

すくうインタフェース O-Key

田部 ちえみ[†] 平川 正人^{††}

[†] 島根大学大学院総合理工学研究科 ^{††} 島根大学総合理工学部

コンピュータは特別な道具から日用品になりつつある。しかし、コンピュータは多くのユーザにとっていまだ使いにくい道具のままである。この原因はコンピュータ操作のための動作が、人間の日常生活での動作とかけ離れていることにある。本論文では新しいインタフェース O-Key を提案する。O-Key ではユーザがモノを「すくいとる」動作によってコンピュータ操作を可能にしている。システムの基本構成と応用例を述べ、評価実験によって有効性を示す。

O-Key: A Scoopable Interface

Chiemi Tanabe[†] Masahito Hirakawa^{††}

[†] Interdisciplinary Graduate School of Science and Engineering, Shimane University

^{††} Interdisciplinary Faculty of Science and Engineering, Shimane University

Nowadays computers are considered a common tool for everyday activities, including document management and message communication. This in turn requests us to carefully design computers so that anyone can use them without any special skill. Augmented/mixed reality and tangible interface would be promising toward the development of such daily use interface systems.

In this paper, we propose a new interface which is named O-Key. O-Key allows the user with his/her hands to scoop up objects which are virtually floating on the water in a bucket. The user can enjoy catching objects for their selection and activation of associated computer application/commands. Applications of the interface system are presented, and an experiment showed that the proposed interface works effectively.

1 はじめに

コンピュータは特別な道具から日用品になりつつある。しかしながら、コンピュータは使いにくい道具であると感じている利用者はいまだ多い。なぜなら、テレビなど日常生活で利用される電器製品と比較して操作が難しく、使用方法を覚えるのに時間がかかるからである。この原因は、コンピュータ操作のための動作が、人間の日常生活での動作とかけ離れていることにある。

身近な例として、日常生活では「つまみを回して目覚まし時計をセットする」「スイッチを押してテレビを見る」など、操作対象を利用者自らの手で直接操作している。しかしながらコンピュータ上では、どちらもマウスやキーボードによる操作としてそれらが実現される。入力操作に汎用性をもたせた結果、モノのある場所へ移動して操作する必要はなくなったが、煩雑な手続きが必要となった。

この問題を解決し、コンピュータが本当の意味で日用品となるには、ユビキタス・コンピューティングに代表される、日常生活に溶け込む形でのコンピューティングが期待される。そのための試みとして、複合現実感やタンジブル・インタフェース¹⁾、ジェスチャインタフェースなど様々な提案がされている。

本研究では、手でモノをすくう動作を可能にするインタフェースシステム「O-Key (オーケイ)」を提案する。すくう動作は日常生活で人間がとる基本的な動作の一つである。この動作をコンピュータ操作に適用することで、直感的で親しみやすいインタラクションの実現を目指す。さらに、このシステムにおける有意義なアプリケーションを検討した。

以下、2章で提案するシステムの機能を述べ、3章ではすくうインタフェースのためのアプリケーションについての説明を行う。4章ではシステムの評価

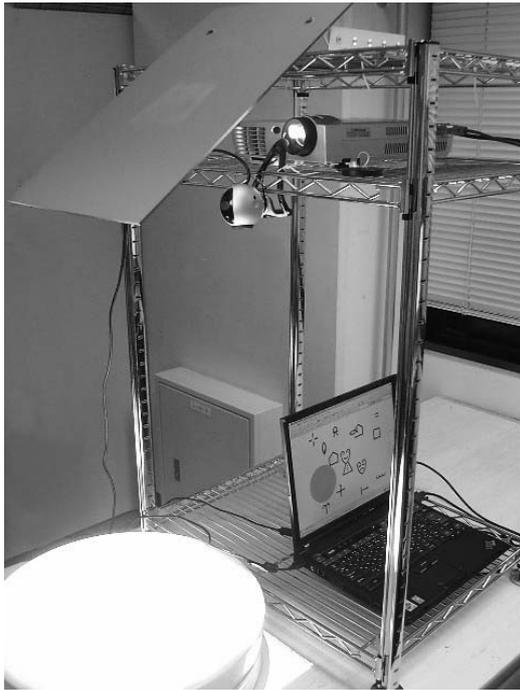


図1 システム外観

を行った結果とその考察を述べる。5章で関連研究について簡単に整理し、6章でまとめと今後の課題を述べる。

2 すくうインタフェース

2.1 基本アイデア

本システムは家庭や公共施設など、使いやすさが特に重要となる場面での利用を想定している。GUIでコンピュータを操作する際、利用者はディスプレイ上のアイコンの中から所望のものをマウスで指示・選択し、クリックして決定する場合がほとんどである。一方、日常生活の中でこれらの操作と同じ働きをする動作には様々なものがある。例えば、「指差す」「見つめる」「掴む」「すくう」などがある。

「指差す」、「見つめる」動作では所望の要素を一度に一つだけしか選択できないことが多い。一方、「掴む」では複数の要素を選択することができるが、選択範囲は手の大きさに制限される。これに対し、「すくう」では複数の要素を手の大きさに制約を受けず広範囲に選択することが容易である上、選択される要素は手の内側に位置するものであり、感覚的にも自明である。そこで本研究では選択する動作の中でも特にコンピュータの操作に適していると考えられる「すくう」動作に注目した。

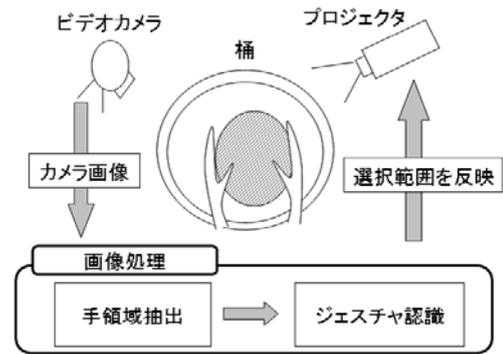


図2 処理の流れ

提案するシステムの外観を図1に示す。本システムは、PC、スクリーン代わりとして操作対象の要素が投影される桶、利用者の手の動きをキャプチャするためのビデオカメラ、プロジェクタ、反射鏡で構成されている。

対象要素が投影される桶には水が張られている。プロジェクタによって投影される要素は桶の底に映し出されるが、それらの要素はあたかも水面に浮かんでいるかのように見える。桶に張った水は視覚的な演出を行うのみではなく、水とのふれあいを通じた楽しさを提供する役割も担う。さらに、人間の生活に欠かせない物質である水をインタフェースに使用し、より日常生活になじむインタフェースシステムを実現できるとも考えた。

2.2 手領域の判定

ジェスチャ認識処理の流れを図2に示す。PCでは利用者の手の動きをカメラでキャプチャして画像処理を施し、どの要素がすくい取られたのかをリアルタイムで検出する。検出結果は桶に投影される映像に逐次反映される。

システムを実現するためのポイントは、すくう動作の開始タイミングと終了タイミングの判断である。このタイミングの判断基準を求めるため、予備実験を行った。実験に先立ち、カメラに映る手の面積によってすくう動作の開始と終了が判断できるのではないかと仮説を立てた。人間がすくう動作をする場合の手の動きをビデオカメラで撮影し観察を行った。

予備実験で明らかになった事項を以下にまとめる。また、手の領域が変化の様子を図3に示す。

- すくい始める時点では、利用者は手を水面に対してほぼ垂直に差し入れる。

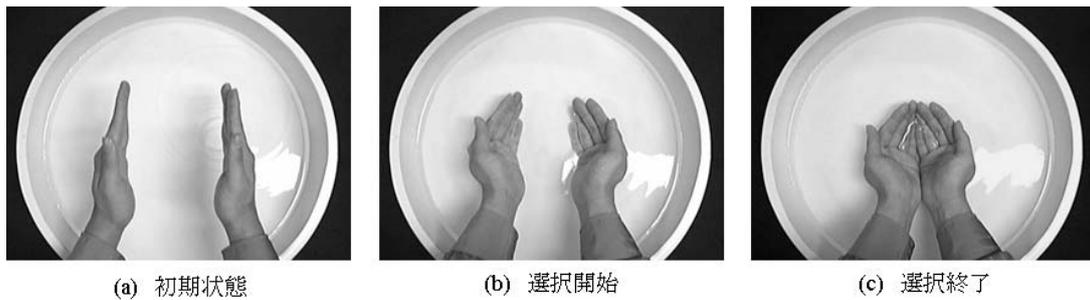


図3 すくうジェスチャ

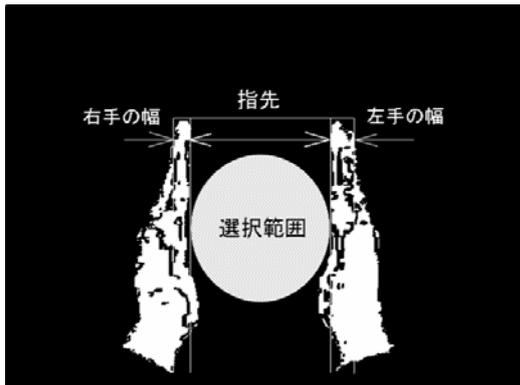


図4 手領域の判定

- すくい終わりに近づくと手が内側は上を向き、カメラに映る手の幅は広がる。
- すくい終わったところで両手が合わさり、一塊の領域となる。

以上の結果より、ビデオカメラに映る手の幅の変化から、すくい始めとすくい終わりを判定できることが明らかになった。そこでカメラ画像中の手の幅をジェスチャの判定基準として採用することとした。さらに、予備実験で撮影した映像から、すくい始めとすくい終わりの判定基準となる閾値を求め、後述するジェスチャ認識処理に用いた。

実際の処理では「手領域抽出処理」、「ジェスチャ認識処理」の2段階に分けて画像処理を行った。具体的な処理の流れを以下に述べる。

手領域抽出処理

1. 初期画像と現在の入力画像を比較して背景差分を取り、動物体領域を抽出する。
2. 動物体領域の画像をRGB表色系からHSV表色系に変換し、肌色領域（手の領域）を抽出する。

3. ラディアンフィルタを適用し、雑音処理を施す。
4. 画像中の右側と左側に存在するセグメントをそれぞれ右手、左手領域とする。
5. 両手の位置と幅および指先の位置を検出する（図4）。

ジェスチャ認識処理

1. 両方の手の幅が閾値を超えたら選択開始と判断する。選択範囲はその時点の両手内側の距離を直径とする円とする。選択領域をこのように定義した理由は、日常生活でのすくい動作と同じ感覚での操作を実現するためである。
この時、選択されている範囲を明示し、選択対象を利用者が確認できるようにしている。
2. 両手が合わさり両方の肌色領域が一つになった場合、選択終了とする。なお、この選択終了状態から再び手のひらが離れた動作をもって選択領域の確定を行い、選択された要素をアプリケーション側に受け渡す。
3. すくい始めと認識した以降で両手間の距離が拡大した場合あるいは手の幅が閾値以下になった場合には選択モードを中断し、改めて選択開始状態になるのを待つ。

ところで、一般的にシステムで利用者のジェスチャを獲得する場合、認識精度の高い磁気センサや赤外線カメラが利用されることも多いが、それらは装置が高価でO-Keyの利用シーンとして想定する生活現場での利用には向いていない。本研究では市販のUSBビデオカメラを使用し、安価にシステムを構築している。さらに、小さな子どもの手から大人の手まで対応可能であるよう配慮してシステムを構成している。



図5 きっかけツール



図6 発想支援システムへの応用例

3 すくうインタフェースを用いたアプリケーション

3.1 アイディア生成支援ツールの実装

筆者らは絵シンボルを用いて人間のアイデア生成支援を行う「きっかけツール」²⁾ (図5)を提案している。このツールは図5のように画面上に提示した絵シンボルあるいはそれらの組み合わせから利用者の忘れていたことを思い起こさせたり、または新たなひらめきを引き起こすきっかけを与える機能と、利用者がきっかけを感じた絵シンボルをメモしてお

く機能を持つ。きっかけ情報を提示する絵シンボルには多摩美術大学の太田によって考案された LoCoS 言語³⁾にて提供されているものを用いている。それぞれの絵シンボルは2次元ウィンドウ内を漂うように位置を移す。絵シンボルは移動しながら時として他と交差し、しかも背景透過になっているため、絵シンボルの重なり具合に応じて無限のきっかけシンボルの生成を可能にしている。

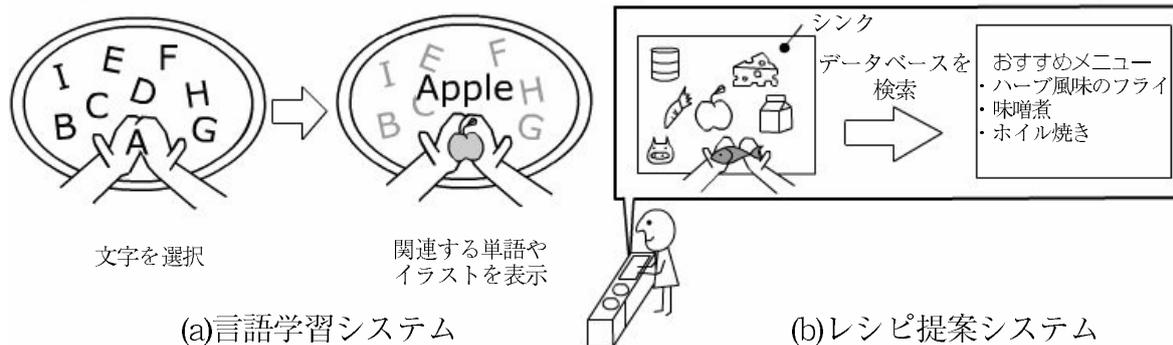
例えば図5のように、「帰りに時計屋に寄る必要があったな」とか「美容院を予約しよう」といった具合に、近隣にある複数の絵シンボルの集合でもって、あるひとつの意味が形成されてもよい。しかも一群として認識する範囲は固定されておらず、見る側の判断に委ねられているところに特徴がある。

このきっかけツールを本論文で提案するすくうインタフェースシステムに適用した。使用風景を図6に示す。利用者は発想のきっかけを感じたシンボル群を一塊として選択し、手桶領域に保存することができる。なお、取りこぼしがないように、すくい始めと判断した時点で手の内側にある絵シンボルオブジェクトの移動を停止させるようにしている。

3.2 他のシステム活用シナリオ

本節では、日常生活における O-Key の応用シナリオを示す。以下のように、O-Key は様々なシーンで利用可能である。

- **音楽演奏・作曲ツール:** メロディーを関連付けたオブジェクトをすくって音楽の演奏や作曲を行うシステム。利用者が所望のオブジェクトをすくって手桶に移すと、オブジェクトに関連づけられたメロディーが演奏される。複数のオブジェクトを手桶に移せば、それらのメロディーが組み合わせられ、利用者は簡単な作曲を行うことができる。
- **フォトアルバム:** 桶に写真を投影して一種のインテリアとして楽しんだり、同じシステムを持つ家族や友人と写真を交換しあったりできるシステム。具体的には、写真を手桶に移すと遠く離れて暮らすもの同士で写真を交換し、それらの写真をアルバムとして管理できる機能を持つ。
- **語学学習ツール:** 桶に投影されたアルファベットと触れ合いながら語学を学べるシステム。利用者が桶の中からすくい上げたアルファベットの映像は、利用者の手の中でそれを含む単語や単語に関連するイラストの映像に変化する。(図

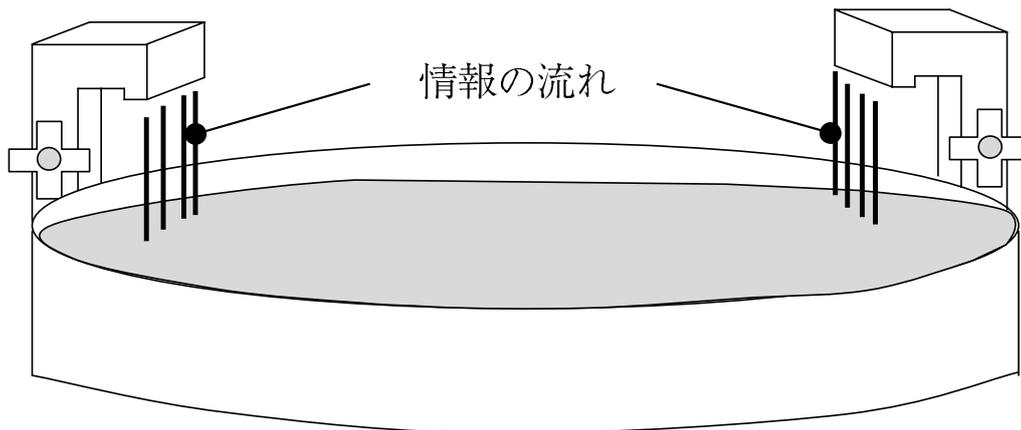


文字を選択

関連する単語や
イラストを表示

(a)言語学習システム

(b)レシピ提案システム



(c)蛇口をメタファとした例

図7 応用例

7(a)). 遊びのながら学習できる、子供向けのアプリケーションとして期待できる。

- **レシピ検索ツール:** 冷蔵庫内の食材リストを投影し、それらをすくって選択することで食材に見合ったレシピを検索し、利用者に提示するシステム。映像を投影する場所を桶ではなく台所の流し台に移せば、このようなシステムも実現可能である。(図7(b)).

以上に挙げた例では桶に張った水に動きをつけないことを想定しているが、桶の中に水の動きを作れば、より多様なシナリオを描くことができる。例えば、桶に蛇口を取り付けて「情報の流入口」のメタファとしたり、排水口を付けて「不要な情報を捨てる」メタファとすることが考えられる(図7(c)). 前述のアイデア生成支援ツールの場合、桶に蛇口を取り付け、ウェブ上のキーワードや画像など最新の情報の供給口として扱い、桶の中の既存の情報と組み合わせ新規のアイデア生成を促進することが期

待できる。さらに、複数の蛇口を取り付ければ異なる情報ソースからの情報供給を行うことも可能になる。そして排水口から桶の中の古い情報を捨てて桶の中の情報量をコントロールすることが考えられる。

加えて、手を使って対象をすくう(選択する)という本システムの特徴から、すくい上げる領域の中央部は重要で、周辺部になるほど重要度を低く捉えるなど、選択領域内の位置によって選択対象に優先度を付加するといったアナログ的な操作感覚の付加もシステム機能を強化する場合の可能性として挙げられる。この特徴を生かした応用例として、障害者向けのシステムが考えられる。障害者向けには、選択・非選択が明示的に区別される従来の対話技法と比べて、柔軟で実用度の高いシステムが実現されよう。

なお、O-Keyは屋内での適用を想定している。屋内の複数の場所で実現する場合、複数のプロジェクタやカメラを設置する必要があるが、プロジェクタについては、コンピュータ制御の反射鏡の利用によって必要とする機器の個数を減らす手法⁴⁾などの



図8 実験に用いた絵シンボル

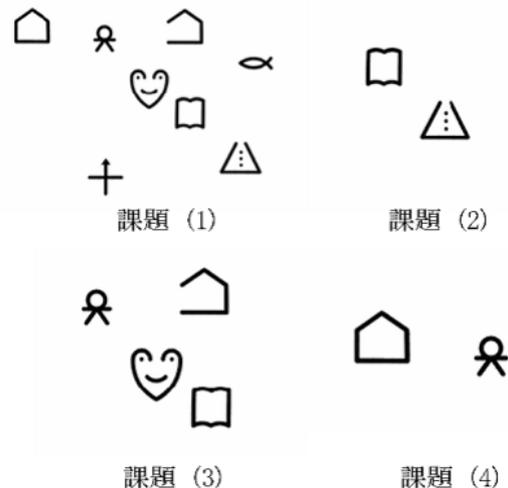


図9 選択対象とした絵シンボル

利用も考えられる。

4 システムの評価

4.1 評価実験

試作した O-Key のジェスチャ認識精度を確認するため、大学生 4 名を被験者とする評価実験を行った。実験には 3.1 節で説明したアイデア生成支援アプリケーションを使用し、水を張った直径 43cm の桶に絵シンボルを投影した。

実験方法を以下に示す。

- 水を張った桶に投影されている絵シンボル群 (図 8) の中から、指定した課題 (図 9) の絵シンボル群を選択する作業を被験者に依頼した。なお、絵シンボルの投影されている位置によって認識精度が異なるかどうかを調べるため、桶に表示する絵シンボルの位置は固定した。
- 利用者に選択動作の開始 (すくいはじめ) を声に出させ、その時点から実際に絵シンボルを選択し終わるまでの時間を測定する。
- 被験者はそれぞれの課題について 3 回ずつ選択作業を行う。被験者が取り組む課題の順序はランダムに決定した。
- 上記実験と併せて操作性についてのアンケートを行った。

4.2 結果と考察

それぞれの課題について、絵シンボルをすくい始めてからすくい終わるまでに要した時間の最大値、最小値、平均値を図 10 に示す。

実験結果において、課題 (1) は最も短い時間で選択でき、平均 6.3 秒で選択が完了した。反面、選択

に最も長い時間を要した課題 (4) の平均所要時間は 19.7 秒であった。このように課題によって選択時間に差が生じた原因は、選択範囲を両手内側の距離を直径とする円としていることにあると考えられる。課題 (4) の絵シンボルは桶に投影されている絵シンボルの端に位置しており、しかも 2 つの絵シンボルは水平方向に並んでいる。このような状態にある絵シンボルを、現在のシステムで選択しようとする、選択範囲が円で指定されるため、周囲の余計な絵シンボルまで選択されるという自体が発生する。このため絵シンボルを選択し直す動作が必要となり、選択完了までに時間を要することになった。この結果から、選択範囲を円だけでなく楕円等で指定できるようにし、より柔軟に選択範囲を指定できるアルゴリズムを実装することが課題として明らかになった。今回は投影されている絵シンボルを固定して実験を行ったが、絵シンボルを移動させた場合での評価も今後の課題である。

なお、操作性に関するアンケートでの被験者からのコメントには、「自然に使えて楽しいインタフェースである」、「すくうという動作でコンピュータを操作できる意外性があり、とても楽しい」などエンタテインメント性について高い評価が目立った。本システムが利用者にとって興味の持てる対話能力を備えていることがアンケートから明らかになった。

絵シンボル選択に必要な時間については、目的や応用領域が異なるため、マウス等の従来のポインティングデバイスを用いた場合と単純な比較はでき

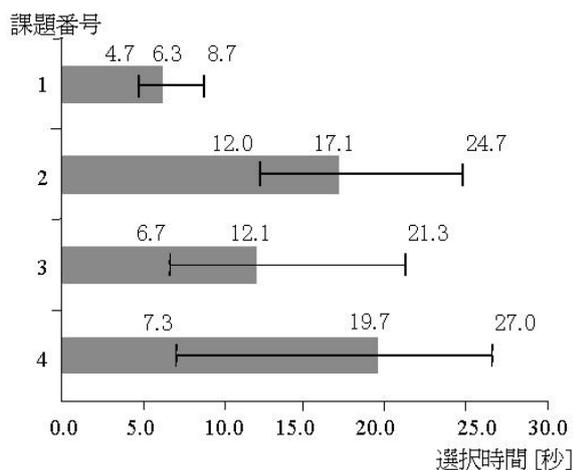


図 10 実験結果

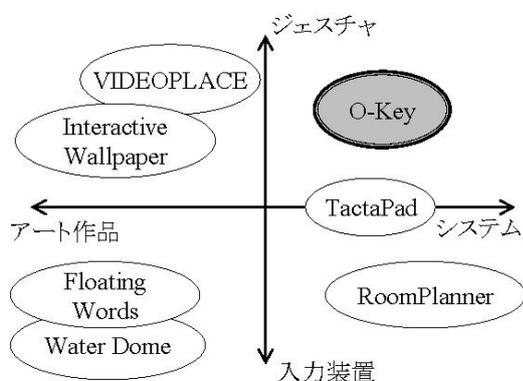


図 11 本研究の位置づけ

ないが、ジェスチャ認識の精度を高めることで、より高速かつスムーズに対象物を選択できるようになり、応用によっては十分に利用可能であると判断される。特別の努力をすることなく誰もがコンピュータを使える、直感的で自然なインタラクションのひとつの形を本システムの試作で示すことができたと言えよう。

5 関連研究

両手を使ったインタラクション手法については既に数々の試みがなされている。

テーブル上での手の動きを用いてオブジェクトを操作するアプリケーションシステムの一例として、RoomPlanner⁵⁾がある。これはタッチパネルを内蔵したテーブルに投影された家具のイメージを、両手のジェスチャによってテーブル上の仮想の部屋に配置するシステムである。

TactaPad⁶⁾は、タッチパネルとカメラを組み合わ

せたシステムであり、あたかも PC のデスクトップを直接手で操作しているかのような感覚にさせる入力装置である。TactaPad⁶⁾ではタッチパネルに触れることで PC 操作を行うが、併せてタッチパネル上部に取り付けられたカメラで利用者の手の形をキャプチャし、それがデスクトップ上に表示される。

RoomPlanner⁵⁾及び TactaPad⁶⁾は共にタッチパネルを用いており、操作領域の大きさに制約がある。一方、O-Key はカメラで捉えられる範囲ならば、どのような大きさの投影領域でも対応可能である点で優れている。

人間の動きをカメラで捉えてインタラクションを図る研究の例としては、VIDEO PLACE⁷⁾、InteractiveWallpaper⁸⁾がある。VIDEO PLACE⁷⁾は利用者のシルエットをディスプレイに表示し、遠隔地の相手とコミュニケーションをとることを目指したシステムである。それでは対象物の移動・拡大・縮小を主たる機能としているが、本研究は対象物を選択するためのインタフェースとして考案している。

一方、周囲の環境や人間の動きの変化によって壁面に映し出される模様（壁紙）が変化するインタラクティブ・アート作品として作製された Interactive-Wallpaper⁸⁾は、日常生活に溶け込むコンピュータの形として壁紙を利用している。実世界のモノにコンピュータからの映像を投影し、人間の動きによって変化を与えるという点で本研究と類似しているが、これはアート作品として作製されたものであり、本研究のシステムとは方向性が異なる。

本研究のように、すくう動作に注目したインタラクティブ・アート作品として、The FloatingWords⁹⁾がある。観客が作品に備え付けられたマイクに向かって発した言葉が音声認識され、水を張った小さなプールにその言葉が水滴のように落ちる。さらに、ひしゃくで水中の文字を突いたりかき混ぜたりすることで文字に変化を与えることができる。この作品では、ひしゃくという特別の器具を用いてのみコンピュータへ指示を与えられるようになっており、素手で操作できる本提案と異なっている。

本研究と類似したインタラクティブ・アート作品に Unconscious Flow¹⁰⁾がある。観客 2 名は心電計を身につけ、水を張った桶をはきんで向かい合う。水には観客の分身である人魚達が投影されている。観客はこの人魚に触れることによってノンバーバルコミュニケーションを体験する。水を張った桶にオブジェクトを投影し、そのオブジェクトとユーザが

触れ合うことでインタラクションを試みる点が本研究と類似している。この作品では人魚を通して人間同士のコミュニケーションを補助しており、システムに他のアプリケーションを搭載することは考慮されていない。これに対して、本システムは日常生活に溶け込むインタフェースシステムを提供することが目的であり、実装するアプリケーションによって様々な場面に対応することが可能である。

Water Dome¹¹⁾ はスクリーンに水を用いたディスプレイである。利用者の頭上にあるノズルから噴水のように水を流し、その水でできた膜に映像を投影する。しかしながら Water Dome¹¹⁾ においては、水と触れ合うことによるインタラクションは実現されていない。

以上に挙げた関連研究と O-Key の相関を、対話に用いる手段ならびにシステムが目指す方向の 2 軸から整理すると図 11 のようになる。

6 まとめ

人間がモノをすくう動作をヒューマン・コンピュータ・インタラクションに用いたインタフェースシステム「O-Key」とそのアプリケーションの提案・試作を行った。水面に投影したオブジェクトに対するユーザの操作として、すくうという行為を可能にした。このシステム上で機能するアプリケーションとして、筆者らが先に提案したきっかけツールを実装した。これは LoCoS 言語で用いられている絵シンボルを使って、利用者に何らかの発想を引き出すきっかけを提供するアプリケーションである。システムのジェスチャ認識精度を確かめるための評価実験を行い、本システムがユーザとの対話能力を備えていることを確認した。

現在のところ O-Key のアプリケーションは一人のユーザによる使用を想定しているが、大きな桶を用意すれば複数の利用者が桶を囲んで一緒に作業することも可能である。複数ユーザが利用できる O-Key では、単なるインタフェースシステムとしてだけでなく、コミュニケーションツールとしての機能も持ち合わせたアプリケーションが実現可能である。

さらに、現在は桶の上に投影されているオブジェクトはすべて仮想の物体であるが、対話に用いる要素は現実のモノでも構わない。現実のモノを物理的アイコン (Phicons)¹⁾ として仮想オブジェクトと共存させることで、より人間の自然な営みに近いイン

タフェースを構築したい。今後は「すくう」動作が有効に機能するアプリケーションの実装ならびにジェスチャ認識精度向上、およびシステム機能拡張が望まれる。

参考文献

- 1) Ishii, H. and Ullmer, B.: Tangible bits: Towards seamless interfaces between people, bits and atoms, *Proc. of SIGCHI Conference on Human Factors in Computing Systems*, pp. 234–241 (1997).
- 2) Tanabe, C. and Hirakawa, M.: A LoCoS-based tool capable of providing a clue for idea generation, *Proc. of 2005 International Conference on Active Media Technology*, pp. 51–56 (2005).
- 3) 太田幸夫: 新しい絵ことば LoCoS(ロコス), 講談社 (1973).
- 4) Pingali, G. et al.: Steerable interfaces for pervasive computing spaces, *Proc. of IEEE International Conference on Pervasive Computing and Communications*, pp. 315–322 (2000).
- 5) Wu, M. and Balakrishnan, R.: Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays, *Proc. of 16th ACM Symposium on User Interface Software and Technology*, pp. 193–202 (2003).
- 6) Tactiva: TactaPad, <http://www.tactiva.com/>.
- 7) Krueger, M. W., Gionfriddo, T. and Hinrichsen, K.: VIDEOPLACE artificial reality, *Proc. of SIGCHI Conference on Human Factors in Computing Systems*, pp. 35–40 (1985).
- 8) Huang, J. and Waldvogel, M.: Interactive Wallpaper, *Proc. of ACM SIGGRAPH 05 Electronic Art and Animation Catalog*, pp. 172–176 (2005).
- 9) Moroi, S.: The Floating words, ACM SIGGRAPH2001 Art Gallery: N-SPACE, 2001.
- 10) Tosa, N. and Nakatsu, R.: Interactive art for zen: 'Unconscious flow', *Proc. of IEEE International Conference on Information Visualization*, pp. 535–541 (2000).
- 11) Sugihara, Y. and Tachi, S.: Water dome - An augmented environment-, *Proc. of IEEE International Conference on Information Visualization*, pp. 548–553 (2000).