシステムを提案した。スクリーンとユーザとの距離に応じて情報提示

方法を 4 つのフェーズに分けることにより、ユーザとシステムの暗黙的なインタラクションから明示的なインタラクションへの自然な誘導を可能にした。ユーザとのインタラクションを可能にすることで、従来システムの問題であった一方的な情報の提示を解決した。また、iPhone や iPad に搭載されているカメラを用いているため、システム全体のコストを抑えることが可能になる。

実験の結果から、提案システムは広告コンテンツ表示に対してユーザに興味を持たせる効果があると考えられる。

今後は、作成したプロトタイプがどのくらい人の興味を惹きつけられているのかについて、多数の歩行者がいる場所で検証していきたい。そして、カメラの撮影範囲を広くすることで、ユーザの手の認識範囲を広げ、システムの操作性を高めていくことについても検討していくつもりである。

参考文献

- [1] Maria Karam, Terry R. Payne, and Esther David. Evaluating bluscreen: Usability for intelligent pervasive displays. In The Second IEEE International Conference on Pervasive Computing and Applications(ICPCA '07), pp. 18–23, 2007.
- [2] Kento Miyaoku, Suguru Higashino, and Yoshinobu Tonomura. C-blink: a hue-difference-based light signal marker for large screen interaction via any mobile terminal. Proceedings of the 17th annual ACM symposium on User interface software and technology, pp. 147–156, 2004.
- [3] 電通 2021 年日本の広告費. <https://www.dentsu.co.jp/news/release/2022/0224-010496.html>
- [4] Yasuto Nakanishi, Yoichi Sato, and Hideki Koike. Enhanceddesk and enhancedwall: Augmented desk and wall interfaces with real-time tracking of user's motion. In Workshop on Collaboration with Interactive Walls and Tables, UbiComp 2002, pp. 27–30, 2002.
- [5] 深澤哲生, 福地健太郎, 小池英樹. 壁型ディスプレイを用いた非接触対話型電子広告システム. 第 14 回インタラクティブシステムとソフトウェアに関するワークショップ (WISS 2006) 論文集, pp. 65–70, 2006.
- [6] 佐藤竜也, 村田雄一, 志築文太郎, 田中二郎. 立つだけインタラクション-自然に使い方がわかる公共大画面向けインタラクション”. In WISS2008, 2008.
- [7] D. Vogel and R. Balakrishnan. Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. In Proceedings of the 17th annual ACM symposium on User interface software and technology(UIST2004), pp. 137–146, 2004.
- [8] Apple developer CoreML. <https://developer.apple.com/jp/machine-learning/core-ml/>
- [9] Apple developer CreateML. <https://developer.apple.com/jp/documentation/createml/>
- [10] WWDC20 videos detect body and hand pose with vision. <https://developer.apple.com/videos/play/wwdc2020/10653/>
- [11] ポケモンずかん ポッチャマ. <https://zukan.pokemon.co.jp/detail/393/>