

個人のコマンド空間に基づく複数人が操作可能な家電操作システム

Multi-User Home Appliance Operation System via Individual Command Spaces

眞嶋宗一郎（中央大） ○学 横田雅恵（中央大）
Sarthak Pathak（中央大） 正 梅田和昇（中央大）

Soichiro MAJIMA, Chuo University, majima@sensor.mech.chuo-u.ac.jp
Masae YOKOTA, Chuo University
Sarthak PATHAK, Chuo University
Kazunori UMEDA, Chuo University

This study proposes a gesture-based home appliance operation system that can support multiple users. In recent years, increasing complexity of home appliances has led to interest in gesture-based operation. Many systems have been developed to operate home appliances by recognizing gestures from camera images. However, conventional systems require an operator to memorize a large number of gestures and have problems of misoperation. To solve these problems, the proposed system introduces an operator specific command space and a start gesture. The effectiveness of the proposed system was verified by experiments, with an average recognition rate of around 95 %. This ensures that the system can be used by multiple operators.

Key Words: Intelligent room, Human machine interface, Gesture recognition, Control system of appliances

1. 序論

我々の日常生活で欠かすことのできない家電製品は、多機能化、高性能化が進んでいる。しかし一方では、操作ボタンの縮小化などによる操作の難化、複雑化という問題も生じている。近年これに対し、手振りなどの簡単なジェスチャを用いて家電製品を操作する研究が行われている。

従来研究では、複数種類のジェスチャを認識し、それぞれのジェスチャに家電操作のコマンドを設定することでジェスチャによる家電の操作を行うシステムが構築されている[1]-[3]。しかし操作可能なコマンド数はジェスチャの種類により限りがあるため、より多くの操作がしたい場合は覚えるジェスチャの数が増え、複雑になるという課題がある。これに対し、顔らは家電操作のコマンドを特定の空間（コマンド空間）に設定し、そこでの手かざしにより家電を操作するシステムを構築した[4]。しかしながら、各コマンド空間に設定されるコマンドは一つのため、一つの階層での操作可能なコマンド数は空間の数に限られてしまう。そこで、覚えやすい数の複数ジェスチャの認識とコマンド空間を組み合わせることで、多数の操作が簡単に行えるシステムを構築した[5]。しかし、ジェスチャとコマンド空間の場所の組み合わせによって、コマンド空間内でジェスチャを行う時に不自然な姿勢を取ってしまうことが原因で認識率が下がることが考えられる。また、コマンド空間の数が一部屋につき一つとなっており、固定して設置されるため、家電操作が可能な場所が限られてしまう。これに対し、操作者を基準とした相対座標系を設定し、操作者の周辺にコマンド空間を設置した手法がある[6]。しかし、コマンド空間が操作者の周辺に常に設置されるため、ふとした時の誤操作が起りやすいことが考えられる。

これらの課題を踏まえ、本研究では、操作者の位置推定を行い、追跡可能なコマンド空間を設置し誤操作が起りづらい複数人が操作可能なシステムを構築することを目的とする。また、操作開始ジェスチャの認識とそれぞれの操作者専用の個人のコマンド空間の導入を提案する。

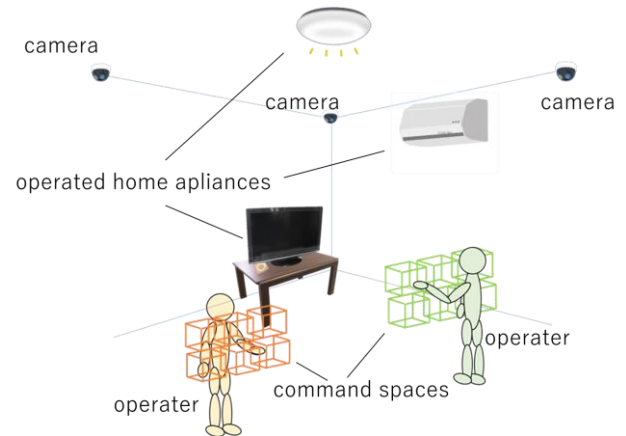


Fig. 1 Conceptual diagram of the system

2. 提案システムの概要

本研究で扱うシステムは、室内の家電を操作者のジェスチャのみで操作可能にしたものである。Fig.1に示すように、部屋を模した空間の天井の四隅に CCD カメラ (AXIS233D) を設置して画像を取得する。

提案するシステムの操作手順を次に示す。まず、操作開始のジェスチャを認識し操作者を特定する。次に、操作者の骨格点の三次元座標に基づいて相対座標系を定義する。本研究では、Fig.2のように四つのカメラに OpenPose[7]を適用して骨格点の三次元情報を取得する。コマンド空間は、相対座標系に基づいて構築され、操作者が動いてもコマンド空間が追跡する。操作者はコマンド空間内でジェスチャを行い、コマンド空間とジェスチャの組み合わせによって紐づけられた家電操作のコマンドを実行して家電操作を行う。なお、本研究では操作者が家電操作を行う時には静止しているとする。

3. 操作者の特定・追跡

操作開始のジェスチャを行った人物を操作者として特定する。また、ふとした時の誤操作と他人による誤操作を抑制するために、操作開始のジェスチャ後の一定フレームの間に追跡された操作者のみが家電操作を可能とする。操作開始のジェスチャは手を挙げるジェスチャである。

操作者の特定の手順を以下に示す。OpenPose を適用してひじと胸の中心点の二次元座標の候補を取得する。次に、候補の中でエピソード拘束を満たすものからステレオ計測により、ひじの三次元座標(x_1, y_1, z_1)と胸の中心点の三次元座標(x_2, y_2, z_2)を求める。一定フレーム連続して $z_1 - z_2 > 0$ となる時に操作開始のジェスチャとして認識する。ジェスチャ認識を行った人物を操作者として特定する。操作者として特定後、胸の骨格点座標を基準に操作者の体領域を一部切り取りテンプレート画像として保存する。また、操作者の胸の三次元座標から四台のカメラ画像内での二次元座標を求め、各カメラ画像での操作者の座標として追跡に用いる。

追跡の手順を以下に示す。まず、人物と前フレームの操作者の胸の中心点間の距離を求める。求めた距離が一定のしきい値以内の場合、操作者の候補として胸の骨格点座標を基準に体領域を一部切り取る。切り取った画像とテンプレート画像の色相の類似度を SAD (Sum of Absolute Difference) によって算出する。類似度が一定以上の場合に操作者として追跡する。類似度の高い候補者が複数人存在する場合は、一番類似度が高い候補者を操作者として追跡する。なお、操作者の位置情報を利用して追跡を行うため、間違った人物を操作者として追跡してしまった場合でも、二台以上のカメラで誤追跡が起きていなければ、その後のフレームでの操作者追跡手法によって追跡復帰が可能である。

4. 相対座標系によるコマンド空間の設置

4.1 相対座標系

本研究での相対座標系は、操作者の三次元情報を基準として定義される座標系である。定義後、操作者が移動すると相対座標系も共に移動する。操作者の特定後、相対座標系を展開し、その相対座標系を基準に操作者専用のコマンド空間を設置する。これにより部屋内の任意の位置での家電操作が可能になる。相対座標系は、OpenPose によって得られる骨格点情報のうち胸の中心点と左右の肩の三点の三次元座標を用いて定義される。胸の中心点を原点とする。また、左肩と右肩



Fig.2 Detection of skeleton points using OpenPose

の骨格点を結び胸の中心点上に平行移動したものを Y 軸とし、高さ方向を Z 軸、Y 軸と Z 軸の外積を X 軸とする。なお、Y 軸を水平にするために左肩と右肩の z 座標の値を等しくする。Fig. 3 に操作者に対する相対座標系の概略図を示す。

4.2 相対座標の計算

左肩、右肩、胸の中心点の三次元座標をそれぞれ $L(x_1, y_1, z_1)$, $R(x_2, y_2, z_2)$, $C(x_3, y_3, z_3)$ とする。ここで、コマンド空間を水平に設置するため、 z_1, z_2 の値を z_3 と等しくする。絶対座標(x, y, z)と相対座標(x', y', z')の関係を次式に示す。

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \mathbf{R} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} x_3 \\ y_3 \\ z_3 \end{pmatrix} \quad (1)$$

$$\mathbf{R} = \begin{pmatrix} \frac{y_2 - y_1}{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}} & \frac{x_1 - x_2}{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}} & 0 \\ \frac{x_2 - x_1}{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}} & \frac{y_2 - y_1}{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (2)$$

$\mathbf{R}$ は絶対座標系と相対座標系間の回転行列である。

4.3 コマンド空間の設置

コマンド空間とは、家電操作のコマンドが紐づけられた仮想空間である。空間に情報を埋め込む空間メモリ[8]の考え方をを用いて、家電操作のコマンドをコマンド空間に関連付ける[4]-[6]。コマンド空間内でジェスチャを行うことで、コマンド空間とジェスチャの組み合わせに紐づけられた家電操作のコマンドを実行する。

なお、コマンド空間内でジェスチャを行う時に不自然な姿勢を取ってしまうことが原因で認識率が下がる[5]ことが考えられる。相対座標系にコマンド空間を設置すると操作しやすい位置に体を動かすことが困難なため、本研究においてこの課題は重要であると考え。これに対し、コマンド空間を手の可動域を参考に弧を描くように配置し、操作する手の方向にずらすことを提案する。具体的なコマンド空間の配置を Fig.4 に示す。なお、操作に使用する手は開始ジェスチャが行われた手とする。

5. ジェスチャの認識

本研究では、コマンド空間における家電のコマンド操作を Fig.5 に示す 2 種類のジェスチャによって行う。いずれも手をコマンド空間にかざすジェスチャであるが、手首とひじの高さの差がジェスチャ 1 は大きいのにに対し、ジェスチャ 2 は小さい。2 種類のジェスチャの判別には骨格点情報を利用する。四つのカメラから取得した画像に OpenPose を適用して手首とひじの座標の候補を取得する。次に、候補の中でエピソード

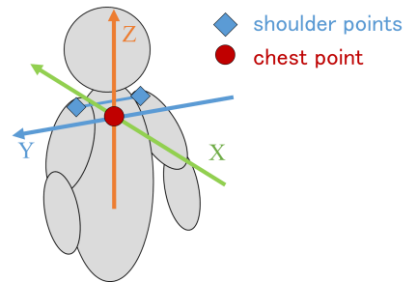
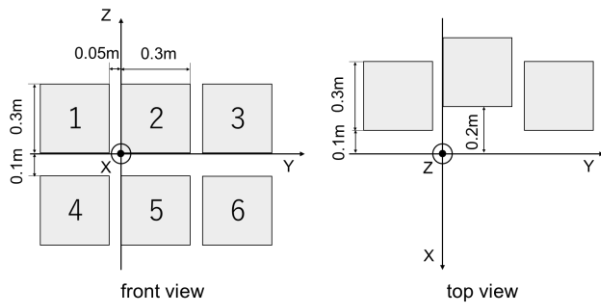
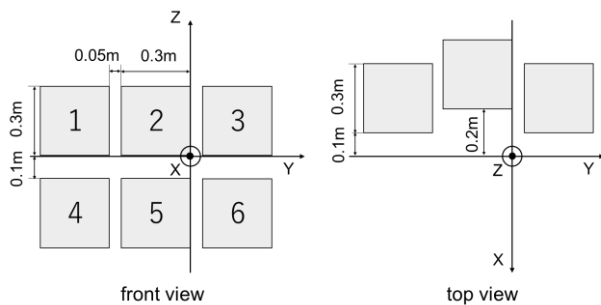


Fig.3 Relative coordinate system based on the operator



(a) Using right hand



(b) Using left hand

Fig.4 Arrangement of the command spaces

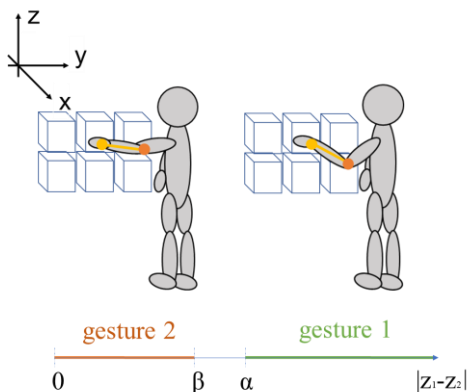


Fig.5 2 gestures

ラ拘束を満たす画素を手首とひじの画素として特定する。ステレオ計測により、手首の三次元座標(x_1, y_1, z_1)とひじの三次元座標(x_2, y_2, z_2)を求める。手首の三次元座標がいずれかのコマンド空間内にある時に手かざしのジェスチャと認識する。さらに、 $|z_1 - z_2| > \alpha$ の時にジェスチャ 1, $|z_1 - z_2| < \beta$ の時にジェスチャ 2 と判断する。 α, β は実験的に定めるしきい値である($\alpha > \beta$)。同じジェスチャが同じコマンド空間内で一定以上連続して認識されたときにコマンドの操作を行うことで、誤操作を防ぐ。 $\beta \leq |z_1 - z_2| \leq \alpha$ の時はジェスチャとして認識しないが、次フレームで認識されるジェスチャが前フレームのジェスチャと同じであれば、連続して同じジェスチャが行われたと判定する。

6. 実験

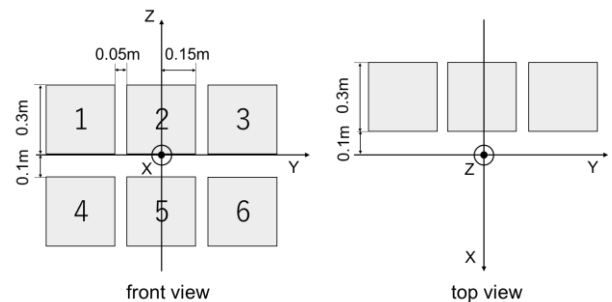
6.1 操作ジェスチャ認識率の評価実験

2 種類のジェスチャの六つのコマンド空間内での認識率を検証した。本システムの操作経験がある操作者一人を対象とした。操作者は同じ位置で各コマンド空間内で 2 種類のジェスチャを 30 回ずつ行い、手首の位置が正しいコマンド空間内にあると認識されたか (位置の認識率), さらにそれに加えて 2 種類のジェスチャの判別も行えたか (認識率) を調べた。

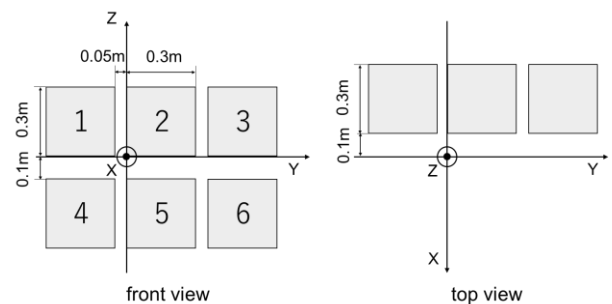
二つのジェスチャの認識率を Table 1, Table 2 に示す。なお、ジェスチャの認識には 2~4 [s]ほどかかった。位置の認識率は、ジェスチャ 1 が 98%, ジェスチャ 2 が 99% となった。ジェスチャが行われたコマンド空間はほぼ正しく認識できたことがわかる。平均認識率は、ジェスチャ 1 が 94%, ジェスチャ 2 が 96% であり、ほとんどのコマンド操作で認識率 90% 以上の結果が得られた。結果から操作者に対して相対座標系の設置が正確に行われ、ジェスチャの認識をほぼ正確に行えることが確認できた。また、上の段のコマンド空間でジェスチャ 2 を行うと腕が疲れてしまうので、何回でも楽に行えるジェスチャを導入する必要がある。また、1 回のジェスチャに 2~4 [s]ほど必要のため高速化も行う必要がある。

6.2 異なる配置のコマンド空間内での認識率の比較実験

提案したコマンド空間の配置による認識精度の違いを調べるために、異なる配置のコマンド空間の位置認識率の比較実験を行った。異なるコマンド空間は Fig.6 に示す 2 種類とし、配置 1 と配置 2 と提案するコマンド空間の配置を比較することで、コマンド空間を操作する手の方向にずらすことと弧を描くように設置することの有用性を検証した。本システムの操作経験がある操作者一人を対象とした。操作開始のジェス



(a) Arrangement 1



(b) Arrangement 2

Fig.6 Other arrangement of the command spaces

チャを行った後、手かざしジェスチャを 10 回ずつ行った。なお、提案するコマンド空間の認識率は前述の実験結果の通りである。

三つの異なる配置方法での位置認識率を Table3 に示す。位置認識率は、配置 1 が 78%, 配置 2 が 97%, 提案する配置が 99%となった。この結果から、コマンド空間を操作する手の方向にずらすことがコマンド空間での操作性向上に有用であることが確認できた。

Table 1 Recognition rates of gesture 1 [%]

Command	Recognition	Position recognition
1	100	100
2	100	100
3	90	100
4	87	90
5	93	100
6	93	100
Average	94	98

Table 2 Recognition rates of gesture 2 [%]

Command	Recognition	Position recognition
1	90	100
2	97	97
3	97	100
4	93	100
5	100	100
6	100	100
Average	96	99

Table 3 Position recognition rates for each arrangement of the command spaces [%]

Command	Arrangement 1	Arrangement 2	Proposed arrangement
1	30	90	100
2	100	100	98
3	90	100	100
4	20	100	95
5	100	100	100
6	100	90	100
Average	73	97	99

7. 結論

本研究では、操作開始ジェスチャと操作者専用のコマンド空間を導入し、誤操作が起こりづらい複数人が操作可能なシステムを構築した。操作精度とコマンド空間の配置についての評価実験を行った。その結果、操作者に合わせてコマンド空間の配置をずらすことで操作性が向上することが確認できた。コマンド空間内での 2 種類のジェスチャ認識は 9 割以上の認識率が得られた。

今後は、動きながらでもコマンド空間内での家電操作がスムーズに行えるシステムの構築を目指す。また、手かざしジェスチャ以外の覚えやすく簡単な複数ジェスチャの導入も目指す。

参考文献

- [1] 入江耕太, 若林直弘, 梅田和昇, “ジェスチャ認識に基づくインテリジェントルームの構築”, 日本機械学会論文集 C 編, vol.73-725, pp.258-265, 2007.
- [2] 松原孝志, 徳永竜也, “テレビ向けジェスチャ UI の試作開発”, 映像情報メディア学会年次大会公演予稿集, pp.4-5, 2010.
- [3] 米澤朋子, 堅川颯土, “ハンドジェスチャの文法的組み合わせによる複数家電操作手法の検討”, 日本人間学会第 63 回大会, vol.58, p2D5-02, 2022.
- [4] 顔世荀, 池勇勲, 梅田和昇, “任意の位置での簡単なジェスチャによる家電操作システム”, 日本機械学会論文集, vol.87-898, p. 20-00310, 2020.
- [5] 眞嶋宗一郎, 顔世荀, Pathak Sarthak, 梅田和昇, “コマンド空間での複数ジェスチャにより家電を操作するシステムの構築”, 精密工学会画像応用技術専門委員会サマーセミナー2021, 2021.
- [6] 鹿野巧, 顔世荀, 梅田和昇, “RGB-D センサを用いた簡易に構築可能な手振りによる家電操作システム”, 日本機械学会ロボティクス・メカトロニクス講演会講演概要集, 2018.
- [7] Cao, Z., Simon, T., Wei, S., Sheikh, Y., “Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields”, CVPR, pp.7291-7299, 2017.
- [8] Niitsuma, M., Kobayashi, H., Shiraishi, A., “Enhancement of Spatial Memory for Applying to Sequential Activity”, Journal of Advanced Sciences, vol.9-1 pp.121-137, 2012.