

身体動作を用いて商品サイズを絞込む 検索インタフェースの設計と評価

大江 龍人[†] 蕪澤 賢三[†] 岩淵 志学[†]
益子 宗^{††} 田中 二郎[†]

既存の e-コマースにおいては、テキストを用いた商品検索が行われる。テキストを用いてテレビや家具などを検索する場合、ユーザは自身が購入したい商品のサイズを予め把握する必要がある、商品サイズをリアルタイムに把握・実感して絞込むことは困難であった。この問題を解決するために、我々は身体動作を用いて商品サイズを絞込む検索インタフェースである「これくら e-コマース」を開発してきた¹⁾。これくら e-コマースを用いることにより、ユーザは自身の身体動作を用いたサイズ表現が可能になり、絞込まれた商品をリアルタイムに閲覧しながら商品検索を行うことが可能になる。本論文ではこれくら e-コマースのインタラクション手法とその実装を示す。また、これくら e-コマースの認識精度の評価と、提示させる視覚フィードバックの評価についてそれぞれ述べる。

Design and Evaluation of The Search Interface for Product Specifications using Hand Gestures

TATSUHITO OE,[†] KENZO NIRASAWA,[†] SHIGAKU IWABUCHI,[†]
SOH MASUKO^{††} and JIRO TANAKA[†]

In traditional e-commerce, we search commodities using a text-based search interface. When a user search TVs, furnitures, etc. using a text, he needs to know item's size that he wants to buy. So it is difficult for a user to search commodities with grasping it's size in real-time. To resolve this problem, we have developed "Korekura e-commerce", that is the search interface using hand gestures. Korekura e-commerce enables a user to describe a size using hand gesture and search commodities while browsing filtered commodities in real-time. In this paper, we show the interaction technique and the implementation of Korekura e-commerce. Moreover, we describe the evaluation of the recognition accuracy and the visual feedback.

1. はじめに

既存の e-コマースにおいては、テキストを用いた商品検索が行われる。テキストを用いた検索においてテレビや家具などを検索する際、ユーザは「商品名 サイズ」と検索クエリを作成し検索を行う場合がある。

例えば、テレビを検索する場合ユーザは「テレビ ○○インチ」、筆筒を検索する場合「筆筒 ○○mm×○○mm×○○mm」と検索クエリを作成する。このように検索クエリを作成して検索を行う場合、ユーザは購入したい商品のサイズを予め把握する必要がある、商品サイズをリアルタイムに把握・実感して絞

むことは困難であった。

商品サイズを直感的に把握・実感した絞込みを実現するために、我々は身体動作を用いて商品サイズを絞込む検索インタフェースである「これくら e-コマース」を開発してきた¹⁾。これくら e-コマースを用いることにより、ユーザは自身の身体動作を用いたサイズ表現が可能になり、さらに絞込まれた商品をリアルタイムに閲覧しながら商品検索を行うことが可能になる。

本研究では、身体動作を用いてサイズを表現するジェスチャを「これくらいジェスチャ」と定義する。これくらいジェスチャとは、サイズを表現する場合に一般的に用いられる両手による身体ジェスチャである。例えば、「これくらいの幅」、「これくらいの高さ」、「これくらいの長さ」などを表現する場合に用いられる。これくらいジェスチャを商品の絞込みに応用することにより、商品のサイズを実感した絞込みを行うことが可能になると我々は考えた。

[†] 筑波大学 大学院システム情報工学研究科

Graduate School of Systems and Information Engineering, University of Tsukuba

^{††} 楽天株式会社 楽天技術研究所

Rakuten Institute of Technology, Rakuten, Inc.

2. 関連研究

身体動作を用いて直感的に検索を行う研究がこれまでになされてきた。Holz ら²⁾ は、ジェスチャにより 3 次元形状のオブジェクト検索を行うインタフェースを示した。また、栗原³⁾ は Kinect を用いて物探しを支援するシステムを示した。Holz らの研究は 3 次元物体の形状を検索対象とし、また、栗原の研究は実世界の物を検索対象とする。これらに対して我々の研究は e コマースにおいて扱われている商品を検索対象とする。

また、拡張現実 (AR) 技術を用いて購買を支援するシステムの開発がなされている⁴⁾⁵⁾⁶⁾。これらのシステムはユーザに商品を重畳表示させることにより、実際の商品を実感させることを目的としている。我々の研究においても、これくらいジェスチャを行った際に、検索された商品がユーザの上に重畳表示される。我々の研究はこの重畳表示と、これくらいジェスチャを用いたサイズ指定とを合わせて、ユーザが商品を実感する経験をさらに強化させる。

3. これくら e コマース

本節ではこれくら e コマースのインタラクション手法について、実際にユーザが商品を検索する流れに即して示す。さらに、これくら e コマースのシステム構成についても示す。

3.1 インタラクション手法

これくら e コマースを用いた商品検索は、(1) 検索する商品ジャンルの指定、(2) 指定したジャンルにおける商品の絞り込み、(3) 商品の選択、(4) 商品の販売ページ閲覧、の 4 つの手順から成る。以下 4 つの手順をそれぞれ述べる。

- (1) 図 1 に示すように、検索する商品のジャンルを指定する。図 1 左には扇風機、図 1 右にはテレビのアイコンがジャンルの例として表示されており、ユーザは手をアイコン上に位置させて検索したい商品のジャンルを指定する。ユーザの両手には視覚フィードバックとして赤いカーソルが表示されており、ユーザはこのカーソルを視認しながらジャンルの指定を行う。
- (2) 図 2 に示すように、これくらいジェスチャを用いて商品のサイズを指定し絞り込みを行う。広げた両手の内側には、両手間の距離から求めたサイズ (ここではテレビのインチ) に該当する商品が表示される。商品の詳細をさらに閲覧したい場合、ユーザは両手を数秒間静止させシステムの状態を



図 1 検索する商品のジャンルを指定する。
Fig.1 Select an item's genre to search.

(3) に遷移させる。



図 2 これくらいジェスチャを用いて商品のサイズを指定し絞り込みを行う。
Fig.2 Filter commodities by specifying the size using Korekurai gesture.

- (3) 手の位置を変更させて図 3 に示すように商品を選択する。図 3 において、ユーザは手を画面左右に位置させることにより、商品の画像を遷移させ目的の商品を選択する。また、手の位置を画面中央に位置させ数秒間静止させることにより、商品が選択されその販売ページが表示される。



図 3 絞込まれた商品から選択する。
Fig.3 Select from filtered commodities.

- (4) web ブラウザを用いて選択された商品の販売ページを閲覧する。

3.2 システム構成

図 4 に示すように、これくら e コマースは計算機、ディスプレイ、距離カメラから構成される。距離カメラは手のジェスチャを検出するために用いられる。また、ディスプレイは絞込んだ商品を表示させるために用いられる。

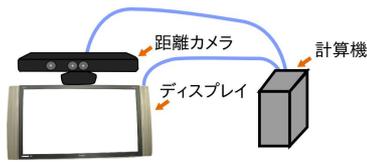


図 4 これくら e-コマースのシステム構成.
Fig. 4 System configuration of Korekura e-commerce.

4. これくら e-コマースの実装

本節ではこれくらいジェスチャの実装と、商品検索部の実装についてそれぞれ述べる。

4.1 これくらいジェスチャの実装

本研究ではこれくらいジェスチャを認識するための距離カメラとして Microsoft 社の Kinect[☆]を用い、Kinect から得られる距離画像などのデータを処理するために Microsoft 社の Kinect for Windows SDK beta^{☆☆}を用いた。

本システムでは、ユーザが実世界において広げた手の間隔を商品の絞込みに用いるため、実世界における両手間の距離を計算する必要がある。我々はこれを実現するために、(1) ユーザの手と Kinect の間の距離 (深度情報) をミリメートル単位に変換、(2) 手順 (1) で変換した距離と Kinect の画角から、1 ピクセルあたりのミリメートルを算出し両手間の距離を求める、(3) 両手間の距離単位をミリメートルから求めるサイズの単位に変換する、の 3 つの手順を実装した。以下、3 つの手順についてそれぞれ述べる。なお、今回は主にテレビの検索を行うことを想定する。

- (1) ユーザの手と Kinect の間の距離をミリメートル単位に変換する。変換された値を d とする。
- (2) d と Kinect の画角を用いてカメラ画像上の 1 ピクセルあたりの mm を計算する。Kinect の画角 (図 5 の θ の 2 倍) は 57° であり、Kinect からユーザの手までの距離における、カメラ画像に写る mm 単位の横幅 (以下 a とする) は、 $a = 2 \times d \times \tan(28.5 \div 180 \times \pi)$ から算出できる。また、縦幅も同様に算出可能である。Kinect からユーザの手までの距離とカメラ画像に映る横幅の関係を図 5 に示す。この結果を画素数で割ることにより、1 ピクセルあたりの mm を算出する。縦幅に関しても同様に mm を算出することによ

り、両手間の距離が求められる。

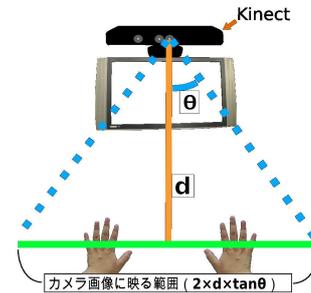


図 5 カメラ画像に写る範囲の計算.
Fig. 5 Calculate a range taken by the camera.

- (3) 手順 (2) にて得られたユーザの両手間の距離 (mm) を 25.4 で割り、インチに変換する。

4.2 商品検索部の実装

これくらいジェスチャにより絞込まれた商品は全て楽天市場^{***}の商品である。我々は、これくら e-コマースにおいて楽天市場の商品画像をリアルタイムに表示させるために以下の実装を行った。

- 楽天商品検索 API^{☆4}を用いて、ジャンル毎にあらかじめ商品のサムネイル画像と商品 URL を取得し、ローカルに保存する。
- 取得した商品のサムネイル画像をリアルタイムに表示させる。

楽天商品検索 API には、商品名とそのサイズをキーにして検索クエリを発行する。例えばテレビの場合は「テレビ ○○インチ」と検索クエリを生成する。この○○インチは 10 インチから 60 インチに設定され、あらかじめ 1 インチずつ検索が行われる。

5. これくらいジェスチャの認識精度に関する評価実験

ユーザがこれくらいジェスチャを行った際の、システムによる大きさの認識精度を検証するため、評価実験を行った。

5.1 実験内容

本実験では、以下の 3 つの項目を評価した。

- (1) 認識精度はどの程度か。
- (2) ユーザと Kinect の距離によって認識精度に差が出るか。
- (3) 実際にサイズを用いて商品検索を行う際にどの程度の粒度で検索が可能であるか。

5.1.1 実験方法

実験の手順を以下に示す。

- (1) 被験者は指定した長さのメジャーに合わせて手を広げ (図 7a)、その大きさを保持したままディ

☆ <http://www.xbox.com/kinect/>

☆☆ <http://research.microsoft.com/en-us/um/redmond/projects/kinectsdk/>

☆☆☆ <http://www.rakuten.co.jp/>

☆4 <http://webservice.rakuten.co.jp/api/itemsearch/>

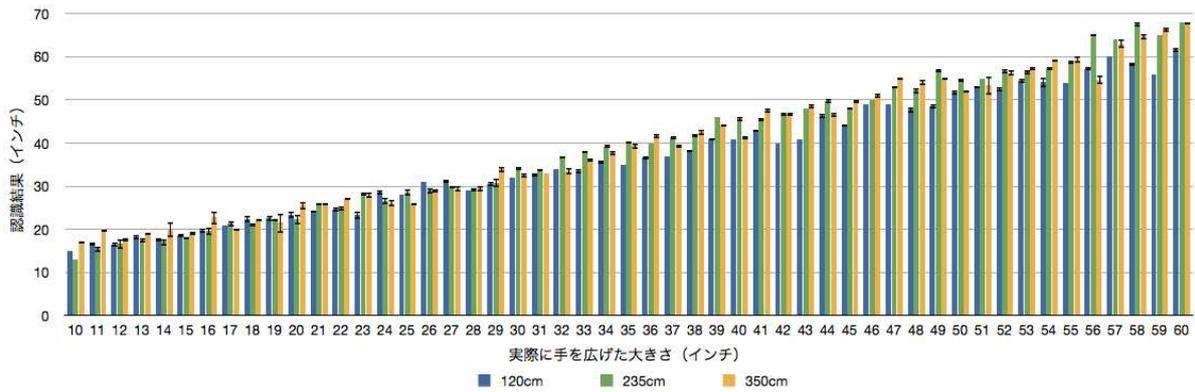


図 6 認識精度に関する評価実験の結果.

Fig. 6 The result of the evaluation about the recognition accuracy.

スプレイの正面に立つ (図 7b).

- (2) 被験者は 1.5 秒間 (15 フレーム分) 手を静止させ、システムはその間に認識されたインチ数を毎フレーム記録する。

以上の手順を 10 インチから 60 インチまでそれぞれ 1 インチずつ行った。この一連の手順を、被験者と Kinect の距離が 120cm, 235cm, 350cm の場合についてそれぞれ行った。これらの距離は、システムが被験者の両手を安定して取得可能な最短距離と最長距離、およびその中間の距離である。なお、筆者 1 名を被験者として実験を行った。

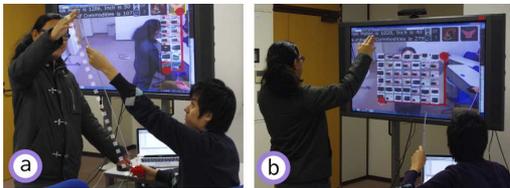


図 7 認識精度の評価実験を行っている様子。a) メジャーに合わせて手を広げる。b) 手を広げたままシステムの正面に立つ。

Fig. 7 Situation of the evaluation about the recognition accuracy. a) the user spread his arms to fit the length of a measuring tape. b) he stands in front of the system with holding arms.

5.2 実験結果

実験結果のグラフを図 6 に示す。グラフの横軸は被験者が提示した両手間の大きさであり、縦軸はシステムが認識した結果 (15 フレームの平均値) である。なお、エラーバーは標準偏差を示す。10 インチから 60 インチそれぞれの試行について、システムが認識したインチ数 (15 フレームの平均値の小数点を切り捨てた値) と実際に手を広げた大きさが一致する確率は、120cm では約 17.6%, 235cm では約 5.9%, 350cm では約 2.0%であった。

また、距離毎による実際の大きさと認識された大きさの差の平均値を図 8 に示す。なお、エラーバーは標準偏差を示す。120cm では差が平均値 1.87 ± 2.01 インチ、235cm では平均値 4.47 ± 1.70 インチ、350cm では平均値 4.40 ± 2.07 インチとなった。

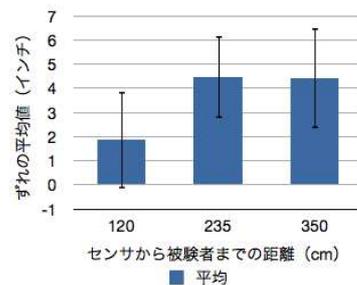


図 8 実際の大きさと認識された大きさの差。

Fig. 8 A gap between the actual size and the recognized size.

5.3 考察

本節では、認識精度、ユーザと Kinect の距離による違い、認識可能な大きさの粒度についてそれぞれ実験結果を受けた考察を示す。

5.3.1 認識精度について

実験から、実際のサイズと認識されたサイズの一致確率が低いことがわかった。これには以下の要因が考えられる。

Kinect による手の検出の問題 Kinect による手の認識は結果がぶれることが多く、指先が認識されることや手首付近が認識されることがある。図 8 のグラフから、実際に手を広げた大きさと認識結果の誤差は最大 7 インチ、つまり 177.8mm である。被験者の指先から手首までの長さは 190mm であったため、この誤差は手の認識結果によって現れ得るものである。

Kinect を取り付ける角度の問題 実験では Kinect は大型ディスプレイの上部に取り付けられており、被験者を見下ろす角度であった。しかし我々の実装では、Kinect とユーザは平行となることを前提としている。そのため、両手間の距離が正確に計測出来なかった可能性がある。

5.3.2 ユーザと Kinect の距離による認識精度の変化について

図 8 のグラフから、Kinect と被験者の距離が 235cm と 350cm のときの認識結果に大きな差がなく、距離を変更しても認識結果が大きく変わることは無いことがわかる。しかし、120cm の場合の認識の誤差は 235cm と 350cm の場合のおよそ半分という結果になった。この結果には実験方法が少なからず影響している。120cm の場合被験者はディスプレイに近く、自然と Kinect に向けて上向きにジェスチャを行っていた。この結果として Kinect と被験者の手が平行な位置関係となった。235cm と 350cm の場合被験者はディスプレイ正面に向けてジェスチャを行っていたため、誤差が若干大きくなったと考える。

5.3.3 認識可能な大きさの粒度について

実験結果から、6~7 インチ刻みで検索するシステムとすれば、十分正確に検索が可能であると考えられる。

ただし、テレビのサイズには商品数が多いサイズ(以下ボリュームゾーン)がある。実際に本システムを用いて検索したところ、40 インチのテレビが約 2800 件見つかったのに対し、31 インチのテレビは 1 件も見つからなかった。そのため、単純に一定のインチ間隔を用いるのではなく、ボリュームゾーンを考慮して検索に適切なインチ間隔を調査することが今後の課題となる。

6. これくらいジェスチャの視覚フィードバックに関する評価実験

本研究では商品のサイズを直感的に把握・実感した絞込みの実現を目的としており、そのためにはユーザに提示する視覚フィードバックが重要になると我々は考える。視覚フィードバックの提示方法は、ユーザが商品サイズを把握する程度に影響を与え、さらに指定したサイズにおける商品個数の把握、商品の多様性の把握にも影響する。

我々はこれくらいジェスチャにおける適切な視覚フィードバックを調査するため、図 9 に示す 5 つの視覚フィードバックの評価を行った。この視覚フィードバックは図 9 左上から、a) 認識された大きさと検索された商品数を文字を用いて提示、b) 検索された

商品のサムネイル画像をアプリケーション画面一面に提示、c) 広げた手の枠内に商品画像を引き伸ばして提示、d) 広げた手の枠内に商品のサムネイル画像を提示、e) c の提示結果の上に d を提示、となる。

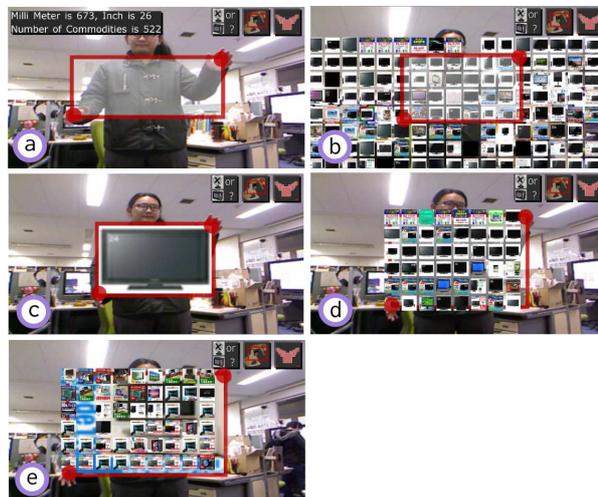


図 9 評価を行った 5 つの視覚フィードバック。
Fig.9 Five visual feedbacks for evaluation.

6.1 実験内容

実験タスクは、5 つの視覚フィードバックにおけるこれくらいジェスチャの実行と、アンケートへの回答からなる。まず被験者は、図 9a~e の 5 つの視覚フィードバックごとにこれくらいジェスチャを行う。5 つのタスクの順番はバランスを取るため、被験者ごとそれぞれランダムに設定した。タスク終了後、被験者は以下に示す 4 つのアンケート項目を 5 段階のリッカート尺度により回答した。

- (1) 商品サイズは把握しやすいものであったか。
- (2) 商品の個数は把握しやすいものであったか。
- (3) 商品の多様性は把握しやすいものであったか。
- (4) 視覚フィードバックは見やすいものであったか。

被験者はまた、a~e いずれの視覚フィードバックが最も好ましいかについて、その理由と共に回答した。この実験は被験者 5 人に対して行った。なお、被験者は 21~24 歳の大学生、大学院生 5 名であった。

6.2 実験結果

被験者が 5 段階のリッカート尺度により回答したアンケートの結果を図 10 に示す。このアンケート結果は、点数が高いほど評価が高いものとなる。

被験者が選んだ最も好ましい視覚フィードバックは、5 人中 3 人が e、2 人が d となった。e が好ましいと答えた理由は以下のものであった。

- 検索にヒットした商品全体が大体見られた。

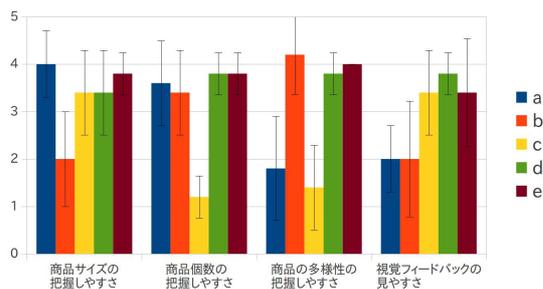


図 10 各視覚フィードバックごとのリッカート尺度による評価。
Fig.10 The evaluation of each visual feedbacks with Likert scale.

- 商品の多様性が分かりやすい。
- そのサイズの商品の特徴を把握するため c や e のように引き伸ばされた画像が欲しい。

また、d が好ましいと答えた理由は以下のものであった。

- e の提示方法は引き伸ばされた画像が隠れて見えずフラストレーションがたまる。

以上の他に被験者からは、「サイズの把握は a の文字提示が一番わかりやすかった」とコメントがあった。

6.3 考 察

実験結果から、d や e のように手を広げた枠内にサムネイル画像を提示させることは効果的であることがわかった。その一方で e に提示される引き伸ばされた画像には、指定したサイズにおける商品の特徴をユーザに把握させる利点があるが、e においては背景画像として隠れてしまう欠点がある。

この問題を解決するために、図 11 に示す視覚フィードバックを実装した。この視覚フィードバックにお



図 11 改善させた視覚フィードバック。
Fig.11 The improved visual feedback.

いてユーザが商品のサイズを指定すると、まず図 11 左に示すように引き伸ばされた画像が提示される。さらにユーザがそのサイズを保持すると、図 11 右に示すようにより小さいサムネイル画像へと提示される画像が遷移する。この視覚フィードバックにより、ユーザはまず商品のおおまかな特徴を把握し、その後に商品の個数や多様性を把握することが可能となる。今後はこの視覚フィードバックについても同様に評価実験を行う。

7. まとめと今後の課題

商品サイズを直観的に把握・実感して絞込むことを目的とし、我々は身体動作を用いて商品サイズを絞込む検索インタフェースである「これくら e-コマース」を開発してきた。本論文ではこれくら e-コマースのインタラクション手法と実装を示した。また、認識精度の評価実験を通して、商品サイズを指定する際の適切な粒度を示した。さらに視覚フィードバックの評価実験を通して、提示させる視覚フィードバックとしてより適切なものを示した。

今後は、商品のボリュームゾーンごとにサイズを指定させることにより、認識結果のぶれをより少なくする。さらに、ユーザに提示させる視覚フィードバックについても引き続き実装・評価を行いたい。

謝辞 本研究は「大規模情報コンテンツ時代の高度 ICT 専門職業人育成」事業の支援を受け、筑波大学と楽天株式会社との教育研究一体型プロジェクトにより実施されたものである。

参 考 文 献

- 1) 大江龍人, 蕨澤賢三, 岩淵志学, 益子宗, 田中二郎: これくら e-コマース: 商品サイズを直観的に絞込める検索インタフェース, 第 4 回楽天研究開発シンポジウム (2011).
- 2) Holz, C. and Wilson, A.: Data Miming: Inferring Spatial Object Descriptions from Human Gesture, CHI '11, ACM, pp.811-820 (2011).
- 3) 栗原竜矢, 岡部誠, 尾内理紀夫: Kinect センサを用いた物探し支援システムの試作, WISS '11, 日本ソフトウェア科学会, pp.218-220 (2011).
- 4) AR(拡張現実)を用いた試着システム. http://www.j-line.co.jp/service/tryon_ar.php (参照 2011 年 11 月 10 日).
- 5) Zoff Mirror. <http://www.zoff.co.jp/mirror/> (参照 2011 年 11 月 10 日).
- 6) 着たいあの服を瞬時に…中国でバーチャル試着室. <http://news.tv-asahi.co.jp/ann/news/web/html/211205015.html> (参照 2011 年 12 月 6 日).